



Conference Proceedings

*Third Conference for Engineering Sciences and
Technology (CEST-2020)*

1-3 December 2020 - AL-Khoms/ Libya

Table of Contents

Conference Tracks	4
Track No. 1: Architecture & Urban Planning Engineering	6
Track No. 2: Civil Engineering	7
دراسة تأثير ألياف البولي بروبيلين والألياف الزجاجية على الخواص اللدنة و الصلدة للخرسانة ذاتية الدمك	8
العوازل الحرارية ودورها في تحسين البيئة الداخلية لمباني المناطق الصحراوية	17
تأثير نوع الملدنات الفائقة على التشغيلية والخواص الميكانيكية للخرسانة ذاتية الدمك	24
Performance of Recycled PVC Aggregates in Concrete – Comparative Study.....	31
Effect of Crude Oil Pollution on Some Geotechnical Properties of Disturbed Sandy Soil.....	40
Effect of Sawdust on The Mechanical Properties of Mortar Using Local Materials.....	51
Experimental Study of Drowned Hydraulic Jump Characteristics Through Different Counterflow Dimensions	59
The Influence of Ageing on Asphalt Mastic Properties Incorporating Calcium Carbonate.....	77
Track No. 3: Chemical & Oil Engineering.....	86
Simulation of Butanol Production through Hydrogenation of Butanal: Effects of Different Reactor Schemes and Operating Conditions.....	87
Evaporation Process Control using MIMO MPC	107
Lammps as Nano-scale Lab to Estimate Fluid Thermal Properties from Molecular Dynamics	117
Application of the TDS Technique in Gas Reservoirs.....	127
Transient Pressure Analysis and Productivity Index Estimation in Horizontal Wells	136
Removal of Hexavalent Chromium from Aqueous Solution using Magnetic Sepiolite Nano-composite as an Adsorbent	148
Track No. 4: Electrical & Computer Engineering	159
A Novel Denoising Method Based on Discrete Linear Chirp Transform.....	159
Optimized Maximum Loadability of Power Systems using an Enhanced Dynamic JAYA Algorithm	167
Optimum Backoff in CSMA-CA for Ad-Hoc Vehicles Network in Motorway	173
Database for Arabic Speech Commands Recognition	180
Decision-Making Rules in Cooperative Cognitive Radio Networks: Evaluation and Comparison	188
Investigation of Renewable Electricity Generation from Solar-Hydrogen Hybrid System in Tripoli	198
An Energy Efficiency Evaluation of MIMO Based LTE RANs with DTX Operation.....	209
Maximum Area Aggregation Approach For Cumulant-Based Probabilistic Optimal Power Flow studies.....	221
Face Recognition with Symmetrical Face Training Samples Based on Histograms of Oriented Gradients	230
Comparative Study of the Electromagnetic Performance of Yokeless and Segmented Armature YASA Machine with Different Rotor Pole Combinations	239
Design of a Non-Isolated Solar PV Inverter for Household Applications.....	247

Quick-Response Fuzzy-Controlled Induction Motor Drive.....	254
Electricity Load Profiling for Coastal Housing Complexes based on the Measurements of Flats Actual Load	262
Track No. 5: Information Technology	272
المحاكاة Packet Tracer 7.3 و لوحة الاردوينو Arduino board باستخدام برنامج المحاكاة	273
The Virtual University: Trends and Challenges in Libya Elmergib University as Case Study	280
Performance Evaluation of LSB Technique into visual Objects based on Steganography.....	290
Developing a mobile game app themed about Libyan culture using Unity engine	300
Comparison of BER Performance of Cooperative Wireless Systems Based on D-OSTBC, D- EOSTBC and DF Relaying Protocol	311
Classifying Various Bacteria Genera by Transfer Learning Model	318
An Extensive Study on Online and Mobile Ad Fraud	326
Design of New Secret Key to Increase the Security of LSB Algorithm.....	338
Internet of Things for Libya Healthcare System: Challenges and Issues	348
Enhancement of VoIP Performance in MANET using Fuzzy Logic.....	354
Comparison of BER Performance of SISO, MISO and MIMO Systems Based on OSTBC and EO- STBC Techniques	364
Track No. 6: Engineering Management and Quality.....	376
تقدير الكلفة الخفية لإنحراف خيوط الصوف الخالص المستخدم في صناعة السجاد (شركة الإنماء للصناعات الصوفية بني وليد)	376
دراسة أثر استخدام استراتيجيات التنفيذ المختلفة للمشروعات الانشائية على أداء الاطراف الرئيسية في مشاريع الانشاءات باستخدام نموذج الأداء العام.....	383
Track No. 7: Mechanical & Industrial Engineering	389
Effects of a Curved Surface of the Blade on the Performance Characteristics of Axial Flow Rotors	390
Dimensions Effect of the Rotating Fluid Zone on the Results when CFD Modeling of Friction Stir Welding.....	399
Risk-Based Maintenance (RBM) Approach for Identifying the Optimum Time of Whole Shutdown (SD) for Gas Liquid Recovery Unit; Processing Columns as a Case Study	410
Frequency Response Function (FRF) Technique for the Diagnosis of Suspension System	416
SCARA Robot Links Length Optimization by Using MATLAB and Verification with SimMechanics and Solidworks	429
تخفيض الاحمال الكهربائية باستبدال السخانات الكهربائية بأنظمة تسخين المياه بالطاقة الشمسية.....	439
Study the Performance and Behavior of the Hybrid System.....	449

Conference Tracks

1. Architecture & Urban Planning Engineering

- Architecture and Urban Planning
- Design Flexibility and its Impact on Architecture and Urbanism
- Digital Architecture and its Effects on Design
- Human Interaction in General Voids
- Historic City Centers and their Role in Achieving Urban Sustainability in Libyan Cities
- Environmental Design and its Impact on Building Performance in Libya
- Architectural Spaces: Analysis and Criticism
- Architecture and Climate Change
- Green Building Design
- Smart & Sustainable Cities
- Building Performance Visualization

2. Civil Engineering

- Design, Analysis and Rehabilitation of Structures
- Sustainable Building Materials and Concrete Technologies
- The Role of Transportation Engineering in Sustainable Development
- Geomatics Engineering and its Applications
- GIS Techniques and its Implications in Society
- Hydraulics, Water Resources and Environmental Engineering.
- Water & Wastewater Treatment and Air Pollution Control

3. Chemical & Oil Engineering

- Separation Technologies
- Catalysis and Chemical Reaction Engineering
- Fluid Mechanics and Heat Transfer Operations
- Energy, Fuels, and Petrochemicals
- Drilling Engineering
- Reservoir Engineering
- Gas Processing
- Environmental Pollution and Engineering
- Polymer Engineering and Nanotechnology
- Modelling and Design in Chemical and Petroleum Engineering

4. Electrical & Computer Engineering

- Broad Band Communication
- Mobile, Wireless and Optical Communication
- Antenna and wave Propagation
- Microwave Engineering and Radar system
- Spectrum Management and Cognitive Radio
- Signal and Image Processing
- Control System Technologies
- Real time control and transportation traffic control
- Robotics and Mechatronics
- Smart Grid Systems and Applications
- Instrumentations
- Power Quality Improvement Techniques
- Power System Modeling and Simulation
- Power System Generation and Distribution



5. Information Technology

- Computer Science and Algorithms
- Software Engineering and Cloud Computing
- Computer Networks and Security
- Multimedia Technologies
- E-Government and E-Commerce
- Education in Information Technology
- Intelligent and Decision Support Systems
- Information Systems
- Human-Computer Interaction
- Database Systems
- Internet Technologies

6. Engineering Management and Quality

- Design and Implementation Management
- Quality Control and Process Monitoring
- Human Resource Management
- Product and process optimization
- Test and Evaluation Techniques
- Risk Management
- Planning and scheduling

7. Mechanical & Industrial Engineering

- Mechanical Power Engineering
- Industrial Engineering
- Production Engineering
- Applied Mechanics
- Materials science Engineering



Track No. 1: Architecture & Urban Planning Engineering



Track No. 2: Civil Engineering

دراسة تأثير ألياف البولي بروبيلين والألياف الزجاجية على الخواص اللدنة و الصلدة للخرسانة ذاتية الدمك

م. خالد محمد عمرو¹، د. عبدالباسط عبدالله التركي²، م. عبدالله أبو القاسم عمرو³، م. يخلف زكري يخلف⁴
khaled200473@gmail.com¹, et.turki@hotmail.com², abdo84a@yahoo.com³,
yakhlefzkri@yahoo.com⁴

¹ قسم الهندسة المدنية، كلية الهندسة، جامعة غريان، غريان، ليبيا، ² قسم الهندسة المدنية، كلية الهندسة، جامعة طرابلس، طرابلس،
ليبيا، ³ قسم الهندسة المدنية، المعهد العالي للعلوم والتقنية، سوق الجمعة، ليبيا، ⁴ قسم الهندسة المدنية، المعهد العالي للتقنيات الهندسية
، طرابلس، ليبيا

ABSTRACT

من المتعارف عليه في مجال تقنية الخرسانة أن الخرسانة ذاتية الدمك (Self-Compacting Concrete) المحتوية على الألياف هي إحدى تطبيقات المشاريع الحديثة المدنية المختلفة و المتنوعة.

والهدف الأساسي في هذا البحث هو إضافة بعض المواد المضافة لتحسين بعض خواص الخرسانة ذاتية الدمك باعتبارها تستعمل بكثرة في تنفيذ المشاريع، وخاصة المباني الحديثة التي تكون بها كثافة تسليح عالية و تكون عالية الارتفاع، ويجب الإشارة إلى أن هناك عدة أنواع من الألياف مثل الألياف الحديدية و الألياف الزجاجية و ألياف الكربون.

في هذا البحث تم دراسة إضافة ألياف البولي بروبيلين والألياف الزجاجية إلى الخرسانة و بنسب 0 %، 0.25 %، 0.50 %، 0.75 %، 1 % من حجم الخرسانة حيث أجريت عدة اختبارات على الخرسانة الطرية و الصلدة، منها اختبار حساب مقاومة الضغط و الشد.

و تبين النتائج بأنه عند إضافة مادة الألياف تبدأ مقاومة الضغط في النقصان حيث كانت أعلى مقاومة ضغط في حدود 51 نيوتن / مم² عند نسبة ألياف 0 %، و أقل مقاومة كانت عند 33 نيوتن / مم² عند نسبة ألياف 1.0 %، بينما مقاومة الشد تتناسب طردياً مع الألياف حيث كانت أقل مقاومة شد بدون ألياف عند 3.16 نيوتن / مم² بينما أعلى قيمة عند 3.72 نيوتن/مم².

Keywords:

خرسانة ذاتية الدمك،
ألياف البولي بروبيلين،
المادة المضافة،
الألياف الزجاجية،
مقاومة الضغط،
مقاومة الشد.

1- المقدمة

الخرسانة ذاتية الدمك (Self-Compacting Concrete) SCC هي الخرسانة التي لها القدرة على الانسياب ذاتياً تحت تأثير وزنها ولا تحتاج إلى دمك أو هز، كما لها خاصية تعبئة كل الفراغات الموجودة في الشدات الخشبية والتحرك في العناصر الخرسانية متغيرة الأبعاد والتي تحتوي على تسليح كثيف بحيث لا يحدث لها انفصال حبيبي (Segregation) ولا نضح (Bleeding) ولكي يتم الحكم على الخرسانة ذاتية الدمك لابد من تحقيق الخواص الثلاثة في الحالة الطرية: قابلية المرور (Passing ability) وقابلية التعبئة (Filling Ability) واستقرار الخرسانة أو مقاومة الانفصال (Resistance of Segregation).

تنفذ الخرسانة ذاتية الدمك (SCC) بنفس مكونات الخرسانة التقليدية مع الفارق بضرورة الإكثار من المواد الناعمة (الرمال) المستخدم والتقليل من الركام الخشن واستخدام إضافات اللزوجة (Viscosity Modifying Agent) و الملدنات الفائقة الفاعلية (Super-plasticizer) التي تؤدي إلى زيادة سيولة الخرسانة [1].

صنفت الخرسانة ذاتية الدمك (SCC) على أنها من أكبر التطورات التي حصلت في علم تقنية الخرسانة في العقود الثلاثة الماضية؛ ومن مميزات الخرسانة ذاتية الدمك أنها سريعة التنفيذ، التقليل من حجم العمالة في الموقع، سهولة الصب، سهولة التشغيل، جيدة الديمومة، تقلل من الضوضاء بالتقليل من الأجهزة والمعدات الخاصة بالدمك في الموقع، زيادة المقاومة [1]

2 - المواد المستخدمة و البرنامج العملي

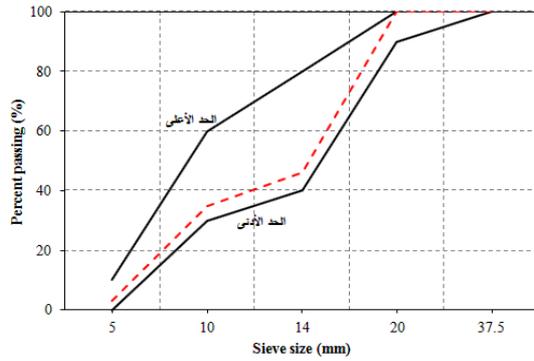
1.2- الماء (Water) : حيث تم استعمال الماء الصالح للشرب و الخالي من الشوائب [2] .

2.2- الإسمنت (Cement) : تم استخدام الإسمنت البورتلاندي العادي في تجهيز جميع الخلطات ويعتبر هو المادة الرابطة و اللاصقة وذلك يعتبر هو المادة المسؤولة عن مقاومة الضغط في الخرسانة [3].

3.2- السليكا فيوم (Silica Fume) : تم استخدام السليكا فيوم في جميع الخلطات الخرسانية , وكانت النسبة المستخدمة هي 10% من وزن الاسمنت , الوزن النوعي للسليكا فيوم 2.2 حيث إن المادة الأساسية في تركيبها هي أكسيد السليكون و تزيد عن 90% من وزنها الكلي، ولها نعومة عالية جداً تصل إلى 200000 سم²/جم أي أنعم من الإسمنت البورتلاندي العادي بأكثر من 100 مرة [4].

4.2- الركام الخشن (Coarse Aggregate) : تم استخدام ركام مقاسه الاعتراري الأكبر 20 ملم حيث تم استخدام هذا النوع مع جميع الخلطات الخرسانية وكان منحني التدرج الحبيبي لهذا الركام داخل حدود التدرج طبقاً للمواصفة الليبية 2002/49 كما هو مبين في الشكل (1). كذلك النتائج الميكانيكية الخاصة بالركام المستخدم في هذه الدراسة موضحة أسفل في الجدول (1) [6,5].

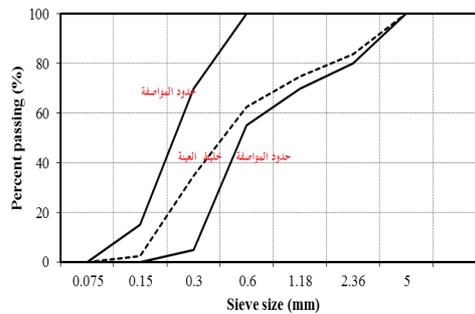
جدول رقم (1): الخواص الميكانيكية للركام الخشن



الشكل رقم(1): منحنى التحليل المنخلي للركام الخشن

الاختبار.	النتيجة	مواصفة الاختبار المواصفة الليبية رقم 49
الوزن النوعي	*2.62	-
نسبة الامتصاص%	%2.3	%3
معامل التهشيم%	%25	لا يزيد عن 45%
معامل الصدم%	%25.7	لا يزيد عن 45%

5.2- الركام الناعم (Fine Aggregate) : تم استخدام نوعين من الركام الناعم في هذه الدراسة، وذلك لأن الركام الناعم (الرمل) ذو معامل نعومه عالي و الرمل متوسط النعومة (قرينيليا) الذي تم استخدامه في هذا البحث هي من القرينيليا المتوفرة بالسوق الليبي وكانت نتائج التحليل المعمل لتلك العينة مبينة في جدول (2)، وعند إجراء اختبار التدرج الحبيبي لعينة من (الرمل) كانت ذات نعومة عالية بالنسبة لحدود المواصفة , وكذلك من خلال دراسات سابقة تبين أنه أثناء استخدام ركام ناعم ومعامل نعومته عالية مما يزيد من صعوبة تدفق الخرسانة ذاتياً ولتفادي حدوث هذه المشاكل مثلاً تم خلط الركام الناعم (الرمل) مع نوع آخر من الركام الناعم الأكثر خشونة ، و كانت نتيجة التحليل المنخلي بعد ما تم الخلط كلاً من الركام الناعم (الرمل) إلي القرينيليا بنسبة 1:1.5 كما في الشكل (2) [6].



الشكل رقم (2): التحليل المنخلي للركام الناعم (الرمل) مع متوسط النعومة بنسبة (1:1.5)

جدول رقم (2): الخواص الفيزيائية للركام الخشن

الاختبار	النتيجة	مواصفة الاختبار (م ق ل 49) [5]
الوزن النوعي	*2.61	-
نسبة الامتصاص (%)	2.2	لا تزيد عن 3

6.2- الإضافات الكيميائية (Admixtures): تم استخدام في هذه الدراسة نوع الملدن (Glenium51) من منتجات شركة BASF حيث يسمى معدل اللزوجة (Viscosity modified) كما يعتبر هذا الملدن مصنع من البوليمرات الصناعية المخصصة خصيصاً لحالات التشغيلية المنخفضة واكتسابها خواص لزوجة عالية , حيث تمنع حدوث عملية النضح (Bleeding) وزيادة في التشغيلية ،حيث إنه تم استخدام عدة نسب مختلفة في كل خلطة على الأخرى بناءً على النتائج الانسيابية للخلطات التجريبية عند زيادة نسبة الألياف بحيث تعطي انتشاراً يتراوح بين 550–800 ملم [8,7].

7.2 - الألياف البولي بروبيلين (Polypropylene Fibers): الألياف البولي بروبيلين تعتبر من أكثر أنواع الألياف الصناعية انتشاراً، وهي من أصل عضوي ومن أرخص المواد المبلعمة، فالمادة الخام لإنتاجها وهي البولي بروبيلين تعتبر منتج ثانوي من عملية تكسير النفط، والشكل (3) يوضح شكل الألياف البولي بروبيلين المستخدم في هذه الدراسة وكذلك الجدول رقم (3) أدناه يوضح مواصفات هذه الألياف .

جدول رقم (3): مواصفات ألياف البولي بروبيلين

الكثافة kg/m ³	طول الألياف mm	القطر mm	المساحة السطحية m ² /kg	قوة الشد Mpa	معامل المرونة Mpa
0.91	12	0.02	250	400-300	8500-5000



الشكل رقم (3): الألياف البولي بروبيلين المستخدمة في

3 – البرنامج العملي:

بعد ما يتم تحضير نسب مواد الخلطة المراد اختبارها ، يتم تحضير الخلطة الكهربائية ومن ثم تبدأ عملية وضع المواد بالخلطة، ويتم وضع نصف كمية الركام الخشن ثم توضع كمية الإسمنت و يليه الرمل الناعم ثم باقي كمية الركام الخشن بالخلطة ويبدأ في الخلط على الجاف من (2-3) دقائق تقريباً ومن ثم تحضير العينات اللازمة لإجراء التجارب على الخرسانة الطرية ومن هذه الاختبارات ، اختبار قابلية التعبئة و قابلية العبور و L-Box و Funnel -V و كذلك تجهيز بعض العينات الخاصة باختبارات الخرسانة الصلدة في القوالب و إجراء عملية المعالجة اللازمة عليها و تحضيرها لاختبار كل عينة حسب الاختبار المطلوب على الاجهزة المتوفرة بالمعمل ، و من هذه الاجهزة ، جهاز اختبار مقاومة الضغط و جهاز اختبار مقاومة الشد و الانحناء .

1.3- الخلطات الخرسانية المختارة:

بناءً على ما تم التوصل إليه من إعداد الخلطات التجريبية من ناحية تحقيق الخرسانة ذاتية الدمك بتطبيق شروطها حيث تم إعداد عدة خلطات خرسانية في هذا البحث لدراسة تأثير ألياف البولي بروبلين على خواص الخرسانة ذاتية الدمك حيث تم الاعتماد على طريقة المعادلة الحجمية لتصميم الخلطات [8].

وكذلك تم دراسة إضافة ألياف البولي بروبلين مع الألياف الزجاجية وخلطها بنسبة (0.25 ، 0.50 ، 0.75) % و تم تقسيم الخلطات الخرسانية إلى جزئين.

الجزء الأول:- دراسة تأثير ألياف البولي بروبلين (PPF) على خواص الخرسانة ذاتية الدمك (SCC) و يبين الجدول (4) الأوزان لكل متر مكعب للخلطات الخرسانية المختارة ذاتية الدمك المسلحة بألياف البولي بروبلين (PPFSCC) حيث تم استخدام نسبة الرمل الناعم الى الركام الخشن بنسبة 1:1 بناءً على نتائج الكثافة النسبية (Bulk Density).

جدول رقم (4): الأوزان لكل متر مكعب لخلطات الخرسانة المختارة للجزء الأول خرسانة ذاتية الدمك المسلحة بألياف البولي بروبلين (PPFSCC)

Number of mix	W/B Ratio	PPF %	CA kg	FA kg	C Kg	SF Kg	W Kg	PPF Kg	HRWR* Kg
M ₅₋₁	0.40	0	892	892	376.2	41.8	167	0	4.19
M ₅₋₂	0.40	0.25	888	888	376.2	41.8	167	2.27	5.23
M ₅₋₃	0.40	0.50	884	884	376.2	41.8	167	4.54	6.28
M ₅₋₄	0.40	0.75	878	878	376.2	41.8	167	6.81	6.81
M ₅₋₅	0.40	1	874	874	376.2	41.8	167	9.07	7.33

الجزء الثاني:- دراسة تأثير ألياف البولي بروبلين مع الألياف الزجاجية على خواص الخرسانة ذاتية الدمك (SCC) في الحالة الصلدة و يبين الجدول (5) الأوزان لكل متر مكعب للخلطات الخرسانية المختارة ذاتية الدمك المسلحة بألياف البولي بروبلين مع الألياف الزجاجية (SyF+ PPF).

جدول رقم (5): الأوزان لكل متر مكعب للخلطات الخرسانية المختارة للجزء الثاني خرسانة ذاتية الدمك المسلحة بألياف البولي بروبلين مع الألياف الزجاجية SCC (SyF + PPF)

Number of mix	W/B Ratio	F %	CA Kg	FA kg	C Kg	SF Kg	W Kg	PPF Kg	SyF kg	HRWR* Kg
M ₅₋₁	0.40	0	892	892	376.2	41.8	167	0	0	4.19
M ₅₋₇	0.40	0.25	888	888	376.2	41.8	167	1.70	0.57	4.19
M ₅₋₈	0.40	0.50	884	884	376.2	41.8	167	2.27	2.29	6.28
M ₅₋₉	0.40	0.75	878	878	376.2	41.8	167	1.70	5.16	6.81
M ₅₋₁₀	0.40	1	874	874	376.2	41.8	167	0	9.9	7.33

*HRWR = High Range Water Reduce.

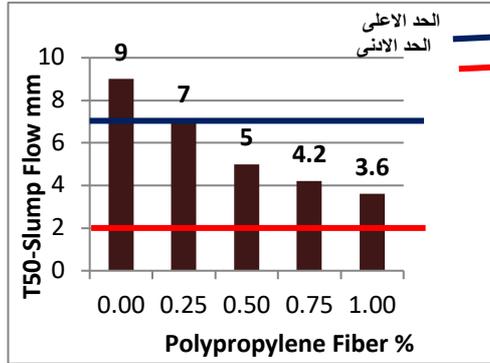
4 - الاختبارات و النتائج للخرسانة ذاتية الدمك في الحالة الطرية و الصلبة:

بعد ما يتم تحضير نسب مواد الخلطة المراد اختبارها و الانتهاء من عملية الخلط ، بعد ذلك يتم عمل الاختبارات الخرسانية في الحالة اللدنة الخاصة بتحقيق شروطها الخرسانة ذاتية الدمك ويتم تعبئة عدد 6 مكعبات (150×150×150مم) و ذلك لإجراء اختبار مقاومة الضغط ، وتعبئة عدد 6 اسطوانات مقاس (100×200مم) لإجراء اختبار مقاومة الشد عليها [10,11]

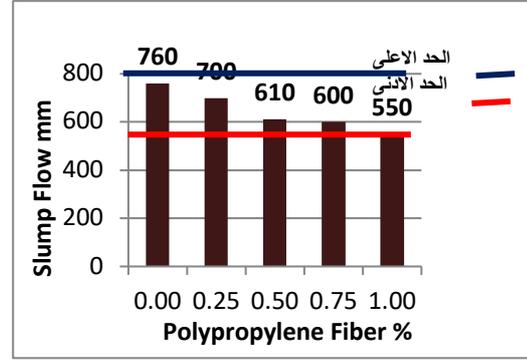
1.4 - قابلية التعبئة (Filling ability) :

يتم التحقق من هذه الخاصية بإجراء اختبار الانتشار (Slump flow) وكذلك اختبار (T50) والشكل (4) يوضح تأثير ألياف البولي بروبلين على الهبوط ، والشكل (5) يوضح تأثير ألياف البولي بروبلين على اختبار T50-Slump وحسب

الدليل لأوروبي الخاص بالخرسانة ذاتية الدمك فإن جميع القيم لهذا الاختبار داخل الحدود المسموح بها وذلك ما بين 550 – 750 حيث نلاحظ أنه أعلى قيمة كانت عند نسبة الألياف 0% في حدود 760 مم أما أقل قيمة فكانت عند نسبة 1.0% للألياف في حدود 550 مم وهي الحد الأدنى للدليل الإرشادي الأوروبي وبالنسبة لاختبار T50-Slump فكانت نتائج الاختبار لجميع النسب داخل حدود الدليل وهي ما بين 2 – 7 مم عند إضافة مادة الألياف ما عدا عند نسبة 0% من الألياف فإنها كانت خارج حدود القيم المطلوبة بالدليل الأوروبي [12]



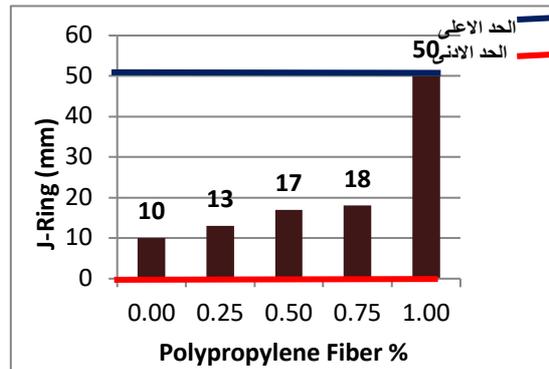
الشكل رقم (5): تأثير ألياف البولي بروبيلين على نتائج اختبار T50 slump



الشكل رقم (4): يوضح تأثير ألياف البولي بروبيلين على نتائج اختبار الانتشار

2.4- قابلية العبور (Passing ability):

يمكن تعيين قابلية العبور للخرسانة ذاتية الدمك المسلحة بالألياف في حالة وجود العوائق باختبار J-Ring كذلك يمكن معرفة مدى قدرة انفصال الحبيبي والشكل (6) يبين نتائج تأثير ألياف البولي بروبيلين على قابلية المرور حيث توضح النتائج أنه كلما زادت نسبة الألياف كلما زادت نسبة العبور ونلاحظ عندما كانت نسبة الألياف 0% كانت أقل نسبة عبور وهي 10 مم وبذلك فإنه عند أعلى قيمة للألياف 1.0% كانت أعلى قيمة للعبور وهي 50% هذا ما يدل على أن الألياف تزيد من قيمة التشغيل للخرسانة و كانت جميع النتائج داخل حدود المواصفات حسب الدليل الإرشادي الأوروبي الخاص بالخرسانة ذاتية الدمك [12]

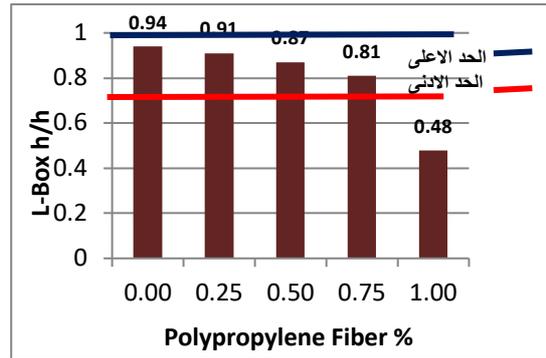


الشكل رقم (6): تأثير ألياف البولي بروبيلين على نتائج

اختبار J-Ring بناء على المواصفة ASTM

3.4 - اختبار L-box :

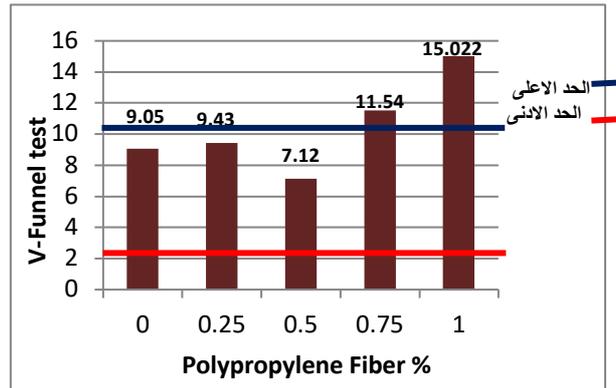
اختبار L-box يعتبر من الاختبار المهمة في تقييم ومتابعة قدرة الخرسانة على الاستواء ذاتياً في حالة صب المنشآت الخرسانية ذات التسليح الكثيف ومتغيرة المقاطع والشكل (7) يوضح تأثير ألياف البولي بروبيلين على نتائج اختبار L-box وذلك حسب الدليل الإرشادي للخرسانة ذاتية الدمك حيث أنه تنص على ان الحدود ما بين 0.7 – 1.0 و نلاحظ من النتائج بأنه جميع القيم داخل الحدود المطلوبة باستثناء القيمة عند زيادة نسبة الألياف الى 1.0% حيث كانت خارج الحدود عند 0.48 [12]



الشكل رقم (7): تأثير ألياف البولي بروبيلين على نتائج اختبار L-Box

4.4 اختبار V-Funnel:

اختبار V-Funnel يعتبر من الاختبارات المكتملة لتقييم سلوك الخرسانة أثناء تدفقها خلال المنشآت متغيرة المقاطع والأبعاد والشكل (8) يبين نتائج تأثير ألياف البولي بروبيلين على اختبار V-Funnel حسب الدليل الاوروبي لاختبارات الخرسانة ذاتية الدمك فإن الحدود المسموح بها في هذا الاختبار ما بين 3 – 12 ثانية ، و نتائج الاختبار من خلال الشكل أدناه كانت داخل الحدود ماعدا عند أعلى قيمة للألياف فكانت خارج الحدود عند زمن 15.02 ثانية [12].

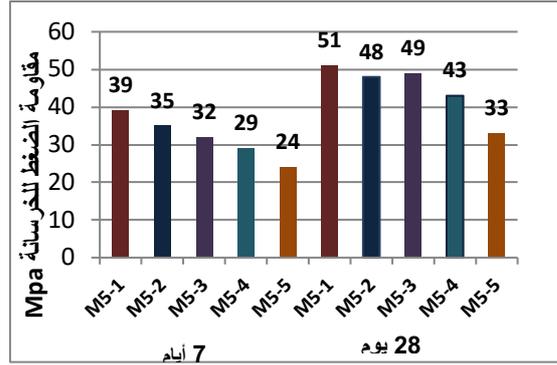


الشكل رقم (8): تأثير ألياف البولي بروبيلين

على اختبار V-Funnel

5.4- نتائج مقاومة الضغط للخرسانة (Compressive Strength):

تم اختبار مقاومة الضغط بعد 7 ، 28 يوم و النتائج كما هو موضحة في الشكل (9) حيث كانت النتائج لهذه المقاومة لمتوسط عدد ثلاثه مكعبات في كل اختبار اجري على مكعبات ذات مقاس (150x150x150) ملم وتم الاختبار بناءً على المواصفة البريطانية BS881-part116:1983 [10] ويمكن ملاحظة مقاومة الضغط للخرسانة المسلحة بالألياف من خلال الشكل أدناه لجميع الخلطات أن المقاومة تقل كلما زادت نسبة الألياف حيث كانت أعلى مقاومة عند 7 أيام في حدود 39 نيوتن/مم² عند نسبة ألياف 0 % بينما كانت أقل مقاومة عند نسبة ألياف 1.0 % في حدود 24 نيوتن /مم²، وكذلك عند مقاومة الخرسانة بعد 28 يوم كانت أعلى مقاومة عند نسبة ألياف 0% في حدود 51 نيوتن / مم²، وكانت أقل مقاومة عند أعلى نسبة ألياف في حدود 33 نيوتن / مم² [7].



الشكل رقم (9): تأثير ألياف البولي بروبيلين

6.4- نتائج مقاومة الشد للخرسانة المسلحة بالألياف (Tensile Strength):

بالنسبة لمقاومة الشد للخرسانة ذاتية الدمك المسلحة بالألياف كانت النتائج كما هي في الشكل (10) ، تم اختبار مقاومة الشد للخرسانة بعد 7، 28 يوم على اسطوانات ذات ابعاد قياسية (100 X 200) ملم ، حيث تم اختبار مقاومة الشد للخرسانة (Strength Tensile Splitting) بمركز البحوث الصناعية تاجوراء و أجري هذا الاختبار بناءً على المواصفة القياسية الأمريكية ASTM C496/C496M-04 ومن خلال النتائج المتحصل عليها في هذه الدراسة كما هي موضحة بالشكل (11) لوحظ أن مقاومة الشد للخرسانة ذاتية الدمك تزداد كلما زادت نسب الألياف حيث تعمل الألياف على مقاومة توسع التشققات أثناء التحميل وهو ما يعرف بنظرية (Crack Bridging) وبالتالي تزداد مقاومة الشد في الخرسانة، هذه النظرية نفس نظرية حديد التسليح المضاف إلى الخرسانة في منطقة الشد^[13].

حيث أنه حسب النتائج المتحصل عليها في الخلطة الخرسانية (M₅₋₁) كانت نسبة الألياف 0% وكانت مقاومة الشد في عمر 28 يوم في هذه الخلطة 3.16 MPa بينما في الخلطة الخرسانية (M₅₋₅) والتي تحتوي على 1% من الألياف كانت مقاومة الشد للخرسانة في عمر 28 يوم في حدود 3.72 MPa .

مقاومة الشد للخرسانة تم حسابها بناءً على المواصفة القياسية الأمريكية ASTM 496/C496 M-01 على الاسطوانات ذات الابعاد القياسية (100X 200) ملم بعد اتمام عملية المعالجة خلال 7 أيام و28 يوم. مقاومة الشد للخرسانة تم حسابها بعد إيجاد اقصى سعة تحمله لحمل الكسر بناءً على المعادلة التالية :

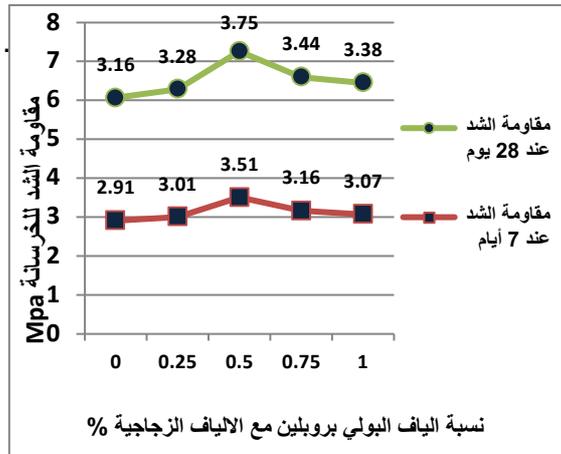
$$T = \frac{2P}{\pi L d} \quad (1)$$

T : مقاومة الشد للخرسانة (N/mm²)

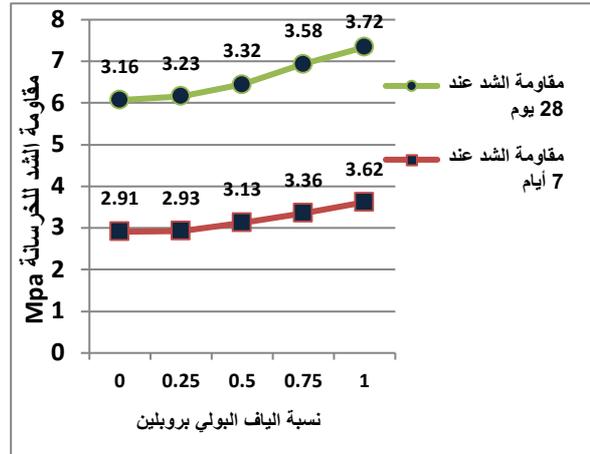
P = أقصى حمل الكسر N

L = الطول mm

d = قطر الأسطوانة (mm)



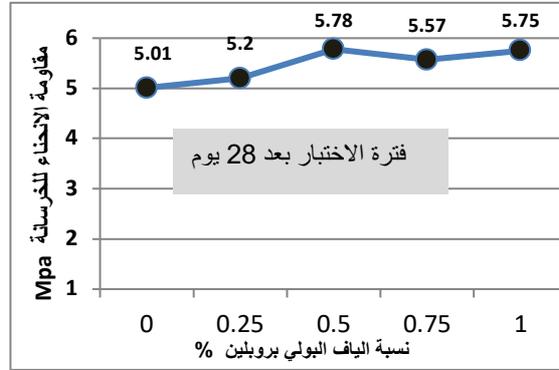
الشكل رقم (11): تأثير الياق البولي بروبيلين مع الالياف الزجاجية على مقاومة الشد للخرسانة ذاتية الدمك



الشكل رقم (10): يوضح تأثير الياق البولي بروبيلين على مقاومة الشد للخرسانة ذاتية الدمك

7.4- مقاومة الانحناء للخرسانة (Pending Strength):

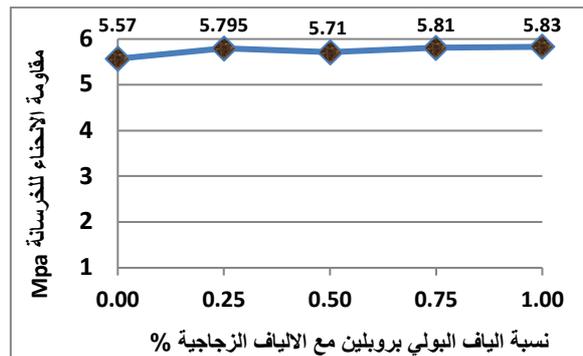
تم اختبار مقاومة الانحناء على الخرسانة ذاتية الدمك المسلحة بالألياف على كمرات (Prisms Beam) ذات أبعاد (100×100×500) ملم وأخذ متوسط لعدد ثلاثة عينات بعد (28، 7) يوم من إتمام عملية المعالجة، حيث تم إجراء اختبار الانحناء للخرسانة ذاتية الدمك بناءً على المواصفة الأمريكية ASTM 78 لجميع الخلطات والشكل (12) يوضح نتائج اختبار مقاومة الانحناء للخرسانة ذاتية الدمك المسلحة بألياف البولي بروبيلين (PPFSCC) [13].



الشكل رقم (12) : تأثير ألياف البولي بروبيلين على مقاومة الانحناء للخرسانة ذاتية الدمك

ونلاحظ من خلال الشكل السابق أن في الخلطة الخرسانية M5_2 عند نسبة الياف في حدود 0.25% كانت مقاومة الانحناء للخرسانة 5.2 MPa، بينما في الخلطة الخرسانية M5_3 كانت نسبة الياف البولي بروبيلين المضافة 0.50% كانت مقاومة الانحناء للخرسانة 5.78 MPa. من خلال النتائج السابقة يتضح أن مقدار الزيادة لمقاومة الانحناء للخرسانة يتراوح بين 4% إلى 16%.

و الشكل (13) يوضح نتائج مقاومة الانحناء للخرسانة ذاتية الدمك المسلحة بألياف البولي بروبيلين مع الألياف الزجاجية حيث نلاحظ أن في الخلطة M5_1 عند نسبة ألياف 0% كانت هذه الخلطة بدون ألياف وكانت مقاومة الانحناء للخرسانة 5.57 MPa بينما في الخلطة M5_7 نسبة الألياف كانت (SYF = 0.0625%، PPF = 0.1875%) وكذلك عند نسبة الألياف المضافة 0.25% كانت مقاومة الانحناء للخرسانة ذاتية الدمك هي 5.795 MPa حيث إن الخلطة الخرسانية M5_10 نسبة الألياف الزجاجية بها 1% كانت مقاومة الانحناء للخرسانة في هذه الخلطة 5.83 MPa. من هنا يتضح أنه كلما زادت نسبة الألياف في الخرسانة ذاتية الدمك تزداد مقاومة الانحناء للخرسانة ويعزي ذلك أن الألياف لها القدرة على الحد من الشروخ واتساعها.



الشكل رقم (13): تأثير ألياف البولي بروبيلين و الألياف الزجاجية على مقاومة الانحناء للخرسانة ذاتية الدمك

5 - الاستنتاجات و التوصيات:

- يمكن تنفيذ خرسانة ذاتية الدمك وبخصائص مقبولة وبمواد محلية مما يقلل من التكلفة الاقتصادية وذلك بالاستغناء عن بعض المواد المضافة.
- نوصي بإجراء دراسة على ديمومة الخرسانة ذاتية الدمك المسلحة بالألياف.
- إضافة الألياف إلى الخرسانة ذاتية الدمك يمكن أن تحقق شروط الخرسانة ذاتية الدمك عن طريق استخدام المدونات، قابلية المرور، (Passing ability)، مقاومة الانفصال الحبيبي (Resistance Segregation).
- كلما زادت نسبة ألياف البولي بروبيلين كلما زادت مقاومة الشد للخرسانة، حيث تزداد المقاومة تدريجياً و بنسبة بسيطة كلما زادت نسبة الألياف بها.

- أدت إضافة محتوى الألياف بنسبة 0.50% إلى نقص مقاومة الضغط بنسبة 4% بينما تزداد المقاومة كلما قلت نسبة الألياف في الخرسانة .
- كلما زادت نسبة ألياف البولي بروبلين مع الألياف الزجاجية للخرسانة كلما زادت مقاومة الانحناء للخرسانة ذاتية الدمك حيث إن نسبة الزيادة تتراوح بين 4% إلى 15% .
- يوصى بدراسة خواص الخرسانة التي لم يتم دراستها في هذا البحث و على سبيل المثال لا الحصر تأثير الألياف على نفاذية الخرسانة ونسبة الامتصاص والانكماش و درجات الحرارة المرتفعة.
- ألياف البولي بروبلين والألياف الزجاجية تؤثر على قابلية التعبئة (Filling ability)، و ذلك كلما زادت نسبة الألياف كلما زادت قابلية التعبئة و التشغيلية.

6 – المراجع:

- [1]. Douglas, R. P. (2004). "Properties of self-consolidating concrete containing type F fly ash: with a verification of the minimum paste volume method," M.Sc. Thesis, Northwestern University, USA.
- [2]. المواصفة الليبية رقم " 294 المياه المستعملة في الخرسانة، 1988.
- [3]. المواصفة القياسية الليبية 340 الاسمنت البورتلاندي العادي 1997.
- [4]. K. D. Raithby, J. W. Galloway and R. I. T. Williams, "Polypropylene-reinforced cement composites for surface reinforcement of concrete structures," International Journal of Cement Composites and Lightweight Concrete, vol. 3, no. 4, pp. 237-246, 1981.
- [5]. المواصفة الليبية رقم 255 "ركام الخرسانة من المصادر الطبيعية طريقة تعيين معامل الصدم 2006.
- [6]. المواصفة القياسية الليبية 49 ركام الخرسانة من المصادر الطبيعية 2002.
- [7]. British Standard Institution "BS1881 testing concrete part 104 method for determination of Vebe time " , London, 1983.
- [8]. The European Project Group " The European Guidelines for Self-Compacting Concrete Specification production and use " May, 2005.
- [9]. ACI Committee 211.4R-08.(2008)." Guide for Selecting Proportions for High-Strength Concrete Using Portland cement and Other Cementations Materials". American concrete institute Farmington Hills, USA, pp.1-29.
- [10]. BS1881: Part 116: 1983: Method for determination of compressive strength concrete cubes.
- [11]. BS 1881: Part 110: 1983: Method for making test cylinders from fresh concrete..
- [12]. European research project "Guide Lines for testing fresh self –Compacting Concrete", Sep 2005.
- [13]. American Society for Testing and Material , "ASTM C494 Type B, D and G: Standard Specification for Chemical Admixtures for Concrete", 2013.

العوازل الحرارية ودورها في تحسين البيئة الداخلية لمباني المناطق الصحراوية

* رواد احمد كريم¹، عز الدين حسن الشبوكي²، نزار أبو بكر غميص³
^{3,1} قسم الهندسة المدنية، كلية الهندسة رقدالين، جامعة صبراتة، صبراتة، ليبيا
² قسم الهندسة الصناعية، كلية الهندسة رقدالين، جامعة صبراتة، صبراتة، ليبيا
Rawad.krayem@sabu.edu.ly*

ABSTRACT

تشير الدراسات إلى أن نسبة الحرارة المنتقلة من الحوائط وأسقف المباني الواقعة في المناخ الصحراوي الجاف بحوالي 60-70% بينما تأتي البقية عن طريق فتحات النوافذ والأبواب ، وبالتالي فإن الحرارة المتسربة من الحوائط والأسقف تمثل الجزء الأكبر المراد التخلص منه عن طريق أجهزة التكييف ، ولذا فإن العزل الحراري يمثل أهمية كبيرة في تخفيض الحرارة المتسربة إلى داخل المبنى وبالتالي التقليل في استهلاك الطاقة الكهربائية المستخدمة في تبريده ، حيث إن الحصول على الكمية المطلوبة من الطاقة الكهربائية تعتبر من أهم المشاكل التي يعاني منها سكان مناطق الجنوب هذه السنوات.

Keywords:

المناخ الصحراوي ،
العزل الحراري ،
التوصيل الحراري ،
الامتصاص

وتهدف هذه الورقة إلى توضيح أهمية العزل الحراري في المباني التي انشأت في المناخ الصحراوي ودوره في تخفيض الطاقة الحرارية المستخدمة في أغراض التكييف من حيث معرفة العزل الحراري وفوائده والتعرف على مواد العزل الحراري الشائعة الاستعمال وخصائصها ثم معرفة كيفية استخدامها في المباني بطريقة سليمة بحيث تؤدي إلى ترشيد الطاقة . حيث استنتجنا من خلال هذا البحث إلى أن مواد العزل الحراري تمثل أهمية كبيرة في المنشآت الواقعة في المناخ الصحراوي الجاف لأنها تؤدي إلى خفض استهلاك الطاقة إذا استعملت في المكان المناسب من المبنى وبالكمية المناسبة.

1. المقدمة

إن العزل الحراري هو عبارة عن استخدام مواد لها خواص عازلة للحرارة ، بحيث تساعد في الحد من تسرب وإنتقال الحرارة من خارج المبنى إلى داخله صيفاً ومن داخله إلى خارجه شتاءً ، وتنقسم الحرارة التي تخترق المبنى والتي من المفترض التخلص منها باستعمال أجهزة التكييف للحفاظ على درجة الحرارة الملائمة إلى ثلاثة أنواع وهي الحرارة التي تخترق الجدران والأسقف والحرارة التي تخترق النوافذ والتي تنتقل كذلك عبر فتحات التهوية الطبيعية، وتكمن فوائد العزل في ترشيد استهلاك الطاقة الكهربائية حيث أثبتت التجارب العلمية أن تطبيق استخدام العزل الحراري في المباني السكنية والمنشآت الحكومية والتجارية والصناعية يقلل من استهلاك الطاقة الكهربائية ويحتفظ بدرجة الحرارة المناسبة للمبنى لمدة أطول دون الحاجة إلى تشغيل أجهزة التكييف لفترات زمنية طويلة ويرفع مستوى الراحة لمستخدمي المبنى، وإن من أهم العوامل التي تؤثر على اختيار مواد العزل الحراري المناسبة أن تكون المادة العازلة ذات مقاومة توصيل حراري منخفض وعلى درجة عالية لمقاومة لنفاذ الماء والإشعاع وأن تكون على درجة عالية في مقاومتها لامتصاص بخار الماء،

وكذلك يجب أن تكون ذات درجة عالية في مقاومتها للإجهادات الناتجة عن الفروقات الكبيرة في درجات الحرارة، وأن تكون ذات خواص ميكانيكية جيدة كارتفاع معامل المقاومة الإنضغاطية.

1.1 أهمية البحث

تكمن أهمية البحث في توفير مناخ يبعث الشعور بالراحة الحرارية داخل المباني وتحسين جودة البيئة الداخلية بواسطة زيادة كفاءة العزل الحراري والتي تساهم في خفض استهلاك الطاقة الكهربائية ، بالإضافة إلى حماية المسكن من التصدعات والتلف نتيجة الاجهاد المتبادل لارتفاع وانخفاض درجات حرارة عناصره.

2.1 مشكلة البحث

يشعر سكان المناطق الصحراوية إلى عدم الارتياح الحراري داخل المباني وذلك نتيجة لارتفاع الحرارة صيفا أو انخفاضها شتاء ويرجع السبب في ذلك إلى سهولة انتقال الحرارة من خلال الغلاف الخارجي المكون للمبني والذي ادي إلى استهلاك الطاقة الكهربائية بشكل كبير لاستعمالها في وسائل التكييف والتدفئة ، مما جعلنا في هذه الدراسة ان نحث علي اهم تقنيات العزل الحراري للمباني في المناطق الصحراوية من اجل توفير الراحة الحرارية داخل المبني والتقليل من معدلات استهلاك الطاقة الكهربائية.

3.1 المساهمة العلمية المتوقعة.

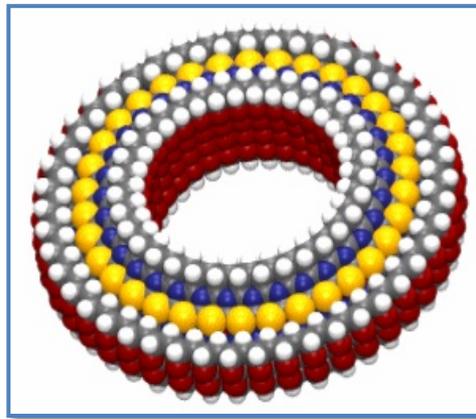
نساهم من خلال هذا البحث في التعريف بتقنيات العزل الحراري والتي لا نستخدمها كثيرا في بلادنا الحبيبة وطرق استخدامها علي اسطح المباني والتي ترجع بالنفع علي الترشيد في استهلاك الطاقة الكهربائية والتي نعاني منها كثيرا هذه السنوات ، بالإضافة إلى التقليل من التصدعات والتلف الموجود علي اسطح المباني نتيجة الاجهاد المتبادل لارتفاع وانخفاض درجات حرارة العناصر الانشائية.

2. مواد العزل الحراري

1.2 استخدام تقنية النانو في عزل المباني وحمايتها من المؤثرات الخارجية

تكنولوجيا النانو من أهم التقنيات الحديثة التي تعتبر من نتاج الثورة الرقمية وخاصة في مجال البناء والتشييد ، وهو العلم الذي يهتم بدراسة معالجة المادة علي المقياس الذري والجزيئي الذي ساهم في ابتكار تقنيات ووسائل جديدة تقاس أبعادها بالنانومتر وهو جزء من الألف من الميكرومتر أي جزء من المليون من المليمتر ، وعادة تتعامل تقنية النانو مع قياسات بين 0.1 إلى 100 نانومتر أي تتعامل مع تجمعات ذرية تتراوح بين خمس ذرات إلي ألف ذرة كما مبين في شكل (1) ، وهي أبعاد أقل كثيرا من أبعاد البكتريا والخلية الحية [1].

وبالتالي تسهم هذه التقنية في إنتاج مواد بناء ذات مميزات وخصائص حرارية وكهربائية وفيزيائية وكيميائية وميكانيكية فريدة ، فسوف تتمكن مباني النانو من مقاومة درجات الحرارة العالية والإشعاعات الضارة والحماية من الحرائق وأيضا ستمكن هذه المباني من صيانة ومعالجة أي تشققات وتصدعات بها مبكرا ، حيث ستدخل هذه التكنولوجيا في إنتاج مواد البناء لتحسين خصائصها ووظائفها مثل المواد المستخدمة في طلاء المباني والمواد المضافة للخلطة الإسمنتية والجسبية والبلاط وتحسين صناعة الزجاج والأخشاب وصناعات حديد الصلب ورفع كفاءة الطاقة في المباني لتجعلها خفيفة الوزن وأكثر قوة ومتانة وحماية للأسطح والجدران من الحرارة العالية والرطوبة والأشعة فوق البنفسجية بالإضافة ان هذه التقنية تعتبر صديقة للبيئة حيث ستساعد مواد البناء في التقليل من انبعاث غاز ثاني اكسيد الكربون في البيئة [1].



الشكل رقم (1) يوضح جزيئات النانو[1]

1.1.2 دور دهان النانو في العزل الحراري لأسطح المباني

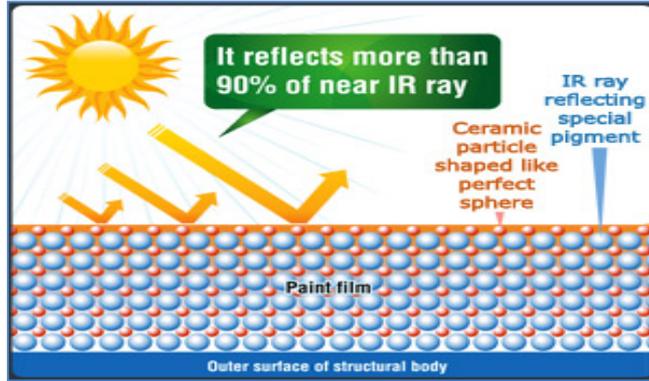
دهانات النانو تتكون من جسيمات متناهية الصغر جدا لها خصائص لامثيل لها مثل المرونة علي الاسطح والالتصاق السهل وله تقنية عالية بالتنظيف الذاتي بمجرد تعرضه للشمس ، حيث تشمل فكرة تطبيق النانو تكنولوجيا في الطلاء على تكثيف الأبخرة الكيميائية من اجل إنتاج طبقة تكون مرتبطة بالمادة الأساسية من اجل إنتاج سطح يحتمى على الحماية المطلوبة أو الخواص الوظيفية المحددة [2].

ولقد تم ابتكار نوع من الدهان يتكون من مزيج من مركبات حبيبات السيراميك الكروية الدقيقة الجوفاء والتي تعمل عمل العازل الحراري ، و هذا الدهان يستخدم على الجدران الداخلية والخارجية على حد سواء ولكنه اكثر فاعلية على الاسطح الخارجية حيث انه يعكس اشعة الشمس و يشتتها ، لذلك فإن دهان السيراميك يعتبر مادة فريدة لتوفير الطاقة الحرارية حيث

انها ذات تأثير مزدوج فهي تعكس و تشتت الحرارة بدهانها على الاسطح الخارجية كما انها تحتفظ بحرارة الفراغ الداخلي في ذات الوقت عند طلاء الأسطح الداخلية بها ، و من امثلة دهانات السيراميك ما يلي [2] :

1.1.1.2 دهانات ANZ السيراميك :

تعمل دهانات " ANZ " كدهان خارجي وداخلي للأسطح عامة سواء كانت حوائط داخلية أو خارجية المعرضة لأشعة الشمس المباشرة و غير المباشرة (شكل رقم 2) فيعمل الدهان على تكوين طبقة ذات مواصفات مصممة و مطورة بتكنولوجيا النانو تتكون من بلورات و جزيئات بالغة في الدقة كروية الشكل بدون فراغات مرتبة بشكل هندسي دقيق تعمل على تشتيت و انعكاس أشعة الشمس و حرارتها عن السطح المدهون ، وبالتالي تنخفض درجة الحرارة الداخلية للسطح المدهون عن الدرجة الطبيعية بفارق يزيد عن (20 درجة مئوية)، حيث لا يتطلب أي دهانات خارجية أو داخلية بطانة عند استخدام دهان السيراميك ANZ ، كما أنه في حالة استخدام دهان السيراميك ANZ كعازل للحرارة والرطوبة والأملاح فإنه يغني عن استخدام دهان خارجي عليه كما في حالة الدهان العازل للحرارة العادي ، لأن دهان ANZ يتميز بتنوع ألوانه وتغطيته السريعة [2].



الشكل رقم (2) دهان ANZ يعمل علي تشتيت الأشعة الشمسية[2]

2.1.1.2 دهان السيراميك سوپر ثيرم Super Therm [2]:

دهان بسمك 250 ميكرون (ربع ميللتر) يعمل كعازل حراري لجميع الاسطح وتم استخدامه لما فيه من مواصفات وخصائص فريدة من نوعها وهي :

1. مادة خزفية مصممة للعمل كغطاء عازل و عكس الحرارة وتخفيض تكلفة الطاقة.
2. تعكس 95% من أشعة الشمس فتعمل علي توفير الطاقة من 20% إلى 70% .
3. الدهان الخزفي الوحيد الحاصل علي شهادة الجودة لكود البناء الأمريكي كمادة عازلة.
4. مادة صمغية متعددة تحتوي علي مادة مطاطية لتأكيد المتانة وطول العمر.
5. مقاومة الحريق من الدرجة الاولى وفي حالة اندلاع حريق تعمل علي عدم انتشاره.
6. تعمل علي عزل 68% من الموجات الصوتية.
7. لا تتأثر بتمدد وانكماش الاسطح المعدنية مع تغيير درجات الحرارة كما مبين في شكل (3).
8. تستخدم علي الاسطح المعدنية والخرسانية.
9. صديقة للبيئة مختبرة ومصروح بها من وزارة الزراعة والصحة الامريكية USDA كمادة آمنة.
10. عمرها الافتراضي من 15 إلى 20 سنة علي الأسطح والوجهات الخارجية في الظروف العادية [2].



الشكل رقم (3) يوضح مدي تماسك جزيئات الدهان المعالج بتقنية النانو[2]

2.1.2 استخدام تقنية النانو في تقليل كمية الحرارة الداخلة من زجاج النوافذ

- إن أغلبية الزجاج المستخدم في الإنشاءات يستخدم علي السطح الخارجي من المباني ولذلك فإن التحكم بكمية الحرارة والضوء النافذين إلى المبني الزجاجي يعتبر من الأمور الهامة الواجب مراعاتها، وبالتالي فإن هناك العديد من تطبيقات النانو في مجال خامة الزجاج وذلك بهدف مشاركة هذه المادة في التحكم في البيئة الداخلية للمبنى وتعتبر مادة ثاني أكسيد السيليكا (SiO_2) من أهم المواد التي تستخدم كطبقة طلاء بين طبقات الزجاج وتعمل على [3]
1. منع التسرب الحراري لأشعة الشمس عبر النوافذ وذلك باستخدام تكنولوجيا " الثروموكروماتيك " و هي عبارة عن طلاء طبقة رقيقة جدا تعمل على العزل الحراري مع توفير الاضاءة المناسبة مما يترتب على ذلك من توفير في استهلاك الطاقة كما مبين في شكل (4).
 2. اختزان قدر من الحرارة ثم بثها مرة اخرى الى داخل الفراغ الداخلى في حالة انخفاض درجة الحرارة فيعمل كجهاز تدفئة.
 3. عاكس لأشعة الشمس الضارة فوق البنفسجية .
 4. خاصية التنظيف الذاتي بتحليل الإتساخات العضوية و مقاومة الضباب .
 5. تقنية التحكم في تغيرات شدة الاضاءة .

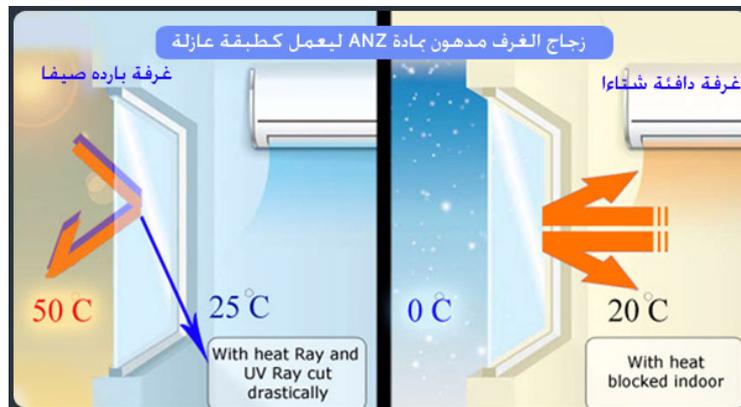


الشكل رقم (4) يوضح الفرق بين زجاج استخدم فيه تقنية النانو وزجاج لم تستخدم فيه تقنية النانو[3]

1.2.1.2 استخدامات دهان السيراميك ANZ -G للزجاج لتخفيض تكلفة استهلاك الكهرباء [2]:

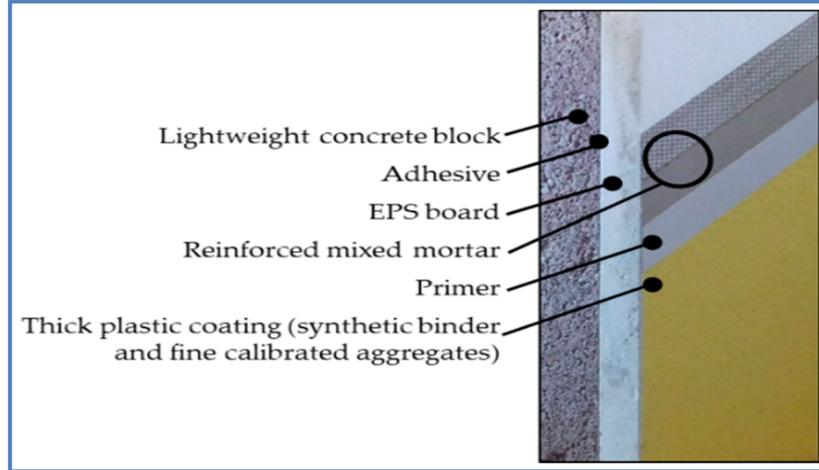
يساهم دهان السيراميك ANZ في تحقيق التصميم الداخلى المستدام لتمييزه بالمميزات الآتية :

1. توفير أكثر من 20% من الطاقة المستخدمة في الفراغات الداخلية في حالة استخدامه في دهان زجاج النوافذ او الحوائط الزجاجية الساترة ،حيث يعمل كطبقة عازلة حرارياً ، ويظهر الأثر الاقتصادي الكبير في الأماكن التي تعتمد اعتمادا كليا على الطاقة ، علماً بان انخفاض درجة التبريد أو التدفئة درجة واحدة فقط يؤدي إلى انخفاض 10% من استهلاك الطاقة الكهربائية كما مبين في شكل (5) .
2. باستخدام دهانات ANZ يتم تثبيت أكثر من 80% من أشعة الشمس فوق الحمراء قريبة المدى (NIR) .



الشكل رقم (5) دهان السيراميك ANZ يعمل كطبقة عازلة للحرارة عند دهانه على الاسطح الزجاجية

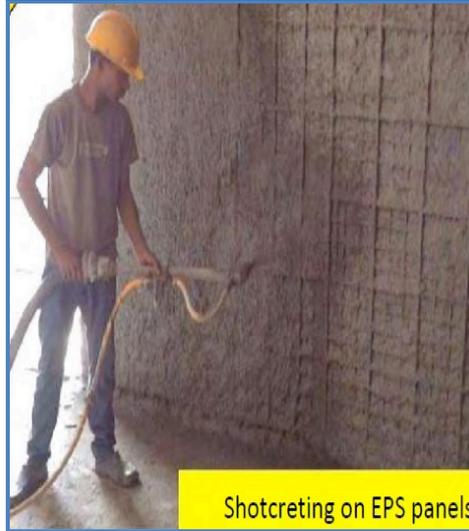
2.2 استخدام نظام مركب العزل الحراري الخارجي (ETICS) باستخدام لوح الفلين العازل (ICB) أو البوليسترين المدد Extender Poly Styrene (EPS)



الشكل رقم (6) نظام العزل الحراري الخارجي المركب (ETICS) [4]

يوجد العديد من المواد التي يتم استخدامها في العزل الحراري ومن أهمها مادة الفلين (ICB) وهي مادة متجددة وطبيعية 100٪ وقابلة لإعادة التدوير بالكامل مصنوعة من الفلين الطبيعي دون مواد لاصقة كيميائية أو إضافات يمكن تطبيقها على غلاف المباني الجديدة والمجددة لتحسين كفاءتها في استخدام الطاقة، ويمتاز الفلين (ICB) بالمرونة والتوصيل الحراري المنخفض وتتراوح بين (0.040 - 0.045 واط / مترمربع لكل درجة مئوية) أي انه يمتاز بخصائص العزل حراري والصوتي، أما مادة البوليسترين (Polystyrene) فهي تصنف من البوليمرات وهي من أهم التقنيات المتداولة عالميا في أعمال التغليف والتعبئة والعزل الحراري للمباني ويتم إنتاج هذه المادة من بلمرة الستايرين الخام و هو مركب عضوي من فصيلة البتروكيماويات وتصنف مادة البوليسترين من عوازل الدرجة الأولى عالميا ، ومن أهم خصائص مادة البوليسترين أنه عازل جيد للحرارة نسبة لتكوينها الخلوي الذي يعمل على طرد الحرارة وذلك حسب كثافته ، ويعتبر البوليسترين مادة خفيفة الوزن وسهلة الحمل والنقل كما أنها مادة تتحمل قوة الانضغاط وقد أثبتت الدراسات انه كلما زادت الكثافة زادت قوتها وزادت قوة انضغاطها، أما عن العزل في المباني فهي تصون المبني من عدة ظواهر طبيعية وكيميائية، ويتم تصنيع مادة البوليسترين الحبيبي الممدد بع على ثلاثة مراحل وهي [4]:

1. مرحلة التمدد الأولى للحبيبات.
 2. مرحلة إنضاج الحبيبات الممددة (إحلال جزئي للهواء بدلاً من الغاز).
 3. مرحلة القولية والتي يجري فيها تعبئة قوالب الإنتاج النهائي بالحبيبات الممددة، ثم يتم حقن القوالب المغلقة ببخار الماء والذي يعمل على تمدد الحبيبات وتمييع سطوحها مما يؤدي إلى التحامها.
- ويعد نظام العزل الحراري الخارجي المركب (ETICS) أحد حلول العزل الحراري الخارجي المستمر لجدران المباني، حيث يقوم بتقليل الجسور الحرارية للمبنى من خلال تحسين معامل النقل الحراري في جميع أنحاء الواجهة الذي يعمل على تحسين الراحة الحرارية الداخلية في الشتاء والصيف على حد سواء وكذلك يرفع من كفاءة المباني في استهلاك الطاقة، ويشتمل نظام ETICS على لوح عزل مطبق على الركيزة (مثبت ميكانيكياً) ، حيث يتم تطبيق طبقة رقيقة واحدة أو طبقتين من التقوية المسلحة كما مبين في شكل (6) ، ويمتاز هذا النظام بخفض وزن الجدران والأحمال الميتة على المبنى وتحسن نفاذية الواجهات وله أيضا مقاومة عالية للتشقق والتفتت، بالإضافة إلى وجود سماكات وتركيبات وألوان مختلفة [4].



الشكل رقم (7) كيفية تثبيت مادة العزل الحراري البوليميرين المقوي بالشبكة المعدنية علي الحائط [5]

3. الاستنتاجات والتوصيات

1.3 الاستنتاجات

1. يجب أن تكون المادة العازلة المستخدمة ذات مقاومة توصيل حراري منخفض وعلى درجة عالية لمقاومة لنفاذ الماء والإشعاع وأن تكون على درجة عالية في مقاومتها لامتصاص بخار الماء.
2. إن تطبيق استخدام العزل الحراري في المباني السكنية والمنشآت الحكومية والتجارية والصناعية يقلل من استهلاك الطاقة الكهربائية.
3. تعمل تقنية النانو على تعديل خواص المواد الطبيعية والصناعية عن طريق تغيير الترتيب الذري للمادة مما يؤدي الى التخلص من الخصائص السلبية او إضافة خصائص اخرى لم تكن موجودة من قبل.
4. تعمل تقنية النانو على تحقيق التصميم الداخلى المستدام من خلال تحسين كفاءة اداء المباني القائمة وتطوير و تحسين خواص الخامات التي تستخدم في تنفيذ التصميم الداخلى.
5. يعتبر استخدام منتجات النانو العازلة من أهم الحلول التي تعمل على تحسين كفاءة المباني القائمة بالفعل فهي عبارة عن طبقة رقيقة جداً غير مرئية و غير سامة يمكن رشها على أى خامات تتوفر مستوى عالي من العزل الحراري او العزل المائي كما تعمل على توفير بيئة داخلية مريحة.
6. تقنية النانو تساهم بشكل كبير في ترشيد استهلاك الطاقة الكهربائية.
7. منع التسرب الحراري لأشعة الشمس عبر النوافذ وذلك باستخدام تكنولوجيا " الثرموكمروماتيك " و هي عبارة عن طلاء طبقة رقيقة جدا تعمل على العزل الحراري مع توفير الاضاءة المناسبة.
8. نظام العزل الحراري الخارجي المركب (ETICS) أحد حلول العزل الحراري الخارجي المستمر لجدران المباني ويرفع من كفاءة المباني في استهلاك الطاقة.

2.3 التوصيات

1. يجب الاهتمام بالعزل الحراري في المباني فهي تصون المبني من عدة ظواهر طبيعية و كيميائية.
2. يجب اختيار مواد العزل الحراري المناسبة بحيث تكون المادة العازلة ذات مقاومة توصيل حراري منخفض وعلى درجة عالية لمقاومة لنفاذ الماء والإشعاع.
3. استبدال العديد من تقنيات المباني المستدامة و التي يتم استخدامها حالياً بتقنيات النانو لما توفره من امكانيات و تطبيقات خاصة و انه في المستقبل القريب سوف تصبح خامات النانو اقتصادية في التكاليف.
4. استخدام خامات النانو العازلة في معالجة المباني القائمة بالفعل لحمايتها و العمل على ترشيد استهلاك الطاقة بها.
5. تطوير هذا البحث في الدراسات القادمة التي تختص بهذا النوع من المواضيع , وإجراء الاختبارات المعملية في ليبيا ومقارنتها بنتائج الأبحاث العالمية.

4. الخلاصة

إن استخدام العزل الحراري في المباني يعتبر من أهم الإجراءات التي ينبغي الأخذ بها لضمان تحقيق الراحة الحرارية للسكان والاستخدام الأمثل للطاقة الكهربائية في المباني التي تتطلب التكييف وذلك لتفادي استهلاك الطاقة، و من أهم مزايا هذا استخدام العزل الحراري الاثار الإيجابية التي تنعكس على مختلف قطاعات المجتمع والدولة، حيث توفر علي الدولة تكاليف الاستثمار الباهظة في قطاع الكهرباء والمتمثل في بناء محطات توليد الطاقة وزيادة ساعات شبكات نقل وتوزيع الكهرباء. هذا بالإضافة إلى تقليل الاعتماد على مصادر الطاقة الأولية التي تزداد الصعوبة في توفيرها مع مرور الوقت . كما يساهم البرنامج بوضوح في تقليل الانبعاثات الغازية الملوثة ، كما ان برنامج العزل الحراري في المباني يعمل علي

خلق فرص عمل جديدة وتشجيع السوق المحلية في مجال المواد والأنظمة العازلة للحرارة مما يساهم في دفع عجلة الاقتصاد بشكل إيجابي.

5. المراجع

- [1] رزين فهد الفوزان . تقنية البناء بالنانو . كلية الهندسة – جامعة القصيم-المملكة العربية السعودية . 2012 , 2 .
- [2] علا محمد سمير إسماعيل . استخدام تكنولوجيا النانو الخضراء في تحقيق التصميم الداخلي المستدام . المؤتمر الدولي الثاني لكلية الفنون التطبيقية – جامعة حلوان (لبنان) . 2012 , (11 , 13) .
- [3] Will Soutter, "Nanotechnology in Green Construction," Stevens Institute of Technology, Manchester, UK, Article 2012.
- [4] José D. Silvestre, André M. P. Castelo, José J. B. C. Silva, Jorge M. C. L. de Brito, and Manuel D. Pinheiro "Retrofitting a Building's Envelope: Sustainability," Applied Sciences, vol. 9, no. 7, p. 1285, March 2019.
- [5] Ministry of Housing and Urban Poverty, "Manual for Expanded Polystyrene," CSIR – Central Building Research Institute, Roorkee , Article S.E(G)/ 0605, 2017.

تأثير نوع الملدنات الفائقة على التشغيلية والخواص الميكانيكية للخرسانة ذاتية الدمك

*أ. عبدالرؤوف أبوذينة¹، م. أيمن رجب محمد عبدالحفيظ²، محمد النفاتي الهادي³

¹عضو هيئة تدريس - قسم الهندسة المدنية، كلية الهندسة، جامعة الزيتونة، ترونة، ليبيا

^{2,3}طالبة بكالوريوس - قسم الهندسة المدنية، كلية الهندسة، جامعة المرقب، القره بولي، ليبيا

¹Raouf.aboudina@yahoo.com, ²ar4440809@gmail.com, ³Hmady.tribole@gmail.com

ABSTRACT

تعتبر الخرسانة ذاتية الدمك من الانواع الحديثة وتمتاز هذه النوعية بالانسيابية العالية واللزوجة وليست في حاجة للدمك، وهي قابلة للانسياب تحت تأثير وزنها الذاتي. تتعكس هذه المميزات في أداية الخرسانة ذاتية الدمك حيث تجعلها قادرة على التشكل والمرونة خلال القوالب الإنشائية الضيقة أو المقاطع الإنشائية ذات التسليح المكثف مع تحقيق درجة دمك عالية دون الحاجة إلى استخدام الهزازات اثناء عملية الدمك الخارجي وكذلك بدون حدوث انفصال حبيبي أو نضوح في الخرسانة .

Keywords:

تتناول هذه الدراسة استخدام انواع مختلفة من الملدنات الفائقة (Super plasticizer) المتوافرة في السوق الليبي لإنتاج الخرسانة ذاتية الدمك. حيث تم استخدام 5 أنواع من الملدنات الفائقة المتوافرة في السوق الليبي وهي (FOSROC-AURAMIX 326C) و (SIKA VISCOCRETE-TEMPO 12) و (agel-Fx6) و (agel-Technohyper N) و (ليبييا لكيماويات البناء- لييومنت أف أف) بنسب متقاربة للحصول على نتائج غير متفاوتة. وذلك لدراسة تأثير هذه الملدنات الفائقة على الخرسانة ذاتية الدمك في حالتها الطرية والمتصلبة. حيث وجد أن الخلطة الخرسانية التي تحتوي على الملدن الفائق (FOSROC-) AURAMIX 326C أعطت أكبر قطر انسياب مقارنة بالخلطات الأخرى. بينما تبين أن الخلطة الخرسانية ذاتية الدمك التي تحتوي على الملدن الفائق (agel-Technohyper N) قد حققت أفضل النتائج في كل من اختبار صندوق حرف L، و اختبار مقاومة الضغط للمكعبات الخرسانية، واختبار مقاومة الانحناء للمنشآت الخرسانية مقارنة بباقي الخلطات.

1. المقدمة

الخرسانة ذاتية الدمك هي نوع من الخرسانات ذات سيولة عالية مما يسمح لها بالدمك الذاتي والتوغل في المساحات الضيقة كثيفة التسليح [1] وعلى الرغم من زيادة لدونة الخرسانة فإنها تحافظ على الجودة العالية التي تتطلبها صناعة البناء اليوم. مما يتيح المرونة بالإضافة إلى المقاومة العالية. ترجع مرونة الخرسانة ذاتية الدمك إلى أن هناك العديد من الإضافات يمكن إضافتها عند تصميم الخلطة الخرسانية لعمل منتج نهائي يناسب النتيجة المطلوب الوصول إليها من ناحية اللدونة والمقاومة. وقد جري تحليل العديد من الحالات المتعلقة بتطبيقات الخرسانة ذاتية الدمك في الفترة من عام 1993 إلى الوقت الحاضر، حيث تزايد استخدام الخرسانة ذاتية الدمك في العديد من البلدان. أثبتت هذه التحاليل انه يجب اعتبار الخلطات الخرسانية ذاتية الدمك من الخلطات ذات مجال واسع الاستخدام ومناسبة للعديد من الاستخدامات [2]. إضافة الملدن المتفوق تستخدم في جميع انواع الخرسانة ذاتية الدمك، لأنه المكون الأساسي الذي يمنح الخرسانة ذاتية الدمك لدونتها العالية، وهذه الإضافة تقلل من كمية الماء المستخدم وتحافظ على نفس كثافة الخرسانة التقليدية [3].

في دراسة سابقة تم دراسة تأثير إضافة ثلاثة أنواع مختلفة من المواد الفعالة للملدن المتفوق وهي: السلفونيت النفثالين الفورمالديهايد (SNF) (Sulphonated Naphthalene Formaldehyde) ، بولي كربوكسيل الأثير (PCE) (Poly Carboxylate Ether) و بولي كربوكسيل الأثير المعدل (MPCE) Modified (Carboxylate Ether) على قابلية التشغيل والخواص الميكانيكية للخلطات الخرسانة ذاتية الدمك (SCC) ، قد لوحظ في إختبارات الخرسانة في حالتها اللدنة والمتصلبة أن أفضل خلطة هي المحتوية على المادة الفعالة (MPCE) تليها بفارق بسيط الخلطة المحتوية على المادة الفعالة (PCE). بينما كانت أضعف نتيجة من الخلطة المحتوية على المادة الفعالة (SNF) [9].

2. أهمية الدراسة :

تتناول هذه الدراسة استخدام انواع مختلفة من الملدنات الفائقة (Super plasticizer) المتوافرة في السوق الليبي لإنتاج الخرسانة ذاتية الدمك، حيث انه لا بد من استخدام الملدن الفائق لإنتاج الخرسانة ذاتية الدمك. حيث يؤدي إستخدام

الملدنات إلى زيادة تشغيلية الخرسانة وسيلولتها، وتخفيض نسبة الماء إلى الإسمنت في الخلطة، وتقليل امكانية حصول انفصال حبيبي أو نضح.
3. أهداف الدراسة :

تهدف هذه دراسة لمعرفة تأثير انواع مختلفة من الملدنات الفائقة (Super plasticizers) على الخواص اللدنة والمتصلية للخرسانة ذاتية الدمك. بالإضافة إلى التحقق من إمكانية الاستفادة من مزايا هذه الملدنات الفائقة في تحسين الخصائص الهندسية للخرسانة ذاتية الدمك ومعرفة تأثيرها على الانسيابية ومقاومة الضغط والانحناء. يتميز هذا العمل بتقديم بعض المعلومات والبيانات المتعلقة بتأثير تغير نوع الملدن الفائق على خواص الخرسانة ذاتية الدمك، أيضا للفت الانتباه وإمكانية استخدام الخرسانة ذاتية الدمك في صناعة التشييد.
4. البرنامج العملي :

في هذه الدراسة استخدمت الملدنات الفائقة بنسبة 1.8% من وزن الإسمنت لجميع الخلطات بإستثناء الخلطة الخرسانية المحتوية على الملدن (AURAMIX 326C)، حيث تم إستخدام نسبة 1.4% من وزن الإسمنت. ذلك لأنها أعلى نسبة مسموح استخدامها حسب النشرة الفنية الخاصة بهذا الملدن الفائق، بينما نسبة 1.4% لم تعطي نتائج داخل المواصفات بالنسبة للخلطات الأخرى، لهذا تم إعتداد نسبة 1.8% لهذه الخلطات. الجدول رقم (1) يوضح نسب ومقادير الملدنات المستخدمة في كل خلطة.

جدول (1) يوضح المقادير المستخدمة لكل ملدن

اسم الملدن	رمز الخلطة الخرسانية	نسبة الملدنات الى وزن الاسمنت في الخلطة الخرسانية	مقدار الملدنات في الخلطة الخرسانية	السعر في السوق الليبي
TEMBO 12	TEMP-SCC	1.8%	9 كجم/م ³	-
AURAMIX 326 C	AURA-SCC	1.4%	7 كجم/م ³	7 دينار / لتر
ليبومنت أف-أف	LIBO-SCC	1.8%	9 كجم/م ³	7 دينار / لتر
FX6	FX6-SCC	1.8%	9 كجم/م ³	-
Technohyper N	TECH-SCC	1.8%	9 كجم/م ³	7.5 دينار / لتر

تم الحصول على الملدنين الفائقين (Sika Viscocrete-TEMPO 12) و (ageI-Fx6) مجاناً من شركة الاتحاد العربي للمقاولات، قد تمت معالجة جميع الخلطات في الماء للفترتين 7 و 28 يوم، وتم عمل اختبار مقاومة الضغط عند الفترتين 7 و 28 يوم، واختبار مقاومة الانحناء عند الفترة 28 يوم.

1.4 مكونات الخلطة الخرسانية:

الإسمنت:

الإسمنت البورتلاندي العادي المستخدم في هذه الدراسة من إنتاج مصنع الاتحاد العربي للمقاولات بمدينة زليتن، وأجريت الاختبارات الخاصة بالتركيب الكيميائي والمعدني و الخواص الفيزيائية، وكانت جميعها تقع ضمن الحدود المسموح بها في المواصفات البريطانية (BS12:1996) [4].

الركام الناعم:

الركام الناعم المستخدم في هذه الدراسة تم توريده من محاجر بمنطقة غريان، وأجريت الفحوصات الخاصة بالرمل وحسب حدود المواصفات البريطانية (BS882:1992) [5]، تشير الفحوصات إلى ان الركام الناعم مطابق للمواصفة من حيث التدرج. علماً بأن الوزن النوعي ونسبة الامتصاص ونسبة المواد الناعمة والكثافة هي 2.631 و 0.1% و 1% و 1500 كجم / م³ على التوالي.

الركام الخشن:

تم توريد الركام الخشن من محاجر في مدينة غريان. تم اختبار التحليل المنخلي للركام الخشن والنتائج كانت ضمن الحدود والمواصفات البريطانية المعتمدة (BS 1992:882) [5]. علماً بأن الوزن النوعي ونسبة الامتصاص ونسبة المواد

الناعمة ومعامل الصدم و لوس انجلوس والكثافة هي 2.579 و 2.5 و 0.55% و 18.4% و 26.3% و 1430 كجم/م³ على التوالي.

الماء:

تم استخدام الماء الصالح للشرب في اعداد الخلطات الخرسانية والذي تم اختباره في مركز البحوث الصناعية بتاجوراء، وأثبتت نتائج الاختبار مطابقتة لحدود المواصفات الليبية رقم 294 لسنة 1988 ف [6].

الإضافات:

تم استخدام عدد 5 من المدونات الفائقة المتواجدة في السوق الليبي كإضافات كيميائية لتوفير خواص التشغيلية للخرسانة ذاتية الدمك. الجدول (2) يوضح بعض خصائص المدونات الفائقة المستخدمة.

الجدول (2) يوضح بعض الخصائص للمدونات الفائقة المستخدمة

اسم الشركة المصنعة	اسم المدون الفائق	نسبة الاستخدام الى وزن الاسمنت	الأساس الكيميائي	اللون
SIKA	Tembo 12	تتراوح بين 0.2-3%	acrylic copolymer	سائل بني فاتح
FOSROC	AURAMIX 326C	تتراوح بين 0.8-1.4%	ether polycarboxylate	سائل بني
ليبيا لكيماويات البناء	ليبومنت – أف أف	تتراوح بين 0.6-3%	Naphthaline Formaldehyde Sulfonate	سائل بني داكن
Agel	Fx6	تتراوح بين 0.6-2%	polycarboxylate	سائل بني داكن
Agel	technohyper N	تتراوح بين 0.6-2%	polycarboxylate	سائل بني داكن

2.4 تصميم وخط الخلطة الخرسانية:

اعتمدت هذه الدراسة على اعداد خلطات خرسانية ذاتية الدمك باستخدام اسمنت بورتلاندي عادي وركام ناعم وخشن وماء وكانت نسبة الخلط حسب المعايير لتصميم الخلطة ونسبة الماء إلى الإسمنت 0.488% وكان المتغير هو نوع ونسبة المدون وتم استعمال نسب مدونات تتراوح من 1.4% إلى 1.8% من وزن الإسمنت، والجدول رقم (3) يوضح نسب وكميات الخلطة الخرسانية المستخدمة. تم خلط المكونات بأسلوب مشابه لما قام به (Grunewald & Walraven) [7] حيث قاما بخلط الاسمنت والرمل لمدة 10 ثواني ثم يضاف ماء الخلط والمدون المتفوق ويستمر الخلط لمدة 110 ثانية وبعد ذلك يضاف الركام ويستمر الخلط لمدة 60 ثانية.

جدول (3) يوضح المقادير التي استخدمت في الخلطة الخرسانية

الاسمنت كجم/م ³	نسبة الماء الي الاسمنت	الركام الخشن كجم/م ³	جرينيليا كجم/م ³	الركم الناعم كجم/م ³
500	880.4	295.8	549.3	845

5. نتائج اختبار الخرسانة في الحالة اللدنة:

الخواص اللدنة لخلطات الخرسانة ذاتية الدمك تم اختبارها باستخدام اختبار الانسياب لقياس الانسيابية واختبار (صندوق حرف L) لقياس قابلية المرور.

تأثير نوعية الملدنات على الخرسانة في الحالة اللدنة:

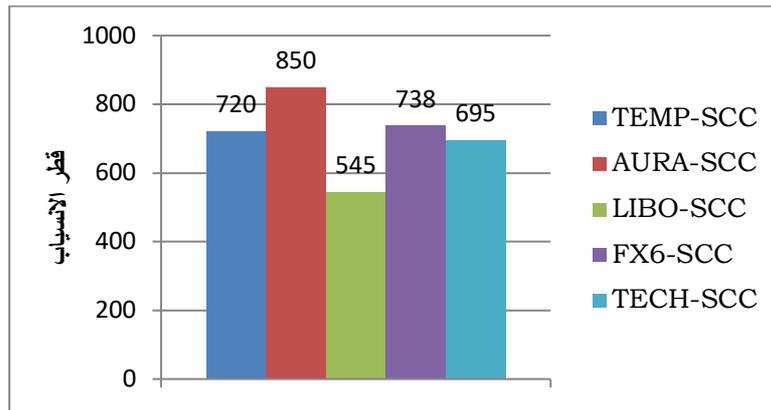
تأثير تغير نوع الملدن المستخدم على الخرسانة ذاتية الدمك أعطى نتائج مختلفة على الخواص اللدنة للخرسانة، فكانت أغلب الخلطات الخرسانية ذات سيولة وانسيابية من الجيدة الى الممتازة، بينما كانت احدى الخلطات الخرسانية ذات سيولة وانسيابية ضعيفة جدا، جعلتها خارج حدود المواصفات المستخدمة (EFNARC) [8] . وتم تدوين النتائج في جدول (4).

الجدول 4 نتائج اختبارات الخرسانة في الحالة اللدنة

رقم الخلطة	رمز الخلطة	الانسياب ملم	حدود المواصفة ملم	صندوق L - H2/H1	حدود المواصفة
1	TEMP-SCC	720	850-550	0.77	H2/H1
2	AURA-SCC	850	850-550	0.75	H2/H1
3	LIBO-SCC	545	850-550	0.69	H2/H1
4	FX6-SCC	738	850-550	0.77	H2/H1
5	TECH-SCC	695	850-550	0.85	H2/H1

اختبار الانسياب:

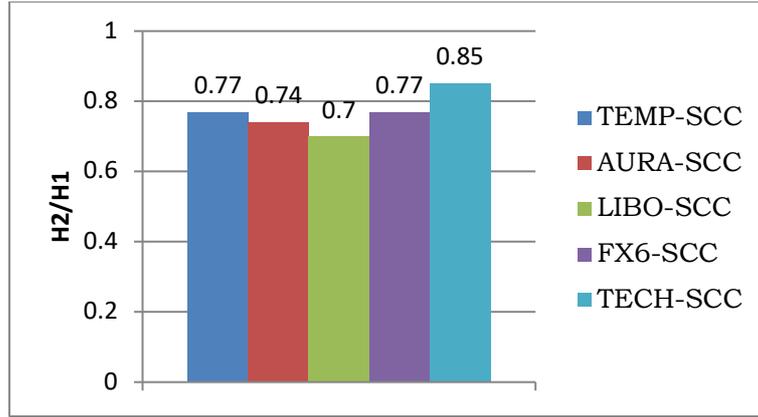
الشكل رقم (1) يوضح نتائج اختبار الانسياب لجميع الخلطات. سجلت جميع الخلطات قطر انسيابي كان ضمن حدود المواصفات الأوروبية (EFNARC) [8] إلا الخلطة (LIBO-SCC) كان خارج المواصفات الأوروبية (EFNARC) [8] واطهرت الخلطة (AURA-SCC) أكبر قطر انسياب مقارنة بالخلطات الأربعة الأخرى المستهدفة في هذه الدراسة. حيث أعطت قطر اكبر من الخلطات (TEMP-SCC و FX6-SCC و TECH – SCC و LIBO-SCC) بنسبة زيادة قدرها (115% و 118% و 122% و 156%) على التوالي.



شكل 1 تأثير نوع الملدن على اختبار الانسياب

اختبار صندوق حرف L:

جميع الخلطات اظهر نسبة ميل (H2/H1) ضمن حدود المواصفات الأوروبية إلا الخلطة (LIBO-SCC) كان خارج المواصفات الأوروبية (EFNARC) [8]. الشكل (2) يوضح نتائج اختبار صندوق L، وكانت الخلطة (TECH-SCC) قد اعطت اقل ميلان (H2 / H1) مقارنة بالخلطات الأربعة الأخرى المستهدفة في هذه الدراسة. حيث أعطت زيادة في الميل أكبر من الخلطات (TEMP-SCC و AURA-SCC و LIBO-SCC و FX6-SCC) بنسب قدرها (110.4% و 113.3% و 123.2% و 110.4%) على التوالي.



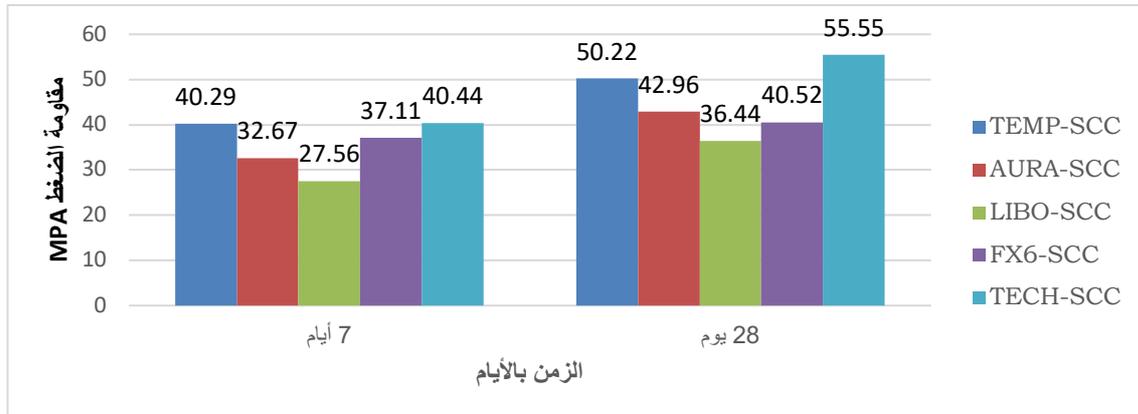
شكل 2 تأثير نوع الملدن على اختبار صندوق-L

نتائج اختبار مقاومة الضغط:

لقياس مقاومة الضغط فقد تم اختبار عدد 30 مكعب ، والجدول رقم (5) و الشكل رقم (3) يوضح نتائج تأثير تغير نوع الملدن على مقاومة الضغط.

جدول 5 يوضح نتائج اختبار مقاومة الضغط لجميع الخلطات

مقاومة الضغط لوحدة (Mpa)			رمز الخلطة	رقم الخلطة
نسبة الزيادة من 7 أيام إلى 28 يوم	28 يوم	7 أيام		
%19.77	50.22	40.29	TEMP-SCC	1
%23.95	42.96	32.67	AURA-SCC	2
%24.37	36.44	27.56	LIBO-SCC	3
%8.42	40.52	37.11	FX6-SCC	4
%27.20	55.55	40.44	TECH-SCC	5



شكل 3 يوضح تأثير الملدن على مقاومة الضغط لجميع الخلطات

من المعاينة يتضح أن الخلطة (TECH-SCC) قد اعطت اعلى مقاومة ضغط خلال فترتي المعالجة 7 أيام و28 يوم، بينما كانت الخلطة (LIBO – SCC) هي الأقل مقاومة بين جميع الخلطات التي تمت دراستها خلال فترتي المعالجة 7 و 28 يوم. حيث تم احتساب قيم مقاومة الضغط من متوسط عدد 3 عينات لكل خلطة.

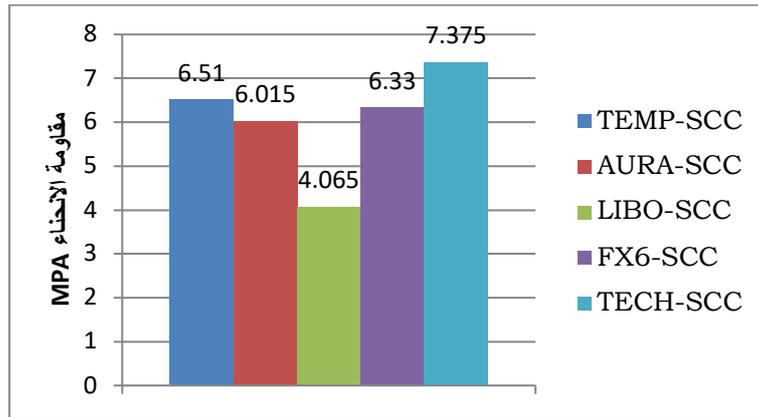
نتائج اختبار مقاومة الانحناء:

عدد 15 المنشور تمت معالجتها في الماء حتى موعد الاختبار المقرر في هذه الدراسة وهو 28 يوم حيث تم احتساب قيم المقاومة الانحناء من متوسط عدد 3 عينات لكل خلطة الجدول (5) يوضح نتائج اختبار مقاومة الانحناء.

جدول 5 يوضح نتائج اختبار مقاومة الانحناء

مقاومة الإنحناء (Mpa) 28 يوم	رمز الخلطة	رقم الخلطة
6.51	TEMP-SCC	1
6.015	AURA-SCC	2
4.065	LIBO-SCC	3
6.33	FX6-SCC	4
7.375	TECH-SCC	5

النتائج المتحصل عليها توضح أن الخلطة (ECHN-SCC) اظهرت أعلى قيمة لمقاومة الإنحناء مقارنة مع جميع الخلطات الأخرى خلال فترة المعالجة 28 يوم، حيث أعطت قيمة أكبر من الخلطات (TEMP-SCC و AURA-SCC و LIBO-SCC و FX6-SCC) بنسبة زيادة قدرها (113.3% و 122.6% و 181.4% و 116.5%) وشكل (4) يوضح نتائج اختبار مقاومة الانحناء.



شكل 4 تأثير نسب الملدنات على مقاومة الإنحناء

6. الخلاصة :

من خلال الدراسة العملية التي أجريت وتحليل النتائج يمكن تلخيص النقاط التالية:

- 1- تباين جودة الملدنات الفائقة الموجودة في السوق الليبي.
- 2- الخلطة AURA-SCC المحتوية على المادة الفعالة PCE اعطت افضل قيمة انسياب بين جميع الخلطات المستهدفة في هذه الدراسة، مما يتوافق مع نتائج الدراسة السابقة [9].
- 3- الخلطة TECH-SCC المحتوية على المادة الفعالة PCE اعطت افضل قيمة ميل في اختبار صندوق L-، وافضل نتيجة في اختبارات مقاومة الضغط ومقاومة الانحناء مما يتوافق مع الدراسة السابقة [9].
- 4- الخلطة LIBO-SCC المحتوية على المادة الفعالة SNF أعطت اقل النتائج في اختبارات الخرسانة في حالتها اللدنة والمتصلبة، مما يتوافق أيضا مع الدراسة السابقة [9].
- 5- الملدن لبيومنت – أف أف، تم به عمل الخلطة LIBO-SCC التي أعطت نتائج خارج المواصفات في إختبارات اللدونة للخرسانة ذاتية الدمك، وأقل النتائج في كل من إختباري مقاومة الضغط والانحناء. لهذا نوصي بإختبار هذا الملدن للتحقق من جودة استخدامه في الخرسانات المختلفة.

6- يوصى بالقيام بعدد من الاختبارات الأخرى للخرسانة في حالتها اللدنة، مثل اختبار U-BOX و J-RING و V-FUNNEL وغيرها.

المراجع :

- [1]. Alireza K, John L et al. Effect of molecular architecture of polycarboxylate ethers on plasticizing performance in alkali-activated slag paste [J]. J Mater Sci, 2014(49):2761–2772.
- [2]. Alone M, Palacios M, et al. Compatibility between polycarboxylate-based admixtures and blended-cement pastes [J]. Cement and Concrete Composites, 2013, (35):151-162.
- [3]. Yamada K, Takahashi T, Hanehara S, et al. Effects of the Chemical Structure on the Properties of Polycarboxylate-type Superplasticizer[J]. Cement and Concrete Research, 2000, 30(2): 197-207.
- [4]. BS 12: 1996 "Specification for Portland Cement". British Standards Institution, 389 Chiswick high road, London, W4 4AC, UK
- [5]. (11) BS 882:1992 "Specification for aggregates from natural sources for concrete". British Standards Institution, 389 Chiswick high road, London, W4 4AL, UK, 1992.
- [6]. المواصفات القياسية الليبية رقم 294 لسنة 1988 ف، الخاصة بالمياه المستعملة في الخرسانة، المركز الوطني للمواصفات والمعايير القياسية • طرابلس
- [7]. Grunewald. S. (2004) "Performance Based -Design of Self Compacting Fiber Reinforced Concrete", PH. D Thesis. Department of Civil And Mining Engineering, Delft University of Technology, Published And distributed By Delft University Press, Netherlands.
- [8]. "The European Guidelines for Self-Compacting Concrete". Specification, Production and Use (2005).
- [9]. Evangeline.K1, Dr.M.Neelamegam2, IOSR Journal of Mechanical and Civil Engineering (IOSR-JMCE) e-ISSN: 2278-1684, p-ISSN: 2320-334X. PP 18-29 www.iosrjournals.org, (Department of Civil Engineering, Easwari Engineering College, India).

Performance of Recycled PVC Aggregates in Concrete – Comparative Study

Salhin Alaud^{1*}, Suvash Chandra Paul²

¹ s.alaud@elmergib.edu.ly, ² suvashpl@iubat.edu,

¹ Civil Engineering Department, Garaboulli Engineering Faculty, Elmergib University, Libya

² Department of Civil Engineering, International University of Business Agriculture and Technology, Dhaka 1230, Bangladesh

ABSTRACT

Keywords:

Recycled plastic,
PVC plastic,
PVC Concrete
properties.

Plastic is the most used material in the world due to its easiest manufacturing and shaping, low density, and low cost. Their accumulation poses environmental problems due to no biodegradable. In recent times, researchers have concerned with the reuse of waste and recycled plastic materials in the concrete mix. Many studies have been published for the behavior of concrete having recycled plastic materials. This paper summarizes and compares the published literature studies on PVC recycled as a partial fine and coarse aggregate form replacement in concrete. The characteristics of the material, size and shape, replacement ratio, and the influence of PVC materials on the concrete properties have been discussed. The resultant concrete properties such as slump and workability, density, compressive strength, and durability of the PVC concrete were compared in the available literature. The results are not convergent due to many factors such as replacement of PVC ratio, shape, size, and the case of the PVC waste. In general, concrete containing PVC reduces the density, mechanical properties (compressive strength, splitting tensile strength, and flexural strength), and dry shrinkage as the PVC replacement ratio increase. The lower mechanical strength of concrete may be attributed to weak of ITZ in the interface between the aggregates and cement paste. Resistance of chloride ion penetration is better in PVC mixed concrete than the reference concrete. The PVC concrete may classify the lightweight based on the density of the concrete.

1. Introduction

The polymer products are widely used in every field, especially in packaging, industries, agriculture, and others. In the last half of the 20th century, plastics have been used increasingly in a broad range of products because of ease of design and manufacture, and cheaper cost. Over 275 million metric tons (MT) of plastic waste in 2010 were produced in 192 coastal countries, and from 4.8 to 12.7 million MT may enter the ocean [1]. In 2017, the production of the plastic has grown to 348 MT [2]. The produced from raw materials was 360 in addition to 30 MT from recycled according to a pilot study of six associations and organizations concerned in the plastics in 2018 [3]. Although of substantial amounts of plastic production, the recycling of the waste presents only one-third of the waste while its quarter goes to landfill, and the rest for energy recovery [4]. The amount of recycled plastic waste was doubled from 2006 to 2018.

Polyvinyl chloride (PVC) is one of the plastic types uses in the construction such as windows, frames, profiles, floor, and wall coverings, pipes, cable insulation, garden hoses, inflatable pools, etc. The first discovery of PVC was in 1835 by the French physicist Victor Regnault and developed manufacturing processes enabling its industrial progress after that by the German professor Fritz Klatter in 1912 [2]. By polymer types in 2016, the PVC presents 10 % of the European plastics converter demand [4]. PVC recycled has been partially replaced in concrete as a fine or coarse aggregate over the past few decades. This article summarizes some related studies of PVC recycled waste materials in concrete and supplies a space for details by comparing the results effect of the PVC additive in concrete from these studies.

2. Material properties of PVC plastics used in concrete

Polyvinyl chloride (PVC) is a basic element in engineering construction and supplies a series of services such as all types of pipes for fluids and gas distribution, doors and windows, profiles, floor and wall coverings, pools, etc. The ground PVC scrap works well as a partial substitute for concrete aggregate rather than being sent to landfill or incinerated. The PVC is replaced in concrete as a part of fine or coarse natural aggregates. Using recycled plastic in concrete reduces the weight of concrete, thus reducing transportation costs [5]. Because the properties of bulk density, compressive and tensile strength of concrete containing waste plastic in general or PVC in any form are less than conventional concretes, the plastic concrete can be used as lightweight concretes for structures with lower strength requirements [6, 7, 8]. Significant differences in properties can occur, based on the source of the waste and whether is mono/a mixed plastic, is the plastic clean from the inorganic components, and are the composing polymers and their respective ratios in the mix is a known [9]. In general, the plastics in concrete have good characteristics as insulation for hot, cold, sound, and saving energy [10]. The chloride ions resistance in the structure is influenced by PVC aggregate, and the results show that the PVC aggregate can be added to the concrete exposed to the aggressive environment [11].

3. PVC Plastic aggregates

There is no specific standard classify the granule size of the plastic added to the concrete or replacement ratio. The PVC recycled concrete may be divided according to the size of the plastic added into dust particles with a size less than 1 mm, fine aggregate with a size of 1-4 mm, and with a size of more than 4 mm for coarse. Radimir et al [6] summarized the classification of the form into dust particles with a size up to 1 mm, for flakes with a size of 1-10 mm, pellets with a size of 10-25 mm and length 25-50 mm for fibers. Figure 1 shows the variety of forms of PVC aggregates used by the researchers. PVC collected directly from the waste PVC pipe by crashing typically has an angular and rough surface [11, 12, 13]. Some PVC aggregates are also processed from the waste by melting and have a smooth surface as seen in Figure 1d. Numerous publications have concerned the properties of concrete containing PVC recycled as aggregate in different forms and sizes.

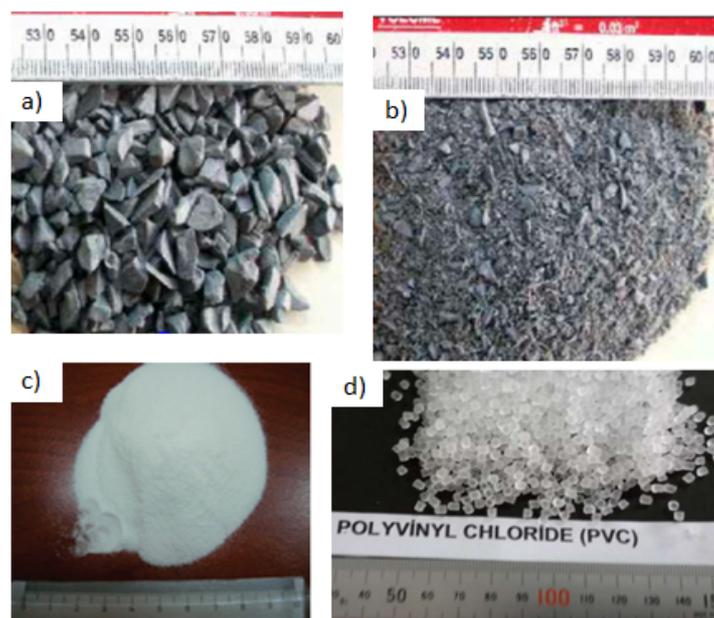


Figure 1: Different forms of PVC aggregates used in concrete such as (a) coarse aggregates, (b) fine aggregates, (c) fine powders and (d) processed PVC aggregates [11, 12, 13].

4. Properties of concrete containing PVC aggregates

This section integrates the information on various fresh, mechanical, and durability properties of concrete mixed with PVC aggregates in the available researches.

4.1 Fresh properties of PVC mixed concrete

Concrete is a type of material that can be poured into the mold to create any shape. However, the fresh properties of concrete such as workability, yield stress, viscosity, consistency, plastic shrinkage, etc. are the crucial factors that can influence both the mechanical and durability properties of hardened concrete [14]. The air content, bleeding, and segregation of fresh concrete are also crucial factors that need to control for having the maximum benefits of using concrete.

In a study [12], PVC was used as both powder (size is between 0 and 0.25 mm) and fine (size is between 2 and 4 mm) forms at 10, 20, and 30% replacement level by volume of aggregate in concrete. The slump value of the concrete was reduced for both powder and fine form of PVC. The maximum reduction of the slump was noticed by about 17% when 10% of PVC fine replaced by the aggregates. However, this reduction was reduced as the percentages of PVC were increased. At 30% replacement of PVC, the slump value was as same as the reference concrete. This trend of a similar slump or even slightly higher slump in concrete was also noticed for the coarse (size > 4 mm) PVC at a higher level of percentages (30, 35, and 40%) [13]. Because of the non-absorbent nature of PVC, the concrete mix may have more water and thus higher slump can be achieved. This behavior was opposite where a systematic reduction in the slump value of concrete was reported for an increased amount of PVC fine aggregates in concrete [15]. Najjar et al. [16] investigated the effect of the percentage of rigid PVC with a particle size of about 4 mm on the slump also. The result showed that there is an increase in the value of slump about 17%, 33%, and 87% in concrete at the PVC waste ratios of 5%, 10%, and 20%, respectively compared with the reference concrete mix.

Azad et al. [17] find that the slump was not changed when fine aggregate is replaced with PVC aggregate not more than 15%, and with coarse aggregate replaced by up to 30% PVC aggregate.

Sudden changes in slump value when PVC coarse aggregate increased to 45% and PVC fine aggregate to 65%. Overall, the slump value of concrete depends on the shape of the PVC aggregates. The smooth shape of processed PVC granule aggregates may lead to a higher slump of concrete, while the angular shape of PVC aggregates increases the friction between the aggregates particles and thus reduce can reduce the slump. Slump value was gradually reduced even at a lower level of PVC replacement (2.5, 5, 7.5, 10, 12.5, and 15%) as fine aggregates in concrete [18]. At these levels, the slump values decreased by about 4%, 6%, 11%, 12%, 15%, and 18%, respectively, in comparison to the reference concrete. Figure 2 shows the workability of PVC mixed concrete reported by the researchers in the existing literature [15, 19, 20]. Inconsistency of the slump results can also be seen from the figure. The density of fresh concrete also decreased as the percentages of PVC aggregates increased [12, 13, 18]. Maximum 6% and 10% reduction in density were reported for the PVC replacement of 30% and 40%, respectively [13]. The lower density of PVC mixed concrete can be attributed to its lower compaction factor as reported in [21]. The air content of concrete was found to be increased with the creased percentages of PVC aggregates [22].

For the PVC replacement levels of 10, 20, 30, 50, and 100%, the air content in concrete was increased to 10%, 10%, 110%, 167%, and 386%, respectively, when compared with the reference concrete.

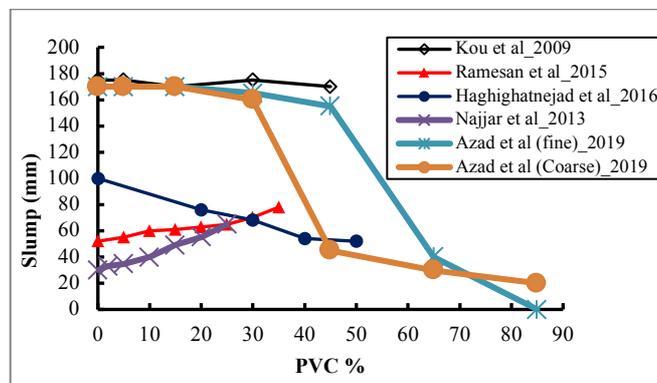


Figure 2: Slump value of concrete with different percentages of PVC.

4.2 Hardened properties of PVC mixed concrete

Concrete ability to withstand the structural and service loads is defined by its hardened properties. The hardened properties such as dry density, compressive strength, tensile strength, modulus of elasticity, etc. were investigated in researches [11, 13, 16, 17, 19, 23].

Results obtained by Hussein et al [18] show that 15% of fine PVC replaced in concrete mix can be reduced the dry density of reference concrete (2331kg/m^3) to (1892kg/m^3), which mean this density is in between the limitation of lightweight structural concrete density according to (ACI 213R 2014) [24]. Different results by Azad et al. [17], where the significant change in density of concrete with PVC aggregate at high PVC content of 85%. Even at this content of PVC replacement, the concrete became lightweight only when the coarse PVC Aggregate was replaced. This may be attributed to the relatively high-density of PVC aggregate used in this study in comparison with that used by the last investigation. The dry density of the concrete decreases by 2.00%, 3.20%, and 4.73% for the concrete containing 10%, 20%, and 25% PVC waste aggregate, respectively in a study by Najjar et al. [16]. The reduction in the dry density is attributed to the PVC density is lower than the natural aggregate density.

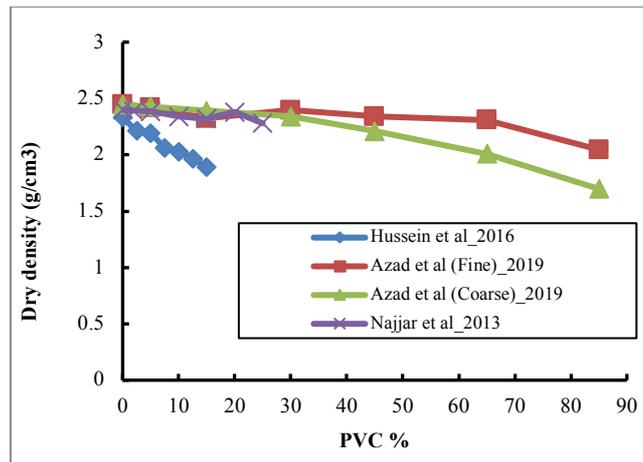


Figure 3: Effect of PVC replacement on the dry density of concrete of PVC.

In every research, it was found that the partial replacement of natural aggregates by PVC aggregates reduced the compressive strength of concrete at any remedial age. At 28 days, the compressive strength of concrete with 20, 30, 40 and 50% PVC fine aggregate mixed concrete was about 10%, 13%, 18%, and 34% lower than that of the control mix [23]. In another study [19], 28 days compressive strength of concrete with the inclusion of PVC granule at replacement level of 5, 15, 30, and 45% reduced about 9%, 19%, 22%, and 47%. The difference in concrete strength was attributed to the weak bond strength between the PVC aggregates and the cement paste. The lower elastic modulus of PVC aggregates compared to the cement paste also form cracks around the PVC particles and hence lower the compressive strength [19]. When comparing the PVC in powder and granule form, the strength of powder PVC concrete was lower than the concrete with granule PVC[12]. Research also showed that the surface treatment of plastic aggregates by chemicals such as hydrogen peroxide solution (H_2O_2) and calcium hypochlorite solution ($Ca(ClO)_2$) can improve the mechanical properties of concrete [25]. Chemical treatment reduces the hydrophobicity of the surface of the plastic surface, allows a stronger bond to be formed between the cement paste and plastic. Figure 3 shows the compressive strength of concrete at 28 days with different percentages of PVC found by the different researchers. In all studies, compressive strength was found to be decreased as the PVC content increased. Although the compressive strength of concrete is reduced as the percentages of PVC content increased, however, 10% PVC replacement was found to be acceptable as the overall strength reduction was below 10%, which is still acceptable for many applications.

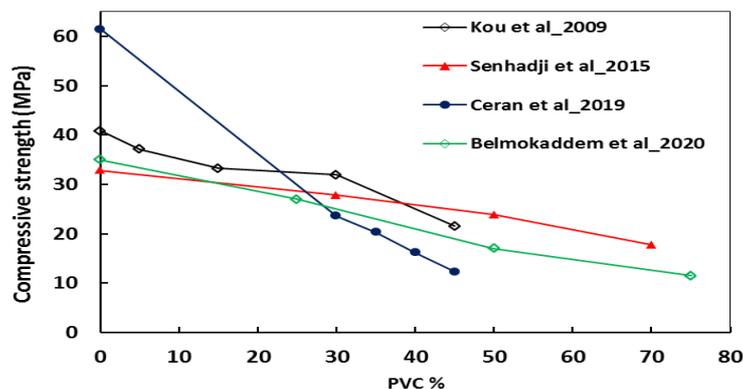


Figure 4: 28 days compressive strength of concrete with different percentages of PVC aggregates.

Like the compressive strength, the modulus of elasticity of concrete was also reduced as the percentages of PVC content increased. At 28 days, about 7%, 14%, 19%, and 62% lower modulus of elasticity was found in concrete with PVC replacement of 5, 15, 30, and 45% [19]. In another study where 30 and 45% natural aggregates were replaced by the PVC aggregates, a maximum reduction in modulus of elasticity was reported to be 8% and 32%, respectively [13]. The splitting tensile strength of concrete was also found to be decreased as the PVC content in the mixes increased [19, 26]. For 10, 20, and 30% of PVC replacement in concrete, at 28 days, the splitting strength of concrete was reduced to about 24%, 30%, and 34% [26]. Lower flexural and tensile strength of different percentages of PVC mixed concrete was also found in [18]. On the contrary, almost similar flexural and tensile strengths were reported for various percentages of PVC dust mixed concrete [27].

From the above discussion, different researches have concluded the different range of optimum PVC content as well as its impact on the concrete properties. Nevertheless, the variation in the results of different studies can be attributed to their different remedial conditions, water-cement ratio (w/c), physical properties of PVC and concrete mix compositions.

4.3 Durability properties of PVC mixed concrete

In its total service life, concrete structures undergo various kinds of environmental loads or actions. The durability of concrete defines by its ability to sustain in those environmental actions without any major damages. This section aims to summarize the different durability tests performed on PVC mixed concrete in the existing researches.

Chloride ion penetrability in different percentages of PVC mixed concrete was investigated by Kou et al [19]. It was found the resistance in chloride ion penetration of concrete increased as the percentages of PVC content increased. Concrete specimens after curing of 28 days showed that the total charges (coulombs) passed through 5, 15, 30, and 45% PVC mixed concrete were reduced by about 12%, 19%, 27%, and 36%, respectively, in comparing to the control concrete specimens. The resistant nature of PVC aggregates can block the passage of the chloride ion and thus increase the durability of concrete.

The water absorption in PVC mixed concrete was also investigated in [15]. For this purpose, cube specimens (100 mm) were dried in the oven and then submerged in tap water. The weight gain was measured at a regular interval. Concrete specimens with the inclusion of 20, 30, 40, and 50% PVC aggregates showed lower absorption of water than the control specimens. The inclusion of different percentages of PVC specimens showed the final absorption values of about 55%, 47%, 35%, and 70% lower than the concrete specimens.

The different percentages of PVC mixed concrete specimens were also tested in ultrasonic pulse velocity (UPV) [17]. At lower percentages of PVC (5% and 10%), there was no noticeable difference in the results of PVC concrete specimens when compared with reference specimens. However, at 40% PVC replacement level, the reduction of UPV values was about 13%, 13%, and 14% for concrete ages of 7, 28, and 56 days, respectively. Therefore, at a lower level of PVC replacement in concrete, the effect on residual UPV was not found to be significant. It was concluded that the similarity of concrete mix improves when adding PVC aggregates which can lead to reduce the permeable pores in the matrix and improve the UPV [12]. In another study, it was concluded that the concrete with 50% PVC, the UPV value is still at good range and acceptable for many applications [11].

The rate of drying shrinkage of concrete also decreased as the percentages of PVC aggregates increased [19, 28]. Figure 4 shows the drying shrinkage results at 60 days of PVC mixed concrete

found in existing studies [19, 28]. Water loss from the concrete induces capillary tension force and thus drying shrinkage occurs in concrete. The intensity of this capillary tension depends on the pore volume, size, and structure of the concrete [29]. Since PVC is resistant and does not absorb water when compared to natural aggregate and hence would be able to reduce the overall drying shrinkage [19].

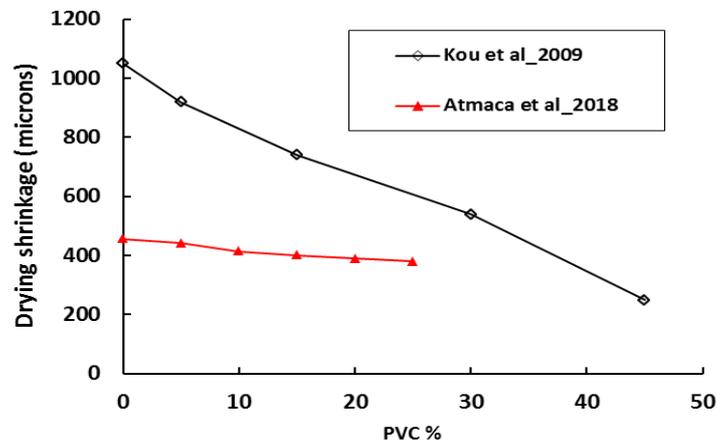


Figure 5: Drying shrinkage of PVC mixed concrete at the age of 60 days [19, 28].

Microstructures of PVC mixed concrete

The effect of PVC on the microstructure of cement paste was also investigated through a scanning electron microscope (SEM) image analysis [11, 17]. As shown in Figure 6a, a weak interfacial transition zone (ITZ) forms between the PVC aggregates and cement paste due to the hydrophobic nature of plastic. On the other hand, a strong ITZ can be formed in the interface between the natural aggregates and cement paste. Weak bonding between aggregates and cement paste also forms micro-cracks under any kind of loading as can be seen in Figure 6b. As the plastic content increases, free water around the aggregates weakens the interface between the plastic and cement paste resulting in a less dense zone with large voids in the matrix [30]. This is a factor contributing to a decrease in the mechanical strength of PVC mixed concrete [11].

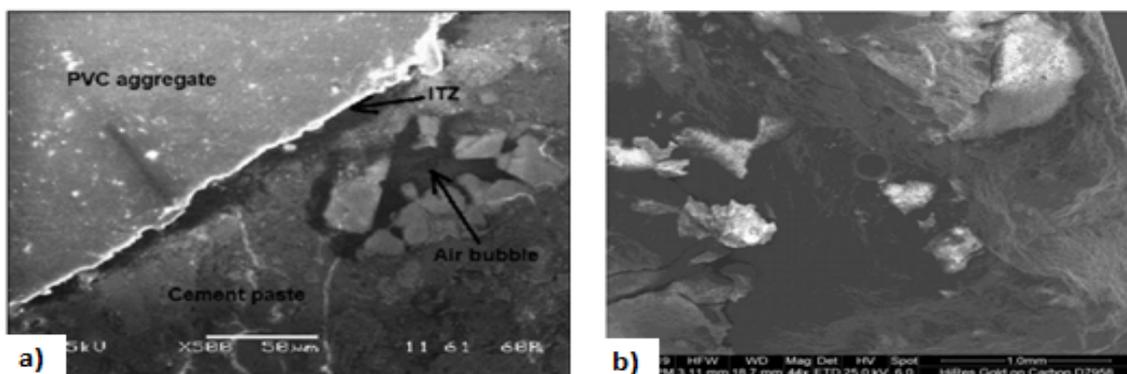


Figure 6: SEM images of microstructure of concrete with (a) ITZ between PVC aggregate and cement paste and (b) cracks in PVC mixed concrete [11, 17]

Concluding remarks

Over the last few decades, there has been much research work performed on the different properties of concrete with various percentages of plastic waste. This paper aimed to summarize the fresh, mechanical, durability, and microstructural properties of PVC mixed concrete from the available literature. From the above discussion, the following conclusions can be drawn:

- The resistant nature of PVC aggregates may supply better workability of concrete. However, the size and shape of the PVC aggregates play a key role in workability. The angular shape of PVC aggregates may increase the friction between the particles and thus reduce the workability.
- The mechanical properties of concrete reduce as the percentages of PVC aggregates increase. And 10% PVC replacement is acceptable for many applications.
- Water absorption and shrinkage of PVC mixed concrete are lower than the reference concrete. At a lower replacement level of PVC (5-10%), no noticeable difference can be found in UPV of concrete. Chloride ion penetration resistance is also better in PVC mixed concrete when compared with reference concrete.
- The hydrophobic nature of PVC aggregates forms weak ITZ in the interface between the aggregates and cement paste, which can lead to the lower mechanical strength of concrete.

For future studies, the mechanism of lower mechanical strength of PVC mixed concrete must be studied. Also, the long term performance of concrete with PVC aggregates should be investigated to grow the confidence of using this in concrete.

Reference

- [1]. J. Jambeck, R. Geyer, C. Wilcox, and T. Siegler, "Plastic waste inputs from land into the ocean," *Science*, 2015.
- [2]. P. Chalmin, "The history of plastics: from the Capitol to the Tarpeian Rock," *The journal of field actions*, 2019.
- [3]. K-2019, "Avoiding environmental littering, pushing forward with a circular economy – Plastics industry creates transparency for global plastic flows," PDF of Global Plastics Flow Study Sumurized from www.k-online.com/GPFStudy: Düsseldorf, 2019.
- [4]. E. Plastics, "An Analysis of European Latest Plastics Production," *the Facts 2019*, 2019.
- [5]. S. Rafat, K. Jamal, and K. Inderpreet, "Use of recycled plastic in concrete: A review," *Waste Management*, vol. 28, p. 1835–1852, 2008.
- [6]. R. Novotny, J. Sal, and M. Ctibor, "Environmental use of waste materials as admixtures in concrete," in *WMCAUS 2019*, 2019.
- [7]. N. Puri, B. Kumar and H. Tyagi, "Utilization of Recycled Wastes as Ingredients in Concrete Mix," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 2(2), 2013.
- [8]. M. Seghiri, D. Boutoutaou, A. Kriker and M. Hachani, "The Possibility of Making a Composite Material from Waste Plastic," in *International Conference on Technologies and Materials for Renewable Energy, Environment and Sustainability*, 2017.
- [9]. P. Sambhaji, "Use of Waste Plastic in Concrete Mixture as Aggregate Replacement," *International Journal of Advanced Engineering Research and Science (IJAERS)*, vol. 3(12), 2016.
- [10]. R. Kim, L. Delva and K. V. Geem, "Mechanical and chemical recycling of solid plastic waste," *Waste Management*, vol. 69, pp. 24-58, 2017.
- [11]. Y. Senhadji, G. Escadeillas, A. Benosman, M. Mouli, H. Khelafi and S. O. Kaci, "Effect of incorporating PVC waste as aggregate on the physical, mechanical, and chloride ion penetration behavior of concrete," *Journal of Adhesion Science and Technology*, vol. 29(7), pp. 625-640, 2015.
- [12]. H. Bolat and P. Erkus, "Use of polyvinyl chloride (PVC) powder and granules as aggregate replacement in concrete mixtures," *Science and Engineering of Composite Materials*, vol. 23 (2), pp. 209-216, 2009.

- [13]. B. Ceran, B. Şimşek, T. Uygunoğlu and N. Şara, "PVC concrete composites: comparative study with other polymer concrete in terms of mechanical, thermal and electrical properties," *Journal of Material Cycles and Waste Management*, vol. 21(4), pp. 818-828, 2019.
- [14]. A. Babafemi, Šavija, C. Paul and V. Anggraini, "Engineering properties of concrete with waste recycled plastic: a review," *Sustainability*, vol. 10(11), p. 3875, 2018.
- [15]. N. Haghghatnejad, Y. Mousavi, J. Khaleghi, A. Tabarsa and S. Yousefi, "Properties of recycled PVC aggregate concrete under different curing conditions," *Construction and Building Materials*, vol. 126, pp. 943-950, 2016.
- [16]. A. Najjar, E. Basha and M. Milad, "Rigid polyvinyl chloride waste for partial replacement of natural coarse aggregate in concrete mixture," *International Journal of Chemical and Environmental Engineering*, vol. 4(6), pp. 399-403, 2013.
- [17]. M. Azad, M. Ilham and M. Shuaaib, "Some properties of concrete with plastic aggregate derived from shredded PVC sheets," *Construction and Building Materials*, vol. 201, pp. 232-245, 2019.
- [18]. H. Hussein, A. Eedan and K. Ahmed, "Mechanical, thermal and acoustical properties of concrete with fine polyvinyl chloride (PVC)," *Iraqi Journal of Civil Engineering*, vol. 11(2), pp. 81-91, 2017.
- [19]. S. Kou, G. Lee, C. Poon and W. Lai, "Properties of lightweight aggregate concrete prepared with PVC granules derived from scraped PVC pipes," *Waste Management*, vol. 29(2), pp. 621-628, 2009.
- [20]. A. Ramesan, S. Babu and A. Lal, "Performance of light-weight concrete with plastic aggregate," *International Journal of Engineering Research and Applications*, pp. 105-110.
- [21]. D. Osei, "Experimental investigation on recycled plastics as aggregate in concrete," *International Journal of Structural and Civil Engineering Research*, vol. 3(2), pp. 168-174, 2014.
- [22]. C. Chen, N. Jaffe, M. Koppitz, W. Weimer and A. Polocoser, "Concrete mixture with plastic as fine aggregate," *International Journal of Advances in Mechanical and Civil Engineering*, vol. 2(4), pp. 49-53, 2015.
- [23]. M. Belmokaddem, A. Mahi, Y. Senhadji and Y. Pekmezci, "Mechanical and physical properties and morphology of concrete containing plastic waste as aggregate," *Construction and Building Materials*, vol. 257(119559), 2020.
- [24]. ACI-213R, *Guide for Structural Lightweight-Aggregate Concrete*, American Concrete Institute, 2014.
- [25]. Z. Lee, S. Paul, S. Kong, and S. Susilawati, "Modification of waste aggregate PET for improving the concrete properties," *Advances in Civil Engineering*, 2019.
- [26]. B. Seshaiyah and K. Lalitha, "Experimental study on the performance of concrete with polymer (PVC) as filler material," *International Journal of Research Sciences and Advanced Engineering*, pp. 74-80, 2017.
- [27]. S. Priyadarshika and M. Rajkannan, "Experimental Investigation on Strength Properties of Concrete by Partial Replacement of PVC Powder with Cement and Flyash," *International Journal of Scientific Research and Engineering Development*, vol. 2(3), pp. 781-789, 2019.
- [28]. N. Atmaca, A. Atmaca, M. Aljumaili and A. Özçetin, "Strength and shrinkage properties of self-compacting concretes incorporating waste PVC dust", *Int. Journal of Energy and Engineering Sciences*, vol. 3(1), pp. 47-57, 2018.
- [29]. K. Kim and S. Lee, "Prediction of differential drying shrinkage in concrete," *Cement and Concrete Research*, vol. 28(7), pp. 985-994, 1998.
- [30]. S. Yang, X. Yue, X. Liu, and Y. Tong, "Properties of self-compacting lightweight concrete containing recycled plastic particles," *Construction and Building Materials*, vol. 84, pp. 444-453, 2015.

Effect of Crude Oil Pollution on Some Geotechnical Properties of Disturbed Sandy Soil

Abir Ahmed Elazzabi

abirelazabi@gmail.com

Department of Civil Engineering, Collage of Engineering, Tripoli

ABSTRACT

Recently, the effects of soil pollution with crude oil have received much attention from many researchers. Most of the results showed contradictions in their results because of adding crude oil to the soil. Some of the results showed an increase in the maximum dry density while other results indicated a decrease as well as for the shear resistance. The present study seeks further research to evaluate the physical and mechanical properties of loose sandy soils contaminated with crude oil obtained from the Nalut region in southwestern Libya.

Keywords:

Crude Oil

Specific Gravity

Permeability

Shear Strength

Maximum Dry Density

The soil samples used in this research were classified according to the Unified Soil Classification System (USCS) as poor graded sandy soil (SP.) And were classified (A-3) according to the American Association of Highway and Transportation Officials for Classifying Soil System (AASHTO).

The contaminated material is crude oil from Abu Al-Tifl field located in Gallo city. The soil was polluted in a laboratory by manually mixing the soil with the pollutant material in percentages (10, 7, 4) of the dry soil weight after the soil was well mixed with the pollutant and after confirming the uniformity of the distribution The pollutant was in the soil and left for three days in closed containers to study the effect of crude oil on the physical and mechanical properties of the soil by conducting conventional tests to measure these properties and through the results of laboratory tests showed that crude oil has a great effect on some properties and little effect on other properties. As the increase in the pollutant percentage in the soil led to a slight effect on the specific gravity, while there is a significant effect on the permeability coefficient as it decreased with the increase in the percentage of pollutant, the optimum water content also decreased and the maximum dry density increased with the increase in the percentage of pollutant, while the value of the California Bearing Ratio (CBR) increased at 4% and then starts to decrease as the crude oil increases.

1- Introduction

The correct design of any engineering facility requires adequate knowledge of the engineering properties of the soil and its effect on the surrounding conditions in addition to knowledge of its impact on the materials presented to it from different sources, which is very important in determining the change that occurs to its engineering properties and among the most likely of these materials to seep into the soil and penetrate it is Crude oil and its derivatives and their high percentages of hydrocarbons, as hydrocarbon leakage and spillage has become a major concern in most oil-producing countries in the world, as oil pollution of the soil is responsible for the majority of soil pollution with hydrocarbons [1].

The impact of these leaks and spills on the environment cannot be overlooked, as they are considered a geo-environmental problem that negatively hinders the quality of soil, groundwater and the atmosphere. Crude oil, or one of its products, when it seeps into the soil through the unsaturated area under the influence of the forces of gravity and capillary property, and through the pores, a part of it is kept there where it adheres to the soil particles, while the remaining part reaches the groundwater level, causing water pollution. The evaporation of the remainder in the atmosphere pollutes the air, vegetation, etc. [2, 3].

Soil pollution above a certain level causes the loss of some geotechnical soil properties, so the problem is complicated by the arrival of the effects of pollution to the soil supporting the foundations and working to weaken it, i.e. losing some of the bearing capacity of the soil and its effect appears in the form of deformations in parts of the structures and sometimes leads to collapse as a result of cavities in the event of neglect and lack of control. Therefore, it is important to determine the effects of soil pollution based on existing structures. This change also has a far-reaching impact on the proposed structures that the polluted soil must support. It can lead to structural or functional collapse of existing structures, especially when pollution causes a significant increase in soil plasticity; Loss of bearing capacity; increasing subsidence, and / or preventing drainage of water or other fluids.

For the proposed structures, this could cause abandonment of the site with the contaminated soil, a decrease in the project potential, or an increase in the project cost. This cost is represented in geotechnical and chemical analyzes of the soil to determine the extent of its contamination, and the deterministic choice of the most stable type of structural foundation or the cost of applying techniques Soil remediation or stabilization. However, the use of some previously verified wastes as soil stabilizers - marble powder [4], steel slag [5,6], modified asphalt pavement (reclaimed asphalt pavement) [7]- May reduce the percentage increase in project cost for proposed facilities on contaminated soil.

In most cases, petroleum hydrocarbons are accidentally caused by an oil spill during transport as a leak from pipelines and storage tanks, or during oil drilling operations, so an oil spill can be considered an inevitable consequence. And the lack of static electrical systems that would protect the tank structures and pipelines from corrosion, and if any, they may not receive follow-up and periodic tests to ensure their efficiency is one of the biggest causes of oil spillage or one of its derivatives into the soil, an example of this is the pipeline transporting oil derivatives linking the Zawiya oil depot.(Which feeds from the Zawiya refinery for oil refining) and the Tripoli oil depot, the airport road of the Brega Oil Marketing Company, where a leakage of this pipeline occurred in the Ghiran area, Janzour, in front of the General Water Authority building, as it negatively affected the water wells in the area and some drops occurred in some unpaved places from here The idea of research came, in addition to the fact that all the oil fields in Libya are located in the center and south of the country and the export ports are on the Libyan coast in the north of the country, which makes the pipelines to transport crude oil from the far south to the far north and it is known that the northern region of the country has 80% of the population. It makes it capable of reconstruction and construction in the future. From here came the idea of studying the effect of studying crude oil on the soil to avoid any future problems in the foundations.

In order for the treatment of the leakage problem to be based on scientific foundations, it is necessary to first study the effect of these fluids on the engineering properties of the soil in preparation for proposing treatment and fixing methods.

2- REVIEW OF LITERATURE:

Researchers have reported that the engineering properties of oil-polluted soils are radically altered and make them unsuitable to support engineering facilities, while others have pointed to the idea that soil pollution with an appropriate proportion of crude oil or its products will modify some geotechnical properties of soil.

During a study conducted in India that pollutant containing percentages of petroleum derivatives lead to reducing soil resistance. Also stabilizing this soil by using a number of stabilizers, including cement, in addition to a mixture of other compositions, and the soil showed a remarkable improvement in its properties [8].

It is investigated the effects of four types of pollutants on the geotechnical properties of clay soils. Pollutants are kerosene, ammonium nitrate, copper and lead. Then each one of them was mixed with the soil at rates of 10% and 25%. The results showed various effects of these pollutants on the geotechnical characteristics of clay soil [9].

The effect of diesel oil on the geotechnical properties of soil has been studied. The results of this study showed that the addition of diesel oil has an effect on the geotechnical properties of the studied soil samples [10]. The effects of petroleum hydrocarbons on the geotechnical properties of clay soil samples were investigated, and the results showed several effects on the geotechnical properties of soil samples [11].

Also, some oil derivatives such as emulsified asphalt and containing large quantities of solvents such as (gasoline and kerosene) have studied its effect on the soil and proved the possibility of using it as an improved (proven) substance for some engineering properties, where it was found that the cohesion value of soil increases with increasing percentage. The asphalt emulsion then decreases with the increase in the ratio from its optimum value, and the internal angle of friction was found to decrease because these materials work to reduce the friction between the particles) [12,13].

A presentation was made on the use of industrial water containing petroleum products in road layering, as it was used as an additive for well-graded sandy soil, and the study showed improvement in most of the engineering properties of soil [14].

The effect of three types of petroleum derivatives ranging from 0% to 6% on sand has been studied. It was found that the bearing capacity is greatly affected as a result of the shear effect. Soil parameters [15].

The amount of increase in the concentration of benzene in the soil as a result of leakage from the oil pipelines has been studied. It was found that the concentration of benzene increased by about 800 to 5000 times the amount of concentration before the leakage, and that this increase was the cause of many construction problems as a result of the differential decrease. In the foundations [16].

The effect of crude oil contamination on compacting sand properties has been studied and it is concluded that when using crude oil as a porous liquid, the maximum unit weight of dry sand was about 6 percent higher compared to water as a porous liquid. The explanation for this is that the oil is more effective in reducing the friction between soil particles resulting in reduced spacing between soil particles; Thus increasing the dry unit weight at a certain compaction effort. The maximum shear strength, internal friction angle, is found to decrease with increasing oil saturation. They concluded by increasing the relative density of sand with the increase in the viscosity of the petroleum pollutant [17].

The effect of crude oil on the geotechnical properties of Kuwaiti sand has been studied, and it has been concluded that the compressibility of sand as a result of adding crude oil was improved by adding oil by up to 4% by weight. The permeability coefficient decreased by 20% and the internal friction angle decreased when the oil was added [18].

The effect of residual basalt soil pollution with motor oil on the geotechnical properties of soil has been studied [19]. The effect of diesel oil pollution on the microstructural changes of clay soil has also been studied [20].

The effect of contamination of clay soil with methanol, ethanol, isopropyl alcohol, and acetic acid on plasticity, consolidation, and shear resistance has been investigated. [21] Laboratory studies were conducted to determine the effect of crude oil contamination on laterite soils on shear resistance [22].

The effect of crude oil contamination has been studied on the plasticity and compression properties and other of geotechnical properties of loose soils, but their results were different, making it difficult to obtain a general description of the effect of crude oil contamination on clay soils [23].

3- MATERIALS:

3.1 WATER: Distilled water was used in this research, except for the compaction test, which used regular water.

3.2 SOIL The soil used in this study was Collected from depth of 50cm below natural ground surface from the Nalut region in southwestern Libya It is a sandy soil that has been classified according to the Unified Soil Classification System (USCS) as (PS) and (A-3) according to the American Association of Highway and Transportation Officials of Soil System (AASHTO), table 1 shows the natural properties of the soil. And Figure 1 shows the results of the sieve analysis test for soil free of contamination.

Table 1: Physical and engineering properties of tested soil

Properties of soil	Values	Specification
Physical Properties		
M.C (%)	1.5%	(ASTMD-2216)
G.S	2.645	(ASTM C 127).
Classification of soil (AASHTO)	A-3	AASHTO T99-01
Classification of soil (USCS)	SP	(ASTMD-2487)
Engineering Properties		
Unsoked CBR (%)	17%	(ASTM D1883-16)
K (cm/sec)	0.043	(ASTM D 2434-68).
Max Dry Density (gm/cm ³)	1.84	(ASTM D698-12).
C(KN/m ²), Ø	0KN/m ² , 31.09°	(ASTM D3080).

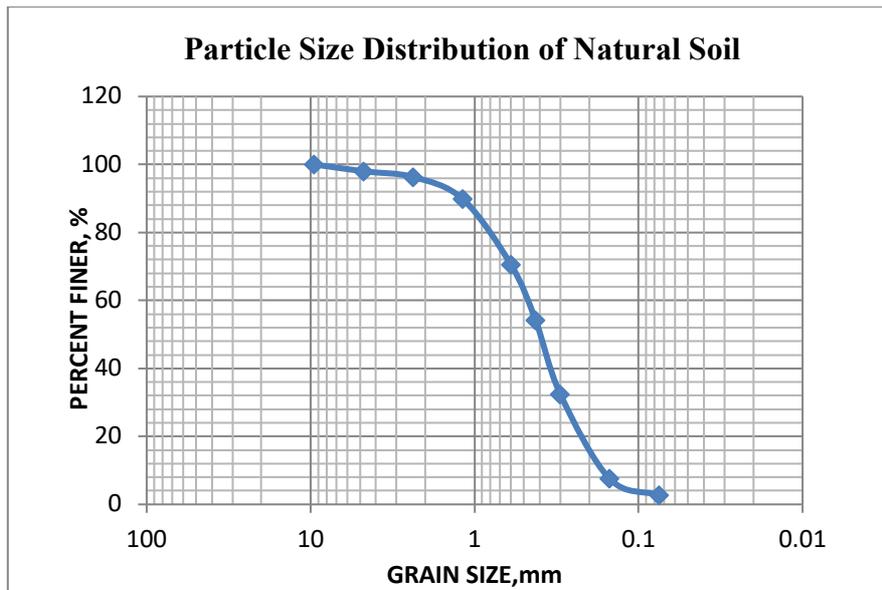


Figure 1: Grain size distribution for the tested soil

3.3 CRUDE OIL (ADDITIVE SUBSTANCE)

Some properties of Crude Oil of Buattifel field are shown in Table2. The analysis of Buattifel crude oil sample was carried out using well recognized standard procedures given in ASTM .IP and UOP methods.

Table2: Physical and chemical properties of Buattifel crude oil

DESCRIPTION	VALUES	METHOD
, g/ml Density @ 15C°	0.8121	ASTM D-4052
Specific gravity@60/60 15F°	0.8129	Calculation
API gravity	42.6	Calculation
Kinematic viscosity @210 F°	2.1512	ASTM D-445

4- METHODOLOGY:

The soil sample was divided into four parts. Crude oil added to each of the parts in 0%, 4%, 7%, and 10% proportions by dry weight of the soil sample, respectively. The soil-crude oil mixtures were thoroughly mixed manually and stored in closed containers for three days to ensure full absorption of contaminants into the soil. A set of laboratory tests carried out on the uncontaminated (0% crude oil content) soil sample and contaminated (4%, 7%, and 10% crude oil contents) soil samples. The procedures for laboratory tests were according to the specifications of the American Society for Testing and Materials (ASTM), the following tests are:

Soil Physical Tests

These included specific gravity test, and Sieve analysis test.

Soil Engineering Tests

These included standard compaction test, direct shear test, unsoaked California bearing ratio test, and permeability test.

5- RESULTS AND DISCUSSIONS:

We have conducted a set of laboratory tests on both natural and polluted sandy soil samples by adding crude oil in different proportions (4%, 7%, 10%), and the results as shown in table3:

Table 3: The characteristics of the soils contaminated with crude oil

Property	Max Dry Density	Gs	O.M.C	CBR	K	\emptyset	C	
	gm/cm ³	-	%	%	cm/sec	°	KN/m ²	
specification	(ASTM D698-12).	(ASTM-C127).	(ASTM D698-12).	(ASTM D1883-16)	D (ASTM 2434-68).	(ASTM D3080 -11)	(ASTM D3080 -11).	
Pollution percentage	0%	1.83	2.645	12.93	17	0.0426	31.09°	0
	4%	1.855	2.279	11.1	19.74	0.0298	26.09°	17.28
	7%	1.86	2.288	8.3	13	0.0087	27.75°	1.84
	10%	1.863	2.184	8.5	10.15	0.0074	4.00°	8.09

5.1 RESULTS OF SPECIFIC GRAVITY TEST:

A graphical illustration of the results of specific gravity tests on the soil admixed with varying percentages of crude oil content is presented in Figure 2. The specific gravities of the contaminated soil samples were found to be lower than that of the uncontaminated soil sample. As the crude oil content in the soil increased, the specific gravity of the contaminated soil progressively decreased. This is attributed to the lower specific gravity of the crude oil.

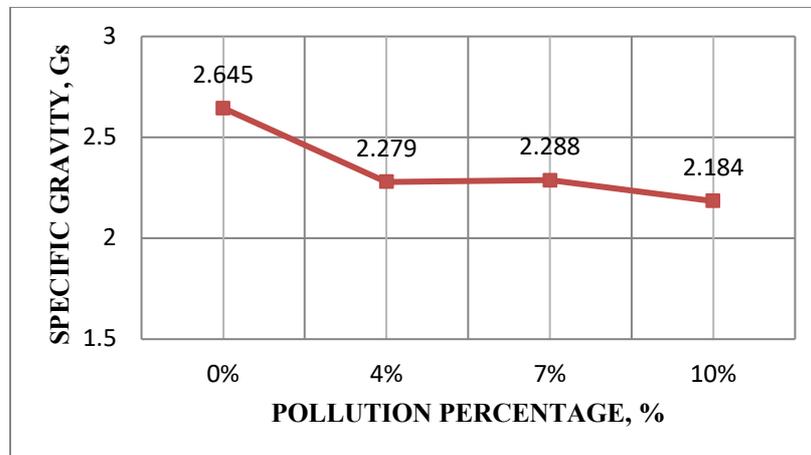


Figure 2: Variation of specific gravity with oil content

5.2 RESULTS OF STANDARD COMPACTION TEST:

The results of the standard compaction test on the soil samples are presented in Figure 3. As crude oil content in the contaminated soil increased, the optimum moisture content (OMC) of the soil decreased and the maximum dry density of the soil increased. The explanation for this is that the oil is more effective in reducing the friction between soil particles resulting in reduced spacing between soil particles; Thus increasing the dry density at a certain compaction effort.

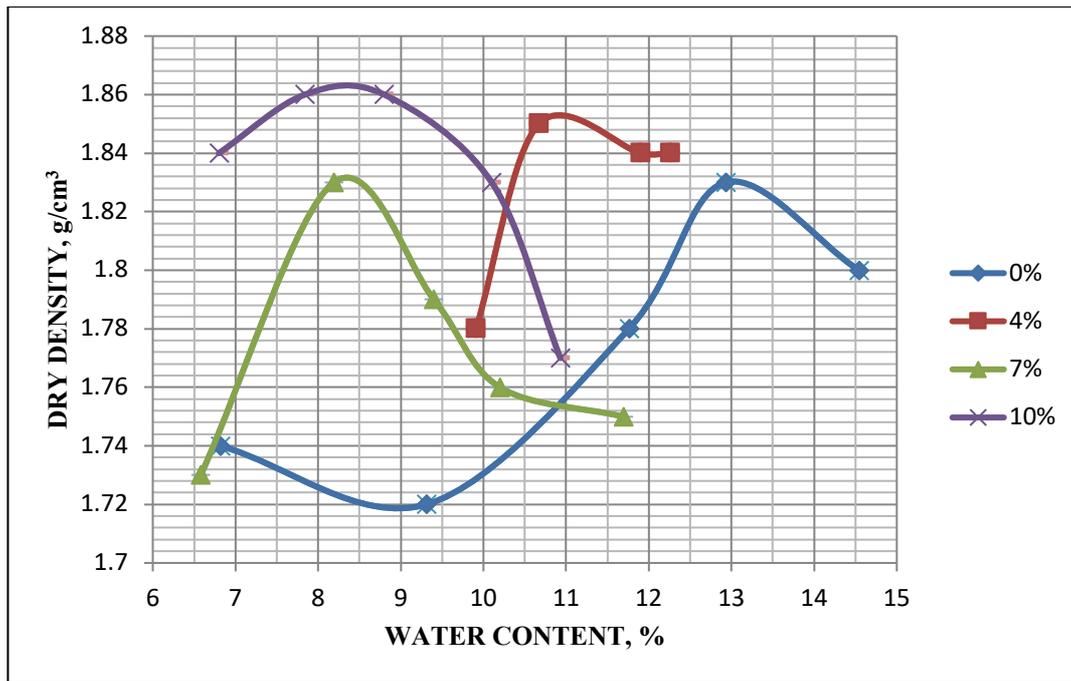


Figure 3: Variation of optimum moisture content and maximum dry density with oil content

5.3 RESULTS OF DIRECT SHEAR TEST:

The results of the Direct shear test for all proportions of contaminated and free of contamination soils are shown in table 4. When studying the table, it was found that the highest value of the internal friction angle when contaminated was 0%, reaching 31.09°, and the lowest value of the internal friction angle was 10%, which is (24.0°). And the highest value for soil cohesion when contaminated was 4%, reaching (17.28KN/m²), and the lowest for soil cohesion by 0%, reaching (0 KN/m²).

Table 4: Results of the direct shear test with oil content

Property		Pollution percentage			
Normal stress, σ KN/m ²	Shear stress, τ KN/m ²	0%	4%	7%	10%
		136.25	61.78	86.060	66.200
272.5	161.09	152.261	163.294	133.504	
408.75	226.18	219.564	208.53	189.774	
Parameters	C, KN/m ²	0	17.28	1.84	8.09
	Φ	31.09	26	27.75	24

5.4 RESULTS OF CALIFORNIA BEARING RATIO TEST:

The differences in unsoaked CBR values for soil with the addition of varying percentages of crude oil are graphically illustrated in Figure 4 and illustrated as a percentage of crude oil in soil, the unsoaked CBR increased slightly and subsequently decreased. The slight increase in the proportion of unsoaked CBR in contaminated soil is thought to have resulted from agglomeration of sand particles facilitated by crude oil, which may have caused an increase in the shear resistance between particles of the soil. The decrease in the non-soaked CBR after adding 4% crude oil indicates that the addition of 4% crude oil is to reduce the increase in the shear strength between the particles. In addition to adding 4% of the crude oil, the lubricating effect of the oil is thought to cause easy soil particles to slide over each other, which represents the decrease in the unsoaked CBR.

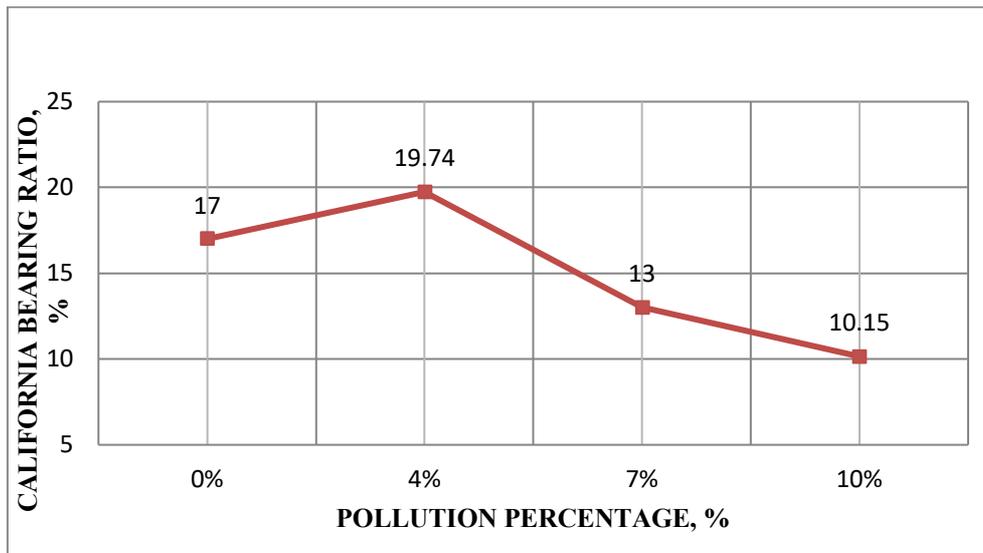


Figure 4: Variation of unsoaked CBR with oil content

5.5 RESULTS OF PERMEABILITY TEST:

A graphical illustration of the results of the permeability tests on the soil admixed with varying percentages of crude oil content is presented in Figure 5. As the crude oil content increased, the permeability of the contaminated soil decreased. Crude oil becomes entrapped in the pore spaces that forms the pathway for water within the contaminated soil and consequently, reduced the pore sizes. The decrease in the permeability of the contaminated soil is attributed to the reduction in the pore space.

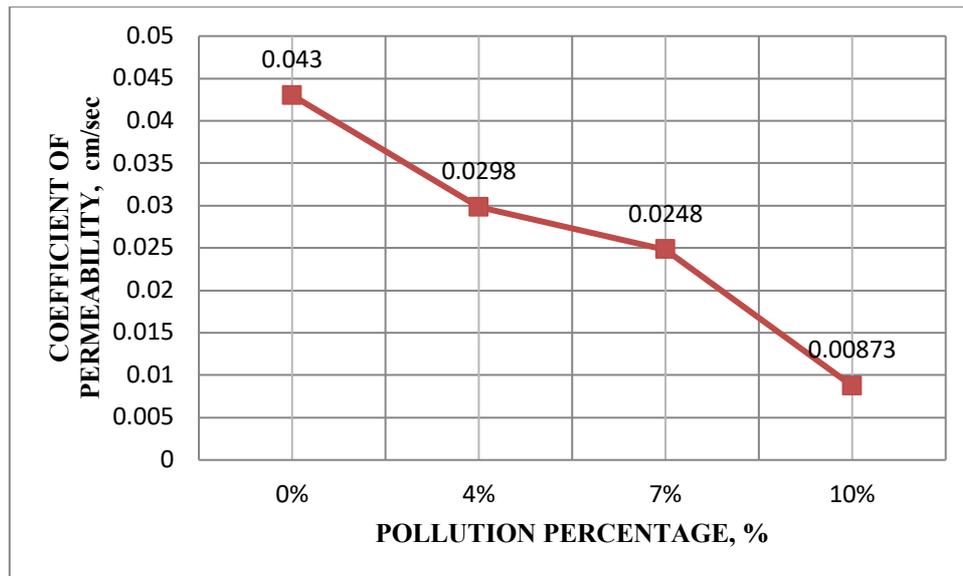


Figure 5: Variation of permeability with oil content

6- CONCLUSIONS AND RECOMENDATIONS:

6.1 CONCLUSIONS:

As a result of the test program that was conducted in this study and the discussions, the following can be concluded:

1. Oil contamination of soil affects the physical properties of soil, as its specific gravity decreased with an increase in the percentage of contamination by crude oil.
2. The compaction tests showed that the increase in the oil content led to an increase in the maximum dry density and a decrease in the optimum water content. The explanation increased the maximum dry density is that the oil is more effective in reducing the friction between soil particles resulting in reduced spacing between soil particles; Thus increasing the dry density at a certain compaction effort.
3. The unsoaked CBR of the soil increases to a certain extent, so its highest value is at 4% of the pollutant. After that, a decrease occurs with an increase in the percentage of crude oil in the soil.
4. A similar behavior was also observed on cohesion (C) when crude oil was increased to a certain limit, so the highest value was at 4% of pollution, and the lowest value was at 0% of pollution. While the internal friction angle (ϕ) decreases with the increase in the proportion of crude oil.
5. Adding crude oil to the soil reduced its permeability, which is due to the confinement of crude oil inside the soil pores. The highest permeability value was when the soil was free of pollution and the lowest value was obtained at 10% of pollution.

6.2 RECOMMENDATIONS:

1. To obtain an effective soil treatment with positive results, it is preferable to start working on soil improvement immediately after stopping the leakage.
2. Since the viscosity of oil and some of its other properties varies according to the different places of exploration and exploration, from field to field and from one region to another according to the materials that the oil is made of over hundreds of years, the change in viscosity may have different effects on the properties of the polluted soil.

3. Construction is not recommended on untreated polluted soils due to the change in its properties.

7- ACKNOWLEDGMENT:

We thank Al-Alya Engineering and Technical Consulting Company for the assistance of sample preparation and soil testing.

REFERENCES:

- [1]. American Petroleum Institute, Sampling and analysis of gasoline range organics in soils. American Petroleum Institute Publication, No. 45 16, 1991.1.5
- [2]. Hinchee, R.E. and Olfenbuttel, R.F. Butterworth-Heinemann, in Situ Bioreclamation: Application and Investigations for Hydrocarbon and Contaminated Site Remediation. Ed. by. London. pp. 125-142.
- [3]. Dineen. D., Slater, J.P., Jicks, P. and Holland, In situ biological remediation of petroleum hydrocarbons in unsaturated soils, Hydrocarbon Contaminated Soil and Ground Water: Analysis, Fate, Environmental and Public Health Effects, Remediation. Ed by KostECKI, P.T. and Calabrese, E.J. Vol. 1. Lewis Publishers, Chelsea, Mich., pp. 177-187, 1990.
- [4]. Okagbue, C. O., and Onyeobi, T. U. S. 1999. Potential of marble dust to stabilise red tropical soils for road construction. Engineering Geology, 53, 371-380. DOI:10.1016/S0013-7952(99)00036-8
- [5]. Akinwumi, I. I., Adeyeri, J. B., Ejohwomu, O. A. 2012. Effects of steel slag addition on the plasticity, strength and permeability of a lateritic soil. Proceedings of Second International Conference on Sustainable Design, Engineering and Construction, ASCE, Texas, pp.457-464. DOI:10.1061/9780784412688.055
- [6]. Akinwumi, I. I. 2012. Utilization of steel slag for stabilization of a lateritic soil. M.Eng. Dissertation, Covenant University, Ota.
- [7]. Akinwumi, I. I. 2014. Plasticity, strength and permeability of reclaimed asphalt pavement and lateritic soil blends. International Journal of Scientific and Engineering Research, 5(6), pp.631-636.
- [8]. Sanjay J.Shah, Shroff A.V, Jignesh v., Tiwari K.C. & Ramakrishnan D.(2002)"Stabilization of Fuel Oil Contaminated Soil"Journal of Geotechnical &Geological Engineering,Vol.21,No.4.
- [9]. Karkush, M. O, Zaboon, A. T. and Hussien, H. M., 2013, Studying the effects of contamination on the geotechnical properties of clayey soil, Coupled Phenomena in Environmental Geotechnics, Taylor & Francis Group, London, pp. 599-607.
- [10]. George, S., EA, A., Sabu, B., NP, K. and George, M., 2015, Effect of diesel oil on the geotechnical properties of soil, International Journal of Civil and Structural Engineering Research, Vol. 2, No. 2, PP. 113-117.
- [11]. Al-Harbawy, A.F.Q (2002)"The effect of Emulsified Asphalt on Some Engineering Properties for Expansive Soil" M.Sc. Thesis, Department of Civil Engineering, College of Engineering, Mosul University Iraq.
- [12]. Ingles O. G. & Matcalf J.B. (1972)"Soil Stabilization Principle & Practice).
- [13]. Ramzi Taha, Amer Al-Rawas, Salim Al-Oraimi, Hossom Hassan& Mohammed Al-Aghbari (2005)"The Use of Brackish &Oil Contaminated Water in Road Construction" Environmental &Engineering Geoscience, Vol.11,No.2.
- [14]. Das. B.M & Eun chal shin (2001)"Bearing Capacity of(- 111 - (Unsaturated Oil-Contaminated Sand" International Journal of Off share &Polar Engineering ,Vol.11, No.3.

- [15]. Yong R.N ,Taheri E. ,Khodadai a.(2007)"Evaluation of Remediation Methods for Soil Contaminated with Benzo" International Journal of environmental research ,Vol.1,No.4.
- [16]. Purj, V.K., B.M. Das, E.C. Cook and E.C. Shin, "Geotechnical Properties of Crude Oil-Contaminated Sand". ASTM Special Technical Publication, 1221, 7 Y 5-88, 1994.
- [17]. H. Al-Sanad, W.K. Eid, and N.F. Ismael, "Geotechnical Properties of Oil- Contaminated Kuwait Sand", Journal of Geotechnical Engineering, No. 5, Vol. 121, pp. 407-412, May 1995.
- [18]. Rahman, Z. A., Hamzah, U., and Taha, M. R., 2010, Influence of oil contamination on geotechnical properties of basaltic residual soil, American Journal of Applied Sciences, Vol. 7, No. 7, PP. 954–961.
- [19]. Izdebska-Mucha, D., and Trzcinski, J. 2008. Effects of petroleum pollution on clay soil microstructure. Geologija,50 Supplement P),S68-S74. DOI: 10.2478/V10056 -008-0027-0
- [20]. Olgun, M., and Yildiz, M. 2010. Effect of organic fluids on the geotechnical behavior of a highly plastic clayey soil. Applied Clay Science, DOI: 10.1016/clay .2010.03.015.
- [21]. Oyegbile, O. B., and Ayininuola, G. M. 2013. Laboratory studies on the influence of crude oil spillage on lateritic soil shear strength: A case study of Niger Delta Area of Nigeria. Journal of Earth Science and Geotechnical Engineering, 3(2), pp.73-83.
- [22]. Khamsehchiyan, M., Charkhabi, A. M., and Tajik, M., 2007, Effects of crude oil contamination on geotechnical properties of clayey and sandy soils, Engineering Geology, Vol. 89, PP. 220-229.
- [23]. Kermani, M., and Ebadi, T. 2012. The effect of oil contamination on the geotechnical properties of fine-grained soils. Soil and Sediment Contamination, DOI:10.1080/15320383.2012.672486.
- [24]. "Chemical and Physical Properties of Crude Oils", www.nap.edu, Retrieved 28-9-2018. Edited.
- [25]. Anne Marie (2018-3-23) " Chemical Composition of Petroleum. Thoughtco Retrieved 2018-7-14. Edited.

Effect of Sawdust on The Mechanical Properties of Mortar Using Local Materials

Nurdeen Mohamed Altwair^{1*}, Lamen Saleh Sryh²

¹nmaltwair@elmergib.edu.ly ²lssryh@elmergib.edu.ly

^{1,2} Department of Civil Engineering, College of Engineering, Elmergib University, Libya

ABSTRACT

Keywords:

sawdust;
cement mortar;
sound conductivity;
mechanical properties,
air content.

In recent years, there has been a great tendency, especially in the field of structural engineering, to use environmentally friendly materials. Therefore, such these materials should be combined with concrete and cement mortar to obtain an environmental and cheaper structural component that has distinctive properties such as lower thermal and sound conductivity and produce lightweight members. In this study, the sand in the cement mortar of the reference mix was replaced by residues of sawdust for different levels ranging between (5, 10, and 15%) by volume. Two methods were used to cure the specimens of this study; the first by immersing the specimens in water, and the second by exposing them to the air under laboratory conditions. Different tests were conducted in order to assess; the air content of fresh mortar, consistency of fresh mortar, compressive strength, flexural strength, and conductivity. From the obtained results, as the percentage of sawdust increased, a significant improvement in sound conductivity was observed. Although the compressive and flexural strength decreased with increasing the percentage of sawdust (up to 15%), this decrease can be considered acceptable for several construction applications. Moreover, the increasing in the percentage of sawdust was noted to be accompanied by an increasing in the percentage of air content. This allows obtaining a lightweight cement mortar with suitable mechanical properties which can be used for various constructions.

1. Introduction

Many researches have been done on the possibility of using locally available materials to partially or fully replacement to the cement or aggregate in the concrete or mortar. Most of these materials are generally noted to be more expensive constituents. The problem of stacking of unmanaged wastes, especially in developing countries, has resulted in increasing the environmental concerns. Recycling of such wastes appears to be applicable solution not only for pollution problem but also for the economic of buildings design. The growing in the tendency to use environmentally friendly, low-cost and lightweight construction materials in building industry has brought about the need to investigate how this can be achieved and to how this can be benefit to the environment [1].

At the beginning of the 1990s, several studies had been conducted to use the industrial wastes in various forms of concrete productions. For instance, the use of paper waste sludge, glass powder, waste rubber and palm oil fuel ash in concrete mixes has received remarkable interest for the study [2-7]. The waste of sawdust is one of these materials which has used in concrete or cement mortar productions. Sawdust is an organic waste resulting from the mechanical milling or processing of wood into various shapes and sizes [8]. The use of sawdust wastes can not only solve the environmental concerns, but also can protect the natural resources of the construction materials. It has considerable features, such as low bulk density, superior heat preservation and lower thermal

and sound conductivity, and can reduce the environmental pollution that resulting from dumping such wastes in close places to the residential communities. The physical and chemical properties of sawdust vary significantly depending on several factors according to the type and specification of the wood [9].

The accumulation of fine sawdust waste in many carpentry workshops, especially in developing countries such as Libya, may cause certain serious environmental problems and health hazards [10]. This study presents an experimental work which investigates the possibility of using the sawdust waste to produce a suitable lightweight mortar. Some of the mechanical properties of mixes that having various levels of sawdust waste have been investigated. The possible successful results in using the sawdust in the construction applications could be open the way for using other agricultural wastes such as the straw, which would achieve more benefits for protecting the environment.

2. Materials and Methods

2.1 Materials

Ordinary Portland cement (OPC) that complies with the requirements of BS EN 197-1:2011 was used in the experimental work of this study. The physical properties and chemical compositions of OPC is provided in Table 1.

Table 1. Chemical compositions and physical properties OPC.

Chemical composition (mass %)		Physical properties	
Component	Value	Property	Value
Silicon dioxide (SiO ₂)	21.14	Specific gravity	3.15
Aluminum oxide (Al ₂ O ₃)	5.91	Specific surface area(m ² /g)	2926
Ferric Oxide (Fe ₂ O ₃)	3.99	Strength activity index at 3 days (MPa)	25
Calcium oxide (CaO)	61.9	Strength activity index at 28 days MPa	42
Magnesium Oxide (MgO)	2.59		
Sodium oxide (Na ₂ O)	0.18		
Potassium oxide (K ₂ O)	0.88		
Sulfur oxide (SO ₃)	1.11		
Phosphorus oxide (P ₂ O ₂)	0.9		
LOI	0.41		

Natural sand with maximum size of 1.2 mm was used as a fine aggregate which collected from quarries of Zlitan City. The fine sand has a fineness modulus of 2.8, specific gravity of 2.62 and water absorption of 0.83%.

The sawdust was obtained from carpentry workshops in Al-Khums City, considering their purity from stones, dust and the large sizes of husks. The used sawdust has a maximum size of less than 600µm. For the purpose of knowing the grain size distribution, the used sawdust was sieved through a set of sieves of 0.6, 0.425, 0.25, 0.15mm (Figure 1). In this study, sawdust was cured before being used as a fine aggregate in producing the cement mortar. Firstly, it was immersed in water and saturated with hydrated lime for the purpose of getting clear of harmful substances. Then, it was submerged in Kerosene for 24hrs to prevent its liability to water absorption (Figure 2).

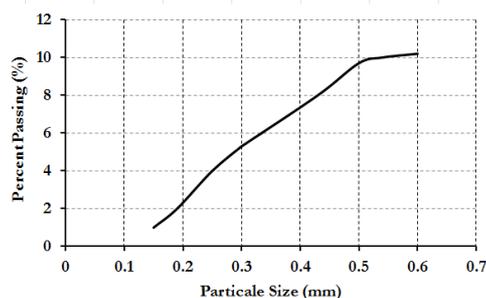


Figure 1: Grain size distribution of sawdust.

Figure 2: Sawdust waste after treatment.

2.2 Mixes Proportions

Due to the remarkable differences in the specific gravities of sand and sawdust, the process of replacing sand by sawdust was done by volume. The replacement percentages of sand by sawdust were 5,10, and 15%. This was done to find out the optimal mix that would give the most appropriate results. The mix of 0% replacement (C) was to serve as reference mix for the other mixes (S1, S2 and S3). The details of mixes are given in Table 2. A 5-liter mortar mixer was used to mix the mortar ingredients at a constant mixing speed of 30rpm. After mixing and casting the mortar into the molds, all specimens were kept in room temperature for 24 hours. After that, the specimens were demolded and cured by following two methods, the first by immersing the specimens in water and the second by exposing them to the air under laboratory conditions. Then, the specimens were left until the day of the test.

Table 2. Mixes proportions.

Mix ID	Cement	Water	Sand	Sawdust
C	1	0.4	1.5	0
S1	1	0.4	1.425	0.075
S2	1	0.4	1.35	0.15
S3	1	0.4	1.275	0.225

2.3 Tests Procedures

The Flow Table test aims to find out the effect of using the sawdust on the fresh mortar consistency. Through this test it would be possible to give an idea about the absorbed water by the sawdust. This test was performed according to ASTM C230/C230M- 15 [11].

For determining the percentage of air voids that existing in the fresh mortar which containing sawdust, the Air Content test was confirmed, this test was conducted according to ASTM C185 – 15 [12].

The adopted geometry for the compressive strength specimens was a cube of 7.05×7.05×7.05cm. The cubes were tested at 7, 14 and 28 days, and the specimens were tested under compression by using the standard procedure for testing the cubes of compressive strength test according to ASTM C109/C109M [13]. The test was performed under a load rate of approximately 1.8 KN/m.

The flexural strength of mortar is determined by conducting the test method of ASTM C78/C78M [14] (Third-point loading or 4-point bending test). Prisms with size of 16x4x4cm were tested. The flexural strength at 7 and 28 days were determined by using four-point loading test. The specimens rest on two supports and the half of the load is applied at each of the third of the span length by using a rigid piece of steel. The specimen was loaded at constant rate until the failure.

In this study, Ultrasonic Pulse Velocity test was approved, with a view to assess the effect of replacing sawdust by sand on the acoustic properties such as sound transmission. This test is conducted by passing a pulse of ultrasonic through specimens of cubes and measuring the time taken by pulse to get through the specimens. Higher velocities indicate higher sound conductivity and continuity of the material, while slower velocities may indicate mortar with lower sound conductivity. The calculation of the ultrasonic velocity average was taken according to the age of specimens, and the test was carried out according to ASTM C597- 09 [15].

3. Tests Results and Discussion

Fresh Mortar Consistency and Air Content

The results of Flow table test of the tested mortars are presented in Figure 3. For all the tested mortars, the flow decreased with increasing the sawdust content. The lower flow in the mortars containing sawdust compared to the control mortar may be attributed to a portion of the mixing water that has been absorbed by the sawdust (although sawdust is soaked in Kerosene to prevent the absorption). In addition, the reason may be due to the roughness of the surface of the sawdust grains which influences the friction between the components of the mortar and causing a lack of flow.

Figure 4 illustrates the effect of replacing the sand by sawdust with different percentages on the air content of the fresh mortar. It can be observed that the increasing in the percentage of sawdust has a significant effect on increasing the air content. By comparing the results of this test, it was observed that the air content in the reference mortar (C) was about 0.36%, and after replacing the sand by 15% of sawdust (S3), the air content became about 1.8%, causing an increase of about 500%. The reason may be due to the difference in the specific gravity of sawdust and sand (S.G of sand 2.66 is much greater than that of sawdust 0.9), this may contribute significantly to the formation of air voids in the fresh mortar containing the sawdust.

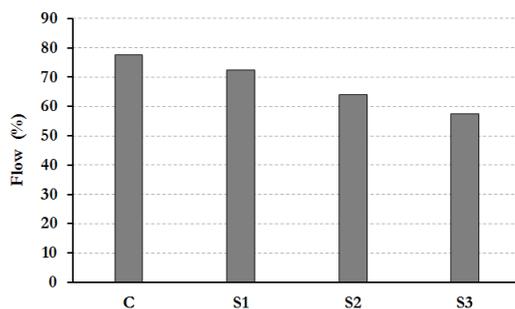


Figure 3: Flow table test results of mortars with different proportions of sawdust.

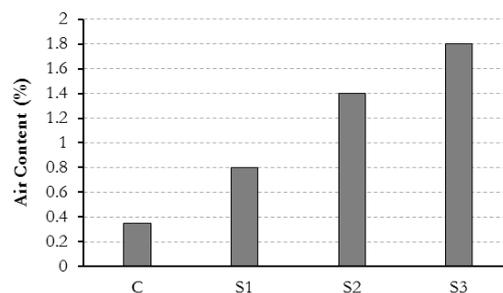


Figure 4: Air content results of mortars with different proportions of sawdust.

Compressive Strength Comparison

Figures 5 and 6 show the results of the compressive strength values obtained from the tests. The average compressive strength values are inversely proportional with the replacement percentage of sawdust cured by water and air under laboratory conditions respectively. The strength dramatically decreases with an increase in the replacement level of sawdust.

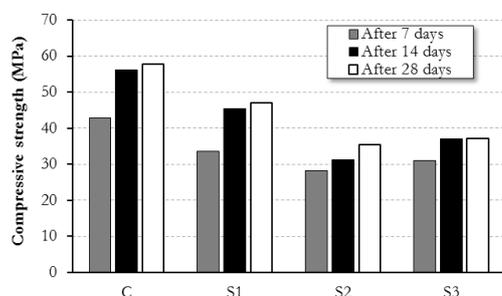


Figure 5: Compressive strength results of mortars cured by water immersion with different proportions of sawdust.

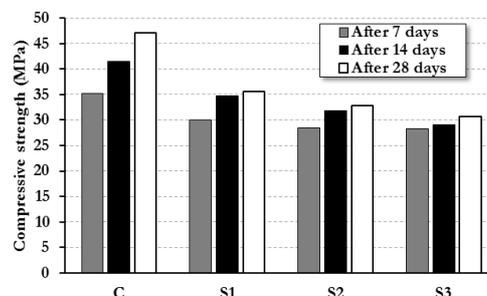


Figure 6: Compressive strength results of mortars cured by exposing to air with different proportions of sawdust.

In both methods of curing (air or water), the strength decreases with an increase in the replacement level of sawdust. According to the comparison of compressive strength results that is shown in Figure 5, the strength of the standard mortar that were cured by immersing in water after 7 days was 42.95 MPa, while it was 57.75 MPa after 28 days. However, the specimens with the highest replacement percentage (15%) were 31 MPa and 37.06 MPa after 7 and 28 days, respectively. Wherever, the percentage of decrease in compressive strength was about 39% and 56% after 7 and 28 days, respectively. Through Figure 6, it can be noted that the effect of replacing sawdust on compressive strength for air-cured specimens is not much different from water-cured specimens, but it is lower in the compressive values. Comparatively, the results indicate that sawdust as a partially replacement of sand mortar can achieve the same order of strength as conventional mortars at same curing periods.

Flexural Strength Comparison

Figure 7 and 8 show the relationship between the average of flexural strength and the percentage of replacing the sawdust with sand at different ages for the tested mortars that cured by air-curing and water immersion. Similar behavior to that noted for the compressive strength was also observed for the flexural strength. However, the flexural values were decrease slightly for the all replacement levels. At the 7 and 28 days water curing period, the range of flexural strength obtained was between 1.63-2.16MPa (for 0% sawdust content) and 1.35-1.34MPa (for 15% sawdust content) as shown in Figure 7. In other words, the flexural strength decreased by 21% and 61% at 7 and 28 days curing period, respectively. However, despite the difference in the water curing period, it can be observed that as the (10% and 15%) of sand was replaced by sawdust, the strength was approximately equal.

The variations of flexural strength for specimens cured by exposing to air under laboratory conditions are presented in Figure 8. The effect of the replacing of sand by sawdust is almost similar for both curing methods. However, the flexural strength was lower compared to that obtained from specimens cured in the water. Generally, the flexural strength test results indicate that the strength decreases with increasing in sawdust content for all the ages and curing methods.

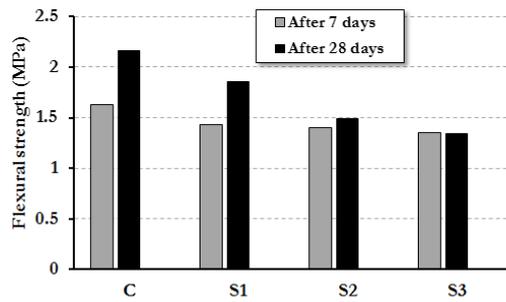


Figure 7: Flexural strength results of mortars cured by water immersion with different proportions of sawdust.

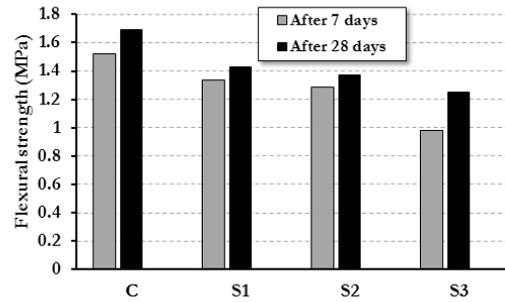


Figure 8: Flexural strength results of mortars cured by exposing to air with different proportions of sawdust.

The decrease in flexural strength can be attributed to poor adhesion is present at the interface due probably to increasing quantity of water absorption by the sawdust grains which causes low efficient hydration reaction on the cement matrix. However, the mortars produced by replacing up to 15% sawdust, the strength is still permissible for using in plain construction applications and that can be used as lightweight mortar. According to the conditions of this study, it is expected that using higher contents of sawdust (more than 15%) seems to deteriorate the flexural strength.

Ultrasonic Pulse Velocity Comparison

Figures 9 and 10 illustrate the effect of sawdust on the ultrasonic pulse velocity of the mortars cured by water and air under laboratory conditions respectively. The test results confirmed that the velocity values are inversely proportional with the replacement percentage of sawdust. For all studied mortars, whether submerged in water or exposed to air, the ultrasonic pulse velocity dramatically decreases with an increase in the replacement level of sawdust. It was observed that for specimens cured by immersing in the water, approximately 19% and 73% as a reduction in the velocity of control mix is obtained from the 15% of sawdust replacement at 7 and 28 days, respectively. On the other hand, the specimens cured by exposing to the air, the reduction was 31% and 30%.

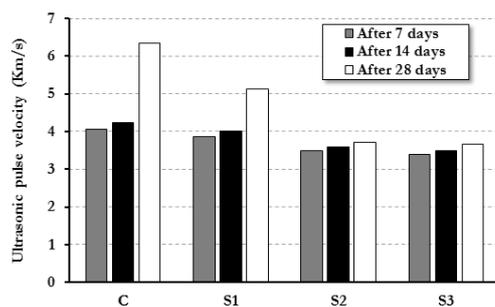


Figure 9: Ultrasonic pulse velocity results of mortars cured by water immersion with different proportions of sawdust.

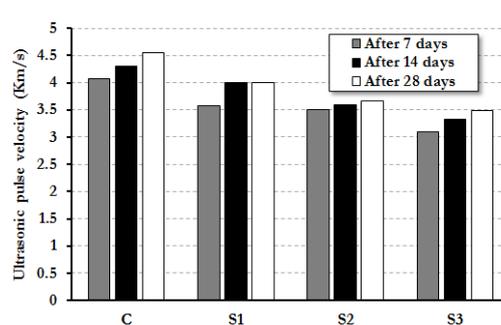


Figure 10: Ultrasonic pulse velocity results of mortars cured by exposing to air with different proportions of sawdust.

The decrease in the ultrasonic pulse velocity values can be attributed to the fact that part of the waves was absorbed by the grains of sawdust. In addition, the presence of air voids that increase the cause of the delay in the transmission of the waves. It can also be observed that the curing

method had a great role in the effect on the ultrasonic pulse velocity results, i.e. the curing by immersing in water increased the velocity more than the curing by exposing to air. For all the curing ages, it is also interesting to note that there is a convergence in ultrasonic pulse velocity values for mortars containing sawdust and cured by exposing to air. However, the results provided by this test indicated that mortars less sound conductivity could be manufactured by adding sawdust to the mixture.

4. Conclusion

From the obtained results of the conducted tests in this study, the following conclusion can be drawn:

- An increase in the percentage of sawdust is accompanied by a noticeable increase in air content, which in result may increase the chance of obtaining lighter weight mortars. Moreover, this can reduce the transmission of sound and increase the possibility of using this type of mortar in various constructions that need to be loaded with light loads.
- Comparatively between the mortars containing sawdust and the standard mortar, it was concluded that the percentage of flow decreases as the replacement percentage of sawdust by sand increases. This could be explained by the fact that when the sawdust increases, the more mixing water is absorbed by the sawdust grains, which reduces the workability, and leads to an increase in friction hence reduce the flow.
- The flexural and compressive strengths of the prisms and cubes respectively which made with local materials and containing sawdust, decrease with the increase of sawdust level in the mix. However, the strength at a replacement level of 15% by the volume of sand is considered reasonable for various construction applications.
- By analyzing the results obtained from the ultrasonic pulse velocity test, it can be concluded that an increase in the percentage of sawdust leads to a decrease in the velocity. This can reduce the possibility of sound transmission through the structural elements which made with mortars containing sawdust.
- Overall, it can be concluded that adding sawdust to cement mortar produces a mortar with less sound conductivity, lightweight and more environmental and economical.

Acknowledgment

The authors would like to thank the Public Works Company in Al-Khums city for its support by providing the equipment and materials for the experimental work.

References

- [1]. T. Paki, M. Halil, "Limestone dust and wood sawdust as brick material," *Building and Environment*, vol. 42, no. 9, pp. 3399–3403, Sep. 2007. Access online on 24 August 2020 at <https://doi.org/10.1016/j.buildenv.2006.08.012>
- [2]. W. Shao, T. Moras, S. Moras, D. Rodriguez "Studies on concrete containing ground waste glass," *Cement and Concrete Research*, vol. 30, no. 5, pp. 91-100, Jan. 2000. Access online on 30 August 2020 at <https://www.sciencedirect.com/science/article/abs/pii/S0008884699002136>.
- [3]. B. Topcu, M. Canbaz "Properties of concrete containing waste glass," *Cement and Concrete Research*, vol. 32, no. 2, pp. 267-273, Feb. 2004. Access online on 28 August 2020 at <https://www.sciencedirect.com/science/article/abs/pii/S000888460300262X>.
- [4]. B. Ahmadi, W. Al-Khaja "Utilization of paper waste sludge in the building construction industry," *Resources, Conservation and Recycling*, vol. 32, no. 2, pp. 105-113, Jun. 2001. Access online on 21 August 2020 at <https://www.sciencedirect.com/science/article/abs/pii/S0921344901000519>.

- [5]. Sofi, "Effect of waste tyre rubber on mechanical and durability properties of concrete – A review," *Ain Shams Engineering Journal*, vol., 9 no. 4, pp :2691-2700, Dec. 2018. Access online on 30 August 2020 at <https://www.sciencedirect.com/science/article/pii/S2090447917301132>.
- [6]. N. M. Altwair, M. A. Megat Johari, S. F. S. Hashim, "Strength Activity Index and Microstructural Characteristics of Treated Palm Oil Fuel Ash," *International Journal of Civil & Environmental Engineering*, vol. 5, no. 11, pp. 100-107, Oct. 2011. Access online on 30 August 2020 at <http://ijens.org/IJCEE%20Vol%2011%20Issue%2005.html>
- [7]. N. M. Altwair, M. A. Megat Johari, S. F. S. Hashim, "Influence of treated Palm Oil Fuel Ash on Compressive Properties and Chloride Resistance of Engineered Cementitious Composites," *Materials and Structures*, Vol, 47, no. 4 pp. 122-138, May 2014. Access online on 25 August 2020 at <https://link.springer.com/article/10.1617/s11527-013-0087-4>
- [8]. O. Adekunle, A.S. Daramola, "Optimisation for the Use of Rice Husk Ash and Sawdust As Alternative Binder For Concrete," *International Journal Of Engineering And Science*, Vol. 2, no. 10, pp. 39-42, Oct. 2013. Access online on 30 August 2020 at <http://www.theijes.com/Vol,2,Issue,10.html>.
- [9]. Yong Cheng, Y. Wen, C. Zhang, H. Li, J. Hu, "The Implementation of Waste Sawdust in Concrete," *Advanced Materials Research*, Vol. 5, no. 12 pp. 943-947, Jun. 2013. Access online on 27 August 2020 at <https://www.scientific.net/AMR.941-944.849>.
- [10]. L. M. H. Bdeir, "Study Some Mechanical Properties of Mortar with Sawdust as a Partially Replacement of Sand," *Anbar Journal for Engineering Sciences*, Vol. 5, no. 1, pp. 22-30, Apr. 2012. Access online on 29 August 2020 at <https://www.iasj.net/iasj?func=article&aId=41133>.
- [11]. ASTM C230/C230M- 15 "Standard specification for flow table for use in tests of hydraulic cement," ASTM International: West Conshohocken, PA, USA, 2015.
- [12]. ASTM C185 – 15, "Standard Test Method for Air Content of Hydraulic Cement Mortar," ASTM International: West Conshohocken, PA, USA, 2015.
- [13]. ASTM C109 / C109M – 15, "Standard Test Method for Compressive Strength of Hydraulic Cement Mortars," ASTM International: West Conshohocken, PA, USA, 2015.
- [14]. ASTM C78/C78M – 16, "Standard Test Method for Flexural Strength of Concrete (Using Simple Beam with Third-Point Loading) ," ASTM International: West Conshohocken, PA, USA, 2016.
- [15]. ASTM C597 – 09 "Standard Test Method for Pulse Velocity Through Concrete," ASTM International: West Conshohocken, PA, USA, 2009.

Experimental Study of Drowned Hydraulic Jump Characteristics Through Different Counterflow Dimensions

Amad.Deen Abdusalam Alghwail^{1*}

¹ civileng_amad@yahoo.com

¹ Department of civil Engineering, Faculty of Engineering/Alkhums, Elmergib University, Libya

ABSTRACT

Hydraulic jump is a phenomenon that occurs whenever the flow profile changes from supercritical to subcritical. In this transition, the water surface suddenly raised, surface rollers are formed, intense mixing occurs, air is entrained, with a considerable amount of water energy is dissipated. This paper presents the results of an experimental investigation on the characteristics of the formed drowned hydraulic jump on a horizontal slotted bed, (counterflow). In the current study, the counterflow was used as an energy dissipation method. Experiments were carried out to study the effect of different counterflow dimensions, represented in slot position, inclination angle of slot and slot width on both the length and the submergence ratio of the drowned jump. Graphical presentations were given, describe the relations between the percentage reduction in the length of the hydraulic jump and Approach Froude number due to variation of slot dimensions, for a Froude number ranges from 8.74 to 13.45. The results show that both the percentage reduction in the jump length and the submergence ratio increase as Froude number increases.

Keywords:

Hydraulic Jump,
Energy Dissipation,
Counterflow, Froude
Number

1- Introduction

Hydraulic jump is a phenomenon that occurs whenever the flow profile changes from supercritical to subcritical. In this transition, the water surface suddenly raised [1], surface rollers are formed, intense mixing occurs, air is entrained, with a considerable amount of water energy is dissipated. Hydraulic jump can be classified in two types according to their bed characteristics. The first type is a classical hydraulic jump with a smooth bed, and the second one is a forced hydraulic jump. The first type has been extensively studied by some Authors [2-6].

Conditions in a channel, such as downstream controls, that can change where the conjugate depths form, in addition, tailwater depth can play a very effective role on the location of the hydraulic jump in a channel, and changes in this depth can move the formed jump either upstream or downstream. In a situation where a downstream control, such as sluice gate, forces the tailwater elevation to a depth above the original conjugate depth (subcritical conjugate depth), as shown in Figure 1. so the jump is pushed upstream. In this case, the sluice gate inhibits the movement of the jump upstream so that the upstream conjugate cannot be attained. This leads to a situation known as a drowned or submerged hydraulic jump [7].

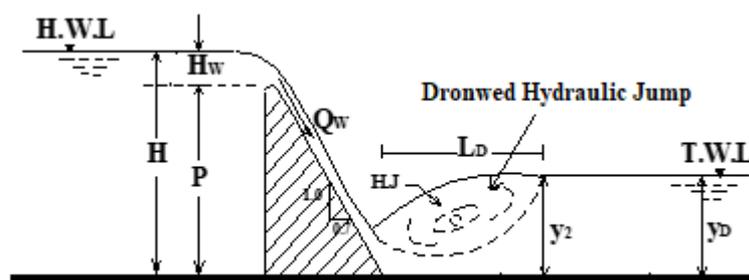


Figure 1: Formation of a drowned hydraulic jump

1.1 Dissipation of Water Energy and Hydraulic Jump

Hydraulic structures such as dams and barrages, spillways, sluice gates are commonly constructed in streams having erodible bed materials. In severe weather conditions, dam spillway is used to prevent the dam from being overtopped. A large amount of water to be released through spillway from the reservoir over a short space of time, resulting in flow of very high discharge and velocity. This normally takes the form of highly turbulent supercritical flow in the spillway and its immediate downstream of the channel. The high flow velocity leads to a bed shear stress greatly higher than in the absence of the structure, causing significantly increased sediment transport downstream of the structure. The channel bed level becomes eroded as a result and this is commonly known as the local scour.

Numerous of comprehensive studies, whether experimental or theoretical have been conducted to investigate the problem of water energy dissipation of hydraulic structures, with the aim of dissipating a larger amount of water energy, in order to protect the downstream bed of such structures against scour process, consequently, avoiding potential structural damage risks. Creation of hydraulic jump in open channel is a useful phenomenon. It is generally used to dissipate of excess kinetic energy downstream hydraulic structures, such as drops, spillways, chutes and gates, increasing weight on an apron and thus reduce uplift pressure under control structures.

Dissipation of water energy downstream hydraulic structures has been dealt with in different ways of approach, based on the two following concepts:

- (1) Provoking large velocity gradient lead to increasing turbulence in the stream.
 - (2) Creating extended and turbulent interfaces between the flowing water and the surrounding air.
- Accordingly, several methods and techniques have been utilized, either to improve and increase the efficiency of the existing dissipators or to find new ones, that satisfy both high effectiveness and minimum cost of construction. These methods or techniques might be classified as follows:
- (1) Energy dissipation by using stilling basins,
 - (2) Energy dissipation by counter flow and
 - (3) Energy dissipation on slopping surfaces.

Tailwater is generally characterized by small velocities resulting in a subcritical flow condition. The flowing water either over weirs, under sluices or through pipes is characterized by a high velocity jet. When a supercritical flow issuing downstream hydraulic structures, impacts the slow tailwater, a free hydraulic jump zone is formed having a large velocity gradient, Figure 2-A. When the flow level at the tailwater is too low, the free jump will be shooting, extending to a larger length; then the velocity gradient becomes small, consequently, the stilling basin must be lengthened so that jump and eddies zones will finish within the length of the stilling basin, which increases the construction costs.

To restrain the free jump, energy dissipators must be applied, to help reducing the size of stilling basins by moving the hydraulic jump forward, Figure 2-B, the formed jump in this case is then called a forced jump. The efficiency of such energy dissipation methods can have a significant impact on the overall cost of a project. As a result, there have been great efforts amongst engineers and researchers towards developing efficient but also cost effective solutions, Therefore, most of the classical methods of water energy dissipation, are mainly aimed at creating a forced jump, whether perfect or drowned.

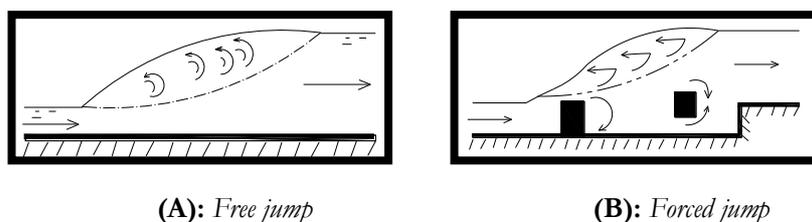


Figure 2: Formation of free and forced hydraulic jump

2- Experimental Study

Experimental work was conducted in the experimental set-up, specially prepared to study the effect of the main parameters concerned with the cross jet on the characteristics of the formed hydraulic jump considering the three cases: free perfect jump, without cross jet, forced perfect jump, with cross jet and drowned jump, with cross jet.

The model was installed in the experimental set up prepared at the laboratory of fluid mechanics, **Faculty of Engineering, Elmergib University**. Figure 3 shows the components of the experimental set-up which consists of the following parts [8]:

2.1 The Testing Flume

The testing flume comprises a rectangular section of channel (1) with inlet (2) and discharge (3) tanks. The testing flume is rested on a pair of rigid pedestals (4). A service module (5) incorporating a sump tank (6) and submersible pump (7) provides a source of water which is continuously recirculated through the channel section making a closed circuit of water supply.

Laboratory experiments are carried out on a rectangular open channel flume 5 m long and rectangular cross section 25 cm high by 7.6 cm wide. The sides are fabricated from 10 mm transparent Perspex sheets which are bonded to a bed fabricated from painted aluminium alloy. The end tanks are constructed from glass reinforced plastic with a smooth gel coat on the inside. Water enters the working section via inlet tank (2). The sides of the inlet tank are profiled with a smooth contraction towards the working section. To reduce the turbulence of the water entering the inlet tank and to produce a smooth flow of water, the pipe supplying the inlet tank (8) has diffused outlet and is covered by perforated plate and glass marbles.

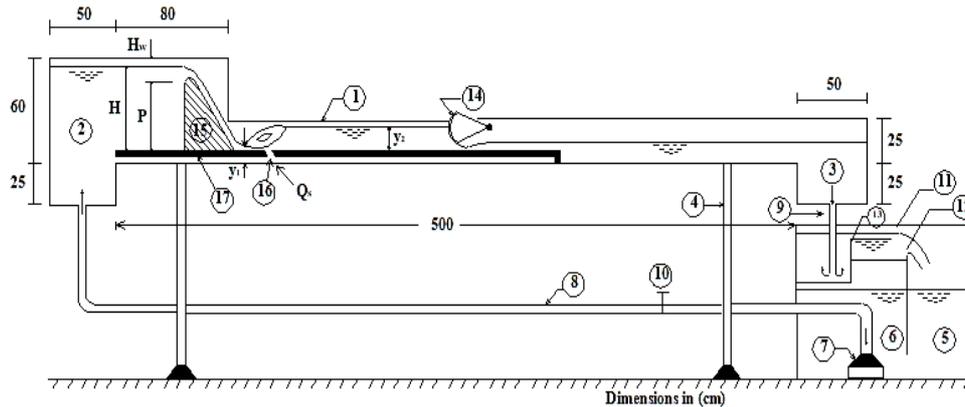


Figure 3: General layout of the experimental set-up

Legend: -

- | | | |
|---------------------|------------------------|-----------------------|
| (1) Testing flume. | (7) Submersible pump. | (13) Screen. |
| (2) Inlet tank. | (8) Supplying pipe. | (14) Control gate. |
| (3) Discharge tank. | (9) Discharge pipe. | (15) Spillway model. |
| (4) Pedestal. | (10) Control valve. | (16) Slot. |
| (5) Service module. | (11) Moulded channel. | (17) structure floor. |
| (6) Sump tank. | (12) Rectangular weir. | |

Water exiting from the flume enters the discharge tank (3) where it returns by gravity, through the discharge pipe (9), to the service module (5).

The service module is constructed from glass reinforced plastic. Water is drawn from the sump tank in the base of the service unit by the submersible pump (7). The water is delivered to the flume through the supplying pipe (8) which has a flow control valve (10). Water returns from the flume discharge tank to a moulded channel (11) on the top of the service module. The water then flows over rectangular weir (12) to the sump tank by gravity. The moulded channel has a screen (13) to damp any disturbance caused by the discharge pipe, so that water surface being smooth before falling over the weir. A radial gate (14) is installed in the working channel to control the water depth downstream of the spillway.

2.2 Dissipation Model

The tested model is made of Perspex sheets; it comprises of the following two components:

The inclined surface, of the spillway, is represented by an inclined plate (15) joining the crest of the ogee weir to the testing flume bottom. It has 10 mm thick, 1.0 (Vertical) to 0.7 (Horizontal) slope and 43.5 cm height, as shown in the above slope of the back face of the spillway was chosen according to the previous studies conducted by some Authors [9-12]. Such a given slope prevents separation of flow from the inclined surface.

The slot (16) was formed through a Perspex sheet (17) of 15 mm thickness, represents the structure floor. Slot's angle, width, and position are selected according to the considered values listed in Table 1. It should be noticed that the slot angle gives the direction of the cross jet flow with respect to the horizontal direction, as shown in Figure 4. Photo 1 shows the different components of the experimental set up [8].



Photo 1: *Components of the experimental set up*

2.3 Measuring Devices

2.3.1 Water Depths Measurement

The headwater depth H was measured using a Piezometric tube, fixed on a vertical scale of 0.50 mm accuracy and connected to the bottom of the testing flume by a rubber tube. The contracted depth y_c or the initial water depth y_1 was measured using a point gauge provided with a Vernier to obtain an accuracy up to 0.10 mm. The gauge reading at the channel bed line was taken firstly, then the gauge is adjusted on the water surface at the contracted section. The difference between the two readings gives the initial water depth, $y_1 (y_1 = y_c)$.

Because of the frequented wave surface at the end of the hydraulic jump, Piezometric tube was used to measure the tailwater depth y_2 . The Piezometric tube is fixed on a vertical scale, of 0.50 mm accuracy. The Piezometric tube is connected to the bottom of the flume by a rubber tube, as shown in photo 2.



Photo 2: *Piezometric tube*

2.3.2 Discharge Measurement

The discharge was measured by a sharp edged rectangular weir of width 17.0 cm and height 5.0 cm. The measuring weir is connected to the moulded channel which has 70 cm long and 25.0 cm

wide. The head above the weir was measured using a point gauge having Vernier of accuracy up to 0.10 mm. Photos 3 and 4 show the components of the measuring weir.

The weir was calibrated using the volumetric method [13]. As a result, the obtained discharge equation may be expressed as;

$$Q = 0.27485 * h^{1.578} \quad (1)$$

Where: Q =discharge, (lit/sec) and h =head over weir.

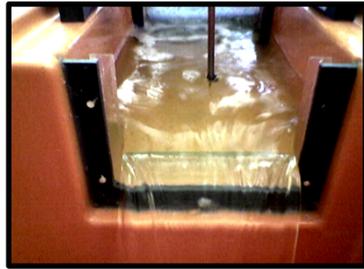


Photo 3



Photo 4

Photos 3 and 4: Components of the measuring weir

3- The Experimental Procedure

3.1 Case of Perfect Free Jump

(1) The pump is turned on and the control valve (10) is opened to a certain limit to obtain constant values of the discharge passing over the weir spillway, where $Q_w = 0.5, 1.0, 1.50, 2.0$ and 2.50 (lit/sec), corresponding to a headwater depth $H = 45.70, 46.98, 48.05, 48.98$ and 49.85 cm, respectively.

(2) For each discharge, the contracted section is singled and the contracted depth y_c and it's distance from the weir toe x_c were measured.

(3) Using the tailgate, the position of the hydraulic jump is adjusted so that the front of jump, immediately being at the contracted section to obtain a perfect jump. In this case, the initial depth of jump y_1 equals the contracted depth y_c .

(4) The tailwater depth y_2 was measured using the Piezometric tube and the length of the jump L_j was measured using a horizontal scale.

3.2 Case of Drowned Jump

(1) Considering constant values of both slot width $b = 0.15$ cm and inclination angle $\theta = 15^\circ$, the slot location was fixed at distances $x_s = 5, 10, 15, 20, 25$ and 30 cm.

(2) Considering $x_s = 5.0$ cm, the pump is turned on and the control valve is adjusted to give the same headwater depth H used in case of free jump. In this case, the discharge passing over the weir spillway Q_w remains constant as considered in case of free jump (step 1).

(3) Due to the effect of the discharge issuing from the slot Q_s , the tailwater depth increases creating a drowned jump. After that, the tailwater depth y_2 and the length of the drowned jump L_D , were measured.

(4) Using the tailgate, the tailwater depth was gradually reduced until the jump front is being immediately at the contracted section. Here the initial depth y_1 (or the contracted depth y_c) is still at the same value found in case of free jump, since Q_w is not changed.

(5) The tailwater depth y_2 and the length of forced jump L_J are then measured.

(6) The head on the rectangular weir (12) is measured and the total discharge Q_T is then estimated using Equation. (1). The discharge issuing from the slot Q_s is then found,

$$Q_s = Q_T - Q_w.$$

(7) Steps from (2) to (6) are repeated for other values of the headwater depth H .

(8) Considering another values of the slot distance x_s steps from (1) to (7) are repeated.

(9) Fixing the slot location at distance $x_s = 15.0$ cm and considering slot width $b = 0.15$ cm, the inclination angle of the slot θ is taken equal to $15^\circ, 30^\circ, 45^\circ, 60^\circ, 75^\circ, 90^\circ$.

For each of the above values of θ , the procedure is repeated from step (2) to (7).

(10) Fixing the slot location at distance $x_s = 15.0$ cm, and considering the inclination angle of slot $\theta = 45^\circ$, the width of slot b is varied as $b = 0.15, 0.20, 0.25, 0.30$ cm.

(11) Considering each of the above values of slot width b steps from (2) to (7) are repeated.

It should be noticed that, a sufficient time was allowed to satisfy a steady state condition of flow before recording the measured values.

3.3 Range of Froude Number

In the present study, Froude number ranges from 8.74 to 13.45. In this range, the jump is well established, the roller and jump action is fully developed to cause appreciable energy loss. However, the water surface downstream of the jump is rough and wavy.

3.4 Range of Discharge

The discharge passing over weir ranges from 500 cm³/sec to 2500 cm³/sec, while the discharge passing through slot ranges from 72.50 to 412.40 cm³/sec. The total discharge ranges from 500 to 2912.40 cm³/sec. The relative slot discharge ranges from 0.039 to 0.439 cm³/sec.

3.5 Range of Experiments

In this work, seventeen experiments were carried out in the laboratory. For each experiment, five runs corresponding to five values of Froude number were carried out, then the total number of experiments equals, $5 \times 17 = 85$ runs.

3.6 Main Parameters Involved in The Current Problem

Referring to Figure 4, the parameters affect on the hydraulic jump characteristics may be grouped as follows;

3.6.1 Boundary Parameters

- (i) The height of the inclined surface p .
- (ii) The width of the tailwater channel B .
- (iii) The distance of slot position x_s .
- (iv) The width of slot b .
- (v) The inclination angle of slot θ .

3.6.2 Flow Parameters

- (i) The headwater depth H .
- (ii) The head over the weir crest H_w .
- (iii) The conjugate depths, y_1 and y_2 .
- (iv) The mean velocity of flow at the contracted section v_1 .
- (v) The discharge passing over the weir Q_w .
- (vi) The discharge passing through the slot Q_s .

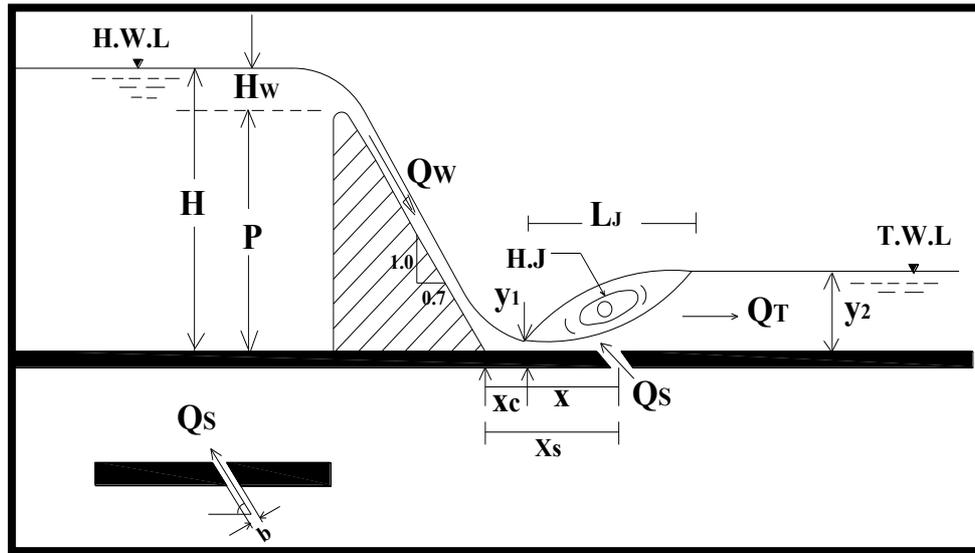


Figure 4: Details of dissipater model

All over the experimental work, the height of the inclined surface p was kept constant at 43.50 cm. For each experiment, the headwater depth H was changed five times since $H = 45.70, 46.98, 48.05, 48.98$ and 49.85 cm. Since $H = p + H_w$, therefore the head over the weir crest H_w equals 2.2, 3.48, 4.55, 5.48 and 6.35 cm, respectively. The width of tailwater channel also was kept constant in the experiments. It should be noticed that the mean velocity v_1 , includes the effect of both H and Q_w .

4- Analysis of Results

The main parameters involved in the current problem are; the distance of slot location x_s the angle of slot with the horizontal θ the width of slot b and Froude number of the incoming flow (supercritical) F_1 . Different values of x_s , θ , b and F_1 were experimentally investigated as indicated in Table 1.

Table 1: *The tested values of the considered parameters*

Parameter	Fixed parameters			Variable parameters		
	x_s , cm	θ°	b , cm	x_s , cm	θ°	b , cm
Distance of slot location, x_s , cm	-	15	0.15	5	-	-
				10		
				15		
				20		
				25		
				30		
Inclination angle of slot, θ°	15	-	0.15	-	15	-
					30	
					45	
					60	
					75	
					90	
Width of slot, b cm	15	45	-	-	0.15	
					0.20	
					0.25	
					0.30	

The case of free jump, in which the cross jet is not existed, was also investigated to be a reference data used to indicate the effect of the above parameters on the hydraulic jump characteristics and to compare the results obtained from the case of forced jump formed due to the cross jet flow. Comparison is based on the obtained results considering the case of the perfect hydraulic jump, either free or forced, and the case of drowned jump as shown in Figure 5. Referring to Figure 5, cases of perfect free, perfect forced and drowned jump have the same conditions of flow upstream the contracted section, wherever $x_s > x_c$; the head water depth H , the head over the weir crest H_w and discharge passing over the weir Q_w .

Downstream the contracted section, the above cases have different conditions; the total discharge Q_T and the tailwater depth y_2 . In case of free jump, where $Q_S = 0$, the total discharge Q_T equals to the weir discharge Q_w or $Q_T = Q_w$ while $Q_T = Q_w + Q_S$ in case of forced perfect or drowned jumps, where Q_S is the discharge passing through the slot.

Analysis of results in this paper was only focused on effect of the considered parameters on the drowned jump.

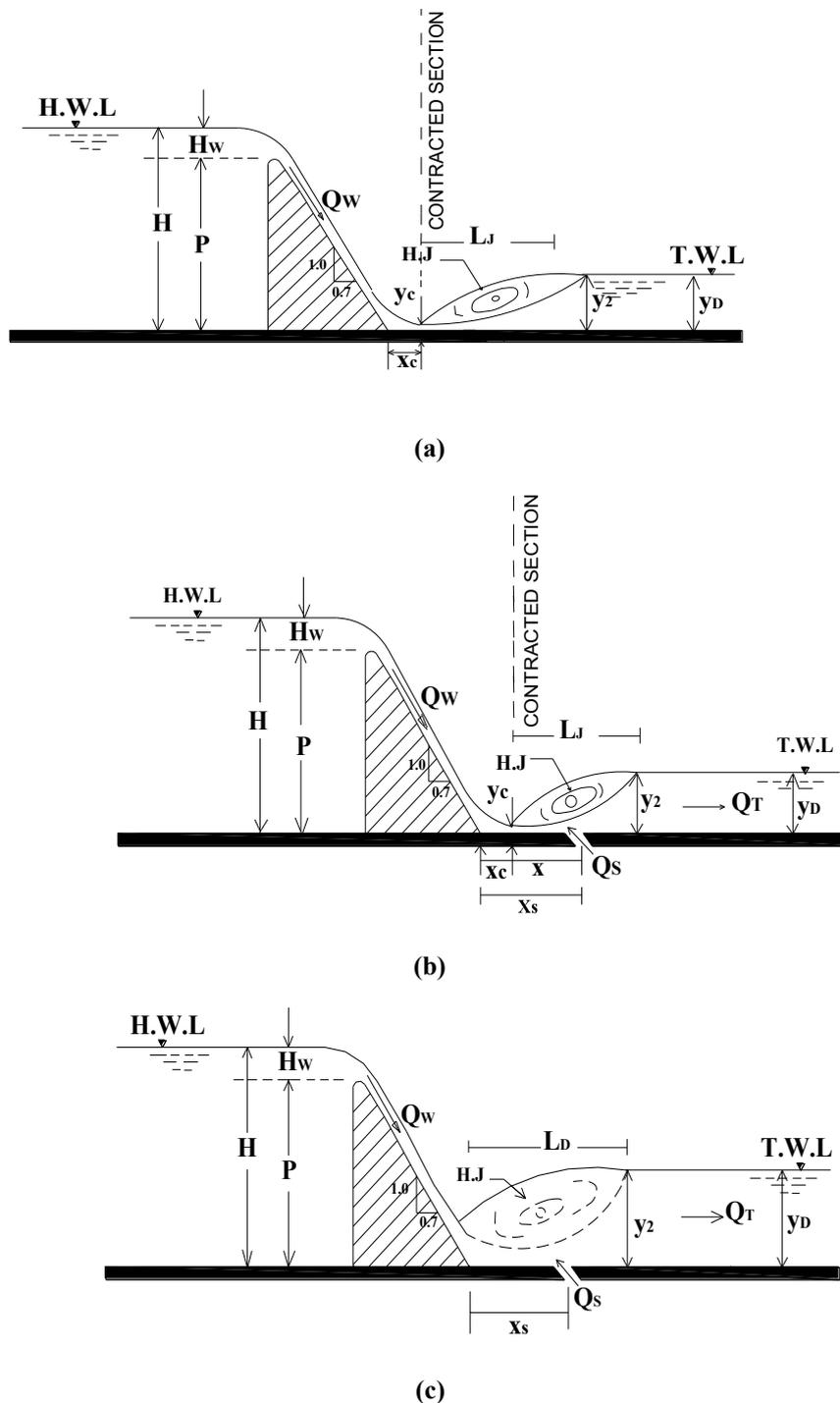


Figure 5: Definition sketch for the formed hydraulic jump; (a) Free perfect jump, (b) Forced perfect jump and (c) Drowned jump

4.1 The Effect of Counterflow on The Drowned Jump

The effect of the cross jet on the drowned jump is examined using the same tested values of slot position x_s , inclination angle θ and slot width b given in Table 1. As illustrated in section 3, after

establishing a free perfect jump, once the cross jet is allowed to act, the tailwater depth y_2 increases and the jump becomes drowned.

Both the length of the drowned jump L_D and the submergence ratio S_D are affected by the variation of slot position x_s , inclination angle of slot θ and slot width b . Here, the submergence ratio S_D is defined as the ratio between the tailwater depth for a free jump y_2 to the same depth creating drowned jump y_D or;

$$S_D = y_2/y_D \quad (2)$$

To indicate the effect of cross jet on the drowned jump, the length of drowned jump L_D is compared to the length L_f which expresses the sum of contraction section distance x_c and length of free perfect jump L_J or; $L_f = x_c + L_J$

The resulting relative difference Δ_L between the length L_f and L_D is expressed as a percentage value, where;

$$\Delta L\% = \frac{L_f - L_D}{L_f} \times 100 \quad (3)$$

The effect of variation of slot location x_s , angle θ and width b on both the submergence ratio S_D and the percentage reduction $\Delta L\%$ of length L_f is discussed below.

4.1.1 Effect of Slot Location x_s

To show the effect of slot location, the distance x_s is varied as $x_s = 5, 10, 15, 20, 25$ and 30 cm while the inclination angle of slot θ and the width of slot b are kept constant at 15° and 0.15 cm, respectively. Table 2 indicates both measured and calculated data, describing the effect of slot position x_s on the characteristics of the drowned jump, while Figure 6 shows the relation between the percentage difference $\Delta L\%$ and Froude number F_1 considering the tested values of slot location x_s . Both Table 2 and Figure 6 show that the reduction in free jump length $\Delta L\%$ increases as F_1 increases for all values of x_s . On the other hand, the maximum values of $\Delta L\%$ are obtained when $x_s = 15$ cm, while the minimum value occur when $x_s = 5$ cm. The maximum value of $\Delta L\%$ is 34.62% occurs when $x_s = 15$ cm and $F_1 = 13.45$, while the minimum value equals 18.75% when $x_s = 5$ cm and $F_1 = 9.22$.

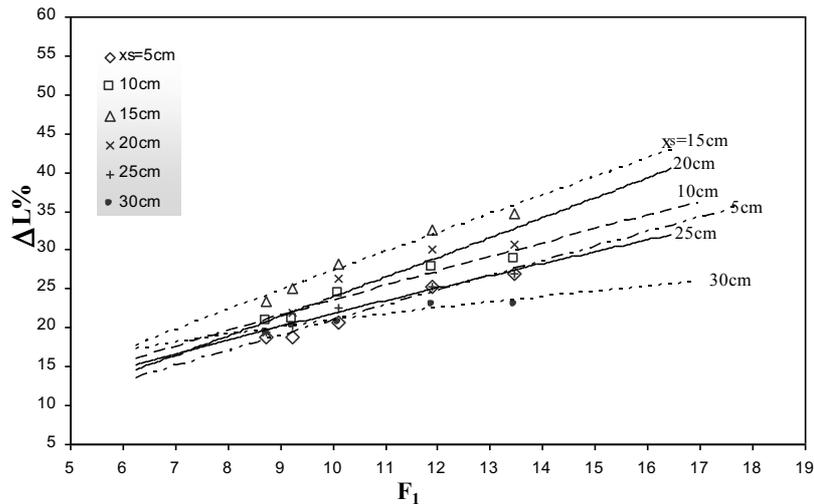


Figure 6: Relation between percentage difference $\Delta L\%$ and Froude number F_1 due to variation of slot location x_s

It is found also from Table 2, that the submergence ratio S_D increase as F_1 increases and has the maximum values when $x_s = 15$ cm. Generally, the submergence ratio S_D ranges from 1.01 to 1.74.

4.1.2 Effect of the Inclination Angle of Slot θ

Considering the effect of the variation of angle of inclination θ on the drowned jump, the slot location x_s and slot width b are kept constant at 15 cm and 0.15 cm, respectively. In the same time the inclination angle θ was varied as $\theta=15, 30, 45, 60, 75, 90^\circ$. Table 3 and Figure 7 show the resulting effect, on the drowned jump due to variation of angle θ . It is found that, the percentage difference $\Delta L\%$ increases as Froude number F_1 increases. The maximum values of $\Delta L\%$ equals 40.38% is obtained when $\theta = 45^\circ$ and $F_1 = 13.45$.

As for the submergence ratio S_D , it has the same trend of $\Delta L\%$, since it increases when F_1 increases, and has its maximum value 1.88 when $F_1 = 13.45$ and $\theta = 45^\circ$.

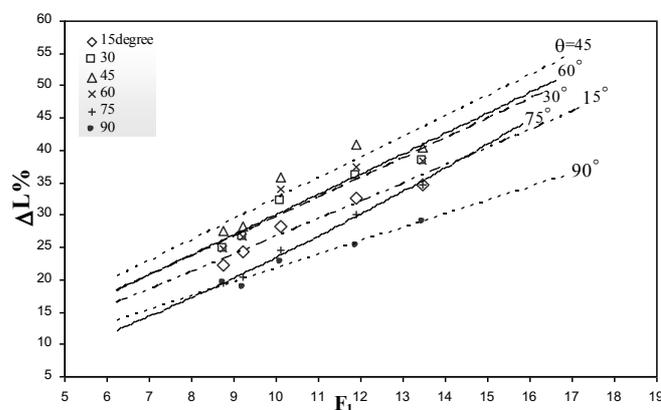


Figure 7: Relation Between $\Delta L\%$ and F_1 Due to Variation of Inclination Angle of Slot θ

4.1.3 Effect of Slot Width b

In this case, the slot location x_s and slot inclination angle θ are kept constant at 15 cm and 15° , respectively, while the slot width b is varied as $b = 0.15, 0.20, 0.25,$ and 0.30 cm. Table 4 and Figure 8 show the data describing the resulting effect, on the percentage difference $\Delta L\%$ and the submergence ratio S_D due to variation of slot width b .

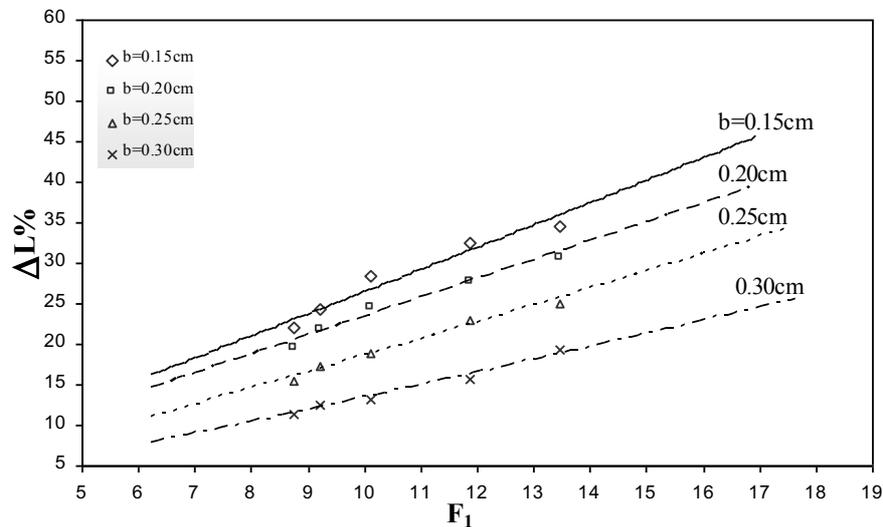


Figure 8: Relation Between $\Delta L\%$ and Froude Number F_1 Due to Variation of Slot Width b

It is found from Table 4 and Figure 8 that both $\Delta L\%$ and S_D increase as F_1 increases and the maximum value of $\Delta L\%$ and S_D are obtained when $b = 0.15$ cm. This refers to that when the slot width increases, an overhead jump occurs and falling in the downstream of the slot which increases the length of drowned jump L_D . Therefore, the slot width $b = 0.15$ cm gives the minimum values of L_D compared to other slot widths.

5- Conclusions

A comprehensive experimental study had been conducted in the present work, to investigate some characteristics of the formed drowned hydraulic jump, when the counter flow used to dissipate the energy of the flow falling over an ogee weir spillway. The obtained results insure that control of the hydraulic jump formed downstream of an ogee weir spillway, is possible using the reversed cross jet dissipator (counterflow). The suggested cross jet dissipator can be used to convert the repelled hydraulic jump not only to a perfect jump but also to a drowned jump and hence reduces the length of solid floor to a large extent.

Graphical presentations were given for a drowned hydraulic jump, for a Froude number ranges from 8.74 to 13.45.

The characteristics of the drowned hydraulic jump are represented by three variables, namely; the length of drowned jump L_D , the reduction in free jump length $\Delta L\%$ and the submergence ratio S_D .

Based on the analysis and discussion of the experimental results, obtained in the present study, the following conclusions may be given as follows:

- (1) In general, the results show that both the percentage reduction in the jump length and the

submergence ratio increase as Froude number increases.

(2) The maximum value of $\Delta L\%$ is 34.62 % occurs when $x_s = 15$ cm and $F_1 = 13.45$, while the minimum value equals 18.75% when $x_s = 5$ cm and $F_1 = 9.22$.

Table 2: Effect of Slot Locations x_s on the Drowned Jump, When $b = 0.15$ cm and $\theta = 15^\circ$

x_s cm	F_1	x_c cm	L_J , cm	L_f cm	L_D cm	y_2 cm	y_D cm	$\Delta L\%$	$\Delta y\%$	$S_D = \frac{y_2}{y_D}$
5	13.45	3.00	23.00	26.00	19.00	5.00	8.10	26.92	62.00	1.62
	11.88	6.50	35.00	41.50	31.00	7.50	9.10	25.30	21.33	1.21
	10.10	8.00	45.00	53.00	42.00	9.60	10.30	20.75	7.29	1.07
	9.22	11.00	53.00	64.00	52.00	11.25	11.53	18.75	2.49	1.02
	8.74	13.50	61.00	74.50	60.50	12.75	12.85	18.79	0.78	1.01
10	13.45	3.00	23.00	26.00	18.50	5.00	8.50	28.85	70.00	1.70
	11.88	6.50	35.00	41.50	30.00	7.50	9.50	27.71	26.67	1.27
	10.10	8.00	45.00	53.00	40.00	9.60	10.70	24.53	11.46	1.11
	9.22	11.00	53.00	64.00	50.50	11.25	11.80	21.09	4.89	1.05
	8.74	13.50	61.00	74.50	59.00	12.75	13.20	20.81	3.53	1.04
15	13.45	3.00	23.00	26.00	17.00	5.00	8.70	34.62	74.00	1.74
	11.88	6.50	35.00	41.50	28.00	7.50	9.70	32.53	29.33	1.29
	10.10	8.00	45.00	53.00	38.00	9.60	10.90	28.30	13.54	1.14
	9.22	11.00	53.00	64.00	48.00	11.25	12.10	25.00	7.56	1.08
	8.74	13.50	61.00	74.50	57.00	12.75	13.40	23.49	5.10	1.05
20	13.45	3.00	23.00	26.00	18.00	5.00	8.60	30.77	72.00	1.72
	11.88	6.50	35.00	41.50	29.00	7.50	9.50	30.12	26.67	1.27
	10.10	8.00	45.00	53.00	39.00	9.60	10.70	26.42	11.46	1.11
	9.22	11.00	53.00	64.00	50.00	11.25	11.70	21.88	4.00	1.04
	8.74	13.50	61.00	74.50	60.00	12.75	13.10	19.46	2.75	1.03
25	13.45	3.00	23.00	26.00	19.00	5.00	8.30	26.92	66.00	1.66
	11.88	6.50	35.00	41.50	31.00	7.50	9.30	25.30	24.00	1.24
	10.10	8.00	45.00	53.00	41.00	9.60	10.40	22.64	8.33	1.08
	9.22	11.00	53.00	64.00	51.00	11.25	11.50	20.31	2.22	1.02
	8.74	13.50	61.00	74.50	60.00	12.75	12.90	19.46	1.18	1.01
30	13.45	3.00	23.00	26.00	20.00	5.00	8.20	23.08	64.00	1.64
	11.88	6.50	35.00	41.50	32.00	7.50	9.10	22.89	21.33	1.21
	10.10	8.00	45.00	53.00	42.00	9.60	10.30	20.75	7.29	1.07



	9.22	11.00	53.00	64.00	51.00	11.25	11.40	20.31	1.33	1.01
	8.74	13.50	61.00	74.50	60.00	12.75	12.90	19.46	1.18	1.01

- (3) Both $\Delta L\%$ and S_D increase as F_1 increases and the maximum values of $\Delta L\%$ and S_D are obtained when $b = 0.15$ cm.
- (4) The submergence ratio S_D has the same trend of $\Delta L\%$, since it increases when F_1 increases, and has its maximum value 1.88 when $F_1 = 13.45$ and $\theta = 45^\circ$.
- (5) The cross jet flow can shorten the length of the solid floor, by creating a drowned jump instead of a free perfect jump, to about 65% with a submergence ratio of about 1.74.
- (6) The value of the counterflow width b (the slot), should not exceed the value of $b = 0.15$ cm, otherwise an overhead jump will occurs and falling in the downstream of the slot which increases the length of drowned jump L_D . Therefore, the slot width $b = 0.15$ cm gives the minimum values of L_D in comparison to other slot widths.

Table 3: Measured and Calculated Data for the Drowned Jump Due to Variation of Inclination Angle θ
 When $b = 0.15 \text{ cm}$ and $x_s = 15 \text{ cm}$

θ°	F_1	x_c cm	L_J, cm	L_f cm	L_D cm	y_2 cm	y_D cm	$\Delta L\%$	$\Delta y\%$	$S_D = \frac{y_2}{y_D}$
15°	13.45	3.00	23.00	26.00	17.00	5.00	8.70	34.62	74.00	1.74
	11.88	6.50	35.00	41.50	28.00	7.50	9.70	32.53	29.33	1.29
	10.10	8.00	45.00	53.00	38.00	9.60	10.90	28.30	13.54	1.14
	9.22	11.00	53.00	64.00	48.50	11.25	12.10	24.22	7.56	1.08
	8.74	13.50	61.00	74.50	58.00	12.75	13.40	22.15	5.10	1.05
30°	13.45	3.00	23.00	26.00	16.00	5.00	8.90	38.46	78.00	1.78
	11.88	6.50	35.00	41.50	26.50	7.50	9.90	36.14	32.00	1.32
	10.10	8.00	45.00	53.00	36.00	9.60	11.20	32.08	16.67	1.17
	9.22	11.00	53.00	64.00	47.00	11.25	12.30	26.56	9.33	1.09
	8.74	13.50	61.00	74.50	56.00	12.75	13.50	24.83	5.88	1.06
45°	13.45	3.00	23.00	26.00	15.50	5.00	9.40	40.38	88.00	1.88
	11.88	6.50	35.00	41.50	24.50	7.50	10.60	40.96	41.33	1.41
	10.10	8.00	45.00	53.00	34.00	9.60	11.60	35.85	20.83	1.21
	9.22	11.00	53.00	64.00	46.00	11.25	12.70	28.13	12.89	1.13
	8.74	13.50	61.00	74.50	54.00	12.75	13.90	27.52	9.02	1.09
60°	13.45	3.00	23.00	26.00	16.00	5.00	9.10	38.46	82.00	1.82
	11.88	6.50	35.00	41.50	26.00	7.50	10.40	37.35	38.67	1.39
	10.10	8.00	45.00	53.00	35.00	9.60	11.30	33.96	17.71	1.18
	9.22	11.00	53.00	64.00	47.00	11.25	12.50	26.56	11.11	1.11
	8.74	13.50	61.00	74.50	56.00	12.75	13.60	24.83	6.67	1.07
75°	13.45	3.00	23.00	26.00	17.00	5.00	8.90	34.62	78.00	1.78
	11.88	6.50	35.00	41.50	29.00	7.50	10.30	30.12	37.33	1.37
	10.10	8.00	45.00	53.00	40.00	9.60	11.20	24.53	16.67	1.17
	9.22	11.00	53.00	64.00	51.00	11.25	12.20	20.31	8.44	1.08
	8.74	13.50	61.00	74.50	60.00	12.75	13.30	19.46	4.31	1.04
90°	13.45	3.00	23.00	26.00	18.50	5.00	8.40	28.85	68.00	1.68
	11.88	6.50	35.00	41.50	31.00	7.50	9.70	25.30	29.33	1.29
	10.10	8.00	45.00	53.00	41.00	9.60	10.70	22.64	11.46	1.11
	9.22	11.00	53.00	64.00	52.00	11.25	11.70	18.75	4.00	1.04
	8.74	13.50	61.00	74.50	60.00	12.75	12.90	19.46	1.18	1.01

Table 4: Data Describing the Effect of Variation of Slot Width b on the Drowned Jump When $x_s = 15$ cm, $\theta = 15^\circ$

b cm	F_1	x_c cm	L_j, cm	L_f cm	L_D cm	y_2 cm	y_D cm	$\Delta L\%$	$\Delta y\%$	$S_D = \frac{y_2}{y_D}$
0.15	13.45	3.00	23.00	26.00	17.00	5.00	8.70	34.62	74.00	1.74
	11.88	6.50	35.00	41.50	28.00	7.50	9.70	32.53	29.33	1.29
	10.10	8.00	45.00	53.00	38.00	9.60	10.90	28.30	13.54	1.14
	9.22	11.00	53.00	64.00	48.50	11.25	12.10	24.22	7.56	1.08
	8.74	13.50	61.00	74.50	58.00	12.75	13.40	22.15	5.10	1.05
0.20	13.45	3.00	23.00	26.00	18.00	5.00	8.60	30.77	72.00	1.72
	11.88	6.50	35.00	41.50	30.00	7.50	9.50	27.71	26.67	1.27
	10.10	8.00	45.00	53.00	40.00	9.60	10.70	24.53	11.46	1.11
	9.22	11.00	53.00	64.00	50.00	11.25	12.00	21.88	6.67	1.07
	8.74	13.50	61.00	74.50	60.00	12.75	13.30	19.46	4.31	1.04
0.25	13.45	3.00	23.00	26.00	19.50	5.00	8.40	25.00	68.00	1.68
	11.88	6.50	35.00	41.50	32.00	7.50	9.30	22.89	24.00	1.24
	10.10	8.00	45.00	53.00	43.00	9.60	10.40	18.87	8.33	1.08
	9.22	11.00	53.00	64.00	53.00	11.25	11.90	17.19	5.78	1.06
	8.74	13.50	61.00	74.50	63.00	12.75	13.10	15.44	2.75	1.03
0.30	13.45	3.00	23.00	26.00	21.00	5.00	8.20	19.23	64.00	1.64
	11.88	6.50	35.00	41.50	35.00	7.50	9.50	15.66	26.67	1.27
	10.10	8.00	45.00	53.00	46.00	9.60	10.20	13.21	6.25	1.06
	9.22	11.00	53.00	64.00	56.00	11.25	11.90	12.50	5.78	1.06
	8.74	13.50	61.00	74.50	66.00	12.75	12.90	11.41	1.18	1.01

References

- [1]. Chow VT. Open channel hydraulics. London: McGraw-Hill Book Co. Ltd.; 1959
- [2]. Peterka AJ. Hydraulic design of stilling basins and energy dissipaters engineering monograph No. 25, U.S. Bureau of Reclamation, Denver Colorado; 1958
- [3]. Rajaratnam N. Hydraulic jumps. In: Advances in hydrosience, vol. 4. Edtion Ven Te Chow; 1967. p. 254.W. S.
- [4]. McCorquodale JA. Hydraulic jumps and internal flows. In: Cheremisionoff NP, editor. Encyclopedia of fluid, vol. 2. Houston: Gulf Publishing; 1986. p. 120–73

- [5]. Hager WH. Energy dissipaters and hydraulic jump. 3rd ed. Dordrecht: Kluwer Academic; 1992
- [6]. Sholichin M, Akib S. Development of drop number performance for estimate hydraulic jump on vertical and sloped drop structure. Int J Eng Sci 2011;5(11):1678–87
- [7]. Chaudhry, M. H. (2008). Open-Channel Flow, Springer Science+Business Media, LLC, New York, NY.
- [8]. Alghwail, A. D. A., Stevović, S., Abourohiem, M. A. (2018). Dissipation of mechanical energy over spillway through counter flow, GRAĐEVINAR, 70 (5), pp.377-391, doi: <https://doi.org/10.14256/JCE.1691.2016>
- [9]. Stephenson, D. (1991), "Energy dissipation down stepped spillways", Water power & Dam construction 42 (9), pp. 27-30
- [10]. Chanson, H. (1994 a). "Comparison of energy dissipation between nappe and skimming flow regimes on stepped chutes", ASCE, J. Hydraulic Research, Vol.32, No.2, pp.213-218
- [11]. Chamani, M. R. and Rajaratnam N. (1999-b), "Onset of skimming flow on stepped spillways". J. Hydraulic Engineering.125 (9) PP. 969-971.
- [12]. Chamani, M. R. and Rajaratnam N. (1999-a), "Characteristics of skimming flow over stepped spillways" J. Hydraulic Engineering. 125 (4) PP. 361-368
- [13]. USBR hydraulic design of stilling basins and energy dissipators, Engineering Monograph, p 92. No.25.
- [14]. Alghwail, A. (2016) Alleviation the scouring problem downstream of dam spillways through a reversed cross-jet flow dissipator. In:ICSE 2016 (8th International Conference on Scour and Erosion), 12-15 September 2016, Oxford, UK.

The Influence of Ageing on Asphalt Mastic Properties Incorporating Calcium Carbonate

Foad Mohamed Elkut^{1*}, Meor Othman Hamzah²

¹ foad.elkut@yahoo.com, ² cemeor@yahoo.com

¹ Department of Civil Engineering, Elmergib University, Al-komes, Libya

² School of Civil Engineering, Universiti Sains Malaysia, 14300 Nibong Tebal, Penang, Malaysia

ABSTRACT

Keywords:

Asphalt mastic,
Viscosity,
Complex shear modulus,
Phase angle,
Ageing.

The ageing of asphalt mastic is one of the key factors determining the lifetime of an asphalt pavement. The challenge created by proper selection of asphalt materials motivate researchers to investigate material properties that can further enhanced performance of asphalt pavement. The study investigated the ageing properties of asphalt mastics incorporating calcium carbonate (CaCO_3) combined with Ordinary Portland Cement (OPC). The empirical tests, which include penetration and softening points, were conducted to ascertain the asphalt mastics consistency. The rheological properties of mastics in terms rotational viscosity and Superpave rutting parameter using 60/70 asphalt binder blended with 5, 10, 15, 20 % of CaCO_3 +OPC were calculated to evaluate asphalt mastics properties subjected to different aged conditions. Rotational viscometer (RV) was used to evaluate the properties of mastics. The dynamic shear rheometer (DSR) was used in temperature sweep test to measure the complex modulus (G^*) and phase angle (δ). ANOVA statistical analysis was used to analyze the results .The test results showed that all asphalt mastics exhibited higher viscosity compared to the base binder. The addition of various content of CaCO_3 +OPC increased the G^* and decreased the δ significantly, indicating an increase in binder elasticity and stiffness, hence a better resistance to deformation.

1- Introduction

The development of modern pavement technology is needed to accelerate significant improvement of pavement quality of highways. Pavement surface distress such as cracks is prevalent on pavements due to the action of repeated traffic loading and cyclic environmental conditions. Use of new materials may help to mitigate the problem by improving the properties of asphalt mixtures. This study focus on the effects of asphalt binder modified with CaCO_3 +OPC as a modifier/filler on the rheological properties of asphalt mastics. Asphalt mastics is a combination of asphalt binder and filler that influence the overall mechanical performance of asphalt mixtures as well as placement workability [1].The properties of mineral filler have a significant effect on asphalt properties in terms of permanent deformation, fatigue cracking and moisture susceptibility [2]. Many different types of fillers obtained by processing natural or manufactures or recycled materials can be used for asphalt pavements such as Portland cement, hydrated lime, and ground slag [3]. Binder ageing is one of the principal factors causing the deterioration of asphalt mixtures [4]. There are two basic mechanisms involved in binder ageing, these include an irreversible process like chemical changes of the bitumen, consisting of oxidation of bitumen molecules, and loss of volatile components which subsequently has an impact on the rheological properties of the asphalt binders. Nowadays, most of ageing on the road is still regarded as thermally induced. The rate of thermal oxidation of asphalt binder in service depends to a large extent on pavement temperature [5]. It has been shown that different asphalt binders have very different increase of viscosity with ageing

time, and the temperature-dependence of ageing is strongly dependent on the asphalt binders [6]. Rutting and fatigue are the most common surface distress on Malaysian roads [7, 8]. The main objective of this paper is evaluate the rheological properties of asphalt mastics subjected to the appropriate ageing conditions.

2- Materials and Methods

A conventional binder grade 60/70 supplied by SHELL Company was used and its properties are shown in Table 1.

Table 1: *Properties of base binder*

Ageing condition	Property	values
Unaged	Penetration [1/10 mm]	63
	Softening Point [°C]	48
	Ductility @ 25 °C [cm]	115
	Relative Density @ 25°C	1.03
	G*/sinδ @ 64°C [Pa]	1621.40
Short term aged	G*/sinδ @ 64°C [Pa]	3584.20
Long term aged	G* sinδ @ 25°C [MPa]	4.51

CaCO₃ plus OPC was used as filler to prepare the mastics. The 60/70 asphalt binder was separately blended with 5, 10, 15, and 20 % of CaCO₃ + OPC by mass of binder as shown in Table 2.

Table 2: *Asphalt mastic preparation*

Content	Description	Designation
0%	0% CaCO ₃ +0% OPC	Control
5%	3.75 % CaCO ₃ +1.25 % OPC	CO5
10%	7.50 % CaCO ₃ +2.50 % OPC	CO10
15%	11.25 % CaCO ₃ +3.75 % OPC	CO15
20%	15 % CaCO ₃ +5 % OPC	CO20

Sample Preparation and Test Methods

The preparing the modified binders, about 800 g of the binder was heated to liquefy in a 2 liter capacity metal container. Upon reaching 145±5°C, the proportion of fillers by 5%, 10%, 15% and 20% of mass of binder were added to the base binder in a high shear mixer for 30 min at 3000 rpm throughout the mixing process. These binder were then artificially short-term aged (STA) according to ASTM D2872 (ASTM, 2012) procedures via the Rolling Thin Film Oven (RTFO) at 163°C for 85 minutes [9]. Subsequently, the Pressure Ageing vessel (PAV) was used to subject the binder to long-term ageing (LTA) according to ASTM D6521 proceedings of the Eastern Asia Society for Transportation Studies, Vol.9, 2013(ASTM, 2013) procedures at 100°C for 20 hours [10]. Conventional tests such as penetration, ring and ball were carried out according to (ASTM D5-97) [11] and (ASTM D36-95) [12] procedures respectively. The rotational viscosity using Brookfield viscometer (RV) was carried out over temperature ranging from 120 to 170°C, spindle No.27, and rotating speed 20rpm. Rutting characterization was tested on parallel plates with 1 mm thickness, 25 mm diameter, loading condition on strain control mode at 1.59 Hz and at temperature

ranging from 46°C to 82°C at 6°C increments using the Dynamic Shear Rheometer (DSR) subjected to different ageing conditions to determine the basic characteristics and rheological properties of the asphalt mastic.

3- Results and Discussion

Penetration and softening point

Figure 1 presents the test results of bitumen 60/70 blended with several CO contents. The results show that compared with control sample, all of asphalt mastics subjected to ageing exhibit lower penetration indicating stiffer asphalt mastics. The penetration changes due to ageing is greatly decreased 25% and 55% for STA and LTA samples, respectively. Incorporation of CO improves the resistance to flow deformation of the asphalt binder. On the other hand, the results show that compared with control sample, all of asphalt mastics exhibit higher softening point subjected to ageing, indicating that incorporation of fillers reduce the temperature susceptibility, increase the plasticity of the asphalt mastics. The softening point changes due to ageing indicates the thermal oxidative reaction of asphalt is reduced by the addition of CO. Sample 15% CO showed the highest softening point compared to other percentage of fillers in asphalt mastics. The reason for this improvement might be due to the bonding strength to restrict the flow of asphalt mastics and made it stiffer. Qin et al., (2014) indicated that the age hardening of asphalt binder is closely correlated and attributed to the compositional changes within the asphalt binder [13]. Stiffer asphalt mastics will be more resistant to rutting. The results show the potential of CO mastics to be an alternative material to be selected by engineers for improving asphalt mixtures properties.

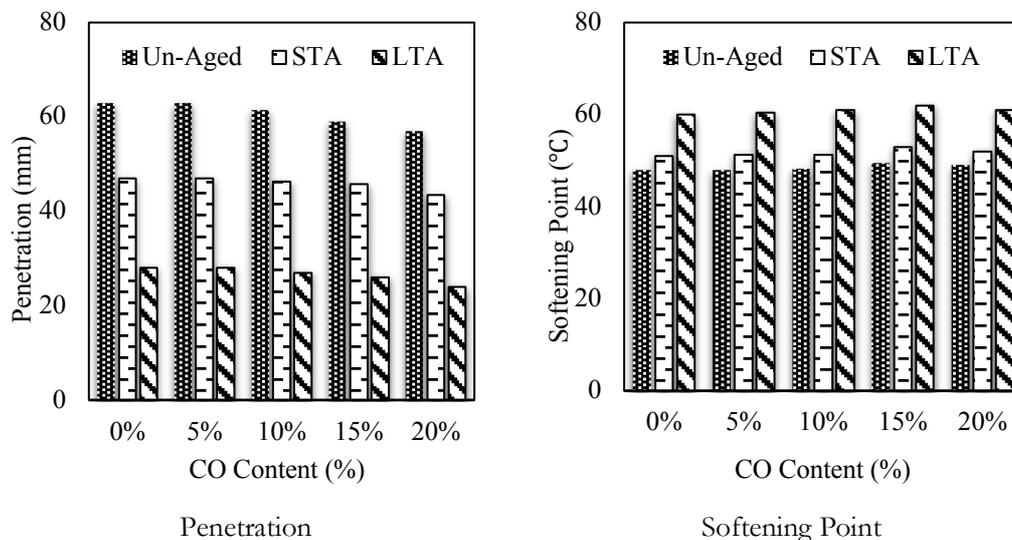


Figure 1: Penetration and Softening Point of Asphalt Mastics Subjected to Ageing

Viscosity

Binder viscosity is important during asphalt production and construction. The relationships between filler contents and viscosity subjected to different temperatures and ageing are shown in Figure 2. At lower temperatures, the viscosities of asphalt mastics containing various CO contents are high. As the temperature increases, the viscosity reduces. Similar trend is shown for asphalt mastics subjected to ageing but with a higher viscosity because ageing causes the binders to stiffen. The increase in filler contents also increases the viscosity compared to unaged base binder.

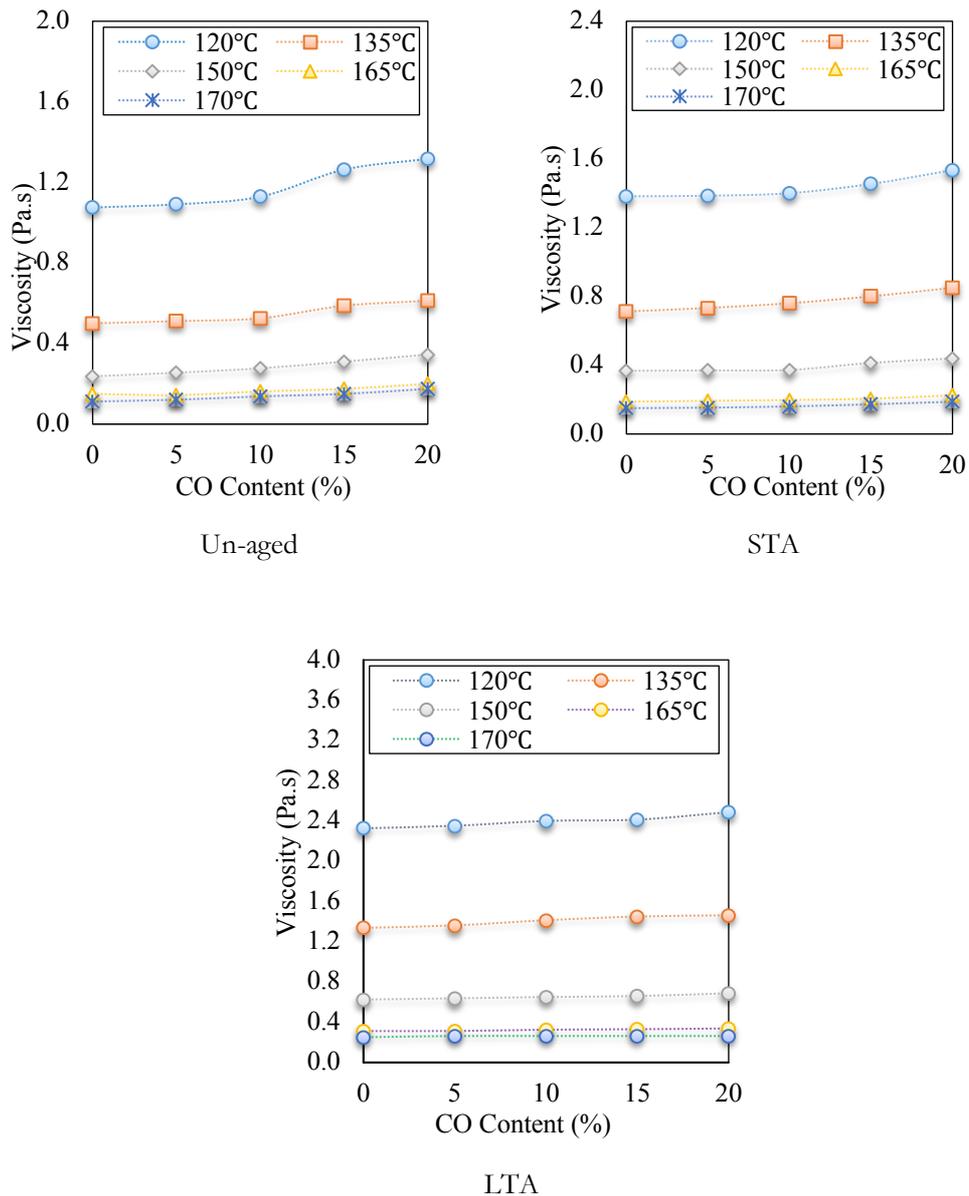


Figure 2: Rotational viscosity values of CO subjected to different ageing condition

Table 3 show the statistical analysis results using two-way analysis of variance (ANOVA) of the viscosity of CO at various ageing conditions at 95% confidence level. The ANOVA was carried out on the experimental data based on the un-aged and STA samples of asphalt mastics containing various filler contents at 120°C, 135°C and 165°C. The result indicates that temperature is the most significant factor that affects the viscosity followed by the effects of ageing and the type of additives. Moreover, the interaction effects between these parameters are found to be significant on the viscosity of asphalt mastics.

Table 3: ANOVA Analysis of Viscosity CO

Source	SS	Df	MS	F	P-Value	Significant
Corrected Model	89.5a	74	1.2	2044.9	<0.001	Yes
Intercept	107.8	1	107.9	182423.8	<0.001	Yes
TT b	69.7	4	17.4	29468.6	<0.001	Yes

AC c	11.8	2	5.9	9973.5	<0.001	Yes
CO d	0.276	4	0.069	116.7	<0.001	Yes
T T* AC	7.5	8	0.940	1589.2	<0.001	Yes
TT * CO	0.094	16	0.006	9.9	<0.001	Yes
AC * CO	0.045	8	0.006	9.5	<0.001	Yes
T * AC * CO	0.053	32	0.002	2.8	<0.001	Yes
Error	0.089	150				
Total	197.5	225				
Corrected Total	89.6	224				

a) R Squared = 0.99 (Adjusted R Squared = 0.99), b) Test Temperature, c) Aging Conditions and d) Asphalt Mastics Containing CO.

The change in viscosity is employed to evaluate the anti-ageing performance of asphalt mastics. The viscosity ageing increment (VAI) was calculated based on Equation (1) and the result is presented in Table 4.3.

$$VAI \% = \frac{\eta_{PAV} - \eta_{RTFOT}}{\eta_{RTFOT}} * 100 \quad (1)$$

Where:

- VAI = Viscosity ageing increment
- η_{PAV} = Viscosity of asphalt mastics after LTA conditions.
- η_{RTFOT} = Viscosity of asphalt mastics after STA conditions.

Table 4: Viscosity Ageing Increment of CO

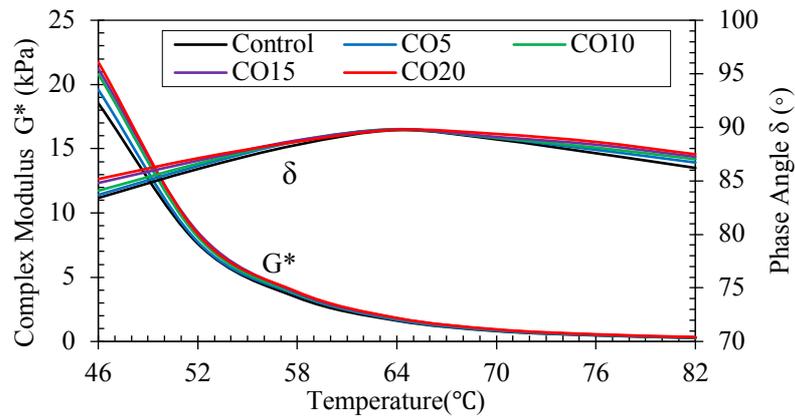
Content (%)	Viscosity (MPa.s) at 135°C		
	STA	LTA	VAI
0	723	1349	86.58
5	733	1362	85.94
10	760	1413	85.92
15	800	1450	81.25
20	850	1462	72.00

The analysed data show that the VAI of asphalt mastics decrease significantly as the filler content increases, which indicated better resistance to ageing, hence less susceptibility to LTA.

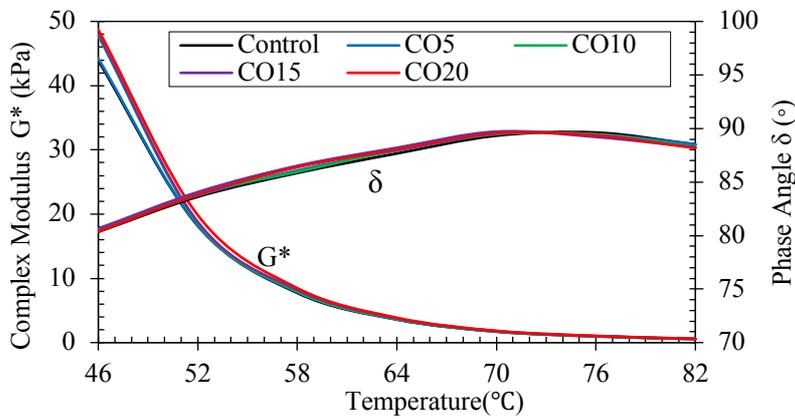
Superpave rutting parameters

The effects of asphalt mastics containing various CO contents on G^* and δ at various temperatures for unaged and aged conditions are illustrated in the master curves shown in Figures 3. This curve presents the variation of G^* and δ versus temperature at 10 rad/sec frequency. To produce a

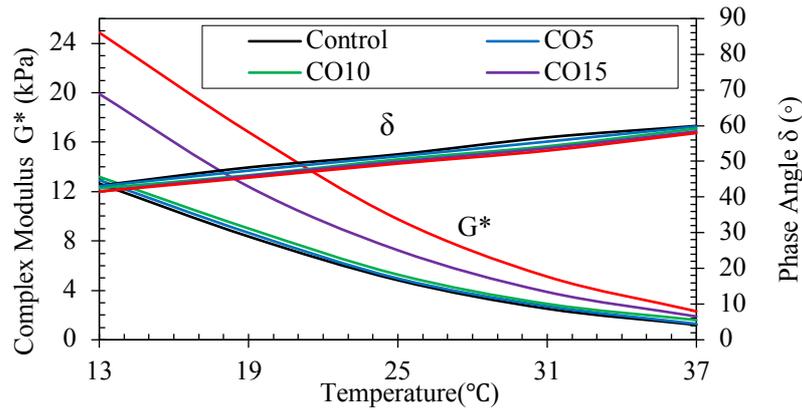
mixture with a higher resistance to rutting, the asphalt mastic should have a higher G^* and lower δ value. As the temperature increases, it has resulted in lower G^* and continuously increased the δ . theoretically, a higher G^* reflects a superior resistance to deformation. Meanwhile, the increase in δ reflects a reduction in elasticity. Higher CO contents increases G^* and decreases δ at high temperatures. The result shows the effects of CO contents on the rheological properties of asphalt mastics samples. Similar trends were observed, the G^* increases and proportionally decreases the δ of samples regardless of STA condition and CO asphalt mastics. It could be due to the structure formed in the asphalt mastics that reduce the temperature sensitivity and increases the elasticity of mastics based on the reduction of δ at elevated temperatures. A linear relationship between δ and temperature for LTA asphalt mastics. An exponential relationship for G^* , with R^2 equal to 0.99 for LTA conditions are presented in the same figure. The addition of various CO contents combined with LTA ageing causes the asphalt mastics to be less susceptible to temperature.



Un-aged



STA



LTA

Figure 3: G^* and δ of Asphalt Mastics subjected to different ageing condition

The two-way ANOVA was performed to analyse the effects of temperature on the increase of G^* and reduction of δ . Tables 4 show the results of the statistical analysis. The results indicated that a higher test temperature has a significant effect on G^* based on the p -value less than 0.05. The result for un-aged samples indicates that the interaction effects between these parameters are found the most significant effects on the complex modulus. The type of asphalt mastics has most significant effects followed by the test temperature for aged samples of asphalt mastics. The addition of various content in asphalt mastics increased the G^* and decreased the δ significantly, indicating an increase in binder elasticity and stiffness, hence a better resistance to deformation.

Table 5: ANOVA Analysis of the complex modulus (G^*)

Sample	Source	Mean Square	Df	F	p -value	Sig.
Un-Aged	TT ^a	6.60E+07	6	1503.26	<0.001	Yes
	CO ^b	5.08E+05	4	11.56	<0.001	Yes
	TT * CO	1.83E+08	24	4158.73	<0.001	Yes
	Error	4.4E+04	70			
	Total		105			
	Corrected Total		104			
			R Squared = 0.999 (Adjusted R Squared = 0.999)			
STA	TT ^a	1.55E+09	6	166.51	<0.001	Yes
	CO ^b	3.52E+08	4	1167.77	<0.001	Yes
	TT * CO	2.7E+06	24	9.012	<0.001	Yes
	Error	9.54E+08	70	3166.49		
	Total		105			
	Corrected Total		104			
			R Squared = 0.999 (Adjusted R Squared = 0.999)			
Sample	Source	Mean Square	Df	F	p -value	Sig.

LTA	TT ^a	11.4E+13	6	99.94	<0.001	Yes
	CO ^b	5.20E+14	4	456.35	<0.001	Yes
	TT * CO	1.52E+13	24	13.37	<0.001	Yes
	Error	1.14E+12	70			
	Total		105			
	Corrected Total		104			
			R Squared = 0.980 (Adjusted R Squared = 0.970)			

(a) TT= Test Temperature, (b) Asphalt Mastics Containing CO

4- Conclusions

The study indicated physical hardness after simulating the base binder to ageing using RTFOT and PAV, reduces binder penetration and increases softening points. This indicates that ageing depends on various conditions relate to field ageing. The viscosity of asphalt mastic at 135°C was lower than 3 Pa.s which fulfils SHRP criteria for provide potential satisfactory result during industrial processing which involved handling, lay down and compaction. The adding 15% CO mastic increased the G* but reduced the phase angle which indicates improve stiffness of mastic asphalt. At high temperature, a good visco-elastic performance was observed for CO asphalt mastics. Hence, the asphalt mastics exhibit better resistance to rutting compared to base binder. This observation was also evident from the statistical analysis. Thus, it can be concluded that there are more advantages to use CO mastic in terms of rheological properties.

5- Acknowledgment

This study was carried out in the of Highway and Transportation laboratory, Faculty of Civil Engineering, Universiti Sincce Malaysia (USM), and all involved in the study are acknowledged.

References

- [1]. W. G. Buttlar, D. Bozkurt, G. G. Al-Khateeb and A. S. Waldhoff. "Understanding asphalt mastic behavior through micromechanics," *Transportation Research Record* 1681, no. 1 (1999): 157-169.
- [2]. A. Zulkati, W. Y. Diew and D. S. Delai. "Effects of fillers on properties of asphalt-concrete mixture. *Journal of transportation engineering*, 138, no. 7 (2012): 902-910.
- [3]. M. C. Liao, G. Airey and J. S. Chen. "Mechanical Properties of Filler-Asphalt Mastics," *International Journal of Pavement Research and Technology*. Sep 2013, Vol. 6 Issue 5, p576-581.
- [4]. J. Read and D. Whiteoak. "The shell bitumen handbook," Fifth Edition, Thomas Telford, London, 2003.
- [5]. Branthaver, Jan Franklin, J. C. Petersen, R. E. Robertson, J. J. Duvall, S. S. Kim, P. M. Harnsberger, T. Mill, E. K. Ensley, F. A. Barbour, and J. F. Scharbron. "Binder characterization and evaluation," Volume 2: Chemistry. No. *SHRP-A-368*. 1993.
- [6]. Lu, Xiaohu, Johann Talon, and Per Redelius. "Ageing of Bituminous Binders–Laboratory Tests and Field Data," In 4th *Eurasphalt Eurobitume Congress*, pp. 1-12. 2008.
- [7]. O. Hamzah and B. Golchin. "A laboratory investigation on the rheological properties of asphalt binder containing rediset," *Journal of the Eastern Asia Society for Transportation Studies* 10.0 1537-1550. (2013)
- [8]. M. O.Hamzah, B. Golchin, A. Jamshidi and E. Chailleux. "Evaluation of Rediset for use in warm-mix asphalts a review of the literatures," *International Journal of Pavement Engineering*. Ahead-of-print (2014) 1-23.



- [9]. ASTM D2872. “Standard test method for effect of heat and air on a moving film of asphalt (rolling thin-film oven test),” USA: *Annual Book of ASTM Standards* (2012).
- [10]. ASTM D6521. “Standard practice for accelerated aging of asphalt binder using a pressurized aging vessel (PAV),” USA: *Annual Book of ASTM Standards* (2013).
- [11]. ASTM D5. “Standard test method for penetration of bituminous materials,” West Conshohocken, USA, *American Society for Testing and Materials*, (2006).
- [12]. ASTM D36. “Standard test method for ring-and-ball apparatus,” West Conshohocken, USA, *American Society for Testing and Materials*, (2009).
- [13]. A Qin, Q., Schabron, J. F., Boysen, R. B. and Farrar, M. J. “Field Ageing Effect on Chemistry and Rheology of Asphalt Binders and Rheological Predictions for Field Aging,” *Fuel*, 121, 86-94. (2014).



Track No. 3: Chemical & Oil Engineering

Simulation of Butanol Production through Hydrogenation of Butanal: Effects of Different Reactor Schemes and Operating Conditions

Mohammed Faraj Saeid^{1*}, Sim Yee Chin²

¹faraj.saeid@yahoo.com, ²chin@ump.edu.my

¹Chemical Engineering Department, Faculty of Engineering, Tobruk University, Tobruk, Libya

²Faculty of Chemical and Process Engineering Technology, Universiti Malaysia Pahang, Lebuhraya Tun Razak Gambang, 26300 Kuantan, Pahang Darul Makmur, Malaysia,

ABSTRACT

The simulation was performed to evaluate the impact of by-passing the second liquid phase reactor in the series due to the catalyst blockage problem leading to the shutdown in a butanol production plant. Butanal hydrogenation for the production of butanol was simulated using the Aspen Plus software package. The process thermodynamics was described by NRTL property model while the kinetic of the main and side reactions was represented by Pseudo-homogeneous model. The attested packed bed reactor model was subsequently used to simulate butanal hydrogenation at different reactor configuration, reactor operating conditions and reactor size. The targeted final total conversion was 99.5% of butanal. The butanal conversion increased with the increase of reaction temperature and residence time. On the other hand, the reactor pressure only affected the performance of the vapour phase reactor but not the liquid phase reactor. Conclusively, the idea of bypassing liquid phase reactor for hydrogenation of butanal in the series could be adopted to achieve the total targeted conversion, provided the plant is allowed to operate at a more severe operating condition. In the modified configuration I, the last reactor should operate at 166.5°C and 16 bar whereas, in the modified configuration II, the reactor should operate at 285°C and 30 bar, considering the safety factor.

Keywords:

Simulation,
butanal hydrogenation,
butanol,
packed bed
reactor,
reactor configuration,
operating condition.

1. Introduction

Butanol is a chemical that has excellent fuel characteristics. It contains approximately 22% oxygen, which will result in more complete fuel combustion when it is used as a fuel extender. Moreover, compared to ethanol, butanol has higher energy density and lower vapour pressure, so butanol is also considered a preferred fuel additive or even a potential replacement for gasoline. The use of butanol as fuel will contribute to clean air by reducing smog-creating compounds, harmful emissions (carbon monoxide) and unburned hydrocarbons in the tailpipe exhaust [1]. Besides, roughly 2×10^6 tons of butanol are produced annually [2] for use as a plasticiser, an industrial solvent, and an

intermediate in the production of butyl acetate, a key ingredient in lacquers and varnishes [3]. The demand for butanol is expected to increase in the future as a consequence of recent studies showing that butanol is a viable alternative to ethanol as an additive to gasoline [4,5].

While many processes exist for the production of butanol, such as the aldol condensation of ethanal [6], oxidation of butane [7], or the enzymatic fermentation of sugars [4,8], the overwhelming majority of butanol is produced in a three-step process involving the homogeneously catalyzed hydroformylation of propylene, separation of the resulting butanal, and subsequent hydrogenation of butanal through a series of high pressure and low pressure hydrogenation packed bed reactor. The Oxo synthesis (hydroformylation) process is based on the reaction of propylene with Oxogas to give a crude mixture of *n*- and iso-butanal to be reacted with hydrogen to form butanol via hydrogenation and distillation [9]. The adoption of a low-pressure rhodium-based catalyst system in place of high-pressure cobalt for the hydroformylation of propylene by reaction with carbon monoxide and hydrogen to produce butanal (an 'oxo' reaction) has brought large cost benefits to oxo producers. The benefits are derived from improved feedstock efficiency, lower energy usage and simpler and cheaper plant configurations [10].

A butanol production plant adopts the technology of using three packed bed reactors in series (1 vapour phase reactor and two liquid phase reactors) for the high pressure and low-pressure butanal hydrogenation respectively. This butanol production plant frequently encountered the catalyst blockage problem in the third reactor, the Low-Pressure Hydrogenation (LPH) reactor, which caused a plant shutdown. In view of minimising the production and revenue loss due to the plant shutdown, the plant operators propose to continue producing the demanded product with the two precede hydrogenation reactors in series, by-passing the third reactor. The proposal of excluding the LPH reactor from the reactor train could be evaluated through simulation analysis. Simulation is a process of designing an accurate operational model of a reactor and conducting investigations with this model for the purpose either of understanding the behaviour of the system or of evaluating alternative strategies for the development or operation of the system [11].

The aim of the present work is to simulate and investigate the effect of the reactor scheme and operating conditions to the reaction performance. The industrial plant data was used to validate the model. Three reactor configurations of butanal hydrogenation were considered for the simulation analysis. Original reactor configuration (single vapour phase reactor followed by two liquid phase reactors), modified configuration I (single vapour phase reactor followed by single liquid phase reactor), and modified configuration II (single vapour phase reactor).

2. Modelling Approach

2.1 Packed bed reactor model

A fixed packed bed reactor consists of a compact, immobile stack of catalyst pellets within a generally vertical vessel. On macroscopic scales, the catalyst bed behaves as a porous media. The fixed beds are thus employed as continuous tubular reactors in which the reactive species in the

mobile fluid (gas or liquid) phase are reacting over the catalyst surface (interior or exterior) in the stationary packed bed [12].

As with the PFR, the PBR is assumed to have no radial gradients in concentration, temperature, or reaction rate. The generalised mole balance on species A over catalyst weight ΔW results in the equation [13].

$$In - Out + Generation = Accumulation \quad (1)$$

$$F_{AW} - F_A(W + \Delta W) + r'_A \Delta W = 0$$

The differential form of the mole balance for a packed-bed reactor:

$$\frac{dF_A}{dW} = r'_A \quad (2)$$

If the pressure drop and catalyst decay are neglected,

$$W = \int_{F_A}^{F_{A0}} \frac{dF_A}{-r_A} \quad (3)$$

where W is the catalyst weight needed to reduce the entering molar flow rate of species A, F_{A0} , to a flow rate F_A .

The simulation of packed bed reactors for the production of butanol via hydrogenation of butanal was carried out using R-Plug in Aspen Plus. RPlug is a rigorous model for plug flow reactors which assumes that perfect mixing occurs in the radial direction and that no mixing occurs in the axial direction. RPlug can model one-, two-, or three-phase reactors. This model is useful when reaction kinetics is known, including reactions involving solids (Aspen Plus guideline 2016).

The process flow diagram of the original reactor configuration (one vapour phase reactor followed by two liquid phase reactors) is shown in Figure 1. The feed stream entered the first reactor with the total flow rate of 1506.26 kmol.hr⁻¹. It comprised of 0.703 mol% of H₂, 0.213 mol% of CH₄, 0.004 mol% of H₂O, 0.0058 mol% of *n*-butanal, 0.017 mol% of *i*-butanal, 0.0003 mol% of *n*-butanol, and 0.002 mol% of *i*-butanol. The reaction temperatures were 213°C, 159°C, and 167°C for the first, second and third reactor respectively. Whereas the reaction pressures for vapour phase and liquid phase reactors were 21 bar and 16 bar respectively.

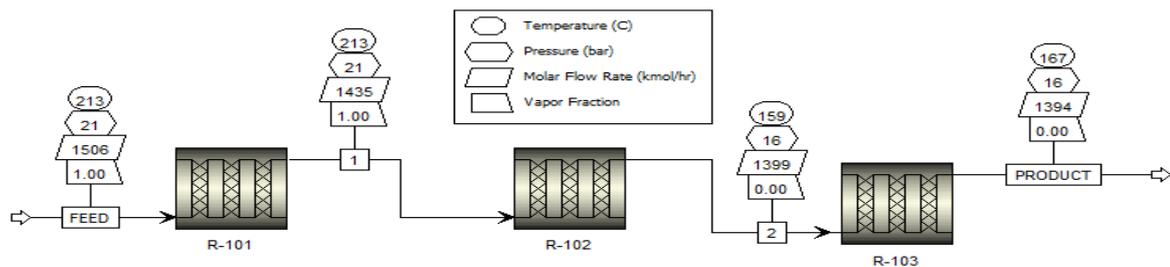


Figure 1: Original reactor configuration for butanal hydrogenation.

2.2 Kinetic model

2.2.1 Butanal hydrogenation

Eq.(4) and Eq.(5) show the chemical equation and kinetic model of *n*-butanal and *i*-butanal for the production of butanol through butanal hydrogenation, which is found as endothermic reaction.



$$r_{bals} = A \exp\left(-\frac{E}{T}\right) P_{bal}^{0.7} P_{H_2}^{0.5} \quad (5)$$

Where r_{bals} = Butanal reaction rate (Moles/liter.hr)

$$A_g = 33.38 \text{ Moles/liter.bar}$$

$$A_l = 1.59345 \times 10^7 \text{ Moles/liter.bar}$$

$$E_g = 2.9481 \times 10^{10} \text{ KJ/kmol}$$

$$E_l = 60520 \text{ KJ/kmol}$$

$$P_{bal} = \text{Partial pressure of butanal (bar)}$$

$$P_{H_2} = \text{Partial pressure of hydrogen (bar)}$$

Where “g” and “l” refer to gas and liquid phase reactions.

2.2.2 Side Reaction

Beside butanal hydrogenation, the elevated temperature in this process also promotes side products such as butyl butyrate. Butyl butyrate is produced from the hydrogenation of *n*-butanol, as shown in eq (6):



The rate equation for the side reaction is shown in eq (7).

$$r_{byprod} = A \exp\left(-\frac{E}{T}\right) P_{bal}^{1.8} \quad (7)$$

Where r_{byprod} = Byproduct reaction rate (Moles/liter.hr)

$$A_g = 3744.539 \text{ Moles/liter,bar}$$

$$A_l = 1.92079 \times 10^9 \text{ Moles/liter.bar}$$

$$E_g = 4.9884 \times 10^{10} \text{ KJ/kmol}$$

$$E_l = 100973.53 \text{ KJ/kmol}$$

$$P_{bal} = \text{Partial pressure of butanal (bar)}$$

$$P_{H_2} = \text{Partial pressure of hydrogen (bar)}$$

2.3 Plant data reconciliation

In order to validate the selected thermodynamics, reactor and kinetic models, the simulated results were compared with the industrial plant data obtained from a packed bed reactor. The present study was limited to the validation of the gas phase reactor due to the unavailability of plant data for the liquid phase reactor. The plant data consists of reaction temperatures and pressure for the vapour phase reaction as well the inlet and outlet mixture composition. The operating conditions are given in Table 1.

Table 1. *Industrial plant data (adopted from [14])*

Parameters	Inlet composition	Outlet composition
Temperature (°C)	125	125
Pressure (bar)	21	21
Flow rate (Kmol.h ⁻¹)	1506.25	1393.18
H ₂ (mol)	0.703	0.679
CH ₄ (mol)	0.213	0.230
H ₂ O (mol)	0.004	0.005
<i>n</i> -butanal (mol)	0.058	0
<i>i</i> -butanal (mol)	0.017	0
<i>n</i> -butanol (mol)	0.003	0.065
<i>i</i> -butanol (mol)	0.002	0.021

2.4 Simulation analysis to investigate the effect of different reactor configurations and operating condition

The effect of different reactor configurations was investigated at a wide range of operating conditions. These reactor configurations included original reactor configuration (single vapour phase reactor followed by two liquid phase reactors), modified configuration I (single vapour phase reactor followed by single liquid phase reactor), and modified configuration II (single vapour phase reactor). Table 2 provides the ranges of reactors operating conditions.

Table 2. *Operating conditions for hydrogenation reactors*

Parameter	Hydrogenation Unit
Reaction temperatures for vapour phase reactor.	135°C - 265°C
Reaction temperature for liquid phase reactors.	128°C - 196°C
Reaction pressure for both vapour and liquid phase reactors.	16 – 21 bar

1. Results and discussion

3.1 Model validation

In order to validate the selected thermodynamic, reactor and kinetic models, the simulated results were compared with the industrial plant data obtained from a packed bed reactor. The present study was limited to the validation of the gas phase reactor due to the unavailability of plant data for the liquid phase reactor. The gas-phase reactor was operated at 125°C and 21 bar. The feed flow rate was 1506.25 kmol/hr with the composition of H₂, H₂O, CH₄, *i*-butanal, *n*-butanal, *i*-butanol, and *n*-butanol.

As shown in Table 3 The butanal hydrogenation process is well described by the kinetic parameters since most of the data points of all components except H₂O are within the marginal error lines of 20%. The H₂O outlet composition is not accurately predicted due to its very low absolute value compared to the other components [15]. The *n*-butanal conversion was 99.77% comparing within the literature which is 99.99%. While the conversion of *i*-butanal was 100% typically the same conversion obtained. The reaction temperature was fixed as 125°C, and the reaction pressure was fixed as 21 bar, based on the literature operating conditions.

In view of the average deviation of less than 10%, the validity of the models was confirmed, and these models were used in the following simulation studies.

Table 3. Comparison of the simulation results with the industrial plant data

Industrial plant data	Simulation results	Relative Deviation (%)	
[15]			
Outlet composition			
N (Kmol.h ⁻¹)	1393.18	1393.48	0.0216
H ₂	0.679	0.679	0
CH ₄	0.230	0.230	0
H ₂ O	0.005	0.004	20
<i>n</i> -butanal	0	0.000143	0
<i>i</i> -butanal	0	0	0
<i>n</i> -butanol	0.065	0.065	0
<i>i</i> -butanol	0.021	0.021	0

3.2 Original reactor configuration

The simulation of a reactor configuration with three (3) packed bed reactors in series (one vapour phase reactor followed by two liquid phase reactors) was done. The effects of different operation conditions on the performance of the third reactor were investigated, and the corresponding simulation results are shown in the following sections.

3.2.1 Effect of reaction temperature

Based on the industrial plant data, the conversion of *n*-butanal in vapour phase reactor should achieve 60%, and for the liquid phase reactors, the conversion of the *n*-butanal should be 85% for the first liquid phase reactor and 99.5% for the second liquid phase reactor.

Figure 2. presents the variations of reactants conversions and product yields versus reaction temperature of the last liquid phase reactor in the series. The reaction was carried out at pressure of 21 bar for the vapour phase reactor and 16 bar for the liquid phase reactors. The reaction temperatures of the vapour phase reactor and first liquid phase reactor were fixed at 213 °C and 158.5 °C respectively. The reactors diameter and length were 2.75 m and 10.92 m, respectively.

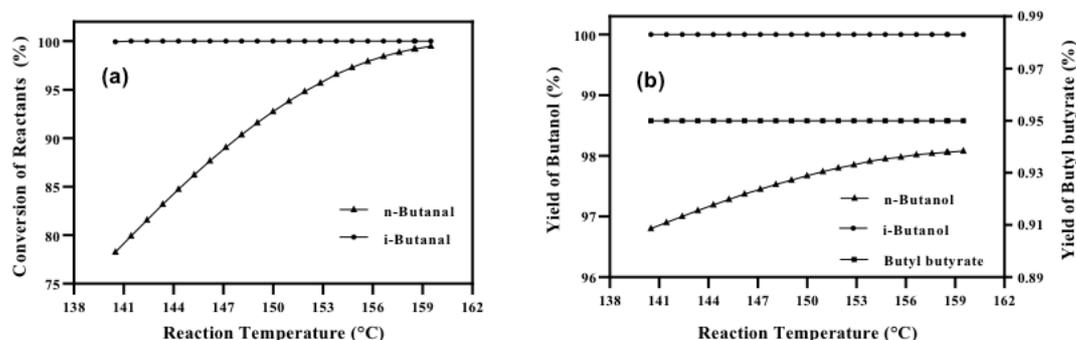


Figure 2: Effect of reaction temperature of the original configuration on (a) Conversion of reactants (b) Yield of products.

The conversion of *n*-butanal increased with the increase of temperature. While the conversion of *i*-butanal was almost constant. With the increase in reaction temperature from 140.5 °C to 159.5 °C, *n*-butanal conversion increased from 78.24% to 99.48%. In addition, the yield of *n*-butanol increased from 96.8% to 98.08%, the yield of *i*-butanol was unvaried and remained at 100%, and butyl butyrate yield was remained constant with a yield of 0.95%. The highest *n*-butanal conversion is 99.99%. and it was obtained at temperature of 167 °C.

3.2.2 Effect of reaction pressure

The effect of pressure on the conversion of reactants of the third reactor in the series can be seen in Figure 3. The operating temperatures of vapour phase reactor and two liquid phase reactors were fixed as 213 °C, 158.5 °C, and 159.5 °C. The reactors diameter and length were 2.75 m and 10.92 m, respectively.

The effect of pressure was investigated in the range of 16 to 21 bar for the second liquid phase reactor. It can be observed that when the pressure increased, the conversion of reactants was not affected negatively or a positively. Therefore, the conversion of *n*-butanal and *i*-butanal are remained constant at 99.48% and 100% as the pressure was increased from 16 to 21 bar. The concentration of reactants is insignificantly affected by even relatively large changes in the total pressure. Consequently, the effect of pressure can be totally ignored. Therefore, the pressure of 16 bar was chosen in the subsequent studies.

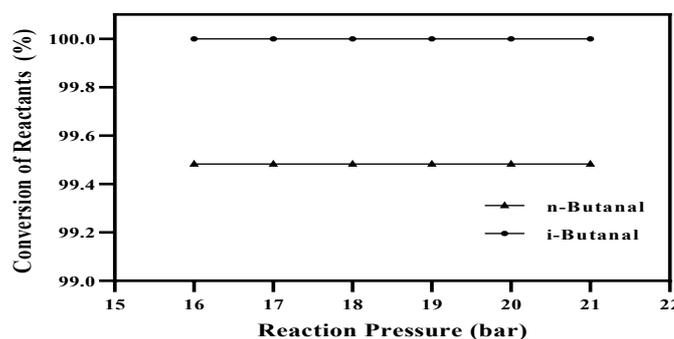


Figure 3: Effect of reaction pressure of the original configuration on the conversion of reactants

3.2.3 Effect of reactor size

The effects of the reactor diameter and length were investigated of the third reactor in the series in the hydrogenation of butanal to produce butanol. The reaction was carried out at 213 °C, 158.5 °C, and 159.5 °C of reaction temperature and 21 bar, 16 bar, and 21 bar of reaction pressure for single vapour and two liquid phase reactors. Figure 4 shows the results of the effect of reactor size.

The reactor volume changes with reactor length and reactor diameter. Reactor volume also affects the reaction performance. Therefore, sizing of the reactor is very important when high conversion and yield values desired. Therefore, the conversion of the reactants and the yield of products changes along the reactor. The reaction rate is a function of concentration; hence, it also varies with reactor length. The diffusion of the reactor is effect by the reactor size, which is the spontaneous intermingling or mixing of atoms or molecules by random thermal motion. It gives rise to motion of the species relative to motion of the mixture. In the absence of other gradients

(such as temperature, electric potential, or gravitational potential). Molecules of a given species within a single phase will always diffuse from regions of higher concentrations to regions of lower concentrations [14].

An increase in reactor length and diameter of the plug flow reactors caused a significant increase in the butanal conversion and in the yield of the products. The allowable L/D ratio is 4 to 10; therefore, the investigation of varies length and diameter was not exceeding the L/D ratio range through all the discussion. As seen in Figure 4 when the reactor diameter was increased from 1.25 to 2.75 m and reactor length was 10.92 m, *n*-butanal and *i*-butanal conversions increased from 24.32 to 99.48% and 35.38 to 100%. While the yield of *n*-butanol, *i*-butanol, and butyl butyrate increased from 49.53 to 98.08%, 64.92 to 100%, and 0.33 to 0.95%. While the reactor length was increased from 10.92 m to 12 m, reactor diameter was 2.75 m, *n*-butanal reaction conversion values increased from 99.48 to 99.98%. While the conversion of *n*-butanal at 12.25 m of reactor length decreased to 99.96% and within increasing of reactor length to 12.5 m, lead to increasing of *n*-butanal conversion with maximum amount of 99.99%. While the conversion of *i*-butanal was remaining constant at 100%. While the yield of *n*-butanol decreases with reactor length, butyl butyrate yield increases. The highest conversion of *n*-butanal was obtained as 99.99% when the reactor diameter and reactor length was 2.75 m and 12.5 m, respectively.

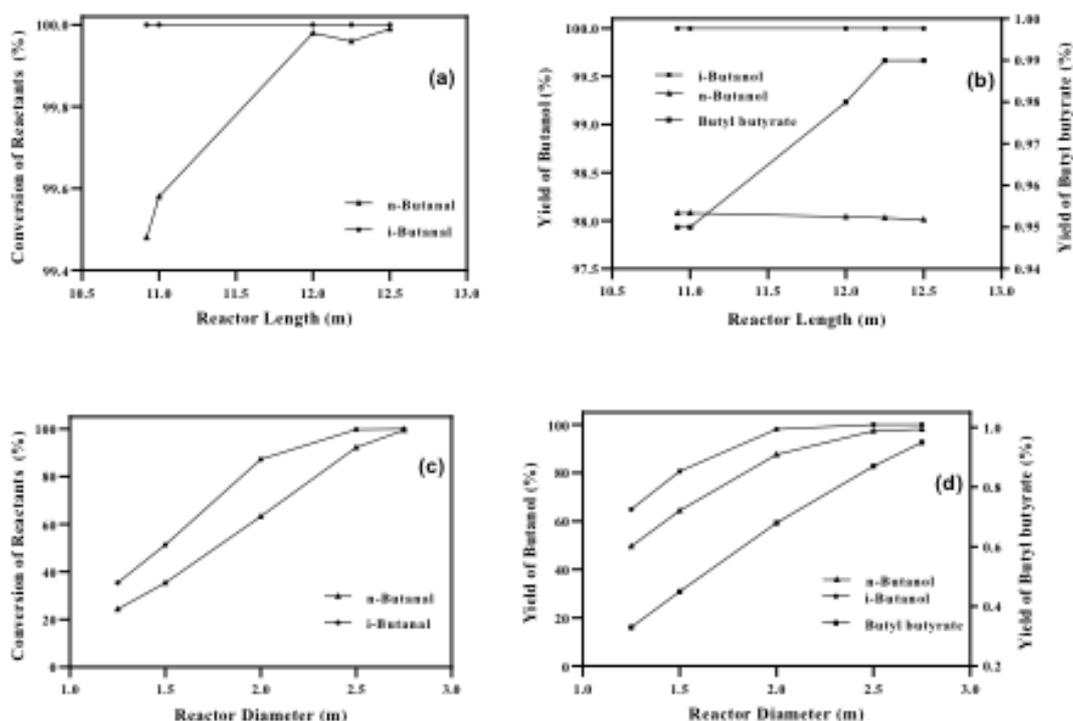


Figure 4: (a) Effect of reactor length of the original configuration on the conversion of reactants (b) Effect of reactor length on the yield of products (c) Effect of reactor diameter on the conversion of reactants (d) Effect of reactor diameter on the yield of products.

As shown in Table 4., the residence time of the second liquid phase reactor was calculated at varies reactor lengths and diameters. The reaction was carried out at 213 °C, 158.5 °C, and 159.5 °C in the first, second and third reactors respectively whereas the reaction pressures of these reactors are 21 bar, 16 bar, and 21 bar correspondingly. When the reactor diameter was increased from 1.25 to

2.75 m with the reactor length of 10.92 m, residence time increased from 5.83 to 28.8 minutes. While when the reactors length was increased from 10.92 to 12.5 m and reactor diameter fixed as 2.75m, residence time increased from 28.8 to 33 minutes. It is to conclude that the residence time of liquid phase reactors increases with increases of reactor length and diameter and resulting in increases of conversion values.

Table 4. *The effect of reactor size of the original configuration on the reactor residence time*
Reactor length fixed as 10.92 (m)

Diameter (m)	Residence time (min)	<i>n</i> -butanal conversion (%)	<i>i</i> -butanal conversion (%)
1.25	5.8	24.32	35.38
1.5	8.5	35.36	51.29
2	15.2	63.22	87.15
2.5	23.8	92.17	99.7
2.75	28.8	99.48	100

Reactor diameter fixed as 2.75 (m)			
Length (m)	Residence time (min)	<i>n</i> -butanal conversion (%)	<i>i</i> -butanal conversion (%)
10.92	28.8	99.48	100
11	29	99.58	100
12	31.6	99.98	100
12.25	32.3	99.96	100
12.5	33	99.99	100

3.3 Modified reactors configuration

Simulation was also performed using the modified reactor configurations. Modified Configuration I comprised of two reactors in series, a vapour phase reactor followed by a liquid phase reactor, while Modified Configuration II constituted by solely the vapour phase reactor. The effects of variation of temperature, pressure and reactor size are reported as follows.

3.3.1 Modified configuration I- bypassing the second low-pressure liquid phase hydrogenation reactor

3.3.1.1 Effect of reaction temperature

Figure 5 shows the effect of reaction temperature on the conversion of reactants and the yields of products. The reaction was carried out at pressure of 21 bar for the vapour phase reactor and 16 bar for the liquid phase reactor. The reaction temperature of the vapour phase reactor was fixed as 255 °C and the reaction pressure were fixed as 21 bar and 16 bar for the vapour and liquid phase reactors. The reactors diameter and length were 2.75 m and 10.92 m, respectively.

The conversion of *n*-butanal increased with the increase of temperature from 77.32% to 99.45% and the conversion of *i*-butanal increased from 99.07 to 100 % with a range of temperature from 146.5 °C to 166.5 °. In addition, the yield of *n*-butanol increased from 93.39% to 96.7%, the yield of *i*-butanol is linear at 100%, and butyl butyrate yield is remained constant with a yield of 1.61%. While the maximum *n*-butanal conversion was obtained at the temperature of 171 °C as 99.99%.

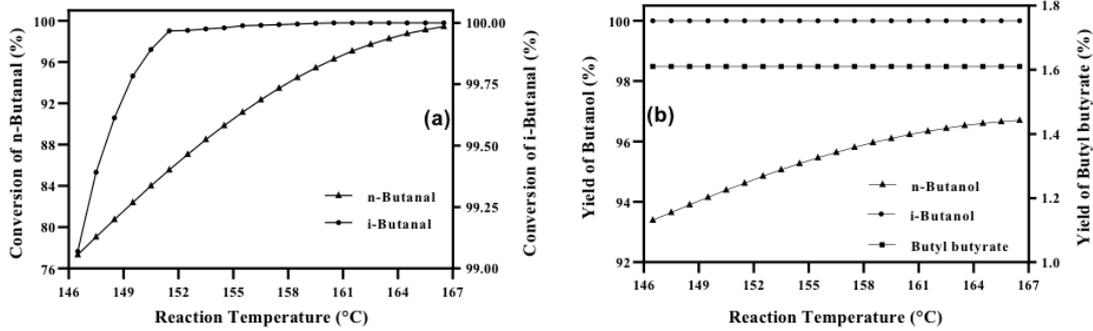


Figure 5: Effect of reaction temperature of the modified configuration I on (a) Conversion of reactants (b) Yield of products

3.3.1.2 Effect of reaction pressure

The effect of pressure on the conversion of reactants can be seen in Figure 6. The operation conditions of the vapour phase reactor and liquid phaser reactor were fixed as 255°C, and 166.5°C of the reaction temperature and the vapour phase reactor reaction pressure was fixed as 21 bar. The reactors diameter and length were fixed as 2.75 m and 10.92 m, respectively.

The effect of pressure was investigated in the range of 16 to 21 bar for the liquid phase reactor. It can be observed that when the pressure is increased, the conversion of reactants does not suffer a negative or a positive effect. Higher pressure or even lower pressure do not encourage butanal conversion. The maximum amount of reactants conversion can be achieved at any certain pressure. Therefore, the conversion of *n*-butanal and *i*-butanal are remained constant at 99.44% and 100% as the pressure was increased from 16 to 21 bar.

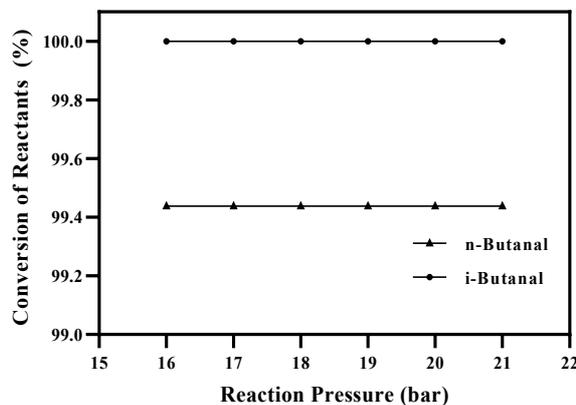


Figure 6: Effect of reaction pressure of the modified configuration I on the conversion of reactants.

3.3.1.3 Effect of reactor size

The reaction was carried out at 255 °C and 166.5 °C for the reaction temperature and 21 bar and 16 bar of reaction pressure for single vapour and liquid phase reactors. The obtained results are shown in Figure 7.

An increase in reactor length and diameter of the plug flow reactors caused a significant increase in the butanal conversion and the yield of the products. As seen in Figure 4 when the reactor

diameter was increased from 1.25 to 2.75 m and reactor length was 10.92 m, *n*-butanal and *i*-butanal conversions increased from 30.07 to 99.44% and 42.41 to 100%. While the yield of *n*-butanol, *i*-butanol, and butyl butyrate increased from 49.53 to 98.08%, 64.92 to 100%, and 0.78 to 1.61%. While the reactor length was increased from 10.92 m to 12 m, reactor diameter was 2.75 m, *n*-butanal reaction conversion values increased from 99.44 to 99.99%. While the conversion of *n*-butanal at 12.5 m of reactor length decreased to 99.98%. While the conversion of *i*-butanal was remaining constant at 100%. The yield of *n*-butanol increases at certain reactor lengths and decreases with others and butyl butyrate yield increases. The highest conversion of *n*-butanal was obtained as 99.98% when the reactor diameter and reactor length was 2.75 m and 12.25 m, respectively.

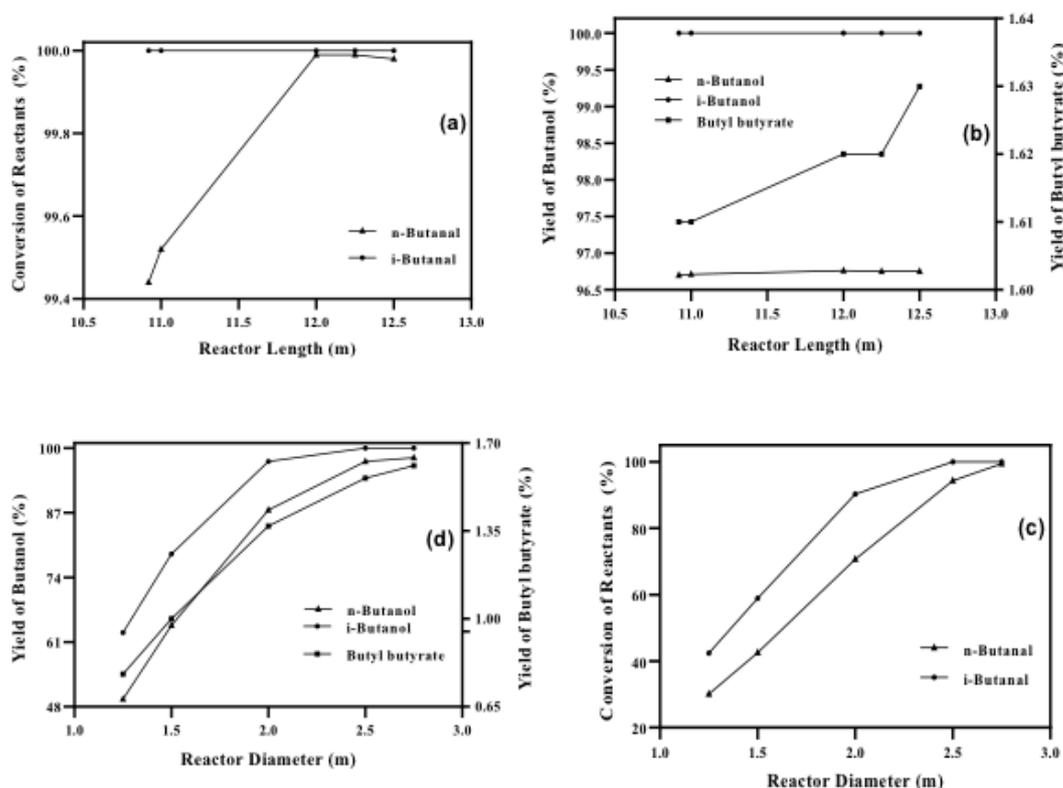


Figure 7: (a) Effect of reactor length of the modified configuration I on the conversion of reactants (b) Effect of reactor length on the yield of products (c) Effect of reactor diameter on the conversion of reactants (d)Effect of reactor diameter on the yield of products.

Table 5. Shows the residence time of the liquid phase reactor at varies reactor lengths and diameters. The reaction was carried out at 255 °C, 166.5 °C temperature and 21 bar, 16 bar of reaction pressure for single vapour and liquid phase reactors. When the reactor diameter was increased from 1.25 to 2.75 m, and reactor length was 10.92 m, residence time increases from 5.83 to 28.8 minutes. While when the reactors length was increased from 10.92 to 12.5 m and reactor diameter fixed as 2.75m, residence time increase from 28.8 to 33 minutes. We concluded that the residence time of liquid phase reactors increases with increases of reactor length and diameter and resulting in increasing of conversion values.

Table 5. The effect of reactor size of the modified configuration I on the reactor residence time

Reactor length fixed as 10.92 (m)			
Diameter (m)	Residence time (min)	<i>n</i> -butanal conversion (%)	<i>i</i> -butanal conversion (%)
1.25	5.8	30.07	42.41
1.5	8.4	42.55	58.95
2	15.2	70.66	90.27
2.5	23.8	94.31	99.98
2.75	28.8	99.44	100
Reactor diameter fixed as 2.75 (m)			
Length (m)	Residence time (min)	<i>n</i> -butanal conversion (%)	<i>i</i> -butanal conversion (%)
10.92	28.8	99.44	100
11	29	99.52	100
12	31.6	99.99	100
12.25	32.3	99.99	100
12.5	33	99.98	100

3.3.2 Benchmarking the simulation results of modified configuration I with that in the original reactor configuration

Based on the plant data, the targeted total conversion from the last reactor in the series should be 99.5%. In the Modified Configuration I, the total conversion of 99.44% from the first liquid phase reactor in the series obtained at reaction temperature of 166.5°C and reaction pressure of 16 bar. While the total conversion of the vapour phase reactor was fixed as 85% obtained at reaction temperature of 255°C and reaction pressure of 21 bar. Comparing with those in the original reactor configuration the total conversion of 99.48% from the second liquid phase reactor in the series obtained at reaction temperature of 159.5°C and reaction pressure of 16 bar. While the reactor length and diameter were fixed as 10.92 (m) and 2.75 (m) in the both cases. A more severe reaction temperature was required in the Modified Configuration I for achieving the total targeted conversion.

3.3.3 Modified configuration II- bypassing the first low-pressure liquid phase hydrogenation reactor

3.3.3.1 Effect of reaction temperature

Figure 8. shows the variations of reactants conversions and product yields versus reaction temperature of a single vapour phase reactor. The reaction was carried out at pressure of 21 bar. The reactor diameter and length were 2.75 m and 10.92 m, respectively.

The total conversion of *n*-butanal and *i*-butanal increased with the increases of temperature. *n*-butanal conversion increased from 54.96% to 89.58% with industrial plant operation conditions limits. While the conversion of *i*-butanal increased from 69.34 to 97.97 %. In addition, the yield of *n*-butanol increased from 53.31 to 86.06%, the yield of *i*-butanol increases from 69.34 to 97.97% and butyl butyrate yield is also increased from 0.82 to 1.76.% In the case of want to exceed the

industrial plant operation conditions, the maximum *n*-butanal conversion was obtained at the temperature of 285 °C and reaction pressure of 30 bar as 99.99%.

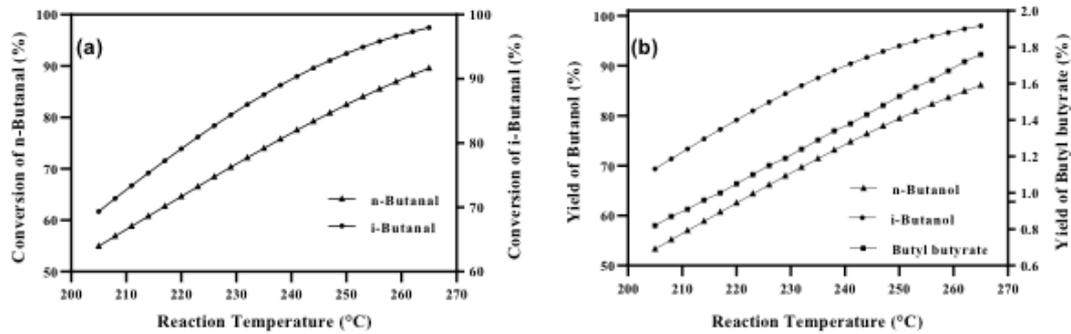


Figure 8: Effect of reaction temperature of the modified configuration II on (a) Conversion of reactants (b) Yield of products

3.3.3.2 Effect of reaction pressure

The effect of pressure on the conversion of reactants and the yields of the products can be seen in Figure 9. The operation temperature of vapour phase reactor was fixed as 265°C. The reactors diameter and length were 2.75 m and 10.92 m, respectively.

The effect of pressure was investigated in the range of 16 to 21 bar for the vapour phase reaction. It can be observed that when the pressure is increased, the conversion of reactants has a positive effect. Higher pressure in vapour phase reaction encourages butanal conversion. However, in gas-phase reactions, the concentration of the reacting species is proportional to the total pressure (Scott Fogler 2016). The maximum amount of reactants conversion can be achieved at 21 bar. Therefore, the conversion of *n*-butanal and *i*-butanal are increased from 76.79 to 89.58% and from 89.8 to 97.97%. While the yields of the products increase from 73.9 to 86.06% of *n*-butanol yield and 89.77 to 97.97% of *i*-butanol, for the yield of butyl butyrate it increased from 1.45 to 1.76% as the pressure was increased from 16 to 21 bar.

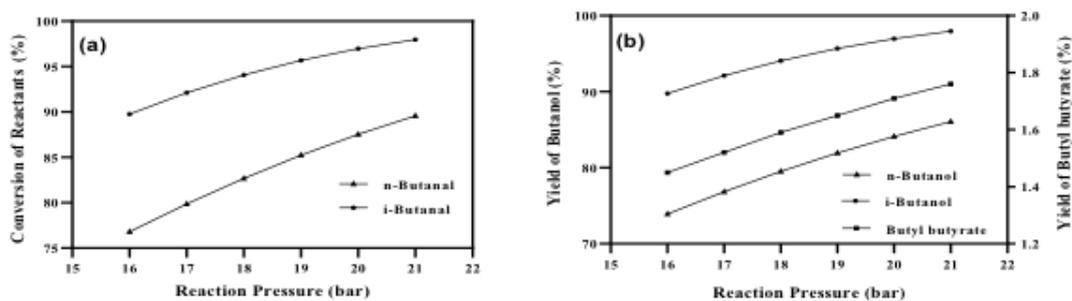


Figure 9: Effect of reaction pressure of the modified configuration II on (a) Conversion of reactants (b) Yield of products

3.3.3.3 Effect of reactor size

The reaction was carried out at 265 °C and 21 bar of reaction pressure for vapour phase reaction. The obtained results are shown in Figure 10. The reactor volume changes with reactor length and reactor diameter. Reactor volume also affects the reaction kinetics. Therefore, sizing of the reactor

is very important when high conversion and yield values desired. Therefore, the conversion of the reactants and the yield of products changes along the reactor. The reaction rate is a function of concentration; hence it also varies with reactor length. An increase in reactor length and diameter of the plug flow reactors caused a significant increase in the butanal conversion and the yield of the products. As seen in Figure 4 when the reactor diameter was increased from 1.25 to 2.75 m and reactor length was 10.92 m, *n*-butanal and *i*-butanal conversions increased from 30.46 to 89.58% and 39.57 to 97.97%. While the yield of *n*-butanol, *i*-butanol, and butyl butyrate increased from 28.62 to 86.06%, 39.57 to 97.97%, and 0.92 to 1.76%. While the reactor length was increased from 10.92 m to 12 m, reactor diameter was 2.75 m, *n*-butanal reaction conversion values increased from 89.58 to 93.66%. While the conversion of *i*-butanal was increasing from 97.97 to 99.45%. The yield of *n*-butanol and *i*-butanol increases from 86.06 to 90.12% and 97.97 to 99.45%. The butyl butyrate yield value is small with a variation of reactor length, which is increased from 1.76 to 1.77%. The highest conversion of *n*-butanal was obtained as 93.65% when the reactor diameter and reactor length were 2.75 m and 12.5 m, respectively.

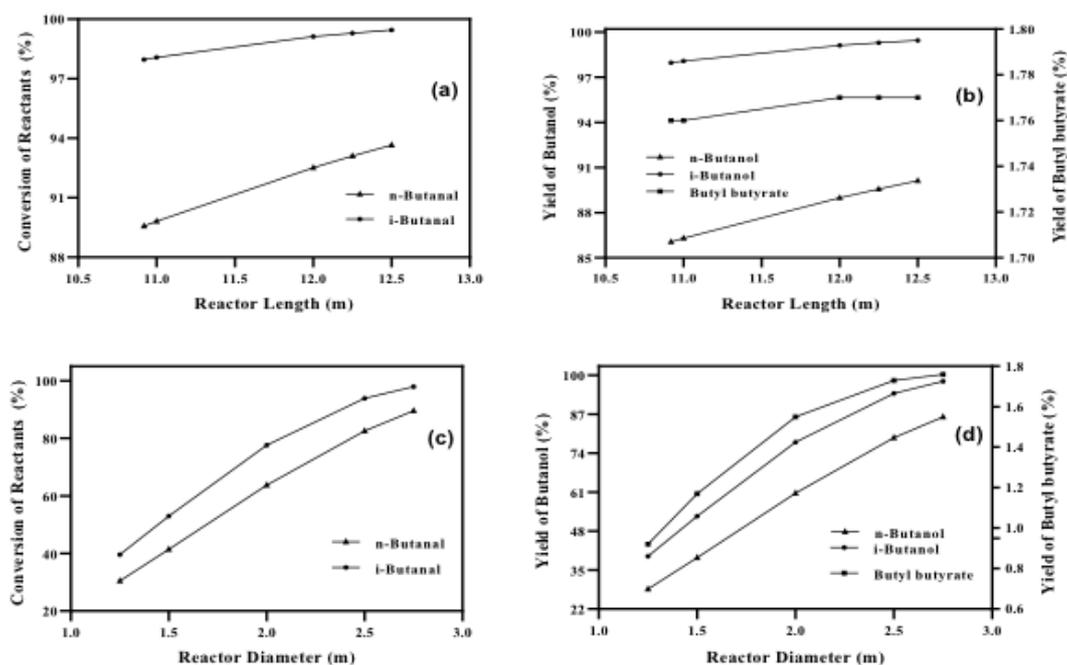


Figure 10: (a) Effect of reactor length of the modified configuration II on the conversion of reactants (b) Effect of reactor length on the yield of products (c) Effect of reactor diameter on the conversion of reactants (d) Effect of reactor diameter on the yield of products

Table 6. shows the residence time of the vapour phase reaction at varies reactor lengths and diameters. The reaction was carried out at 265 °C of reaction temperature and 21 bar of reaction pressure. When the reactor diameter was increased from 1.25 to 2.75 m, and reactor length was 10.92 m, residence time increases from 0.257 to 1.3 minutes. While when the length of the reactors was increased from 10.92 to 12.5 m and reactor diameter fixed as 2.75m, residence time increase from 1.3 to 1.5 minutes. We concluded that the residence time of vapour phase reaction increases with increases of reactor length and diameter and resulting in increases of conversion values.

Table 6. The effect of reactor size of the modified configuration II on the reactor residence time
 Reactor length fixed as 10.92 (m)

Diameter (m)	Residence time (min)	<i>n</i> -butanal conversion (%)	<i>i</i> -butanal conversion (%)
1.25	0.3	30.46	39.57
1.5	0.4	41.45	52.95
2	0.7	63.76	77.63
2.5	1.1	82.59	93.89
2.75	1.3	89.58	97.97

Reactor diameter fixed as 2.75 (m)			
Length (m)	Residence time (min)	<i>n</i> -butanal conversion (%)	<i>i</i> -butanal conversion (%)
10.92	1.3	89.58	97.97
11	1.3	89.82	98.08
12	1.4	92.52	99.12
12.25	1.5	93.12	99.29
12.5	1.5	93.66	99.45

3.3.4 Benchmarking the comparison of the simulation results of modified configuration II with that in the original reactor configuration

In the Modified Configuration II, the total conversion from the vapour phase reactor can be obtained when the reaction temperature fixed as 285 °C and reaction pressure of 30 bar. Comparing with that in the original reactor configuration the total conversion value from the second liquid phase reactor in the series obtained at reaction temperature of 159.5°C and reaction pressure of 16 bar. While the reactor length and diameter were fixed as 10.92 (m) and 2.75 (m) in both cases. The total *n*-butanal conversion in the Modified Configuration II was 99.99 % in the case of the reactor operating conditions exceeded the industrial plant operating conditions. While in the original configuration was 99.48%. In relative to the original configuration and Modified Configuration I, the more severe operating temperature and pressure were needed to attain the desired total conversion.

3.4 Comparison of data with that in the literature

Table 7. Compares the *n*-butanal conversion to *n*-butanol obtained from the present study with the findings of other researches about hydrogenation of butanal. In literature, hydrogenation of butanal was mostly carried out experimentally. Among the limited number of simulation studies found in literature, the effect of temperature was usually investigated.

Table 7. Comparison of conversion values with those in the literature

Temperature (°C)	Pressure (bar)	Catalyst type	Reactor type	Reaction Phase	Conversion of <i>n</i> -butanal (%)	References
139.85	3.1	SX-Rh SILP, and Shvo/SiO ₂ catalysts	Packed bed reactor	Vapour	90	[16]
125	21	G-66 copper base catalyst	Tubular reactor	Vapour	99.99	[14]
210	5.8	Copper-zinc oxide catalysts	Fixed bed reactor	Vapour	99	[17]
159.5	16	Copper-zinc oxide catalysts	Packed bed reactor	Liquid	99.48	This study
166.5	16	Copper-zinc oxide catalysts	Packed bed reactor	Liquid	99.45	This study
265	21	Copper-zinc oxide catalysts	Packed bed reactor	Vapour	89.58	This study

In this study, the conversion of *n*-butanal was mostly higher than in other studies in the different reaction conditions. Also, in the present studies, the high conversion of *n*-butanal was obtained under the higher temperature according to the literature. The differences between the conversion of values in our study and the values presented in the literature are related to reactor type, reactor size, and reaction phase. Packed bed reactor and NRTL thermodynamic model were chosen in this study. These are important points for the conversion values. The most appropriate reactor type, reactor sizing and thermodynamic model is resulted in high *n*-butanal conversion value. In addition, optimization of operating parameters is also important for the high conversion and yield values.

2. Conclusion

Aspen Plus was successfully applied to simulate and investigate the effect of the reactor scheme and operating conditions to the reaction performance. In order to attain the total conversion of *n*-butanal as 99.44%, the best-operating conditions for the last reactor in the modified configuration I were 166.5°C and 16 bar. While a total conversion of *n*-butanal of 99.99% was achieved at the best-operating conditions of the modified configuration II at 30 bar and 285 °C in the expense of a condition that exceeded the industrial operating conditions limits. The best parameters obtained from the simulation study for butanal hydrogenation to produce butanol indicates considerable potential and courage for bypassing the last liquid phase reactor in the original reactor configuration.

3. Acknowledgements

The authors would like to acknowledge the Faculty of Chemical and Process Engineering Technology, University Malaysia Pahang for providing the needed facilities in finishing the present work.

References

- [1] Ezeji TC, Qureshi N, Karcher P, Blaschek HP. Production of butanol from corn. In: Minter SD, editor. Alcoholic fuels.
- [2] A. Chauvel, G. Lefebvre, in: Petrochemical Processes, second ed., Editions Technip, Paris, 1989, pp. 81–90.
- [3] P. Durre, Ann. N. Y. Acad. Sci. 1125 (2008) 353.
- [4] S. Lee, M.O. Cho, C.H. Park, Y. Chung, J.H. Kim, B. Sang, Y. Um, Energy Fuels 2 (2008) 3459.
- [5] E. Christensen, J. Yanowitz, M. Ratcliff, R.L. McCormick, Energy Fuels 25 (2011) 4723.
- [6] J.E. Rekoske, M.A. Barteau, Ind. Eng. Chem. Res. 50 (2011) 41.
- [7] S. Marengo, P. Comotti, G. Galli, Catal. Today 81 (2003) 205.
- [8] Y. Lee, J.H. Park, S.H. Jang, L.K. Nielsen, J. Kim, K.S. Jung, Biotechnol. Bioeng. 101 (2008) 209.
- [9] Brito, M., & Martins, F. (2017). Life cycle assessment of butanol production. Fuel, 208, 476– 482. doi:10.1016/j.fuel.2017.07.050.
- [10] Tudor, R., & Ashley, M. (2007). Enhancement of Industrial Hydroformylation Processes by the Adoption of Rhodium-Based Catalyst: Part I. Platinum Metals Review, 51(3), 116–126. doi:10.1595/147106707x216855.
- [11] Gil Chaves, I. D., López, J. R. G., García Zapata, J. L., Leguizamón Robayo, A., & Rodríguez Niño, G. (2016). Process Analysis and Simulation in Chemical Engineering. doi:10.1007/978-3-319-14812-0.
- [12] Jakobsen, H. A. (2008). Packed Bed Reactors. Chemical Reactor Modeling, 953–984. doi:10.1007/978-3-540-68622-4_11.
- [13] Scott, H. (2016). Elements of Chemical Reaction Engineering, Fifth Edition. Prentice Hall.
- [14] Zhang, H., Zhang, X., & Fang, D. (2005). Mathematical Modeling of the Reactor Synthesizing Butanol from Butyraldehyde by Hydrogenation, Journal of East China University of Science and Technology (Natural Science Edition), 31(1), 5–8.
- [15] Chin, S. Y., Hisyam, A., & Prasetiawan, H. (2016). Modeling and Simulation Study of an Industrial Radial Moving Bed Reactor for Propane Dehydrogenation Process. International Journal of Chemical Reactor Engineering, 14(1). doi:10.1515/ijcre-2014-0148.
- [16] Hanna, D. G., Shylesh, S., Parada, P. A., & Bell, A. T. (2014). Hydrogenation of butanal over silica-supported Shvo's catalyst and its use for the gas-phase conversion of propene to butanol via tandem hydroformylation and hydrogenation. Journal of Catalysis, 311, 52–58. doi:10.1016/j.jcat.2013.11.012.
- [17] John E. Logsdon, Houston; Richard A. Loke, Santa Fe, both of Tex.; Jay S. Merriam, Louisville, Ky.; Richard W. Voight, Houston, Tex. (1989). U.S. Patent No. US4876402A. Washington, DC: U.S. Patent and Trademark Office.

Evaporation Process Control using MIMO MPC

Dr. Yousif A. alsadiq

ysf_826@yahoo.co.uk

Sirte University, Engineering faculty, ChE Department

ABSTRACT

Keywords:

Chemical process
Control,
Proportional-
Integral
Controller, Model
Predictive Control,
MATLAB /
Simulink

In the processing industry, controllers play a crucial role in keeping our plants running—virtually everything from simply filling up a storage tank to complex separation processes, and to chemical reactors. There are some important issues when we design a control system. In the first place, we need to identify the role of various variables. We need to determine what we need to control, what we need to manipulate, what are the sources of disturbances, and so forth. However, chemical processes are highly non-linear in nature, especially when they have multiple inputs – multiple outputs (MIMO) variables with complex interactions. Evaporators usually operate before a major drying process, which require more energy and is difficult to control. It is therefore important to achieve good control in the evaporation stage so that the drying process is operating with steady inputs. In this paper three of main evaporated process variables such as Liquid level in separator, operation pressure, and the product concentration are first controlled using PI controller in the presence of a variables step changes and a load disturbance. Advanced controllers such as model predictive control that is used for a wide range of application in the process industry. The potential utilization of such advanced predictive controllers is to design control systems that give effective control in this multivariable environment. Model predictive control is applied to the evaporated process with same mentioned conditions for variable step changes and load applied in case of PI controller. The objective of this paper is to present and illustrates in a comparatively study to the results obtained by PI controller, the use of MPC in providing an effective control for a MIMO evaporator plant in the presence of step and load disturbances change. The sum of the integral of absolute error (IAE) is used as a criterion for evaluating the controller's performance.

1. Introduction

The principle in the evaporation process is separation by evaporation from a liquid mixture where at least one component is not volatile. The physical process of evaporation requires the input of energy in the form of heat to convert a liquid into vapour. Since all evaporators use the process of evaporation to remove water, every evaporator requires a source of heat to operate [1]. The heat source for almost all evaporators is water vapour, either in the form of boiler steam or waste vapour from another process. Evaporators come in many different

shapes and sizes. Selecting the best evaporator for a given application can sometimes be a confusing and even intimidating task. Technical terms like falling film, forced circulation, and multiple effect can add to the confusion [1,2]. In this article we will take a brief, not-too-technical look at the most common types of evaporators, how they work, and some of their applications. In this paper the study focuses on the forced circulation evaporation type.

2. Evaporation Process

Evaporation is an engineering operation that is usually used to remove a liquid from a solution, suspension or liquor by boiling some of the liquid. It is usually treated as a separation of liquid mixtures into concentrate and vapour. This operation is usually performed in a heating device called an evaporator. The evaporator is made of a heat exchanger for heating up a solution and a

way to remove the vapour from the boiling solution [1]. The forced circulation evaporation process is of our interest case of study in this paper.

Forced Circulation Evaporation Process

The forced circulation evaporator was developed for processing liquors which are susceptible to scaling or crystallizing. Liquid is circulated at a high rate through the heat exchanger, As the liquid enters the separator where the absolute pressure is slightly less than in the tube bundle, the liquid flashes to form a vapour [2]. The forced-circulation evaporator is a common processing unit in milk, several types of food, sugar mills, alumina production and paper manufacture. This process is used to concentrate a dilute liquor by evaporating its solvent (usually water) as shown in the Figure 1.

A feed stream with solute of concentration C_1 (mass percentage) is mixed with high volumetric recycle flow rate and fed into a vertical evaporator (heat exchanger). The mixture solution will pass through the tube. A saturated input steam is used to heat up the mixture by condensing on the outer surface of the tubes. The liquor which passes up inside the tube boils and then passes to a separator vessel. In the separator, the liquid and vapour are separated at constant temperature and pressure. The liquid is recycled with some being drawn off as product with solute concentration of C_2 . The vapour is usually condensed with water. Water is used as the coolant. The case study is the Newell and Lee forced circulation evaporator model [3].

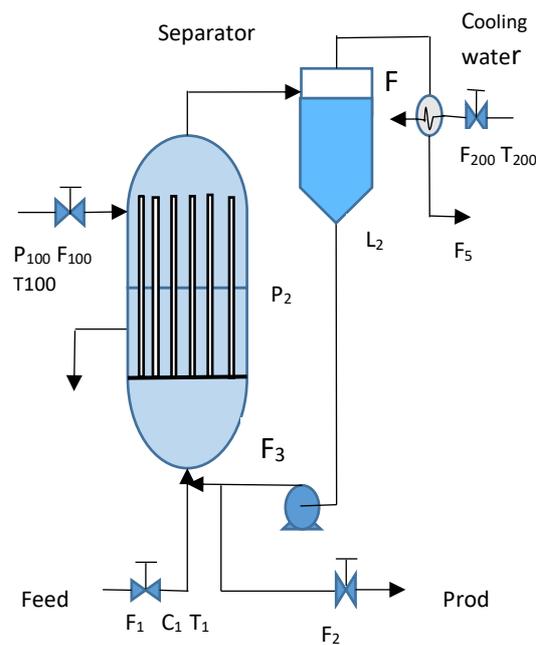


Figure 1: Forced Circulation Evaporator Process

Modelling the process

The process model can be divided to four parts as follow:

2.1.1 The separator

$$\rho A \frac{dL_2}{dt} = F_1 - F_4 - F_5 \quad (1)$$

Where, ρ is the liquid density and A is the cross-sectional area of the separator and ρA is assumed to be 20 kg/m.

2.1.2 The evaporator:

is represented by 5 equations

$$M \frac{dC_2}{dt} = F_1 C_1 - F_2 C_2 \quad (2)$$

$$\beta \frac{dP_2}{dt} = F_4 - F_5 \quad (3)$$

$$T_2 = 0.5616P_2 + 0.3126C_2 + 48.43 \quad (4)$$

$$T_3 = 0.507P_2 + 55.0 \quad (5)$$

where M is a constant liquid hold up in the evaporator of 20 kg. C_p and λ are the heat capacity and the latent heat of evaporation of the process liquid which is assumed constant at 0.07 kW/K(kg/min) and 38.5 kW/(kg/min) respectively. The constant $\beta = 4$ kg/kPa is used to convert a mass of vapour into a pressure in the vessel.

2.1.3 The steam:

The steam side of the evaporator is model with three algebraic equations as the dynamics are assumed to be very fast

$$T_{100} = 0.1538P_{100} + 90.0 \quad (6)$$

$$Q_{100} = 0.16(F_1 + F_3) (T_{100} - T_2) \quad (7)$$

$$F_{100} = \frac{Q_{100}}{\lambda_s} \quad (8)$$

where $\lambda_s = 36.6$ kW/(kg/min) is the latent heat for steam. The term $0.16(F_1 + F_3)$ correlates the flow to the evaporator to the overall heat transfer coefficient times the area, UA1, at the given process conditions.

2.1.4 The condenser:

The condenser is also modelled with the following set of algebraic equations:

$$Q_{200} = \frac{UA_2(T_3 - T_{200})}{1 + \frac{UA_2}{2C_p F_{200}}} \quad (9)$$

$$T_{201} = T_{200} + \frac{Q_{200}}{F_{200} C_p} \quad (10)$$

$$F_5 = \frac{Q_{200}}{\lambda} \quad (11)$$

where $UA_2 = 6.84 \text{ kW/K}$ is the overall heat transfer coefficient times the area and the heat of evaporation, λ is 38.5 kW/(kg/min) .

Pairing Variables

the layout of an evaporator has 8 degrees of freedom which can be categorised as manipulated variables, MV, and variables of disturbance D. The controlled variables CV are seen from the model's differential equations and all of these are evaluated. These three process variables are the system's desired control variable. This provides the following general overview of the model system. The relative gain array (RGA) can be used in order to get the best variable pairing that could achieve the best control performance [4]. The selected manipulated and controlled pairing variables are the product flow rate F_2 to control the product level L_2 in the separator, the cooling flow rate F_{200} to control operation pressure P_2 , and feed flow rate F_1 to control the product concentration C_2 . However, disturbances variables D, used are cooling water inlet temperature T_{200} , and the feed concentration C_1 .

Table 1. Nominal steady state parameters values for the evaporator system.

Variable	Variable Description	Value	Unit
F_1	Feed flowrate	10.0	kg/min
F_2	Product flowrate	2.0	kg/min
F_3	Circulating flowrate	50.0	kg/min
F_4	Vapour flowrate	8.0	kg/min
F_5	Condensate flowrate	8	kg/min
C_1	Feed concentration	5	%
C_2	Product concentration	25	%
T_1	Feed temperature	40.0	0C
T_2	Product temperature	84.6	0C
T_3	Circulating temperature	80.6	0C
L_2	Separator level	1.0	M
P_2	Operation pressure	50.5	kPa
F_{100}	Steam flowrate	9.3	kg/min
T_{100}	Steam temperature	119.9	0C
P_{100}	Steam pressure	194.7	kPa
Q_{100}	Heater duty	339.9	kW
F_{200}	Cooling flowrate	208.0	kg/min
T_{200}	Cooling water inlet temperature	25.0	0C
T_{201}	Cooling water outlet temperature	46.1	0C
Q_{200}	Condenser duty	307.9	kW

3. Controllers

There are many open loop and closed loop (feedback) controls strategies like, Proportional controller, Integral controller, Derivative controller, combination of these, PI, PD, PID, fuzzy logic controllers (FLC) etc which are available and proved its influence in process control.

3.1 PID Controllers

A PID controller represents the simplest form of controller that utilises Derivative and Integral operations on the system. PID controllers have several important functions: they have the ability to eliminate steady-state error through the integral action, and they can cope with actuator saturation, if used with anti-windup. These controllers are also effective for many control problems, particularly where there are benign process dynamics and modest performance requirements [5]. A PID controllers are also effective for particularly where there are benign processes dynamics and modest performance requirements [5]. The simplest form of PID controller can be represented by equation.

$$C(t) = K_c \left(e(t) + \frac{1}{\tau_i} \int_0^t e(t) dt + \tau_D \frac{de(t)}{dt} \right) \quad (12)$$

Where, K_c = Proportional constant, τ_i = integral time constant, τ_D = Derivative time constant, $e(t)$ = error (controller input), $C(t)$ = Controller command (controller output).

3.2 Model Predictive Controller (MPC)

Model Predictive Control (MPC) has a long history in control engineering. It is one of the few fields that has attracted ongoing attention from researchers in both the industrial and academic communities. Four main aspects of model predictive control make the design approach appealing to both practitioners and academics. The first factor is the architecture, which uses a completely multivariable system structure, where the output parameters of the multivariable control system are related to the technical aspects of the system; thus, they can be understood and 'tuned' by engineers. The second factor is the potential of the system to manage both 'soft' constraints and hard constraints in a multivariate control setting. This is especially appealing to industry, where there are expected to be tight profit margins and limitations on the process activity. The third element is the ability to perform on-line optimization of the operation. The fourth element is the simplicity of the framework architecture [6,7,8].

Model Predictive Control Toolbox provides functions, an app, and Simulink® blocks for systematically analysing, designing, and simulating model predictive controllers. You can specify plant and disturbance models, horizons, constraints, and weights. The toolbox enables

you to diagnose issues that could lead to run-time failures and provides advice on tuning weights to improve performance and robustness. By running different scenarios in linear and nonlinear simulations, you can evaluate controller performance [9]. The plant model is developed on a Simulink and connected to an MPC block as in Figure [2].

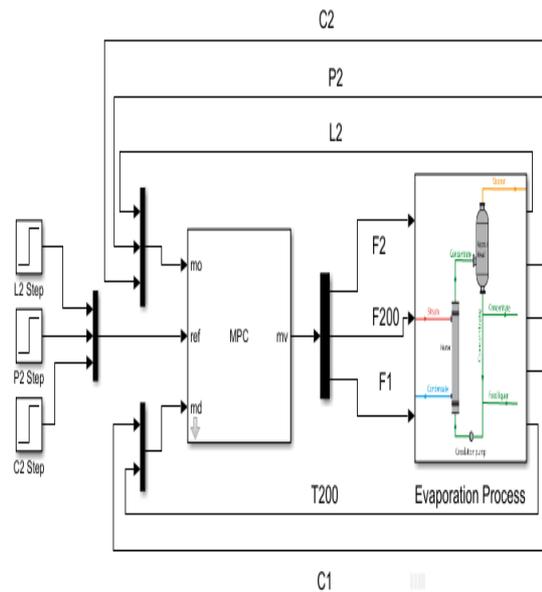


Figure 2. The Overall Control System

3.2.1 General formulation of MIMO MPC

The MPC controller performs all estimation and optimization calculations using a discrete-time, delay-free, state-space system with dimensionless input and output variables. Assume that the plant has m inputs and also denote to m^{th} future sampling discrete., q outputs and n_1 states. We also assume that the number of outputs is less than or equal to the number of inputs (i.e., $q \leq m$). In the general formulation of the predictive control problem, we also take the plant noise and disturbance into consideration. The multi-input multi-output (MIMO) model predictive control system is represented as follow [6].

$$x_m(k+1) = A_m x_m(k) + B_m u(k) + B_d \omega(k) \quad (13)$$

$$y(k) = C_m x_m(k) \quad (14)$$

where u is the manipulated variable or input variable, y is the process output, x_m is the state variable vector, and $\omega(k)$ is the input disturbance, assumed to be a sequence of integrated white noise. This means that the input disturbance $\omega(k)$ is related to a zero mean, white noise sequence $\epsilon(k)$ by the difference equation

$$\omega(k) - \omega(k-1) = \epsilon(k) \quad (15)$$

Note that from (13), the following difference equation is also true:

$$x_m(k) = A_m x_m(k-1) + B_m u(k-1) + B_d \omega(k-1) \quad (16)$$

By defining $\Delta x_m(k) = x_m(k) - x_m(k-1)$ and $\Delta u(k) = u(k) - u(k-1)$, then subtracting (16) from (13) leads to

$$\Delta x_m(k+1) = A_m \Delta x_m(k) + B_m \Delta u(k) + B_d \epsilon(k) \quad (17)$$

where $\Delta y(k+1) = y(k+1) - y(k)$

Choosing a new state variable vector $x(k) = [\Delta x_m(k)^T \ y(k)^T]^T$, we have:

$$\begin{bmatrix} \Delta x_m(k+1) \\ y(k+1) \end{bmatrix} = \begin{bmatrix} A_m & o_m^T \\ C_m & I_{q \times q} \end{bmatrix} \begin{bmatrix} \Delta x_m(k) \\ y(k) \end{bmatrix} + \begin{bmatrix} B_m \\ C_m B_m \end{bmatrix} \Delta u(k) + \begin{bmatrix} B_d \\ C_m B_d \end{bmatrix} \epsilon(k)$$

$$y(k) = \begin{bmatrix} o_m & I_{q \times q} \end{bmatrix} \begin{bmatrix} \Delta x_m(k) \\ y(k) \end{bmatrix} \quad (18)$$

where $I_{q \times q}$ is the identity matrix with dimensions $q \times q$, which is the number of outputs and o_m is a $q \times n_1$ zero matrix. In (18), A_m , B_m and C_m have dimension $n_1 \times n_1$, $n_1 \times m$ and $q \times n_1$, respectively. For notational simplicity, we denote (18) by

$$\begin{aligned} x(k+1) &= Ax(k) + B\Delta u(k) + B_\epsilon \epsilon(k) \\ y(k) &= Cx(k) \end{aligned} \quad (19)$$

where A , B and C are matrices corresponding to the forms given in (18).

3.2.2 Controllers Tuning

By default, MPC uses a static Kalman filter (KF) to update its controller states, which include the plant model states, disturbance model states, and measurement noise model states. This KF requires two gain matrices. By default, the MPC controller calculates them during initialization. They depend upon the plant, disturbance, and noise model parameters, and assumptions regarding the stochastic noise signals driving the disturbance and noise models. More details about state estimation in MPC are in [6, 9]. However, PID controller used is with a built-in tuning algorithm which by default, chooses a crossover frequency (loop bandwidth) based on the plant dynamics, and designs for a target phase margin to tune the PID gains.

4. Simulation Study

The performance of the Model predictive control (MPC) in controlling the evaporation process is examined, by applying a step response for the three controlled variables L_2 , P_2 and C_2 at $t=100$, $t=200$, $t=300$ respectively, and applying a disturbance step for C_1 and T_{200} at $t=400$ and $t=500$ respectively, the simulation is run for 600 min. Same operation conditions were applied using PI controllers. However, three PI controllers were needed to control the three parameters, therefore, the tuning process were not easy hence each loop is tuned with keeping the other two fixed and repeating this until an acceptable response is achieved. Figures 3 illustrates the performance obtained. Therefore, using a multi-input multi-output (MIMO) MPC is easier and reliable.

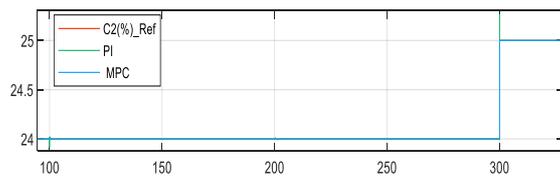
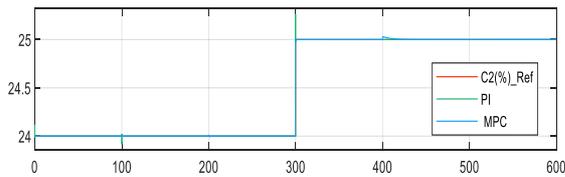
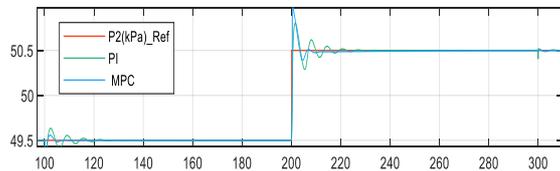
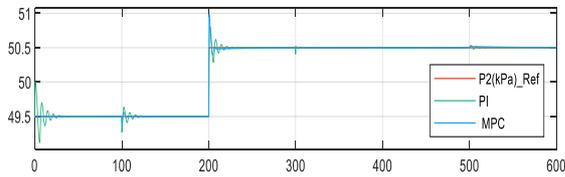
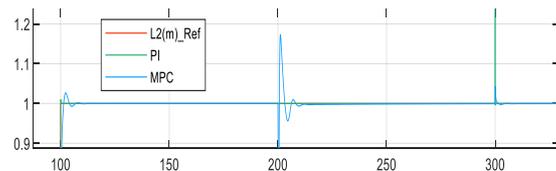
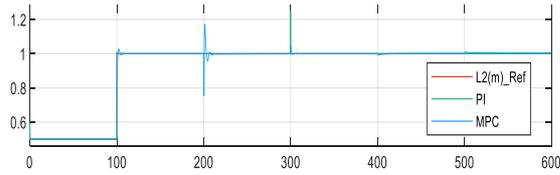


Figure 3: System response using both PI and MPC

Figure 4: enlargement of figure 3

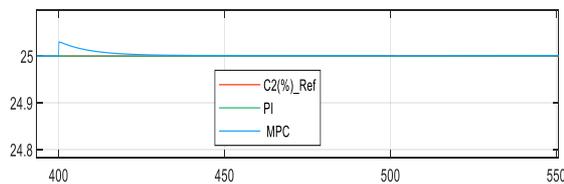
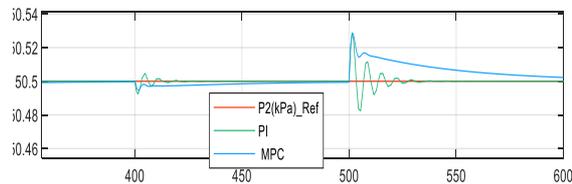
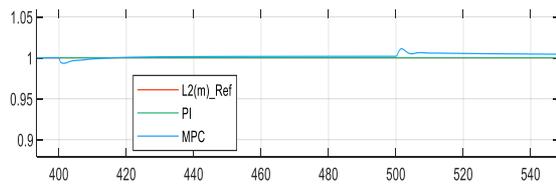


Figure 5: 2nd Enlargement of figure 3

Effect of step change

An Enlargements of the Figure 3, shows the comparison of the responses obtained by PI and MPC controllers. Moreover, examining the figure below, it is seen that using PI the L_2 response has a very small overshoot for the step change of L_2 and with no overshoot at P_2 step change while it has got a quite bigger for C_2 step change with settling time of 0.01 min. In contrast, of the MPC response which has a small overshoot at L_2 step at $t=100$ with a settling time of $t=4$ min, a bigger overshoot at the step of P_2 giving a settling of $t=6$ min, and a very small overshoot at C_2 step

change showing a settling time of $t=1$ min. However, for the P_2 response, the performance of both controllers is quite similar, where, it has a small overshoot at the step change of L_2 , but more oscillatory in PI response, with a settling time of $t=17$ min in case of PI and $t=5$ min in case of MPC. A bigger overshoot at the P_2 step change with settling time of $t=22$ and 10 min for PI and MPC respectively. A very small overshoot for C_2 step change with settling time of $t=5$, and 8 min for PI and MPC respectively. For the response of C_2 both controllers, show a quite negligible effect at the step changes of L_2 , and P_2 , while an obvious overshoot at the step change of C_2 , with a settling time of $t=0.1$ and 0.2 min for PI and MPC.

Effect of disturbances

Figure 5 shows the enlargement of Figure 3 from $t=400 - 500$ with the effect of disturbances C_1 and T_{200} respectively on the three responses. PI controller has a good effluence on rejecting the disturbances as can be seen on L_2 response, whereas a small downshoot and overshoot are apparent in the case of MPC in response to the two disturbances with a settling time of approximately $t=11$ min at C_1 and $t=25$ min at T_{200} disturbance. For P_2 response a very small overshoot for both controllers with a settling time $t= 6$ min at C_1 disturbance for PI, and $t=4$ min for MPC respectively and $t=18$ min for PI and $t=31$ min for MPC at T_{200} disturbance. For C_2 response, the only obvious overshoot is that caused by MPC at C_1 with a settling time $t=22$ min. whereas the PI controller shows a good rejection for the two disturbances.

Table 2: The values of Integral absolute error obtained (IAE)

Controller	IAE			Total
	L_2	P_2	C_2	
PI	0.02	6.27	0.0005	6.2905
MPC	1.358	3.136	0.589	5.0830

5. Conclusion

Following the foregoing results obtained by both controller PI and MPC, it can be insured that the performances of both controllers were very good and they slightly differ in some points according to the structure of each of them. three PI controllers were needed to control the three variables which made it a time consuming as each PI control loop is dealt with by its own in tuning process. Hence, to tune each PI controller the other two will be fixed until a good response is obtained, then the same process is repeated with the next controller by fixing the other two, and keep repeating the tuning process for several times until a fairly good response is achieved by the three controllers. However, MPC are a multiple input multiple output controller, and all variables are controlled at once. Therefore, just one run might be sufficient after adjusting the MPC parameters (Control horizon, and prediction horizon). Moreover, the two controllers show a very good control tracking for the step changes applied at different time of the simulation for the three variables and a quite similar responses were obtained even when introducing the disturbances, despite a negligible difference. It is realised that the responses obtained using PI controller are more oscillatory than those obtained using MPC, whereas the responses obtained using MPC have less settling time in most cases compared to those obtained from PI controller. Therefore, the best comparison can be concluded from the values of the Integral absolute error (IAE) obtained for each controlled variable and the overall IAE. Hence L_2 and C_2 have very small values using PI, but a bigger value for P_2 , while it is reduced up to its half value using MPC, but with a relative L_2 and C_2 values. However, as the optimisation process is based on the sum of IAE values at each iteration, then it is obvious that MPC has better performance on this respect.



References

- [1]. A.S. Yusuff, and A. Giwa. Investigating the Effects of Some Input Variables on the Operation of an Evaporator through Dynamics. International Journal of Advanced Research (2015), Volume 3, Issue 1, 670-683.
- [2]. Kroschwitz, J.I. and Howe-Grant, M. (Eds). 1994, Kirk-Othmer Encyclopedia of Chemical Technology, 4th Edition, John Wiley & Sons, New York, USA, 9: 960-964.
- [3]. R. B. Newell and P. L. Lee. Applied Process Control - A Case study. Prentice Hall, 1989.
- [4]. William L. Luyben, Micheal L. Luyben. Essentials of Process Control. McGraw-Hill 1997.
- [5]. WG da Silva; PP Acarnley; JW Finch; 'Tuning of brushless dc drive speed controller with an on-line genetic algorithm', European Power Electronics Conference, Lausanne, 1999.
- [6]. Liuping Wang; Model predictive Control System Design and Implementation Using MATLAB. Springer, 2009
- [7]. J. K. Huusom, N. K. Poulsen, S. B. Jørgensen, and J. B. Jørgensen. Tuning offset-free model predictive control based on arx models. Journal of Process Control, 22:1997–2007, 2012.
- [8]. J. K. H. Knudsen, J. K. Huusom, and J. B. Jørgensen. Practical implementations of advanced process control for linear processes. In Proceedings of the 18th Nordic Process Control Workshop - NPCW, 2013.
- [9]. Alberto Bemporad, Manfred Morari, N. Lawrence Ricker; Model Predictive Control Toolbox™ Getting Started Guide © COPYRIGHT 2005–2015 by The MathWorks, Inc.

Lammps as Nano-scale Lab to Estimate Fluid Thermal Properties from Molecular Dynamics

Abdusalam Alkhwaji^{1*}, Salem Elbahloul², Mohamed S. Farhat³

¹alkhwaji@asmarya.edu.ly, ²s.elbahloul@eng.misuratau.edu.ly, ³m.farhat@asmarya.edu.ly

¹Mechanical Engineering Department, AL Asmarya Islamic University, Zliten, Libya

²Mechanical Engineering Department, Misurata University, Misurata, Libya

³Chemical Engineering Department, AL Asmarya Islamic University, Zliten, Libya

ABSTRACT

On the one hand the complexity of nano-scale thermofluid system considered as an expensive to investigate experimentally, on the other there are available discrete models such as Molecular Dynamics model offers an affordable and accurate choice for researchers to perform nano-scale investigation to look at different fluid aspects. In this paper we demonstrate how engineers can benefit from the Molecular Dynamics simulation to estimate water thermal properties. Lammps software was used in this study to calculate water thermal properties. The atomistic model type of water which has been used is the four-site transferable intermolecular potential water molecules (TIP4PEW). The thermal properties of water to be estimated are thermal conductivity, viscosity, density, specific isochoric and isobaric heat capacities, Prandtl Number, and Volumetric thermal expansion coefficient. Results compared to experimental data, and showed very good agreement. That prove the usefulness of molecular dynamics simulation as an engineering research tools to investigate thermal properties of any fluid. This study suggests Lammps as an excellent simulation tools to perform computational studies of Thermofluid properties and consider Lammps as a cost affordable nanoscale lab, since it is an opensource software. The accuracy of Lammps depends on the quality of the force field. The results of these study were compared to experimental published water properties at temperatures of 288, 300, 312 and 324 K and pressure 1 atm.

Keywords:

Molecular dynamics;
water thermal properties;
specific isochoric &
isobaric thermal
capacities; dynamic
viscosity, thermal
conductivity; Prandtl
Number; Volumetric
thermal expansion
coefficient; Lammps.

1- Introduction

Literatures published on modelling fluid flow in nanochannels/nanopores at nano-scale level indicate how important is nanoscale studies to both biomedical, thermofluid and energy-related fields. Nano-channels and nanofluids are two examples of the ongoing worldwide nanoscale thermofluid studies which have tremendous improvement in the field of heat exchanger design, solar thermal energy and other thermal management fields [1-4]. Water is another topic which have been studied by many researchers at the nanoscale size. Molecular Dynamics approach is an effective tool to study fluids at nanoscale level. The most commonly used Molecular Dynamics softwares are Lammps, Gromacs, Amber, Charmm and Namd. Each one has some good features such as being an open source software (Lammps and Gromacs). Gromacs software is used commonly for biological studies. Lammps is a great MD-simulation tools for material and fluids. Using GROMACS package, Meguel Angel [5] evaluated shear viscosity of 500 water molecules of four water types, results indicate variant performances of tested models. Pekka Mark [6] performed MD-simulation using CHARMM simulation software, to determine self-diffusion coefficients and radial distribution functions, $g(O-O)$, $g(O-H)$, and $g(H-H)$ for the TIP3P, SPC, and SPC/E water models. The results of this MD-simulation compared versus experimental data, where the SPC/E water model showed best bulk water dynamics and structure. Water thermal conductivity is typically calculated utilizing either with equilibrium molecular dynamics simulations (EMD) or the

nonequilibrium molecular dynamics (NEMD) methods. The EMD method apply the Green-Kubo formula and requires a direct calculation of the heat flux while in the NEMD method a known heat flux is imposed and the temperature gradient is calculated [7]. The goal of the present research is to perform a nanoscale study to evaluate thermal properties of the TIP4PEW water model using Lammmps [8-9], and Moltemplate [10-11]. The water was taken as a case study for using Lammmps to perform a discrete fluid model simulation instead of the macroscopic continuum model. In this paper, thermal conductivity, dynamic viscosity, density, specific isochoric and isobaric heat capacities, Prandtl Number, and volumetric thermal expansion coefficient were determined using EMD method for liquid water under one atmosphere and temperatures of 288, 300, 312 and 324 K. The EMD is the most common used method and it is known as an accurate technique to calculate thermal properties of fluids. The MD-results compared to the experimental published water properties from reference [12] and showed good agreement. This study demonstrates the usefulness of using Lammmps to study many thermofluid phenomena at nanoscale dimension.

2- Basics of Molecular Dynamics

Provided that there is a sufficient number of molecules within the simulated liquid box, the liquid properties can be computed from the discrete particle information by averaging the fluid quantities. The present section discusses molecular-based models. Molecular dynamics is about solving the Newton second law by integration to determine molecules positions, velocity and interaction force [13].

$$m_i \frac{d^2 \vec{r}}{dt^2} = \sum_{j \neq i, j=1}^N \vec{F}_{ij} + \sum_{j_w \neq i, j=1}^{N_w} \vec{F}_{ij_w} + \vec{F}_{sou} \vec{i} \quad (1)$$

The first term of the right-hand side states the molecular force derived by the Lennard-Jones potential, followed by the second term which illustrates the molecular force between particle *i* and other solid nanochannel wall particles denoted as *j* when water molecules interact with the nanochannel wall. The external force represented by the last term in Eq. 1, which can be gravity, or electric force, or pressure. The force can be derived from the derivative of potential energy with respect to atomic distance:

$$\vec{F}_{ij} = - \frac{\partial \varphi_{ij}}{\partial \vec{r}_{ij}} \quad (2)$$

Forces can be computed from a variety of interatomic energy models such pair potentials (e.g. LJ, Morse, Buckingham), multi-body potentials (e.g. EAM, CHARMM, UFF, DREIDING), reactive potentials (e.g. ReaxFF) and others. The LJ potential represented by the function φ_{ij} which is given by:

$$\varphi_{ij} = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (3)$$

Where *r* is the distance between two molecules, σ is a molecular length scale, and ϵ is an interaction strength parameter. Eq. 1 can be integrated using the Verlet algorithm, or the finite difference methods, or velocity Verlet algorithm or any other available methods. Thermostats are algorithms which used to rescale the velocities of molecules to the simulated liquid box temperature, such as the Berendsen thermostat and the Langevin dynamics. The process of preparing Molecular Dynamics simulation starts with building the molecule initial position (*x*, *y*, *z*) of all atoms. The second step is to create the force field. Thus, two steps can be easily generated using Moltemplate

software. Then it might be required to do energy minimization to find an optimal structure of the system. The final step is to calculate the trajectory by solving Newton’s second law Eq.1 for all atoms in the force field at each timestep [13-14].

3- The Fluid Simulation Sitting – (TIP4PEW water model)

The simulation system contains 1000 water molecules in a cubic region as in figure 1. The periodic type of boundary condition was adopted for this Molecular Dynamics simulation. The constant temperature, number of atoms and pressure (NPT) conditions were applied when calculating dynamic viscosities and Isobaric specific thermal capacities. While NVT was applied to determine isochoric specific heat capacities. During calculation of thermal conductivities, the NPT-, NVT- and NVE-ensembles were applied to the water box. Duration of each ensembles was 1 ns (1000000 femtoseconds).

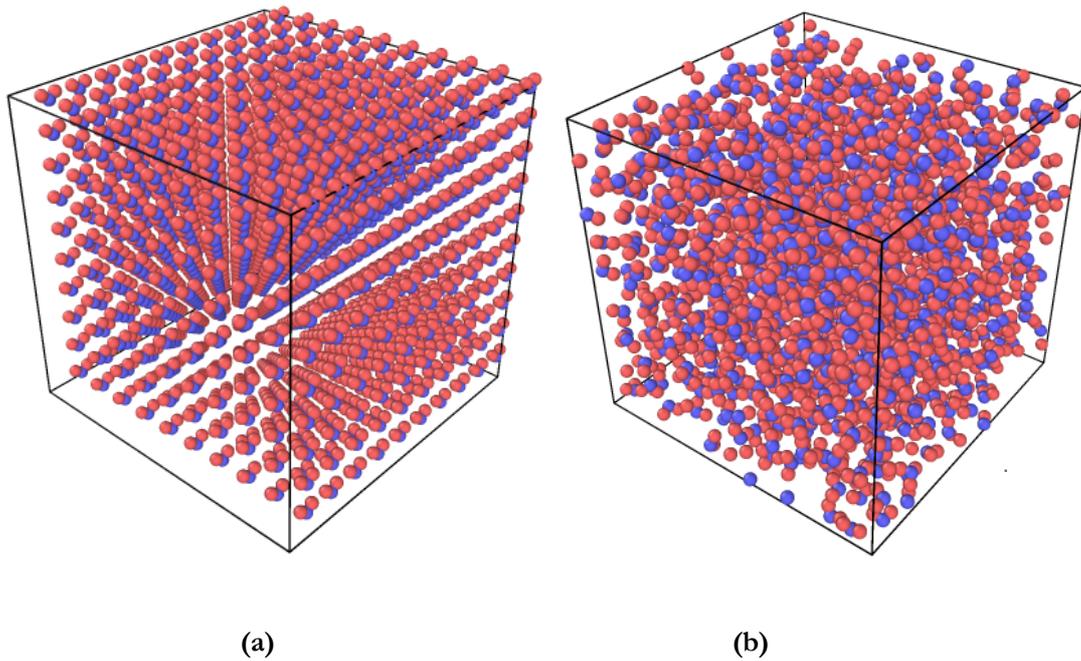


Figure 1: System model initial sitting of 10x10x10 water molecules (a) before and (b) after pressure equilibration (NPT ensemble)

The energy E_{ab} between molecules a and b is represented by the sum of the coulomb and Leonard-Jones potentials centred on the sites, by

$$E_{ab} = \sum_{ij} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{r(ij)} \right] \quad (4)$$

Where r_{ij} is the distance between atoms i and j and q_i and q_j are each site k, the parameters A_{kk} and B_{kk} were given by $A_{kk} = 4\epsilon_k \sigma_k^{12}$ and $B_{kk} = 4\epsilon_k \sigma_k^6$. Where model parameters are given in Table 1.

Table 1: Parameters of TIP4PEW water model [15]

Model	σ	ϵ	q_H	q_O	θ
TIP4-2005	3.16435	0.16275	0.52422	-1.04844	104.52

4- The Green Kubo Formula

The Green Kubo formula is a Molecular dynamic relation relates microscopic fluxes at either NPT, NVT or NVE ensembles, to thermal conductivity or dynamic viscosity. The G-K method uses a correlation function to calculate the transport coefficients during the equilibrium state. [16-17]

$$k = \frac{1}{3Vk_B T^2} \int_0^{\infty} \langle q_i(0)q_i(t) \rangle dt \quad (5)$$

Where k, V, T, and KB, are the thermal conductivity, volume of the simulation box, system kinetic temperature, Boltzmann constant, respectively. q is the per-atom contributions of kinetic, potential and virial from non-bond, bond, and angle interaction.

The heat flux is defined as

$$q = \left[\sum_{j=1}^N v_j E_j - \sum_{\alpha=1}^N H_{\alpha} \sum_{j=1}^{N_{\alpha}} v_{\alpha j} \right] + \frac{1}{2} \left[\sum_{i=1}^N \sum_{j=1, j \neq i}^{N_{\alpha}} r_{ij} (v_j \cdot F_{ij}) \right] \quad (6)$$

Where particle velocity is v_j , total energy per atom is E_j , and the averaged partial enthalpy is H_{α} . Displacement and interatomic forces are r_{ij} and F_{ij} respectively, and N is number of particles. The averaged partial enthalpy is given by;

$$H_{\alpha} = \frac{1}{N_{\alpha}} \sum_{j=1}^{N_{\alpha}} (E_j + r_j \cdot F_j) \quad (7)$$

5- Dynamic Viscosity

Dynamic viscosity is a measure of the momentum flux and it is defined by the gradient of velocity.

$$\tau = -\mu \frac{dv}{dy} \quad (8)$$

where τ is the shear stress, and $\frac{dv}{dy}$ is the gradient of velocity. Dynamic viscosity has units of pressure-time and is calculated using the Green-Kubo formula using the following formula [16-17]

$$\eta_{xy} = \frac{V}{k_B T} \int_0^{\infty} \langle P_{xy}(0)P_{xy}(t) \rangle dt \quad (9)$$

where η is the dynamic viscosity. The stress tensor P is given by

$$P_{xy} = \sum_j m_j v_{xj} v_{yj} + \frac{1}{2} \sum_{i \neq j} r_{xij} F_{yij} \quad (10)$$

Where only the three stress tensor components of off-diagonal P_{xy} , P_{xz} , and P_{yz} are used.

6- Specific Isobaric and Isochoric Heat Capacities

During the NPT ensemble equilibrium, the isobaric specific thermal capacity between temperatures of T1 and T2 and the associated total energies of E_1^{tot} and E_2^{tot} respectively. The quantum contribution of intermolecular vibrational moles to the calculated isobaric specific heat capacity denoted as $\left(\frac{\partial Q}{\partial T}\right)$, which is documented in Table 1 [1], includes the quantum contribution term for temperatures of 288, 300, 312, and 324 K.

$$C_p = \frac{E_2^{tot} - E_1^{tot}}{T_2 - T_1} + \frac{\partial Q}{\partial T} \quad (11)$$

Table 2: Quantum contributions term of equation 11

Temperature (K)	Quantum contributions (KJ/kg-K)
288	-0.4993
298	-0.5166
300	-0.5201
312	-0.5409
324	-0.5617

The isochoric specific heat capacity is given as [18]

$$c_v = \frac{\langle E^2 \rangle - \langle E \rangle^2}{K_B T^2} \quad (12)$$

where at the NVT equilibrium the average energy is $\langle E \rangle$, c_v is the specific isochoric thermal capacity of the entire system, and k_B is Boltzmann constant = $1.3806504 \times 10^{-23}$ J/K.

7- Prandtl Number

The water Prandtl Number describes the ratio of the momentum diffusivity, ν , to the thermal diffusivity α . It is therefore an important thermofluid property. The Prandtl Number indicates the effect of momentum in velocity boundary layer, as well as the effect of heat transfer by diffusion in thermal boundary layer. The Prandtl Number is given by [19].

$$Pr = \frac{c_p \mu}{k} = \frac{\nu}{\alpha} \quad (13)$$

8- Volumetric Thermal Expansion Coefficient

Volumetric thermal expansion coefficient β ($1/^\circ\text{C}$) of water provides a measure of the density changes in response to a change in temperature at constant pressure [19].

$$\beta = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_p = - \left(\frac{\Delta \rho / \rho}{\Delta T} \right)_p = \frac{1}{V(p, T)} \frac{V(p, T + \varepsilon) - V(p, T - \varepsilon)}{2\varepsilon} \quad (14)$$

9- Results and Discussion

Figure 2 shows water thermal conductivity results, Maximum absolute error (the magnitude of the difference between the experiment and the estimated values) recorded was 0.15 kW/m.K at temperature of 300 K, while at the rest of temperatures, errors were less than or equal to 0.1.

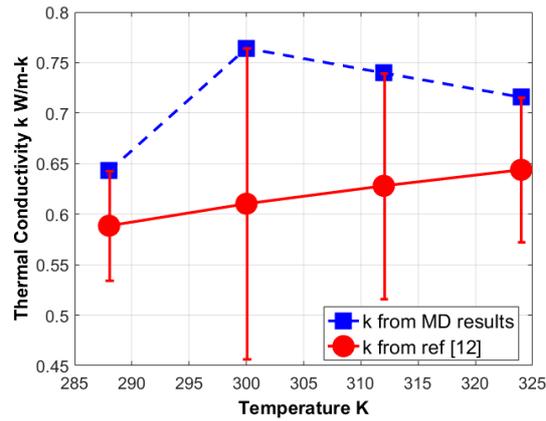


Figure 2: Thermal conductivities from Molecular Dynamics versus experimental values

Figure 3 illustrates dynamic viscosities from Molecular Dynamic simulation versus experimental values and error-bar. The maximum error was 0.00003 Pa. s at temperature of 288 K. At the rest of temperatures, dynamic viscosities have not showed significant errors.

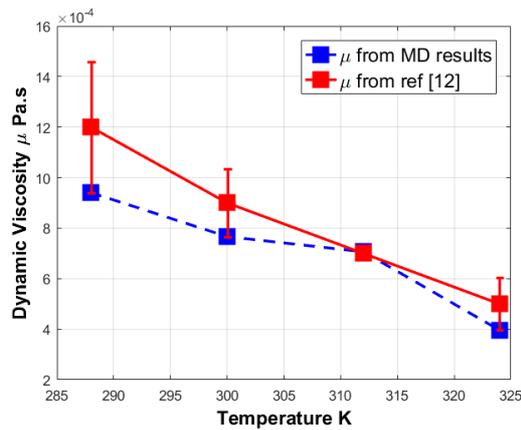


Figure 3: Dynamic viscosities from Molecular Dynamics versus experimental values

Figure 4 shows specific isobaric heat capacities, experimental values at the corresponding temperatures, and the error bars. Maximum error was about 0.53 kJ/kg.K at temperature of 324 K.

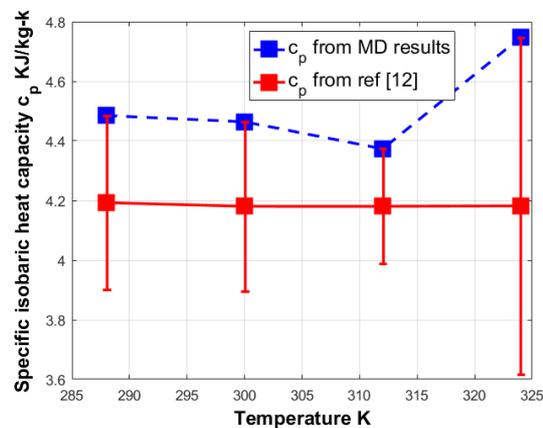


Figure 4: Specific isobaric heat capacities from Molecular Dynamics versus experimental values

The specific isochoric heat capacities in Figure 5 are plotted versus experimental specific isochoric heat capacities and error bars. The maximum error was recorded equal to 0.16 at temperature of 300 K.

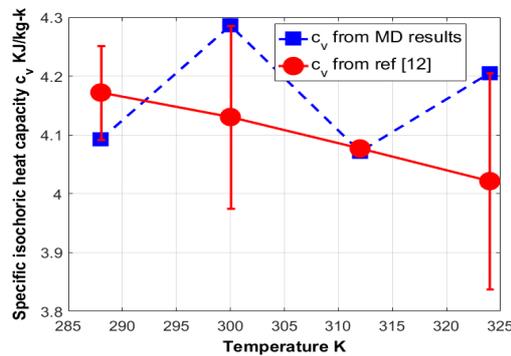


Figure 5: Specific isochoric heat capacities from Molecular Dynamics versus experimental values

In Figure 6, Densities from MD-simulation are underestimated at temperatures of 288, 300, and 324 K. The maximum error was 10 Kg/m³.

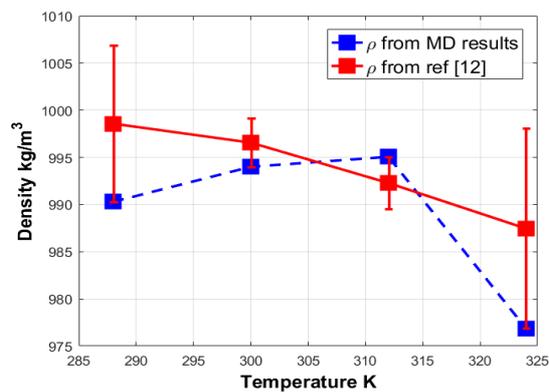


Figure 6: Densities from Molecular Dynamics versus experimental values

The Prandtl number results are plotted in Figure 7 versus experimental Prandtl number for temperatures of 288, 300, 312 and 324 K and pressure of 1 atm. The maximum error between both quantities was 2.

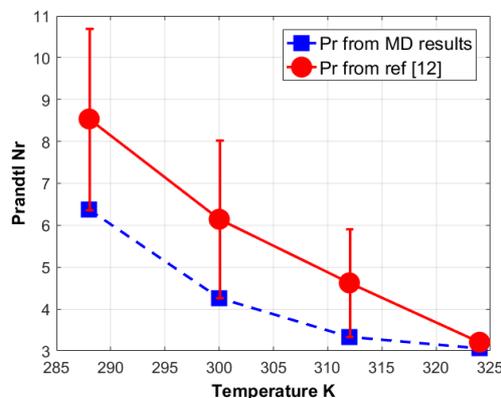


Figure 7: Prandtl numbers from Molecular Dynamics versus experimental values

Figure 8 shows thermal expansivities from MD-simulation versus experimental results. The maximum error between the two quantities was about 0.000125 K⁻¹.

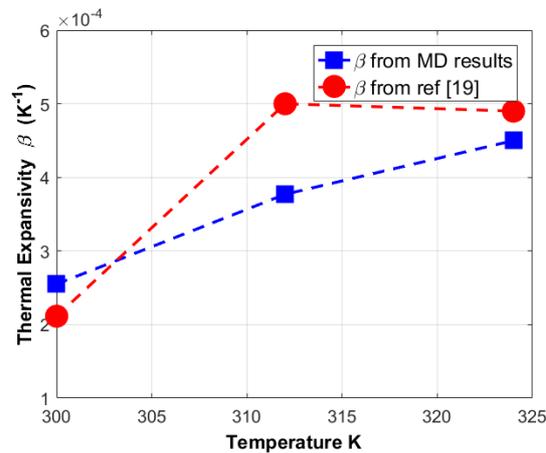


Figure 8: Thermal expansivities from Molecular Dynamics versus experimental values

Figure 9 illustrates absolute relative errors between MD-simulation results and experimental available values at four temperatures, where x-axis indicates all five properties while y-axis is absolute relative errors. The maximum absolute relative error of thermal conductivities was 25 % at temperature of 300 K. For dynamic viscosities, maximum absolute relative error was 23 % at temperature of 288 K. Specific isobaric heat capacities have maximum error at temperature of 324 K equal to 14 %. The maximum absolute relative error of specific isochoric heat capacities was 4.6 % at temperature 324 K. The absolute relative error of densities was negligible.

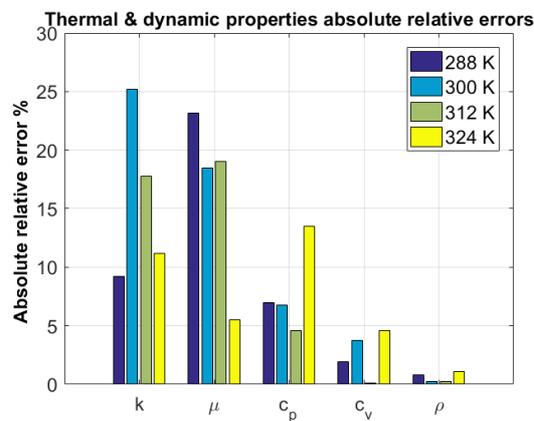


Figure 9: Absolute relative errors between Molecular Dynamics properties and experimental values

Thermal properties of water determined using the TIP4PEW water model and Lammmps, This Molecular Dynamic approach saves the cost associated to an experimental approach.

10- Conclusions

The Molecular Dynamics results provide good argument to support Lammmps as a nano-scale laboratory to investigate fluid thermal properties. In this study, we have used the TIP4PEW water model as an example of how Lammmps can be used to calculate fluid thermal properties. Lammmps software can be used as a Thermofluid lab to study behaviours of most fluids in nano-scale and femtoseconds- or picoseconds-scales to capture more details of interactions between molecules, and avoid approximating most of physical phenomena. From the generated MD-simulation results,

it is clear that thermal properties converged to the experimental properties and that the MD thermal properties were mostly agreed with the published water thermal properties. A main cause of the errors in these Molecular Dynamics calculations is that when integration performed the noise associated with the autocorrelation function changes the physical properties slightly as for the conductivities. To minimize this noise, it is possible to search for an optimal damping factor to reduce noise from the autocorrelation function.

References

- [1]. Abdusalam Al Khwaji, S. Elbahloul, Khairil F. B. Abu Bakar, M. Zulkifly Abdullah, “The Comparison Between Water Models in Predicting Water Thermal And Dynamic Properties From Molecular Dynamics”, International Journal of Scientific & Technology
- [2]. Chuanhua Duan, Wei Wang, and Quan Xie “Review article: Fabrication of nanofluidic devices “, *Biomicrofluidics* 7, 026501 (2013); doi: 10.1063/1.4794973.
- [3]. Mahian, O., Kianifar, A., Kalogirou, S. A., Pop, I., & Wongwises, S. (2013). A review of the applications of nanofluids in solar energy. *International Journal of Heat and Mass Transfer*, 57(2), 582–594. doi: 10.1016/j.ijheatmasstransfer.2012.10.037
- [4]. Huminic, G., & Huminic, A. (2012). Application of nanofluids in heat exchangers: A review. *Renewable and Sustainable Energy Reviews*, 16(8), 5625–5638. doi: 10.1016/j.rser.2012.05.023
- [5]. Miguel Angel González and José L. F. Abascal, (2010) “The shear viscosity of rigid water models”, *The journal of chemical physics* 132, 096101
- [6]. Pekka Mark and Lennart Nilsson, “Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K”, *J. Phys. Chem. A* 2001, 105, 9954-9960
- [7]. Yijin Mao, Yuwen Zhang, “Thermal conductivity, shear viscosity and specific heat of rigid water models”, 2012 Elsevier, *Chemical Physics Letters* 542 (2012) 37–41.
- [8]. LAMMPS molecular dynamics simulator; [accessed 2019 June 10]. Available from: <http://lammps.sandia.gov> 2019.
- [9]. LAMMPS Documentation, 30 Apr 2019 version, [<https://lammps.sandia.gov/doc/Manual.html>]
- [10]. Moltemplate molecular dynamics simulator; [accessed 2019 June 10]. Available from: <https://www.moltemplate.org> 2019.
- [11]. Moltemplate Documentation, 30 Apr 2019 version, [https://www.moltemplate.org/doc/moltemplate_manual.pdf]
- [12]. W. M. Haynes, ‘Liquid Properties’, *Handbook of Chemistry and Physics*, 92nd edition, Taylor & Francis, New York (2012)
- [13]. Li, Y., Xu, J. & Li, D. Molecular dynamics simulation of nanoscale liquid flows. *Microfluid Nanofluid* 9, 1011–1031 (2010). <https://doi.org/10.1007/s10404-010-0612-5>
- [14]. Gad-el-Hak Mohamed, “The Fluid Mechanics of Microdevices – The Freema Scholar Lecture”, 1999. *Journal of Fluids Engineering—ASME*
- [15]. [http://www1.lsbu.ac.uk/water/water_models.html]
- [16]. William L. Jorgensen, Jayaraman Chandrasekhar, Jeffrey D. Madura, Roger W. Impey, and Michael L. Klein, “Comparison of simple potential functions for simulating liquid water”, *The Journal of Chemical Physics*.
- [17]. H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, “The missing term in effective pair potentials”, *Journal of Physical Chemistry* 91 (1987) 6269-6271.)
- [18]. M.C. Gupta, “Statistical thermodynamics”, (1993), New York: John Wiley & Sons
- [19]. Theodore L. Bergman, Adrienne S. Lavine, Frank P. Incropera, David P. Dewitt, “Fundamentals of Heat and Mass Transfer”, 7th edition, John Wiley & Sons, Inc.
- [20]. Timothy W. Sirk, Stan Moore, Eugene F. Brown, “Characteristics of thermal conductivity in classical water models”, (2013) *The Journal of Chemical Physics* 138, 064505.
- [21]. Y. Wu, H. L. Tepper and G. A. Voth, Flexible simple point-charge water model with improved liquid state properties, *Journal of Chemical Physics*, 124 (2006) 024503.
- [22]. Yunus A. Cengel, Afshin J. Ghajar, “Heat and Mass Transfer Fundamentals & Applications”, Fourth edition, McGraw-Hill, New York, (2011)



-
- [23]. E M Kirova and G E Norman, "Viscosity calculations at molecular dynamics simulations", Journal of Physics : Conference Series 653 (2015) 012106
 - [24]. Haibo Yu, Tomas Hansson, and Wilfred F. van Gunsteren, "Development of a simple, self-consistent polarizable model for liquid water", Journal of Chemical Physics, 118 (2003).
 - [25]. Simon Stephan, Jens Staubach, Hans Hasse, Review and Comparison of Equations of State for the Lennard-Jones Fluid, Fluid Phase Equilibria (2020), doi: <https://doi.org/10.1016/j.fluid.2020.112772>

Application of the TDS Technique in Gas Reservoirs

Laila D. Saleh^{1*}, Muhanad Ikaybah², Ahmed Moftah Abulqasem³, Hayat K. Alhaj⁴

¹ladasa08@gmail.com, ²Mohaned123tkd@gmail.com, ³ahmedaboalqasem0@gmail.com,

⁴hayatkhalifa@yahoo.com

^{1,2,3,4} Department of Petroleum Engineering, University of Tripoli, Libya

ABSTRACT

Keywords:

TDS technique,
Pressure,
Derivative,
Wellbore storage.

Tiab's Direct Synthesis (TDS) is a direct technique to interpret transit well pressure tests. It is analysis of pressure and pressure derivative without type-curve matching. This method uses log-log Plot of the pressure and pressure derivative versus time to compute reservoir parameters such as permeability, wellbore storage, skin factor, and average reservoir pressure.

The main objective of this study is to apply TDS technique for gas reservoirs "long test", and show the advantage of TDS technique, where this technique is particularly useful when the late-time infinite acting radial flow is not observed, which is called "short test".

Two cases are presented in the study to analyze pressure data using TDS technique, a build-up test is conducted on the well (4U11) from Faregh field and well (5A5) from Gialo field.

Based on the results and comparing between the long and short test; the results were converged, which indicates that TDS can be used if some flow regimes are not available. The technique is useful method for analyzing pressure test.

1- Introduction

One of problem in pressure derivative on log-log plot that the radial flow region is not observed, TDS solved this issue by analyzing well bore storage region and read direct specific values from the plot substituting in equations to determinate parameters that cannot be estimated from radial flow such like permeability, and skin this advantage is called (Short test) [1] . Another problem faced in this paper, the huge raw data from DST test, which takes reading for pressure every ten seconds, solving this problem comes with filtering or screening the best build up region to be fit and clear enough for TDS calculations.

The objective of this paper is to analyze pressure transient in gas reservoirs using Tiab's direct synthesis method and use advantage of this technique to determine reservoir properties

2- Methodology

The pressure test data was collected and screened by using Kappa software. Then the pressure data was analyzed by TDS technique using Excel software.

2.1 TDS technique

Direct synthesis method uses a log-log plot of pressure and pressure derivative data versus time to calculate various reservoirs and well parameters. The technique uses the pressure derivative technique to identify reservoir heterogeneities. In this method, the values of the slopes, intersection

points, and beginning and ending times of various straight lines from the log-log plot can be used in exact analytical equations to calculate different parameters [2, 3].

TDS technique procedure

In this paper, TDS technique uses to analyze pressure test data in two cases as following;

- **The wellbore storage and infinite acting radial flow is observed (Long test):**

The following steps are presented with greater detail in Tiab [1, 3].

Step 1- Converting pressure data (P) to real gas pseudo pressure data $m(P)$ to linearize the gas diffusivity equation.

$$m(P)_i = \left(\frac{2P}{\mu Z}\right) i^* + m(P)_i \quad (1)$$

Step 2- Calculate $\Delta m(P)$ and Δt using the following equation:

$$\Delta m(P) = m(P)_i - m(P) \quad (2)$$

Step 3- Calculate Derivative of each values using equation:

$$t \times \Delta m(P)' = \Delta t \times \left(\frac{\Delta m(P)_{i+1} - \Delta m(P)_{i-1}}{(\Delta t)_{i+1} - (\Delta t)_{i-1}} \right) \quad (3)$$

Step 4- Plot $\Delta m(P)$ and $t \times \Delta m(P)'$ versus time in log-log graph.

Step 5- Draw the unit-slope line corresponding to the wellbore storage flow regime using early-time pseudo pressure and derivative points.

$$C = \left(\frac{0.42 q_{sc} T}{\mu i} \right) \left(\frac{t_N}{t \times \Delta m(P)_{N'}} \right) \quad (4)$$

Step 6- Draw the infinite acting radial flow line using late-time pseudo pressure derivative points.

This line is horizontal.

Step 7- Read the coordinates of the infinite-acting horizontal line intersect this value is $(t \times \Delta m'(P))$.

Step 8- Calculate permeability (K):

$$K = \frac{711.26 q_{sc} T}{h(t \times \Delta m'(P))} \quad (5)$$

Step 9- Calculate the total skin factor (S'): The skin factor is determined from a relationship between the pseudo pressure and its derivative during the infinite acting radial flow.

$$S' = S + Dq_{sc} = 0.5 \left[\frac{(\Delta m(P))_R}{(t \times \Delta m'(P))_R} - \ln \left(\frac{K t_R}{\phi (\mu C_t) i (r_w^2)} \right) \right] \quad (6)$$

Step 10- Calculate the average reservoir pressure:

First, calculate the average pseudo pressure using equation

$$m(\bar{P}) = m(p_i) - 2 \left((t \times \Delta m'(p))_r \left[\left(\frac{(t \times \Delta m'(p))_{pss}}{(\Delta m(p))_{pss} - (t \times \Delta m'(p))_{pss}} \right) \left(\ln \frac{r_e}{r_w} - \frac{3}{4} + S' \right) \right] \right) \quad (7)$$

Then using the equation from the relation between pressure (P) and pseudo pressure $m(P)$ to calculate average reservoir pressure.

- **The infinite acting radial flow is not observed (Short test).**

The following procedures are applied if the pressure test is too short to observe the infinity acting radial flow or is not well defined. Therefore, it cannot read the value of $(t \times \Delta m'(P))_R$; because the horizontal line of the infinite acting radial flow is not observed.

Step 1- Plot $\Delta m(P)$ and $t \times \Delta m'(P)$ versus time on a log-log graph .

Step 2- Read the value of t_x , and $t \times \Delta m'(P)_x$ from the log.log plot as shown in case study.

Step 3- Calculate $(t \times \Delta m'(P))_R$

$$(t \times \Delta m'(P))_R = \left(0.151 \frac{qT}{\mu C}\right) t_x - (t \times \Delta m'(P))_x \quad (8)$$

The value of $(t \times \Delta m'(P))_R$, can be used to estimate the permeability if the infinite acting line is not available(short test) as shown in the case study .

Step 4- Calculate permeability (K) by using equation 5.

Step 5- Calculate total skin factor(S'): The skin factor is estimated from:

$$S' = 0.171 \left(\frac{t_x}{t_{RiUS}}\right)^{1.24} - 0.5 \ln \left(\frac{0.8935C}{\phi h C_t r_w^2}\right) \quad (9)$$

3- Case Study

This paper presents two Libyan gas wells from different fields. These wells are well (4U11) in Gialo field and well (5A5) in Faregh field.

3.1 Case 1- Well (4U11)

Relevant information concerning well 4U11 reservoir and fluid is given in Table 1.

Table 1: General information of Well (4U11)

Field:	Gialo
Well:	4U11.
Well Radius(rw):	0.34 ft.
Drainage Area Radius(re):	745 ft.
Net Thickness(h):	31 ft.
Initial Pressure (Pi):	5250 Pisa.
Reservoir Temperature:	263°F.
Total Compressibility	1.074×10^{-4} psi-1
Gas viscosity(μ_g)	0.028 cp.
Specific gravity (γ_g)	0.69
Gas flow rate (qg)	15.807MMscf/D
Porosity (ϕ)	18

Step 1. Convert the pressure data (P) to m(P) using the equation 1, as shown in table 1.

Table 2: Calculation of $m(p)$ for Well (4U11).

Pressure	Gas Z Factor	gas Viscosity	$2p/z*\mu$	$(2P/\mu z)m$	$(2P/\mu z)m*\Delta p$	$m(p), \text{psi}^2/\text{cp}$
Pisa		cP	Psi/cp	Psi/cp	psi^2/cp	psi^2/cp
1000	0.911	0.015	146327.2	73163.59	7.32E+07	7.32E+07
1300	0.912	0.016	178101.7	162214.4	4.87E+07	1.22E+08
1700	0.915	0.017	218531.5	198316.6	7.93E+07	2.01E+08
2100	0.921	0.018	253430.4	235980.9	9.44E+07	2.96E+08
2600	0.924	0.020	281294	267362.2	1.34E+08	4.29E+08
3100	0.934	0.021	316202.3	298748.1	1.49E+08	5.79E+08
3700	0.939	0.024	328538.4	322370.4	1.93E+08	7.72E+08
4700	0.991	0.030	316242.8	322390.6	3.22E+08	1.09E+09
5228	1.036	0.034	296871.7	306557.2	1.62E+08	1.26E+09

Figure 1 is a linear plot of $m(P)$ versus P . The equation of trend line is

$$y = 38.876x^2 + 150671x - 1E+08 \quad (10)$$

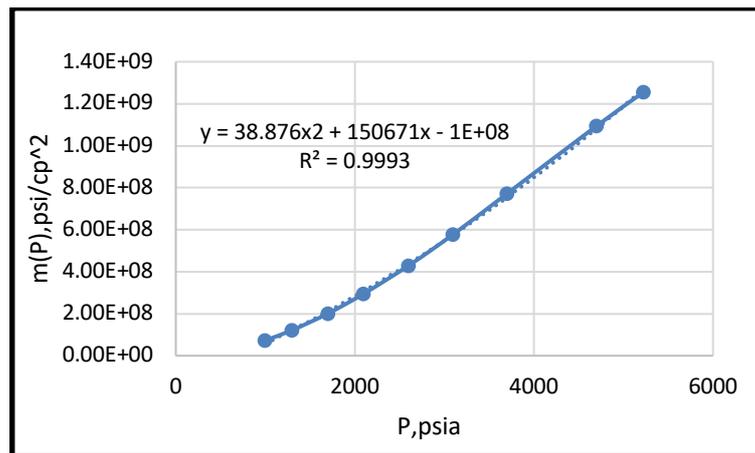


Figure 1: Linear plot of P and $m(P)$ for Well (4U11)

By using equation 10, calculate the initial reservoir pressure:

$$y = (38.876 \times 52502) + (150671 \times 5250) - (1E+08) = 1.76 \times 10^9 \text{ psi}^2/\text{cp}$$

$$\text{Then, } m(P_i) = 1.76 \times 10^9 \text{ psi}^2/\text{cp}$$

After convert all pressure to pseudo pressure, calculate the $\Delta m(P)$ using equation 2 and Δt . Then calculate the derivative of pseudo pressure by using equation 7.

3.1.1 Infinite acting radial flow line is observed (Long test) for well 4U11:

Draw the unit-slope line corresponding to the well bore storage flow regime as shown in figure 2, then calculate the well bore storage effect (C) by using equation 4.

$$C = \left(\frac{0.42 \times 15.807 \times 10^3 \times (263 + 460)}{0.028} \right) \left(\frac{0.05}{2.85 \times 10^8} \right) = 0.03003 \text{ bbl/psi}$$

From a log-log plot of pseudo pressure versus time, figure 2, the infinite acting radial flow is observed at $(t \times \Delta m'(P))_r = 4 \times 10^6 \text{ Psi}^2/\text{cp}$ (the coordinates of the infinite-acting horizontal line intersect). The average permeability is computed by using equation 5.

$$K = \frac{711.26 \times 15.807 \times 10^3 \times (263 + 460)}{31 \times 4 \times 10^6} = 6.55 \text{ md.}$$

From figure 2, read the values of t_R , and $\Delta m(P)$ and calculate the total skin factor (S') by using the equation 6

$$S' = 0.5 \left[\frac{1.72 \times 10^9}{4 \times 10^6} - \ln \left(\frac{5.66 \times 1.666667}{0.18 \times 0.028 \times 1.074 \times 10^{-4} \times 0.34^2} \right) \right] = 15.64$$

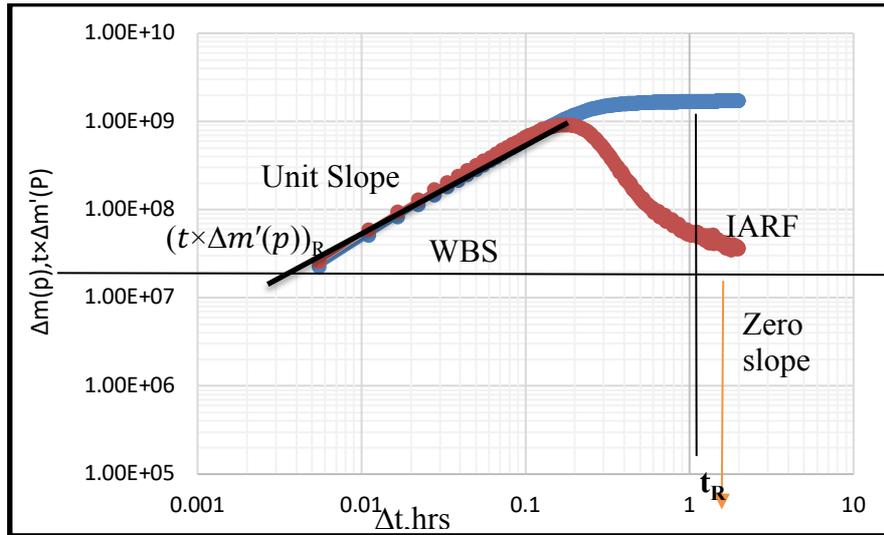


Figure 2: Analysis of infinite acting radial flow for Well (4U11).

3.1.2 Infinite acting radial flow line is not observed (short test) for Well (4U11):

In this case, the infinity acting radial flow line was cut to apply the analysis (short test) when infinite acting radial flow is not observed as shown in figure 3. From figure 3 read the value of t_x , and $(t \times \Delta m'(P))_x$ and calculate $(t \times \Delta m'(P))_R$ using equation 8.

$$(t \times \Delta m'(P))_R = \left(0.151 \frac{15.807 \times 10^3 \times (263 + 460)}{0.028 \times 0.03003} \right) \times 0.3 - 5.01 \times 10^8 = 3.81 \times 10^7 \text{ psi}^2/\text{cp.}$$

Calculate the permeability from equation 5.

$$K = \frac{711.26 \times 15.807 \times 10^3 \times (263 + 460)}{31 \times 3.81 \times 10^6} = 6.88 \text{ md.}$$

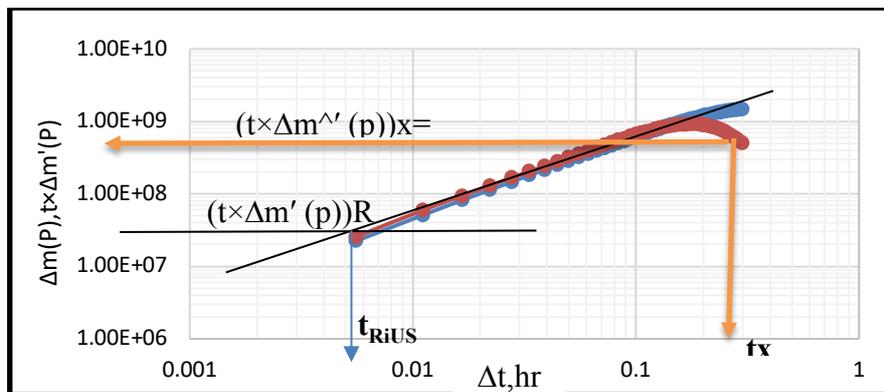


Figure 3: Short test for Well (4U11).

Read the value of t_{RiUS} ($t_{RiUS}=0.0065$ hrs.) from figure3. , then calculate total skin factor using equation 9.

$$S' = 0.171 \left(\frac{0.3}{0.0065} \right)^{1.24} - 0.5 \ln \left(\frac{0.8935 \times 0.03003}{0.18 \times 31 \times 1.074 \times 10^{-4} \times 0.34^2} \right) = 16.75.$$

3.2 Case 2- Well (5A5)

Relevant information concerning well 5A5 reservoir and fluid is given in Table 3.

Table 3: General information of Well (5A5)

Field:	Faregh.
Well:	5A5.
Well Radius(r_w):	0.34 ft.
Drainage Area Radius(r_e):	1500 ft.
Net Thickness(h):	33.5 ft.
Initial Pressure (P_i):	5100 psi.
Reservoir Temperature:	240 F.
Total Compressibility	1.13×10^{-4} psi ⁻¹ .
Gas viscosity(μ_g)	0.0282 cp.
Specific gravity (γ_g)	0.6685
Gas flow rate (q_g)	24.05 MMscf/D.
Porosity (ϕ)	13.9.

Step 1. Convert the pressure data (P) to $m(P)$ using the equation 1, the part of calculation is shown in table 4.

Table 4: $m(p)$ for Well (5A5).

P,psi	$m(P)$,psi ² /cp
1014.7	7.49E+07
1064.7	8.23E+07
1114.7	9.01E+07
1164.7	9.82E+07
1214.7	1.07E+08
1264.7	1.15E+08
1314.7	1.24E+08
1364.7	1.34E+08
1414.7	1.43E+08
1464.7	1.53E+08
1514.7	1.64E+08
1564.7	1.74E+08
1614.7	1.85E+08
1664.7	1.96E+08

Figure 4 is a linear plot of $m(P)$ versus P . The equation of trend line is

$$y = 37.513x^2 + 99041x - 7E+07 \quad (11)$$

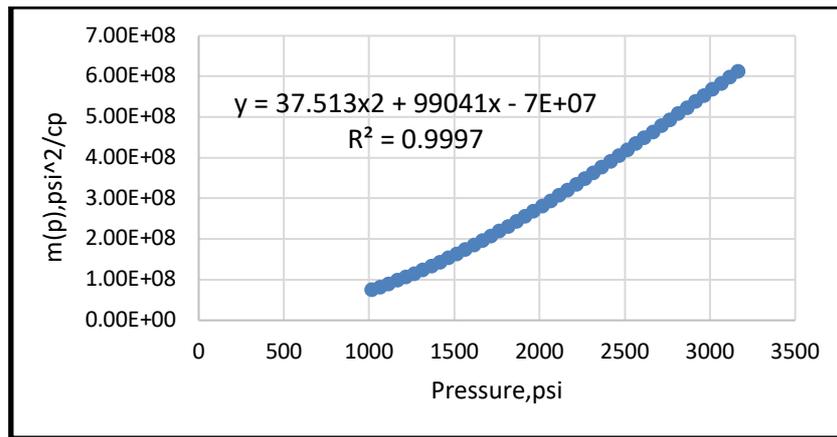


Figure 4: Linear plot of P and m(P) for Well (5A5)

By using equation 10, calculate the initial reservoir pressure:

$$y = (37.513 \times 51002) + (99041 \times 5250) - (7E+07) = 1.411 \times 10^9 \text{ psi}^2/\text{cp}$$

$$m(P_i) = 1.411 \times 10^9 \text{ psi}^2/\text{cp}$$

After convert all pressure to pseudo pressure, calculate the $\Delta m(P)$ using equation 2 and Δt . Then calculate the derivative of pseudo pressure by using equation 7.

3.2.1 Infinite acting radial flow line is observed (Long test) for well 5A5:

After screening of (54143) data point, became (257) data point for analysis.

Draw the unit-slope line corresponding to the well bore storage flow regime as shown in figure 5, then calculate the well bore storage effect (C) by using equation 4.

$$C = \left(\frac{0.42 \times 24.05 \times 10^3 \times (240 + 460)}{0.0282} \right) \left(\frac{0.025}{1.557 \times 10^8} \right) = 0.0402 \text{ bbl/psi}$$

From a log-log plot of pseudo pressure versus time, figure 5, the infinite acting radial flow is observed at $(t \times \Delta m'(P))_r = 2.68 \times 10^7 \text{ Psi}^2/\text{cp}$.

(the coordinates of the infinite-acting horizontal line intersect) . The average permeability is computed by using equation 5.

$$K = \frac{711.26 \times 24.05 \times 10^3 \times (240 + 460)}{33.5 \times 2.68 \times 10^7} = 13.33 \text{ md.}$$

From figure 5, read the values of t_R , and $\Delta m(P)$ and calculate the total skin factor (S') by using the equation 6

$$S' = 0.5 \left[\frac{4.08 \times 10^8}{2.68 \times 10^7} - \ln \left(\frac{13.33 \times 0.35}{0.139 \times 0.0282 \times 1.13 \times 10^{-4} \times 0.34^2} \right) \right] = 2.15$$

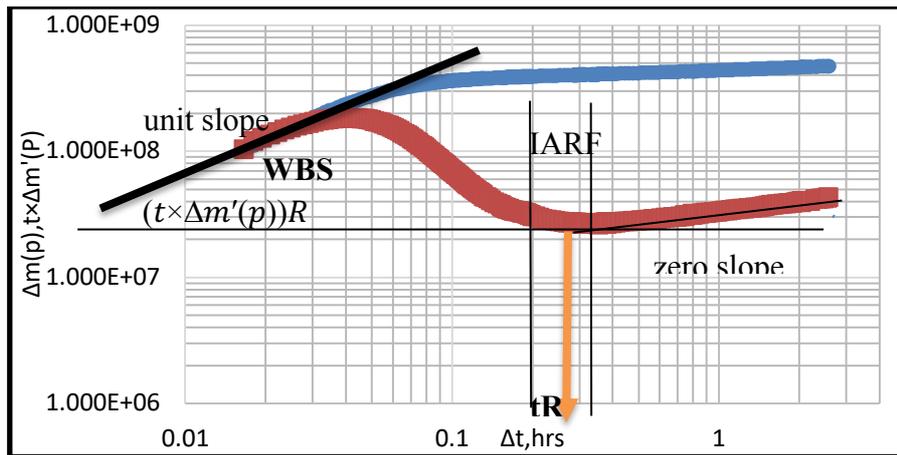


Figure 5: Analysis of infinite acting radial flow for Well (5A5).

3.2.2 Infinite acting radial flow line is not observed (short test) for Well (5A5):

In this case, the infinity acting radial flow line was cut to apply the analysis (short test) when infinite acting radial flow is not observed as shown in figure 6. From figure 6 read the value of t_x , and $(t \times \Delta m'(P))_x$ and calculate $(t \times \Delta m'(P))_R$ using equation 8.

$$(t \times \Delta m'(P))_R = \left(0.151 \frac{24.05 \times 10^3 \times (240 + 460)}{0.0282 \times 0.0402} \right) \times 0.069 - 1.29 \times 10^8$$

$$= 2.67 \times 10^7 \text{ psi}^2/\text{cp.}$$

Calculate the permeability from equation 5.

$$K = \frac{711.26 \times 15.807 \times 10^3 \times (263 + 460)}{31 \times 2.67 \times 10^7} = 6.88 \text{ md.}$$

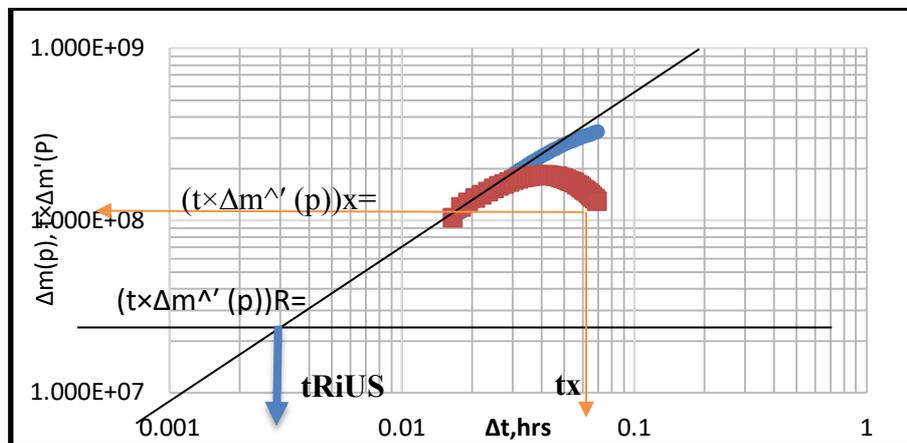


Figure 6: Short test for Well (5A5).

Read the value of t_{RiUS} ($t_{RiUS} = 0.043$ hrs) from figure 3, then calculate total skin factor using equation 9.

$$S' = 0.171 \left(\frac{0.069}{0.043} \right)^{1.24} - 0.5 \ln \left(\frac{0.8935 \times 0.0402}{0.139 \times 33.5 \times 1.13 \times 10^{-4} \times 0.34^2} \right) = 2.15$$

3.3 Summary of the Results:

Table 5 shows the comparing of the main results between the long test and short test in Wells.

Table 5: Comparing the results of Well (4U11) and (5A5).

Parameter	Well 4U11	
	Long test	Short test
$(t \times \Delta m' (p)) R_p \text{psi}^2 / cp$	2.68×10^7	3.67×10^7
K, md	13.33	13.35
Total skin(S^2)	2.15	2.15
	Well 5A5	
	Long test	Short test
$(t \times \Delta m' (p)) R_p \text{psi}^2 / cp$	4×10^7	3.81×10^7
K, md	6.55	6.88
Total skin(S^2)	15.64	16.75

4- Conclusions

Based on the results, TDS technique was shown to be useful, accurate and effective. The technique is particularly useful when the infinite acting radial flow or wellbore storage line have not been observed. The TDS method showed accurate results compared to conventional software matching, the values of permeability by using (Long test) was 6.55 md in Well (4U11), 13.33 md in Well (5A5), where the values of permeability by using (Short test) was 6.88 md in Well (4U11), 13.35 md in Well (5A5), and the value of total skin factor using (Long test) was 15.64 in Well (4U11), 2.15 in Well (5A5), where the value of total skin factor using (Short test) was 16.75 in Well (4U11), 2.15 in Well (5A5). The TDS is an effective method for calculating the average reservoir pressure from well test data for a vertical gas wells, where the average reservoir pressure was (6656) psi in Well (4U11) and (4991) psi in Well (5A5).

References

- [1]. Tiab, D: "Analysis of Pressure and Pressure Derivative without Type-Curve Matching — Skin and Wellbore Storage," Journal of Petroleum Science and Engineering (1995) 12, 171 _ Originally, paper SPE 25423 presented at the 1993 SPE Production Operations Symposium, Oklahoma City, OK, Mar. 21-23.
- [2]. Tarek Ahmed, *Advance Well Testing Analysis*. Introductory well testing 1st edition 2013 Tom Aage Jelmert & bookboon.com ISBN.
- [3]. Tiab, D: "Analysis of Pressure and Pressure Derivative without Matching: Vertically Fractured Wells in Closed Systems," Journal of Petroleum Science and Engineering (1994) 11, 323. Originally, paper SPE 12345 presented at the 1990 SPE Annual Technical Conference and Exhibition, New Orleans, IA, Sept. 23-26.

Transient Pressure Analysis and Productivity Index Estimation in Horizontal Wells

Haiat K. Alhaj^{1*}, Laila D. Saleh², Rabia Hunky³

¹ hayatkhalifa@yahoo.com, ² ladasa08@gmail.com, ³ rabmoh2005@yahoo.com

^{1,2,3} Department of Petroleum Engineering, University of Tripoli, Libya

ABSTRACT

The Analysis of pressure build-up tests in horizontal wells are known as complicated due to changing of flow regimes, formation thickness, well horizontal length....etc. The main objective of study is presents an interpretation method for horizontal well pressure transient testing that is applied to a buildup test from a horizontal well The use of transient well testing for determining reservoir parameters and productivity of horizontal wells has become common because of the upsurge in horizontal drilling. Initially, horizontal well tests were analyzed with the conventional techniques. During the last decade, analytic solutions have been presented for the pressure behavior of horizontal wells. New flow regimes have been identified, and simple equations and flow regime existence criteria have been presented for them [1]. The flow regimes are now used frequently to estimate horizontal and vertical permeability of the reservoir, wellbore skin, and reservoir pressure. Where result of The Giger-Reiss-Jourdan and Joshi was considered more representative result as compared with actual the productivity index and flow rate for isotropic and anisotropy reservoir. One objective of this work is to recall the proper way to use these formulae and to recall the assumptions made that may limit their use.

Keywords:

Productivity,
horizontal wells,
Pressure,
test,
flow

1- Introduction

The technology of drilling and production of horizontal wells has been recognized as one of the most important technical achievements in the oil and gas industry in the last twenty five years. The industry demand of horizontal drilling technology has produced a variety of new applications and techniques.

During the period of 1980 to1984, only one or two horizontal wells were drilled worldwide. In 1988 that number of horizontal wells jumped to over 200 wells[1]. Since, a gradual increase in wells has been noticed, with 1570 wells drilled in 1994. Industry projections in the year 2000 over 5000 wells were drilled horizontally.

A typical horizontal well project is different from a vertical well project because productivity of a well depends upon the well length. Moreover, well length depends upon the drilling technique that is used to drill the well. Therefore, it is essential that reservoir and drilling engineers work together to choose the appropriate drilling technique, which will give the desired horizontal well length.

The other important consideration is the well completion scheme. One can either have an open hole, insert a slotted liner, insert a liner with external casing packers, or case the hole and perforate the casing, depending upon local completion needs and experience. The type of completion affects horizontal well performance and certain types of completions are possible with certain types of drilling techniques and in certain formations [2].

Well length, the well's physical location in the reservoir, the tolerance in drilling location, and the type of completion that can be achieved strongly affects well performance[3]. Therefore, it is very

important for reservoir engineers to understand different drilling and completion techniques and their advantages and disadvantages.

Due to this fact, we present an overview of horizontal well technology. This includes the advantages and disadvantages of horizontal wells, the suitable environment to drill horizontal wells (applications of horizontal wells), and the drilling and completion techniques.

2- Materials and Methods

This Pressure Transient Analysis in Horizontal Wells

The dramatic increase in horizontal drilling activity has made the use of transient well testing common practice in determining the productivity of horizontal wells. In the past, horizontal wells were analyzed using the techniques which had been developed for vertical wells. Over the last years however, new solutions have been presented for horizontal wells. Transient pressure analysis of horizontal wells is considerably more complicated than it is for vertical wells, due to new flow regimes identified in horizontal wells. Identification of the flow regimes is necessary for proper estimation of horizontal and vertical permeabilities of the reservoir, and of wellbore skin.

Horizontal wells pose two special problems for the reservoir engineer. The most obvious is the large wellbore storage effect associated with horizontal sections which may be thousands of feet in length. Wellbore storage effects are pressure effects related to the volume of fluids in the wellbore before the test begins. This potential problem can be overcome by downhole shut-in and downhole flow measurements. The second problem is the more complex nature of the transient and the existence of overlapping flow regimes.

Before discussing the analysis procedure, it is appropriate to state the goals of the well test analysis. In general, a well test analysis of a horizontal well is conducted to achieve the following objectives:

1. To obtain reservoir permeabilities, (k_x , k_y , k_z),
2. To determine whether all the drilled length of a horizontal well is also a producing length,
3. To estimate mechanical skin factor (S_m) or damage related to drilling and completion of the horizontal well. Based upon magnitude of the damage a decision regarding well stimulation can be made.
4. Horizontal well performance (Productivity index).

Because of the 3D nature of flow geometry geometry complications of horizontal wells many authors presented analytical solutions for the pressure response in horizontal well test.

These methods resulted from solving the three dimensional Diffusivity Equation by different assumption at the wellbore condition or at the boundary. The following section describes the suggested equations which considering the physical model shown in Figure (1).

Theory and Calculation

A Theory section should extend, not repeat, the background to the article already dealt with in the Introduction and lay the foundation for further work. In contrast, a Calculation section represents a practical development from a theoretical basis.

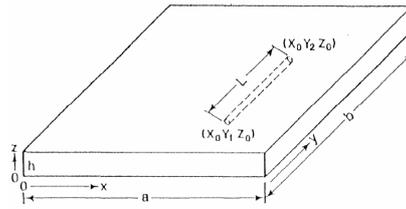


Figure 1: Babu and Odeh physical model

3- Theory and Calculation

Mathematical First Radial Flow Period:

Babu and Odeh(6) derived an equation Eq.(1) describes the flow behavior of a horizontal well producing at a constant rate during this period as the following:

$$P_i - P_{wf} = \left(\frac{162.6 q \mu B_o}{\sqrt{k_x k_z} L} \right) \left[\log \left(\frac{t \sqrt{k_x k_z}}{\phi \mu c_t r_w^2} \right) - 3.23 + 0.87 S_m \right] \quad (1)$$

A plot of pwf vs. t for draw down test data and plot of pws vs. (tp+Δt)/ Δt for build up test data on semi-log paper gives a straight line with slope m_{r1}, from this slope can calculate the geometric average permeability $\sqrt{k_x k_z}$ and the mechanical skin S_m as the following:

$$\sqrt{k_x k_z} = \frac{162.6 q \mu B_o}{m_{1r} L} \quad (2)$$

$$S_m = 1.151 \left[\frac{\Delta P_{1hr}}{m_{1r}} - \log \left(\frac{\sqrt{k_x k_z}}{\phi \mu c_t r_w^2} \right) + 3.23 \right] \quad (3)$$

Where: $\Delta p_{1hr} = (p_i - p_{1hr})$ for drawdown test, and $\Delta p_{1hr} = (p_{\Delta t=1} - p_{wfo})$ for buildup test.

First Linear Flow Period:

A requirement of L for this early linear flow period to occur is given by the following equation:

$$L > 3.33 D_z \sqrt{\frac{k_y}{k_z}}; \text{ Where: } D_z = \max(z_o, h - z_o)$$

Babu and Odeh derived an equation Eq.(4) describes the flow behavior of a horizontal well producing at a constant rate during this period as the following:

$$P_i - P_{wf} = \left(\frac{8.13 q \mu B_o}{L h} \right) \left(\sqrt{\frac{t}{\phi \mu c_t k_x}} + \frac{17.37 h}{\sqrt{k_x k_z}} (S_z + S_m) \right) \quad (4)$$

A plot of pwf vs. t for drawdown test data and pws vs. $(\sqrt{t_p + \Delta t} - \sqrt{\Delta t})$ for buildup test data on linear paper should result in a straight line with a slope m_{L1}, from this slope the horizontal permeability in x-direction k_x can be calculated by using Eq.(5), and the mechanical skin from Eq.(6).

$$k_x = \left(\frac{8.13 q \mu B_o}{m_{1L} L h} \right)^2 / \phi \mu c_t \quad (5)$$

$$S_m = \left(\frac{L\sqrt{k_x k_z}}{141.2 q \mu B_o} \right) \Delta P_{0hr} - S_z \quad (6)$$

Where: $\Delta p_{0hr} = (p_i - p_{0hr})$ for drawdown test, and $\Delta p_{0hr} = (p_{\Delta t=0} - p_{wfo})$ for buildup test.

S_z is the pseudo skin may be visualized as the skin resulting from partial penetration in the vertical direction, given by:

$$S_z = \ln\left(\frac{h}{r_w}\right) + 0.25 \times \ln\left(\frac{k_x}{k_y}\right) - \ln\left(\sin\frac{180^\circ(d_z)}{h}\right) - 1.838 \quad (7)$$

Second Radial Flow Period:

For this period to occur, the penetration ratio should be $L/b < 0.45$. Babu and Odeh derived the following equation Eq.(8) to describe the flow in this period:

$$P_i - P_{wf} = \left(\frac{162.6q \mu B_o}{h \sqrt{k_x k_y}} \right) \left[\log\left(\frac{k_y t}{\phi \mu c_t L^2}\right) - 1.76 + 0.87 \sqrt{\frac{k_y}{k_x}} \frac{h}{L} (S_z + S_m) \right] \quad (8)$$

A plot of p_{wf} vs. t for the drawdown test data and p_{ws} vs. $(tp + \Delta t) / \Delta t$ for buildup test data on semi-log paper gives a straight line with slope m_{r2} , from this slope can be calculate the geometric average permeability $\sqrt{k_x k_y}$ in horizontal plane and the mechanical skin S_m as the following:

$$\sqrt{k_x k_y} = \frac{162.6q \mu B_o}{m_{2r} h} \quad (9)$$

$$S_m = \left(1.151 \sqrt{\frac{k_z L}{k_y h}} \right) \left[\frac{\Delta P_{1hr}}{m_{2r}} - \log\left(\frac{k_y}{\phi \mu c_t L^2}\right) + 1.76 \right] - S_z \quad (10)$$

Where: $\Delta p_{1hr} = (p_i - p_{1hr})$ for drawdown test, and

$\Delta p_{1hr} = (p_{\Delta t=1} - p_{wfo})$ for buildup test.

S_z as in Eq.(7).

Second Linear Flow Period:

Babu and Odeh derived the following equation Eq.(11) to describe the flow in this period:

$$P_i - P_{wf} = \left(\frac{8.13q \mu B_o}{b h} \right) \left(\sqrt{\frac{t}{\phi \mu c_t k_x}} + \frac{17.37h}{\sqrt{k_x k_z}} (S_z + S_t) \right) \quad (11)$$

A plot of p_{wf} vs. t for drawdown test data and p_{ws} vs. $(\sqrt{t_p + \Delta t} - \sqrt{\Delta t})$ for buildup test data on linear paper should result in a straight line with a slope m_{L2} , from which the horizontal permeability in x-direction k_x can be calculated and also the total skin S_t as the following:

$$k_x = \left(\frac{8.13q \mu B_o}{m_{2L} b h} \right)^2 / \phi \mu c_t \quad (12)$$

$$S_t = \left(\frac{L\sqrt{k_x k_z}}{141.2 q \mu B_o} \right) \Delta P_{0hr} - S_z \quad (13)$$

This is the only flow period that reflects the total skin, S_t Where:

$$S_t = S_m \left(\frac{b}{L} \right) + S_R \quad (14)$$

Where:

S_R = Skin due to partial penetration in all directions.

S_z as in Eq.(9).

To calculate the S_m , we need calculate S_R as shown below, once S_R is calculated, then

$$S_m = (L/b) (S_t - S_R).$$

Calculation of S_R :

As known, $S_R = 0$ when $L = b$. If $L < b$, then the value of partial penetration skin factor S_R depends upon the following two conditions:

Case (a):

$$S_R = P_{XYZ} + P'_{XY} \quad (15)$$

The P_{XYZ} Component is a result of the degree of penetration (L/b), and the P'_{XY} component is a result of the location of the well in x-y plane. The skin component resulting from the z location is negligible.

$$P_{XYZ} = \left(\frac{b}{L} - 1 \right) \left[\ln \left(\frac{h}{r_w} \right) + 0.25 \ln \left(\frac{k_x}{k_z} \right) - \ln \left(\sin \frac{180^\circ z_o}{h} \right) - 1.84 \right] \quad (16)$$

$$P'_{XY} = \left(\frac{2b^2}{Lh} \sqrt{\frac{k_z}{k_y}} \right) \left[F \left(\frac{L}{2b} \right) + 0.5 \left[F \left(\frac{4y_{mid} + L}{2b} \right) - F \left(\frac{4y_{mid} - L}{2b} \right) \right] \right] \quad (17)$$

Where pressure computations are made at $y_{mid} = (y_1 + y_2)/2$. (i.e. the midpoint along the well length).

$$F \left(\frac{L}{2b} \right) = - \left(\frac{L}{2b} \right) \left[0.145 + \ln \left(\frac{L}{2b} \right) - 0.137 \left(\frac{L}{2b} \right)^2 \right] \quad (18)$$

The evaluation of $F[(4y_{mid}+L)/2b]$ and $F[(4y_{mid}-L)/2b]$ depends on their arguments; i.e. $(4y_{mid}+L)/2b$ and $(4y_{mid}-L)/2b$. If the argument < 1 , Eq.(18) is used. In this case, $(L/2b)$ is replaced by $(4y_{mid}+L)/2b$ and/or $(4y_{mid}-L)/2b$. On the other hand, if the argument > 1 , then the following equation is used:

$$F(x) = (2-x) \left[0.145 + \ln(2-x) - 0.137(2-x)^2 \right] \quad (19)$$

Where $F(x)$ = Function used to describe effects of well location in horizontal plane, $x = (4y_{mid}+L)/2b$ or $(4y_{mid}-L)/2b$, with $x > 1$.

Case (b):

$$S_R = P_{XYZ} + P_Y + P_{XY} \quad (20)$$

The P_{XYZ} Component is given by Eq.(16).

$$P_Y = \left(\frac{6.28b^2}{ah} \right) \left(\frac{\sqrt{k_x k_z}}{k_y} \right) \left[\left(\frac{1}{3} - \frac{y_{mid}}{b} + \frac{y_{mid}^2}{b^2} \right) + \frac{L}{24b} \left(\frac{L}{b} - 3 \right) \right] \quad (21)$$

$$P_{XY} = \left(\frac{b}{L} - 1 \right) \left(\frac{6.82a}{h} \sqrt{\frac{k_z}{k_x}} \right) \left(\frac{1}{3} - \frac{x_o}{a} + \frac{x_o^2}{a^2} \right) \quad (22)$$

3.1 Mathematical Expressions and Symbols

This section may each be divided by subheadings or may be combined. A combined Results

Well x1 Test Analysis:

Pressure Build-up Test Analysis Calculations:

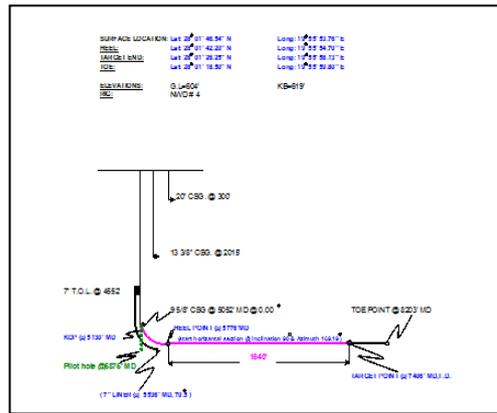


Figure 2: A Schematic of a wellbore diagram for Well x1

Step (1): General Data Required for the Test Analysis:

Table 1: Well Information

Well Orientation	Horizontal
Well Completion	Open Hole
Oil Production Rate, q	3792 STB/D
Producing Time, tP	24 hrs
Well Drilling Length, L	1640 ft
Well Radius, rw	0.250 ft
Vertical Well Location, zw	30 ft
Nearest Upper or Lower Boundary, dz	30 ft
Vertical Section Spacing, A	104 acres

Table 2: Reservoir Rock and Fluid Data.

Formation Thickness, h	60 ft
Formation Porosity, ϕ	26.9 %
Total Compressibility, ct	8.6×10^{-6} psi ⁻¹
Oil Formation Volume Factor, Bo	1.24 Bbl/STB
Oil Viscosity, μ_o	0.930 cp

Step (2): Calculate Pressure Drop and Pressure Derivative:

The pressure drop and derivative versus time data as shown in figure (3).

Step (3): Identification of Wellbore Storage Effect and Flow Periods:

- 1- Prepare a log-log plot of pressure drop $[(P_{ws} - P_{wfo}) \text{ vs. } \Delta t]$, as shown in figure (4)
- 2- From the log-log plot of $[(P_{ws} - P_{wfo}) \text{ vs. } \Delta t]$, unit slope line is not evident then there is no wellbore storage effect.
- 3- Prepare a pressure derivative $[(d(P_{ws})/d(\log \Delta t) \text{ vs. } \Delta t)]$ on a log-log graph. The plot is shown in figure (5).
- 4- From this plot, three flow periods can be clearly identified:
 - The first radial flow appearing as horizontal line during the period (0.167 to 0.333) hours.
 - The first linear flow period appearing as a $\frac{1}{2}$ unit slope line approximately during the period (0.583 to 2.750) hours.
 - The second radial flow appearing as horizontal line during the period (3.250 to 10) hours.

Step (4): Analysis of the First Radial Flow Period:

- 1) Plot pressure data versus Horner time function, $[P_{ws} \text{ vs. } ((t_p + \Delta t) / \Delta t)]$ on semi-log paper as shown in figure (6).
- 2) The semi-log plot of $[P_{ws} \text{ vs. } ((t_p + \Delta t) / \Delta t)]$ and of its slope, shows clearly that the Horner time from $(t_H = 145)$ to $(t_H = 73)$ can be fitted to a semi-log straight line as shown in figure (6). This could be interpreted as the effect of an early-time radial flow. (i.e., First radial flow period)
- 3) Read the slope directly from the plot, $m_{1r} = 87.91$ psi/cycle.
- 4) From this flow period the equivalent permeability in vertical plane (x-z directions), $\sqrt{k_x k_v}$, can be calculated:

$$\sqrt{k_x k_v} = 162.6 \left(\frac{q \mu B_o}{m_{1r} L} \right) = \sqrt{k_x k_v} = 4.93 \text{ md.}$$

$$162.6 \left(\frac{3792 \times 0.930 \times 1.24}{87.91 \times 1640} \right)$$

Step (5): Analysis of the First Linear Flow Period:

- 1) Plot pressure data versus $(\sqrt{t_p + \Delta t} - \sqrt{\Delta t})$ on linear paper as shown in figure (7).
- 2) The linear plot of $[P_{ws}$ vs. $(\sqrt{t_p + \Delta t} - \sqrt{\Delta t})]$, shows clearly that the square root time from $[(\sqrt{t_p + \Delta t} - \sqrt{\Delta t}) = 4.194]$ to $[(\sqrt{t_p + \Delta t} - \sqrt{\Delta t}) = 3.514]$ can be fitted to a linear straight line as shown in figure (7). This could be interpreted as the effect of an early-time linear flow. (i.e., First linear flow period)
- 3) Read the slope directly from the plot, $m_{L1} = 141.67$ psi/hr^{0.5}.

4) From this flow period the horizontal permeability in x-directions (k_x), can be calculated:

$$k_x = \left(\frac{8.13 q \mu B_o}{m_{L1} L h} \right)^2 \left(\frac{1}{\phi \mu c_t} \right) k_x = \left(\frac{8.13 \times 3792 \times 0.930 \times 1.24}{141.67 \times 1640 \times 60} \right)^2 \times \left(\frac{1}{0.269 \times 0.930 \times 8.6 \times 10^{-6}} \right) = 3.02 \text{ md.}$$

5) Combining results of the analysis of the early-time radial flow and early-time linear flow,

$$k_x = 3.02 \text{ md, and } \sqrt{k_x k_v} = 4.93 \text{ md. } \therefore k_v = 8.05 \text{ md}$$

Step (6): Analysis of the Second Radial Flow Period:

1) The semi-log plot of (P_{ws} vs. Horner time function) and its slope as shown in figure (6), shows clearly that the Horner time from ($t_{H1} = 8.385$) to ($t_{H2} = 3.400$) can be fitted to a semi-log straight line. This could be interpreted as the effect of the late-time pseudo-radial flow. (i.e., Second radial flow period)

2) Read the slope directly from the plot, $m_{2r} = 273.30$ psi/cycle.

3) From this flow period the average horizontal permeability in x, y directions $\sqrt{k_x k_y}$, can be calculated:

$$\sqrt{k_x k_y} = 162.6 \left(\frac{q \mu B_o}{m_{2r} h} \right) \sqrt{k_x k_y} = 43.36 \text{ md.}$$

4) Combining this result with the calculated value of $k_x = 3.02$ md (in step 4) we calculate the value of $k_y = 622.0$ md.

The following table summarized the permeability estimation from each flow period.

Table 3: summarized the permeability estimation from each flow period

Flow Period	Permeability Estimation	Result	Unit
First Radial Flow	$\sqrt{k_x k_v}$	4.93	md
First Linear Flow	k_x	3.02	md
	k_v	8.05	md
Late-time Pseudo-Radial Flow	$\sqrt{k_x k_y}$	43.36	md
	k_y	622.00	md

Step (7): Skin Factor Calculations:

1) Calculate the pseudo-skin caused by partial penetration in the vertical direction.

$$S_z = \ln\left(\frac{h}{r_w}\right) + 0.25 \times \ln\left(\frac{k_x}{k_v}\right) - \ln\left(\sin\frac{180^\circ(d_z)}{h}\right) - 1.838$$

S_z = 3.51

2) Evaluate the mechanical skin using the early-time radial results:

$$S_m = 1.151 \left[\frac{\Delta P_{1hr}}{m_{r1}} - \log\left(\frac{\sqrt{k_x k_v}}{\phi \mu c_t r_w^2}\right) + 3.23 \right]$$

S_m = - 3.14

3) Evaluate the mechanical skin using the early-time linear results:

Extrapolate the straight line on figure (6) to $\sqrt{\Delta t} = 0$, (i.e., to $(\sqrt{t_p + \Delta t} - \sqrt{\Delta t}) = 4.90$) read $P_{ws}(\Delta t = 0)$, and read P_{wfo} from actual measured test values, and then calculate ΔP_o .

$$S_m = \left(\frac{L \sqrt{k_x k_v}}{141.2 q \mu B_o} \right) \times (P_{ws}(\Delta t = 0) - P_{wfo}) - S_z$$

S_m = - 3.19

4) Evaluate the mechanical skin using the second radial flow results:

Extrapolate the straight line in figure (5) to $\Delta t = 1$ hour, (i.e., $((t_p + \Delta t) / \Delta t) = 25$), and read, $P_{ws}(1hr)$, and then calculate ΔP_{1hr} : $P_{ws}(1hr) = 1623$ psi, and $P_{wfo} = 1519$ psi $\Delta P_{1hr} = P_{ws}(1hr) - P_{wfo} = 1623 - 1519 = 104$ psi

$$S_m = 1.151 \frac{L}{h} \sqrt{\frac{k_v}{k_y}} \left[\frac{\Delta P_{1hr}}{m_{r2}} - \log\left(\frac{k_y}{\phi \mu c_t L^2}\right) + 1.76 \right] - S_z$$

S_m = - 3.12

The following table summarizes the skin estimation from each flow period.

Table 4: summarized the skin estimation from each flow period

Flow Period	Reservoir Parameter	Value
First Radial Flow	S _m	-3.14
First Linear Flow	S _z	3.51
	S _m	-3.19
Second Radial Flow	S _m	-3.12

We note the high consistency in the evaluation of S_m from the three periods, the average S_m being (-3.15)

Productivity Index in Horizontal Well (x1) Calculations:

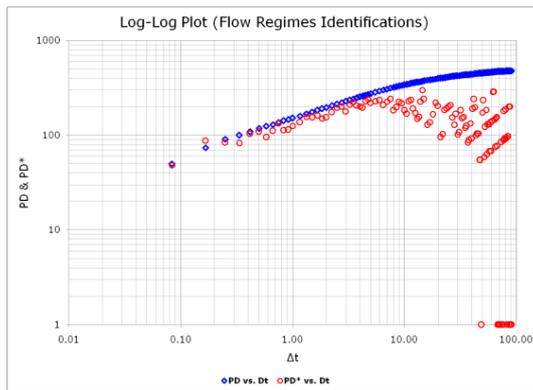


Figure 3: Log-log plot, flow regimes identification

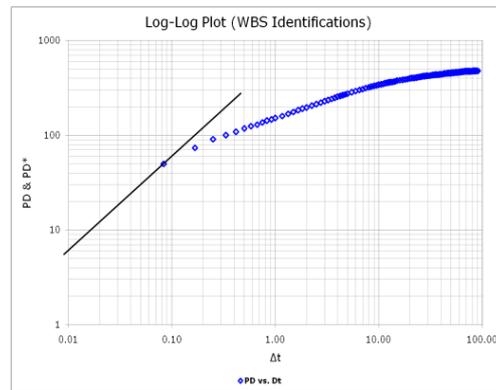


Figure 4: Log-log plot, wellbore storage effect identification

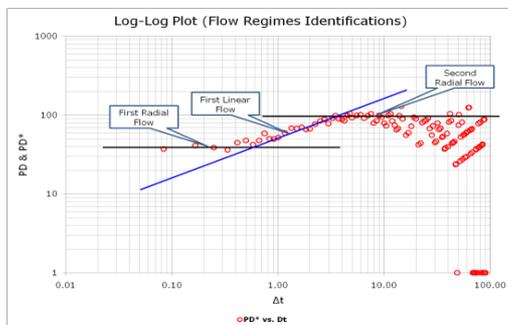


Figure 5: Log-log plot of pressure derivative, flow regimes identification

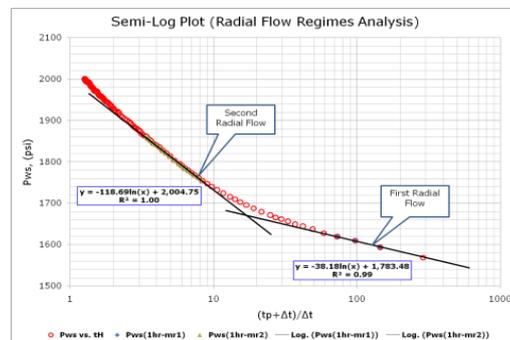


Figure 6: Semi-log plot, radial flow analysis

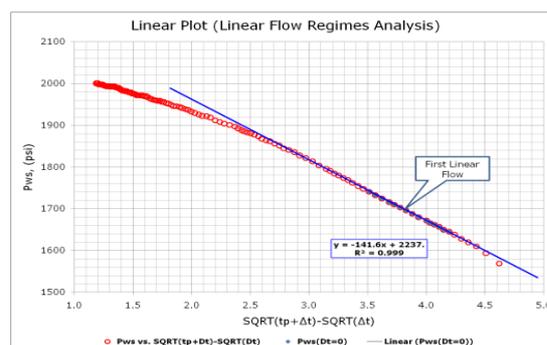


Figure 7: Linear-log plot, linear flow analysis

Calculation of drainage radius of the vertical well (r_{ev}) or (b):

$$r_{ev} = b = \sqrt{\frac{43560 A}{\pi}} = \sqrt{\frac{43560 \times 104}{\pi}} = 1201 \text{ft}$$

Calculation of the drainage area of the horizontal well (r_{eh}):

Method (1): $A_1 = \frac{L(2b) + \pi b^2}{43560} = 194 \text{acre}$

Method (2): $A_2 = \frac{\pi a b}{43560} = 175 \text{ acres}$

Drainage radius of the horizontal well (reh): $r_{eh} = \sqrt{\frac{43560 A_{avg}}{\pi}} = 1600 \text{ ft.}$

Productivity Index Calculation under Steady-State Condition:

The following table summarizes the productivity index and flow rate estimation for isotropic and anisotropic reservoir under steady state condition.

Table 5: summarized the productivity index and flow rate estimation for isotropic and anisotropy reservoir under steady state condition

Method	Steady State Condition			
	Isotropic reservoir		Anisotropic reservoir	
	Jh STB/day/Psi	qoh STB/day	Jh STB/day/Psi	qoh STB/day
Actual	7.88	3792	7.88	3792
Borisov	10.68	5137	#	#
The Giger-Reiss-Jourdan[4]	10.00	4810	6.75	3247
Joshi[8]	10.37	4988	6.92	3328
The Renard-Dupuy[5]	10.67	5132	9.39	4517
---	Pseudo Steady State Condition			
Mutalik et al.[7]	#	#	25.53	12280
Babu and Odeh[9]	#	#	1.56	750
Kuchuk et al.[10]	#	#	6.72	3232

4- Conclusions

Transient pressure analysis of horizontal wells is considerably more complicated than it is for vertical wells because of The existence of three and more flow regimes, in contrast to just one radial flow regime in normal vertical wells and The presence of at least three different types of skins and the non-uniformity of the mechanical skin, the skin value in a homogeneous formation would be minimum at the farthest end of the horizontal section and increasing as we approach the slanted and vertical section. The pressure test provides only an average value of the mechanical

skin along the horizontal section. To obtain a clear semi-log straight line of the late-time radial flow, there should be enough time kept for the production before shutting-in the well for build-up test. It is better and preferable to use Horner time function instead of using shut-in time for calculating the pressure derivative, especially in case of a long time test. In the fortunate case where most of the flow regimes are evident, it would be possible to calculate more than one value for the permeability perpendicular to the horizontal section (k_x) and for the mechanical skin (S_m), and checked against each other. This advantage is not available in case of vertical wells. The Giger-Reiss-Jourdan and Joshi was considered more representative result as compared with actual the productivity index and flow rate for isotropic and anisotropy reservoir In horizontal wells the vertical permeability plays an important role, since it is a main factor in the duration of early radial flow period. The flow rate value obtained by using anisotropy Renard-Dupuy method was confirmed by the actual flow rate that means the reservoir is anisotropic reservoir under steady-state condition

References

- [1] Urayet, A. A.: "Advanced Topics in Transient Pressure Analysis", *Part of the Technical Program Organized for the Petroleum Research Center*, Tripoli, 2004.
- [2] Goode, P. A. and Thambynayagam, R. K. M.: "Pressure Drawdown and Buildup Analysis for Horizontal Wells in Anisotropy Media", *Paper SPE 14250*, presented at the 60th Annual Technical Conference and Exhibition of the Society of Petroleum Engineers, Las Vegas, NV, Sept. 22-25, 1985. A revised version was presented at the SPE Formation Evaluation, pp. 683-697, December 1987.
- [3] Odeh, A. S. and Babu, D. K.: "Transient Flow Behavior of Horizontal Wells, Pressure Drawdown and Buildup Analysis", *SPE Formation Evaluation*, pp.7-15, March 1990.
- [4] Giger, F. M., Reiss, L. H., and Jourdan, A. P.: "The Reservoir Engineering Aspect of Horizontal Drilling", *Paper SPE 13024 presented at the SPE 59th Annual Technical Conference and Exhibition*, Houston, Texas, Sept. 16-19, 1984.
- [5] Borisov, Ju. P.: "Oil Production Using Horizontal and Multiple Deviation Wells", *Nedra, Moscom*, 1964. Translated by J. Strauss, S. D. Joshi (ed.), Phillips Petroleum Co., the R & D Library Translation, Bartlesville, Oklahoma, 1984.
- [6] Babu, D. K. and Odeh, A. S.: "Productivity of a Horizontal Well", *SPE Reservoir Engineering*, pp. 417-421, November 1989.
- [7] Ozkan, E., Raghavan, R., and Joshi, S. D.: "Horizontal Well Pressure Analysis", *SPE Formation Evaluation*, pp. 567-575, December 1989.
- [8] Joshi, S. D.: "Horizontal Well Technology", *Tulsa*, OK: PennWel Publishing Company, 1991.
- [9] Odeh, A. S. and Babu, D. K.: "Transient Flow Behavior of Horizontal Wells, Pressure Drawdown and Buildup Analysis", *Paper SPE 18802 presented at the SPE California Regional Meeting held in Bakersfield*, CA, Apr. 5-7, 1989.
- [10] Kuchuk, F. J.: "Well Testing and Interpretation for Horizontal Wells", *JPT*, pp. 36-40, January 1995.

Removal of Hexavalent Chromium from Aqueous Solution using Magnetic Sepiolite Nano-composite as an Adsorbent

Aysha Ali Ahribesh^{1*}, Atega Said Aljenkawi², Asma Mustafa Milad¹, Hanan Saleh Abosdil², Rada Petrović³

¹aysha.ahribesh@gmail.com, ²atega.aljenkawi@yahoo.com, ³asma.aga2009@gmail.com, ⁴hsabosdil@elmergib.edu.ly, ⁵radaab@tmf.bg.ac.rs

¹Department of Chemical and Petroleum Engineering, College of Engineering, Elmergib University, Libya

²Department of Chemistry, College of Science, Elmergib University, Libya

³University of Belgrade, Faculty of Technology and Metallurgy, Belgrade, Serbia

ABSTRACT

Magnetic nano-composite (MNCs) samples were prepared by co-precipitation method and characterized by XRD, SEM, nitrogen adsorption-desorption. The prepared MNCs evaluated for removal of hexavalent chromium Cr(VI) from its solutions. The adsorption experiments performed by bath method and the effects of solution pH, Cr(VI) concentration and contact time were studied. The adsorption capacity of the MNCs was affected significantly by pH of the Cr(VI) solution and based on the degree of magnetite. The results showed that the adsorption of Cr(VI) decreased with the increase of pH. The highest removal efficiency was achieved with the composite having the highest content of magnetite. The optimum contact time for maximum adsorption capacity was determined to be 480 min. The obtained adsorption data analyzed by the Langmuir, Freundlich and Sips isotherm models. The equilibrium data fitted well to the Sips model at a solution pH value of 2.0. The kinetic study showed the adsorption process followed pseudo-second-order kinetic. The prepared nano-composite materials can be used to remove Cr(VI) from water and industrial wastewater.

Keywords:

Magnetic nano-composites, sepiolite, chromium (VI), adsorption, modeling and kinetics.

1 Introduction

Hexavalent chromium Cr(VI) is a highly toxic element release from various industrial applications such as plating of metals, magnetic tapes, pigments, leather tanning, wood preserving, paints, electrical and electronics' equipment manufactures and catalyst production [1]. In natural waters, Cr(VI) is also presented due to erosion of chromium deposits found in rocks and soils. The concentration of chromium in drinking-water should be less than 1 $\mu\text{g}/\text{dm}^3$ [2]. Repeated exposure to hexavalent chromium compounds can cause damage to the nose, inflammation of the lungs, allergic reactions in the skin, kidney damage and cancer of the lung [3].

Several treatment technologies have been developed to remove chromium from water and wastewater [4]. The common methods include chemical precipitation, ion exchange, membrane separation, ultrafiltration, flotation, electrocoagulation, solvent extraction, reduction, reverse osmosis, dialysis/electrodialysis, adsorption/filtration, flocculation, chelation [3]. Among these methods, adsorption is considered as a promising process for the removal of different ions such as Cr(VI) due to its low cost and high ions selectivity.

In recent years, magnetic nano-composites (MNCs) displayed considerable attention as adsorbent materials to remove undesired chemical components from industrial wastewaters and the magnetic

materials can be separated conveniently from aqueous solutions. Several MNC materials have been investigated for removing of chemical pollutants from wastewater such as magnetite/sepiolite for removal of atrazine, Co^{2+} and Cd^{2+} [5,6,7]. Other types of MNCs for example, magnetite/zeolite-X [8], diatomite-supported/magnetite nanoparticles [9], montmorillonite-supported magnetite nanoparticles [10], magnetic activated carbon composites [11] and magnetic chitosan composites [12] for the removal of Cr(VI). Sepiolite is a hydrated magnesium silicate with the ideal formula of $\text{Si}_{12}\text{Mg}_8\text{O}_{30}(\text{OH})_4(\text{OH}_2)_4 \cdot 8\text{H}_2\text{O}$ which characterized by its fibrous morphology and intracrystalline channels [13]. Due to high specific surface area and nano-sized channels of the magnetite sepiolite particles, it can be dispersed into its structure [7].

The aim of this study was to investigate the adsorption of Cr(VI) ions on pure magnetite, and magnetite/sepiolite nano-composites. The effects of solution pH, Cr(VI) concentration and contact time on the adsorption capacity of Cr(VI) are evaluated. The achieved experimental results of Cr(VI) removal were analyzed by adsorption isotherms and kinetics models.

2 Experimental Procedure

2.1 Materials and Methods

The natural sepiolite (SEP) obtained from Andrići (Serbia). The reagent grade chemicals, such as NaOH, $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$ and $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ and $\text{K}_2\text{Cr}_2\text{O}_7$ were used without any further purification.

2.2 Preparation of Sepiolite-magnetite Nano-composites and Pure Magnetite

The sepiolite-magnetite nano-composite (MSEP) and pure magnetite (Fe_3O_4) were prepared by the co-precipitation method according to procedures described in previous study [7]. The details are: magnetic sepiolite composites were synthesized by adding 5 g of sepiolite into 1 M NaOH solution and the mixture stirred for 30 min under nitrogen atmosphere. Exactly 4.51 g of $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$ and 2.306 g of $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ with Fe(II)/Fe(III) molar ratio of 1:2 were dissolved in 100 cm^3 of deoxygenated deionized water by bubbling N_2 gas then kept at 60 °C in water bath under nitrogen atmosphere. The sepiolite and NaOH suspension was added into the Fe(II)/Fe(III) solution and the mixture was aged for 2 h at 60 °C. The produced solid materials separated and washed three times with deoxygenated deionized water and then with ethanol. The powder dried in vacuum oven at 60 °C until constant mass. Three types of MSEP (MSEP 1, MSEP 2 and MSEP 3) are synthesized using different weights of both $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$ and $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ (see Table 1).

Pure magnetite particles were synthesized by dissolving 4.51 g of $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$ and 2.306 g of $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ in 100 cm^3 of deoxygenated deionized water by N_2 gas. The solution was kept under nitrogen atmosphere in a water bath at 60 °C followed by drop wise adding of a desired concentration of NaOH solution (25 %). The formed black mixture aged at 60 °C for 2 h. The obtained solid material was separated and washed several times with deionized and deoxygenated water and then with ethanol and dried at 60 °C under vacuum until constant weight.

Table 1: Masses of $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$ and $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ salts for the composites preparation and assessment of the magnetization

Sample	$\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ (g)	$\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$ (g)	Magnetization
MSEP 1	2.31	4.51	Yes
MSEP 2	3.46	6.77	
MSEP 3	1.15	2.25	

2.3 Characterization of the Adsorbents

X-ray diffraction (XRD) analysis of the samples was carried out by ITAL STRUCTURES APD2000 diffractometer using Cu KR radiation in the 2θ angle range from 3° to 50° with a 0.02°

step. The specific surface area (S_{BET}), volume of the mesopores and micropores and the pore size distribution of the samples were calculated from the nitrogen adsorption–desorption isotherms using a Micrometrics ASAP 2020 apparatus. Furthermore, the morphology of the samples was examined by a Tescan MIRA3 field emission gun scanning electron microscope (FESEM).

2.4 Adsorption Experiments

Adsorption of Cr(VI) ions from water by the produced pure magnetite and sepiolite-magnetite nano-composites materials is investigated by batch method at room temperature (25 ± 1 °C). The general method used for this study is described as follows: 20 cm³ of chromium solution was placed in 50 cm³ reagent bottle and pH of solution adjusted to a desired value. A constant amount of adsorbent (0.02 g) was added to the solution and shaken for certain contact time. The pH of solutions before and after adsorption measured by using a pH meter (Ino Lab WTW series pH 720) and the initial and final concentrations of Cr(VI) ions were determined by using atomic absorption spectrometer (AAS) (Perkin Elmer 730). The effects of different pH value (2.0 to 5.0), ion concentration (10, 20, 30, 40, 50 mg/dm³) and contact time (0.5 to 24 h) on Cr(VI) ions adsorption capacity are investigated. The equilibrium adsorption capacity (q_e) was calculated by using the following equation [13]:

$$q_e = \frac{C_i - C_e}{m} \cdot V \quad (1)$$

where C_i and C_e are the initial and the equilibrium concentrations of Cr(VI) (mg/dm³), respectively, m is the mass of the adsorbent (g), and V is the volume of the solution (dm³). The quantity of Cr(VI) ions adsorbed after period of time t (q_t) was calculated according to the following equation [13]:

$$q_t = \frac{C_i - C_t}{m} \cdot V \quad (2)$$

where C_t is the concentration of Cr(VI) after period of time t .

2.5 Adsorption Isotherms

The adsorption isotherms play an important role in understanding the mechanism of adsorption process (adsorbate-adsorbent interaction) as well as to estimate the maximum capacity of an adsorbent. Langmuir, Freundlich and Sips isotherm models has been selected to analyze the obtained results from initial chromate concentration at pH 2.0 ± 0.1 . The adsorption isotherms constants were determined by non-linear regression analysis using the OriginPro 8.5. The experimental adsorption kinetic data was fitted with pseudo first-order and pseudo-second-order models.

2.5.1 Langmuir model

Langmuir adsorption isotherm assumed that the adsorption takes place at specific localized sites (homogeneous sites) on the surface of adsorbent and adsorption is complete when a monolayer is formed [14]. The nonlinear form Langmuir isotherm is described by the following equation [14]:

$$q_e = \frac{q_m K_L C_e}{1 + K_L C_e} \quad (3)$$

where q_m is the maximum adsorption capacity (mg/g) and K_L is the Langmuir constant related to the energy of adsorption (dm³/mg).

2.5.2 Freundlich model

In Freundlich model, the adsorbate form multilayer on surface of the adsorbent with non-uniform distribution of adsorption energies over the heterogeneous surface of adsorbent [15], which may include chemisorption if active sites are strong followed by physisorption [16]. Freundlich isotherm model is expressed by the following equation:

$$q_e = K_f C_e^{1/n} \quad (4)$$

where K_f is the Freundlich constant related to the adsorption capacity ($\text{mg}^{(1-1/n)} \text{dm}^{3/n}/\text{g}$) and n is the dimensionless adsorption intensity parameter.

2.5.3 Sips model

Sips isotherm is a combined form of Langmuir and Freundlich models [17]. When C_e approaches at low value, the Sips isotherm effectively reduces to Freundlich, which can be employed to describe the system's heterogeneity, while at high C_e it predicts the Langmuir monolayer adsorption characteristic and it implies a homogeneous adsorption process. The Sips isotherm model is described by the following mathematical equation:

$$q_e = \frac{q_m K_a \cdot (C_e^{n_s})}{1 + K_a \cdot (C_e^{n_s})} \quad (5)$$

where K_a is the Sips equilibrium constant (dm^3/mg) n_s and n_s is the index of heterogeneity.

2.6 Kinetics Study

The effect of contact time on adsorption of Cr(VI) on MSEP 2 adsorbent were studied by changing the time from 0.5 to 24 h using solutions of 40 mg/dm^3 concentration at $\text{pH } 2.0 \pm 0.1$. Pseudo-first order and pseudo-second order models are widely used to examine the adsorption kinetics data to describing the adsorption mechanism [18,19]. The pseudo-first-order and pseudo-second-order models were fitted to the kinetics of Cr(VI) ions adsorbed onto MSEP 2 data. The linearized mathematical pseudo 1st order model is expressed by the following equation [18]:

$$\log(q_e - q_t) = \log q_e - \frac{k_1 t}{2.303} \quad (6)$$

where q_e is adsorption capacity at equilibrium in mg/g , q_t is the adsorption capacity at time t in mg/g and k_1 is the first order rate constant (min^{-1}). By plotting $\log(q_e - q_t)$ versus t , a straight line should be obtained with a slope of $-k_1$ and y-intercept of $\log q_e$. The adsorption is more inclined towards physisorption if the pseudo 1st order model best fits the adsorption kinetic data. The pseudo-second-order kinetic equation is given by the subsequent equation [19]:

$$\frac{t}{q_t} = \frac{1}{k_2 q_e^2} + \frac{t}{q_e} \quad (7)$$

where: k_2 is the rate constant of the pseudo second order model in unit of $\text{g}/(\text{mg} \cdot \text{min})$. If the experimental data fits to pseudo 2nd order model then the adsorption's mechanism is chemisorption controlled.

3 Results and Discussion

The produced adsorbent composites are characterized by XRD patterns, SEM images and textural parameters of MSEP 2 and MSEP 3. In our previous reported work [7], the XRD patterns of SEP and Fe₃O₄ are reported. That reported data is used in this work for the comparison. The XRD patterns of the MSEP 2 and MSEP 3 composite samples are shown in Figure 1. Diffraction characteristic peaks of SEP were seen in these samples and the main diffraction peak of Fe₃O₄ is observed at 2θ of $\sim 35.5^\circ$ partially overlapped with the sepiolite peak at $2\theta \sim 35^\circ$ [7]. However, taking into account that the most intensive magnetite peak is hardly noticeable in the pattern of the MSEP 3 sample, probably because of the low quantity of magnetite in the sample. The sample with higher content of magnetite (MSEP 2) showed a wide peak which indicated the presence of small crystallites as a result of the formation of high number of nuclei at high concentration of Fe(II) and Fe(III).

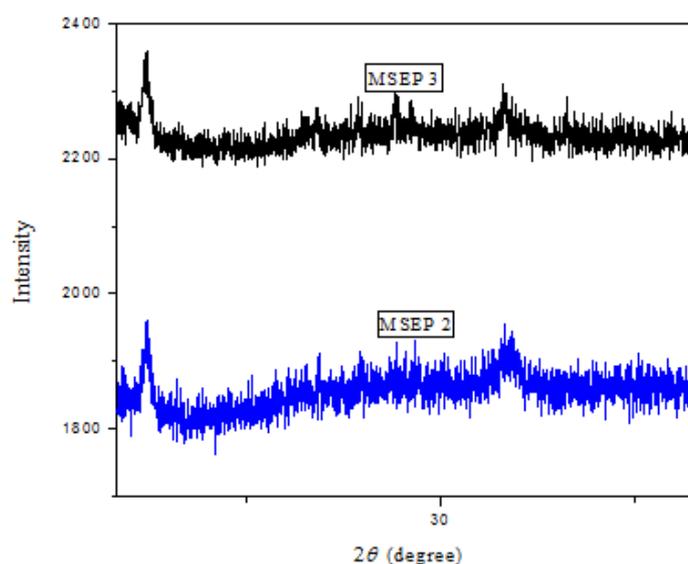


Figure 1: The XRD patterns of MSEP 2 and MSEP 3 composites.

The SEM images of the samples MSEP 2 and MSEP 3 are presented in Figure 2. Both MSEP 2 (Figure 2a) and MSEP 3 (Figure 2b) composite samples showed similar surface morphologies to the surface morphology of the sample MSEP 1 as reported in our previous work [7]. The SEM images of the samples are clearly showed the single and aggregates of magnetic particles along with the sepiolite fibers. Furthermore, the samples MSEP 2 and MSEP 3 had different content of magnetite but the differences in morphologies were insignificant (Figures 2 a and b).

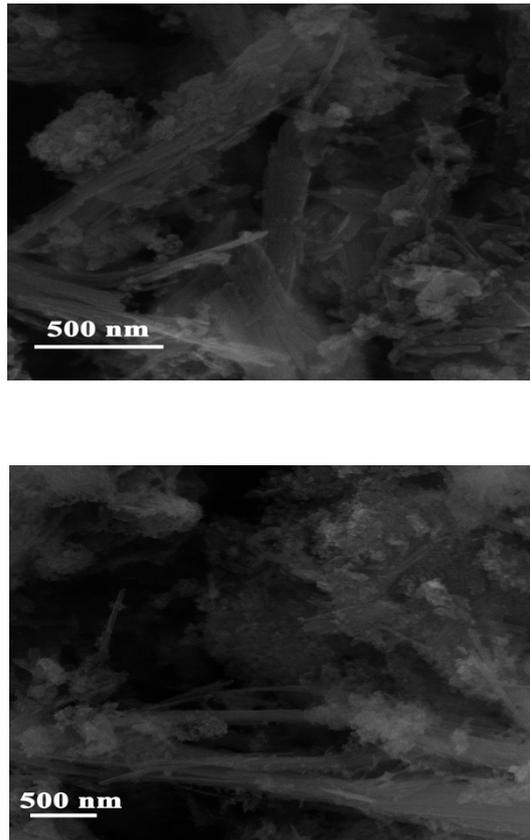


Figure 2: SEM images of (a) MSEP 2 and (b) MSEP 3.

The textural parameters of both MSEP 2 and MSEP 3 composites are showed in Table 2 while textural parameters of Fe_3O_4 and SEP are reported previously [7]. It can be seen from the Table 2, that the values of the specific surface area of the composites are between the values of pure compounds of Fe_3O_4 and SEP. The mesopore volume of the composites is slightly higher than SEP and the micropore volume is lower than of SEP. This may be due to the coating of SEP surfaces by Fe_3O_4 which obstructs some of the main micropore channels of SEP and the formation of new mesopores between Fe_3O_4 particles in the composites. The specific surface area, volume of mesopores and volume of micropores of MSEP 3 composite are slightly higher than MSEP 2 composite. It is obvious that the higher quantity of iron salts lead to lower S_{BET} , V_{micro} and V_{meso} , due to the higher magnetite content and aggregation.

Table 2: The textural parameters of MSEP 2 and MSEP 3 composite samples.

Sample	Specific Surface Area S_{BET} (m^2/g)	Micropore Volume V_{micro} (cm^3/g)	Mesopore Volume V_{meso} (cm^3/g)	Maximum Pore Diameter D_{max} (nm)	Mean Pore Diameter D_{mean} (nm)
MSEP 2	182.6	0.059	0.354	3.7	7.9
MSEP 3	215.1	0.081	0.498	4.0	10.6

3.1 Effect of Solution pH

The pH plays major role in the adsorption process because it affects the solution chemistry of adsorbate and surface charge of adsorbent. The influence of initial pH of solution on the adsorption capacity of Cr(VI) from different solution concentrations for the sample MSEP 1 was

studied and the results presented in Figure 3. The results showed that the adsorption capacity sharply decreases when the increase of pH value above pH 2. Therefore, pH 2.0 has been considered as the optimum pH for further Cr(VI) adsorption experiments. General, Cr(VI) is present in solution as anions, CrO_4^{2-} , $HCrO_4^-$ or $Cr_2O_7^{2-}$, depending on its concentration and solution pH. In acidic solution, $HCrO_4^-$ exists at low concentrations, and $Cr_2O_7^{2-}$ and $HCrO_4^-$ exist in solutions of high concentration, whereas CrO_4^{2-} becomes dominant species at $pH \geq 7$. For wide range of pH values, the surface charge of natural sepiolite is negative, therefore adsorption of Cr(VI) anions on sepiolite was negligible [20]. The decrease in the adsorption capacity with pH value could be attributed to the increasing in the negative charges of the adsorbent surface [9,10, 21].

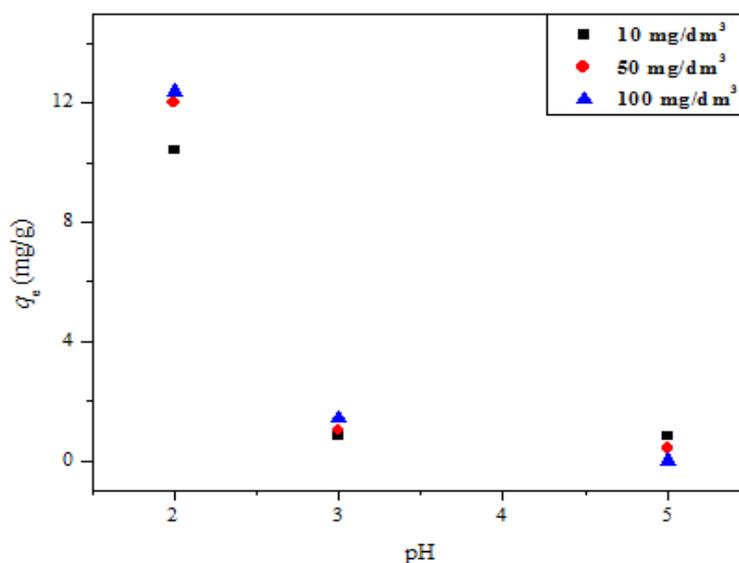


Figure 3: The effect of the initial pH on Cr(VI) adsorption by sample MSEP 1 for the initial Cr(VI) concentrations of 10, 50 and 100 mg/dm³.

3.2 Effect of Cr(VI) Concentration

Figure 4 presents the relationship between the maximum adsorption capacity of Cr(VI) ions on Fe₃O₄, MSEP 1, MSEP 2 and MSEP 3 and the concentration of Cr(VI) solution. It is very clear that the maximum adsorption capacities of the different adsorbents have similar trends with different values. The maximum capacity of Cr(VI) ions (17.3 mg/g) on pure magnetite is higher than the other adsorbent materials (MNCs). Furthermore, the sample with higher content of magnetite MSEP 2 displays highest capacity of 15.6 mg/g than MSEP 1 (11.8 mg/g) and lower content of magnetite MSEP 3 (9.0 mg/g). It is obvious that the content of magnetite is the major factor controlled the adsorption capacity due to the limit capacity of the sepiolite [20]. The adsorption capacity of the sample MSEP 2 is less than the pure magnetite by about 10%, while the theoretical content of magnetite in this sample was 36.5 weight percentages. It can be concluded that the presence of sepiolite in composites has improved the adsorption properties of the magnetite.

According to the literature [22], Cr(VI) can be adsorbed onto magnetite and magnetite composites by formation of inner and outer sphere complexes. In addition, the possible mechanism is also the reduction of Cr(VI) by magnetite to Cr(III) ions, which can be adsorbed on the surface of the adsorbent or precipitate as Cr(OH)₃ or Cr₂O₃. The precipitation is more likely at higher pH values, but reduction of Cr(VI) at higher pH values is difficult to take place. It can be supposed that the main mechanism of chromate adsorption on magnetite and magnetite composites is the electrostatic interactions of chromate ions and protonated surface functional groups (formation of outer-sphere complexes).

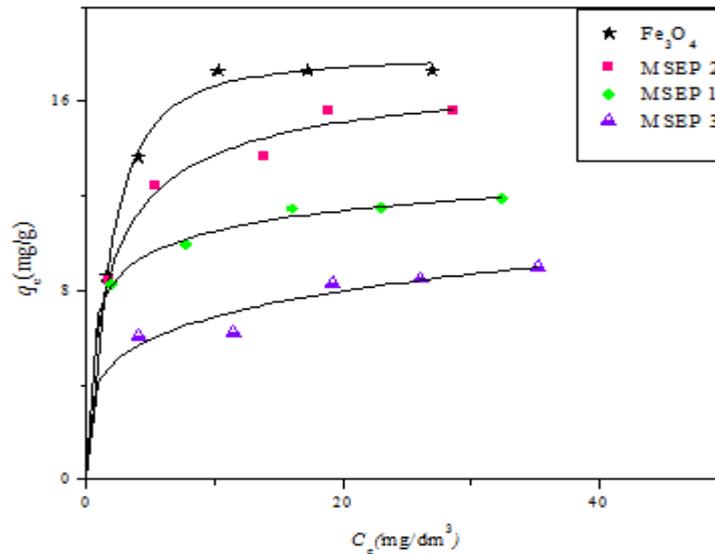


Figure 4: Adsorption isotherms for Cr(VI) onto Fe₃O₄ and the MNCs (adsorbent dosage = 0.02 g/20 cm³, pH 2.0±0.1).

3.3 Adsorption Isotherm Study

The adsorption isotherms constants are summarized in Table 3. The correlation coefficients (R^2) of the fitted data of the studied isotherm models proved that the Sips model showed the highest R^2 (0.998) and is the best model to explain the adsorption behavior of Cr(VI) for all the samples. Both Langmuir and Freundlich isotherm models were also fitted the results well. Sips isotherm is a combined form of Langmuir and Freundlich expressions deduced for predicting the heterogeneous adsorption systems and circumventing the limitation of the rising adsorbate concentration associated with Freundlich isotherm model. At low adsorbate concentrations, it reduces to Freundlich isotherm while at high concentrations it predicts a monolayer adsorption capacity characteristic of the Langmuir isotherm. According to the suitability of Sips model for the chromium adsorption, it can be supposed that the ions adsorbed on the heterogeneous surface of the adsorbents until the monolayer is formed.

The maximum adsorption capacities (q_m) and K_L parameter calculated by Langmuir model are found in the following order MSEP 1 > MSEP 2 > Fe₃O₄ > MSEP 3. The high values of K_L due to strong interactions between Cr(VI) ions and surface of the samples [17]. The Freundlich constant K_F is indicator of the adsorption capacity, while n is a adsorption intensity parameter (the strength of adsorption bonds is high when n is high) [17]. Moreover, the calculated maximum adsorption capacities by Sips model were slightly higher than that obtained by Langmuir model. Sips constant (K_a) is indicator of strength bond between the Cr(VI) ions and active sites on the surface of samples [17]. The decrease in K_a values is in the order of MSEP 1 > MSEP 2 > Fe₃O₄ > MSEP 3. The value of n_s in Sips model for a homogeneous material is 1 but it is less than one for heterogeneous materials [13]. Finally, the correlation coefficient (R^2) of isotherm models has very high value for all samples (near to the unity as shown in Table 3), which indicated that theoretical values (q_m) are in good agreement with experimental values ($q_{e,exp}$).

Table 3: Langmuir, Freundlich and Sips isotherms constants and coefficients of correlation (R^2) for the adsorption of Cr(VI) onto the Fe_3O_4 and the magnetic composites at $pH\ 2.0 \pm 0.1$

Sample	Adsorption model									
	Langmuir			Freundlich			Sips			
	q_m (mg/g)	K_L (dm^3/mg)	R^2	k_f ($mg^{(1-1/n)}$ $dm^{3/n}/g$)	$1/n$	R^2	q_m (mg/g)	K_a (dm^3/mg) ^{ns}	n_s	R^2
MSEP 1	11.91	1.087	0.994	7.668	0.131	0.997	12.43	0.827	0.359	0.998
MSEP 2	16.31	0.608	0.992	8.209	0.203	0.987	17.79	0.621	0.738	0.994
MSEP 3	9.26	0.364	0.963	4.237	0.211	0.981	9.83	0.022	0.219	0.981
Fe_3O_4	17.82	0.579	0.993	9.503	0.206	0.958	17.96	0.481	1.419	0.997

3.4 Adsorption Kinetics Study

To study the effect of contact time on adsorption capacity of Cr(VI) ions, the sample with the highest adsorption capacity has been selected (MSEP 2). The adsorption capacity as a function of contact time is presented in Figure 5. The adsorption process required relative long time to reach the equilibrium. The optimum contact time for maximum adsorption capacity was determined to be about 480 min in which the chromium ions formed outer-sphere complexes and concentrated near the charged surface with weak bonds. All steps of adsorption process which are external film diffusion, internal particle diffusion and adsorption step need long time to be completed. The overall rate of adsorption is controlled by the slowest step, which is film diffusion or pore diffusion.

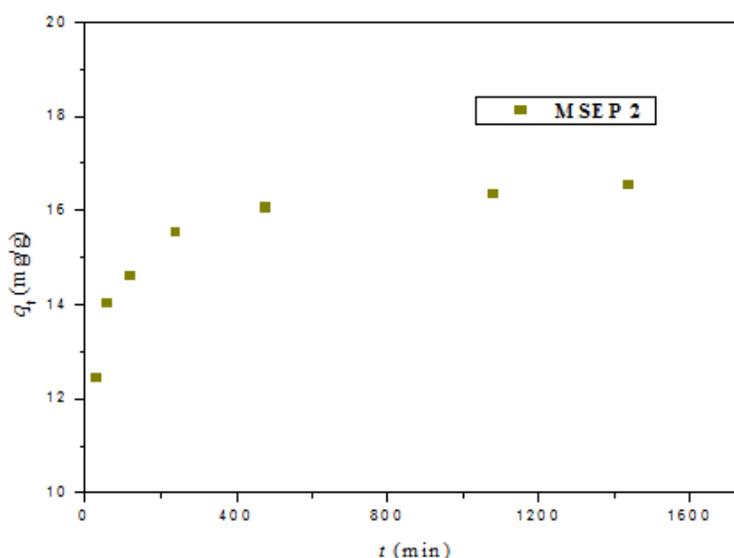


Figure 5: The effects of adsorption time on the adsorbed quantity of Cr(VI) at $pH_i = 2.0 \pm 0.1$ onto MSEP 2, (Cr(VI) concentration was $40\ mg/dm^3$).

In order to evaluate the adsorption kinetics of Cr(VI) on MSEP 2, the conventional kinetic models, pseudo-first and pseudo-second order models, were applied to analyze the experimental data. The adsorption kinetics constants, obtained by linear fitting and correlation coefficients (R^2) are summarized in Table 4. The correlation coefficients suggest that the adsorption kinetic data was fitted much better to pseudo-second order rate equation than by pseudo-first order equation. The R^2 value for the pseudo-first order model is 0.858 while for pseudo-second model is close to 1. Moreover, the value of equilibrium adsorption capacity (q_e) of 16.66 mg/g calculated by pseudo-second mode was close to the experimental maximum capacity (16.53 mg/g). These results

illustrated that the rate-limiting step in Cr(VI) is a chemisorption which involve valency forces through sharing or exchanging of electrons between adsorbent and adsorbate.

Table 4: The kinetic parameters and correlation coefficients for the adsorption of Cr(VI) onto the MSEP 2 composite.

Adsorbent	Pseudo-first order model			Pseudo-second order model		
	k_1 (min^{-1})	q_e (mg/g)	R^2	q_e (mg/g)	k_2 ($\text{g/mg}\cdot\text{min}$)	R^2
MSEP 2	$2.30\cdot 10^{-3}$	1.53	0.858	16.66	0.074	0.999

4 Conclusions

In this study, natural sepiolite (SEP) was used to synthesize the magnetic nano-composites (MNCs) for the removal of Cr(VI) ions from solution. The adsorption capacity of the sepiolite for Cr(VI) was very low and magnetite sepiolite improved the Cr(VI) adsorption capacity. Adsorption capacity of composites decreased when pH of the solution increased. The optimum pH value of the Cr(VI) solution for the maximum adsorption of Cr(VI) onto MNCs, was 2.0 under the experimental conditions. The magnetic nano-composite sample (MSEP 2) showed the highest adsorption capacity among other prepared MNCs samples. The optimum contact time for maximum adsorption capacity was 480 min. Adsorption isotherm models study of Cr(VI) onto the prepared nano-composites showed the best fitted data with the Sips model and suggested the adsorption onto all adsorbents occurred via chemisorption mechanisms. Pseudo-second order model has given the good fitting of the data and suggested the rate-limiting step mechanism.

5 Acknowledgements

The authors would like to acknowledge the financial support of the Ministry of Education, Science and Technological Development of the Republic of Serbia, Project No. III 45019.

Reference

- [1] N.K. Chandra Babu, K. Asma, A. Raghupathi, R. Venba, R. Ramesh, S. Sadulla, Screening of leather auxiliaries for their role in toxic hexavalent chromium formation in leather-posing potential health hazards to the users, *J. Clean.Prod.* 13, 1189-1195. 2005.
- [2] Controlling Exposure to Hexavalent Chromium in Aerospace and Air Transport Painting. OSHA Fact Sheet (Publication 3650), (2013) https://www.osha.gov/Publications/OSHA_FS-3650_Aerospace_Paint.pdf. Access online on 5 OCT 2020.
- [3] M. Kumari, CU. Pittman Jr, D. Mohan, Heavy metals [chromium (VI) and lead (II)] removal from water using mesoporous magnetite (Fe_3O_4) nanospheres, *J. Colloid Interf. Sci.* 442, 120–132. 2015.
- [4] D. Mohan, C. Pittman Jr, Activated carbons and low cost adsorbents for remediation of tri- and hexavalent chromium from water, *J. Hazard. Mater.* 137, 762-811. 2006
- [5] H. Liu, W. Chen, C. Liu, Y. Liu, C. Dong, Magnetic mesoporous clay adsorbent: Preparation, characterization and adsorption capacity for atrazine, *Micropor. Mesopor. Mat.* 194, 72-78. 2014.
- [6] S. Yu, L. Zhai, S. Zhong, Y. Qiu, L. Cheng, X. Ren, Synthesis and structural characterization of magnetite/sepiolite composite and its sorptive properties for Co(II) and Cd(II), *J. Taiwan Inst. Chem. E.* 59, 221-228. 2016.

- [7] A. A. Ahribesh, S. Lazarević, I. Janković-Častvan, B. Jokić, V. Spasojević, T. Radetić, Đ. Janačković, R. Petrović, Influence of the synthesis parameters on the properties of the sepiolite based magnetic adsorbents, *powder Technol.* 305, 260-269. 2017.
- [8] M-E. Kouli, G. Banis, M. Savvidou, A. Ferraro, E. Hristoforou, A study on magnetic removal of hexavalent chromium from Solutions using magnetite/zeolite-X composite particles as adsorbing material, *Int. J. Mol. Sci.* 21, 2707. 2020.
- [9] P. Yuan, D. Liu, M. Fan, D. Yang, R. Zhu, F. Ge, J. Zhu, H. He, Removal of hexavalent chromium [Cr(VI)] from aqueous solutions by the diatomite supported/unsupported magnetite nanoparticles, *J. Hazard. Mater.* 173, 614-62. 2010.
- [10] P. Yuan, M. Fan, D. Yang, H. He, D. Liu, A. Yuan, J. Zhu, T. Chen, Montmorillonite-supported magnetite nanoparticles for the removal of hexavalent chromium [Cr(VI)] from aqueous solutions, *J. Hazard. Mater.* 166, 821-829. 2009.
- [11] S. Bayazit, Ö. Kerkez, Hexavalent chromium adsorption on super paramagnetic multi-wall carbon nanotubes and activated carbon composites, *Chem. Eng. Res. Des.* 92, 2725-2733. 2014.
- [12] Y. Wei-Chun, T. Qiong-Zhi, D. Shu-Yu, C. Li-Yuan, W. Hai-Ying, Single-step synthesis of magnetic chitosan composites and application for chromate (Cr(VI)) removal, *J. Cent. South Univ.* 23, 317-323. 2016.
- [13] A.J. Habish, S. Lazarević, I. Janković-Častvan, B. Potkonjak, Đ. Janačković, R. Petrović, The effect of salinity on the sorption of cadmium ions from aqueous medium on Fe(III)-sepiolite, *Chem. Ind. Chem. Eng. Q.* 21, 295-303. 2015.
- [14] I. Langmuir, The adsorption of gases on plane surfaces of glass, mica and platinum, *J. Am. Chem. Soc.* 40, 1361-1403. 1918.
- [15] H. Freundlich, Concerning adsorption in solutions, *Zeitschrift für Physikalische Chemie*, 57, 385-470. 1906.
- [16] A.W. Adamson, A.P. Gast, Physical Chemistry of Surfaces, sixth ed., Wiley - Interscience, New York, 1997.
- [17] K.Y. Foo, B.H. Hameed, Insights into the modeling of adsorption isotherm systems, *Chem. Eng. J.* 156, 2-10. 2010.
- [18] S. Lagergren, K. Sven, *Vetenskapsakad. Handl.* 24, 1-39. 1898.
- [19] Y.S. Ho, G. McKay, *Process Biochem.* 34, 451-465. 1999.
- [20] V. Marjanović, S. Lazarević, I. Janković-Častvan, B. Potkonjak, Đ. Janačković, R. Petrović, Chromium(VI) removal from aqueous solutions using mercaptosilane functionalized sepiolites, *Chem. Eng. J.* 166, 198-206. 2011.
- [21] S. Rajput, C. U. Pittman Jr, D. Mohan, Magnetic magnetite (Fe₃O₄) nanoparticle synthesis and applications for lead (Pb²⁺) and chromium (Cr⁶⁺) removal from water, *J. Colloid and Interface Sci.* 468, 334-346. 2016.
- [22] G. Gallios, M. Václavíková, Removal of chromium (VI) from water streams: a thermodynamic study, *Environ Chem Lett* 6, 235-240. 2008. doi 10.1007/s10311-007-0128-8.



Track No. 4: Electrical & Computer Engineering

A Novel Denoising Method Based on Discrete Linear Chirp Transform

Osama A S Alkishriwo *

o.alkishriwo@uot.edu.ly

Department of Electrical and Electronic Engineering, University of Tripoli, Libya

ABSTRACT

Denoising of chirp based signals is a challenging problem in signal processing and communications. In this paper, we propose a suitable denoising algorithm based on the discrete linear chirp transform (DLCT), which provides local signal decomposition in terms of linear chirps. Analytical expression for the optimal filter response is derived. The method relies on the ability of the DLCT for providing a sparse representation to a wide class of broadband signals like chirp signals. Simulation results show the efficiency of the proposed method, especially for mono-component chirp signals.

Keywords:

Discrete linear chirp transform,
Denoising,
Chirp signals,
Nonstationary signal filtering.

1. Introduction

Chirp signals, known as linear frequency modulated signals, are important class of nonstationary signals which frequently encountered in many practical applications, such as radar, sonar, and telecommunications. Noise is the main factor which influences the transmission and recognition of signals in communications, radar, sonar and optics. In order to reduce the influences caused by noise, several methods and algorithms have been proposed for denoising. Unfortunately, most of these algorithms were not efficient for nonstationary signals, such as chirp signals, which are characterized with time–frequency identities. In general methods or algorithms that are based on frequency domain only are not effective for such type of signals [1].

To deal with nonstationary signals, several algorithms are presented in the literature of denoising chirp based signals. Most of them are based on extending Fourier representations capable of providing instantaneous–frequency information for multi–component signals. These algorithms can be achieved by considering polynomial–phase transform [2], or second–order polynomial transforms [3-4]; however, the latter is preferable due to computational viability. Furthermore, a parametric characterization of the instantaneous frequency of each of the components [5] provides a realistic view of the evolving nature of the signal. Although procedures based on the chirplet transform [6–7], and polynomial chirplet transform (PCT) [8] have been proposed, their numerical implementation is difficult to obtain because of no straightforward way to solve non–convex optimization problems with multiple extrema.

In one hand, the fractional Fourier transform (FrFT) proposed by Namias [9] in 1980 has drawn a considerable amount of attention in analysing and processing of nonstationary chirp based signals. It is a generalization of the conventional Fourier transform. It has been applied to different problems in signal processing including signal separation and filtering [10–15]. The FrFT provides a continuous representation of a signal from the time to the frequency domain at intermediate domains by means of the fractional order of the transform. On the other hand, the DLCT is also an extension of the discrete Fourier transform (DFT) and provides a parametric modelling of the

instantaneous frequencies of the components. It is introduced to represent a signal as a combination of linear chirps [4]. The DLCT is implemented efficiently using the fast Fourier transform (FFT) algorithm and can be applied for compression of nonstationary signals [16]. Unlike the FrFT in doing joint time–frequency representation, the DLCT is a joint chirp–rate frequency transformation, which can do a better job in denoising chirp based (nonstationary) signals. It has been shown in [17] that the discrete linear chirp transform has better performance than the fractional Fourier transform in terms of sparsity, computation time, and peak location.

In [18], a DLCT denoising algorithm which is used to obtain an estimate for the desired signal is proposed. The algorithm relies on the ability of the DLCT to decompose a signal iteratively into its components locally. Each of these components is filtered separately and then synthesized with the other filtered components to give the filtered signal. Since each segment of the signal has different components with different bandwidths, the filter has to be time–varying filter. The results show that the DLCT algorithm provides better performance than the FrFT algorithm.

In this paper a novel denoising method based on the discrete linear chirp transform is proposed. The designed filter approach uses the minimization criterion between the desired signal and the filtered signal to obtain optimal filter coefficients. The presented method shows an improvement in the performance compared with the conventional DLCT algorithm given in [18] for mono–component chirp signals.

The rest of the article is organized as follows. Section II presents the discrete linear chirp transform. In Section III, the proposed filter design technique is introduced. Simulation results are given in Section IV. Conclusions and future work are shown in Section V.

2. Discrete Linear Chirp Transform

Consider a discrete time chirp based analytic signal $x(n)$, where $n = 0, 1, \dots, N - 1$ and N is the number of samples embedded in a complex white Gaussian noise $\eta(0, \sigma^2)$. Then the observation signal $y(n)$ can be modeled as

$$y(n) = x(n) + \eta(n) \quad (1)$$

The signal $x(n)$ can be well estimated using the DLCT [2], which its pairs are given as follows

$$X(k, m) = \sum_{n=0}^{N-1} x(n) \exp\left(-j\frac{2\pi}{N}(c m n^2 + k n)\right) \quad (2)$$

$$x(n) = \frac{1}{N} \sum_{m=-L/2}^{L/2-1} \sum_{k=0}^{N-1} X(k, m) \exp\left(j\frac{2\pi}{N}(c m n^2 + k n)\right) \quad (3)$$

where C is the resolution of the transform in the chirp–rate domain and L is an even integer number as defined in [2]. The DLCT attempts to decompose a signal using discrete linear chirps characterized by a discrete frequency $2\pi k/N$, and a chirp–rate $\beta = C m$.

For certain chirp–rate domain, we can define the DLCT pairs in the matrix form as

$$\mathbf{X} = \mathbf{D}_\beta \mathbf{x} \quad \text{and} \quad \mathbf{x} = \mathbf{D}_{-\beta} \mathbf{X} \quad (4)$$

where \mathbf{X}, \mathbf{x} are $N \times 1$ vectors and $\mathbf{D}_\beta, \mathbf{D}_{-\beta}$ are $N \times N$ matrices given as

$$\mathbf{D}_\beta(k, n) = \exp\left(-j\frac{2\pi}{N}(kn + \beta n^2)\right) \quad (5)$$

and

$$\mathbf{D}_{-\beta}(k, n) = \exp\left(j\frac{2\pi}{N}(kn + \beta n^2)\right) \quad (6)$$

Since the DLCT is a unitary transformation, the DLCT matrices in (5) and (6) are related by $\mathbf{D}_\beta = \mathbf{D}_{-\beta}^H$, where $(\cdot)^H$ denotes the conjugate transpose operation, $\mathbf{D}_\beta \mathbf{D}_{-\beta}^H = \mathbf{I}$, and \mathbf{I} is the identity matrix.

3. Proposed Filter Design Technique Using the DLCT

It is well known that the optimal filter design for stationary process time-invariant signals can be achieved with Wiener filter which can be implemented efficiently using fast Fourier transform. However, for time-varying nonstationary signals such as chirp signals, we need a transform that can handle these time variations in a more efficient way. Therefore, the DLCT will be used to represent such signals. In this paper, we will adopt a similar criteria used in Wiener filter design which is the minimization of the average square error.

Our design problem requires that we find the filter response, \mathbf{g} , that minimizes the average square error over M realizations as follows

$$\mathbf{g} = \arg \min_{\mathbf{g}} \frac{1}{M} \sum_{i=1}^M \|\hat{\mathbf{x}}_i - \mathbf{x}\|^2 \quad (7)$$

where $\hat{\mathbf{x}}_i$ is the DLCT inverse for the multiplication of the observation signal \mathbf{y}_i with the filter response \mathbf{g} as

$$\hat{\mathbf{x}}_i = \mathbf{D}_{-\beta} \mathbf{G} \mathbf{D}_\beta \mathbf{y}_i \quad (8)$$

and \mathbf{G} is an $N \times N$ diagonal matrix whose elements are $\mathbf{g} = \text{diag}(\mathbf{G}) = [g_0, g_1, \dots, g_{N-1}]$. If we let $\mathbf{s}_i = \mathbf{D}_\beta \mathbf{y}_i$ and since \mathbf{G} is a diagonal matrix, then we can write $\mathbf{G} \mathbf{s}_i = \mathbf{S}_i \mathbf{g}$, where \mathbf{S}_i is an $N \times N$ diagonal matrix defines as $\mathbf{s}_i = \text{diag}(\mathbf{S}_i)$, given that $\bar{\mathbf{S}}_i = \mathbf{D}_{-\beta} \mathbf{S}_i$. Thus, we can rewrite our quadratic cost function $J(\mathbf{g})$ such as

$$J(\mathbf{g}) = \frac{1}{M} \sum_{i=1}^M (\bar{\mathbf{S}}_i \mathbf{g} - \mathbf{x})^H (\bar{\mathbf{S}}_i \mathbf{g} - \mathbf{x}) \quad (9)$$

or equivalently,

$$J(\mathbf{g}) = \frac{1}{M} \sum_{i=1}^M (\mathbf{g}^H \bar{\mathbf{S}}_i^H \bar{\mathbf{S}}_i \mathbf{g} - 2\mathcal{R}(\mathbf{x}^H \bar{\mathbf{S}}_i \mathbf{g}) + \mathbf{x}^H \mathbf{x}) \quad (10)$$

where $\mathcal{R}(\cdot)$ is the real part. Equation (10) can be expressed in the following form

$$J(\mathbf{g}) = \mathbf{g}^H \mathbf{Q} \mathbf{g} - 2\mathbf{b}^T \mathbf{g} + a \quad (11)$$

where,

$$\mathbf{Q} = \frac{1}{M} \sum_{i=1}^M \bar{\mathbf{S}}_i^H \bar{\mathbf{S}}_i,$$

$$\mathbf{b} = \frac{1}{M} \sum_{i=1}^M [2\mathcal{R}(\mathbf{x}^H \bar{\mathbf{S}}_i)]^T, \text{ and}$$

$$a = M \|\mathbf{x}\|^2$$

The optimal response of the filter \mathbf{g}_{opt} is the solution of minimization the cost function $J(\mathbf{g})$ and this can be done by computing the gradient of (11) with respect to \mathbf{g} , which yields

$$\mathbf{g}_{opt} = \mathbf{Q}^{-1}\mathbf{b} \tag{12}$$

Since \mathbf{Q} is a diagonal matrix, its inverse always exists. So far we consider only the minimization of the average square error with respect to the filter response \mathbf{g} . Now, how do we choose the other parameter in the cost function which is the chirp-rate β ? Trying to find the optimal β analytically is not an easy problem to solve. In this paper, we use an approach based on the DLCT. We compute the DLCT of the observation signal and find the chirp-rate that maximizes it.

In (12), the optimal solution \mathbf{g}_{opt} depends on the desired signal \mathbf{x} . An estimate for the desired signal can be obtained using the algorithm given in [18].

4. Results and Discussion

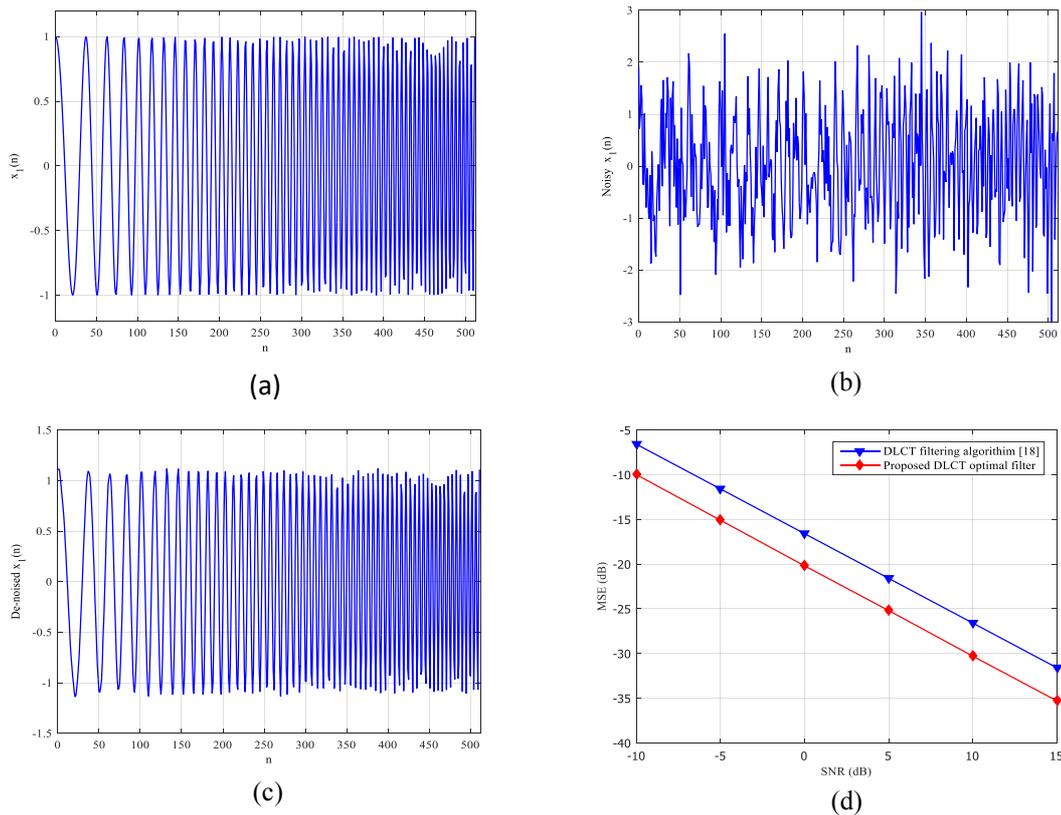


Figure 1: Linear chirp signal: (a) the chirp signal, (b) the noisy chirp signal with SNR=0 dB, (c) the filtered chirp signal using the optimal DLCT filtering method, (d) the mean square error for the DLCT and optimal DLCT methods.

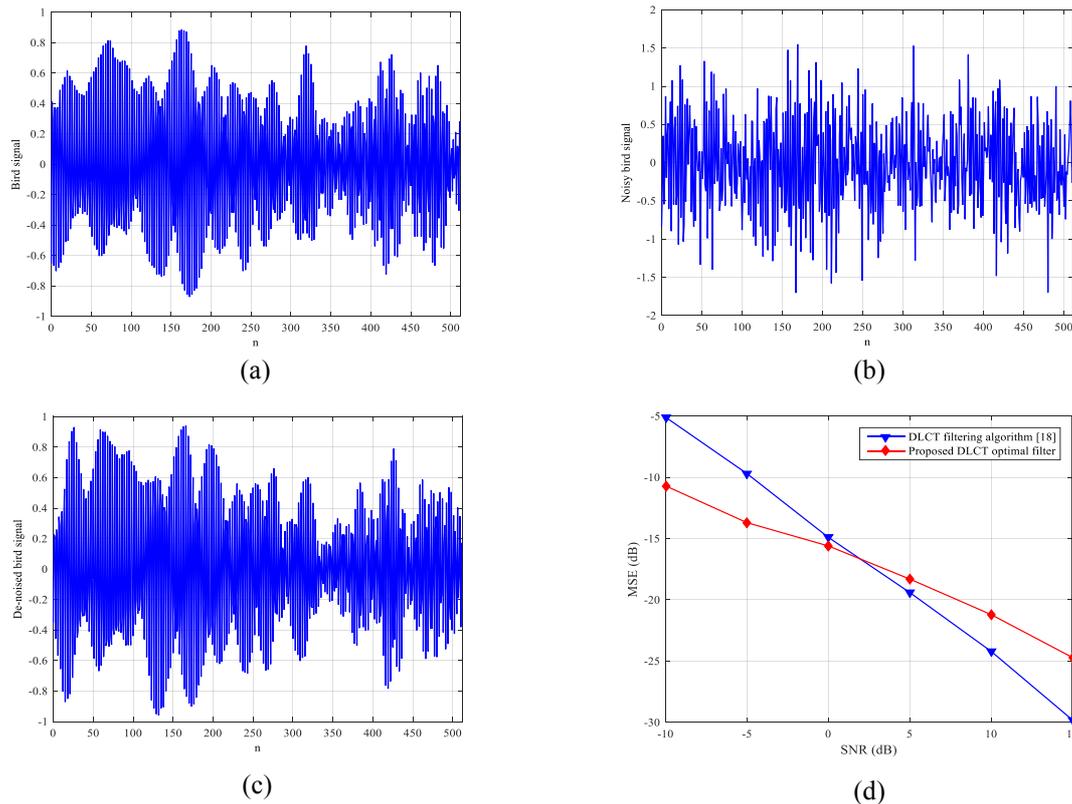


Figure 2: Real-world signal: (a) the bird chirping signal, (b) the noisy bird chirping signal with SNR=0 dB, (c) the filtered bird chirping signal using optimal DLCT filtering method, (d) the MSE for various methods.

To demonstrate the performance of the proposed DLCT denoising algorithm, two examples are performed using a synthetic signal and a real-world signal. In each example, we compare the mean square error (MSE) of the proposed method with the DLCT filtering algorithm given in [18]. The estimated signal using the DLCT algorithm is used as the desired signal for the proposed method with a set of 20 realizations.

For the case of synthetic signal, we use the signal given in Figure 1(a). Figure 1(b) and (c) show the noisy chirp signal, and the denoised chirp signal using optimal DLCT filtering method, respectively. In Figure 1(d), we present the MSE for various algorithms against signal-to-noise ratio (SNR). It can be seen that the proposed optimal filtering algorithm outperforms the performance of the DLCT filtering algorithm.

To quantify the MSE improvement, a real-world signal (bird chirping signal) with varying noise level is also simulated. The noiseless and corresponding corrupted bird chirping signals are presented in Figure 2(a) and (b). The denoised bird chirping signal based on the optimal DLCT filtering algorithm is shown in Figure 2(c) at SNR=0 dB. Similar to the previous example, Figure 2(d) depicts the MSE for the proposed method compared with the conventional DLCT filtering algorithm [18] as a function of the input SNR, where input SNR is varied from -10 dB (severely poor SNR) to 15 dB (high SNR). The DLCT filtering algorithm greatly enhances the MSE for low SNR levels. It is shown that for proposed filtering method, the performance is not optimal since it achieves worse MSE at high SNR. This is because the signal has many components with different chirp rates.

5. Conclusions

In this paper, an optimal filtering method based on the discrete linear chirp transform is proposed. The design of the filter is carried out and a closed form solution for the impulse response of the filter is derived. The performance of the proposed filter is compared with other algorithms. Simulation results show that the DLCT optimal filtering method outperform the performance of the conventional DLCT filtering algorithm for mono–component chirp signals. However, for multi–component chirp signals, the proposed filter gives better performance in low signal–to–noise ratio environment. As a future work, we will look for another approach that can deal with multicomponent signals contaminated in noise more optimally.

References

- [1]. O. A. Alkishriwo, L. F. Chaparro, and A. Akan, “Signal separation in the Wigner distribution using fractional Fourier transform,” European Signal Processing Conf., EUSIPCO, Spain, September 2011, pp. 1879-1883.
- [2]. S. Peleg and B. Friedlander, “The discrete polynomial-phase transform,” IEEE Transactions on Signal Processing, vol. 42, no. 8, pp. 1901-1914, August 1995.
- [3]. M. Z. Ikram, K. Abed-Meraim, and Y. Hua, “Fast quadratic phase transform for estimating the parameters of multicomponent chirp signals,” Digital Signal Processing, vol. 7, no. 2, pp. 127-135, 1997.
- [4]. O. A. Alkishriwo and L. F. Chaparro, “A Discrete Linear Chirp Transform (DLCT) for Data Compression,” in Proc. of the IEEE International Conf. on Information Science, Signal Processing and their Applications, Montreal, Canada, July 2012, pp. 1283-1288.
- [5]. Y. Li, H. Fu, and P. Y. Kam, “Improved, approximate, time–domain ML estimators of chirp signal parameters and their performance analysis,” IEEE Transactions on Signal Processing, vol. 57, no. 4, pp. 1260-1272, April 2009.
- [6]. Q. Yin, S. Qian, and A. Fing, “A fast refinement for adaptive Gaussian chirplet decomposition,” IEEE Transactions on Signal Processing, vol. 50, no. 6, pp. 1298-1306, June 2002.
- [7]. Y. Lu, R. Demirli, G. Cardoso, and J. Saniie, “A successive parameter estimation algorithm for chirplet signal decomposition,” IEEE Transactions on Signal Processing, vol. 53, no. 11, pp. 2121-2131, November 2006.
- [8]. Z. K. Peng, G. Meng, F. L. Chu, Z. Q. Lang, W. M. Zhang, and Y. Yang, “Polynomial chirplet transform with application to instantaneous frequency estimation,” IEEE Transactions on Instrumentation and Measurement, vol. 60, no. 9, pp. 3222-3229, September 2011.
- [9]. V. Namias, “The fractional order Fourier transform and its application to quantum mechanics,” IMA Journal of Applied Mathematics, vol. 25, , no. 3, pp. 241-265, 1980.
- [10]. L. B. Almeida, “The fractional Fourier transform and time–frequency representations,” IEEE Transactions on Signal Processing, vol. 42, no. 11, pp. 3084-3091, November 1994.
- [11]. H. M. Ozaktas, Z. Zalevsky, and M. A. Kutay, The Fractional Fourier transform with applications in optics and signal processing, John Wiley, New York, USA, 2001.
- [12]. E. Sejdic, I. Djurovic, and L. Stankovic, “Fractional Fourier transform as a signal processing tool: an overview of recent developments,” Signal Processing, vol. 91, no. 6, pp. 1351-1369, 2011.
- [13]. L. Durak and S. Aldirmaz, “Adaptive fractional Fourier domain filtering,” Signal Processing, vol. 90, no. 4, pp. 1188-1196, April 2010.
- [14]. K. Prajna and C. K. Mukhopadhyay, “Fractional Fourier transform based adaptive filtering techniques for acoustic emission signal enhancement,” Journal of Nondestructive Evaluation, vol. 14, pp. 1-15, January 2020.
- [15]. L. Wu, Y. Zhao, L. He, S. He, and G. Ren, “A time–varying filtering algorithm based on short–time fractional Fourier transform,” International Conference on Computing, Networking, and Communications (ICNC), Big Island, Hi, USA, February 2020, pp. 555–560.



-
- [16]. O. A. Alkishriwo and L. F. Chaparro, "Signal compression using the discrete linear chirp transform (DLCT)," in Proc. of IEEE 20th European Signal Processing Conference (EUSIPCO 2012), Bucharest, Romania, Aug. 2012, pp. 2128-2132.
 - [17]. O. A. Alkishriwo, "The discrete linear chirp transform and its applications," PhD Thesis, University of Pittsburgh, 2013.
 - [18]. O. A. Alkishriwo, A.A. Elghariani, A. Akan, "Iterative time-varying filter algorithm based on discrete linear chirp transform," First Conference for Engineering Sciences and Technology (CEST-2018), Garaboulli, Libya, pp. 86-92, September 2018.

Optimized Maximum Loadability of Power Systems using an Enhanced Dynamic JAYA Algorithm

Ibrahim A. Farhat^{1*}, Abdullah O. Hawal²

¹ ibrahimfarhat@yahoo.com, ² aohawal@elmergib.edu.ly

¹ Electrical and Computer Engineering Department, Faculty of Engineering, Al-Asmarya University, Libya

² Electrical and Computer Engineering Department, Faculty of Engineering, Elmergib University, Libya

ABSTRACT

The problem of Maximum Loadability of power systems is addressed in this paper using a proposed dynamic JAYA algorithm. The maximum loadability problem is a typical optimization problem in which the maximum loadability point is to be determined optimally. Voltage stability of power systems is maintained by determining the estimated margin between the system operating point and the maximum loadability limit. The basic JAYA algorithm has been introduced to solve foremost optimization problems with small-scaled nature. However, when applied to large-scale, nonlinear and non-convex constrained problems, it showed a poor convergence characteristics. In order to deal with these weaknesses, the original algorithm has been improved by adding some dynamic features to its convergence behavior. The modified algorithm has been presented and validated when applied to well-known typical power systems. The obtained results were compared to the results achieved by other equivalent optimization techniques.

Keywords:

JAYA algorithm,
maximum
loadability limit,
Power system
optimization,
Voltage stability.

1. Introduction

Voltage instability is one of the common operation problems accompanying power system networks as a result of various operating conditions. To deal with this issue and maintain voltage stability after the occurrence of disturbances in the system, the state of equilibrium has to be restored successively. One of the most effective reasons behind the voltage instability is the extensively stressed and heavily loaded systems. Shortage of reactive power supply that does not satisfy the demand is another possible reason. Switching problems, unscheduled outages in addition to poor system voltage profiles can also cause voltage instability and lead the system to loose equilibrium. Obviously, voltage collapse can be experienced in different parts of the network [1]. Although they are linked to each other, maximum loadability and voltage stability concepts should not be mixed up. One should think of maximum loadability problem considering the system static characteristics and not limited by the voltage stability aspects. Nevertheless, the closer the system operating point to the maximum loadability limit, the more likely the voltage instability to occur [2]. Figure 1 shows the P-V curve demonstrating the nose point of the maximum loading point.

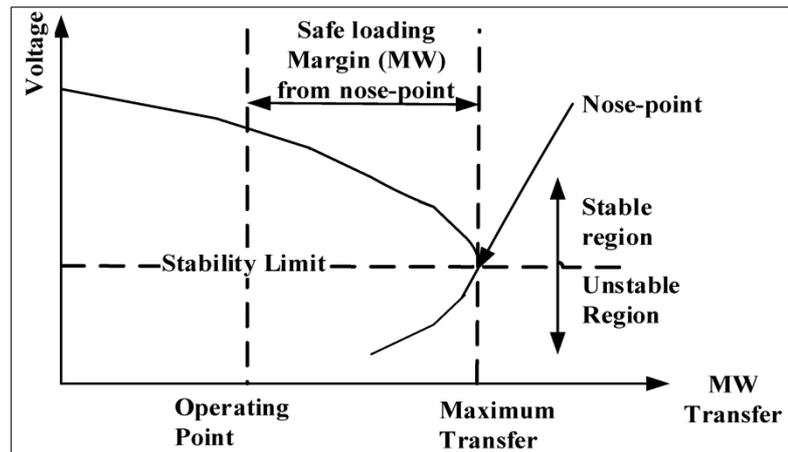


Figure 1: P-V curve and maximum loadability limit

As shown in the figure, the system operates in the upper part of the curve with static and dynamic stability characteristics are feasible. In this region, the curve has a high voltage-low current profile. The maximum loadability limit point defines the voltage collapse point behind which the system loses equilibrium [1]. Optimization methods in general can be classified as classic deterministic and heuristic non-deterministic techniques. Many of these have been applied to solve the maximum loadability optimization problem. Among the deterministic calculus-based approaches are the Interior Point method and the Sequential Quadratic Programming algorithms [3, 4]. More recent heuristic methods have also been utilized successfully to solve this problem. These include Particle Swarm Optimization (PSO) [5], Ant Colony Optimization (ACO) [6] and Genetic Algorithm (GA) [7]. The JAYA algorithm is one of the heuristic non-calculus-based approaches that have been applied for optimization problems. This algorithm has recently introduced as a simple and non-deterministic optimization technique [8]. In spite of its popularity and simplicity due to the limited number of parameters it requires, the basic JAYA algorithm suffers from significant convergence issues. This poor convergence behavior can be observed when applying the basic algorithm to nonlinear, nonsmooth high-scaled optimization problems with nonlinear constraints. The major weakness is the divergence to local minima instead of converging to the global. It was observed through the experience and trials that some population diversity preservation issues were the reason behind this deficiency. In this paper, a modified adaptive JAYA algorithm, MAJAYA, is presented and applied to determine the maximum loadability limit of power systems. The remainder of the paper is organized as follows: Section 2 provides the formulation of the problem. In Section 3, the MAJAYA is described. Simulation results are demonstrated in Section 4. The conclusion is drawn in Section 5.

2. Maximum Loadability

The optimization problem of the Maximum Loadability is formulated as a nonlinear constrained optimization problem [3].

2.1 Objective Function

The objective of the problem is to find the maximum loading of a power system taking into consideration the operational constraints of the system [3]. The maximum loadability problem is formulated as follows:

$$\text{Max } f = \lambda \quad (1)$$

where λ is the load incremental parameter with reference to the current operating point of the system. This parameter is bounded by its *initial* value which is 0 and its upper limit $\lambda_{collapse}$ at the voltage collapse point. Accordingly, the power demand of the buses increases instantaneously as follows:

$$P_{di} = P_{dio} + \lambda P_d \quad (2)$$

$$Q_{di} = Q_{dio} + \lambda Q_d \quad (3)$$

P_{di} and Q_{di} are the active and reactive power at the i^{th} load bus, while P_{dio} and Q_{dio} are the *initial* active and reactive power at the i^{th} load bus.

2.2 Constraints

The system constraints include bus voltage magnitude limits, power generation upper and lower limits, switchable capacitor limits and transformer tap changer limits.

- Load balance

$$P_{gi} - P_{di} = |V_i| \sum_{j=1}^N |Y_{ij}| |V_j| \cos(\delta_i - \delta_j - \theta_{ij}) \quad (4)$$

$$Q_{gi} - Q_{di} = |V_i| \sum_{j=1}^N |Y_{ij}| |V_j| \sin(\delta_i - \delta_j - \theta_{ij}) \quad (5)$$

where the number of buses is N and the voltage profile of the i^{th} bus is $|V_i|$ and δ_i ; while the i^{th} element of the system's Y_{bus} matrix is $|Y_{ij}|$ and θ_{ij} .

- Power generation upper and lower limits

$$P_{gi,min} \leq P_{gi} \leq P_{gi,max} \quad (6)$$

$$Q_{gi,min} \leq Q_{gi} \leq Q_{gi,max} \quad (7)$$

- Bus voltage magnitude and angle limits

$$|V|_{min} \leq |V|_i \leq |V|_{max} \quad (8)$$

$$\delta_{i,min} \leq \delta_i \leq \delta_{i,max} \quad (9)$$

- Transformer tap changer limits

$$T_{k,min} \leq T_k \leq T_{k,max} \quad (10)$$

- Switchable capacitor limits

$$Q_{ci,min} \leq Q_{ci} \leq Q_{ci,max} \quad (11)$$

In the above equations; the number of generating units is N_g , the number N_b is the number of system buses, N_t is the number of tap changing transformers and N_s is the number of switchable shunt capacitors.

3. The JAYA Algorithm

This section is divided into two parts. In the first the basic JAYA is explained. The modified JAYA algorithm is demonstrated in the second part.

3.1 The Basic JAYA Algorithm

The original JAYA algorithm is a deterministic heuristic optimization method that was introduced by R. Venkata [8] recently. JAYA is a Sanskrit word which means victory. This approach has been effectively applied to solve a number of constrained optimization problems. The very limited number of control parameters required for this algorithm is one of its good features. A short explanation on how the JAYA algorithm can be implemented is shown as follows:

- 1- The population size, number of variables for decision and stopping criteria are defined and the process is initialized.
- 2- Population of size P (candidate solution) \times q (decision variables) is generated.
- 3- An initial solution is determined.
- 4- Best and worst solutions ($X^{k_i, best}$, $X^{k_i, worst}$) are computed.

- 5- The solution vector is updated as follows:

$$X_{i,j}^{k+1} = X_{i,j}^k + r_1(X_{i,best}^k - |X_{i,j}^k|) - r_2(X_{i,worst}^k - |X_{i,j}^k|) \quad (12)$$

where, r_1 and r_2 are random numbers within the interval [0, 1].

- 6- Solution candidates are compared to check if the updated solution is better than the previous one or not. The update is accepted if the new candidate is better, otherwise it is rejected.
- 7- The stopping criterion is checked and applied so that the algorithm is terminated when satisfied or otherwise return to update step 2.

The JAYA algorithm requires identifying only the maximum iteration in addition to the size of population. It can also be seen from Equation (12) that the candidate moves closer to the best solution and pushes away from the worst one.

3.2 The Modified Adaptive JAYA Algorithm

In order to deal with the hitches linked to the basic algorithm, some changes are suggested in this section. These modifications are introduced to improve the algorithm's behaviour when applied for large nonlinear nonsmooth objective functions.

The proposed adjustment is to dynamically update the population size. Accordingly, once the initial population size is chosen, then the population size is updated adaptively according to the following mechanism [9]:

$$P_{new} = P_{old} \times round(1 + s) \quad (13)$$

In the above equation s is an arbitrary variable such that $-1 < s < 1$. Consequently, in the modified adaptive algorithm (MAJAYA), all the population elements will be updated to the next population vector if the new size is larger than the old one and the optimal solution will be assigned to the remaining candidates. On the other hand, only the best solution will be moved to the next population if the old size is larger than the new one. Logically, no action is needed if there is no change in the size.

4. Simulation Results

The proposed MAJAYA was employed to determine the maximum (optimal) loadability for two well-known IEEE 30 bus and IEEE 118 bus power systems. The algorithm was implemented and coded in MATLAB and executed on an Intel Core i7-8750H 2.20GHz personal computer with 8 GB RAM. In order to check for consistency, 50 independent runs were conducted with different random initial solution for each run. Results obtained were compared with those of some other methods. The various algorithm parameters were tuned independently since they were problem-dependent. A considerable number of preliminary runs were executed individually so that the optimal parameter combination was obtained.

The IEEE 30 bus system consists of six generating units, four transformers and 41 branches. The system configuration can be found in [10]. Results obtained for maximum bus voltage amplitude and angle at the maximum loading point are shown in Table 1.

Table 1: 30 bus system: Bus voltages at the maximum loading.

Bus No.	Bus Voltage		Bus No.	Bus Voltage	
	V pu	δ°		V pu	δ°
1	1.0000	0.0000	16	0.9888	-12.0034
2	0.9998	-3.4534	17	0.9829	-12.7768
3	0.9758	-4.5678	18	0.9890	-12.8876
4	0.9897	-4.7856	19	0.9756	-12.8899
5	0.9876	-5.8783	20	0.9987	-12.8987
6	0.9843	-5.9723	21	0.9991	-12.6589
7	0.9887	-7.7685	22	0.9978	-13.0000
8	0.9954	-8.3334	23	0.8936	-13.1432
9	0.9865	-9.7683	24	0.9778	-12.8798
10	1.0000	-10.8721	25	0.9878	-11.5678
11	0.9788	-10.9867	26	0.9788	-11.5453
12	0.9878	-10.9999	27	1.0000	-11.3485
13	0.9877	-11.4638	28	0.9789	-11.5456
14	0.9856	-11.4537	29	1.0000	-12.7776
15	0.9867	-11.6665	30	0.9819	-12.6799

The proposed algorithm was once again applied to the IEEE 118 bus test system to confirm its effectiveness. The diagram of the testing system can be found in [10]. This system is made up of 54 generation buses, 64 load buses, and 186 branches. It also has 9 transformers and 14 switchable capacitor bank.

The results were compared to those obtained by some other evolutionary techniques previously employed to solve the problem. These are Differential Evolution (DE) [11], Multi Agent-based Hybrid Particle Swarm Optimization (MAHPSO) [12], Cuckoo Search Algorithm (CSA) [13] and Dynamic Modified Bacterial Foraging Algorithm (DMBFA) [14]. The comparison is demonstrated in Table 2.

Table 2: Comparison of the results.

Method	Maximum power (pu)	
	30-bus system	118-bus system
MAHPSO	2.6081	56.45
DE	2.6709	56.543
CSA	2.8396	62.5671
DMBFA	2.9284	62.9865
MAJAYA	2.9307	62.9888

5. Conclusions

In this paper, a modified adaptive JAYA algorithm was applied to determine the maximum loadability of power systems. This problem was addressed as an optimization problem and

formulated as so. The system's operation constraints were taken into account. These included voltage magnitude limits, power generation upper and lower limits, switchable capacitor limits and transformer tap changer limits. The adaptive algorithm was modified to update the solution vector and enhance the convergence properties of the algorithm. Simulation results demonstrated the robustness and effectiveness of the algorithm for the maximum loadability problem. The algorithm was tested and applied to two well-known IEEE power systems. Comparison with some selected heuristic optimization methods showed that the applied algorithm has accomplished good and satisfying results compared to those obtained by the other techniques.

References

- [1]. P. Kundur, "Power System Stability and Control", McGraw –Hill, 1994.
- [2]. K. R. Padiyar, "Power System Dynamics: Stability and Control", BSP, 2nd ed., 2008.
- [3]. G. D. Irisarri, X. Wang, J. Tong, S. Moktari, "Maximum loadability of power systems using interior point non-linear optimization method," IEEE Trans. Power Systems, vol. 12, no. 1, Feb. 1997, pp. 162–172.
- [4]. V. Ajarapu and C. Christy, "The continuation power flow: A tool for steady state voltage stability analysis", IEEE Trans. Power Systems, vol. 7, no. 1, Feb. 1992, pp. 416-423.
- [5]. L. D. Arya, S. C. Choube, M. Shrivastava, and D. P. Kothari, "Loadability margin enhancement using coordinated aggregation based particle swarm optimization (CAPSO)" Int. J Electr Power Energy Syst., vol. 32, no. 9, 2010, pp. 975-984.
- [6]. M. R. Kalil, I. Musirin, and M. M. Othman, "Ant colony optimization for maximum loadability search in voltage control study," Conference 2006 IEEE International Power and Energy, Nov. 2006 , Putra Jaya , pp. 240 – 245.
- [7]. P. Acharjee, "Identification of Maximum Loadability Limit and Weak Buses using Security Constraint Genetic Algorithm", International Journal of Electrical Power and Energy Systems, vol. 36, pp. 40-50, 2012.
- [8]. R. Venkata Rao, "Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems," International journal of industrial computations vol.7, pp. 19-34, 2016.

Optimum Backoff in CSMA-CA for Ad-Hoc Vehicles Network in Motorway

Sami F. Alahmar ^{1*}, Omar O. Aldawibi ²

¹ sami.alahmer1989@gmail.com , ² omar.aldawibi@hpiz.edu.ly

¹ Department of Electrical and Computer Engineering , Faculty of Engineering,
Elmergib University, Libya

² Department of Electrical and Electronics Technology, the Higher Institute for Engineering
Technology, Zliten – Libya

ABSTRACT

With current technological developments, wireless networks are becoming popular. VANET is a type of MANET that allows data to be transferred between nearby vehicles. These types of communications can help prevent accidents and investigate post-crash accidents or traffic jams by allowing vehicles to share and broadcast safety information with other vehicles to alert drivers. VANETs offer many possibilities for many new applications. This paper provides an evaluation of mobility influence on the initial backoff contention windows performance. The study evaluates basic performance metrics such as packet delivery ratio, throughput and average end-to-end-delay by using the network simulator (NS-2). They propose four dynamic value initial backoff of contention windows mechanisms to alleviate network performance degradation due to high mobility. The nodes are running in the same direction and at constant speeds in terms of a varying number of initial backoff of contention windows and to determine the influence routing protocol.

Keywords:

Ad-hoc Networks;
Routing protocols;
Network
Simulator.

1. Introduction

Mobile Ad-hoc Network (MANET) is collected of wireless mobile nodes interconnected with each other through an autonomous configuration in the absence of any infrastructure. Vehicular Ad-hoc network is a particular kind of MANET, where smart vehicles act as nodes [1], [2], [3].

Vehicular Ad-hoc Network (VANET) has become one of the most important research areas in the field of wireless communication. Vehicle to Vehicle data transfer is one of the main challenges within the design of VANET because it needs to design a dynamic routing protocol. Routing in traditional MANET is different from the VANET routing because of extremely dynamical topologies [4], [5].

This paper discusses the Optimum backoff window for Ad-hoc Vehicles Network in Motorway in terms of the mobility influence and of the total packet delivery ratio, throughput and the minimum end-to-end delay. Also, two ray ground channel has been used in the NS-2 simulator.

2. Media Access Protocols

The Medium Access Control sublayer (MAC) forms the lower half of the data link layer. It directly interfaces with the physical layer. It provides services such as addressing, framing and Medium Access Control. Unlike the LLC, these services vary with the physical medium in use. Of these, medium access control is considered to be the most important service. It is relevant to networks (such as LANs) where a single broadcast transmission channel needs to be shared by multiple competing machines [6].

2.1 ALOHA

The ALOHA protocol [7] is one of the oldest multiple access mechanisms. ALOHA had its origins in the Hawaiian Islands. It was borne out of the need for interconnecting terminals at the campuses of the University of Hawaii, which were located on different islands. It was devised in the 1970s at the University of Hawaii. The original ALOHA system is referred to as pure ALOHA. Roberts extended this system and developed what is called the slotted ALOHA system, which doubled the throughput of the pure ALOHA system [8].

2.2 Carrier Sense Media Access (CSMA)

The maximum achievable throughput in the ALOHA protocols is low because of the wastage of bandwidth due to packet collisions. Packet collisions could be reduced by having the nodes sense for the carrier signal on the channel before they actually start transmitting. Carrier sense multiple access (CSMA) protocols are those in which nodes, before transmitting, first listen for a carrier (i.e., transmission) on the channel, and make decisions on whether or not to transmit based on the absence or presence of the carrier.

2.3 CSMA with Collisions Avoidance (CSMA/CA)

The Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) was designed for the popular Wi-Fi wireless network technology (IEEE 802.11). CSMA/CA also senses the transmission channel before transmitting a frame. Furthermore, CSMA/CA tries to avoid collisions by carefully tuning the timers used by CSMA/CA devices [8].

3. Backoff Window Algorithm

In this paper, compare the capabilities of varied initial backoff contention window algorithms and better performance-based on metrics.

The proposed exponential backoff mechanism employed in the CSMA protocol is a slotted binary algorithm [9]. Any node desiring to send data utilizes both the physical and virtual carrier sense functions to determine the state of the medium. The node senses the channel; if the channel is free it will start sending, otherwise, the node should defer transmission. After DIFS time, the node generates a random backoff period. This procedure minimizes the probability of collisions during contention. The backoff time can be calculated using the following equation:

$$\text{Backoff_time} = \text{int}(\text{CW} * \text{Random}(0,1) * \text{slot_time}) \quad (1)$$

Where Backoff_time is the time that must be deferred before the node starts the transmission, CW is the contention window which is an integer between CWmin and CWmax, Random (0,1) is a pseudo-random number between 0 and 1, slot_time is the duration time, and int is the integer part of the equation [10].

The CW starts initially at CWmin and takes its next value every time there is an unsuccessful transmission. When CW reaches CWmax, it remains at this value until it is reset. Any node that performs backoff senses the channel during each slot_time; if the channel state sensed is idle during that slot_time, the backoff procedure decreases its backoff time by the amount of slot_time. If the channel state sensed is busy after that, the backoff procedure is frozen for that time slot until the channel is to be sensed again. The channel should be sensed as idle for DIFS time before the backoff procedure is allowed to resume [11]. The node should start transmission whenever the backoff timer reaches zero, as shown in Figure1.

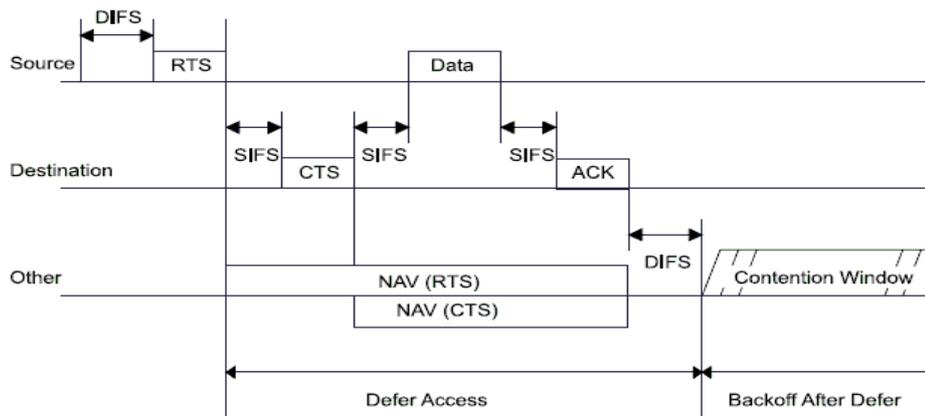


Figure 1: CSMA/CA Protocol Mechanism

4. Network Simulation

The network simulator [NS-2] is discrete event simulation software for network simulation.

However, NS-2, an open software that has been built by a number of different developers, suffers from a number of known and unknown bugs [12].

5. Model Parameters

The parameters were different routing protocols like as AODV, DSDV and DSR are chosen for simulation using the performance metrics such as packet delivery ratio, throughput and average end-to-end delay in different scenarios i.e., for 8_16_32 and 64 slots in the initial backoff connection window as show in the Table 1.

Table 1: Simulation Parameters

Parameters	Details
Simulator	NS2.35
Area of simulation	250 m * 250 m
Node Placement	Moving at Same
MAC protocol	802.11
Radio Propagation	Two Ray Ground
Routing Protocol	AODV,DSR,DSDV
Simulation Time	150 sec
Packet Size	1000 Byte
Number of Vehicles	8
Transmission Rate	1 Mbps
Min Speed	15 m/s
Max Speed	45 m/s
Traffic Type	TCP
Initial Energy	1000 J

6. Performance Metrics

The simulations were performed using Network Simulator (Ns-2), which is popularly used for Ad-hoc networking community [13], [14]. The routing protocols were compared based on the following 3 performance metrics:

- Packet Delivery Fraction (PDF), The ratio of data packets delivered to those generated by the sources [%].
- Total throughput is the total number of packets successfully received per unit of time [bps].
- Average end-to-end delay : is the delay calculated by averaging the time that needed for each data packet to be transmitted from the source to its final destination [sec].

7. Results and Discussion

This section discusses the influence of different initial backoff contention windows in motorway on different time slots for mobile Ad-hoc networks.

Figures 2,3,4,5,6,7,8,9 and 10 show the performance of different routing protocols for two sources versus the packet delivery ratio [%], the total throughput [bps], average end to end delay [sec] respectively; in a variety of number of speed.

Figure 2 and 3 shows the packet delivery ratio and throughput of the AODV routing is the best performance with the initial backoff contention window (equal 15 slots), and varying number of speed (from 15 to 45 m/s) at the same direction.

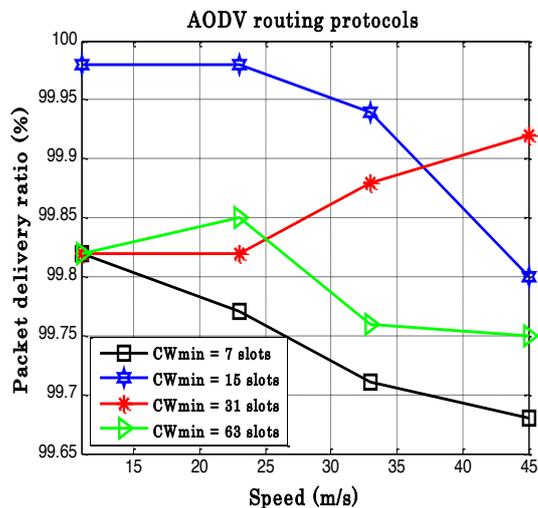


Figure 2: Packet Delivery Ratio (%)

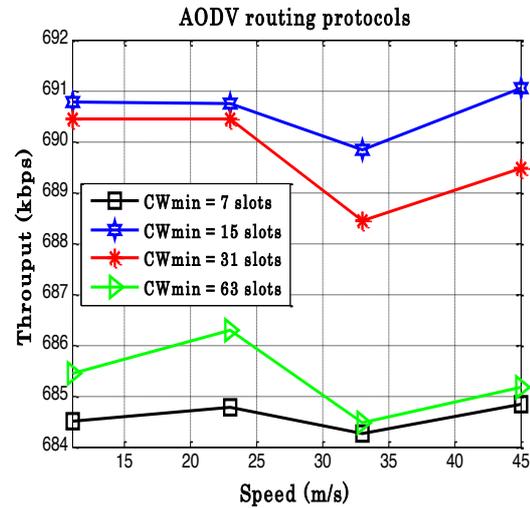


Figure 3: Throughput (kbps)

Figure 4 shows the average end to end delay of AODV routing is the best performance with the initial backoff contention window (equal 7 slots).

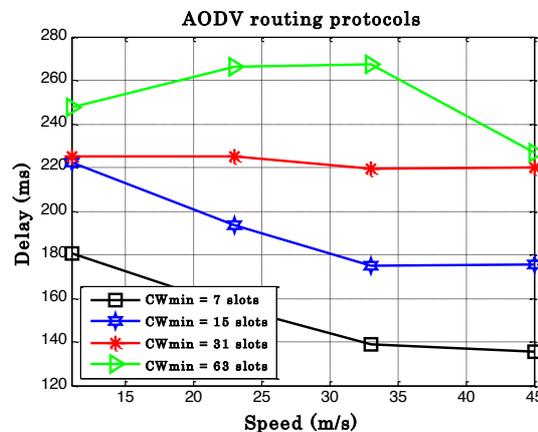


Figure 4: End to End Delay (ms)

Figure 5 shows the AODV routing outperformance DSR routing in low mobility situation with the initial backoff contention window (equal 7 slots). On the other hand in high mobility situation, the DSR routing outperformance AODV routing with the initial backoff contention window (equal 31 slots).

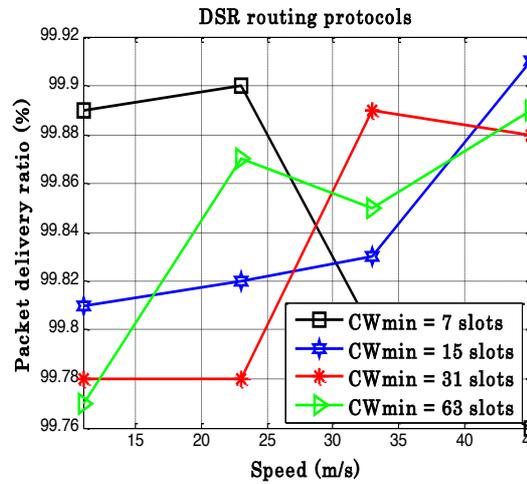


Figure 5: Packet Delivery Ratio (%)

Figure 6 shows the throughput of DSR routing is the best performance with the initial backoff contention window (equal 31 slots).

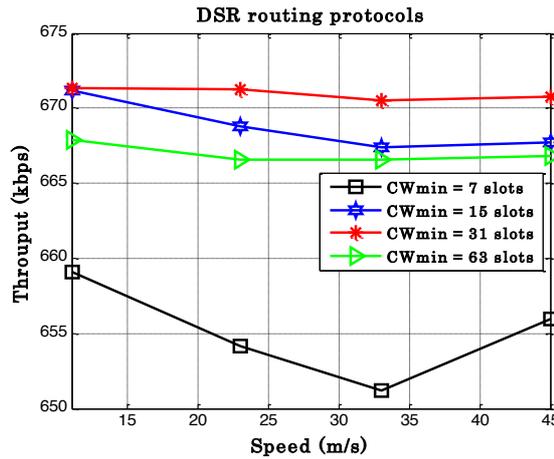


Figure 6: Throughput (kbps)

Figure 7 shows the average end to end delay of DSR routing is the best performance with the initial backoff contention window (equal 15 slots).

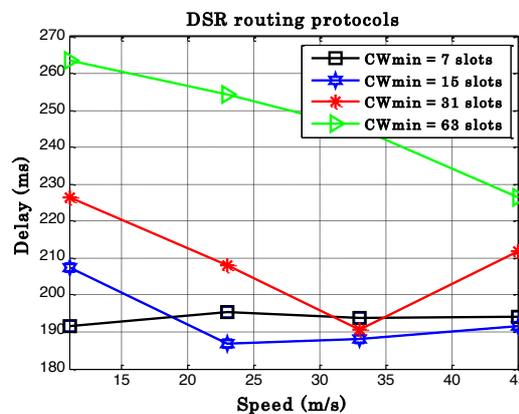


Figure 7: End to End Delay (ms)

Figure 8 and 9 shows the packet delivery ratio and throughput of the DSDV routing is the best performance with the initial backoff contention window (equal 15 slots).

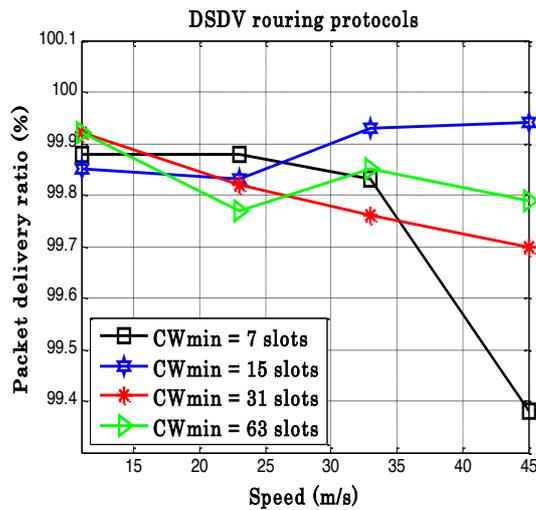


Figure 8: Packet Delivery Ratio (%)

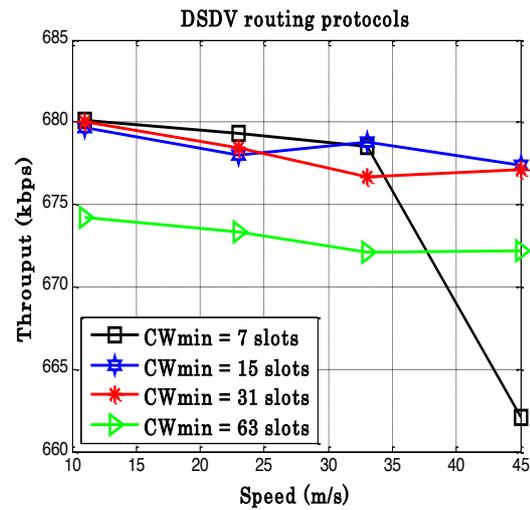


Figure 9: Throughput (kbps)

Figure 10 shows the average end to end delay of DSDV routing is the best performance with the initial backoff contention window (equal 7 slots).

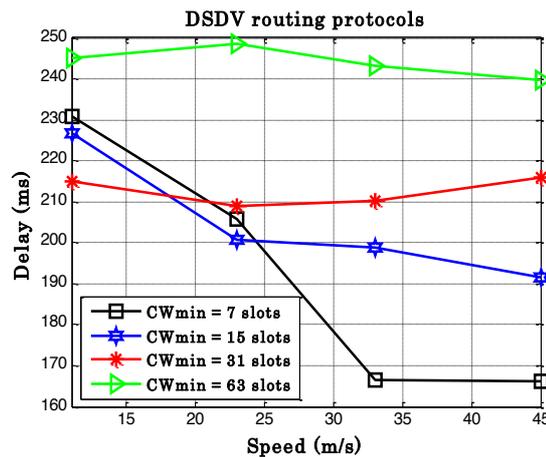


Figure 10: End to End Delay (ms)

8. Conclusion

This paper evaluates the performance of three routing protocols namely: AODV, DSR and DSDV using the NS-2 simulator. This involves examining the influence of the initial backoff contention windows ($CW_{min} = 8, 16, 32$ and 64 slots) on the routing protocols, the evolution of the density of the network while the nodes are in mobility. The results lead to important conclusions about the mechanism for backoff contention windows when random access is disabled.

When the value of the initial backoff contention windows equal $CW_{min} = 15$ slots, the throughput is higher in the AODV and DSDV routing protocols. But when initial backoff contention windows reduced to $CW_{min} = 7$ slots at a varying number of nodes reduces the average end-to-end delay in the AODV and DSDV routing protocols. Although when the value of the initial backoff contention windows $CW_{min} = 31$ slots, the DSR routing protocol gains higher throughput than others.

Using initial backoff contention windows $CW_{min} = 15$ slots at a varying number of nodes reduces the average end-to-end delay in the DSR routing protocol. DSR protocol is influenced by initial

backoff contention windows relative to other routing protocols. In addition, as the initial backoff contention windows increase, the DSR end-to-end average delay increases. Conspicuously, the DSR is affected by the initial backoff contention windows.

AODV and DSDV performance is best when the initial backoff contention windows are equal to 15 slots. Further, in the case of the small end-to-end average delay in the initial backoff contention windows, equal to $CW_{min} = 7$ slots.

References

- [1] T. Ali, L. Khalil and Y. Majeed, "Review and performance comparison of VANET protocols: AODV, DSR, OLSR, DYMO, DSDV & ZRP," 2016 Al-Sadeq International Conference on Multidisciplinary in IT and Communication Science and Applications (AIC-MITCSA), Baghdad, pp. 1-6, May 2016.
- [2] H. Bello, A. Aibinu, Z. Wang, A. Onumanyi, E. Onwuka, J. Dukiya, "An optimized routing algorithm for vehicle Ad-hoc networks" Engineering Science and Technology, an International Journal Vol. 22, Issue 3, PP. 754-766, June 2019.
- [3] Z. Houssaini, I. Zaimi, M. Oumsis, and S. Ouatik, "Comparative Study of Routing Protocols Performance for Vehicular Ad-hoc Networks" International Journal of Applied Engineering Research" Vol. 12, N. 13, pp. 3867-3878, 2017.
- [4] G. Campus, K. Malaysia, "Vehicular Ad-hoc Network (VANET): Review" IEEE International Conference on Innovative Research and Development (ICIRD), Bangkok Thailand, pp. 11-12, May 2018.
- [5] M. Saravanan, P. Ganeshkumar, "Routing using reinforcement learning in vehicular ad hoc networks" Computer Science and Engineering, KPR Institute of Engineering and Technology, Coimbatore, India, January 2020.
- [6] 802.11 Standard –wireless LAN medium access control and physical layer specifications," June 1997.
- [7] S. Kumar, V. S. Raghavan, and J. Deng, "Medium access control protocols for ad hoc wireless networks: A survey," Ad Hoc Networks, vol. 4, pp. 326-358, 2006.
- [8] K. J. Kumar, B. Manoj, and C. S. R. Murthy, "RT-MuPAC: Multi-power architecture for voice cellular networks," in International Conference on High-Performance Computing, pp. 377-387, 2002.
- [9] M. Natkaniec and A. Pach, "An analysis of the backoff mechanism used in IEEE 802.11 networks," in Computers and Communications, Proceedings. ISCC 2000. Fifth IEEE Symposium on, pp. 444-449, July 2000.
- [10] B. Manoj and C. Murthy, "Throughput enhanced wireless in local loop (TWiLL): the architecture, protocols, and pricing schemes," ACM Sigmobile Mobile Computing and Communications Review, vol. 7, pp. 95-116, 2003.
- [11] G. Bianchi, "performance analysis of the IEEE 802.11 distributed coordination function," IEEE Journal on Selected Areas in Communications, vol. 18, pp. 535-547, March 2000.
- [12] Network Simulator 2 (NS-2) <http://mohit.ueuo.com/NS-2.html>
- [13] S. Alahmar, O. Aldawibi, "Optimum Routing Protocol for Static Wireless Ad-hoc Network", Libyan International Conference on Electrical Engineering and Technologies (LICEET2018), Tripoli, 4-7 March 2018.
- [14] S. Alahmar, O. Aldawibi, "Influence of Initial Backoff Contention Window on Mobile Ad-Hoc Network", Journal of Electrical and Electronic Engineering and Information Technology, Bani-Walid, Vol.1, pp 55-58, July 2020.

Database for Arabic Speech Commands Recognition

Lina Tarek Benamer^{1*}, Osama A.S. Alkishriwo²

¹ l.benamer@uot.edu.ly, ² o.alkishriwo@uot.edu.ly

^{1,2} Department of Electrical and Electronics, College of Engineering, University of Tripoli, Libya

ABSTRACT

Keywords:

Speech Recognition, Arabic Speech Command Recognition, Wavelet Time Scattering, Support Vector Machine (SVM), Long Short-Term Memory (LSTM), Mel-Frequency Cepstrum Coefficients (MFCC), K-Nearest Neighbor (KNN).

Technology is all around us and it's changing rapidly, expanding Internet access has had huge impacts on everyday lives as people do everything on their phones and computers. The widespread growth in the use of digital computers, have an increasing need to be able to communicate with machines in a simpler manner. One of the main tasks that can simplify communication with machines is speech recognition. In this work, we introduce the Arabic speech commands database that contains six Arabic control order words and Arabic spoken digits. The created database is used to analyze and compare the recognition accuracy and performance of three recognition techniques which are, Wavelet Time Scattering feature extraction with Support Vector Machine (SVM) classifier, Wavelet Time Scattering feature extraction with Long Short-Term Memory (LSTM) classifier, and Mel-Frequency Cepstrum Coefficients (MFCC) feature extraction with K-Nearest Neighbor (KNN) classifier. Finally, the experimental results show that the most accurate prediction of the database commands was 98.1250% given by Wavelet Time Scattering feature extraction and LSTM classifier and the fastest training time for the database was 144 minutes given by MFCC and KNN classifier.

1. Introduction

Speech recognition is an automatic identification of speech by machine using some characteristics of the speaker's voice [1,2], and it is an important technique especially that the world is passing through the era of information and communication technology, where we use speech recognition in many application areas such as human–robot interaction, human–computer interaction, and telephone applications. Speech-based interaction [3,4] is performed using natural human voice and also has many difficulties such as noise, behaviour of humans, and accent of spoken words. Therefore, it is an open challenge for the researchers to develop speech recognition techniques that can recognize different words correctly.

Many speech recognition techniques exist and have been developed, several researchers did several research works related to this area. Truong et al. [5] presented a novel multi-speaker segmentation method that makes use of the wavelet analysis and support vector machine to separate various speech signals of the speakers through multi-dialog approach. Pawan and Raghunath [6] presented a text-independent speaker recognition technique with Fourier transform by means of MFCC and SVM. Mohamed and Ramachandran [7] described the improvement in property of the normalization by using HMM and artificial neural networks. M.A. Anusuya et al. [8] proposed a PCA-based Kannada speech recognition technique where discrete wavelet transforms are used for calculating wavelet coefficients. Current research in the computer science field focuses on deep learning for monitoring change in speech patterns, speech recognition and classification [9,10]. The authors in [11] present an expert neural network based on dynamic selection of classifiers for

application in a speech signal pattern recognition system. M. Imtiaz et al. in [12] proposed an approach of speech recognition system based on isolated word structure using Mel-Frequency Cepstral Coefficients (MFCC's), Dynamic Time Warping (DTW) and K-Nearest Neighbour (KNN) techniques. A combination of a hidden Markov model (HMM) and a deep long short-term memory (LSTM) network for speech recognition is given in [13]. An automatic language identification system using Mel-Frequency Cepstral Coefficients (MFCCs) for feature extraction with K-means clustering and Support Vector Machine (SVM) for classification is introduced in [14]. A comparative study of MFCC-KNN and LPC-KNN for Hijaiyyah letters pronunciation classification system is presented in [15].

As speech technology has developed, the number of individuals who would like to train and evaluate recognition models has grown rapidly, but the availability of datasets has not widened. The main contribution of this paper is to create a database of Arabic Speech Commands as an attempt to build a standard training and evaluation dataset for a class of simple speech recognition tasks. Its primary goal is to provide a way to build and test small models that detect when a single word is spoken, from a set of targeted words, with the fewest possible failures from background noise or unrelated speech. The database contains six Arabic control order words and Arabic spoken digits made with different users. Broadening access to databases will certainly encourage collaboration across groups and enables apples-for-apples comparisons between different approaches, helping the whole field move forward. The created database is used to study, implement, and compare the performance of three different speech recognition techniques.

2. Proposed Arabic Speech Commands Recognition System

The proposed speech recognition system is shown in Figure 1. The features of spoken Arabic commands that were created to contain Arabic control words and Arabic digits are extracted then classified using three machines and deep learning techniques. The classification is performed using Wavelet Time Scattering with a Support Vector Machine (SVM) also with a Long Short-Term Memory (LSTM) network. In addition, an approach using machine learning to identify people based on features extracted from the recorded speech is also presented in this work. In this approach, the features are the Mel-Frequency Cepstrum Coefficients (MFCC), which used to train the K-Nearest Neighbor (KNN) classifier.

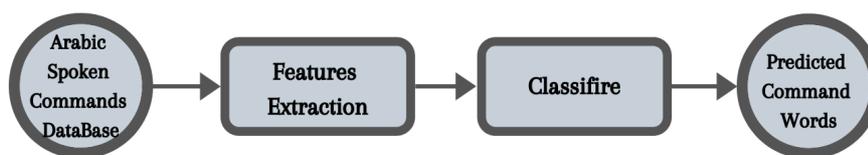


Figure 1: Proposed Arabic speech commands recognition system

2.1 Database

Data is very important to carry out tests, and when it comes to machine learning and deep learning tests we require having a lot of data to obtain the most accurate results. In this paper the database presented is created in Arabic language to support the researchers work in the field of Arabic speech recognition, carry out their tests using the Arabic commands database.

The created speech command database consists of 1600 recordings in Arabic of 6 control words: Add (إضافة), Back (رجوع), Cancel (إلغاء), Confirm (تأكيد), Continue (متابعة), Delete (حذف), and the

digits 0 through 9 obtained from forty different speakers. The frequency of human voice ranges from 20Hz to 14,000Hz (typically from 300Hz to 4,000Hz). The frequency of a sound wave determines the human tone and pitch. In general, the frequencies, which have the most significant part of speech, lie between about 100Hz and 4,000Hz.

As a first step, the voice samples were collected together in one record from each volunteer reading the six words and the ten digits using the recording app on a mobile phone in a normal environment with as minimum background noise as possible. After that using Audacity software the collected data was divided so each audio file is for one specific command and converted to the WAV format so MATLAB could read and use the audio files. The presented database is available online at [16].

The database consists of 16 balanced classes with 100 recordings sampled at 48000Hz, managed to ensure the random division of the recordings into 80% training and 20% test sets as shown in Table 1.

Table 1: Commands Database Labels and Count

Name in English	Name in Arabic	Label	Total Count	Male Count	Female Count
Zero	Sefr "صفر"	0	100	59	41
One	Wahed "واحد"	1	100	37	63
Two	Ethnan "إثنان"	2	100	37	63
Three	Thalatha "ثلاثة"	3	100	41	59
Four	Arbaa "أربعة"	4	100	45	55
Five	Khamsa "خمسة"	5	100	40	60
Six	Seta "ستة"	6	100	40	60
Seven	Sabaa "سبعة"	7	100	44	56
Eight	Tamania "ثمانية"	8	100	40	60
Nine	Tesaa "تسعة"	9	100	43	43
Add	Edafa "إضافة"	A	100	34	66
Back	Rojou "رجوع"	B	100	36	64
Cancel	Elgha "إلغاء"	C	100	31	69

Delete	Hadef “حذف”	D	100	34	66
Confirm	Takeed “تأكيد”	E	100	33	67
Continue	Motaba “متابعة”	F	100	63	37

The recordings in the database are not of equal durations, and are not prohibitively large, so through reading the database files a histogram of the signal lengths was constructed as given in Figure 2. The histogram shows that the distribution of recording lengths is positively skewed. Classification uses a common signal length of 8192 samples, and it's considered as a conservative value that ensures that truncating longer recordings does not cut off speech content. Meaning that if the signal is greater than 8192 samples (1.024 seconds) in length, the recording is truncated to 8192 samples. And if the signal is less than 8192 samples in length, the signal is pre-padded and post-padded symmetrically with zeros out to a length of 8192 samples.

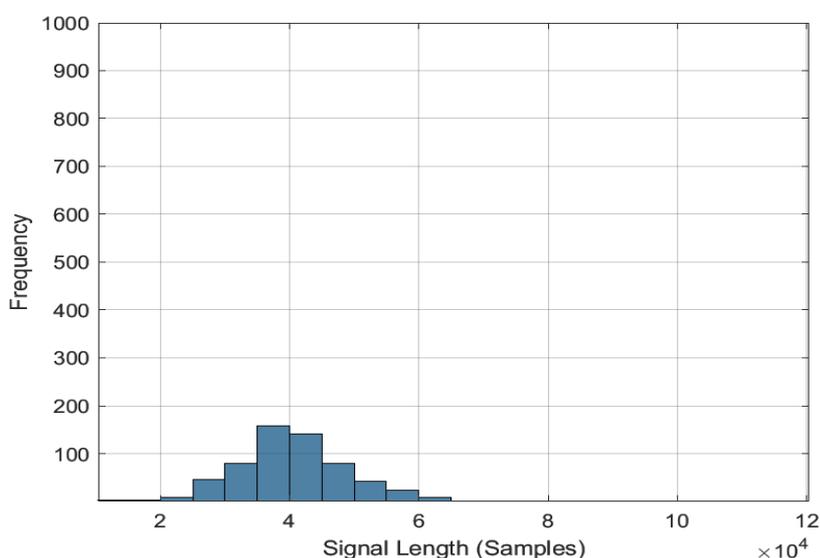


Figure 2: Histogram of the signal lengths

2.2 Feature Extraction Block

The process of dimensionality reduction by which an initial set of raw data is reduced to more manageable groups for processing is called feature extraction. The characteristic of large data sets which has large number of variables, require a lot of computing resources to process. By extracting features from the input data the feature extraction increases the accuracy of learned models. This phase of the general framework reduces the dimensionality of data by removing redundant data. Of course, it increases training and inference speed while still accurately and completely describing the original data set.

2.2.1 Wavelet Time Scattering

Wavelet transform is a mathematical approach that is widely used for signal processing applications. It can be a decompose of special patterns hidden in the mass of data. Regarding the prediction issue through time series and neural networks, we need a modelling task. Neural networks as a general estimator in the estimation of extremely nonlinear systems have limited capability. Wavelet transform has the ability to simultaneously display functions and manifest their local characteristics in the time-frequency domain. The use of these characteristics facilitates the training of neural networks with accuracy to model extremely nonlinear signals. Wavelet techniques in general are

effective tools for good data representations and feature extractions which can be used with most available classification algorithms. The wavelet scattering transform allows us to produce reliable and locally stable features to small deformations which can be used in conjunction with a deep neural network. Convolution, nonlinearity, and averaging, three successive main and required operations to produce a wavelet scattering transform of a time series input signal.

2.2.2 Mel-Frequency Cepstrum Coefficients (MFCC)

This is a method of extracting frequency information in speech signals and converting them into coefficients. Based on a linear cosine transform of a log power spectrum on a nonlinear Mel scale of frequency the representation of the short-term power spectrum of a sound is made. Mel-frequency cepstral coefficients are a parametric representation of the speech signal that is commonly used in automatic speech recognition and they are calculated by applying a Mel-scale filter bank to the Fourier transform of a windowed signal. Subsequently, a DCT (discrete cosine transform) transforms the logarithmized spectrum into a cepstrum. The Mel filter banks consist of overlapping triangular filters with the cut off frequencies that can be determined by the centre frequencies of the two adjacent filters. The filters have linearly spaced centre frequencies and fixed bandwidth on the mel scale. The logarithm has the effect of changing multiplication into addition. It converts the multiplication of the magnitude in the Fourier transform into addition. MFCCs simulate the properties of the human auditory system, that's one of the reasons they are widely applied in speech processing.

2.3 Classification Block

An algorithm that sorts data into labelled classes, or categories of information. And classification in general is the process of predicting the class of given data points. Classes are sometimes called targets or labels or categories. Classification predictive modelling is the task of approximating a mapping function from input variables to discrete output variables.

2.3.1 Support Vector Machine (SVM)

SVM is a binary classification algorithm that determines the decision boundary between feature vectors of two classes. It can be scaled well for multi-class classification and can be generalized well in high dimensional feature spaces. It is a supervised machine learning algorithm that is recognized as an easy-to-use and robust technique for classification, regression, and other learning tasks. It has a better classification performance on a small number of training samples using a technique called the kernel trick to transform data and then based on transformations, an optimal boundary between the possible outputs can found.

2.3.2 Long Short-Term Memory (LSTM)

An artificial recurrent neural network (RNN) architecture used in the field of deep learning, LSTMs are designed to overcome the vanishing gradient problem that restricts the memory capabilities of traditional RNNs, increases the chance of facing a gradient problem and losing information since too many time steps have been added. LSTMs are designed to allow the retention of information for longer periods compared to traditional RNNs. LSTMs continue learning over numerous time-steps and back propagate through time and layers because they can maintain a constant error. And they can also use gated cells to store information outside the regular flow of the RNN. With these cells, the network can manipulate the information in many ways, including storing information in the cells and reading from them. The cells are individually capable of making decisions regarding the information and can execute these decisions by opening or closing the gates. Since there can be lags of unknown duration between important events in a time series, LSTM networks are well-suited to classifying, processing, and making predictions based on time series data.

2.3.3 K-Nearest Neighbor (KNN)

An unsupervised learning technique that is used for classifying objects based on closest training examples in the feature space. KNN is a type of lazy learning or instance-based learning where the function is only approximated locally and all computation is deferred until classification. The KNN is considered one of the simplest classification techniques when there is little or no prior knowledge about the distribution of the data. The followed up rule simply retains the entire training set during learning and assigns to each query a class represented by the majority label of its k-nearest neighbors in the training set. In this method, each sample should be classified similarly to the surrounding samples. Therefore, if the classification of a sample is unknown, then this could be predicted by consideration of the classification of its nearest neighbor samples. In this classifier, we use Euclidean distance between the classes which can be further classified to know which class the data belongs to.

3. Results and Discussion

In order to obtain the recognition accuracy and analyse the performance of the Arabic speech commands database with the proposed techniques, experiments were carried out on the created database using MATLAB Audio Toolbox, Statistics and Machine Learning Toolbox, Deep Learning Toolbox and Wavelet Toolbox. The computed results were given using Intel Core i5 4250U at 1.3 GHz with 4GB RAM. The MATLAB used a single CPU with constant 0.0021988 learning rate of over 7500 iterations.

Table 2: Comparison of different speech recognition techniques in terms of accuracy and computation

Techniques	Training Time	Test Accuracy
Wavelet Time Scattering feature extraction with Support Vector Machine (SVM) classifier	261 min	96.2500%
Wavelet Time Scattering feature extraction with Long Short-Term Memory (LSTM) classifier	526 min	98.1250%
Mel-Frequency Cepstrum Coefficients (MFCC) feature extraction with K-Nearest Neighbor (KNN) classifier	144 min	94.89%

As shown in Table 2, we can see that each technique had its pros and cons, which would make the technique that showed the most accurate prediction of the database commands the Wavelet Time Scattering feature extraction with Long Short-Term Memory (LSTM) classifier giving us 98.1250% test accuracy and the fastest training time for the database by the Mel-Frequency Cepstrum Coefficients (MFCC) feature extraction with K-Nearest Neighbor (KNN) classifier given by 144 minutes of total training time. Based on these results we can use the technique that will serve the purpose of our application in which we would need the best prediction accuracy or the fastest training time.

4. Conclusions

This paper presents the analysis of Arabic speech recognition via Wavelet Time Scattering feature extraction with Support Vector Machine (SVM) classifier, Wavelet Time Scattering feature extraction with Long Short-Term Memory (LSTM) classifier, and Mel-Frequency Cepstrum Coefficients (MFCC) feature extraction with K-Nearest Neighbor (KNN) classifier. The test experiments provided a recognition accuracy and training time performance of these techniques

using the created Arabic speech commands database. Experiments with the three speech recognition techniques showed that the recognition accuracy for the Wavelet Time Scattering feature extraction with Long Short-Term Memory (LSTM) classifier was the best, although the difference between all of the techniques recognition accuracy was small. On the other hand, the best result in the training time performance was by Mel-Frequency Cepstrum Coefficients (MFCC) feature extraction with K-Nearest Neighbor (KNN) classifier, and the difference between the training times was large. The Arabic Speech Commands dataset has shown to be useful for training and evaluating a variety of models.

As a future work, we recommend further investigation of Arabic speech recognition performance by different recognition techniques using this Arabic speech commands database to move forward with developing more speech control applications that support the Arabic language.

Acknowledgment

Massive thanks to everyone who volunteered with their recordings from my family, friends, and colleagues to create this database.

References

- [1]. C. S. Kumar and P. M. Rao, "Design of an automatic speaker recognition system using MFCC, vector quantization and LBG algorithm", *International Journal on Computer Science and Engineering*, vol. 3, no. 8, pp. 2942-2954, Aug. 2011.
- [2]. S. K. Gaikwad and B.A. Gawali, "A review on speech recognition technique," *International Journal of Computer Applications*, vol. 10, no. 3, pp. 16–24, Nov. 2010.
- [3]. Y. Lee and K. W. Hwang, "Selecting good speech features for recognition," *ETRI Journal*, vol. 18, no. 1, Apr. 1996.
- [4]. D. Y. Genoud, D. Ellis, and N.Morgan, "Combined speech and speaker recognition with speaker adapted connectionist models," *IEEE Workshop on Automatic Speech Recognition and Understanding*, Keystone, Colorado, , Dec. 1999.
- [5]. T. K. Truong, C. L. Chien, C. Shihuang, "Segmentation of specific speech signals from multi dialog environment using SVM and wavelet," *Pattern Recognition Letters*, vol. 28, no. 11, pp. 1307-1313, Aug. 2007.
- [6]. K. A. Pawan, S.H. Raghunath, "Fractional Fourier transform based features for speaker recognition using support vector machine," *Computers and Electrical Engineering*, vol. 39, no. 2, pp. 550-557, Feb. 2013.
- [7]. M. A. Ramachandran and K. N. Nair, "HMM/ANN hybrid model for continuous Malayalam speech recognition," *Procedia Eng.*, vol. 30, pp. 616-622, 2012.
- [8]. M.A. Anusuya and S.K. Katti, "Mel frequency discrete wavelet coefficients for Kannada speech recognition using PCA," *In Proceedings of International Conference on Advances in Computer Science (ACEEE)*, pp. 225-227, 2010.
- [9]. A.B. Nassif, I. Shahin, I. Attili, M. Azzeh, and K. Shaalan, "Speech recognition using deep neural networks: A systematic review," *IEEE Access*, vol. 7, pp. 19143-19165, Feb. 2019.
- [10]. D. O'Shaughnessy, "Recognition and processing of speech signals using neural networks," *Circuits, Systems, and Signal Processing*, vol. 38, pp. 3454-3481, Mar. 2019.
- [11]. P. Rocha, W. Silva, and A. Barros, "Hierarchical expert neural network system for speech recognition," *Journal of Control, Automation and Electrical Systems*, vol. 30, pp. 347-359, Mar. 2019.
- [12]. M. A. Imtiaz and G. Raja, "Isolated word automatic speech recognition (ASR) system using MFCC, DTW & KNN," *The 2016 Asia Pacific Conference on Multimedia and Broadcasting*, Bali, Indonesia, pp. 106-110, Nov. 2016.
- [13]. W. Ying, L. Zhang, and H. Deng, "Sichuan dialect speech recognition with deep LSTM network," *Frontiers of Computer Science*, vol. 14, pp. 378-387, Aug. 2019.
- [14]. V. K. Verma and N. Khanna, "Indian language identification using K-means clustering and support vector machine (SVM)," *2013 Students Conference on Engineering and Systems (SCES)*, Allahabad, India, Apr. 2013.



-
- [15]. Adiwijaya, M. Nur Aulia, M. S. Mubarok, W. U. Novia, and F. Nhita, “A comparative study of MFCC-KNN and LPC-KNN for hijaiyyah letters pronunciation classification system,” *2017 5th International Conference on Information and Communication Technology (ICoIC7)*, Malacca City, Malaysia, May 2017.
- [16]. https://github.com/ltkbenamer/AR_Speech_Database.git

Decision-Making Rules in Cooperative Cognitive Radio Networks: Evaluation and Comparison

Magda A. Abunab¹, Mohamed A. Elalem^{2*}

¹pg.maabunab@elmergib.edu.ly, ²maelalem@elmergib.edu.ly

^{1,2} Department of Electrical and Computer Engineering, Faculty of Engineering, Elmergib University, Libya

ABSTRACT

Cognitive radio CR is a communication technology developed to solve the problem of spectrum scarcity. Energy detection based on cooperative spectrum sensing represents a solution to enhance the throughput of CR since the information about primary signal presence are collected using many sensing nodes with different channel conditions. Each node reports its own reports to the network centre to make its decision. However, the throughput cannot be maximized unless efficient decision rules are used to combine the collected information and produce right final judgment. In this paper, a centralized cooperative spectrum sensing scheme is used, and basic decision rules are presented. New decision-making rule based on statistical average of the node reports is proposed. Closed form expressions for probability of detection and false alarm probability for the different decision rules are given. Comparison on the throughput performance of each decision rule is studied simulated via a system model and MATLAB programming. Fading channel is assumed for data transmission, while the reporting channels are assumed to be free of errors.

Keywords:

Cognitive radio,
cooperative spectrum
sensing,
decision rules,
energy detection.

1. Introduction

Emerging wireless devices and applications further accelerates the development of wireless systems. Such an exponential growth of wireless communication also imposes huge demands on radio spectrum. As a natural resource, radio spectrum is scarce and limited. Nowadays, the spectrum is managed by government agencies such as the Federal Communications Commission (FCC), and assigned to licensed users on a long term basis to avoid interference among wireless systems. Although this static allocation approach worked well in the past, it cannot serve the ever increasing demand for wireless communication well because of the problem of spectrum scarcity. Recent studies reveal that the allocated spectrum is underutilized. Some parts of spectrum remain largely underutilized; some parts are sparingly utilized, while the remaining parts of the spectrum are heavily occupied.

It is recognized that this kind of static allocation police has resulted in poor spectrum utilization. Furthermore, spectrum underutilization by licensed users exacerbates spectrum scarcity. The main reason of spectrum underutilization is that licensed users typically do not fully utilize their allocated bandwidths for most of the time, while unlicensed users are being starved for spectrum availability. To deal with this dilemma, cognitive radio is a paradigm created in an attempt to enhance spectrum utilization, by allowing unlicensed users to coexist with licensed users and make use of the spectrum holes. The spectrum holes are defined as the spectrum bands owned by licensed users, which are unused at a particular time and specific geographic location. Cognitive radio is the key enabling

technology that enables next generation communication networks, also known as Dynamic Spectrum Access (DSA) networks, to utilize the spectrum more efficiently in an opportunistic fashion without interfering with the primary users. It is also defined as a radio that can change its transmitter parameters according to the interactions with the environment in which it operates. It differs from conventional radio devices in that a cognitive radio can equip users with cognitive capability and reconfigurability.

Cognitive capability defines the ability to sense and gather information from the surrounding environment, such as information about transmission frequency, bandwidth, power, modulation, etc. With this capability, cognitive users (CU) can identify the best available spectrum. Reconfigurability is the ability to rapidly adapt the operational parameters according to the sensed information in order to achieve the optimal performance. By utilizing the spectrum in an opportunistic fashion, cognitive radio allows secondary users to sense the portion of the spectrum if available, select the best available channel, co-ordinate spectrum access with other users, and leave the channel when a primary user reclaims the spectrum usage right [1].

Spectrum sensing process is needed to achieve this detection in cognitive radio. CR users must be able to detect the signal of the primary user. Individual spectrum sensing is sometimes difficult since the fundamental characteristics of wireless channels such as multipath fading, shadowing, can degrade the signal. To overcome these issues of individual spectrum sensing, cooperative spectrum sensing is proposed, where CRs send their local sensing information to Fusion Centre (FC) where the final decision can be made [2].

In cooperative spectrum sensing, the cognitive cycle which include sensing operation of each cognitive user, transmits the sensing information to FC and takes a final decision about licensed users signal presence. All these operations should be done as fast as possible and with a high probability of correct decision. The performance measure of these two requirements is the throughput. The throughput of CR is defined as the ratio of the total transmission time to the total frame time after successful final decision is taken. The throughput is normally deteriorated because of the existence of channel noise, fading and because of the use of inefficient fusion rules at FC. Therefore, there is always a need for developing efficient approaches to handle this deterioration. In this paper, most fusion rules will be analyzed and evaluated furthermore, one new fusion rule is proposed.

The remainder of this paper is organized as follow. Section 2 presents the system model. The spectrum sensing analysis is given in Section 3. The cooperative sensing and decision-making rules are covered in Section 4. Section 5 introduces the concept of throughput in CR network. Simulation results are given in Section 6. Finally the paper is concluded in Section 7.

2. System Model

The system model is set up as illustrated in Figure 1. Energy detection technique is adopted to detect whether primary user is transmitting or not. It is assumed that number of cognitive radio users are random uniformly distributed around the area where the primary user is in operating. The sensing procedure is that the CR users locally sense the PU. Then, they collaboratively forward either its decision or observation to the fusion center. An error-free transmission is assumed in reporting channel such that the fusion center receives exactly the same information as sent. The fusion center makes final decision and inform all CR users.

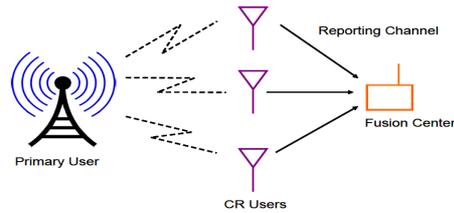


Figure 1: The system model.

3. Spectrum Sensing Analysis

Noncooperative spectrum sensing occurs when only one secondary user performs the primary user detection process. According to this scenario, three different aspects for spectrum schemes are discussed the proceeding subsections. Cooperative spectrum sensing is introduced in the next section.

3.1 Energy Detection

Energy detection has become a widely used technique to sense the primary user signal [3]. A block diagram of a conventional energy detector is illustrated in Figure 2. A band-pass filter (BPF) is first applied, and then its output is squared, integrated, and compared against a threshold to make a decision on the presence of a signal. The energy detection method is attractive because of its implementation simplicity compared to other sensing schemes, as well as fast detection of the primary signals. It has a good resistance against dynamic radio environment where none a prior knowledge about the PUs is available (non-coherent detector). However, the performance of the energy detector is easily affected by channel fading, shadowing, and interferences.

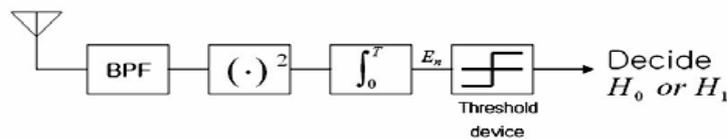


Figure 2: Block diagram of a conventional energy detector.

The signal statistics (the computed energy) are compared to a predetermined threshold. The average total energy detected E , using N_s samples, is defined as

$$E = \frac{1}{N_s} \sum_{i=1}^{N_s} |r(n)|^2 \quad (1)$$

where $r(n)$ is the voltage value received at sample n . If E is more than or equal to γ , this indicates that the spectrum is used (hypothesis H_1) and if E is smaller than γ , this means there is hole in spectrum (hypothesis H_0).

Specifically, the energy of the received signal is collected in a fixed bandwidth W and a time slot duration T and then compared with a pre-designed threshold γ , if $E \geq \gamma$, then the cognitive radio assumes that the primary system is in operation, i.e., H_1 . Otherwise, it assumes H_0 . The average probability of detection, false alarm, and missing of energy detection (The miss detection occurs

when the primary user is in operation but the cognitive radio fails to sense it) over noisy and fading channels can be given by, respectively, [4]:

$$P_D = E[P_r\{H_1|H_1\}]_\eta = e^{-\frac{\gamma}{2}} \sum_{n=0}^{\alpha-2} \frac{1}{n!} \left(\frac{\gamma}{2}\right)^n + \left(\frac{1+\bar{\eta}}{\bar{\eta}}\right)^{\alpha-1} \times \left(e^{-\frac{\gamma}{2(1+\bar{\eta})}} - e^{-\frac{\gamma}{2}} \sum_{n=0}^{\alpha-2} \frac{1}{n!} \left(\frac{\gamma \bar{\eta}}{2(1+\bar{\eta})}\right)^n \right) \quad (2)$$

$$P_{FA} = E[P_r\{H_1|H_0\}] = \frac{\Gamma(m, \frac{\gamma}{2})}{\Gamma(m)} \quad (3)$$

$$P_m = E[P_r\{H_0|H_1\}] = 1 - P_D \quad (4)$$

where $\bar{\eta}$ denotes the average SNR at the cognitive radio. α is a certain margin of protection which is a measure of how much interference above the noise floor the primary user can tolerate (typical value is 5dB). $E[\cdot]_\eta$ represents the expectation over the random variable η (the instantaneous SNR) which is modelled as exponential distributed. $P_r\{\cdot\}$ is the probability of the event. $\Gamma(\cdot)$ is the *gamma function*, and $\Gamma(\cdot, \cdot)$ is the *incomplete gamma function*.

3.2 Matched Filter

The matched filter detection is a linear filter and is used when a secondary user has a prior knowledge of the PU signal properties. This prior information includes carrier frequency, modulation type and pulse shape. This condition makes the matched filter detection impractical. A matched filter maximizes the signal-to-noise ratio (SNR) of the received signal so it is the optimal signal detection. Its performance degrades when there is a reduction of channel knowledge due to rapid changes in the channel state conditions.

A matched-filtering process is equivalent to a correlation scheme; wherein a signal is convolved with a filter whose impulse response is a mirror and time shifted version of the reference signal. The matched filter $h(t)$ convolves the received signal $r(t)$ with a time-reversed version of the known signal as;

$$r(t) \otimes h(t) = r(t) \otimes s(T - t - \tau) \quad (5)$$

where T refers to a symbol time duration and τ is a shift in the known signal $s(t)$, and \otimes refers to the convolution operator. The details of this technique can be found in [5,6].

3.3 Cyclostationary Detection

A signal is said to be cyclostationary if its autocorrelation is a periodic function of time with some period. Cyclostationary feature detection exploits the periodicity of the received signal to identify the presence or absence of primary users. The periodicity is commonly embedded in sinusoidal carriers, spreading code and cyclic prefixes of the primary signals. Due to the periodicity, these cyclostationary signals exhibit the features of periodic statistics and spectral correlation. The complex system depicting this method of detection is also presented in [6].

A signal $s(t)$ is said to be Cyclostationary, if its mean and autocorrelation function $E[s(t)]$, $R_s(t, \tau)$ are periodic, i.e., for any integer k :

$$E[s(t)] = E[s(t + kT_0)] \quad \text{and} \quad R_s(t, \tau) = R_s(t + kT_0, \tau) \quad (6)$$

The details of using cyclostationary analysis as a technique to accomplish signal detection is described in [7].

Compared to energy detection, Cyclostationary feature detection has a better performance when SNR is low. However, it has the same disadvantage as matched filter detection in the sense that it needs prior knowledge of the primary user. Also, it requires a long detection time which makes it less popular than energy detection.

Because energy detection is the most popular and simplest sensing method, it has selected for spectrum sensing in this paper. So, the two disadvantages mentioned above are overcome.

4. Cooperative Sensing and Decision-Making Rules

Cooperation has always benefits. With cooperation, communications can greatly improve the data transmission and reduce the transmission errors. It is defined as the willingness of users in the same network to share information, power and computation with neighbouring nodes and this can lead to savings of overall network resources. In this section, cooperation among cognitive users is considered. This eases to reduce the uncertainty of information recorded by single user detection. The cooperative detection can provide more accurate performance. However, it requires additional operations and overhead traffic to communicate among CR users. As a result, there can be an effect on the performance of resource-constrained networks. Different decision-making rules to collect and combine this information is analysed and evaluated.

In combining rules based cooperative spectrum sensing, CR users forward their local decision to the fusion center to make a final decision. Assuming that the energy observations at each CR user is independent and identically distributed (i.i.d.). All decisions from the cognitive radio users are then sent to the fusion center, where the global decision is made. The probability of detection and false alarm probability at the fusion center is given by [7]:

$$P_D = \sum_{i=k}^{N_c} \binom{N_c}{i} (P_{D,i})^k (1 - P_{D,i})^{N_c-i} \quad (7)$$

$$P_{FA} = \sum_{i=k}^{N_c} \binom{N_c}{i} (P_{FA,i})^k (1 - P_{FA,i})^{N_c-i} \quad (8)$$

where N_c is the number of secondary users sensing the spectrum and $P_{D,i}$ and $P_{FA,i}$ is the probability of detection and false alarm probability of the i^{th} cognitive user respectively, k is set according to the used rule, and $\binom{N_c}{k}$ is binomial coefficient.

Here, different decision combining rules are described.

4.1 OR Rule Decision

In this rule, if any one of the local decisions sent to the decision maker is a logical one, the final decision made by the decision maker is one. The OR rule decides that a hole is present if any of the users detect a hole [8, 9], therefore k is set to 1 in Eq. (3.1) and Eq. (3.2). The probability of detection and probability of false alarm of the final decision of this rule are, respectively [10]:

$$P_{D,OR} = 1 - (1 - P_D)^{N_c} \quad (9)$$

$$P_{FA,OR} = 1 - (1 - P_{FA})^{N_c} \quad (10)$$

4.2 AND Rule Decision

In this rule, if all of the local decisions sent to the decision maker are one, the final decision made by the decision maker is one. The fusion center's decision is calculated by logic AND of the received hard decision statistics. The AND rule decides that a hole is present if all users detect a hole [8] therefore k is set to N_c in Eq. (7) and Eq. (8). The probability of detection and probability of false alarm of the final decision of this rule are, respectively:

$$P_{D,AND} = (P_D)^{N_c} \quad (11)$$

$$P_{FA,AND} = (P_{FA})^{N_c} \quad (12)$$

4.3 MAJORITY Rule Decision

In this rule, if half or more of the local decisions sent to the decision maker are the final decision made by the decision maker is one, the MAJORITY rule decides that a hole is present if half or more of users detect a hole [9] therefore k is set to $N_c/2$ in Eq. (7) and Eq. (8). The probability of detection and probability of false alarm of the final decision are, respectively:

$$P_{D,MAJORITY} = \begin{cases} \sum_{i=N_c/2}^{N_c} \binom{N_c}{i} (P_{D,i})^i (1 - P_{D,i})^{N_c-i}, & N_c \text{ is even} \\ \sum_{i=\text{ceil}(\frac{N_c}{2})}^{N_c} \binom{N_c}{i} (P_{D,i})^i (1 - P_{D,i})^{N_c-i}, & N_c \text{ is odd} \end{cases} \quad (13)$$

$$P_{FA,MAJORITY} = \begin{cases} \sum_{i=N_c/2}^{N_c} \binom{N_c}{i} (P_{FA,i})^i (1 - P_{FA,i})^{N_c-i}, & N_c \text{ is even} \\ \sum_{i=\text{ceil}(\frac{N_c}{2})}^{N_c} \binom{N_c}{i} (P_{FA,i})^i (1 - P_{FA,i})^{N_c-i}, & N_c \text{ is odd} \end{cases} \quad (14)$$

where $\text{ceil}(\frac{N_c}{2})$ rounds the elements of $N_c/2$ to the nearest integers greater than or equal to $N_c/2$.

4.4 Middle Plus One MPO Decision

Middle Plus One MPO is first proposed rule and it is a type of hard decision rules. The mechanism of its operation is as follows: the fusion center makes a final decision of "0" when half plus one or more of the local decisions sent to the fusion center are "0" (indicating the existence of hole) therefore k is set to $(\frac{N_c}{2} + 1)$ in Eq. (7) and Eq. (8) and the fusion center indicates that the hole is present. This rule will increase the throughput since it increases the number of secondary users that give the same decisions which decrease the probability of false alarm. The probability of detection and the probability of false alarm of MPO rule is given in the following equations:

$$P_{D,MPO} = \begin{cases} \sum_{i=\frac{N_c}{2}+1}^{N_c} \binom{N_c}{i} (P_{D,i})^i (1 - P_{D,i})^{N_c-i}, & N_c \text{ is even} \\ \sum_{i=\text{ceil}(\frac{N_c}{2})+1}^{N_c} \binom{N_c}{i} (P_{D,i})^i (1 - P_{D,i})^{N_c-i}, & N_c \text{ is odd} \end{cases} \quad (15)$$

$$P_{FA,MPO} = \begin{cases} \sum_{i=\frac{N_c}{2}+1}^{N_c} \binom{N_c}{i} (P_{FA,i})^i (1 - P_{FA,i})^{N_c-i}, & N_c \text{ is even} \\ \sum_{i=\text{ceil}(\frac{N_c}{2})+1}^{N_c} \binom{N_c}{i} (P_{FA,i})^i (1 - P_{FA,i})^{N_c-i}, & N_c \text{ is odd} \end{cases} \quad (16)$$

4.5 AVERAGE Rule Decision

In this rule, the final decision of H_1 is made only when the average of all CR reports lies above a certain predefined threshold. This threshold represents the long-term observations of the average of probabilities of detection of the primary user presence. Monte Carlo (MC) method, which is a stochastic technique based on the use of random numbers can form the basis of calculating this threshold. The higher the number of Monte Carlo samples, the greater the confidence of this threshold. Therefore k in Eq. (7) and Eq. (8) is set a value where the count of CR users performing spectrum sensing exceed the predefined threshold. The threshold of the probability of detection and false alarm probability can be expressed as

$$P_{D,a} = \text{average}(P_{D,i}; \forall i = 1 \dots N_c) |_{MC} \quad (17)$$

$$P_{FA,a} = \text{average}(P_{FA,i}: \forall i = 1 \dots N_c) |_{MC} \quad (18)$$

where $\text{average}(\cdot)$ is the statistical average function, and MC refers to Monte Carlo algorithm.

Then the value of k will be the minimum value of users those satisfy the following criteria

$$k = \min \begin{cases} \text{count}(P_{D,i}: P_{D,i} \geq P_{D,a} \forall i = 1 \dots N_c) \\ \text{count}(P_{FA,i}: P_{FA,i} \geq P_{FA,a} \forall i = 1 \dots N_c) \end{cases} \quad (19)$$

where $\min(\cdot)$ is the known minimum function, and $\text{count}(\cdot)$ is a function that counts its arguments.

Now, this value of k can be plugged in Eq. (7) and Eq. (8) to find the probability of detection and probability of false alarm of the final decision.

$$P_D = \sum_{i=1}^k \binom{k}{i} (P_{D,i})^i (1 - P_{D,i})^{k-i} \quad (20)$$

$$P_{FA} = \sum_{i=1}^k \binom{k}{i} (P_{FA,i})^i (1 - P_{FA,i})^{k-i} \quad (21)$$

5. Throughput of Cognitive Radio

The throughput in cognitive radio is defined as the ratio of the total transmission time to the total frame time after successful final decision is taken [10]. The frame structure in cognitive radios of duration T consists of sensing time τ and data transmission time $(T - \tau)$. The cognitive user senses the spectrum band for a specific time duration τ . Then, in the case of hole presence, the user starts data transmission over remaining frame time duration $(T - \tau)$. The normalized achieved throughput can be expressed as [11]

$$R = \frac{(T-\tau)}{T} (1 - P_{FA}) \quad (22)$$

5.1 Throughput Improvement in Cognitive Radio

According to Eq. (22) and what previously discussed, by controlling some parameters, the throughput of cognitive radio can be improved. Some of these parameters act directly and others indirectly to the throughput improvement. For example, for sensing time τ parameter, a higher sensing time results in precise spectrum sensing, and avoiding interference with the licensed user. However, on the other hand, an increase in sensing time results a decrease in transmission time, leading to low throughput, and shorter sensing time degrades the sensing process, so the optimal sensing time can maximize the SU's throughput.

Frame duration T is very important parameter, for a given sensing time, the larger frame duration, the longer the data transmission time $(T - \tau)$ and maximum throughput value. Also, the longer the frame duration, the more chances that the PU becomes active, thus more interference between PU and SU, which degrades the throughput. Thus, there exists an optimum frame duration for which interference is minimum and throughput of the CR is maximum.

Number of secondary user performing sensing N_c is also a sensitive parameter to throughput. Using more cooperative SUs can enhance spectrum decision making which leads to increase the throughput.

Last but not least, throughput is affected by the adopted fusion rules make the global decisions. However, using appropriate fusion rule in cooperative spectrum sensing causes improvement in cognitive radio throughput.

6. Simulation Results

In this section, the simulation of four decision rules to enhance throughput in CR are presented. The traditional decision rules: OR rule, MAJORITY rule and Middle Plus One (MPO) rule are simulated. Then the proposed decision rule named as AVERAGE rule is also simulated. For the sake of brevity, AND rule is omitted in this study because it gives the worst performance compared to the other decision rules.

These rules are evaluated for the purpose of comparing the throughput with each other. The scenario consists of sensing stage (local decision), transmission stage (reporting) and decision stage (global decision).

A centralized cooperative spectrum sensing scheme is used, where a number of secondary users sensing for the primary user. The primary user data are generated randomly and QPSK modulated. The sensing stage is performed using the energy detection method, in which each secondary user computes the energy of the sensed spectrum. This requires to transform the primary signal to frequency domain. The sensing process is performed in AWGN channel. After that either hard or soft decision schemes are used. An error free transmission is assumed at the reporting channel where the throughput is calculated. In the simulation, the sampling frequency is chosen to be 6 MHz and there are two bits per symbol. The total frame length is 0.1 sec. Number of cooperative users is varied from 1 to 10 users.

In Figure 3, the probability of detection versus the number of cooperative users is plotted. It shows that the probability of detection increases with the increase of the number of users, for all schemes. This is because that when there are number of devices (or users) are involved in the cooperative communication model, there is a high probability of correct decision making expected. When there is a high density of cooperative mobile users, then there would be high probability to make a correct decision. It can also be observed from Figure 3 that the probability of detection of 'OR Rule' better than the rest of the schemes. For instance, in the case of 7 users, the probability of detection of 'OR Rule' and the probability of detection of 'MPO Rule' is 98%, the probability of detection of 'MAJORITY Rule' is 86% however, AVERAGE Rule is only 80%. It is important to note that the performance of all the schemes matches when there are high number of mobile users performing spectrum sensing. In summary, for a lower number of users, 'OR Rule' outperforms than rest of the schemes.

Figure 4 illustrates the probability of false alarm as a function of number of cooperative users. The probability of false alarm has a higher value for a lower number of users. This is because, when lower number of users are present in the network, there is a high probability that the fusion center will make wrong decisions of existing a channel (false alarm). In the case of fewer users, the fusion center does not have enough statistics information to make the correct decision about the presence of the primary user signal.

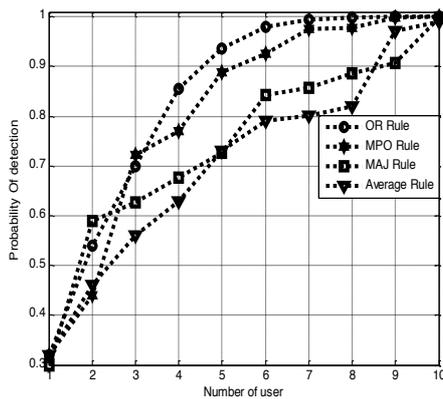


Figure 3: Probability of detection versus number of users.

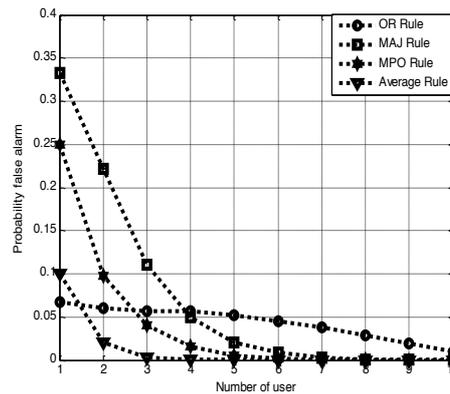


Figure 4: Probability of false alarm versus number of users

Figure 5 compares the performance of all the studied schemes in terms of achieved throughput versus the number of users present in the network. It is obvious that for a higher number of users in the network, there would be higher throughput. It can be depicted from Figure 5 that for a higher number of users, throughput is high as well. However, 'AVERAGE Rule' gives better throughput compared to the other traditional schemes. In conclusion, 'AVERAGE Rule' is the best option when throughput maximization is the key requirement of the network.

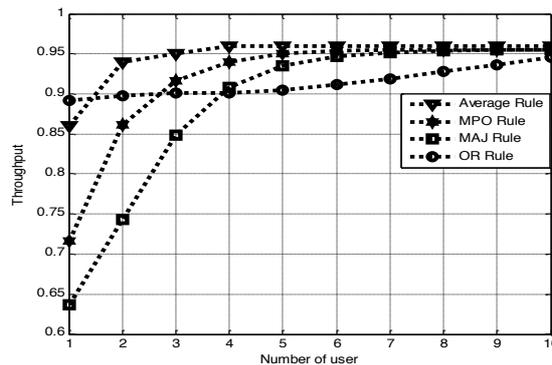


Figure 5: Throughput versus number of users.

7. Conclusions

Static spectrum assignment policy causes low spectrum utilization while the demand for spectrum continues to increase. To solve this issue, Cognitive Radio has been introduced. In order to implement dynamic spectrum access, techniques including spectrum sensing, spectrum analysis and spectrum decision must be adapted. This paper focuses on investigating spectrum sensing techniques by applying energy detection in collaboration sense. Group of user independently perform local spectrum sensing and then report a decision to FC. The FC then makes a final decision. Different decision fusion rules are studied and analyzed. A new rule is also proposed which based on taking the statistical average of all sensing information measured by the individual cooperative users. Monte Carlo algorithm is used to simulate the statistical average for the similar channel environment. This value is used as a threshold value to compare with the measured results.

The comparison between all these rules is done by evaluating the performance of each rule and its response to the detection probability and false alarm. Channel throughput is adopted as a key performance for each rule. The performance of the different all the studied schemes in terms of achieved throughput versus the number of users present in the network are presented. It is

observed that more number of users in the network, higher throughput is achieved. However, the 'AVERAGE Rule' gives better throughput compared to the other traditional schemes. In conclusion, 'AVERAGE Rule' offers the best option when throughput maximization is the key requirement of the network.

References

- [1]. N. Zhang and J. Mark, "Security-aware Cooperation in Cognitive Radio Network", 2014.
- [2]. J. Mitola, "Cognitive radio: An integrated agent architecture for software defined radio", Ph.D. dissertation, KTH Royal Inst. of Technol., Stockholm, Sweden, 2000.
- [3]. D. Sun, T. Song, M. Wu, J. Hu, J. Guo and B. Gu, "Optimal Sensing Time of Soft Decision Cooperative Spectrum Sensing in Cognitive Radio Networks", IEEE wireless communication and networking conference, pp: 4124-4128, 2013.
- [4]. M. Elalem, L. Zhao, and Z. liao, "Interference mitigation using power control in cognitive radio networks", In Proc. IEEE Vehicular Tech. Conf. (VTC), Taipei, Taiwan, May 2010.
- [5]. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," Communications Surveys & Tutorials, IEEE, Volume. 11, no. 1, pp. 116-130, 2009.
- [6]. M. Subhedar and G. Birajdar, "Spectrum sensing techniques in cognitive radio networks: a survey," International Journal of Next-Generation Networks, Volume. 3, no. 2, pp. 37-51, 2011.
- [7]. W. A. Gardner and C. M. Spooner. "Signal interception: performance advantages of cyclic-feature detectors," IEEE Transactions on Communications, 40(1) pp: 149-159, 1992.
- [8]. R. Bharadwaj and H. Goud, "Soft Decision based Spectrum Sensing for Cognitive Radio Networks", International Journal of Digital Application & Contemporary Research, Volume: 03, Issue: 09, 2015.
- [9]. D. Cabric, Sh. Mubaraq, R. Brodersen, "Implementation Issues in Spectrum Sensing for Cognitive Radios", Signals, Systems and Computers, 2004. Conference Record of the Thirty-Eighth Asilomar Conference on, pp: 772-776, 2004.
- [10]. M. Ibnkahla, "Cooperative Cognitive Radio Networks", Queens's University, Kingston, Canada, 2015.
- [11]. Ch. Sita and S.Pallam, "Performance Analysis of Cognitive Radio Networks (IEEE 802.22) for Various Network Traffics", International Journal on Recent and Innovation Trends in Computing and Communication, Volume: 05 Issue: 07, pp: 348-353, 2017.

Investigation of Renewable Electricity Generation from Solar-Hydrogen Hybrid System in Tripoli

Ehsan Altayef^{1*}, Adel Jaber², Fatih Anayi³, Bdereddin Samad⁴, Shaima Dhabaa⁵

¹ Tripoli201919@gmail.com, ² Jaber_ade@yahoo.com,

³ Anayi@cardiff.ac.uk, ⁴ Fobadr@gmail.com, ⁵ Shaima.lutfi@gmail.com

¹ Department of Electrical and Electronics, College of Engineering, University of South Wales, UK

^{2,4} Department of Electrical and Electronics, Azahra Higher Institute of Sciences and Technology, Libya

³ Department of Electrical and Electronics, College of Engineering, Cardiff University, UK

⁵ Department of English Language, Tripoli University, Libya

ABSTRACT

This paper presents a stand-alone solar hydrogen plant to cover the daily electricity demand of a residential unit in Tripoli- Libya. Solar power was obtained through International Global Radiation and photovoltaic (PV) panels, while hydrogen was acquired via water electrolysis. This renewable solar-hydrogen community has demonstrated and clarified that it is possible to be fully reliant on renewable electricity daily. This paper also compares the financial feasibility of supplying the residential unit with electricity using hydrogen, diesel generators, gasoline generators and the grid. Actual economic data from Az-zawiya oil refining company and General Electricity Company of Libya is used. The establishment of an integrated solar-hydrogen power plant to provide daily electrical residential requirements in Tripoli is demonstrated. The data calculation of this development is undertaken using Photovoltaic Geographical Information System (PVGIS) software tool that provides a free and open web access to solar radiation and temperature data and to PV performance assessment tools for any location in Europe and Africa, as well as large part of Asia and America to model the radiance and the amount of sunlight (PVGIS) and excel for modeling the hydrogen demand and production. Capacity and the efficiency of the solar-hydrogen plant to provide a community of electricity for a year without any shortages or deficiency is investigated.

Keywords:

photovoltaics,
electrolyzer,
hydrogen,
Global warming potential,
Fuel cell.

1- Introduction

Mostly papers starts with introduction. It contains the brief idea of work, requirement for this research work, problem statement, and Authors contribution towards their research. Recent references [1] should be included for showing previous work done and importance of current work. This section should be succinct, with no subheadings unless unavoidable [2, 3]. State the objectives of the development and provide an adequate background related to your work, avoiding a detailed literature survey or a summary of the results.

The sun is ultimately the source of all energy sources; even fossil fuels are derived from organic life dependent on the sun. Sun energy arrives in the earth's atmosphere mainly as heat and light, which create wind and biomass. On a macro level of analysis, fossil fuels represent extremely concentrated, pressurized stores of solar energy whose rapid combustion disrupts the earth's more gradualist ecosystem. Conversely, harnessing the daily solar energies of light, heat and wind exerts more gentle impacts on the world [1].

Global warming is accelerating and the most important priority of governments nowadays (if not always reflected in practice) is to reduce carbon emissions. CO₂ mitigation strategies include sun-oriented cookers, water radiators, dryers, biofuel, enhanced cooking stoves and hydrogen technology.

Hydrogen is the most inexhaustible element in the universe, and its potential has long been noted, for a future 'hydrogen economy' replacing the existing carbon economy [2, 3]. Hydrogen-related technologies are generally unsuitable for widespread applications, largely due to safety concerns about storage, as discussed later in this paper, but hydrogen clearly shows promise for base capacity power generation, transportation and household use in cooking and heating. Transportation alone consumes 50% of global petroleum, so hydrogen fuel cell vehicles could make a massive contribution to the reduction of greenhouse gases (GHG) [4, 5].

Hydrogen fuel cell technology essentially replaces energy production from fossil fuel combustion with production from the combustion of hydrogen with oxygen, producing only water as a waste product; it is thus intrinsically an ideal solution to halt the negative impact of transport on the environment. However, the isolation of hydrogen in elemental form requires energy-intensive production.

Ironically, fossil fuels are currently the main resource to deliver hydrogen through steam reforming. In this procedure hydrocarbon feedstock is separated to discharge the hydrogen gas. The significant issue with this procedure, aside from the way that fossil fuel is utilized to power it, is that for every kilogram of hydrogen created, more than eight kilograms of carbon dioxide is likewise delivered, making it counterproductive as an alternative energy source [6, 7]. Conversely, the creation of hydrogen gas by electrolysis, whereby water atoms are split by electrical power derived from solar PV to deliver hydrogen and oxygen, creates no direct carbon emissions, making it essentially a clean renewable technology as well as a source of hydrogen production. Conventional sources of energy production are depleting, and their adverse effects on the environment, such as climate change and global warming, are well known. Therefore, the utilization of renewable energy for energy production is increasing in interest worldwide because of sustainable development and environmental concerns [8, 9].

2- Purpose of This Work

Libya is facing a serious issue with the electricity shortages and blackouts due to increasing demand (related to the increasing population and more energy requirements related to climate and lifestyle factors) and damage to the existing power infrastructure due, as well as interrupted oil and gas supply due to on-going conflict in some areas. To solve this persistent daily problem, this study designs a hydrogen from solar plant to supply a residential unit with renewable electricity. This paper is an attempt of reliance on renewable energy in covering domestic electricity demand.

The application of this development will create a competitive environment and motivate other organizations to invest in renewable energy and hydrogen production. The paper aims to investigate the possibility of hydrogen production from PV solar panels and the economic and technical feasibility of hydrogen production from renewable solar energy. It also provides a comparison between the grid and renewable energy regarding the technical, environmental and economic basis.

The objectives of the paper are to: (1) Estimate the residential demand of hydrogen, (2) Define the methods of energy conversion applied, (3) Calculate the technical and economic feasibility of renewable hydrogen generation.

Different methods utilized for transforming hydrogen into energy are evaluated, including combustion in internal combustion engines, hydrogen/oxygen combustion for steam generation, catalytic combustion, electrochemical conversion in fuel cells. The paper examines electro-chemical conversion in fuel cells in detail and looks at the advantages of electro-chemical processes of producing hydrogen and its reliability, and whether it can wholly rely on renewable energy.

This paper proposes a hydrogen production plant relying on a solar energy system, which would be the first established hydrogen plant in Libyan. This research intends to provide a residential unit consisting of six apartments with electricity and make it fully reliant on renewable energy.

3- Data Analysis Techniques

This paper calculated the amount of hydrogen required to supply a residential unit consisting of six apartments with a measured quantity of electricity ranging between 4.5 kW and 5 kW in peak power demand. The highest power demand periods are recorded when utilizing a higher number of devices and machines around the house depending on the daily requirements of a Libyan household, with particularly high cooling loads in summer and high heating loads in winter related to air conditioning systems.

To work out the amount of hydrogen needed to supply the whole building with electricity some calculations must be undertaken to know the size of the work and the required number of PVs, electrolyzers and ancillary equipment. Some other equally important calculations include those pertaining to the consumed power, power demand and the amount of power to be stored in order to assure that there are no power blackouts during the whole year. These calculations had to be done accurately and correctly. Excel programmer was used to guarantee error-free equations. All the required equations and necessary data were entered in Microsoft Excel to be worked out to get the needed results. Data of global radiation for Tripoli- Libya, the site for the studied building, was obtained from the Photovoltaic Geographical Information System (PVGIS) and the radiation data was entered from sunrise to sunset [10]. The solar radiation was provided to the system every five minutes during daylight hours. PVGIS provides guideline data for latent solar energy from photovoltaic frameworks in Europe, Africa, and South-West Asia. It is a part of the photovoltaic and solar electricity (SOLAREC) work exploring renewable energy for the European Union and private interests. In addition to latent solar radiation, it was essential to determine solar panel efficiency using excel, based on data obtained from the manufacturer. In this paper the chosen PV had an efficiency of 15,75% (Ps-p672305) category from PROPSOLAR manufactory. Each panel can be produce a maximum power of 309.9Wp. The power demand was also considered a key element in the work calculation; in addition to the load demand registered every five minutes for each apartment. And for the comparison with diesel generators, gasoline generators and the grid, actual economic data are used, without factoring in government subsidies of diesel and gasoline.

4- Energy Requirement

The load was assumed to be lighting bedroom, dining room, bathroom, kitchen, living room, the outside area (streetlight); air conditioning; and household appliances such as refrigerator, microwave, toaster, television and computer.

Table 1: Energy requirement for each dwelling

No	Application	No of units	Power per unit (W)	Operating hours per day (h)	Energy consumption W/day
1	Refrigerator	1	24	24h	576
2	Microwave	1	600	20m	200
3	Toaster	1	2200	20m	733
4	Kitchen kettle	1	2200	5m	183
5	Washing machine	1	500	30m	250
6	Water heater (boiler)	1	1200	2h	2400
7	Lights	6	60	5h	1800
8	TV	1	150	5h	750
9	Air conditioning	1	495	12h	5940
10	Computer	2	50	2h	200
11	Home internet	1	5	24h	120
12	Street light	2	100	12h	2400
				Total	15552

Table 1 describes the energy requirements for applications in one apartment. Appliance usage time and the loads were estimated for daytime usage.

Each apartment consumes about 4.3 kW; since that building has a three-floors, with two apartments in each floor, the consumption of the whole building will be 26 kW. The curve below in Figure 1 shows the daily consumption of each apartment for 24 hours, including the time in which the power consumption peaks. Figure 2 shows the electrical devices utilized in each apartment and the time in which they are used most.

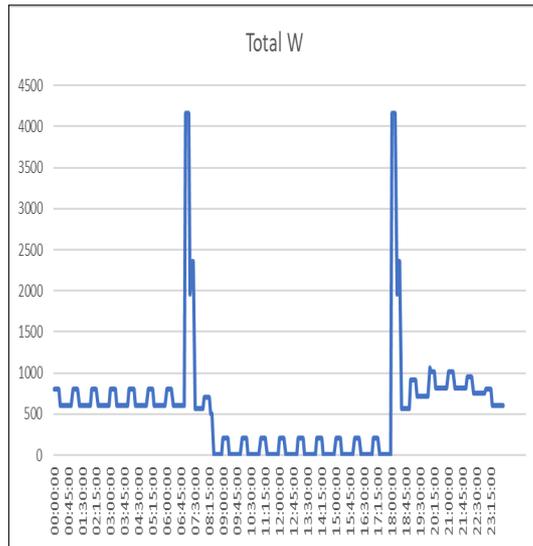


Figure 1: Daily electrical consumption.

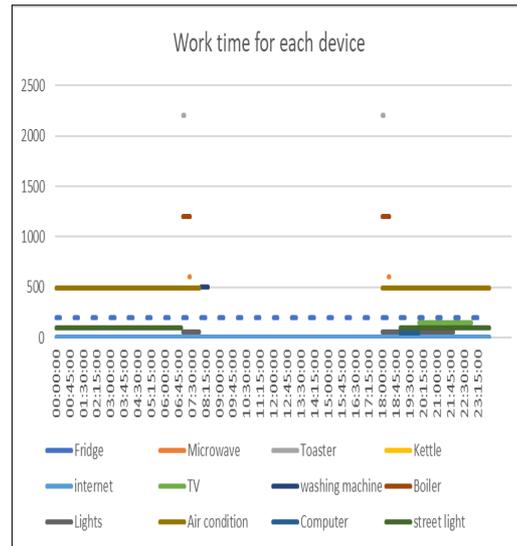


Figure 2: Daily consumption of each device

During daylight, the PV array must be capable of providing about 30 kW for the building requirements plus the electrolysis load, which is 75 kW. Fuel cell output power should not be less than 30 kW to ensure that the electricity demand of the building can be met at night.

To ensure a modest surplus of power, it is suggested that 30 kW of power will feed the building during the day and an additional 75 kW of power will be used to provide for the generation of hydrogen. Thus, a 105 kW solar array will be necessary.

5- Results and Discuss

The results of the paper in terms of three states: when the number of solar cells was 145, 193 and 552. It also examines the economic feasibility of supplying the residential from solar-hydrogen plant. In addition, it assesses the system's need for hydrogen storage, and provides a comparison between the grid and the renewable energy regarding the technical environmental and economic basis.

Figure 3 shows solar irradiance (W/m^2) for a typical day in each month, indicating the latent solar radiance available per m^2 every 5 minutes throughout the day. It is noteworthy that the months June, July and August are the most energy productive as they represent the summer season, in which the sunlight hours are obviously longer, and solar radiance is stronger.

In the winter, particularly December (when the days are shorter), the power production is much less due to the short daylight radiance, which results in smaller hydrogen yield than in other months. Therefore the system was supplied with a considerably higher number of solar cells in order to recuperate the shortage, so that it can generate the sufficient amount of energy to operate the electrolyzer.

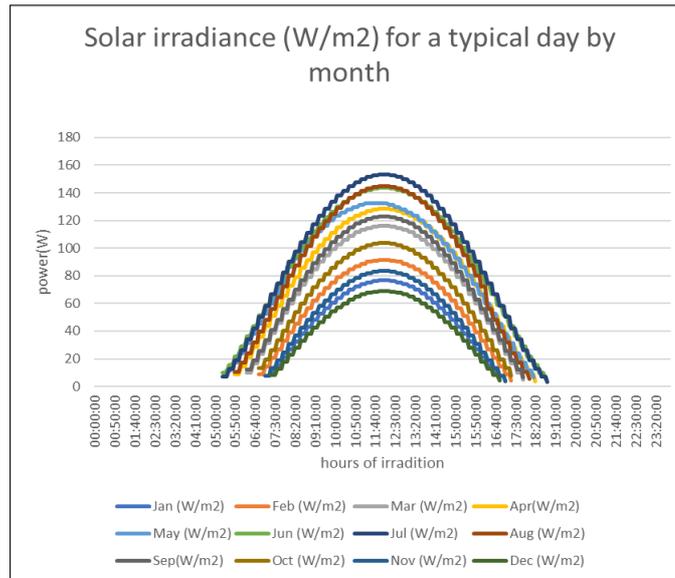


Figure 3: Solar irradiance (W/m^2) for a typical day of the month.

The solar panels' efficiency (15.75%) indicates that a great amount of power can be produced from the solar radiation in summer, ranging from 150-160 W/m^2 for almost 14 hours from 5:00 am to 18:00 pm. In the winter the produced energy is less, ranging between 65-85 Wp for every m^2 . This amount of energy is generated in shorter period, over about 10 hours between 7:00 am and 16:00 pm. This amount of produced energy gives this job a great potential of success with many solar cells (particularly to increase the efficiency of winter collection).

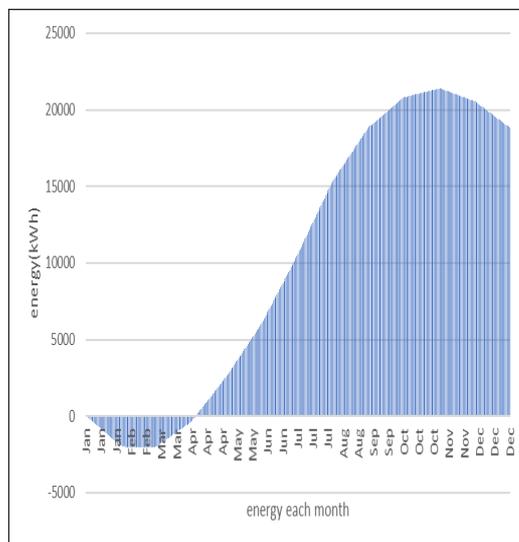


Figure 4: Deficit and surplus of accumulated hydrogen storage

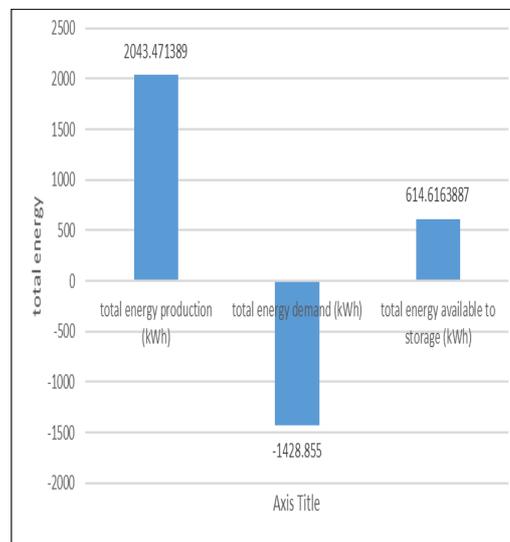


Figure 5: Hydrogen amount.

As shown in Figure 4, 193 PV cells were used in the work. When the project starts operating in January the generated hydrogen is insufficient during the first three months, due to the short daylight period and the high demand, while the rate of the generated power rises significantly at the beginning of April. It is clearly noticed that the captured solar power is very low in winter, yet it starts to increase gradually in spring and summer as the hours of sunlight increase, and the amount of generated energy is sufficient.

It is noteworthy to mention that the increasing rate of hydrogen production will guarantee a total power sustenance for the residential building, and an additional amount of hydrogen that can be stored or even utilized to supply other apartments. Consequently, the optimal month to start operation is April for securing power provision. Figure 5 shows the amount of the generated energy, approximately 2043.471389 kWh, and energy demand, which is -1428.855 kWh, leaving 614.6163887 kWh to be stored or sold.

Figure 6 illustrates the possibility of having 145 solar cells less than the number of cells in the previous cases. In this scenario, when the project starts in July the number of PV cells is decreased to minimize costs. The figure clearly shows that produced energy in the first five months was insufficient to cover the power demand, while it was significantly higher from June onwards because of longer daily sunlight. Figure 7 indicates the amount of the produced energy (1481.723079 kWh), total energy demand (-1428.855 kWh), and the surplus (52.86807899 kWh); stored hydrogen is only enough to supply the building for less than two days in the absence of more production. In this case, the power production will be enough for the whole residential unit, yet there is a small chance of power shortages in case of cloudy weather or sandstorms or high-power load.

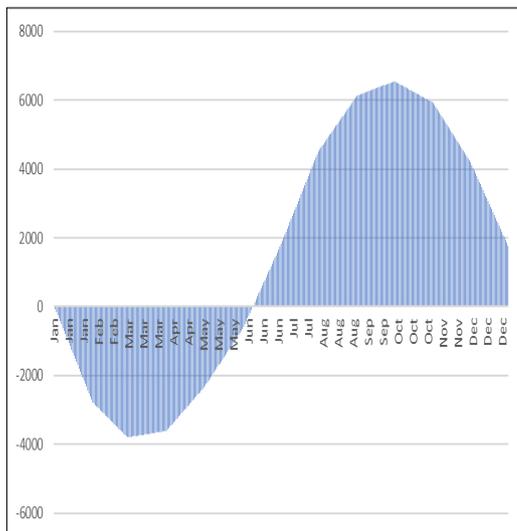


Figure 6:Energy storage content.

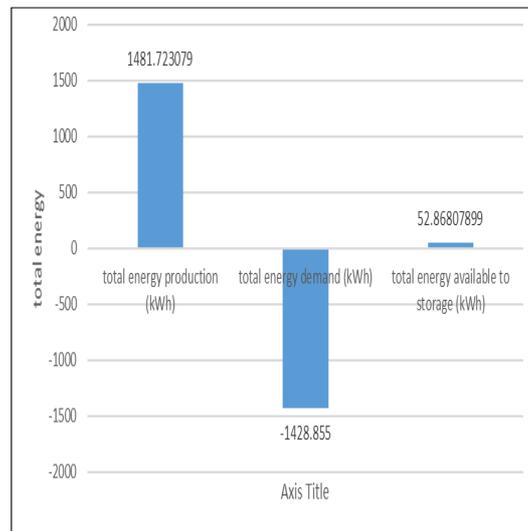


Figure 7:Hydrogen amount.

Figure 8 shows the scenario of the number of solar cells being increased to 552. Clearly the amount of produced energy is enough to supply the building, guaranteeing no shortages throughout the year. Nevertheless, the cost increases because of the increased number of PV cells. However, this offers great potential for selling stored hydrogen (or even direct electrical energy) to other buildings or the grid in the case of a hybrid system.

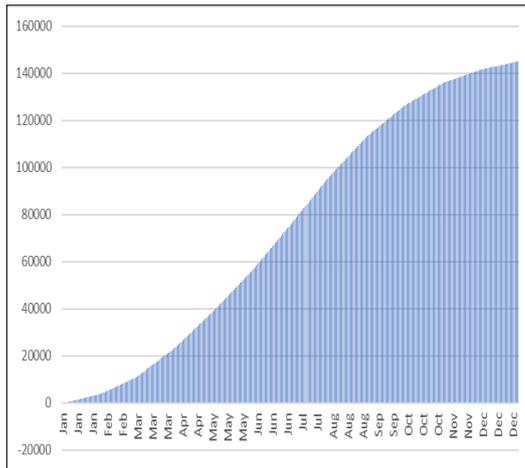


Figure 8:Energy storage content.

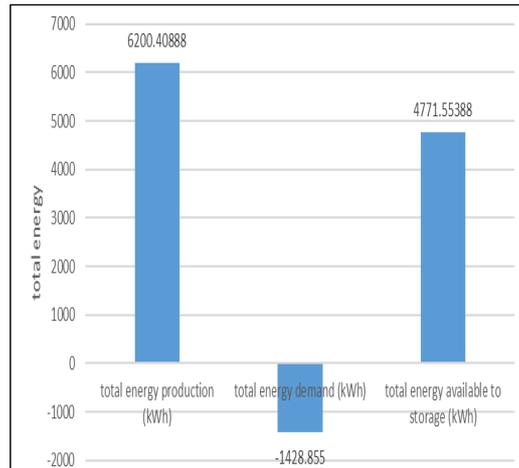


Figure 9:Hydrogen amount.

It can be noticed from figure 9 that the produced energy is 6200.40888 kWh, whereas the power demand was -1428.855 kWh, thus the power demand of the residential unit will be fully covered with a surplus of 4771.55388 kWh for storage. From the given three cases, it can be ascertained that the amount of generated hydrogen increases with the increase of the PV cells. Both methods are renewable and reliable, so producing electricity from hydrogen will be very successful in replacing petrol and gas.

From the three cases, it is evident that hydrogen can be efficiently produced from solar energy. In addition, it is possible to produce hydrogen from a stand-alone PV system in this work. Thus, hydrogen production from solar PV renewable energy source in Tripoli, Libya, is feasible both technically and economically.

6- Comparison

This section compares the financial feasibility of supplying the residential unit with electricity using hydrogen, diesel generators, gasoline generators and the grid. Actual economic data are used, without factoring in government subsidies of diesel and gasoline.

Four diesel generators are necessary to supply building demand, with each generator supplying 15 kw. The total installation cost is 80,000 LYD [12] approximately £43,000. Each needs 3.5L/h to run and the price for one liter of diesel is 0.15 LYD with government subsidy. The real prices of gasoline and diesel in the Libyan market are 0.45LYD and 0.65 LYD per liter, respectively [13]. Every two generators work for 12h/d, requiring 7L of diesel per hour. The volume of diesel required for two generators to work 12h:

$$(L) = (3.5+3.5L/h) * 12h = 84L$$

The volume of diesel required for four generators to work 24h:

$$(L) = (84 + 84) = 168L/Day$$

The cost of diesel for four generators to work 24h:

$$168L * 0.65 LYD = 109.2LYD /day$$

Therefore, the monthly diesel cost = 109.2 * 30 = 3276 LYD/month (approximately £1770).

In addition, the gasoline generator has the same price and number if generators, however it needs 4.5L/h; however, there is a difference in the amount of gasoline per hour and the price of gasoline.

Volume of gasoline required for two generators to work 12h:

$$(L) = (4.5+4.5L/h) * 12h = 108L$$

Volume of gasoline required for four generators to work 24h:

$$(L) = (108 +108) = 216 L/day$$

The cost of gasoline for four generators to work 24h:

$$216 * 0.45LYD = 97.2 LYD/day$$

Therefore, the monthly gasoline cost = $309.0924 * 30 = 2916LYD/month$ (approximately £1576).

Both diesel and gasoline generators need two people to work on their maintenance, one in the morning and the other at night, with a salary of £1500 per month.

On the other hand, electricity from the grid will be the cheapest. In Libya, the price for 1 kWh is 0.02 dirhams [14] with a 150LYD fee to connect to the power supply for lighting and housing apartment buildings consisting of two floors or more according to Libyan General Electric Company. However, the price for kWh without support of government is 0.1677LYD according to the Zawia CCPP Maintenance Manager [15].

The cost of electricity per (kWh) = the energy demand for six flats (30 kWh) * kWh unit price (0.1677LYD) = 5.031LYD

Therefore, the monthly cost of electricity is $5.031LYD * 30 = 150.93LYD$ (£81.583). The cost and efficiency of using a dedicated hydrogen plant is shown in (Table 2), using data from [16], considering the building cost and four staff members (two maintenance workers and two supervisors working two shifts, each with a salary of £1500 per month).

Table 2: Hydrogen system cost

No	Type of device	The cost (£/kW)	Efficiency (%)
1	Electrolyzer	£1177	70%
2	Compressor	£784	90%
3	Hydrogen tank	£690	95%
4	Converter	235	95%
5	Fuel cells	£1569	50%
6	PV	£0.41/W = £410/kW	15.75%
7	The building	£10,000	
8	Total	£14,865	

Figure 10 demonstrates the operating bills for the grid, diesel generator, gasoline generator and hydrogen plant options. It is clear that the prices of using gasoline and diesel generators converge, while the hydrogen cost is slightly substantially higher, while the grid price is negligible. Furthermore, as shown in Figure 11, the equipment cost of hydrogen is of vastly greater magnitude than the diesel and gasoline generators, costing about £450,000.

Utilizing renewable energy sources particularly solar in hydrogen generation is intrinsically feasible both economically and technically, but obviously the only sensible option for a Libyan citizen is to source electricity from the grid if possible, due to the overwhelming cost advantage. However, this could change in the long term if the government follows the example of many Gulf Cooperation Council (GCC) countries and reduces subsidies of national oil and gas, and instead subsidies hydrogen generation, which has a beneficial net effect on the national economy over the long term,

as it is dependent on oil exports [17]. In addition to the more important objectives of being environmentally friendly and reducing Green House Gase (GHG) emissions [18].

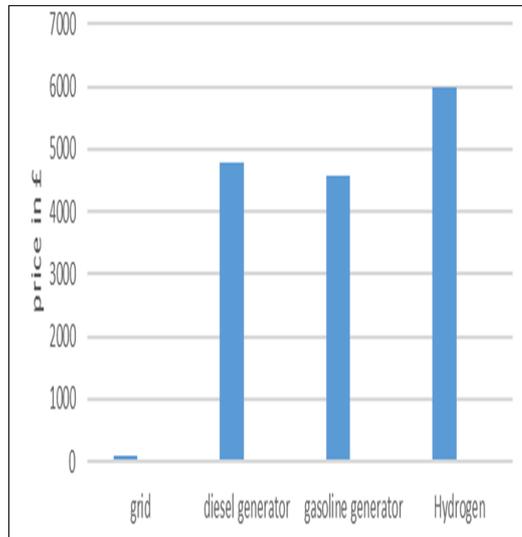


Figure 10:Monthly operating bill.

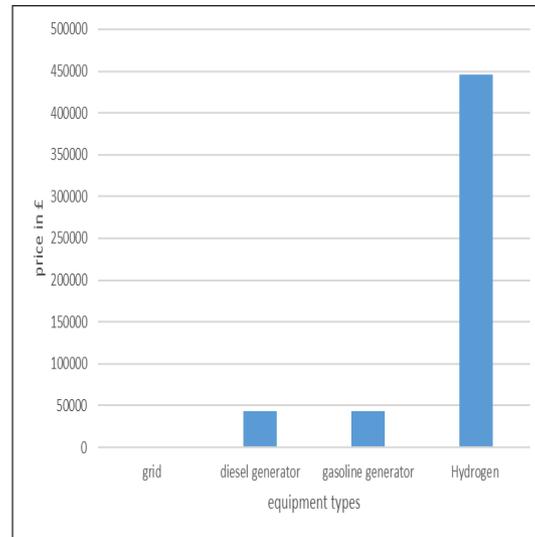


Figure 11:Equipment cost.

7- Conclusions

Hydrogen is a clean fuel that produces only water on combustion or when combined with oxygen in fuel cells to produce electrical power. Like electricity, it is an energy carrier, and it has potential for energy storage, transportation and electricity generation for countless outlets, such as lighting, heating, and powering motor vehicles. This paper focused on generating renewable electricity from combined solar-hydrogen system. This paper has also investigated the capacity and the efficiency of the solar-hydrogen plant to provide a community of electricity for a year without any shortages or deficiency. In order to ensure the successfulness of the development, solar radiance was collected from the Global Radiance Agency along with the manufacturers' data on solar panels' efficiency to calculate the daily solar power generated. The produced solar power that is utilized in electrolysis to generate the hydrogen enables the later back-transformation of hydrogen into electricity.

The development can be completely felicitous throughout the whole year in terms of technically meeting the power demand. Despite the high cost of plant, it is still the best method as it is renewable and environmentally friendly, and it produces energy from sunlight and water. The high cost of the system can be overcome by gaining government support for hydrogen to electricity developments as an infinite energy carrier, preserving oil and gas reserves for export. From the previous comparison it can be said that hydrogen is the fuel of the future and it can be produced from the most abundant sources, sunlight and water. However, the cost of producing hydrogen is high compared to petrol and gas, mainly related to installation and equipment costs, which will decrease over time as technologies improve and market competition increases, and which will become increasingly cost competitive with fossil fuels as the latter diminish, dwindle and ultimately expire. Thus, while there is little in the short term to recommend the utilization of solar-hydrogen renewable energy generation and renewable economically, considering long-term prospects and the reliance of the national economy on oil and gas exports, it is clearly sensible to start preparing for a greater role for renewable.

This paper is considered to be the first hydrogen study in Libya it will gain a lot of public acceptance as it is an environmentally friendly and safe facility demonstrating the feasibility of the use of alternatives to fossil fuels in the country.

8- Acknowledgment

Third Conference for Engineering Sciences and Technology (CEST-2020) Alkhoms - Libya

The Authors thanks Azahra Higher Institute of sciences and Technology and Cardiff University for their support.

References

- [1]. E. COMMISSION, “Hydrogen Energy and Fuel Cells,” 2003.
- [2]. U. S.-E. U Summit, “Hydrogen Economy Fact Sheet,” *Cooperation on the Development of a Hydrogen Economy*, 2003. [Online]. Available: <https://georgewbush-whitehouse.archives.gov/news/releases/2003/06/20030625-6.html>. .
- [3]. G. W. Crabtree, M. S. Dresselhaus, and M. V Buchanan, “*If the fuel cell is to become the modern steam engine , basic*,” no. December, pp. 39–45, 2004.
- [4]. J. K. Dahl *et al.*, “*Rapid solar-thermal dissociation of natural gas in an aerosol flow reactor*,” vol. 29, pp. 715–725, 2004.
- [5]. I. Papagiannakis, “*Studying and Improving the Efficiency of Water Electrolysis Using a Proton Exchange Membrane Electrolyser*,” 2005.
- [6]. P. L. Spath, M. K. Mann, and M. K. Mann, “*Life Cycle Assessment of Hydrogen Production via Natural Gas Steam Reforming*,” no. February, 2001.
- [7]. E. Centurioni, D. Iencinella, R. Rizzoli, and F. Zignani, “*Silicon Heterojunction Solar Cell: A New Buffer Layer Concept With Low-Temperature Epitaxial Silicon*,” vol. 51, no. 11, pp. 1818–1824, 2004.
- [8]. O. Behar, A. Khellaf, and K. Mohammedi, “*A review of studies on central receiver solar thermal power plants*,” *Renew. Sustain. Energy Rev.*, vol. 23, pp. 12–39, 2013.
- [9]. T. Solar *et al.*, “*Performance Improvement and Energy Cost Reduction under Different Scenarios for a Parabolic*,” vol. 2019, pp. 1–23.
- [10]. “PHOTOVOLTAIC GEOGRAPHICAL INFORMATION SYSTEM,” *PV potential estimation utility*, 2016. [Online]. Available: https://re.jrc.ec.europa.eu/pvg_tools/en/tools.html.
- [11]. P. E. Dodds, I. Staffell, A. D. Hawkes, and F. Li, ScienceDirect “*Hydrogen and fuel cell technologies for heating: A review*,” *International Journal of Hydrogen Energy*, vol. 0, 2014.
- [12]. D. Paper, “*Issues in China Power Sector Reform: Generator Dispatch*,” no. July, 2016.
- [13]. Brega petroleum marketing Co. A.Daw@brega.ly.
- [14]. RCREEE, “*PA-CEMP program in Libya*,” 2016. [Online]. Available:<https://www.rcreee.org/content/pa-cemp-program-now-libya>.
- [15]. Gaeer, A. (2016). *The cost of electricity in Libya* [email].
- [16]. O. V Marchenko and S. V Solomin, ScienceDirect “*The future energy: Hydrogen versus electricity*,” *Int. J. Hydrogen Energy*, vol. 40, no. 10, pp. 3801–3805, 2015.
- [17]. Live Population, “*Population of Libya (2019)*,” *live population*, 2019. [Online]. Available: <https://www.livepopulation.com/country/libya.html>.
- [18]. S. G. Elsharkawy and E. S. Elmallah, “*Spatiotemporal investigation of long-term seasonal temperature variability in Libya*,” vol. 179, pp. 535–549, 2016

An Energy Efficiency Evaluation of MIMO Based LTE RANs with DTX Operation

Dr. Abdelrahman Arbi^{1*}, Prof. Timothy O'Farrell²

¹ asarbi@elmergib.edu.ly, ² t.ofarrell@sheffield.ac.uk

¹ Department of Electrical and Computer Eng., Faculty of Eng., Elmergib University, Libya

² Department of Electronic and Electrical Eng., Faculty of Engineering, University of Sheffield, UK

ABSTRACT

This paper presents a system level evaluation of the energy consumption of a 4G LTE radio access network RAN when upgrading the access base stations from SISO to MIMO base stations. A macro only cell deployment scenario is evaluated, and the power consumption of the base stations is estimated by a parametric power consumption model. The impact of the discontinuous transmission feature (DTX) on the energy consumption of 2x2 and 4x4 MIMO radio access network with reference to the SISO radio access network is analyzed. A non-full buffer FTP traffic model is used in the analysis. Our results show that, without DTX, the 2x2 MIMO deployment consumes the least amount of energy. No energy savings are observed when more than 2 transmit antennas is used. When DTX is enabled, and at high values of offered traffic, the 4x4 MIMO deployment becomes the most energy efficient deployment option among the SISO and 2x2 MIMO deployments.

Keywords:

Energy Efficiency,
MIMO radio access
networks,
Energy Consumption Gain,
Power consumption model,
DTX.

1. Introduction

Multiple-In-Multiple-Out (MIMO) techniques are used in RANs to enhance cell capacity through spatial multiplexing or cell coverage through spatial diversity, without increasing the channel bandwidth or the RF transmit power. However, the power consumption of a MIMO base station increases due to the need for more signal processing and more RF chains when compared with a Single-Input-Single-Output (SISO) base station. The question of whether deploying MIMO base stations in a RAN is more energy efficient than SISO base stations or not is addressed in this paper. The energy consumption of SISO and MIMO RANs is evaluated when the base station circuit power consumption is considered and when more realistic non-full buffer FTP traffic model is used. The impact of enabling the fast DTX on the RAN energy consumption is also assessed.

Many previous publications have investigated the issue of MIMO energy efficiency. The authors of [2] have evaluated the MIMO energy efficiency in wireless sensor networks (WSNs) employing Alamouti diversity. Contra to the widely held view that MIMO systems would be more energy efficient than SISO ones, the authors concluded the opposite when the overhead circuit energy consumption was considered. Their research demonstrated that in short-range applications, SISO WSNs can be more energy efficient than MIMO WSNs.

The energy consumption ratio (ECR) measured in J/bit has been investigated for both SISO and 2x2 MIMO Alamouti schemes in [3], for both the Round Robin (RR) and Proportional Fair (PF) scheduling algorithms. When only the RF transmit power is considered, the authors concluded that the 2x2 MIMO transmission mode (i.e. the Alamouti space frequency block code) is both more spectrally and energy efficient than SISO in both the urban micro and urban macro cell deployment

scenarios. Similarly, the authors of [4] investigated the MIMO energy efficiency for different cell types and have observed that a significant SINR improvement can be obtained with MIMO (e.g. 7.1 dB for a 2×2 configuration) independent of the cell type. However, an increase in base station power consumption is observed due to the increased complexity of the transceiver circuitry, and the additional signal processing needed. The authors of [5] have evaluated the MIMO energy efficiency at the link level and have shown that MIMO does not provide any energy efficiency gain when more than two transmit antennas are used if the circuit overhead power consumption is included in the analysis.

In [6], the energy consumption of MIMO and SISO RANs under the same coverage and average traffic load conditions were compared. The results demonstrated that although a single MIMO base station consumes more energy than its SISO counterpart when transmitting at full power, the MIMO RAN can be more energy efficient as smaller number of MIMO base stations than SISO ones is needed to offer the same capacity.

Most of the previous publications considered only the case a full buffer traffic model, which unrealistically assumes that there is always data to be sent. In this paper, the authors compare the energy consumption of MIMO and SISO RANs for a bursty traffic model. Such a non-full buffer traffic model enables the energy efficiency to be investigated as a function of cell offered average traffic load. In addition, the impact of fast discontinuous transmission (DTX) at the base station is included to the analysis [7, 8]. Furthermore, the process of modelling the power consumption of a radio base station is described by adopting the same model developed in [9]. A system level MATLAB simulator of 4G LTE RAN was developed by the authors to determine the RAN average traffic load, and average throughput with users capable of operating in SISO and MIMO modes.

The paper is organized as follows: Section 2 describes the processes of modelling the power consumption of a radio base station. Section 3 evaluates the energy savings due to enabling the DTX in a radio base station. The MIMO channel capacity estimation is covered in Section 4. The analysis and network model are presented in Section 5, followed by the results and the discussion of the results in Section 6. Finally, Section 7 concludes the paper.

2. Radio Base Station Power Consumption

The radio base station (RBS) consumes the largest proportion of energy in a cellular network. In a homogeneous macro cell RAN, RBSs typically consume ~60 % of the total RAN energy [10]. Base station sites usually consist of seven common units identified as: the antennas and feeder cables; the power amplifier (PA); the RF transceiver; the baseband processing unit BBU; the backhaul unit; the cooling unit; and the AC/DC power supply unit. Figures 1 and 2 show these common units in SISO and 2x2 MIMO base stations. The enhanced Green Radio (GR) base station power consumption model which was developed in [9] is used to estimate the power consumption of a base station in this paper. This model predicts the power consumption of the base station versus the cell average traffic load at various number of radio chains (MIMO order).

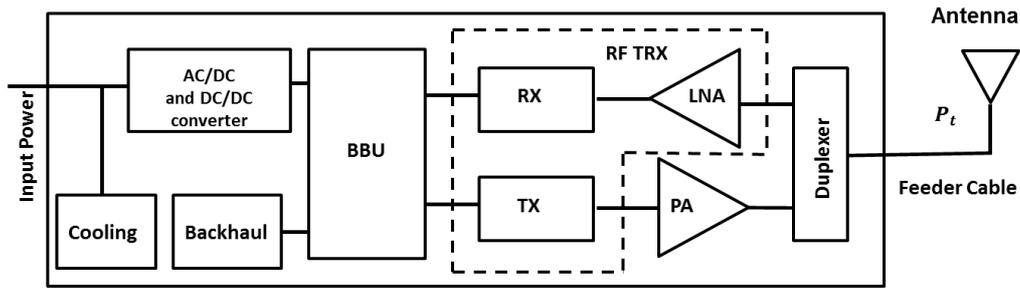


Figure 1: The SISO radio base station simple architecture

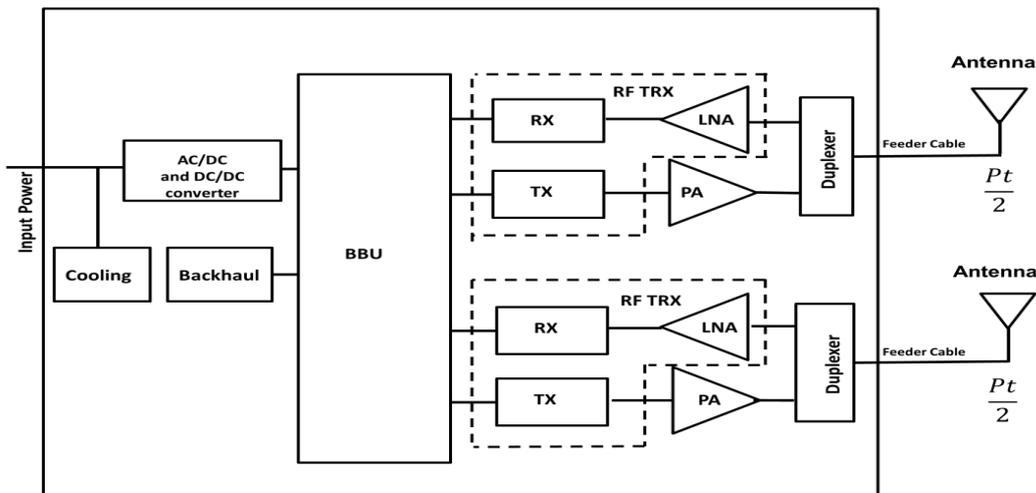


Figure 2: The 2x2 MIMO radio base station simple architecture

The average power consumption of a base station site is calculated by summing the power consumption of the base station units, as expressed in (1) and (2).

$$P_{site} = P_{cool} + P_{bh} + P_{bts} \quad (1)$$

$$P_{bts} = P_{rect} + n_s \cdot n_t (P_{bb} + P_{trx} + P_{pa}) \quad (2)$$

In (1) and (2), the term in P_{site} represents the total base station site power consumption, P_{cool} denotes the cooling system power consumption, P_{rect} is the power consumed by the power supply unit, P_{bh} is the backhaul power consumption, and P_{bts} is the base transceiver station power consumption. The term P_{bb} is the baseband signal processing power consumption, P_{trx} is the RF transceiver power consumption, P_{pa} is the power consumed in the power amplifier, n_s is the number of sectors per site, and n_t is the number of transmit antennas per sector.

2.1 Power Amplifier

The power amplifier (PA) consumes the most power in a macro cell base station and its consumption depends on the average transmit RF power per antenna, the power amplifier efficiency η_{pa} , and the feeder cable losses σ_{feed} as shown in (3).

$$P_{pa} = \frac{\alpha \cdot P_{t,max}}{n_t \cdot \eta_{pa} \cdot (1 - \sigma_{feed})} \quad (3)$$

In (3), $P_{t,max}$ is the maximum transmitted RF power per sector while α denotes the normalised average offered load factor during a nominal observation time and has a value between 0 and 1. The value of the power amplifier efficiency η_{pa} at a specific transmit power can be estimated by using (4) [11], where $\eta_{pa,max}$ is the PA maximum efficiency corresponding to $P_{pa,max}$.

$$\eta_{pa} = \sqrt{\frac{\alpha \cdot P_{t,max}}{P_{pa,max}}} \cdot \eta_{pa,max} \quad (4)$$

2.2 RF Transceiver and Baseband Processing Unit

The RF transceiver TRX contains circuits for clock/carrier generation and distribution, modulators, mixers, filters, buffers, low noise amplifier and the analogue/digital converters. The RF transceiver power consumption per sector can be estimated by (5) as a function of the power consumption of a reference base station configuration [12].

$$P_{trx} = \frac{n_t \cdot B \cdot P_{trx,ref}}{B_{ref}} \quad (5)$$

In (5), the term $P_{trx,ref}$ is the RF transceiver power consumption at a reference bandwidth B_{ref} , assuming a single transmit antenna per sector. The baseband processing unit power consumption is usually modelled as being dependent on the number of transmit antennas and the system bandwidth. The base band power consumption is estimated by (6), where $P_{bb,ref}$ is the baseband unit BBU power consumption at a reference bandwidth B_{ref} , also assuming a single transmit antenna per sector.

$$P_{bb} = \frac{n_t \cdot B \cdot P_{bb,ref}}{B_{ref}} \quad (6)$$

2.3 Power Supply and Cooling Unit

The Power supply unit comprises the AC/DC and DC/DC converters. The power losses in this unit are usually modelled as a fixed percentage (10% to 15%) of the overall power consumption of the base station [13]. Site cooling is required at macro-cell sites to maintain a suitable operating temperature inside the base station cabinet. The power consumption of the cooling unit is estimated using (7) by calculating the amount of heat generated inside the cabinet in British Thermal Units (BTU) per hour. The feeder cable losses σ_{feed} are considered by subtracting the transmitted RF power per sector from the base transceiver station power consumption. The energy efficiency rating EER of a cooling unit indicates the number of BTUs per hour removed for each 1 W of consumed power, while the number 3.4121 is used to convert from BTU to watts.

$$P_{cool} = \frac{3.4121}{EER} \cdot \left(P_{bts} - \frac{n_s \cdot P_t}{1 - \sigma_{feed}} \right) \quad (7)$$

2.4 Overall Base Station Power Consumption

The overall base station site power consumption can be estimated as the sum of the power consumption of the base station subsystems in (1) and (2). Table 1 lists the parameters of the power model used to estimate the total power consumption of the macro base station sites for SISO and different MIMO schemes. These values are obtained from [14] while the PA efficiency values are estimated by (4).

Table 1: Macro cell BTS power model

Parameter	Macro 1x 40 W	Macro 2x20 W	Macro 4x10 W
Sectors	1	1	1
P_t /Antenna (dBm)	46	43	40
Feeder losses (dB)	3	3	3
PA Back-off (dB)	8	8	8
$P_{pa,max}$ (dBm)	57	45	51
$\eta_{pa,max}$	90%	90%	90%
η_{pa} at P_t	25.37%	25.37%	25.37%
Bandwidth (B_{ref} , MHz)	20	20	20
Transceiver Unit ($P_{trx,ref}$, W)	13	13	13
Processing Unit ($P_{bb,ref}$, W)	30	30	30

Figure 3 shows the macro cell base station site power consumption versus the average traffic load factor α . The average traffic load α refers in LTE to the frequency-time resource utilisation factor in the cell. When $\alpha = 1$, the base station transmits at full power which equals to 40 watts per antenna for a SISO base station, and 20 watts per antenna in case of 2x2 MIMO base station. Figure 3 shows also that reducing the macro cell average traffic load results in a decrease in the site power consumption.

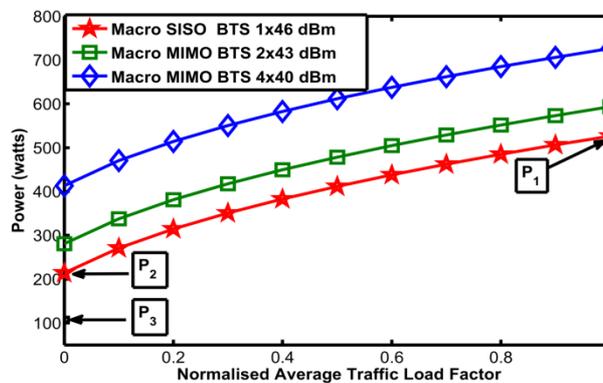


Figure 3: Macro cell power consumption for various MIMO schemes

3. Discontinuous Transmission (DTX)

Large energy savings can be achieved if some or all the base station subsystems are put into fast sleep mode during periods of low traffic intensity [15]. Under normal operation, an LTE base station transmits broadcasting, synchronization, and cell reference control signals, even when there is no traffic in the cell. This results in a small idle time to implement the DTX. One solution is to implement a fast DTX mechanism during the MBSFN (multicast and broadcast single frequency network) sub-frames. The LTE Rel-8 standards allow up to 6 MBSFN sub-frames to be transmitted per LTE frame. Each MBSFN sub-frame carries the CRS (cell specific reference) signal in the first LTE symbol. The sub-frame duration is 1 ms, which represent 10% of the frame duration, and consists of 14 LTE symbols, each symbol of 71.4 μ s duration. By assuming a 30 μ s sleep state transition time in the base station [7], the remaining available idle time for DTX is $(13 \times 71.4 - 30)$ μ s. A maximum of 6 sub-frames can be configured in MBSFN. Therefore, the maximum percentage of idle time $\alpha_{DTX,max}$ in 6 sub-frames over one LTE frame is given by (8).

$$\alpha_{DTX,max} = \frac{6 \times (1000 - 71.4 - 30)}{10000} = 0.5392 \quad (8)$$

When fast sleep mode is enabled, the cell may be in one of three states as follows: 1) transmitting at full power when there is a traffic present in the cell, thereby consuming P_{ON} (i.e. P_1 in Figure 3); 2) transmitting only the control signals, thereby consuming P_{OH} (i.e. P_2 in Figure 3) when there is no traffic in the cell; and 3) the base station is in DTX, thereby consuming P_{DTX} (i.e. P_3 in Figure 3). The P_{DTX} is expressed as a function of P_{OH} as shown by (9), where δ_{DTX} is the base station DTX factor.

$$P_{DTX} = \frac{P_{OH}}{\delta_{DTX}} \quad (9)$$

When DTX is not enabled, $\delta_{DTX} = 1$, and δ_{DTX} is greater than one when DTX is enabled. The larger the value of δ_{DTX} , the more amount of reduction in the power during the DTX. The intensity of the average traffic load in the cell determines the duration of the DTX period in the cell. If the cell average traffic load α_{ON} is known, then the proportion of DTX time in one frame is estimated by (10).

$$\alpha_{DTX} = \min \left(\left(\left\lfloor \frac{1 - \alpha_{ON}}{0.1} \right\rfloor \right) \times 0.0899, \alpha_{DTX,max} \right) \quad (10)$$

The number "0.0899" equals the ratio of maximum idle time in one sub-frame to the LTE frame duration. The term in $\left\lfloor \frac{1 - \alpha_{ON}}{0.1} \right\rfloor$ is equal to the number of idle sub-frames in one LTE frame. Let α_{OH} denote the ratio of the time spent transmitting the reference and control signalling to the total time duration of an LTE frame, as expressed in (11).

$$\alpha_{OH} = 1 - \alpha_{ON} - \alpha_{DTX} \quad (11)$$

If the values of δ_{DTX} , α_{ON} , α_{DTX} , α_{OH} , P_{DTX} , P_{OH} and P_{ON} , are known, the base station site power consumption during DTX mode can be calculated by (12).

$$P_{site} = \alpha_{ON} \cdot P_{ON} + \alpha_{OH} \cdot P_{OH} + \alpha_{DTX} \cdot P_{DTX} \quad (12)$$

By substituting (9) into (12), the base station power consumption during the DTX mode is given by (13).

$$P_{site} = \alpha_{ON} \cdot P_{ON} + P_{OH} \left(\alpha_{OH} + \frac{\alpha_{DTX}}{\delta_{DTX}} \right) \quad (13)$$

Figure 4 shows the P_{site} versus the cell average traffic load and DTX factor δ_{DTX} for a SISO macro base station. The curves show that enabling DTX gives significant power savings at low traffic loads. While higher values of δ_{DTX} lead to more savings in the base station power consumption.

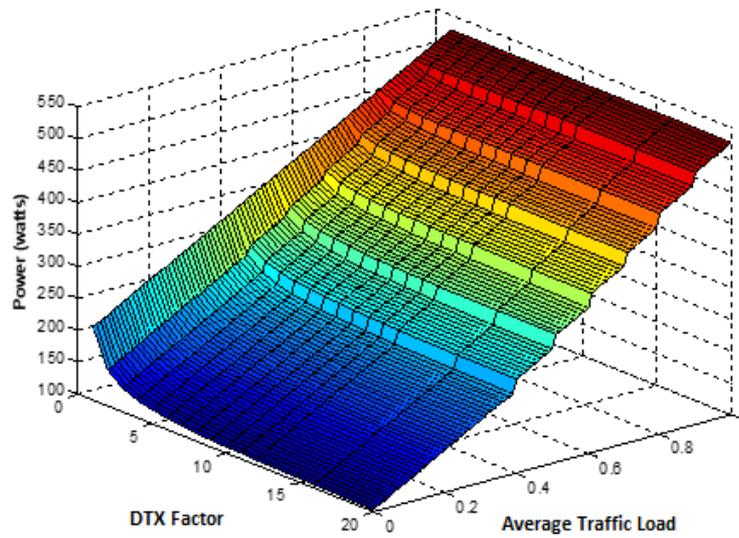


Figure 4: The power consumption of a SISO macro base station with DTX

4. MIMO Channel Capacity

Two MIMO operating modes are considered in this paper: 1) a spatial multiplexing MIMO mode with channel state information at the receiver (CSIR); and 2) a spatial diversity MIMO mode based on dominant eigenmode beamforming. The bit/s/Hz channel capacity of an $N_t \times N_r$ single user CSIR MIMO scheme, with N_t transmit antennas and N_r receive antenna, is given by (14) [16], where H is the MIMO channel matrix.

$$C_{CSIR} = \log_2 \left[\det \left(I_{N_r} + \frac{SNR}{N_t} \cdot HH^H \right) \right] \quad (14)$$

In (15), the CSIR MIMO channel capacity is expressed as the sum of N_t channels in terms of the eigenvalues λ_i of HH^H , where SNR represents the signal to noise ratio at the receive antennas.

$$C_{CSIR} = \sum_{i=1}^{i=N_t} \log_2 \left(1 + \frac{SNR}{N_t} \cdot \lambda_i \right) \quad (15)$$

For dominant eigenmode beamforming, only one information stream is transmitted, and all the power is allocated to the channel which has the largest eigenvalue λ_{max} . The channel capacity in this case is given by (16).

$$C_{eig} = \log_2(1 + SNR \cdot \lambda_{max}) \quad (16)$$

Figure 5 shows the channel capacity versus SINR of the SISO and the two MIMO schemes. The MIMO channels are modelled as unit variance Gaussian Independent and Identical Distributed zero mean complex processes.

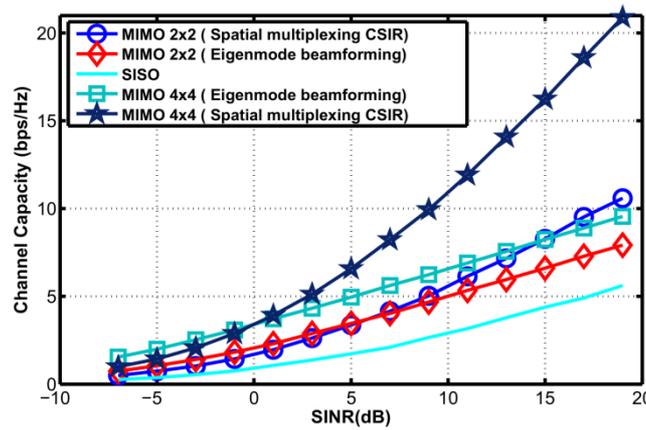


Figure 5: SISO and MIMO channel capacities

5. Energy Analysis and Network Model

5.1 Energy Efficiency Analysis

The average site power consumption P_{site} is estimated by the presented power model in Section 2 as a function of the number of transmit antennas and the cell average traffic load factor. The consumed energy in the base station site is given by (17), where T_{OH} is an observation time.

$$E = T_{OH} \cdot P_{site} \quad (17)$$

We determine the energy efficiency using the energy consumption gain (ECG) [9], which defined as the ratio of the energy consumption E_{SISO} of a reference RAN, here comprised of N SISO base stations, to the energy consumption E_{MIMO} of a test RAN, here also comprised of N MIMO base stations as shown in (18).

$$ECG = \frac{E_{SISO}}{E_{MIMO}} \quad (18)$$

The ECG is evaluated at user arrival rates per cell ranging from 0.5 to 1.5 arrivals per second, whereby each user is considered to download a 2 MB file and then departs the network. The user arrival rate is an indication of the average offered traffic per cell. For example, 0.5 arrivals per second is equivalent of $0.5 \times 2 \times 8 = 8$ Mbit/s. The ECG is estimated for the various considered MIMO RANs with reference to a SISO RAN comparing a DTX enabled sites with the disabled case.

5.2 The Network Model

A homogeneous multicell, multiuser RAN with a hexagonal cell layout and omnidirectional base station antennas is simulated in the developed MATLAB simulator. Performance statistics are collected only from the centre cell. The SINR values are calculated by considering only the six first tier of interfering cells. A static system level simulator is used to estimate the users' average throughput as well as the cell average traffic load for the SISO and MIMO RANs. A non-full buffer FTP traffic model with a variable arrival rate, described in 3GPP document 36.814 is adopted. Users dynamically select between spatial multiplexing or diversity modes to achieve a maximum user throughput. Performances between SISO and MIMO schemes are compared for same values of offered traffic and same number of base stations.

6. Results and Discussion

6.1 RAN Energy Efficiency Without DTX

Table 2 lists the main configuration parameters used in the RAN simulation. The cell average traffic load values for the various deployment schemes, given in Table 3, are used to determine the macro cell RAN power consumption for different cell average offered traffic values. The ECG is calculated by (18) for 2x2 and 4x4 MIMO RANs with the SISO RAN as a reference and plotted in Figure 6 for the non-DTX case. The ECG versus the cell average offered traffic of the 2x2 MIMO RAN is nearly always greater than unity. That means, though the SISO RAN can meet the offered traffic requirement, the 2x2 MIMO RAN meets the same offered traffic requirement while expending less energy. This is not the case for the 4x4 MIMO RAN which exhibits an ECG less than unity for all offered traffic values.

Table 2: Simulation parameters

Parameter	Omni Macro Cell
Frequency (MHz)	2000
Site Range (km)	Macro 0.5
Antenna Gain (dBi)	15
Wall Penetration loss (dB)	20
Traffic Model	FTP (2MB file)
User Distribution	Uniform
Resource allocation	Round Robin
Observation Time	100000 TTIs
Path loss model	3GPP Uma

Table 3: Cell average traffic load

Arrival Rate Per Second Per Cell	Cell Average Offered Traffic (Mbit/s)	Cell Average Load α_{ON}		
		SISO	MIMO 2x2	MIMO 4x4
0.50	8	0.357	0.174	0.100
0.75	12	0.609	0.299	0.172
1.00	16	0.635	0.316	0.182
1.25	20	0.776	0.399	0.232
1.50	24	0.850	0.490	0.295

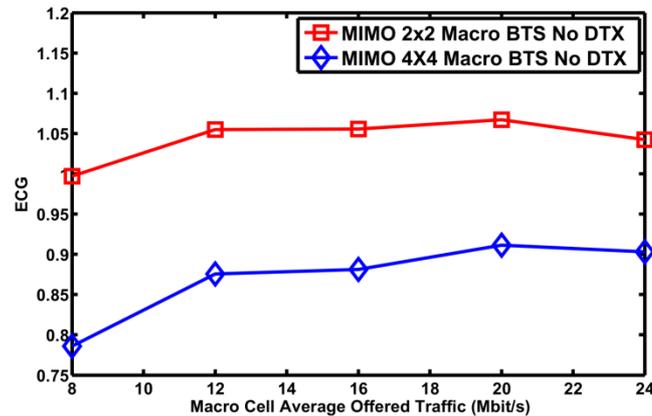


Figure 6: ECG of 2x2 and 4x4 MIMO macro cells

The trends in Figure 6 is attributed to the decrease in the cell average traffic load when increasing the number of transmit antennas. While the reduced load factor saves energy, the increase in number of RF chains consumes more energy. An optimum trade-off occurs for the 2x2 MIMO RAN. Thus, from an operator’s perspective, upgrading from SISO to 4x4 MIMO cannot be justified on the basis of energy consumption alone.

These observations contradict the widely held view that adaptive MIMO muting techniques always improve the RAN energy efficiency [17]. When measured by the ECG figure of merit, our results illustrate that switching from a 2x2 MIMO mode to SISO degrades the RAN energy efficiency. Also, the curves in Figure 6 show that the use of more than 2 transmit antennas at the base station degrades the energy efficiency. As mentioned earlier, the above results correspond to the non-DTX base station mode of operation

6.2 RAN Energy Efficiency With DTX Enabled

When DTX is enabled, the optimum trade-off between the cell average traffic load and the number of transmit antennas shifts to a new position. Figure 7 plots the curves of ECG versus cell average offered traffic for the same SISO and MIMO schemes with DTX enabled. The reference scheme is a SISO RAN without DTX enabled.

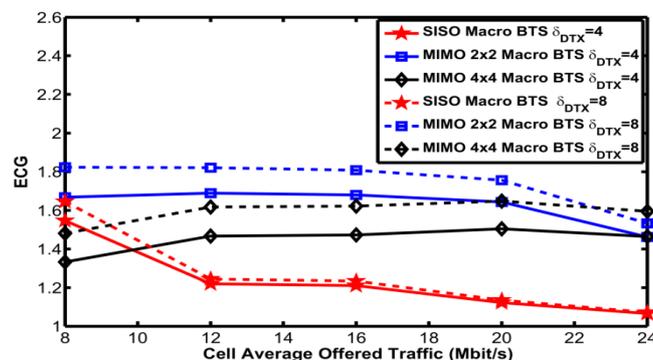


Figure 7: ECG of SISO and MIMO RANs with DTX enabled

The curves show that all the DTX cases save energy. At low cell average offered traffic values, the 2x2 MIMO RAN is the most energy efficient scheme whereas at very high offered traffic, the 4x4 MIMO RAN becomes the most energy efficient. Increasing the DTX factor δ_{DTX} from 4 to 8

leads to more energy saving in MIMO RANs than in SISO RANs which illustrates the benefits of DTX in MIMO schemes.

7. Conclusion

A system level evaluation of the energy efficiency of LTE based MIMO RANs was presented in this paper. Two MIMO schemes were considered: 2x2 and 4x4 MIMO with both the diversity and spatial multiplexing gains included in the evaluation. A non-full buffer FTP traffic model was used to model the cell average traffic load variation and to observe the model's impact on the RAN energy consumption. Moreover, the process of modelling the power consumption in radio base stations with DTX either disabled or enabled was presented.

The key findings of this paper show that for macro cells, the 2x2 MIMO RAN can provide greater energy efficiency than SISO for the same offered traffic when DTX is disabled. When DTX is enabled, both the 2x2 and 4x4 MIMO RANs are more energy efficient than the SISO RAN.

References

- [1]. W. Guo and T. O'Farrell, "Capacity-Energy-Cost Tradeoff in Small Cell Networks," in Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th, 2012, pp. 1-5.
- [2]. C. Shuguang, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," IEEE Journal on Selected Areas in Communications, vol. 22, pp. 1089-1098, 2004.
- [3]. C. Turyagyenda, T. O. Farrell, J. He, and P. Loskot, "SFBC MIMO Energy Efficiency Improvements of Common Packet Schedulers for the Long Term Evolution Downlink," in Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd, 2011, pp. 1-5.
- [4]. F. D. Cardoso, R. Torrea-Duran, C. Desset, and L. M. Correia, "MIMO strategies for energy efficient transmission in LTE pico-cell environments," in Software, Telecommunications and Computer Networks (SoftCOM), 2012 20th International Conference on, 2012, pp. 1-5.
- [5]. Heliot, M. A. Imran, and R. Tafazolli, "On the energy efficiency gain of MIMO communication under various power consumption models," in Future Network & Mobile Summit (FutureNetw), 2011, 2011, pp. 1-9.
- [6]. Arbi and T. O'Farrell, "A comparative study of energy efficiency between MIMO and SISO based LTE RANs," in Communications (ICC), 2015 IEEE International Conference on, 2015, pp. 43-48.
- [7]. P. Frenger, P. Moberg, J. Malmudin, Y. Jading, and I. Godor, "Reducing Energy Consumption in LTE with Cell DTX," in Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd, 2011, pp. 1-5.
- [8]. J. F. Cheng, H. Koorapaty, P. Frenger, D. Larsson, and S. Falahati, "Energy Efficiency Performance of LTE Dynamic Base Station Downlink DTX Operation," in 2014 IEEE 79th Vehicular Technology Conference (VTC Spring), 2014, pp. 1-5.
- [9]. Arbi, A., O'Farrell, T., Zheng, F. C., & Fletcher, S. C. (2017). Toward green evolution of cellular networks by high order sectorisation and small cell densification. In *Interference Mitigation and Energy Management in 5G Heterogeneous Cellular Networks* (pp. 1-28). IGI Global.
- [10]. H. Jianhua, P. Loskot, T. O'Farrell, V. Friderikos, S. Armour, and J. Thompson, "Energy efficient architectures and techniques for Green Radio access networks," in Communications and Networking in China (CHINACOM), 2010 5th International ICST Conference on, 2010, pp. 1-6.
- [11]. D. Persson, T. Eriksson, and E. G. Larsson, "Amplifier-Aware Multiple-Input Multiple-Output Power Allocation," Communications Letters, IEEE, vol. 17, pp. 1112-1115, 2013.
- [12]. H. Holtkamp, G. Auer, V. Giannini, and H. Haas, "A Parameterized Base Station Power Model," Communications Letters, IEEE, vol. 17, pp. 2033-2035, 2013.



- [13]. Desset, B. Debaillie, V. Giannini, A. Fehske, G. Auer, H. Holtkamp, et al., "Flexible power modeling of LTE base stations," in Wireless Communications and Networking Conference (WCNC), 2012 IEEE, 2012, pp. 2858-2862.
- [14]. G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, et al., "How much energy is needed to run a wireless network?," Wireless Communications, IEEE, vol. 18, pp. 40-49, 2011.
- [15]. S. Tombaz, S. w. Han, K. W. Sung, and J. Zander, "Energy Efficient Network Deployment With Cell DTX," IEEE Communications Letters, vol. 18, pp. 977-980, 2014.
- [16]. G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," Wireless personal communications, vol. 6, pp. 311-335, 1998.
- [17]. P. Skillermark and P. Frenger, "Enhancing Energy Efficiency in LTE with Antenna Muting," in Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th, 2012, pp. 1-5.

Maximum Area Aggregation Approach For Cumulant-Based Probabilistic Optimal Power Flow studies

Ali Tamtum^{1*}, Khalid Mohamed Alajel², Khairi Muftah Abusabee³

^{1*} aamtatum@yahoo.com, ² alajaly2005@gmail.com, ³ kmabusabee@elmergib.edu.ly

^{1*,2,3} Department of Electrical and Computer Engineering, Elmergib University, Libya

ABSTRACT

The paper introduces a Maximum Area Aggregation (MAA) approach for Cumulant-Based Probabilistic Optimal Power Flow (P-OPF) studies. The Maximum Area Aggregation (MAA) approach relies on the Cumulant Method (CM) to produce Probability Density Functions (PDFs) in the limited and the original cases, and then combines these PDFs to generate the final PDF for all system variables. The probabilities that system variables reach their limits are computed and the maximum probability is extracted and used to find the final PDF by aggregating the PDFs (the original PDFs and the limited ones). The proposed approach is verified against Monte-Carlo Simulation (MCS) consisting of 10,000 samples and compared with the original Cumulant Method (CM). The results of MAA approach demonstrate significant improvements when compared with traditional CM results.

Keywords:

Maximum Area
Aggregation (MAA), Monte
Carlo Simulation (MCS),
Optimal Power Flow,
Probabilistic Optimal
Power Flow (P-OPF),
Probability Density
Functions (PDFs).

1. Introduction

Power systems are stochastic in nature, this fact made most publications introduce computational methods for solving power flow problems using only deterministic Optimal Power Flow (OPF) approaches [1]–[5]. However, in recent years, Probabilistic Optimal Power Flow (P-OPF) problems have been developed, which address random quantities such as bus loading and generator bids, (for example [6]–[13]). Probabilistic Optimal Power Flow (P-OPF) seeks a distribution for system variables and represents this distribution using Probability Density Function (PDF) [14]–[16]; that is, each individual point on the PDF represents an optimal solution for a certain randomly generated input.

Probabilistic techniques have been used to account for random quantities within power systems since the early seventies [2]. There are many papers published regarding the introduction of random quantities, such as power demand, into the ordinary Power Flow (PF) problems [17]–[19].

Monte Carlo Simulation (MCS) technique is an easy way to account for uncertainty in power systems [17], [20]–[22]. Although this technique provides accurate results, and it is easy to implement, it involves a large number of trials, which makes it computationally expensive and time consuming. The computational expense of MCS has led to the foundation of alternative methods such as the Cumulant Method [9], [10], [23]–[29]. The Cumulant Method has been shown to be well suited for traditional OPF type problems. In [11], the authors introduce an adaptation of a Two-Point Estimate Method (2PEM), which was proposed in [30], to a market based P-OPF. Recently, a P-OPF was studied including wind generator bids as in [31]. In [12], a comparison of the two-point estimate method (2PEM) and the CM to find PDFs of Locational Marginal Prices (LMP) in a deregulated electricity market is introduced. When applied to a problem based on electricity markets (where generator bids are considered as the input uncertainty), the CM method tends to produce poor

results as the CM cannot capture discrete changes in the merit order. In this paper, a new solution approach, called the Maximum Area Aggregation (MAA), to a market-based problem is proposed based on the CM. The Maximum Area Aggregation (MAA) approach relies on the CM to produce Probability Density Functions (PDFs) in the limited and the original cases, and then combines these PDFs to generate the final PDF for all system variables.

2. Cumulant Method For P-OPF.

This section provides a brief background on the adaptation of the Cumulant Method to a Market-Based P-OPF to highlight the limitation of the standard CM when applied to such problems. Further details on the fundamentals of the Cumulant Method can be found in [9], [10], [26]–[29], [32]. The Cumulant Method (CM) is based on linearly mapping cumulants of known random variables (bus loading) into the cumulants of unknown random variables (active and reactive power generation, bus voltage magnitudes, and phase angles) [9], [14].

Mathematical Formulation of CM

Given a random variable x with a Probability Density Function (PDF) $f_X(x)$, the mean μ_x is the expected value of x and is stated as [14]:

$$\mu_x = E[x] = \int_{-\infty}^{\infty} xf_X(x)dx \quad (1)$$

Cumulants are a statistical measure of a random variable and are calculated based on the cumulant generating function $\Psi_X(s)$. The cumulant generating function $\Psi_X(s)$ is defined as the natural logarithm of the moment generating function and is written as:

$$\Psi_X(s) = \ln (\Phi_X(s)) \quad (2)$$

Where the moment generating function $\Phi_X(s)$ is defined as:

$$\Phi_X(s) = E(e^{sX}) \quad (3)$$

The n^{th} order moment is computed by taking the n^{th} derivative of (3) with respect to s and evaluating at $s = 0$.

Therefore, to generate the n^{th} order cumulant λ_n for the random variable x , the n^{th} derivative of equation (2) with respect to s is taken and evaluated at $s = 0$.

Since any given system includes many random inputs, the combination of two or more random variables is essential [32]. The linear combination of m known and independent random variables (x_1, x_2, \dots, x_m) is given by:

$$z = a_1x_1 + a_2x_2 + \dots + a_mx_m \quad (4)$$

where z is a new random variable and a_1, a_2, \dots, a_m are the mapping coefficients.

In market-based systems, prices are no longer modelled as deterministic variables; rather, prices are modelled as random variables. As a consequence, Locational Marginal Prices (LMP) change in response to changes in bids.

given as:

$$\begin{aligned} \min \quad & C^T P_{gen} \\ \text{s.t.} \quad & \mathcal{F}(X) = 0 \\ & X \geq X_{min} \\ & X \leq X_{max} \end{aligned} \quad (5)$$

where $\mathbf{C}^T P_{gen}$ is the objective cost function, \mathbf{C} is a vector of generator bids, P_{gen} is a vector of the active power generation, $\mathcal{F}(X)$ is a vector of AC power flow equations, X represents the unknown variables (active and reactive power generation, voltage magnitudes and phase angles), and X_{min}, X_{max} are the lower and upper limits.

Equation (5) represents the main market-based P–OPF problem (i.e. minimizing the cost of power generation). The formulation of such problems and the CM solution technique follows a similar pattern to that presented in [9], [10], [32]. This technique is based on incorporating bus loading and generator bids as random variables into the OPF problem. The first-order KKT optimality conditions of (5) are found and written as:

$$F(\mathbf{Y}, \mathbf{L}, \mathbf{C}) = 0 \quad (6)$$

where $F(\cdot)$ is the set of equations defining the first-order KKT conditions, \mathbf{Y} is the vector of conventional problem variables including slack variables and Lagrange multipliers, \mathbf{L} is the vector of bus loading, and \mathbf{C} is the vector of generator bids.

Lagrange Multipliers associated with AC power-flow equations as equality constraint, γ_1 , are of special interest. Those Lagrange Multipliers are directly related to the costs \mathbf{C} [33]. Moreover, γ_1 represents a vector of spot prices in the system, known as Locational Marginal Prices (LMP).

In order to find the cumulants of the conventional problem variables, \mathbf{Y} , a linear relationship between \mathbf{Y} , \mathbf{L} , and \mathbf{C} is developed by taking the full derivative of (6) as follows:

$$\mathcal{H}\Delta\mathbf{Y} + \hat{\Delta}\mathbf{L} + \hat{\Delta}\mathbf{C} = 0 \quad (7)$$

where

$$\hat{\Delta}\mathbf{L} = [0 \ 0 \ 0 \ 0 \ \Delta\mathbf{L} \ 0 \ 0 \ 0 \ 0 \ 0]^T \quad (8a)$$

$$\hat{\Delta}\mathbf{C} = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ \Delta\mathbf{C}]^T \quad (8b)$$

and $\Delta\mathbf{L}$ and $\Delta\mathbf{C}$ are vectors of the changes in the input random variables (bus loading and bid prices). The expression (7) can be rearranged into the following form:

$$\Delta\mathbf{Y} = -\mathcal{H}^{-1}\hat{\Delta}\mathbf{L} - \mathcal{H}^{-1}\hat{\Delta}\mathbf{C} \quad (9)$$

Therefore, for areas around the mean solution, changes in system variables \mathbf{Y} are linearly related to changes in bus loading \mathbf{L} and generator bids \mathbf{C} . The negative inverse of the Hessian of the Lagrangian matrix, $-\mathcal{H}^{-1}$, is used to map cumulants of the input random variables into cumulants of system outcomes.

Once the cumulants for the conventional problem variables are found utilizing (4) and (9), PDFs for these variables are then rebuilt using the Gram-Charlier series theory [9].

Comparison Methods

Comparison between the proposed approaches' results and the MCS is done through the use of the Normalized Sum of Square Error (NSSE) and the Averaged NSSE (ANSSE) [34]. In order to calculate NSSE, the sum of square error (SSE) is found using:

$$SSE = \sum_{i=1}^n (f_r(i) - f_c(i))^2 \quad (10)$$

where $f_r(i)$ is the reference value at point i (based on a 10,000 sample MCS), $f_c(i)$ is the computed value at i , and n is the total points taken for comparison. The NSSE is then computed as:

$$NSSE = \frac{\sum_{i=1}^n (f_r(i) - f_c(i))^2}{\sum_{i=1}^n (f_r(i))^2} \quad (11)$$

Finally, the ANSSE is computed using:

$$ANSSE = \frac{\sum_{i=1}^m NSSE_i}{m} \quad (12)$$

where m is equal to the number of generator buses (gb) for active and reactive power error calculation, m is equal to the number of buses (nb) for voltage magnitude error calculation, and m is equal to the number of non-slack buses (nsb) for phase angle error calculation.

3. Maximum Area Aggregation (MAA) Approach.

The original Cumulant Method presented in [9], [10], [32] does not take active limits, away from the mean solution, into account while solving the P-OPF problem. To account for limits becoming active away from the mean using the CM, the P-OPF problem is solved again with the variable being fixed at its active limit. The PDFs obtained with the variable being fixed at its active limit show the behaviour of the system and provide information relevant to a case when the limit is active. The original PDFs, together with the PDFs obtained with the variable being fixed at its active limit, are then aggregated to generate the final corrected PDFs. Probability Density Functions generation and aggregation is summarized in the following steps:

- 1) Probability density functions of all unknown variables are found using the standard CM described in [26] and [32].

$$F_{X_o}(X_o) = \begin{bmatrix} F_{V_o}(V_o) \\ F_{\delta_o}(\delta_o) \\ F_{P_o}(P_o) \\ F_{Q_o}(Q_o) \end{bmatrix} \quad (13)$$

Where $F_{V_o}(V_o) \in \mathcal{R}^{nb \times 1}$, $F_{\delta_o}(\delta_o) \in \mathcal{R}^{(nb-1) \times 1}$, $F_{P_o}(P_o) \in \mathcal{R}^{ngb \times 1}$, $F_{Q_o}(Q_o) \in \mathcal{R}^{ngb \times 1}$ are the probability density functions of voltage magnitudes, phase angles, active power generation, and reactive power generation respectively.

- 2) Active limits in the system are found by computing the areas probabilities [14, 15] that the variables values exceed the imposed limits. This is done by integrating the original PDFs, $F_{X_o}(X_o)$, past the upper limits as follows [14]:

$$\check{\alpha}_i = \int_{X_{max_i}}^{\infty} x_i f_{X_i}(x_i) dx \quad (14)$$

where $\check{\alpha}_i$ is the area under the PDF that extends past the limit X_{max_i} for the random variable X_i . It is noteworthy that if the lower limit is active, the area $\check{\alpha}_i$ (14) becomes:

$$\check{\alpha}_i = \int_{-\infty}^{X_{min_i}} x_i f_{X_i}(x_i) dx \quad (15)$$

where $\check{\alpha}_i$ is the area under the PDF that extends past the limit X_{max_i} for the random variable x_i .

The value of $\check{\alpha}_i$ indicates the ratio, expressed out of 1, that the variable exceeds its limit; that is, the variable value is greater/smaller than the limit for x_i . For m variables, the vector A is written as:

$$A = [\check{\alpha}_1 \quad \check{\alpha}_2 \quad \cdot \quad \cdot \quad \cdot \quad \check{\alpha}_m]^T \quad (16)$$

For the problems considered in this paper, m equals two times the number of generator buses plus two times the number of buses in the system minus 1, since the angle of the slack bus is assumed known (i.e. the voltage angle reference). Since the upper limit in (14) is ∞ and the lower limit in (15) is $-\infty$, none of the elements in A will equal zero. Hence, the value 0.1 was chosen as the largest negligible area; that is, the areas that are smaller than

0.1 will not be considered, since they have no significant effect. Although this value was chosen based on experience, it was found that the sensitivity of the results to this choice is low (for example, setting the threshold of 0.11 or 0.09 has negligible impact on the final results).

- 3) Since the areas presented in the set \mathcal{A} are equal to the probabilities that the variables reach their limits, the variable corresponding to the largest element in \mathcal{A} is most likely to reach its limit first. Hence, the limit corresponding to the largest element in \mathcal{A} is assumed to be the cause of the distortion. Accordingly, for the largest element in \mathcal{A} , $\check{\alpha}_{max}$ where $\check{\alpha}_{max} \geq 0.1$, the CM is then applied again to obtain new PDFs, $F_{X_{\check{\alpha}_{max}}}(X_{\check{\alpha}_{max}})$, corresponding to the case when the limit is active; that is, the P-OPF problem is solved with the variable corresponding to $\check{\alpha}_{max}$ being fixed at its active limit. Solving the P-OPF problem with the variable corresponding to $\check{\alpha}_{max}$ being fixed at its active limit results in new PDFs associated with the limited case. Using this approach, the following PDFs are found:

$$F_{X_{\check{\alpha}_{max}}}(X_{\check{\alpha}_{max}}) = \begin{bmatrix} F_{V_{\check{\alpha}_{max}}}(V_{\check{\alpha}_{max}}) \\ F_{\delta_{\check{\alpha}_{max}}}(\delta_{\check{\alpha}_{max}}) \\ F_{P_{\check{\alpha}_{max}}}(P_{\check{\alpha}_{max}}) \\ F_{Q_{\check{\alpha}_{max}}}(Q_{\check{\alpha}_{max}}) \end{bmatrix} \quad (17)$$

- 4) In this step, the PDFs obtained in Step 1, (13), and in Step 3, (17), are aggregated together to form the final corrected PDFs, $F_{X_c}(X_c)$. Since the scalar $\check{\alpha}_{max}$ presents the probability that the system is limited, the probability that the system is not limited is equal to the scalar $(1-\check{\alpha}_{max})$. Hence, the aggregation to find the final corrected PDFs is done by adding $(1-\check{\alpha}_{max}) F_{X_o}(X_o)$ to $\check{\alpha}_{max} F_{X_{\check{\alpha}_{max}}}(X_{\check{\alpha}_{max}})$ as follows:

$$F_{X_c}(X_c) = (1-\check{\alpha}_{max}) F_{X_o}(X_o) + \check{\alpha}_{max} F_{X_{\check{\alpha}_{max}}}(X_{\check{\alpha}_{max}}) \quad (18)$$

Note that equation (18) produces new PDFs which are combination of the original PDFs and the limited ones. This aggregation, (18), is applied to all unknown variables, except for the variable that corresponds to $\check{\alpha}_{max}$ in the set \mathcal{A} .

- For any variable that reaches a limit, the integration (14) is utilized to correct the original PDFs into the final corrected PDFs. The use of integration to correct the limited PDFs is summarized in the following steps:
 - a) Compute the area under the original PDF that extends past the limit. For the case where the distribution extends past an upper limit, (14) can be used to compute the area to the right of the limit. For the case where the distribution extends past a lower limit, (15) can be used to compute the area to the left of the limit.
 - b) The excess area is compressed to the limit imposed. The overall area under the PDF is still unity but the result is no longer a smooth function. However, the MCS results are also limited by the constraint in the problem and exhibit non-smooth behaviour as well.
 - c) In order to compare the P-OPF result with the MCS histogram, the probability associated with the limited value is divided by the bin width of the histogram so that the amplitudes are comparable. Once this is completed, the final bin of the histogram can be directly compared with the PDF computed by the P-OPF method and corrected through the use of integration.

4. Numerical Results

Numerical results presented in this paper are based on applying the CM using the MAA approach to the economic P-OPF problem to the IEEE 30-bus system [35]. A MATLAB program is used for simulation.

It is noteworthy that the CM results and results of the CM using the MAA approach are exactly the same if the original IEEE 30-bus data file is used; that is, without modifying the lower/upper limits, no limits are encountered. Hence, the upper limit of the active power generation at bus 5 was reduced from 1 p.u to 0.3 p.u to force the upper limit at this particular bus to be active. With this modification to the upper limit of the active power generation at bus 5, there is significant distortion to the results for other system variables.

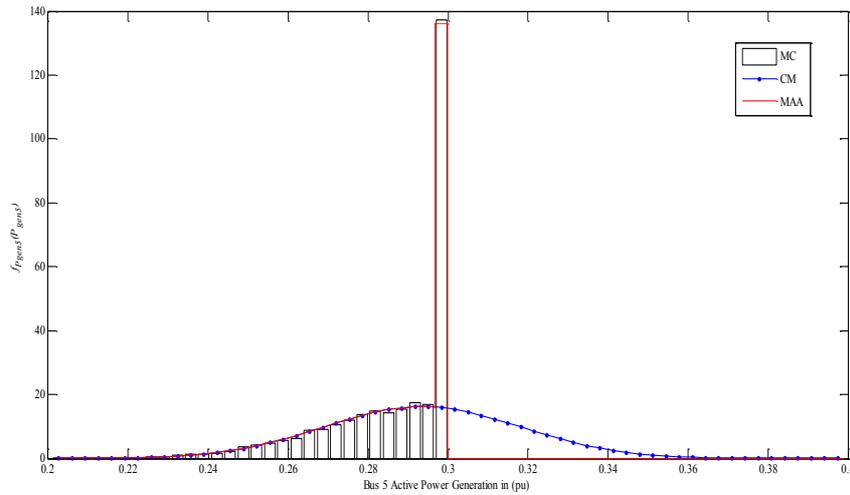


Figure 1: IEEE 30 Bus Test System, Probability Density Function of Active Power Generation at bus 5, $f_{P_{gen5}}(P_{gen5})$, Results of the CM Using Maximum Area Aggregation

Figure 1 depicts the final corrected PDF of active power generation at bus 5, when the upper limit is active. The dotted line represents the final corrected PDF, $F_{X_c}(X_c)$, the solid line represents the original uncorrected PDF, $F_{X_o}(X_o)$, and the histogram represents MCS. From Figure 1, results of the CM using the MAA approach are much closer to MCS than the original CM results.

Figure 2 depicts the final corrected PDF of active power generation at bus 8 when the upper and lower limits are not active. It is noteworthy that the MAA approach accurately estimates

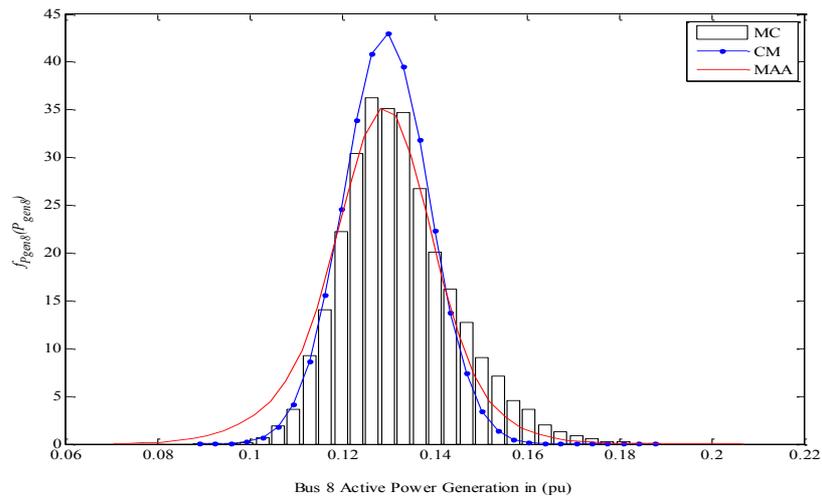


Figure 2: IEEE 30 Bus Test System, Probability Density Function of Active Power Generation at bus 8, $f_{P_{gen8}}(P_{gen})$, Results of the CM Using Maximum Area Aggregation

PDFs of system variables that have active limits, yet fails to accurately estimate PDFs of other system variables. Variables that have active limits are corrected using the integration, (14). Hence, a significant improvement is noticed, see Figure 1. The other variables, however, are corrected using (18).

The way the formula (18) combines the PDFs causes insignificant improvement in the results; that is, (18) combines both the limited and the original curves in their entirety. This approach is likely to produce satisfactory and accurate results compared with the CM against MCSs.

Table 1: IEEE 30 Bus Test System, ANSSE of Optimal Power Flow Variables for the CM and the MAA

Power System Variables	ANSSE (CM)	ANSSE (MAA)
Active Power	0.1964	0.0240
Reactive Power	0.0031	0.0029
Voltage Magnitude	0.0086	0.0066
Phase Angle	0.0296	0.0180

Average error results for the active and reactive power generation, voltage magnitudes, and phase angles, for the CM and the MAA approach. It is noteworthy that the ANSSE of the active power generation improved from 19:64% to 2:40% and for the phase angles from 2:96% to 1:8% for the CM and the MAA respectively. The ANSSE of reactive power generation and voltage magnitudes, for the CM and MAA, are also improved but not as well though.

5. Conclusion

A Maximum Area Aggregation (MAA) approach for Cumulant-Based Probabilistic Optimal Power Flow (P-OPF) studies is introduced. The new approach relies on the Cumulant Method (CM) to produce Probability Density Functions (PDFs) in the limited and the original cases, and then combines these PDFs to generate the final PDF for all system variables. The probabilities that system variables reach their limits are computed and the maximum probability is extracted and used to find the final PDF by aggregating the

PDFs (the original PDFs and the limited ones). A MATLAB program is written and used to verify the proposed approach against Monte-Carlo Simulation (MCS) consisting of 10,000 samples and compared with the original Cumulant Method (CM). The Averaged NSSE (ANSSE) is computed and used to present the improvement. The results of MAA approach demonstrate significant improvements when compared with traditional CM results.

References

- [1] M. Huneault and F. Galiana, "A survey of the optimal power flow literature," IEEE Transactions on Power Systems, vol. 6, May 1991.
- [2] M. T. Schilling, A. L. da Silva, R. Billinton, and M. El-Kady, "Biography on power system probabilistic analysis (1962 - 1998)," IEEE Transactions on Power Systems, vol. 5, no. 1, February 1990.
- [3] G. L. Torres and V. H. Quintana, "Optimal Power Flow by a Nonlinear Complementarity Method," IEEE Transactions on Power Systems, vol. 15, no. 3, pp. 1028-1033, August 2000.
- [4] R. Jabr, A. Coonick, and B. Cory, "A Primal-Dual Interior Point Method for Optimal Power Flow Dispatching," IEEE Transactions on Power Systems, vol. 17, no. 3, pp. 654-662, August 2002.
- [5] V. A. de Sousa, E. Baptistaand, and G. Costa, "Optimal power flow via interior-exterior method," Power Engineering Society General Meeting, 2007. IEEE, June 2007.
- [6] M. Madrigal, K. Ponnambalam, and V. H. Quintana, "Probabilistic optimal power flow," IEEE Canadian Conference on Electrical and Computer Engineering, vol. 1, May 1998.
- [7] T. Yong and L. R.H., "Stochastic optimal power flow: formulation and solution," Power Engineering Society Summer Meeting, 2000. IEEE, vol. 1, pp. 237 - 242, 2000.
- [8] K. L. C. K.A. and D. P.W., "Stochastic opf via bender's method," Power Tech Proceedings, 2001 IEEE Porto, vol. 3, 2001.
- [9] A. Schellenberg, W. Rosehart, and J. Aguado, "Cumulant-based probabilistic optimal power flow (P-OPF) with Gaussian and gamma distributions," IEEE Trans. on Power Systems, vol. 20, no. 2, pp. 773-781, May 2005.
- [10] A. Schellenberg and W. Rosehart and J. Aguado, "Cumulant based probabilistic optimal power flow (P-OPF)," in Proceedings of the 2004 International Conference on Probabilistic Methods Applied to Power Systems, Sept 2004, pp. 506-511.
- [11] G. Verbić and C. A. Canizares, "Probabilistic Optimal Power Flow in Electricity Markets Based on a Two-Point Estimate Method," pp. 1883-1893, Nov 2006.
- [12] G. Verbić, A. Schellenberg, W. Rosehart, and C. A. Canizares, "Probabilistic optimal power flow applications to electricity markets," International Conference on Probabilistic Methods Applied to Power Systems, vol. 1, no. 1, pp. 348-353, June 2006.
- [13] A. Schellenberg, W. Rosehart, and J. Aguado, "Cumulant based stochastic optimal power flow (s-opf) for variance optimization," Power Engineering Society General Meeting, 2005. IEEE, vol. 1, pp. 473 - 478, June 2005.
- [14] A. Papoulis and S. Pillai, Probability, Random Variables, and Stochastic Processes, 4th ed. McGraw Hill, 2002.
- [15] S. M. Ross, Introduction to Probability and Statistics for Engineers and Scientists, 3rd ed. Academic Press, 2004.
- [16] M. Kendall and A. Stuart, The Advanced Theory of Statistics, 2nd ed. Charles Griffin, 1963.
- [17] C. Sufana, G. Heydt, and P. Sauer, "A Linearised Technique and Monte Carlo Simulation for Stochastic Power Flow Studies of Electric Power Systems," in Modeling and Simulation, 1975, pp. 7-11.
- [18] B. Borkowska, "Probabilistic Load Flow," IEEE Transactions on Power Apparatus and Systems, vol. PAS-93, no. 3, pp. 752-759, May-June 1974.
- [19] R. Allan, B. Borkowska, and C. Grigg, "Probabilistic Analysis of Power Flows," Proceedings of the Institution of Electrical Engineers, vol. 121, no. 12, pp. 1551-1556, December 1974.
- [20] K. Timko, A. Bose, and P. Anderson, "Monte Carlo Simulation of Power System Stability," IEEE Transactions on Power Apparatus and Systems, vol. 102, no. 10, pp. 3453-3459, October 1983.
- [21] A. L. da Silva and V. Arienti, "Probabilistic load flow by a multilinear simulation algorithm," IEEE Proceedings C (Generation, Transmission and Distribution), vol. 137, no. 4, pp. 276-282, July 1990.
- [22] R. Billinton and L. Gan, "Use of Monte Carlo Simulation in Teaching Generating Capacity Adequacy Assessment," IEEE Transactions on Power Systems, vol. 6, no. 4, pp. 1571-1577, November 1991.
- [23] G. Viviani and G. Heydt, "Stochastic optimal energy dispatch," IEEE Transactions, vol. PAS-100, pp. 3221-3228, 1981.



- [24] L. Sanabria and T. Dillon, "Stochastic power flow using cumulants and Von Mises functions," International Journal of Electrical Power and Energy Systems, vol. 8, no. 1, pp. 47–60, January 1986.
- [25] P. Zhang and S. T. Lee, "Probabilistic load flow computation using the method of combined cumulants and gram-charlier expansion," IEEE Transactions on Power Systems, vol. 19, no. 1, February 2004.
- [26] F. Tamtum, A. Schellenberg, A. Rosehart, W.D., "Enhancements to the Cumulant Method for Probabilistic Optimal Power Flow Studies" IEEE Transactions on Power Systems, vol. 24, no. 4, pp. 1739–1746, Nov 2009.
- [27] Dadkhah, M., Venkatesh, B. "Cumulant Based Stochastic Reactive Power Planning Method for Distribution Systems With Wind Generators" IEEE Transactions on Power Systems, vol. 27, no. 4, pp. 2351–2359, Nov 2012.
- [28] Ali Abdusalam Tamtum "Cumulant-Based Monte-Carlo Method for Probabilistic Optimal Power Flow studies" 1st International Conference on Electrical and Computer Engineering , 26-28 March, 2013, Benghazi, Libya
- [29] Mr. Ali Abdusalam Tamtum , Dr. Abubkr Ali Elshekhi , Dr. Shukri El Dabar "Probabilistic Optimal Power Flow Studies Considering Network Outages" High Professional Institute For Comprehensive Professions-Alkhums Researches of the 5th Scientific Conference 28/11/2015
- [30] H. P. Hong, "An efficient point estimate method for probabilistic analysis," Reliability Engineering and System Safety, vol. 59, pp. 261–267, 1998.
- [31] Weigao Sun, Mohsen Zamani, Hai-Tao Zhang, Yuanzheng Li Dabar "Probabilistic Optimal Power Flow With Correlated Wind Power Uncertainty via Markov Chain Quasi-Monte-Carlo Sampling" IEEE Transactions on Industrial Informatics, Volume: 15, Issue: 11, Year: 2019
- [32] A. Schellenberg, W. Rosehart, and J. Aguado, "Introduction to cumulantbased probabilistic optimal power flow (P-OPF)," IEEE Trans. on Power Systems, vol. 20, no. 2, pp. 1184–1186, May 2005.
- [33] K. Xie, Y. Song, J. Stonham, E. Yu, and G. Liu, "Decomposition model and interior point methods for optimal spot pricing of electricity in deregulation environments," IEEE Transactions on Power Systems, vol. 15, no. 1, pp. 39–50, February 2000.
- [34] S. Haykin, Adaptive Filter Theory, 4th ed. Prentice-Hall, 2002.
- [35] University of Washington College of Engineering. Power system test case archive. World Wide Web, <http://www.ee.washington.edu/research/pstca/> . Accessed May 1, 2019

Face Recognition with Symmetrical Face Training Samples Based on Histograms of Oriented Gradients

Saad Allagwail^{1*}, Abdulmajed Elbkosh²

¹ saad.elhashmi@gmail.com, ² aalbkosh@elmergib.edu.ly

^{1,2} Department of Electrical and Computer Engineering, Faculty of Engineering, Elmergib University, Libya

ABSTRACT

Face recognition technology is one of the advanced technologies that help to recognize and identify human faces using an image or video clip. Although many face recognition techniques have been proposed in the literature, a robust face recognition system is still a challenging task. It is known that, in general, increasing the number of training images also increases the performance of face recognition systems. In this paper, a new set of training samples is generated from the original samples, using the symmetry property of the face and the recognition performance is improved. The proposed method has three main stages: generating new images, feature extraction and classification. The symmetry property of the face is used to generate new images, the Histograms of Oriented Gradients is used for feature extraction and the Euclidean distance is used for classification. The proposed method is tested and evaluated using AR dataset which is widely used for testing and comparing the accuracy of face recognition systems. The experimental results show that the proposed method has a recognition accuracy rates higher than the traditional methods.

Keywords:

face recognition;
symmetry;
histogram of oriented
gradients

1. Introduction

With the development of the computers and computer-based systems, their application areas for computer vision-based systems have been increased based on demand for the daily life, One of the important demands for the computer-based systems is face detection and recognition for security systems.

Humans are identified by their physiological, behavioral, and biological properties by biometrics. Biometrics are divided into two categories: physiological biometrics, which include the identification of individuals by physiological and biological characteristics such as face, fingerprint, iris, eye, etc. The second category is the behavioral biometrics, which include the identification of individuals through human situations such as handwriting, signature, walking, etc. [1].

Robust and accurate face recognition (FR) is one of the most important problems in computer vision applications. In literature, there are several methods used for FR, including holistic, local, and hybrid methods [2, 3]. However, recent research has revealed that a symmetry-based approach for FR is a useful method to increase the performance of the FR system; thus, it is possible to realize FR using the property of face symmetry [4, 5].

The benefit of using symmetry property in FR is studied by Allagwail et al as presented in [6, 7] however more study is presented in this paper.

There are many methods can be used to extract the features from the face images such as: the Local Binary Pattern (LBP) [8-10], the Gray Level Co-Occurrence Matrix (GLCM) [11], the Gabor Filter [12], and the Histograms of Oriented Gradients (HOG) [13], since these methods perform well for a texture feature extraction that could be used for the FR [14-16].

The performance of FR methods can be tested and examined using some benchmark facial datasets [4, 17, 18] such as: Olivetti Research Laboratory (ORL) [19], Yale [20] and AR datasets [5, 21, 22].

2. Materials and Methods

Although there are many image processing methods, only some of them are able to solve face recognition problem [23]. There are many methods proposed, for instance, Principle Component Analysis (PCA) [24], which is also called as Eigen faces [25], Linear Discriminant Analysis (LDA) [26]. The two most popular methods are eigenfaces [27], and Fisherfaces [28].

In this paper, we have implemented a face recognition system based on the original and symmetrical samples of the face images. In the first stage, more training images, for FR system, are generated using the symmetry property of the face, then in the second stage, the HOG technique is used to extract the features form the training and the testing images, these features are fit to the final stage, the classification stage, which classify the features or images according to their Euclidean distance to the tested images. The images which are used in the experiments are taken from AR dataset.

3. Theory and Calculation

3.1 Generating new images

In order to increase the size of the training data, new training images are generated using the symmetry property of the face, since those images reflect some appearance of the face that is not shown by the original images as illustrated in Figure 1.

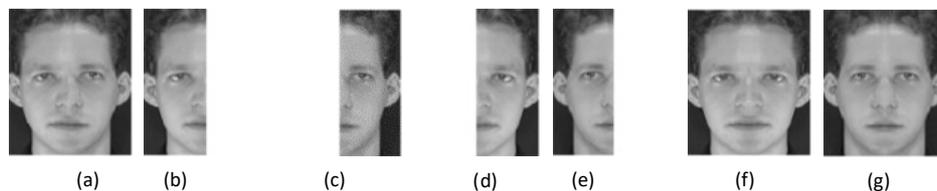


Figure 1: a) Original image, b) left side, c) right side, d) mirror of left side, e) mirror of right side, f) integrating left side with mirror, g) integrating right side with mirror

3.2 Feature extraction using Histograms of Oriented Gradients (HOG)

One of the very popular methods for feature extrication is the Histograms of Oriented Gradients (HOG). It is one type of descriptors that is used a lot in the human detection [13]. The HOG concept is to compute the gradient orientation and the gradient direct magnitude. To obtain the HOG of an image, first, the changes in X and Y are computed, then the magnitude and direction are obtained.

3.2.1 Computing Gradients

The main operation of HOG is the derivative, or the center difference, since, there are two derivatives, the x derivative and the y derivative, once these derivatives are obtained, the gradient magnitude and the gradient orientation can be computed.

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x-h)}{2h} \quad (1)$$

The magnitude is given by:

$$s = \sqrt{s_x^2 + s_y^2} \quad (2)$$

And the orientation is given by:

$$\theta = \arctan\left(\frac{s_y}{s_x}\right) \quad (3)$$

3.2.2 Blocks and Cells

Figure 2 shows a face image, it is assumed that this image is a 64x128 image, if this image is divided to 128 cells, then some blocks are taken, for example the first block is block 1 with 2x2 cells, then the second block is 50 % overlapped, which block 2, so, each block is consist of 2x2 cells with size 8x8 which means 16x16, with 7x15 = 105 blocks in total.

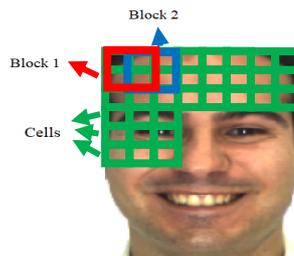


Figure 2: The blocks and the cells

3.2.3 HOG Feature Extraction Steps

To calculate the HOG for an image with 64 x 128, for example, the image is divided onto 16x16 blocks with 50% overlap, so therefore there are 7x15 with total of 105 of blocks, and each block consists of 2x2 cells, and the size is 8x8, then the HOG is quantized with 9 directions or bins, if the direction is not one of the bins then some kind of interpolation can be done, also, the Gaussian can be applied to smooth the histogram, then all the descriptors can be concatenated since there are 105 of these block and each one is 9 dimensional, this gives a very large described, about 3780 dimension descriptor and this for the whole image of the block in the image, as shown in Figure 3.

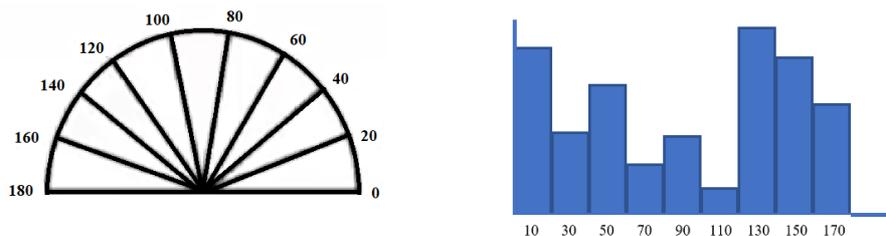


Figure 3: The histogram quantization to 9 bins

The procedure of HOG for feature extraction can be summarized in the following steps:

1. Compute the centered horizontal and vertical gradients with no smoothing.
2. Compute gradient orientation and magnitudes.
 - For color image, pick the color channel with the highest gradient magnitude for each pixel.
3. For 64x128 image,
4. Divide the image into 16x16 of 50% overlap.
 - $7 \times 15 = 105$ blocks in total
5. Each block should consist of 2×2 size 8×8 .
6. Quantize the gradient orientation into 9 bins.
 - The vote is the gradient magnitude.
 - Interpolate votes bi-linearly between neighboring bin center.
 - The vote can also be weighted with Gaussian to down-weight the pixel near the edges of the block.

Concatenate histograms (Feature dimension: $105 \times 4 \times 9 = 3780$)

3.2.4 The Linear Interpolation

The better histograms can be found by doing the interpolation. If there are 9 bins, and the range of the gradient orientation is between $[0^\circ \ 180^\circ]$, this range is quantized into these 9 bins, if the orientation has 85° , and since, there is no bin with 85° , in this case, this is split into couple of bins which are closest to that, these bins are 70° and 90° , since the difference between 70° and 85° is 15° , and the difference between 90° and 85° is 5° , the values is divided proportionally according to this ratio, that means $(5/20 = 1/4)$ and $(15/20 = 3/4)$, and the histogram is distributed according to this concept, as shown in Figure 4.

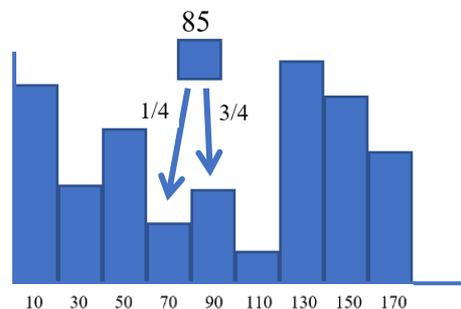


Figure 4: The HOG interpolation

3.2.5 Feature Vector

Each block has its histogram, all the histograms are concatenated to produce the final feature vector of the all image as shown in Figure 5.

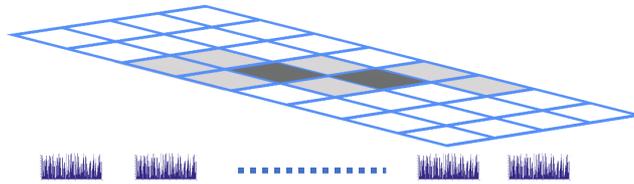


Figure 5: Concatenating the histograms

3.2.6 Visualization

The visualization of the HOG, as in Figure 6, with some blocks and each block has its histogram which corresponds to the face regions, and some blocks give the dominating direction for some certain region, which give the visualization of the representation that represents the face and calculates how much the distance from these different parts of the face.

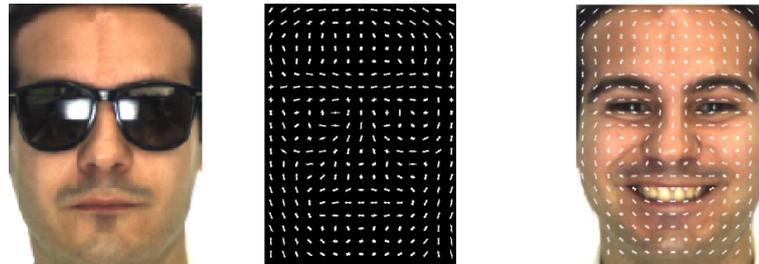


Figure 6: The HOG visualization

By looking to the presentation (middle figure), it is clear it is representing a face, and this is used to recognize the face. Once the descriptor exists, for lots of training examples (faces), then any train techniques such of machine learning can be used for face recognition by classifying the face according to their features.

The HOG is a very strong and popular descriptor, and it is a kind of global descriptor which looks at the whole image. Authors propose this descriptor and they use it for face detection [13] and face recognition [29].

3.3 Classification

The Euclidean Minimum Distance classifier is used for the classification stage since it is considered to be one of the most popular classifiers that can be easily designed and widely used [30]. In general, it is used to examine the similarities between objects.

The Euclidean distance d between two points i and j , where $i = (i_1, i_2, \dots, i_n)$ and $j = (j_1, j_2, \dots, j_n)$, in Cartesian coordinates, is the length of the straightest line between them. This distance is given by the formula:

$$d(i, j) = \sqrt{(i_1 - j_1)^2 + (i_2 - j_2)^2 + \dots + (i_n - j_n)^2} \quad (4)$$

Therefore, if the two points are close to each other, then the value of d is small; otherwise, it is large. The Euclidean vector is the location of a point in a Euclidean n -space, where the length of this vector is measured by the formula of the Euclidean norm, given by:

$$\|I\| = \sqrt{i_1^2 + i_2^2 + \dots + i_n^2} \quad (5)$$

This tool is used to test how similar one object (face) is to another by testing the similarities between their respective feature vectors.

3.4 Dataset

The AR dataset is used in many face recognition papers, it contains over 4000 color face images of 126 people, including 26 frontal views of faces with different facial expressions, illumination conditions and occlusions, like (smile, anger, scream, left light on, right light on, all side lights on, wearing sun glasses, wearing sun glasses and left light on, wearing sun glasses and right light on, wearing scarf, wearing scarf and left light on, wearing scarf and right light on). The images of 120 individuals are captured in two sessions and each session contains 26 color images [21]. Figure 7 shows some images form AR dataset.



Figure 7: AR face database

4. Results and Discussion

These results are carried out using the images from AR dataset, In this part, all AR dataset is used to test the recognition system, the system is trained using 10% of dataset and the rest 90% is used for testing, then 20% for training and 80% for testing, up to 95% for training and 5% for testing. In this experiment, the Histograms of Oriented Gradients HOG is used for feature extraction, this technique extracts a lot of feature from the image and produce a long feature vector that describe the information in the image. The results are shown in Figure 8. The recognition system is examined one time using original training samples (OTS) and the second time using original with symmetrical training samples (OSTS).

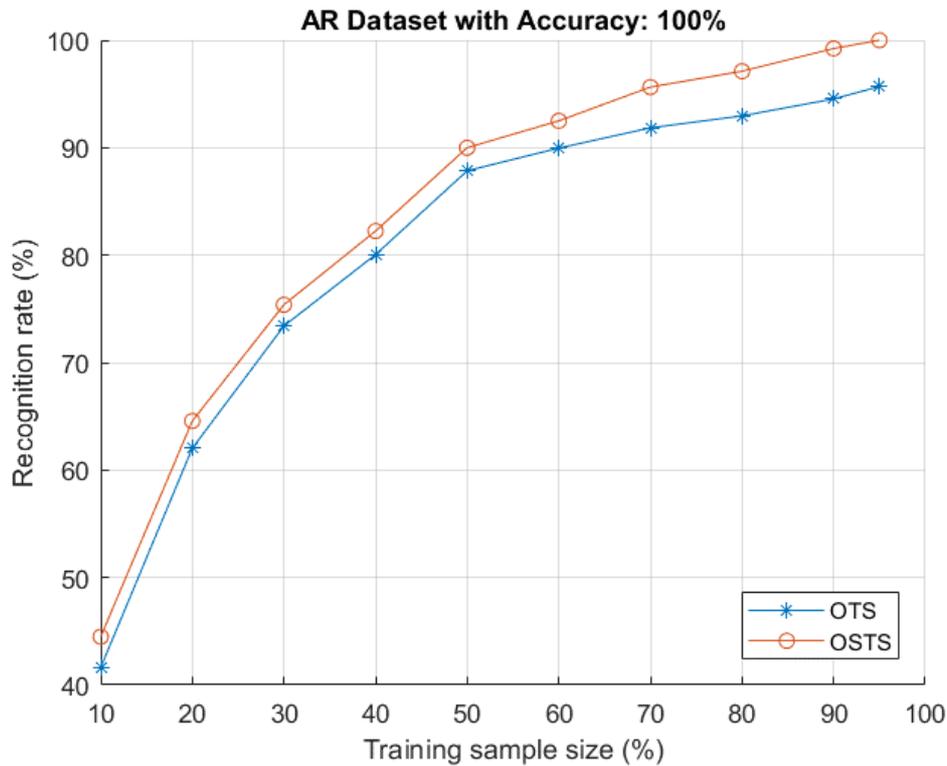


Figure 8: The recognition rate using HOG for OTS and OSTs

The obtained results are summarized in Table 1.

Table 1: The recognition rates on the AR dataset, using the OTS and the OSTs

Type of training samples	Training set %									
	10	20	30	40	50	60	70	80	90	95
	Recognition Rate %									
Original training samples (OTS)	42	63	73	82	89	91	94	96	97	98
Original with Symmetrical Training samples (OSTS)	44	65	75	82	90	93	96	97	99	100

As we can see from Figure 8 and Table 1, as the number of training samples increase, the accuracy increases too. Also, in general, the results using original with symmetrical samples are better than using only the original samples.

5. Conclusions

This paper presents an effective method to overcome the restricted number of the training sets using the symmetry property of the face. First, a new set of face images is generated using the left and right halves of each face. Second, the features of these samples are extracted using HOG method. Finally, the Euclidean classifier is used to obtain the results of the recognition. The experimental results show that the proposed method has a recognition accuracy rates higher than the traditional methods.

References

- [1] H. Vallabh, "Authentication using finger-vein recognition," University of Johannesburg, 2012.
- [2] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with local binary patterns: Application to face recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, pp. 2037-2041, 2006.
- [3] X. Tan, S. Chen, Z.-H. Zhou, and F. Zhang, "Face recognition from a single image per person: A survey," *Pattern Recognition*, vol. 39, pp. 1725-1745, 2006/09/01/ 2006.
- [4] Z. Liu, J. Pu, Q. Wu, and X. Zhao, "Using the original and symmetrical face training samples to perform collaborative representation for face recognition," *Optik-International Journal for Light and Electron Optics*, vol. 127, pp. 1900-1904, 2016.
- [5] Y. Peng, L. Li, S. Liu, T. Lei, and J. Wu, "A New Virtual Samples-Based CRC Method for Face Recognition," *Neural Processing Letters*, vol. 48, pp. 313-327, August 01 2018.
- [6] S. Allagwail, O. S. Gedik, and J. Rahebi, "Face Recognition with Symmetrical Face Training Samples Based on Local Binary Patterns and the Gabor Filter," *Symmetry*, vol. 11, p. 157, 2019.
- [7] S. O. Allagwail and O. S. Gedik, "Face Recognition with Symmetrical Face Training Samples Based on Local Binary Patterns and Gaussian Low-Pass Filter," in *International Conference on Multidisciplinary, Engineering, Science, Education and Technology (IMESET'19)*, Kuala Lumpur, Malaysia, 2019, pp. 22-27.
- [8] L. Liu, P. Fieguth, Y. Guo, X. Wang, and M. Pietikäinen, "Local binary features for texture classification: Taxonomy and experimental study," *Pattern Recognition*, vol. 62, pp. 135-160, 2017.
- [9] B. Julsing, "Face recognition with local binary patterns," *Research No. SAS008-07, University of Twente, Department of Electrical Engineering, Mathematics & Computer Science (EEMCS)*, 2007.
- [10] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns," in *Computer vision-ecv 2004*, 2 ed: Springer, 2004, pp. 469-481.
- [11] P. Yang and G. Yang, "Feature extraction using dual-tree complex wavelet transform and gray level co-occurrence matrix," *Neurocomputing*, vol. 197, pp. 212-220, 7/12/ 2016.
- [12] Y. Sun and J. Yu, "Facial Expression Recognition by Fusing Gabor and Local Binary Pattern Features," in *International Conference on Multimedia Modeling*, 2017, pp. 209-220.
- [13] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2005, pp. 886-893.
- [14] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, pp. 971-987, 2002.
- [15] P. M. Arabi, G. Joshi, and N. V. Deepa, "Performance evaluation of GLCM and pixel intensity matrix for skin texture analysis," *Perspectives in Science*, vol. 8, pp. 203-206, 2016.
- [16] W. Li, K. Mao, H. Zhang, and T. Chai, "Designing compact Gabor filter banks for efficient texture feature extraction," in *2010 11th International Conference on Control Automation Robotics & Vision*, 2010, pp. 1193-1197.
- [17] A. Azeem, M. Sharif, M. Raza, and M. Murtaza, "A survey: Face recognition techniques under partial occlusion," *Int. Arab J. Inf. Technol.*, vol. 11, pp. 1-10, 2014.
- [18] G. Mehta and S. Vatta, "An Introduction to a Face Recognition System using PCA, FLDA and Artificial Neural Networks," *IJARCSSE*, vol. 3, 2013.
- [19] F. S. Samaria and A. C. Harter, "Parameterisation of a stochastic model for human face identification," in *Applications of Computer Vision, 1994., Proceedings of the Second IEEE Workshop on*, 1994, pp. 138-142.
- [20] A. Georghiadis, P. Belhumeur, and D. Kriegman, "Yale face database," *Center for computational Vision and Control at Yale University*, <http://cvc.cs.yale.edu/cvc/projects/yalefaces/yalefaces.html>, vol. 2, p. 6, 1997.
- [21] A. Martinez and R. Benavente, "The AR face database," CVC Tech. Report1998.
- [22] X. Liu, L. Lu, Z. Shen, and K. Lu, "A novel face recognition algorithm via weighted kernel sparse representation," *Future Generation Computer Systems*, vol. 80, pp. 653-663, 2018.
- [23] Ö. Tilki, "PCA based face recognition: An application," 2014.
- [24] K. Kim, "Face recognition using principle component analysis," in *International Conference on Computer Vision and Pattern Recognition*, 1996, pp. 586-591.
- [25] M. Naem, I. Qureshi, and F. Azam, "FACE RECOGNITION TECHNIQUES AND APPROACHES: A SURVEY," *Science International*, vol. 27, 2015.
- [26] S. Meshgini, A. Aghagolzadeh, and H. Seyedarabi, "Face recognition using Gabor-based direct linear discriminant analysis and support vector machine," *Computers & Electrical Engineering*, vol. 39, pp. 727-745, 2013.
- [27] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of cognitive neuroscience*, vol. 3, pp. 71-86, 1991.
- [28] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 19, pp. 711-720, 1997.



-
- [29] O. Déniz, G. Bueno, J. Salido, and F. De la Torre, "Face recognition using histograms of oriented gradients," *Pattern recognition letters*, vol. 32, pp. 1598-1603, 2011.
 - [30] C. Ravat and S. A. Solanki, "Survey on Different Methods to Improve Accuracy of The Facial Expression Recognition Using Artificial Neural Networks," 2018.

Comparative Study of the Electromagnetic Performance of Yokeless and Segmented Armature YASA Machine with Different Rotor Pole Combinations

¹Mohammed F. KHATAB, ²Aiman NOUH

^{1,2}Department of Electrical & Electronic Engineering, Department of Electrical and Electronics Engineering, College of Engineering University of Omar Al-Mukhtar - Al-Bayda, Libya\

¹Mohammed.khtab@omu.edu.ly, ²aiman.nouh@omu.edu.ly

ABSTRACT

In this paper, a conventional axial flux yokeless and segmented armature YASA machine is designed and then simulated. It has been proved that, this machine has a high torque density and has been utilized for wind generation and electric vehicle applications. With 12 stator segments, different rotor pole combinations are designed. To maximize the machine torque, the topology is optimized by means of three-dimensional finite element analysis (3D-FEA). The electromagnetic performance for different pole combinations is then analysed and compared. It has been found that the machine with 16 and 14 rotor poles both have high torques. However, the machine with 14 rotor pole has better torque quality and back EMF compared with the other rotor pole combinations.

Keywords:

PM machines,
axial flux machines,
YASA machine,
electromagnetic
performance.

1- Introduction

With the aim of improving the winding arrangements and eliminating the yoke iron so as to reduce the iron loss in the axial flux TORUS machine, a Yokeless and Segmented Armature (YASA) axial field permanent magnet machine was initially developed and prototyped for automotive applications in [1]. The machine was formed by removing the stator yoke of the TORUS type since the flux of this topology passes axially through the stator poles. Compared to other axial flux machines, the YASA machine has the merits of short end windings, high efficiency, high torque density, high winding packing factor and easily fitted stator poles [2] [3]. However, since the stator is a combination of iron pieces, the mechanical design of the stator pole holder is still vitally important due to the axial force exerted between the stator and the rotors. This is the main drawback of double rotor axial flux machines [4].

Numerous studies have been devoted to the YASA machine performance analysis and manufacturing fabrication for different applications, for example wind generation and electric vehicles. It was first presented and prototyped for a sports vehicle in [1] and [2]. With torque density of 17.6 Nm/Kg, a 10-pole, 12-stator segment prototype was developed, and Soft magnetic composite (SMC) material used for the machine stator segments. For in-wheel direct drive traction applications, 10 poles and 12 stator segments, a 6 kW YASA machine was investigated in [5]. In [3], a 16-pole and a 15-stator segment YASA generator was designed and prototyped. The experimental results show that an efficiency of approximately 87% was obtained at rated speed. In [6], YASA machines were compared to the conventional Radial flux PM (RFPM) and Axial flux PM (AFPM) machines. The comparative investigation found that the YASA machine has a significantly higher torque density as well as low active material weight.

It has been found that internal stator AFPM machine topologies have been extensively developed and investigated for drive applications. Moreover, a YASA machine has superior performance amongst internal stator AFPM topologies. Furthermore, it has been found that based on a performance comparison of AFPM machine topologies, a YASA machine is inherently suited to direct drive application [2] [7]. Therefore, in this paper, since this machine has a unique topology, the machine performance will be investigated in order to provide a baseline of comparison for the subsequent proposed axial field machines. With 12 stator segments, the machine is designed and optimized with different rotor pole combinations. A performance comparison between YASA machine topologies at no-load and on-load conditions will be investigated. Moreover, the torque performance with respect to the current, current density and copper loss will be obtained and compared.

2- YASA Machine Configuration and Principle Of Operation

A YASA machine, as illustrated in Figure , has double rotors with PMs mounted on two opposing rotor discs, and a single stator between both PM rotors. The stator is individually segmented and stacked together with high strength material holder to form magnetically isolated stator poles. Moreover, fractional slot concentrated windings are wound around each of the individual stator poles and connected to form the stator phase windings.

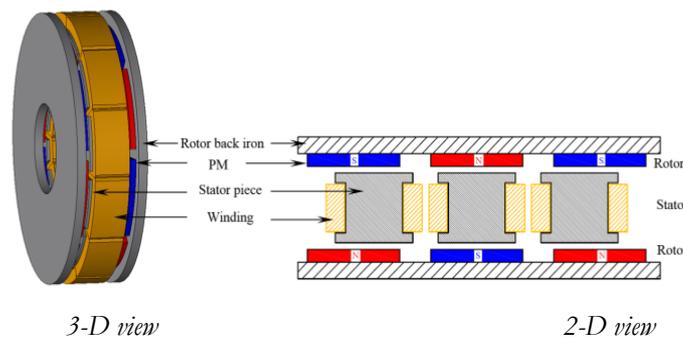


Figure 1: YASA machine configuration.

For a given volume of PM machines and for a specific of phase number, the design is essentially based on obtaining a combination of the stator and rotor pole pairs. The phase number is selected as three in this study, and therefore, for double layer 3-phase windings of YASA machine, the number of stator segments must be a multiple of 3 and the phases are shifted by 120 elec. deg. Moreover, for balanced magnetic pull as well as high winding factor, the PM machine stator and rotor pole numbers should be related by [8] [9]:

$$n_s = 2p \pm 2 \quad (1)$$

A 12 poles with concentrated windings has been chosen for the YASA machine stator. The coils are connected in series to form the stator phase windings. The selection of appropriate combination of slot and pole numbers should ensure the winding distribution balance in the machines. The number of slots per pole per phase (S_{pp}) can be calculated by [10]:

$$S_{pp} = \frac{n_s}{2pm} \quad (2)$$

where m is the number of phases.

Consequently, the most appropriate stator and rotor pole combination can be obtained by (1), when k is even. Table lists the possible rotor pole combinations, their S_{pp} and the fundamental winding factor k_{w1} for n_s of 12 poles. The fundamental winding factor for double-layer concentrated windings is calculated by[11]:

$$K_{w1} = K_{d1} K_{p1} \quad (3)$$

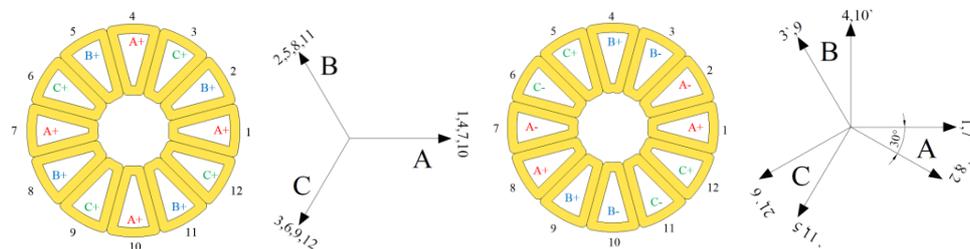
where K_{d1} and K_{p1} are the main harmonic distribution and pitch factors, respectively.

It has been confirmed that high winding factor and balanced concentrated windings are normally adopted when $0.25 \leq S_{pp} \leq 0.5$ [10]. The winding distribution whereby $S_{pp} < 0.25$ has a low fundamental winding factor causing a non-sinusoidal back EMF (i.e. 12/20, 12/22). Moreover, the pole combinations which have $S_{pp} > 0.5$, (i.e. 12/2, 12/4) and where S_{pp} is thus an integer must have distributed winding to obtain a high winding factor. In terms of the rotor pole combinations listed in Table , for YASA machine and due to the focus of this study, this paper only explores YASA with $0.25 \leq S_{pp} \leq 0.5$ in which a high winding factor is gained.

Table 1: Rotor pole combinations for 12-slots, 3-phase concentrated windings

Name	$2p$	S_{pp}	K_{d1}	K_{p1}	K_{w1}
YASA12/22	22	0.18	0.966	0.25	0.25
YASA12/20	20	0.2	1	0.5	0.5
YASA12/16	16	0.25	1	0.866	0.866
YASA12/14	14	0.28	0.966	0.966	0.933
YASA12/10	10	0.4	0.966	0.966	0.933
YASA12/8	8	0.5	1	0.866	0.866
YASA12/4	4	1	1	0.5	0.5
YASA12/2	2	2	0.966	0.25	0.25

For the selected rotor pole combinations, the distributions of the phase coils are determined by the induced back EMF phasor in each coil. The coils distributions for the proposed pole combination are illustrated in Figure 2. Moreover, since the machine has 12 stator poles connected as a three-phase machine, each phase comprises four coils. Furthermore, for the topologies with $0.5 > S_{pp} > 0.25$ (i.e. YASA12/10, YASA12/14), two successive stator segments are wound with coils of the same phase. In this case, a high fundamental winding factor can be obtained. However, the topologies with $S_{pp} = 0.5$ and 0.25 , (i.e. YASA12/8, YASA12/16) have a relatively low fundamental winding factor of 0.866.



(a) YASA12/8 and YASA12/16

(b) YASA12/10 and YASA12/14

Figure 7: Coils arrangements and the EMF phasor for YASA machines with different slot/pole number combinations.

3- Comparison Of Machine Performance For Different Rotor Pole Combinations

With the aim of comparing rotor pole combinations of YASA machine, the optimized machines were designed and analysed with the aid of JMAG 3-D FEA software. The optimal dimensions and the machine parameters are compared and listed in Table 2. The machine topologies were studied at no-load and on-load conditions, details of which can be found in the following sections.

Cogging Torque

Cogging torque exists in PM machines because of the variation of the permeance of air-gap due to stator slots. Therefore, the interaction of the permeance harmonics and the magnet MMF harmonics results in unwanted torque harmonics and thus torque pulsating.

The cogging torque was studied for the selected YASA machine topologies. Figure 3 compares the cogging torques and the corresponding harmonics for YASA machine topologies. YASA12/8 has the highest peak-peak cogging torque while YASA12/10 has the lowest cogging torque. The cogging torque level can be estimated for the slot and pole number combinations of PM machines by the cogging torque factor C_T which can be expressed as indicated in [14] as:

$$C_T = \frac{2p n_s}{N_C} \quad (4)$$

where: N_C is the least common multiple between the number of stator poles and the number of rotor pole.

Higher C_T value indicates higher cogging torque value. However, the minimum value for C_T is unity, which constitutes a proper selection for rotor pole pair combination. For YASA12/8 and YASA12/16, the cogging torque factor is 4. It equals 2 for YASA12/10 and YASA12/14.

Table 2: YASA machine pole combination optimal dimensions and parameters

Parameter	YASA 12/8	YASA 12/10	YASA 12/14	YASA 12/16
Rated speed (RPM)	400	400	400	400
Rotor pole no. ($2p$)	8	10	14	16
Stator slot no. (n_s)	12	12	12	12
Machine inner diameter (mm)	30	30	30	30
Machine outer diameter (mm)	90	90	90	90
Axial length (mm)	25	25	25	25
Air gap length (mm)	0.5	0.5	0.5	0.5
Number of turns of armature coil/phase	80	80	80	80
Packing factor	0.5	0.5	0.5	0.5
Rotor pole pitch (degree)	45	36	25.7	22.5
PM angle (degree)	38.7	36	22.8	18.9
PM thickness (mm)	2.22	2.34	2.7	3
Slot area (mm ²)	48.8	48	52	52
Armature stator axial length (mm)	14.2	13.66	14	14.3

Tip thickness (mm)	1	0.75	0.5	0.5
Rotor axial length (mm)	4.9	5.17	5	4.85
I_{rms} (A)	14.4	14.3	14.8	15
PM volume (mm ³)	13270	16588	16691	17509

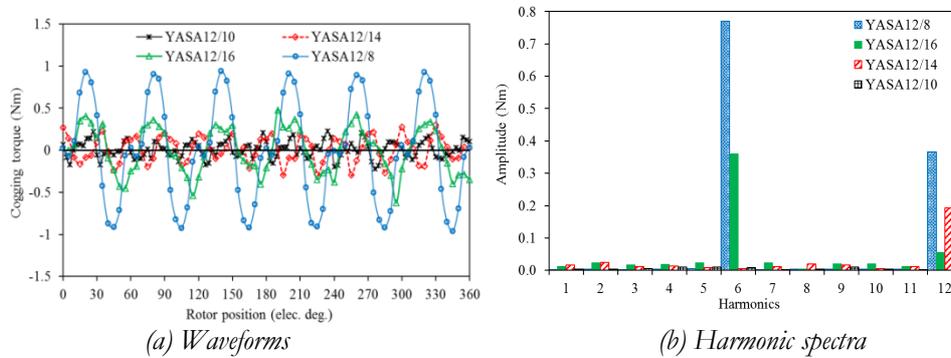


Figure 3: Comparison of cogging torques of YASA machine topologies.

No-Load Flux Density

The no-load flux density produced by the PM is one of the most essential for analyzing the PM machine performance. Figure shows the flux distribution of the stator and rotor of YASA topologies. It can be seen that the topologies in which the rotor pole-arc is bigger than the stator pole-arc have a significant leakage flux since the effective magnet area is smaller than the full rotor pole-arc (i.e. YASA12/8), as indicated in Figure (a). Moreover, the saturation in the back iron is higher in YASA12/8 compared to the other topologies since the flux per pole is higher. Moreover, YASA12/14 has the highest flux density. In this case, the rotor pole-arc is approximately the same as the stator pole-arc without tips. Thus, the leakage flux is minimized, and the flux passed through the stator pole reaches its maximum value, as indicated in Figure (c). Furthermore, YASA12/16 has approximately the same flux density distribution as YASA12/14.

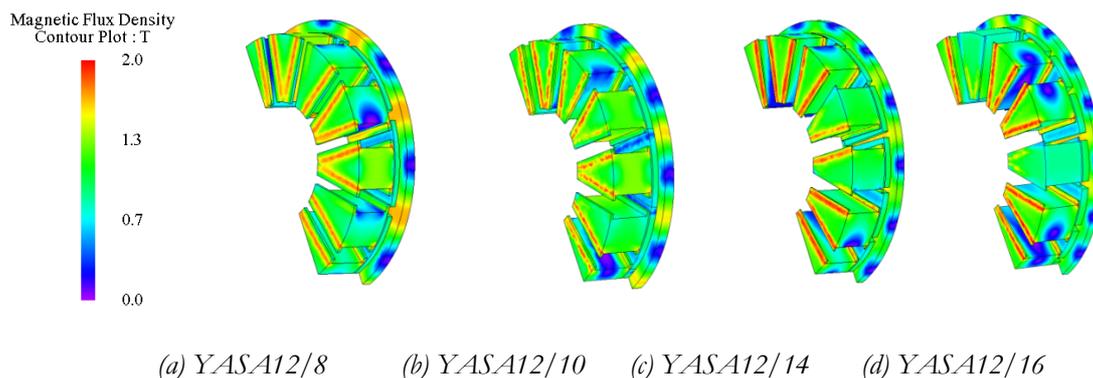


Figure 4: Stator and rotor flux density distributions of YASA machine topologies.

No-Load Flux Linkage and Back EMF

The no-load phase flux linkages of YASA machine topologies are analysed and compared in Figure 5. It is clear that YASA12/8 followed by YASA12/10 have the largest flux linkage amplitude while YASA12/16 has the smallest flux linkage amplitude. However, YASA12/8 flux harmonics include certain odd order harmonics, such as 3rd, 5th, and 7th, which make the flux linkage nonsinusoidal compared to the other topologies. Moreover, the no-load phase back EMFs of the YASA machines are studied and compared at a rotor speed of 400 rpm. Figure 6 shows the comparison of the back

EMF and the corresponding harmonics for YASA topologies. YASA12/14 has the largest back EMF fundamental amplitude of approximately 4.26 V, whereas the EMF fundamental amplitude of YASA12/16 of approximately 4.25 V is just below YASA12/14. On the other hand, the amplitude of the fundamental harmonic of YASA12/10 is approximately 3.9 V, however, YASA12/8 has a non-sinusoidal waveform and contains odd order harmonics in which the amplitude of the fundamental component is approximately 3.3 V, as can be seen in Figure (b).

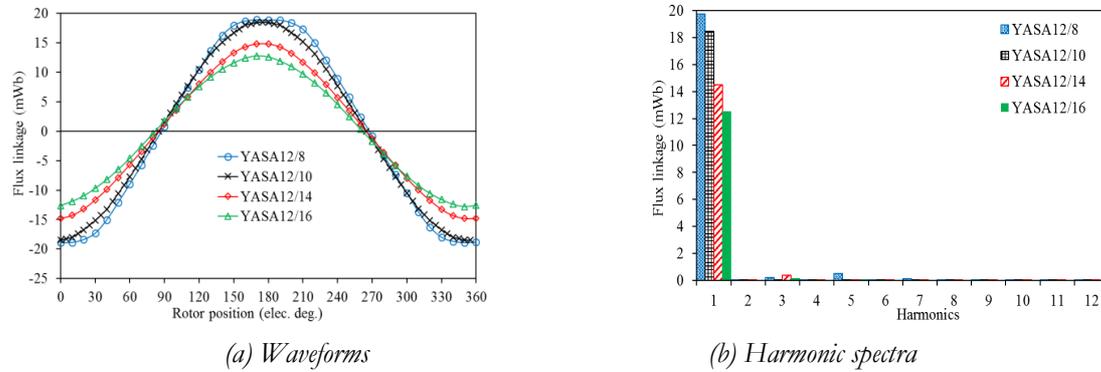


Figure 5: Comparison of phase flux linkages of YASA machine topologies.

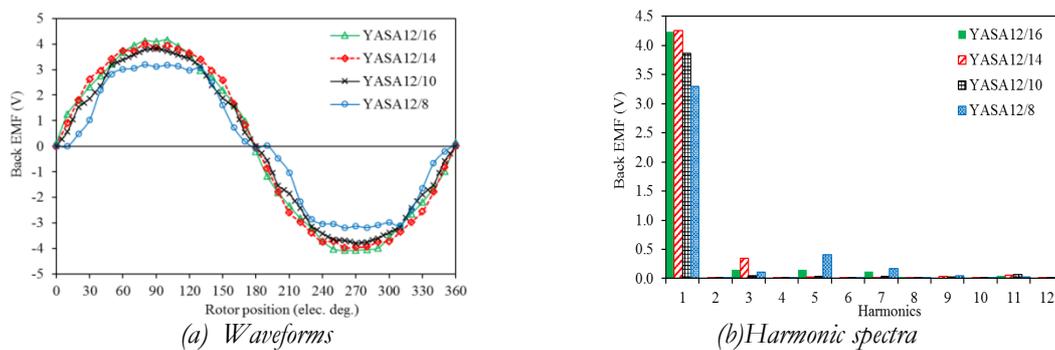


Figure 6: Comparison of phase back EMFs of YASA machine topologies at 400 rpm.

Electromagnetic Torque

The electromagnetic torque is studied and compared at rated current and the same machine speed.

Figure 7 shows a comparison between YASA topologies in terms of the electromagnetic torque and the current angle relation at rated current, and Figure 8 compares the electromagnetic torque for YASA machines with different pole combinations. It is clear that, YASA12/16 and YASA12/14 have nearly the same highest torque while YASA12/8 has the lowest torque in which the leakage flux is high due to the magnet pitch being much bigger than the slot pole pitch. On the other hand, as can be seen in Figure(b), YASA12/14 has the lowest torque ripple followed by YASA12/10. YASA12/8 has the highest torque ripple compared to the other topologies.

Figure 9 shows a comparison between the proposed YASA machine topologies with respect to the torque, torque ripple and cogging torque. The cogging torque and torque ripple are calculated with reference to the average torque by:

$$T_{\text{ripple}}, T_{\text{cog}} = \frac{T_{\text{max}} - T_{\text{min}}}{T_{\text{avg}}} \times 100 \% \quad (5)$$

It should be noted that the preceding formula is used to calculate the cogging torque at no load and the torque ripple at load. The figure shows that YASA12/14 has a superior torque performance compared to the other topologies while YASA12/8 has high cogging torque and torque ripple.

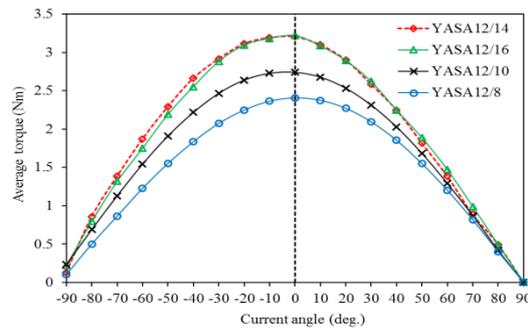


Figure 7: Comparison of torque-current angle curves of YASA machine topologies.

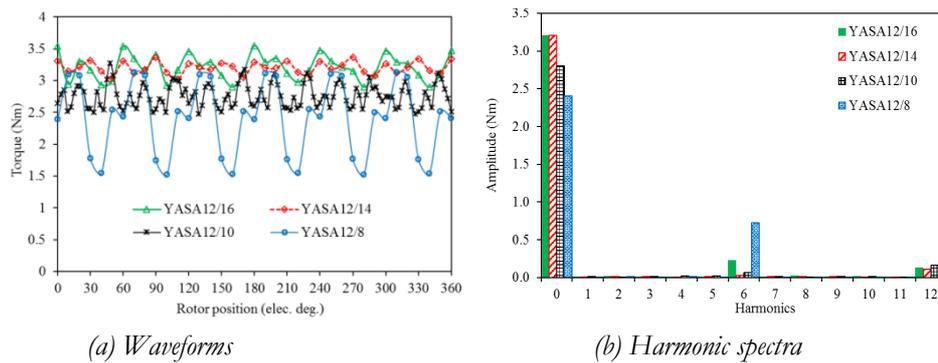


Figure 8: Comparison of electromagnetic torques of YASA machine topologies.

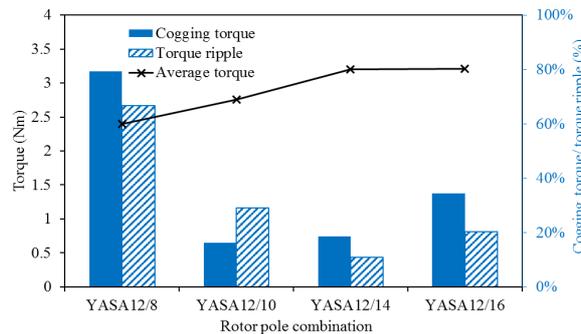


Figure 9: Comparison of average torque, torque ripple and peak-peak cogging torque of YASA machine topologies

4- Conclusion

In this paper, YASA machines with the same number of stator segments and different rotor pole number combinations with a high winding factor were chosen and designed with the aid of 3D-FEA. Moreover, a performance comparison between the YASA machine topologies with different rotor pole numbers has been verified. The results show that both YASA12/14 and YASA12/16 machines have higher no-load and on-load performances. Moreover, such topologies have higher back EMF and average torque; however, YASA12/16 has a noticeably higher cogging torque and torque ripple with the reference of YASA12/14. Furthermore, YASA12/8 has the smallest average torque and highest cogging torque and torque ripple compared to the other topologies. Overall, YASA12/14 and YASA12/16 both have a higher torque density in which both have approximately the same optimal magnet and stator active areas.

References

- [1]. T. J. Woolmer and M. D. McCulloch, "Axial flux permanent magnet machines: A new topology for high performance applications," in *IET - The Institution of Engineering and Technology Hybrid Vehicle Conference 2006*, 12-13 Dec. 2006, pp. 27-42.
- [2]. T. J. Woolmer and M. D. McCulloch, "Analysis of the Yokeless And Segmented Armature Machine," in *2007 IEEE International Electric Machines & Drives Conference*, 3-5 May 2007 2007, vol. 1, pp. 704-708.
- [3]. H. Vansompel, P. Sergeant, and L. Dupré, "Optimized Design Considering the Mass Influence of an Axial Flux Permanent-Magnet Synchronous Generator With Concentrated Pole Windings," *IEEE Transactions on Magnetics*, vol. 46, no. 12, pp. 4101-4107, 2010.
- [4]. F. G. Capponi, G. D. Donato, and F. Caricchi, "Recent Advances in Axial-Flux Permanent-Magnet Machine Technology," *IEEE Transactions on Industry Applications*, vol. 48, no. 6, pp. 2190-2205, 2012.
- [5]. W. Fei, P. C. K. Luk, and K. Jinupun, "A new axial flux permanent magnet Segmented-Armature-Torus machine for in-wheel direct drive applications," in *2008 IEEE Power Electronics Specialists Conference*, 15-19 June 2008 2008, pp. 2197-2202.
- [6]. S. T. Vun, M. D. McCulloch, and C. Y. Leong, "The development of an electromagnetic analytical design tool for megawatt-scale YASA generators," in *IET Conference on Renewable Power Generation (RPG 2011)*, 6-8 Sept. 2011 2011, pp. 1-6.
- [7]. N. J. Stannard, J. G. Washington, and G. J. Atkinson, "A comparison of axial field topologies employing SMC for traction applications," in *2016 19th International Conference on Electrical Machines and Systems (ICEMS)*, 13-16 Nov. 2016 2016, pp. 1-6.
- [8]. D. Ishak, Z. Q. Zhu, and D. Howe, "Permanent-magnet brushless machines with unequal tooth widths and similar slot and pole numbers," *IEEE Transactions on Industry Applications*, vol. 41, no. 2, pp. 584-590, 2005.
- [9]. J. Wang, K. Atallah, Z. Q. Zhu, and D. Howe, "Modular Three-Phase Permanent-Magnet Brushless Machines for In-Wheel Applications," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 5, pp. 2714-2720, 2008.
- [10]. J. Cros and P. Viarouge, "Synthesis of high performance PM motors with concentrated windings," *IEEE Transactions on Energy Conversion*, vol. 17, no. 2, pp. 248-253, 2002.
- [11]. N. Bianchi and M. D. Pre, "Use of the star of slots in designing fractional-slot single-layer synchronous motors," *IEE Proceedings - Electric Power Applications*, vol. 153, no. 3, pp. 459-466, 2006.
- [12]. M. Aydin, Z. Q. Zhu, T. A. Lipo, and D. Howe, "Minimization of Cogging Torque in Axial-Flux Permanent-Magnet Machines: Design Concepts," *IEEE Transactions on Magnetics*, vol. 43, no. 9, pp. 3614-3622, 2007.
- [13]. B. Xia, J. Shen, P. C. Luk, and W. Fei, "Comparative Study of Air-Cored Axial-Flux Permanent-Magnet Machines With Different Stator Winding Configurations," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 2, pp. 846-856, 2015.
- [14]. Z. Q. Zhu and D. Howe, "Influence of design parameters on cogging torque in permanent magnet machines," *IEEE Transactions on Energy Conversion*, vol. 15, no. 4, pp. 407-412, 2000.

Design of a Non-Isolated Solar PV Inverter for Household Applications

Abdallah Suliman Hussein^{1*}

¹ alxash1@gmail.com

¹ Department of Electrical and Electronic Engineering, Faculty of Engineering, Misurata University, Libya

ABSTRACT

Solar energy is taken into consideration as the promising alternative for the future energy demands to reduce carbon dioxide (CO₂) emissions from fossil fuel resources and to slow the depletion of limited energy resources. This paper presents a method on how to design a DC/AC inverter suitable for photovoltaic applications with focuses on minimizing the power losses and the cost.

Keywords:

PV,
H-inverter,
MOSFET,
DC-link and
AC grid.

1- Introduction

The energy demand in Libya is steadily increasing and new types of energy sources must be found in order to cover the existing and future demands and to reduce the energy dependency on oil. Despite the challenges that face the General Electricity Company of Libya (GECOL) to build new fossil fuel plants, renewable energy opportunities are still strategically of high importance and must be adopted by the Libyan government.

One type of renewable energy sources is the photovoltaic (PV) cell, which converts sunlight to electrical current, without any form for mechanical or thermal interlink. PV cells are usually connected to make PV module, which generates a DC voltage and power depending on temperature and solar irradiation. PV modules can therefore not be connected directly to the grid but must be connected through an inverter. The two main tasks for the inverter are to capture the optimal PV module power, in order to harvest the most energy, and to inject a sinusoidal current into the grid.

The initial cost for a PV module is in the moment high compared with other sources. The price for a PV module, including inverter, cables and installation, is approximately 3.2\$ per Watt [1]. This corresponds to a production-price of 0.06\$ per kWh over a time period of 25 years (expected lifetime of PV boards), which cannot yet compete with the electricity supplied by GECOL which has a tariff of 0.02LD per kWh. However, it might be profitable for domestic use, since household and commercial electricity generators as an alternative to the public electricity often of poor quality with harmful impact to environment and increasing the price of diesel in the black market.

Moreover, PV power becomes more competitive by developing inexpensive and reliable inverters. This research has therefore focus on the field of inverter technologies, which is used to interface a PV module to the grid. This paper contains a combined analytical and simulation model of the PV inverter for household applications to determine the operating condition of the H-bridge inverter at no load and maximum power point and the parameters of power semiconductor devices, AC side inductance and DC-link capacitor. In this paper, PSIM software are used in circuit modelling and simulation.

2- Description of the H-bridge Inverter

Among various transformer-less topologies for PV application, the topology chosen due to power factor limitation is a PWM high power factor H-bridge where DC-AC conversion is required. The power circuit is presented in Fig. 1 for grid connected applications; the design values applied in this work are summarized in Table 1. V_{dc} represents solar panels output. Transistors M1-M4 are rated at full DC-link voltage which is in the range of 350-400V for regions such as Libya where RMS grid voltage is 230V. The switching frequency is intentionally chosen at 10kHz to push the first significant harmonic around $2f_{sw}$ seen in the voltage generated by the H-bridge, above audible noise which required to lower the environmental problems in household applications. In this study, the power rating of the converter is 2.5kW, which is a representative figure for residential PV application.

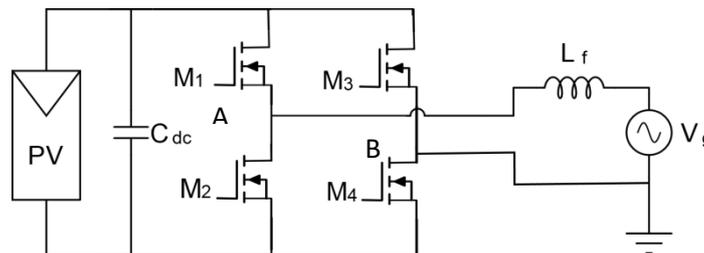


Figure 1: H-Bridge Inverter

Table 1: Converter Parameters

Parameter	Value
P_{max}	2.5kW
V_{dc}	400V
V_g	230V _{rms} /50Hz
f_{sw}	10kHz
PF	1

3- Photovoltaic Module – Inverter Interface

The specifications for the PV module interface correspond to the SPR-210 PV from SUNPOWER [2] are presented in Fig. 2. The voltage across the PV modules when it operates at the maximum power point is 40V at 1000W/m², and a cell temperature of 25°C. In order to avoid overmodulation in PWM scheme and to be able to control the grid current, it is necessary to choose the PV panel voltage at maximum power (smaller DC-link voltage) to be higher than the peak grid voltage plus the voltage drop across the semiconductors and the line side inductance.

The modulation index at full load is chosen to be $m_l = 0.9$ to avoid overmodulation.

$$V_{dc} \geq \sqrt{2} V_g / m_l = \sqrt{2} 230 / 0.9 = 361.4V$$

If we divide this voltage by the max power voltage of a panel (40V), number of 9.04 panels should be connected in series to supply the DC-link voltage. Practically the number of PV panels connected in series are chosen to be 10 PV panels. Therefore, the DC-link voltage at maximum power point is set equal to 400V.

$$V_{dc,MPP} = 10 \times 40 = 400V$$

However, in the worst case at no-load conditions (when maximum sun light falls on the PV panel) the maximum voltage that the DC-link of H-bridge inverter should withstand is:

$$V_{dc,max} = 10 \times 47.8 = 478V$$

Thus, the inverter must withstand at least 478 V without being damaged. This condition of the H-bridge inverter at no load should be considered when selecting the DC-link capacitor rating.

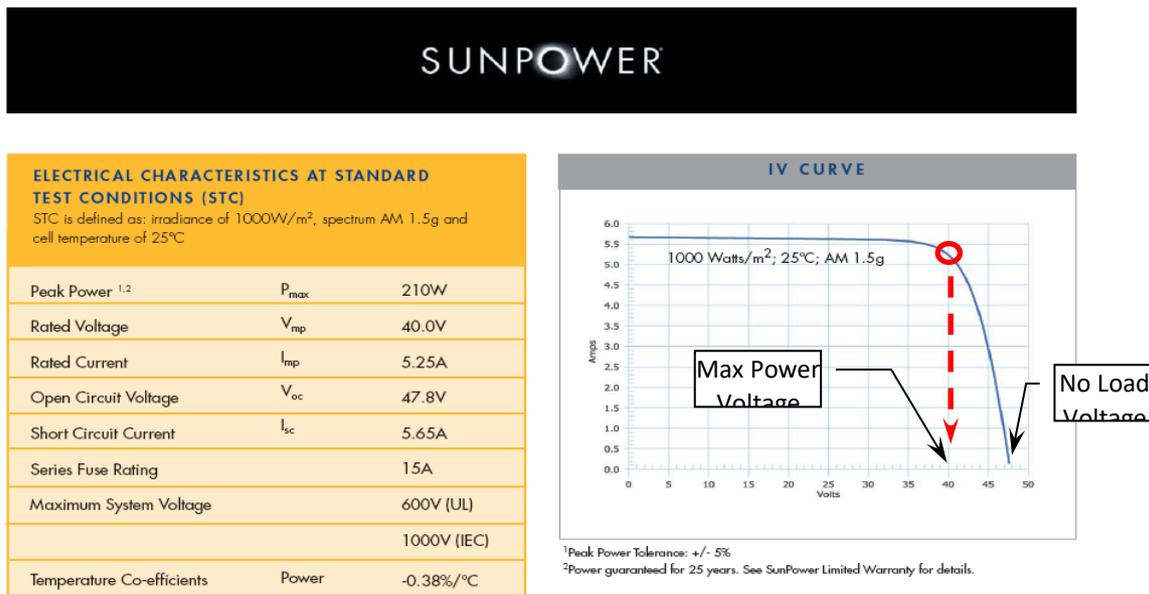


Figure 2: Typical datasheet of a PV panel outlining the no-load and rated voltage/current and the V-I characteristic.

4- Selection of Power Semiconductor Devices

Since there are always safety coefficient that needs to be considered in the power transistor selection: the voltage de-rating $k_v = V_{stress}/V_{rating}$ (typically 0.6-0.7); and the current de-rating $k_I = I_{stress}/I_{rating}$ (typically 0.7-0.8).

The rated voltage of the power transistor devices used in H-bridge inverter with DC-link voltage of 400V is:

$$V_{rating} = V_{stress}/k_v = 400/0.7 = 571V \cong 600V$$

Here, 600V power MOSFETs are used for the two half-bridges. Use of external anti-parallel diodes is optional in the case of power MOSFETs, because of the presence of an intrinsic body diode. Using the body diode will decrease the high cost and the size of H-bridge power cell.

The maximum sinusoidal current that can be processed by these devices assuming a ratio of k_I between the current rating of the switch and the peak value of the RMS current is:

$$I_{rating} = I_{stress}/k_I = 10.9 \times \sqrt{2}/0.8 = 19.2A \cong 20A$$

The selection of the power MOSFET is a trade-off between cost and power losses. The selected MOSFET transistor (IPA60R160P7XKSA1) has specifications of 600V blocking voltage, 20A continuous current, 0.12Ω on-state resistance, and housed in a TO-220 encapsulation. The MOSFET body diode has forward voltage of 0.9V and low reverse recovery charge of 1.7μC [3].

MOSFETs for the DC/AC inverter are selected among the CoolMOS type from Infineon, because they have very good properties in terms of cost versus power losses.

The total conduction losses P_{cond} in the MOSFET at full generation is equal to 7.03W, and the switching losses P_{sw} found equal to 6.14W. The heat-sink thermal resistance is calculated equal to 0.66 °C/W considered the reference heat-sink temperature equal to 60°C for some industrial applications, i.e. PV inverters, and ambient temperature chosen equal to 25°C.

The device junction temperature by considering the device total power loss and ambient temperature is computed to 111.6°C for ambient temperature of 45°C during warm times in summer. This is, however, still seen safe for the power device since the device junction can withstand 150°C without being damaged (from datasheet) [3]

5- Selection of the AC inductance

The output filter of grid-connected inverter is to smooth the output ripple injected to the grid caused by high switching frequency. The filter AC inductance of H-Bridge inverter is determined by selecting the most relevant frequency domain voltage harmonics generated by the H-bridge inverter. As shown in Fig. 3, the FFT of the voltages generated by each of the H-bridge inverter leg with reference to the DC-link mid-point (top two plots) and the result in the FFT of the line-to-line voltage generated by the H-bridge inverter (bottom), revealing the cancelation of the switching frequency harmonic around 10 kHz and the doubling of the fundamental voltage. The result in Fig. 3 shows that the most severe switching harmonic is placed at 20kHz with peak voltage of 100V.

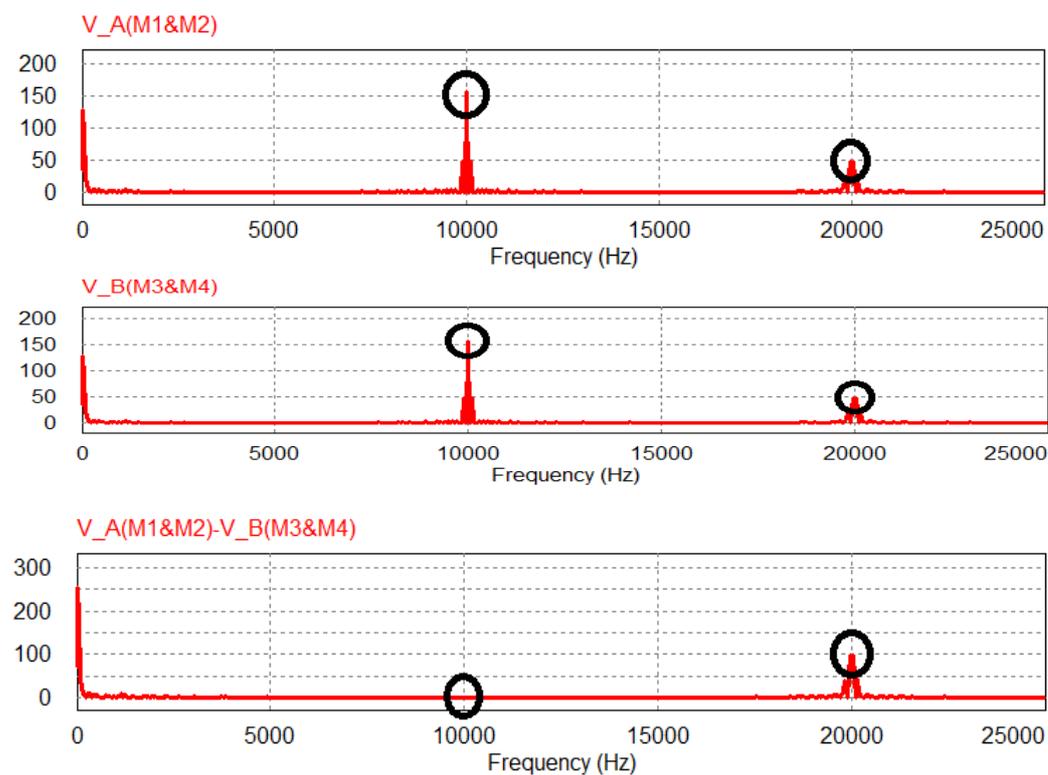


Figure 3: The FFT of the voltages generated by each of the H-bridge inverter legs and the line-to-line voltage (bottom).

According to EN61000-3-2-A standard for current harmonics, any harmonics must be limited below 2 % of the fundamental current (here 218mA) [4].

$$L_{AC,min} = \frac{V_{h-pk}}{2\pi \cdot f_h \cdot \sqrt{2} \cdot \Delta I_h} = \frac{100}{2\pi \cdot 20 \cdot 10^3 \cdot \sqrt{2} \cdot 218 \cdot 10^{-3}} = 2.6mH$$

Magnetic material selection is the first step for inductor design. In comparison, ferrite can achieve low loss at high switching frequency, on the other hand ferrite has low saturation flux density. Although amorphous is characterized by high flux density, but it is prone to high core losses at high switching frequency [5]. Therefore, for this topology and at this power level, ferrite core made from 3C92 material with low losses at high switching frequency, low cost and relatively high saturation (540mT) is the preferred choice in solar inverters [6].

In sizing the inductor, the winding area (A_w) is calculated first based on the current density of 500A/cm², for $I_p=10.9 \times \sqrt{2}=15.4A$, $A_w=0.0308\text{cm}^2$. The core permeance AL (nH/turn²) is determined using $A_L = L/N^2$, and then locating the value on the core selector table; the first

core size located above the calculated value is the smallest core size that can be used in the design. Ferrite core E71/33/32-3C92-G500 with 50 turns was designed with the target to provide inductance of 2.6mH [6].

The core loss in the inductor P_{fe} is estimated equal to 10.2W from the core loss curves normally provided by the manufacturer in the datasheet, while the power loss in the winding is calculated equal to 5.3W and given by:

$$P_{cu} = \rho \frac{N \cdot MLT}{A_w} I^2 = 1.72 \times 10^{-8} \frac{50 \times 160 \times 10^{-3}}{0.0308 \times 10^{-4}} 10.9^2 = 5.3W$$

where MLT is Mean-Length-per-Turn of the core, N is the number of turns of the winding (copper) and I is RMS value of winding current.

The total power loss in the inductor is found:

$$P_{ind} = P_{fe} + P_{cu} = 10.2 + 5.3 = 15.5W$$

6- Determining the DC-link Capacitor

The DC-link capacitor in H-bridge inverter topologies is used to absorb not only the high switching frequency components of the PWM converter current but also for the low frequency current components showing twice mains frequency (i.e. 100Hz for 50Hz mains frequency) due to the pulsating power flow seen in single phase topologies. The DC-link capacitor is distributed into two stages, which are the input bulk capacitors and the stage of high frequency decoupling filters. The input bulk capacitors are based on electrolytic capacitor technology which are good for low frequency ripple. However, high frequency decoupling capacitors based on film technology are responsible for high frequency current ripple.

At full power, the maximum 100Hz current ripple will then be determined [7]:

$$I_{C(rms)} = \frac{I_{dc}}{\sqrt{2}} = \frac{6.25}{\sqrt{2}} = 4.4A@100Hz$$

where I_{dc} is the ripple free DC output current and equal to $I_{dc} = P_{dc}/V_{dc}=2500/400=6.25A$.

The size of DC-link capacitor can be determined by imposing the power into the DC-link and the peak-peak DC-link voltage ripple which is limited to say 2% of the DC-link voltage (or smaller), otherwise an oscillation in the extracted power from the PV panel will occurred due to the voltage ripple in the dc-link.

$$C_{dc,min} = \frac{P_{dc}}{2\omega \cdot V_{dc} \cdot \Delta V_{dc(peak)}} = \frac{2500}{200\pi \cdot 400 \cdot 4} = 2.49mF$$

The selection is made to a standard electrolytic capacitor (ALS70A332MF) of a 3.3mF and ESR=77mOhm and rated at 500V to allow for some over voltages to withstand at no-load voltage condition, without damaging the capacitor.

A final check is performed to ensure that the capacitors can handle the 100Hz current ripple. From the datasheet, this capacitor can handle up to 12.3A@100Hz, which means that the 4.4A@100Hz which is the actual current ripple stresses in the DC-link capacitor can be easily handled.

The DC-link capacitor bank power loss at full power level:

$$P_{C-loss} = ESR \cdot I_{100Hz-RMS}^2 = 0.077 \cdot (4.4)^2 = 1.5W$$

In addition, two film capacitors with 14μF, 630V and ESR=1mOhm each (B32676G6146K) are connected in parallel giving total capacitance of 28μF with negligible power loss.. They are mounted very close to the half bridges of H-bridge inverter to create a low impedance path for the HF ripple current, since the electrolytic capacitors are good only up to some kHz.

Fig.4 shows the waveform of the DC-link voltage. The result is a 396-404 V peak-to-peak voltage ripple of 8V, which is exactly the imposed value (4V peak) in the design stage for the DC-link capacitance.

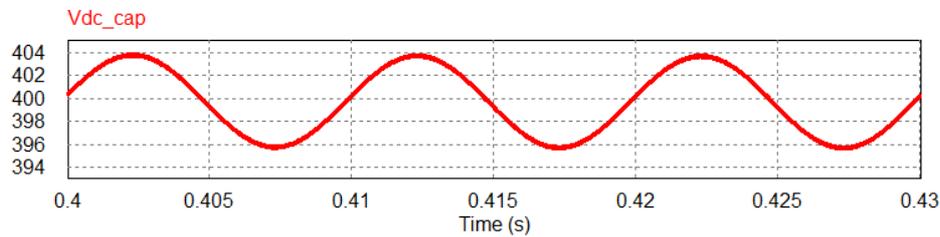


Figure 4: The voltage across the DC-link capacitors.

7- Estimation of the converter losses

The total loss P_{total} of the H-Bridge inverter can be defined as a sum of conduction P_{cond} and switching P_{sw} losses in all MOSFET devices, DC-link capacitor loss P_{C-loss} and inductor loss P_{ind} .

$$P_{total} = 4P_{sw} + 4P_{cond} + P_{C-loss} + P_{ind}$$

$$P_{total} = 4 \times 6.14W + 4 \times 7.03W + 1.5W + 25.7W = 79.9W$$

$$\eta = 1 - \frac{P_{circuit}}{P_{in}} = 1 - \frac{79.9}{2500} = 96.8\%$$

The efficiency of the inverter operating at 10kHz is estimated equal to 96.8% at full power which is satisfactory for this topology and at this power level.

8- Conclusions

In this paper, a combined analytical and simulation model to design a DC-AC inverter for PV applications was developed. Focus has been on determining the operating condition of the H-bridge inverter at no load and maximum power point based on the PV panel characteristics. In addition, the parameters of semiconductor devices, AC side inductance and DC-link capacitor are selected to obtain a low-cost solution with high efficiency.

References

- [1]. "تكلفة الطاقة الشمسية للمنازل"، مركز بحوث ودراسات الطاقة الشمسية <http://csers.ly/ar/news/5-panel-cost>
- [2]. SunPower Corporation, "Sunpower SPR-210-BLK-U (210W) Solar Panel," <http://www.solaradesigntool.com/components/module-panel-solar/Sunpower/1/SPR-210-BLK-U/specification-data-sheet.html>



- [3]. Infineon, “600V CoolMOS P7 Power Device,” https://www.infineon.com/dgdl/Infineon-IPA60R160P7-DataSheet-v02_01-EN.pdf?fileId=5546d462696dbf120169b4543f6d483f
- [4]. “Practical Guide for EN 61000-3-2 - Limits for harmonic current emissions,” <https://www.emcstandards.co.uk/a-practical-guide-for-en-61000-3-2-limits-for-h>
- [5]. W. G. Hurley and W. H. Wölfle, “Transformers and inductors for power electronics: theory, design and applications“. John Wiley & Sons, 2013.
- [6]. FERROXCUBE, “3C92 and 3C93 The high saturation and high temperature low loss power ferrites “, <http://ferroxcube.home.pl/appl/info/3C92&3C93.pdf>
- [7]. J. W. Kolar, and others, “Pwm converter power density barriers,” in 2007 Power Conversion Conference - Nagoya, April 2007, pp. P-9–P-29.

Quick-Response Fuzzy-Controlled Induction Motor Drive

Mohamed Elsharif^{1*}, Tareq Elgargani²

¹ m.elsharif@asmarya.edu.ly, ² tarkargany@gmail.com

¹ Department of Electrical and Computer Engineering, Faculty of Engineering, Al Asmarya University, Libya

² Department of Electrical and Computer Engineering, Faculty of Engineering, Elmergib University, Libya

ABSTRACT

Keywords:

Fuzzy controller,
rule base,
indirect field oriented
control,
induction motor drive,
speed control.

High performance motor drives require high accuracy, fast response, wide range of control, robustness and immunity from the effect of parameter variations. Three phase motors have a complex and highly nonlinear mathematical model associated with interactive parameters. This makes designing a conventional controller for such a system is a hard task. Researchers are paying more attention to fuzzy logic controllers (FLCs) since they can be employed to control complex or nonlinear systems even without knowing their mathematical model. The main task of this paper is to design and implement an FLC for indirect field orientated control of a three phase induction motor drive. The proposed controller is a proportional-derivative (PD) FLC. It uses the speed and its derivative as input and the electromagnetic torque as output. The input and output are coupled with simple linguistic if-then rules. The spread of each input and output is adjusted using a gain block to achieve the best performance in a trial-and-error process. Also, an incremental counter is attached to the output of the controller to yield the desired electromagnetic torque. The design was implemented and tested using MATLAB/SIMULINK. Finally, the simulation results and figures were presented.

1- Introduction

Vector control techniques have been used widely to control three phase induction motor drives. The motor torque and flux are controlled by acting on the space vectors of the stator current [1]. Indirect field oriented control is one of the common used vector control techniques. It has the advantage of using only current and speed sensors. Its principle is based on resembling a separately excited DC motor [2]. Classical proportional-integral (PI) controllers are employed in the control technique [3, 4]. However, designing a PI controller for the induction motor drive is a difficult task since the induction motor model is highly nonlinear, inaccurate and having interactive parameters [5, 6]. Fuzzy logic applications to induction motor include control, performance enhancement, fault detection and motor diagnoses [7, 8]. FLCs may be used in conjunction with other classical controllers to control induction motor drives. The FLC, in this case, is used to adapt classical PI or PID controller parameters according to the variations in the controlled system [9]. Also, it can be used in a standalone form to replace classical controllers [10]. The standalone constant-structure FLC has the advantage for being simple to design, build and tune.

2- Indirect Field Oriented Control

The principle of indirect field oriented control is based on resembling a separately excited DC machine, where the field at some point, is kept constant while the armature current is controlled in order to control the motor speed [2]. Park's transform is widely used in the analysis of induction

motors. It refers stator current and voltage vectors to a fixed frame to the rotor, which in turn has a direct and quadratic (d-q) components [11]. Generally, the generated electromagnetic field of a three-phase induction motor in the d-q fictitious frame is given by:

$$T_{em} = \frac{2}{3} p_p \frac{L_m}{L_r} (i_{sq} \lambda_{rd} - i_{sd} \lambda_{rq}) \quad (1)$$

where p_p is the number of pole pairs; L_m is the mutual inductance; L_r is the rotor inductance; i_{sd} , i_{sq} are the d- and q-axis stator current components; λ_{rd} , λ_{rq} are the d- and q-axis rotor flux linkage components.

In order to make the torque equation similar to that one of the separately excited DC machine, λ_{rq} is set to zero so that $\lambda_{rd} = \lambda_r$. This only can be done by adjusting the angle between the rotor flux linkage and stator current vector to keep λ_r aligned to the d-axis by increasing or decreasing i_{sq} component as shown in Figure 8.

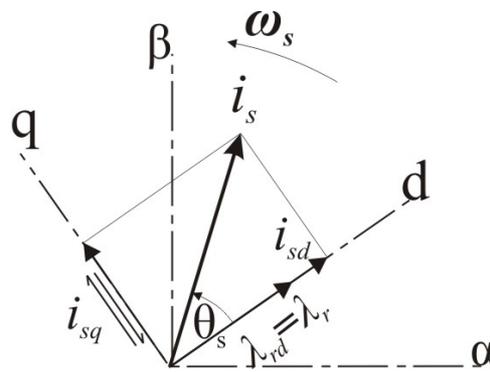


Figure 8: Principle of indirect field oriented control

Both the rotor flux and the synchronous angle (θ_s) are estimated from the measured stator currents and rotor's mechanical speed. The measured mechanical speed and the estimated slip angle are integrated in order to calculate the synchronous angle (θ_s).

$$\theta_s = \int \left(p_p \omega_m + \frac{L_m i_{sq}}{\tau_r \lambda_{rd}} \right) dt \quad (2)$$

where ω_m is the mechanical angular speed and $\tau_r = L_r/R_r$ is the rotor time constant.

Meanwhile the rotor flux component (λ_{rd}) is obtained by solving the following differential equation

$$\frac{\lambda_{rd}}{\tau_r} + \frac{d\lambda_{rd}}{dt} = \frac{L_m}{\tau_r} i_{sd} \quad (3)$$

where $\lambda_{rd} = \lambda_r$ and $\lambda_{rq} = 0$.

The desired (i_{sq}^*) component is calculated from equation (1)

$$i_{sq}^* = \frac{3 L_r T_{em}^*}{2 L_m \lambda_r} \quad (4)$$

where T_{em}^* is desired electromagnetic torque.

Meanwhile, the desired direct-axis stator current (i_{sd}^*) is obtained from equation (3) at steady-state, so that $\frac{d\lambda_{rd}}{dt} = 0$.

$$i_{sd}^* = \frac{\lambda_r}{L_m} \quad (5)$$

The inverse of Park's transform is applied on (i_{sd}^* and i_{sq}^*) components to obtain stator reference current (i_{sabc}^*). The measured stator currents (i_{sabc}) are compared with the reference ones (i_{sabc}^*) in order to generate the required ON/OFF signals for the DC/AC inverter.

3- FLC Structure

The fuzzy logic controller is a microcontroller embedded into the system and designed to work with fuzzy mathematics. The block diagram of a typical FLC is shown in Figure 9. The expert or designer shall have the access to set many parameters including number of inputs and outputs, type and number membership functions, spread, fuzzy rules, aggregation, implication and defuzzification methods.

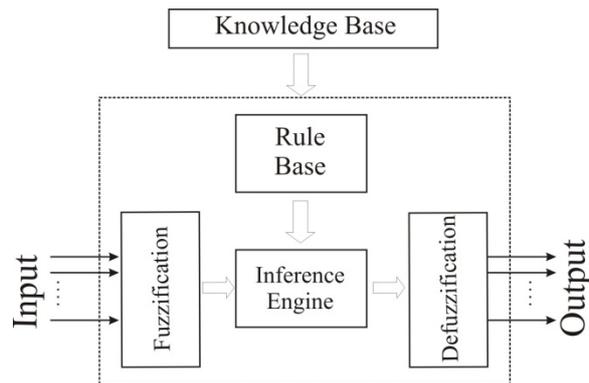


Figure 9: FLC block diagram

In general, an FLC consists of the following components:

The Fuzzification Layer

This layer transforms the inputs, which are in single values (crisp) form, to the correspondent defined fuzzy membership functions. The fuzzy controller may have more than one input.

Inference Engine

The inference engine performs with the fuzzy calculations. First, it implicates the fuzzy inputs based on the provided rules. Then, it collects (aggregates) the implicated membership functions of the output.

Rule Base

The rule base is a set of rules in the form of if-then format that links the membership functions of input and output. The number of rules must cover all the defined membership functions of inputs and outputs. Setting the rules use the human knowledge and expertise in easy linguistic terms.

The Defuzzification Layer

The aggregated fuzzy output from the inference engine is unusable unless it is turned back to a single value (crisp) form. This process is called defuzzification. The aggregated fuzzy shape is weighed and normalized to obtain one crisp value at a time for each output.

4- Simulation

The design and simulation work of the induction motor drive were done by using MATLAB/SIMULNK. The general block diagram of the proposed system is presented in Figure 10. The motor was fed by a DC power source via an inverter while the speed and stator current were measured in order to be used in the control process. The design of the system went through

two stages.

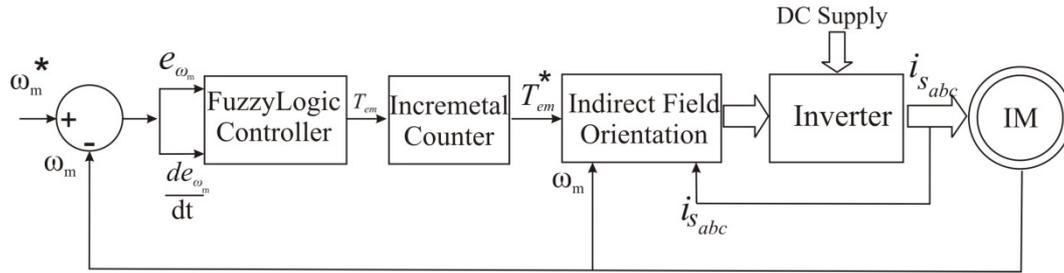


Figure 10: Block diagram of the proposed induction motor drive

Indirect Field Orientation Unit

The modelling equations in section 2 were used to build the indirect field oriented control block as shown in Figure 11. Equation 2 was used to estimate the current angle using the mechanical speed, the estimated rotor flux and the direct component of stator currents. The rotor flux vector (λ_r) was estimated by solving equation 3 while the desired quadratic current component was calculated using the incremental counter output (T_{em}^*) and rotor flux. Park's transform was used to resolve the measured currents into the d-q form, and its inverse for getting the desired stator currents (i_{sabc}^*) from their d-q counterparts. The measured and the desired currents were compared in a hysteresis band controller in order to produce the necessary ON/OFF signals for power switching device.

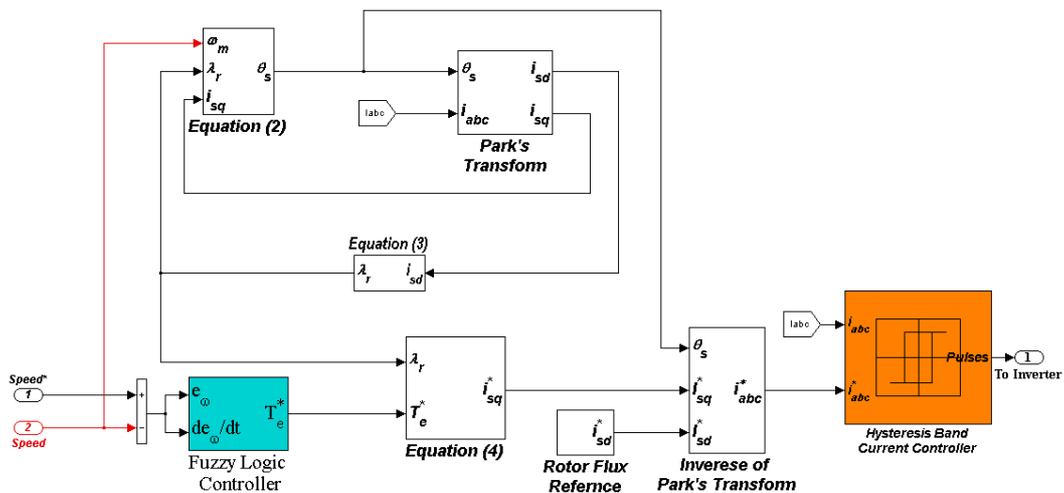


Figure 11: Control unit

A medium size of a 30hp three phase squirrel-cage induction motor was used in simulation with its rated parameters defined in Table 4 [12].

Table 4: Induction motor parameters

Parameter	Symbol	Value
Nominal supply voltage	v_s	400V
Nominal stator current	i_s	39.1A
Nominal frequency	f	60Hz
Pole pairs	p_p	3
Nominal rotor speed	n_m	1146rpm
Stator resistance	R_s	0.294Ω
Stator leakage inductance	L_{ls}	1.4mH
Rotor resistance	R_r	0.156Ω
Rotor leakage inductance	L_{lr}	0.74mH
Mutual inductance	L_m	41mH

Fuzzy Control Unit

The proposed PDFLC was designed using MATLAB fuzzy toolbox. Both speed error and its rate of change were used as input and the electromagnetic torque as an output where five triangular membership functions were assigned to each one of them as shown in Figure 12.

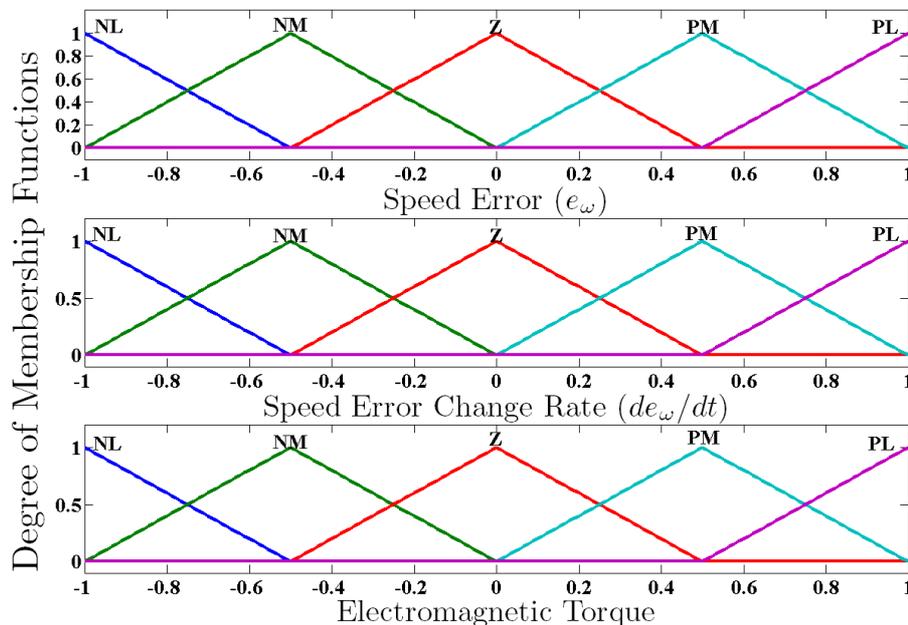


Figure 12: Controller membership functions

The membership functions were linearly distributed and their spreads were adjusted using gain blocks on input and output. The rule base was established based on the expertise knowledge in 25 rules as Table 5 shows.

Table 5: PD Fuzzy logic controller rules

		e_{ω_m}					
		NL	NM	Z	PM	PL	
de_{ω_m}/dt	NL	NL	NL	NM	NM	Z	
	NM	NL	NM	NM	Z	PM	
	Z	NM	NM	Z	PM	PM	
	PM	NM	Z	PM	PM	PL	
	PL	Z	PM	PM	PL	PL	

where **NL** is negative large, **NM** is negative medium, **Z** is zero, **PM** is positive medium, and **PL** positive large.

Also, an incremental counter with delayed positive unity feedback was used on the controller output to generate the desired electromagnetic torque (T_{em}^*) as shown in Figure 13. The counter holds its output constant until the next cycle occurs. The new output of the counter is combined of the summation of the delayed T_{em}^* and the real-time FLC output (T_{em}). The controller output has the effect of gradual increasing or decreasing T_{em}^* . The incremental counter has two functions. First, it stabilizes the drive operation even when the output of the controller is zero. Second, it has a limiter to keep the motor current within the safe limits.

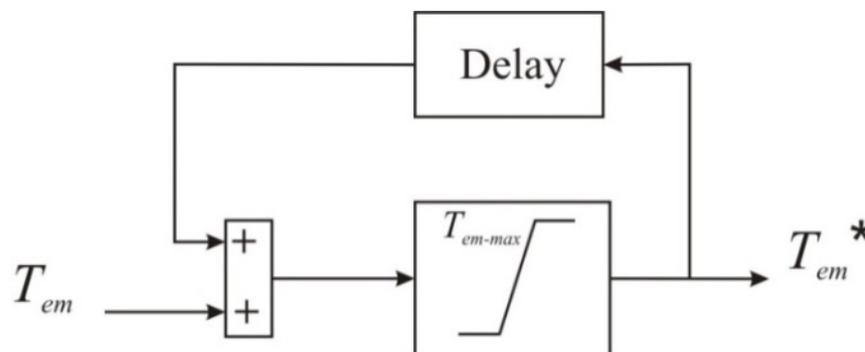


Figure 13: An incremental counter with delayed feedback

5- Results

After the controller gains were tuned to obtain the best results, all simulation data were captured and kept in order to be analysed. The drive was first started to run at 764 rpm and then a speed step change was applied at 0.4 sec to reach 1146 rpm without applying load. At 0.8 sec the load is applied as shown in Figure 14.

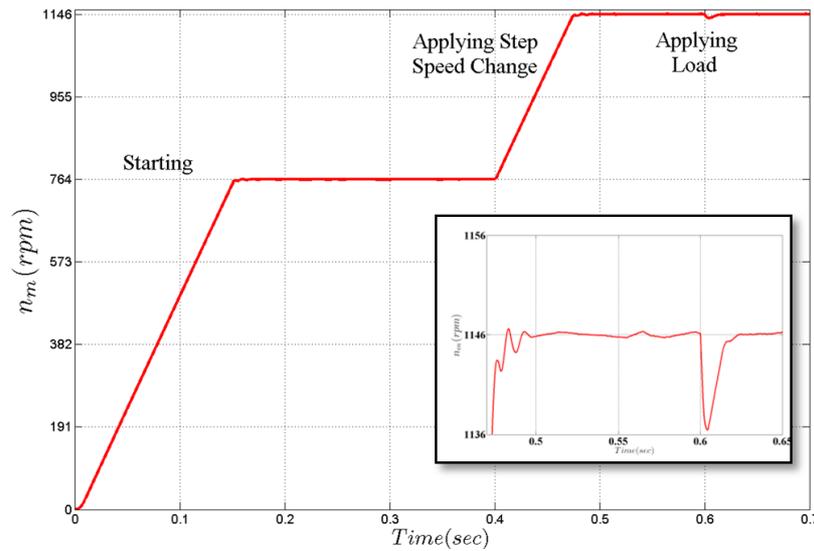


Figure 14: Induction motor drive speed response curve

The peak overshoot for the applied step speed change was almost neglected. However, speed dropped momentarily by 10rpm when load was applied which represents only 0.87 % of the motor speed. The speed response was fast and constant until it became very close to the setting speed.

Finally, the stator motor currents were plotted Figure 15. The plot clearly shows the change of current angles as a variation of speed or torque were applied. This change in the angle was done to keep the rotor flux aligned to direct axis when drifting occurred.

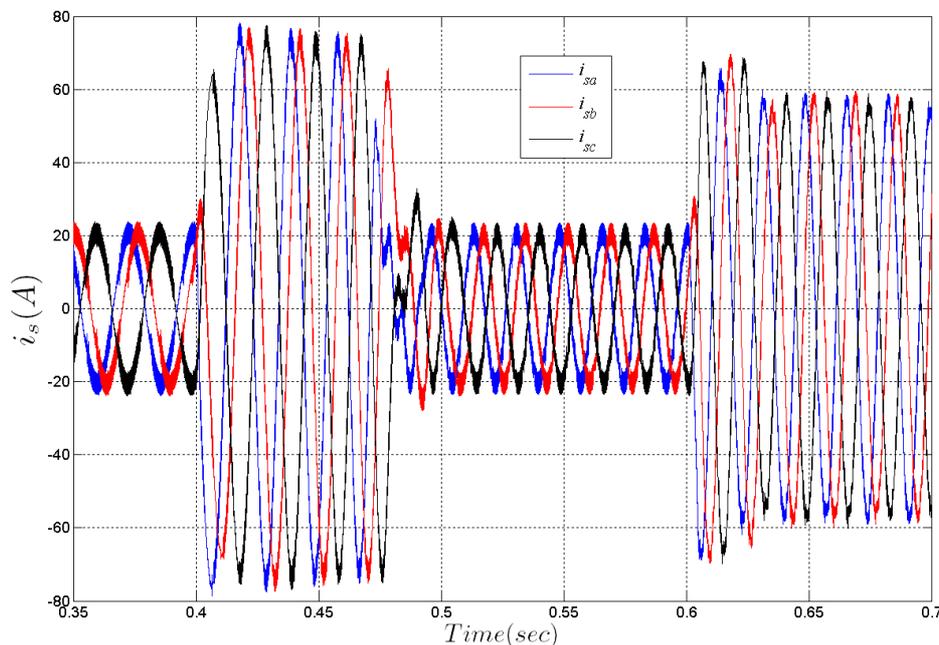


Figure 15: Motor currents

6- Conclusions

This paper investigated the design and implementation of a PD fuzzy logic controller for indirect field orientated control of a three-phase induction motor drive. The speed error and its derivative were used as controller input while the electromagnetic torque as controller output. Each input and output had five triangular membership functions. The cases of input and output were joined together with 25 if-then fuzzy rules. Only the spread of fuzzy sets of each individual input and

output was adjusted to obtain the best performance. In addition, an incremental counter attached to the controller output was used to stabilize the system. Measured stator currents and rotor speed were used to decouple the rotor flux and the current angle in order to be used in conjunction with the incremental counter output to generate the required stator currents for field orientation. The proposed PD fuzzy logic controller achieved fast response, small peak overshoot, wide range of speed control and fast recovery time when variations in speed and load were applied.

References

- [1]. G. S. Buja and M. P. Kazmierkowski, "Direct torque control of PWM inverter-fed AC motors - a survey," *IEEE Trans. Ind. Electron.*, vol. 51, no. 4, pp. 744-757, Aug. 2004.
- [2]. F. Blaschke, "The principle of field orientation as applied to the new transvector closed-loop control system for rotating-field machines," *Siemens Rev.*, vol. 39, no. 5, pp. 217-220, 1972.
- [3]. Ahallianath K G R and W. Beevi M, "Indirect field oriented control of induction motor using predictive current controller," 2015 Int. Conf. on Control Commun. & Computing., India, pp. 248-253, ICCCC. 2015.
- [4]. B. K. Nishad and R. Sharma, "Induction motor control using modified indirect field oriented control," 8th IEEE India Int. Conf. on Power Electron., JAIPUR, India, pp. 1-5, IICPE. 2018.
- [5]. S. Pati, M. Patnaik, and A. Panda, "Comparative performance analysis of fuzzy PI, PD and PID controllers used in a scalar controlled induction motor drive," *Proc. Int. Conf. Circuits, Power Comput. Technol.*, pp. 910-915, ICCPCT. Mar. 2014
- [6]. L. Murari et al., "A proposal of project of PI controller gains used on the control of doubly-fed induction generators," *IEEE Latin Amer. Trans.*, vol. 15, no. 2, pp. 173-180, Feb. 2017.
- [7]. K. Hsu et al., "Enhanced fuzzy sliding mode control to motion controller of linear induction motor drives," *2014 IEEE Int. Conf. on Syst. Sci. and Eng.*, Shanghai, pp. 268-273, ICSSE. 2014.
- [8]. D. A. Surwase, A. S. Jalit and M. D. Chavan, "Fault detection and protection of induction motor using fuzzy logic," *2017 Int. Conf. on Innovations in Elect., Electron., Instrum. and Media Technol.*, Coimbatore, India, pp. 66-70, ICEEIMT. 2017.
- [9]. Kumar R and F. Daya J L, "A novel self-tuning fuzzy based PID controller for speed control of induction motor drive," 2013 In.I Conf. Control Commun. and Comput., India, pp. 62-67, ICCCC. 2013 .
- [10]. M. Pandey and A. Brave, "An intelligent fuzzy logic controlled based induction motor drive," *Int. J. Sci. Res.*, vol. 4, no. 6, pp. 1092-1096, Jun. 2015.
- [11]. R.H. Park, "Two-reaction theory of synchronous machines generalized method of analysis-part I," *Trans. AIEE*, vol. 48, no. 3, pp. 716-730, 1929.
- [12]. Trzynadlowski, *Control of Induction Motors*. San Diego, CA: Academic Press, 2001.

Electricity Load Profiling for Coastal Housing Complexes based on the Measurements of Flats Actual Load

Salem A. S. Ahbil^{1*}, Hamid H. Sherwali²

¹ salemhabel@yahoo.com, ² hsherwa@hotmail.com

¹ General Electric Company of Libya (GECOL)

² Department of Electrical and Electronic Engineering, University of Tripoli, Libya

ABSTRACT

In this paper, a methodology for estimating end-use load shapes using the hourly whole-house metered load data, the household demographic survey data and the weather data (temperature) is presented. End use load shapes presents a method of generating realistic electricity load profile data for some of city of Tripoli domestic buildings. This method could help in predicting the daily load profile from individual flats to community. The results obtained show that the overall methodology provides an effective means for end-use load shape modeling and estimation.

Keywords:

Electricity consumption;
Load profiles;
Domestic buildings;
Appliances;
Occupancy;
Dwelling characteristics;
Metering Equipment.

1- Introduction

Distribution systems obviously exist to supply electricity to the end users, so loads and their characteristics are important. Load data is crucial for planning electricity distribution networks and optimal production capacity [1].

The load types can be divided into five categories: residential, agricultural, industrial, commercial, and public [1]. The identification of the pattern of energy uses for a house and the prediction of the domestic load profile are essential in order to match the load shape to the power generated. The domestic energy-consumption can be divided into three categories: cooling and space heating, domestic hot water, lighting and appliances [2]. The used pattern varies depending on many determinants, such as weather, household composition, income, behaviour patterns of occupants... etc. Occupants influence the use of electricity of the houses by the number of electrical appliances they own and throughout their use of the appliances [2]. The number of occupants and their age influences energy consumption, for example, households where there are no children or where couples work consumes less energy than a household with children or older people [3].

The estimated population of Libya according to the Bureau of Statistics and Census of Libya is approximately 6.2 million [4]. The city of Tripoli, the capital, is the largest city in Libya with a population of just over a million. In addition, it is the country's principal sea port and the commercial and manufacturing centre. The energy consumption per person in this city has increased by 2.6% between 2008 and 2009 [6]. The demand on the energy is increasing every year and there is a shortage in the studies that addressing the load profiles and the energy consumption of the different load types mentioned before.

In November 2009, the Cadmus Group (Cadmus), Navigant Consulting (Navigant), and Global Energy and Technology Consulting (GETCON) has been contracted by the General Electricity Company of Libya (GECOL) to provide an assessment of the demand side management (DSM) potential in Libya. These services were provided throughout 2010 and include winter load (end-use metering), estimating DSM potential, and designing appropriate pilot programs. End-use and load metering were conducted between December 2009 and March 2010 for the following six sectors: residential (housing), street lighting, commercial, governmental, agricultural and light industrial [6]. They provided their data without explaining how they conducted their measurements and obtained their results. This paper is a complementary study to the study conducted in [6] to generating a realistic electricity load profile data for the pure flats complex residential. Direct measurements were taken over a month period for flats in housing complex located in a district named (Zawiyat al-Dahmani) 600m far from seashore near to what is well known as eyes hospital. A sample of 18 flats in bloke number 2 was chosen to be subjected to the load measurements. The total flats in the housing complex were 108 divided between 6 blocks. Measurements were taken at each flat and at the main feeder supply power to the building using reliable measuring equipment every 5 minutes for all the study period. A supporting detailed survey using a questionnaire of the householders sought to obtain an understanding of the appliances used and energy-related behaviour.

2- Methodology

In order to come-up with close-to real end-uses load curve for housing complex, the research consider the following steps:

- 1- Load measurement inside selected flats in the housing complex.
- 2- Supporting detailed survey using a questionnaire of householders sought to obtain an understanding of the occupant behaviour toward their use of electricity and generally their habits toward the use of energy.
- 3- Mathematical and Statistical Analysis for data and measurement analysis.

2.1. Electricity Measurement

Electricity measurements were made using a five-minute load meter installed in series using three different types of meters. The measurements of individual consumers load curves were performed in periods of approximately 20 days (February-March). The meter is the conventional digital type, where each meter rotation produces an electric pulse. These pulses are counted by the electronic equipment that accumulates them in programmed intervals by the user (1, 5 or 15 min). Thus the power (demand) at the interval is determined. In order to achieve a good accuracy, curves were defined at intervals of 5 minutes i.e. 288 points on a daily curve were measured. A meter was located at the feeder of each of the selected flats in the building. The type of the measuring equipment used was a Multichannel meter. Measurements were also carried out close at the main feeder of the building and at the main feeder of the housing complexes. Based on a collection of measurements, the average electricity consumption was calculated by averaging consumption for each day of the month.

2.2. The Region under Study

Zawiyat al-Dahmani included village, homes were detached dwellings, bungalows, and terraced houses; the remaining 40% of homes flats in the blokes. 20% of the homes were less than 15 years old, 30% between 15 and 30 years old, 30% between 30 and 60 years old and 20% of houses were over 60 years old, the weather in this suburb is Mediterranean sea climate. The choice of this district because it contains many of the communities represented in vertical format identical buildings (the 108 flats form the housing complex of 6 blocks) in terms of space and the number of flats. Each block has nine floors; the floor has two flats with area of 180 m² each. In each flat there are three

bedrooms, living room, dining room, two toilets and kitchen, and each block contain two lifts and two three phase pumps. This pattern of buildings can be seen throughout the city of Tripoli and even the entire country.

3- Questionnaire Output

A questionnaire was distributed to consumers of block number 2, nine consumers back it. The 18 flats selected for load measurements were less than 7 years old. 19% of flats had four occupants, 22% had two occupants, 23% had five occupants, 12% had six occupants and 24% had eight occupants. The income of the occupant of these flats, as reflected from the questionnaire, classified as an average. 75% of the occupants between 18 to 45 years old and the remaining are between 1 to 17 years old.

In all the surveyed flats, heating was provided by means of Electrical heaters. The other installations in the surveyed flats which consume electricity are divided broadly into three categories: lighting, kitchen and entertainment. The lighting installations are as follows: 90% of flats have halogen bulbs, 10% fluorescent tubes. The kitchens are as follows: most flats use microwave ovens. Each household has at least one fridge or fridge-freezer and 10% of the households have also separated freezers.

Eighty percent of the flats under test have different air condition, 80% have washing machine, and 90% have Irons. In terms of entertainment, each flat may have on the average 2.7 TV sets, 1.5 video recorders, 1.4 DVD players, 1 stereo system, 0.8 games consoles and 1 computer.

4- Household Types

There are two types of flats in the housing complex under study:

Type1: Single adult household, the occupier is a full time employee and the unoccupied period is considered from 9:00am to 6:00pm. The number of the flats represents a 54% of total flats covered by load measurements.

Type 2: Two adults with children, one occupant have a full time job and the other adult may have a part time job in the morning. The unoccupied period is considered from 9:00am to 1:00pm. The number of the flats represents a 46% of total flats covered by load measurements.

5- Daily Load Curves

Daily load curves are obtained from the measured values of consumed energy for each flat during each day of the measurement period for all types of flats, where measuring equipment were installed. Mathematical and statistical analysis has been done using Excel and Minitab programs in order to generate domestic load profile for different consumption types, which can be found in details in [7] (as this paper cannot fit for the equations and the tables of measured data). Curves were classified for two measure types according to flats occupancy; i.e. working days and off days, and occupant style; i.e. single or family.

5.1 Type1: Single Adult Household

There are 7 flats for a single adult household (1, 4, 5, 10, 11, 17, and 18).

5.1.1 Average Load Curve for Week Days

Figure (1) shows the average daily load curve for one flat of the type one (single adult) on week days.

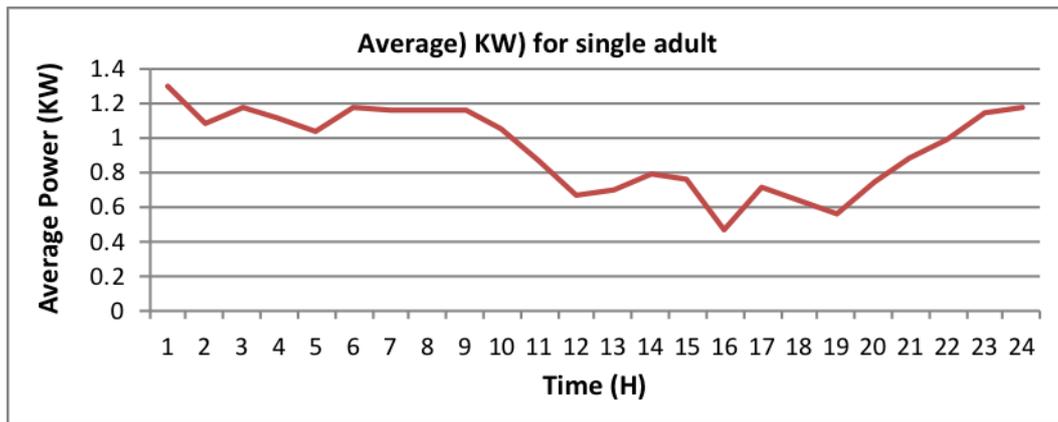


Figure (1): The average daily load curve for single adult flats during week days.

It can be noticed that a gradual increase in the electrical power consumption starting from 05:00 a.m. This is the wake up of some consumers at the dawn time where they began their daily activity and use of different electrical equipment for the preparation of their breakfast and taking shower. Such activities continue disproportionately among the flats which their occupants are ready for going out to work. The load consumption is then declining gradually, as a result of being out of the flats at different times. This slop continues until approximately mid-day where activities start to pick up, when some starts coming back from work. Such coming back reflects the demanding on power needs for preparing lunch, also using heaters due to the drop of temperature, this continues until 14.00 afternoons. One of Libyan habits is to relax after 14:00; as a result, power demand declines during that period, i.e. (14:00 till 16:00). However, as activity is resumed at 16:00, load starts to go up and down for a small period depending upon social activities as some visits and so on. Some of the occupants stay at their flats, but others leave. Power demand increases gradually around 18:30, as shown in the curve. It can be noticed thereafter through the curve starts gradually declining due to the gradual decrease of power demand, as most of the occupants reach sleeping time.

5.1.2 Average Load Curve for Week Days

Figure (2) shows the average daily load curve for one flat with a single adult on weekends.

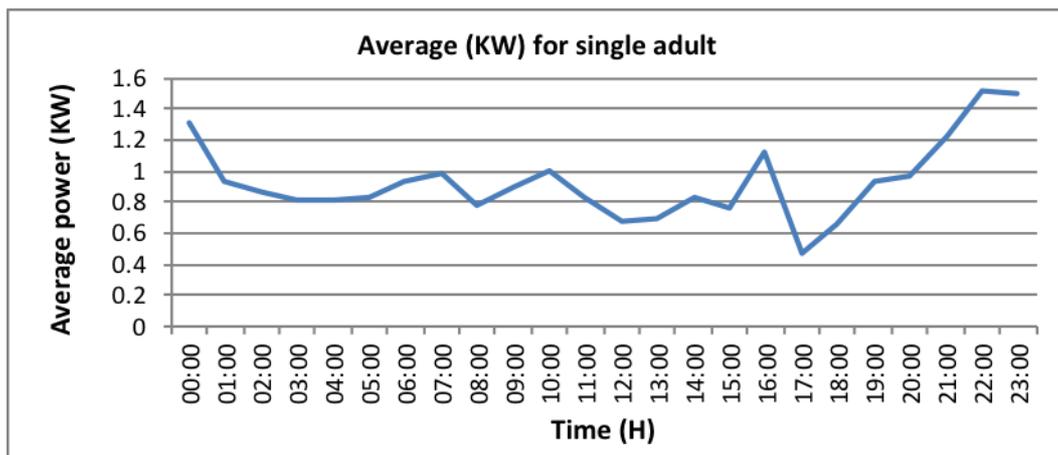


Figure (2): The average daily load curve for single adult during weekends.

In weekends, occupants usually do not wake up early, as they do not go to work. The life style of a single adult leaving in a flat is different than that of a family. This can be seen quite obvious from the details of load curve.

From 05:00 to 16:00, it can be noticed that the curve does not change much from flat line at 0.8 KW reading. The little increase and decrees reflects a minor change in the occupants habits as some

still wake up early. This is due to the wake up of some consumers at the dawn time when they begin their daily activity and use of different electrical utilities such as utensils for preparation of the breakfast, taking shower and washing clothes with automatic washing machine, cleaning with vacuum, preparation of lunch and accompanying use of electrical devices inside the kitchen, such as electrical heaters, electrical ovens and watching T.V. After that the curve starts gradually declining due to the gradual decrease of power demand, due to the siesta of most consumers. Then at 18:00, the curve increases gradually due to the increase of power demand. This may be explained by returning back of some consumers who were out to stay at the flats until the next day. Another reason could be the severe decrease in temperatures, particularly in evening, increase of heating demand until 22:30. At this time the curve starts gradually declining due to the gradual decrease of power demand. Then the curve takes approximately constant shape, increases or decreases slightly, due to some equipment working automatically during sleeping time, such as heaters, water heaters and refrigerators, thus it continues until 5:00. For single adult occupant flats, the maximum consumption recorded between 20:00 to 24:00 is 1.5 KW. This is understood since the occupant usually stays in their flats at these times and uses the electrical equipment. From 14:00 to 17:00, another peak is recorded since this time is common for preparing the lunch and using the washing machines. The minimum average power is around 17.00 because most occupants are outside their flats either to visit friends, families or to go shopping.

5.2 Type2: Two Adults with Children Household

There are 6 flats for two adults with children households (3, 6, 7, 8, 9, and 12).

5.2.1 Average Load Curve for Week Days

Figure (3) shows the average daily load curve for one flat of type two (two adult with children) on week days.

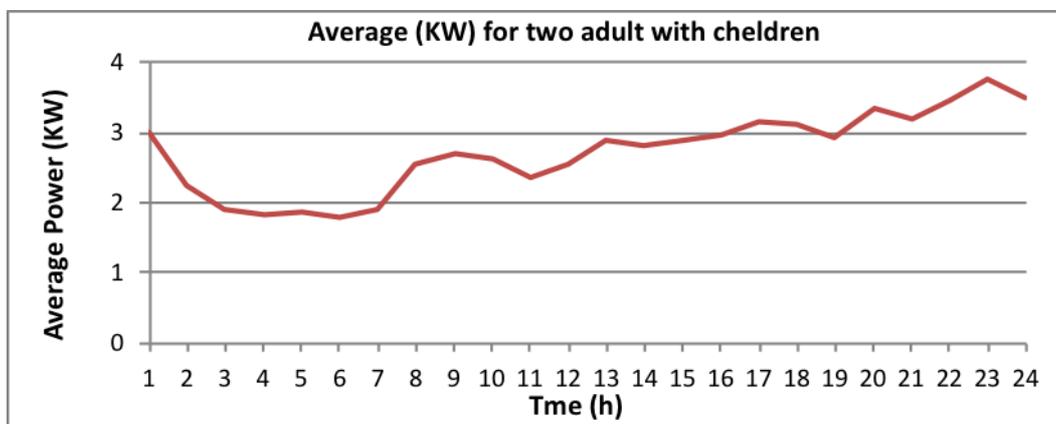


Figure (3): The average daily load curve for two adults with children during week days.

At 06:00, a gradual increase of the demand on electrical power can be noticed. This is due to the wake up of some of occupants living at this type of flats at dawn time when they begin their daily activity as mentioned before. Such activity continues disproportionately among the people who are ready for going out to work until 8:30, when consumers start leaving to work and children to schools. The curve gradually decreases, and continues in such decrease until 11:00, thereafter starts to increase due to the increase of demand on power by the housewives, starting the activities that can increase the consumption of electrical power, such as washing clothes with automatic washing machine, cleaning with vacuum, preparation of lunch and accompanying use of electrical devices inside the kitchen. The curve continues increasing until 18:00. After that, the load curve fluctuates between increase and decrease for short periods. Until 23:00. Then the curve starts gradually declining due to decrease on power demand and consumers go to bed. The curve afterwards takes stable shape and the base load is only on automatic equipment during sleeping time, such as heaters and refrigerators, it continues until 06:00.

5.2.2 Average Load Curve for Weekends

Figure (4) shows the average daily load curve for one flat of type two on weekends. From 06:00 to 16:00, a gradual increase and decrease on the electrical power consumption can be noticed due to the different daily activity of the families as mentioned before, after that the curve starts gradually declining due to the gradual decrease on power demand. Then at 17:00, the curve increases gradually due to the coming back of other consumers who were out, also due to the severe decrease in temperatures, particularly in evening. After 23:00, it can be seen that the curve starts gradually declining due to the gradual decrease on power demand, since most of consumers go to sleep. Then the curve takes approximately constant shape, increases or decreases slightly until 06:00. For two adults with children occupant flats, the maximum consumption recorded between 18:00 and 24:00 was 3.73 KW.

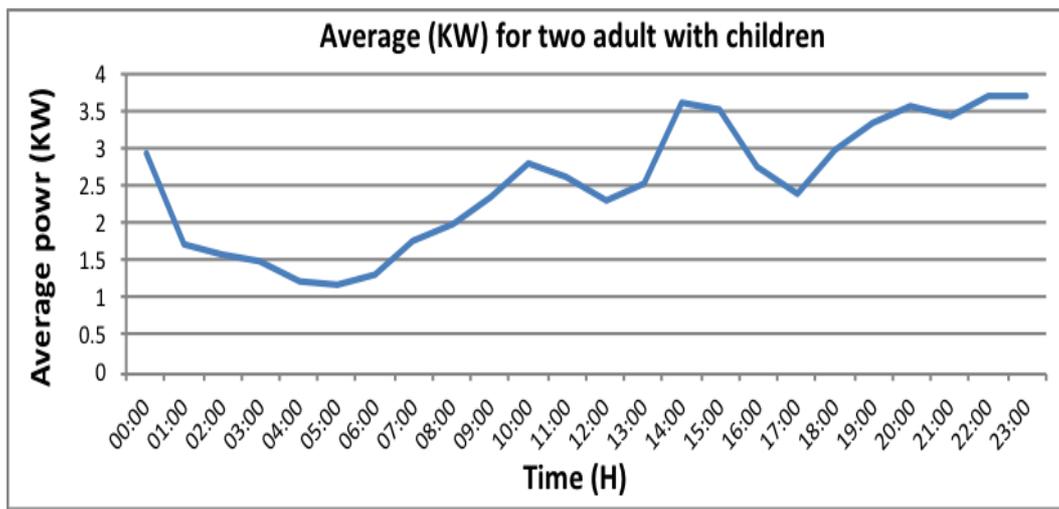


Figure (4): The average daily load curve for two adults with children during weekends.

5.3 All Flats with Two Types Household

5.3.1 Average Load Curve for Week Days

Figure (5) shows the average daily loads curve for all flats with two types on week days. The curve shows similarities in the behavior with the one presented in figure (3) which is average daily load curve of flats with two adults with their children. The similarity of the two curves is due to consumers living in this type of flats. It also may be attributed to the fact that the average always biased to the dominant values, which is in this case the consumption of the flats with bigger family members. The key factors of the behavior of this curve is the same as mentioned in the commentary on fig (3). As the power consumption depends on the size of the family and the available facilities at home, electrical power consumption in modern house increases after sunset, since this time people usually stay at homes and relax having meals and watching T.V. This was notable since the demand of power goes to minimum and load curve start declining since midnight the use of electrical power is limited to heating or a little light.

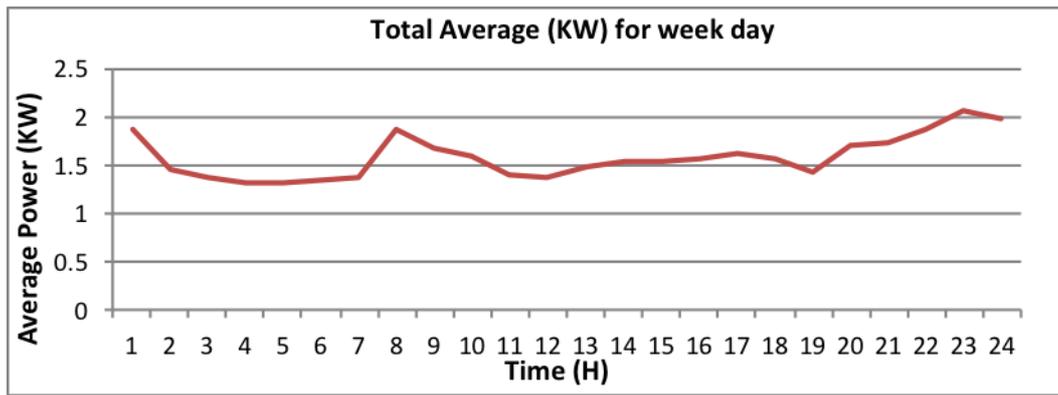


Figure (5): The average daily loads curve for all flats with two types on week days.

5.3.2 Average Load Curve for Weekends

Figure (6) shows the average daily loads curve for all flats with two types on weekends. It can be seen that the behavior of the curve is not more different than the curve of total average (KW) for weekends illustrated in figure (4). The similarity of the two curves is due the same reasons that mentioned in the previous section.

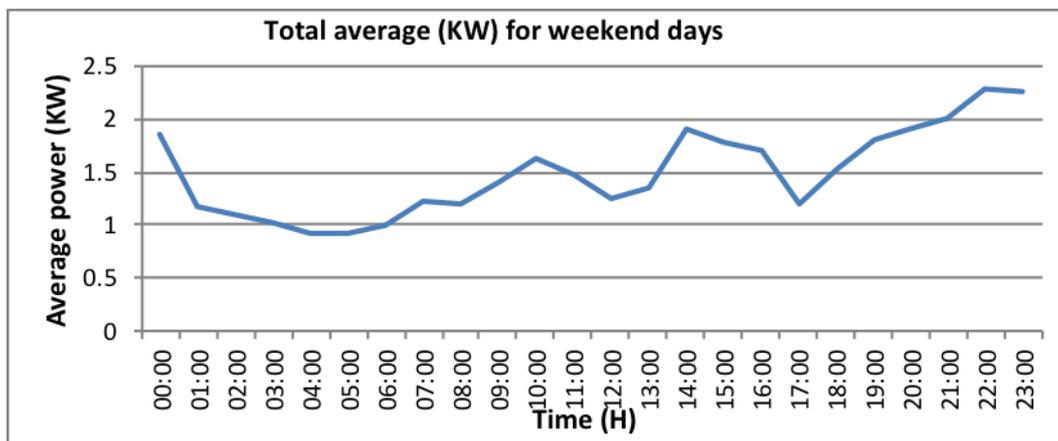


Figure (6): The average daily loads curve for all flats with two types during weekends.

5.4 Daily and Average Load Curves As Measured On the Main Cable

Figure (7) shows the average daily electricity load profiling for the main distribution box (DB). The average load curve records its minimum values between 01:00 and 05:00. At this time the main load is lighting, fridges and freezer (which usually stay on all the time. The maximum consumption was recorded between 23:00 and 24:00 was 34 KW. From 05:00 to 08:00, another peak is recorded since this time is common for preparing the breakfast and taking shower. And the minimum average power is 15 KW around 05.00 because the most occupants are sleeping.

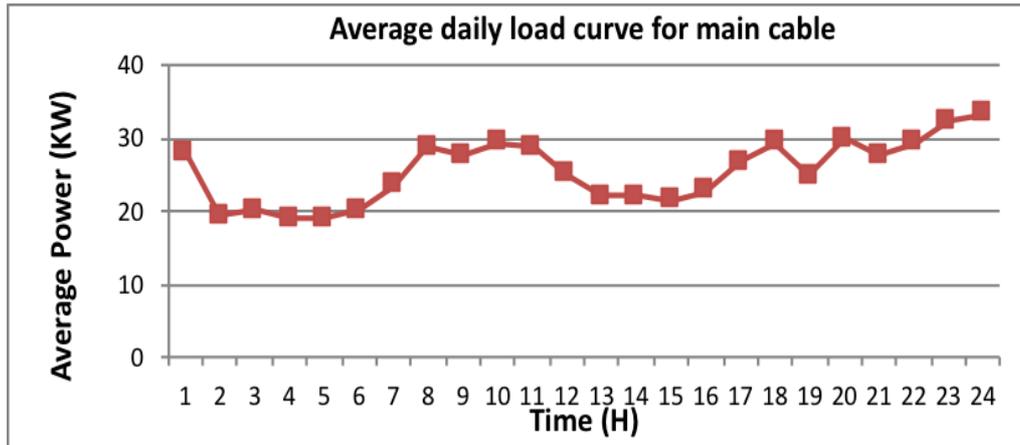


Figure (7): The average daily Power (KW) for the main distribution box during week days.

6. Comparison between Main Cable Load Curve and Average Load Curve of the Flats

The plots of actual load curve for the flats as resulted from individual load of each flat (*Total average KW for (13) flats in block 2 = Max. Average power per flat * No. of Flats in Block 2 (which is 13)*) and the load curve as measured at the main feeder are shown in figure (8). In normal cases, these two curves should be very close to each other and never coincides. This curve represents the sum of average electrical load for all flats of the two previous types of consumers, and the average electrical load measured at the main cable to all flats. The main difference between the two curves is the power consumed in lifts and water pumps. However, four flats (flat13, flat14, flat15, flat16) were not included in the measurements due to social and technical impairments. The common area lights for all block (2) are also not included.

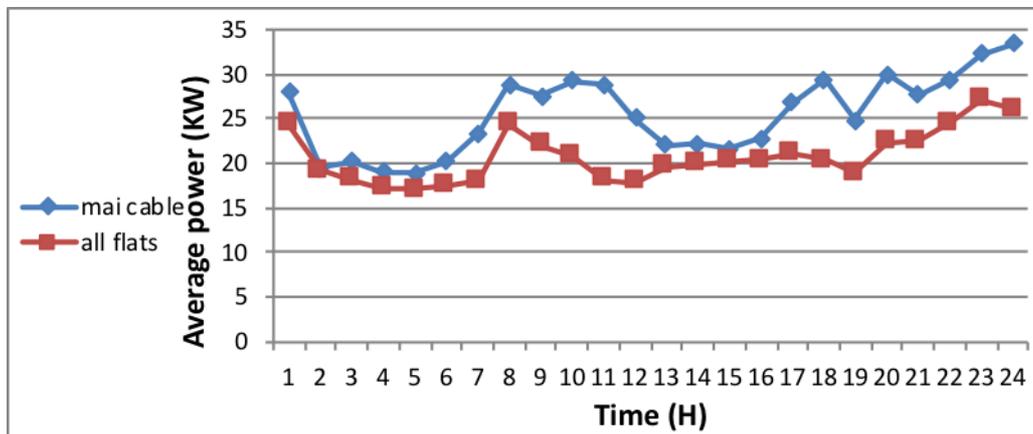


Figure (8): The average daily load curves for the main cable and all flats in building No. 2.

7. Average Daily Flat Load Shape

Figure (9) shows the average Daily Flat Load Shape in one day. The fluctuation in the power consumption is from 2.2 KW max to 1.33 KW min. This means the power consumption by consumers are nearly the same during each day of the week (including weekends), which reflects the power consumption behavior of the residents using different appliances in different times.

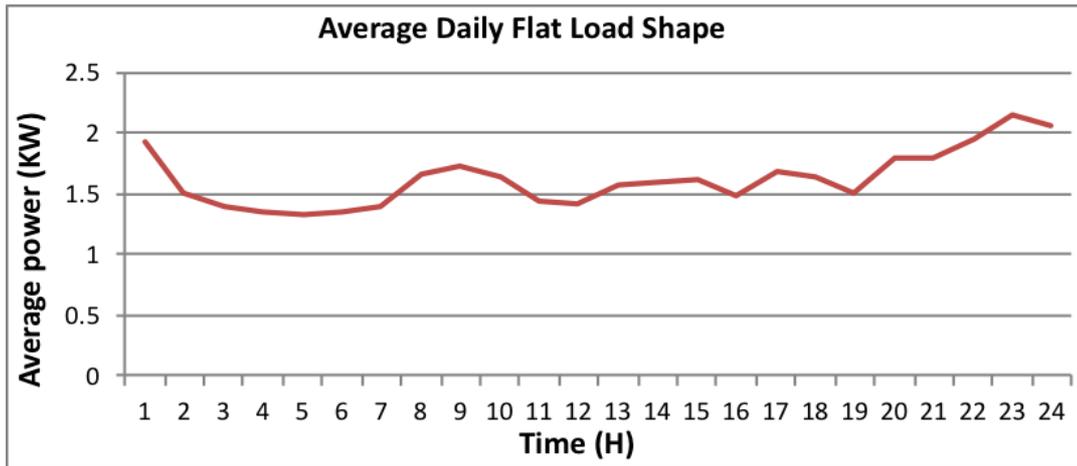


Figure (9): The average Daily Flat Load Shape.

8. Comparison between the Outcomes Obtained from this Study and the Outcomes of the American Consultant Cadmus Study

As mentioned before, the Cadmus Group (Cadmus) with other corporations were contracted by the General Electricity Company of Libya (GECOL) to provide an assessment of the demand side management (DSM) potential in Libya in 2009. Table (1) shows the comparisons of the main outcomes which demonstrate that the outcomes in terms of the average power may differ based on the accuracy and the parameters, such as a pure residential feeder or mixed one, included in the measurements and calculations.

Table 1: A comparison between this study and the Cadmus 2010 study

Study	Flat area	measurement Areas	Time of measurements	Duration between each reading	Measurement period	Daily demand in (KW)		
						Average	Max.	Min.
This Study	180 m ²	Zawet Eddahmani	All the day long for 20 days	5 minutes	All the daylong for 20 days in February and March	2.72	3.77 at 10 p.m.	1.77 at 5 a.m.
Cadmus 2010 Study	200 m ²	Airport road - Azzuhur quarter - Enjila	All the daylong for 15 days	15 minutes	All the daylong for 15 days from 12/2010 to 01/2010	1.8	2.4 at 7 a.m.	1.4 at 5 a.m.

9. Conclusions

This paper presents a method of generating realistic electricity load profile data for the city of Tripoli domestic buildings. This method is based on the information and results of previous investigations and works that is available in the public reports and statistics.

Two household types, which present the behavioural characteristics and common occupancy pattern in the city of Tripoli, have been proposed to regenerate the diversity in electricity consumption between households. The five minutes daily load profiles for proposed community of 108 flats have been generated for the various types and then the average electricity consumption

for individual flat has been calculated. The generated load profiles show a good agreement in flats load profiles compared with the other residential types of the city.

This method has allowed the creation of typical seasonal and lifestyle activity profiles from statistical data. These profiles play an important role in predicting the future electricity demand and in designing and distributing the loads. The daily load profile from individual dwelling to community can be predicted using this method.

The measurements reflected the consumption behaviours of different size of families during working and weekend days. The measurements conducted under this study may be considered as of high accuracy. That is because the measuring units measure at (5) minutes intervals and the selected area is 100% residential, i.e. no other type of loads are fed by the main feeder to the complex.

Acknowledgment

This study was supported in partially by the General Electric Company of Libya (GECOL).

References

- [1]. Turan Gonen, “Electrical Power Distribution System Engineering”, 1986, McGraw-Hill Inc.
- [2]. M. Ihabal, H. S. Rajamani, M. K. Jalboub, M. A. Elgadal and M. M.Hebal, “Electricity load profiling for the UK Domestic Buildings”, the Fifth Libyan Arab International Conference on Electrical and Electronic Engineering, October 23 – 26, 2010, Tripoli, Libya
- [3]. Yigzaw G. Yohanis , Jayanta D. Mondol, Alan Wright, Brian Norton, “Real-life energy use in the UK: How occupancy and dwelling characteristics affect domestic electricity use”, Science Direct Energy and Buildings 40 (2008) 1053–1059.
- [4]. Annual report in year 2009 for General Electric Company of Libya.
- [5]. The CADMUS group, Inc. /Energy Services "Winter Load Research" November 18, 2010.
- [6]. Salem A. S. Ahbil and Hamid H. Sherwali, “Electricity Load Profiling and Diversity Factor Estimation for Coastal Housing Complexes based on the Measurements of Flats Actual Load”, Master of applied Science thesis, University of Tripoli, Tripoli, Libya, 2013.



Track No. 5: Information Technology

إنشاء المنزل الذكي باستخدام برنامج المحاكاة Packet Tracer 7.3 و لوحة الاردوينو Arduino board

د. أبو عجيبة دغمان*، محمد جويلي ، حسين درنية ، و ابوذر رمضان
قسم شبكات الحاسوب، كلية تقنية المعلومات، جامعة الزنتان، ليبيا
*العنوان البريدي (abuajila@uoz.edu.ly)

ABSTRACT

المنزل الذكي Smart Home هو الذي يمكن التحكم فيه آليا من خلال أنظمة ذكية و التي تحتوي على أجهزة تحكم متطورة يتم تركيبها في المنزل ويتم التحكم بها عن طريق شاشات اللمس المثبتة على الجدار أو جهاز iPad أو بواسطة الهاتف الذكي Smart Phone حيث تتم عملية التحكم والمراقبة لجميع الأجهزة الكهربائية والإلكترونية مثل الإضاءة و أجهزة التكييف والتلفاز والكاميرات والأبواب الكهربائية. كما أن نظام المنزل الذكي يخبر عن أي تغيير يحدث في المنزل بواسطة الرسائل النصية SMS أو الاتصال أو أي طريقة يفضل المستخدم استخدامها. كما بإمكان المستخدم إضافة كاميرات المراقبة ومجسات الاستشعار للحركة وذلك لنظام الأمن والحماية بالمنزل [1]. في هذه الورقة البحثية سوف يتم التطرق الى تقنية الشبكات اللاسلكية Wi-Fi وتقنيات انترنت الأشياء IoT حيث أن هاتين التقنيتين لا بد من استخدامهما لإنشاء وتطبيق المنزل الذكي Smart Home. حيث تم تصميم وتطبيق المنزل الذكي باستخدام برنامج المحاكاة Packet Tracer 7.3 و لوحة الاردوينو Arduino board.

Keywords:

انترنت الأشياء (IoT) ،
المحاكاة (Simulation) ،
لوحة الاردوينو (Arduino Board) ،
المنزل الذكي (Smart Home).

1. المقدمة

المنزل الذكي Smart Home هو الذي يمكن التحكم فيه آليا من خلال أنظمة ذكية و التي تحتوي على أجهزة تحكم متطورة يتم تركيبها في المنزل ويتم التحكم بها عن طريق شاشات اللمس المثبتة على الجدار أو جهاز iPad أو بواسطة الهاتف الذكي Smart Phone حيث تتم عملية التحكم والمراقبة لجميع الأجهزة الكهربائية والإلكترونية مثل الإضاءة و أجهزة التكييف والتلفاز والكاميرات والأبواب الكهربائية. كما أن نظام المنزل الذكي يخبر عن أي تغيير يحدث في المنزل بواسطة الرسائل النصية SMS أو الاتصال أو أي طريقة يفضل المستخدم استخدامها. كما بإمكان المستخدم إضافة كاميرات المراقبة ومجسات الاستشعار للحركة وذلك لنظام الأمن والحماية بالمنزل [1]. في هذه الورقة البحثية سوف يتم التطرق الى تقنية الشبكات اللاسلكية Wi-Fi وتقنيات انترنت الأشياء IoT حيث أن هاتين التقنيتين لا بد من استخدامهما لإنشاء وتطبيق المنزل الذكي Smart Home. حيث تم تصميم وتطبيق المنزل الذكي باستخدام برنامج المحاكاة Packet Tracer 7.3 و لوحة الاردوينو Arduino board.

1.1 تقنية الشبكات اللاسلكية (Wi-Fi)

الشبكات اللاسلكية Wireless LAN هي أي نوع من الشبكات الحاسوبية التي تعمل على نقل المعلومات بين الأجهزة المتصلة بالشبكة عبر الأمواج الكهرومغناطيسية بشكل لاسلكي، أي من دون استخدام الأسلاك (الكابلات). إن هذا النوع من الشبكات ينفذ عادةً مع نظم نقل معلومات بالتحكم عن بُعد من خلال استخدام أمواج كهرومغناطيسية كأمواج الراديو وذلك لإرسال واستقبال الإشارات. هذا التنفيذ يتم في الطبقة الفيزيائية من الشبكة. تعتبر أجهزة ال-Wi-Fi والتي تعمل جميعها وفق معيار IEEE 802.11 من الأجهزة الأكثر استخداما في الشبكات اللاسلكية المحلية.

هناك عدة معايير للشبكات اللاسلكية حددها معهد المهندسين الإلكترونيين والكهربائيين Institute of Electrical and Electronics Engineers (IEEE) أكثرها شيوعا هي معايير IEEE 802.11a ، IEEE 802.11b ، IEEE 802.11n ، هذه المعايير متوافقة مع بعضها البعض [2].

2.1 انترنت الأشياء (Internet of Things)

يشير مصطلح إنترنت الأشياء (Internet of Things) إلى ملايين الأجهزة و الأدوات المرتبطة بالإنترنت في جميع أنحاء العالم. حيث يمكن لهذه الأجهزة أن تجمع البيانات وتشاركها فيما بينها. و بفضل توافر المعالجات الرخيصة والشبكات اللاسلكية

WLAN، أصبح من الممكن تحويل مصباح ضوئي، جهاز تكييف، أو أي جهاز كهربائي آخر إلى جزء من عالم «إنترنت الأشياء». ومن هنا تبرز أهم مميزات إنترنت الأشياء، وهي أنها تتيح للإنسان التحرر من المكان، أي أن الشخص يستطيع التحكم في الأجهزة المنزلية من دون الحاجة إلى وجوده في مكان محدّد للتعامل مع جهاز معين.

وفقاً لما تم نشره عبر مؤسسة Cisco Academy، فإن ثلث الأجهزة المتصلة بالإنترنت سيكون أجهزة الكمبيوتر والهواتف الذكية وأجهزة التلفزيون الذكية. أما عن التلّين الأخرين سيكونان أنواعاً أخرى من "الأشياء": مثل المستشعرات والمحركات والأجهزة الذكية المبتكرة حديثاً والتي تراقب وتحكم وتحلّ وتحسّن عالماً. بعض الأمثلة على المستشعرات المتصلة الذكية هي: أجراس الأبواب الذكية، وأبواب المراب، وأجهزة تنظيم الحرارة، والأجهزة الرياضية القابلة للارتداء، وأجهزة تنظيم ضربات القلب، وإشارات المرور، ومواقف السيارات، وغيرها الكثير [3].

وعادة ما يستخدم مصطلح «إنترنت الأشياء» بشكل أساسي للأجهزة التي لا يُتوقع اتصالها بالإنترنت، فلا يُعتبر الكمبيوتر الشخصي ولا الهاتف الذكي جزءاً منها، لكن قد يتم تصنيف ساعة ذكية أو خاتم ذكي متعقب للياقة كجهاز لإنترنت الأشياء [5]. [4]

3.1. مراجعة المؤلفات ذات الصلة

الكثير من الدراسات العلمية انجرت في مجال إنترنت الأشياء (IoT) في السنوات الأخيرة. كانت كلها تهدف إلى تطبيق فكرة IoT في مجالات الحياة العملية المختلفة. فمثلاً على سبيل المثال لا الحصر في المرجع [3] كان الهدف من الدراسة هو تصميم مدينة ذكية حيث يمكن لكاميرات الويب الذكية نقل حزم البيانات عن بُعد. تشير الدراسة إلى أنه عند توصيل كاميرا الويب الذكية عبر وسيط لاسلكي، يمكنها إرسال واستقبال حزم البيانات عبر والتي يمكن للمستخدمين الوصول إليها عبر شبكة الإنترنت.

كذلك قام باحثون باستخدام برنامج المحاكاة Cisco Packet Tracer لإنشاء وتطبيق فكرة المنزل الذكي. كان الهدف من الدراسة معرفة مدى إمكانية تطبيق IoT على المنازل الذكية اخذين في عين الاعتبار توفير الحماية اللازمة للبيانات المرسلّة بين الأجهزة [6].

دراسة أخرى كانت تهدف إلى تطبيق الجامعة الذكية باستخدام تقنيات IoT معتمدين بشكل أساسي على الإصدار 7.3 من برنامج المحاكاة Cisco Packet Tracer والذي بدوره يحتوي على أجهزة IoE مختلفة [7] [8].

على عكس الدراسات السابقة والتي اعتمدت إما على نوع واحد من أجهزة إنترنت الأشياء لتصميم مدينة ذكية كما في [3] أو التي ركزت بشكل أساسي على تصميم المنزل الذكي مقتصرين على استخدام برنامج المحاكاة Cisco Packet Tracer كما في [6] و [7] و [8]، فإن الهدف الأساسي من هذه الورقة البحثية هو استخدام برنامج المحاكاة Cisco Packet Tracer لإنشاء المنزل الذكي حيث تم توصيل أجهزة إنترنت الأشياء المختلفة ومعالجتها بطريقة تسمح بنقل المعلومات القيمة واستخدام جهاز Home Gateway ليتم التحكم فيها عن بُعد في أي مكان آخر بواسطة الهواتف الذكية. الجزء الآخر من الورقة البحثية كان يهدف إلى تحويل ماتمت محاكاته إلى الواقع العملي ولكن نظراً لعدم توفر جميع أجهزة IoT في السوق فقد تم تحقيق جزء ما تمت محاكاته وذلك باستخدام لوحة الأردنيو Arduino Board.

2. الجانب العملي والمنهجية

هدف هذا البحث إنشاء نموذج يحاكي المنزل الذكي Smart Home وذلك باستخدام برنامج المحاكاة Cisco Packet Tracer وأيضاً إنشاء نموذج واقعي للمنزل الذكي وذلك من خلال استخدام لوحة الأردنيو وأدواتها.

1.2. برنامج المحاكاة (Packet Tracer)

برنامج المحاكاة Packet Tracer هو برنامج محاكاة مرئي تم تصميمه بواسطة Cisco حيث يتيح للمستخدمين إنشاء وتصميم شبكات حاسوبية تحاكي الواقع لأنه في كثير من الأحيان لا يمكن تصميم شبكات حقيقية لغرض الدراسات البحثية وذلك بسبب تكاليفها الباهظة. برنامج المحاكاة Cisco Packet Tracer يسمح للمستخدمين بتكوين اعدادات الشبكة باستخدام واجهة سطر أوامر Command Line Interface CLI أو من خلال واجهة المستخدم الرسومية Graphical User Interface GUI مما يسمح للمستخدمين بإضافة وإزالة أجهزة الشبكة المحاكية بالشكل الذي يروونه مناسباً.

الإصدارات الحديثة لبرنامج المحاكاة Cisco Packet Tracer دعمت الكثير من الأجهزة التي تستخدم في تقنيات إنترنت الأشياء مثل الهاتف الذكي، جهاز قياس حرارة الغرفة، أجهزة الاضاءة الذكية، وغيرها الكثير من الاجهزة [9]. يبين الشكل (1) واجهة برنامج المحاكاة Cisco Packet Tracer الإصدار 7.3. هناك الكثير من الخصائص التي يتميز بها برنامج المحاكاة Cisco Packet Tracer عن غيره من برامج المحاكاة الأخرى. حيث ان البرنامج يسمح للطلاب بتصميم شبكات معقدة، يسمح للطلاب باكتشاف مفاهيم إنترنت الأشياء، يتيح للمستخدمين بناء وتصميم وتكوين المدينة الذكية و المنزل

الذكي، كما انه يوفر تصوراً واقعياً ومحاكاة لاجهزة إنترنت الأشياء. برنامج المحاكاة Cisco Packet Tracer تم اختياره في هذه الدارسة بناء على الخصائص سالفة الذكر [10].



شكل(1): واجهة برنامج المحاكاة Packet Tracer 7.3

2.2. لوحة الاردوينو (Arduino board)

لوحة الاردوينو هي لوحة الكترونية مفتوحة المصدر Hardware Open Source تستخدم لتطوير الكثير من المشاريع والأفكار المتعلقة بالتحكم الآلي بصوره سهلة و بسيطة وذلك عن طريق استخدام Arduino Integrated Development Environment (IDE) [11].

تعتبر بيئة التطوير Arduino IDE هي الأداة المستخدمة في كتابة الشفرات البرمجية بلغة C وتحويلها بعد ذلك إلى صيغة تنفيذية يمكن وضعها على المتحكم الدقيقة الموجودة على اللوحة. هناك عدة انواع من لوحة الاردوينو وهي: Arduino Uno، Arduino Mega، Arduino Nano، Arduino Mini، و Arduino Lilypad.

تختلف هذه اللوحات عن بعضها البعض من ناحية عدد المخارج والمداخل والتي تحدد عدد الأجهزة التي يمكن التحكم بها وعدد الحساسات Sensors التي يمكن دمجها مع اللوحة وكذلك نوع المتحكم الدقيقة وسرعة المعالج الموجود بداخلها وإمكانية تبديلها. في هذه الدراسة تم استخدام لوحة الاردوينو من نوع Arduino Uno وذلك لان هذا النوع يعتبر من أكثر الأنواع استخداماً في المنازل الذكية. الشكل (2) يوضح لوحة الاردوينو Arduino Uno.

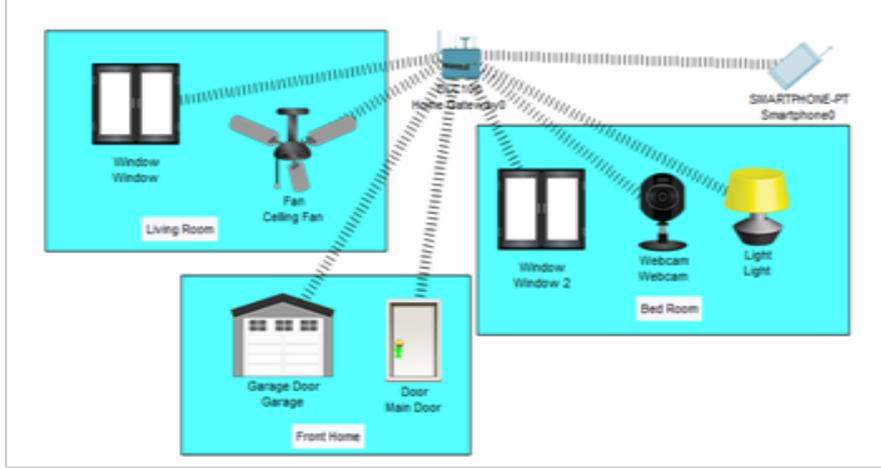


شكل (2): لوحة الاردوينو Arduino Uno

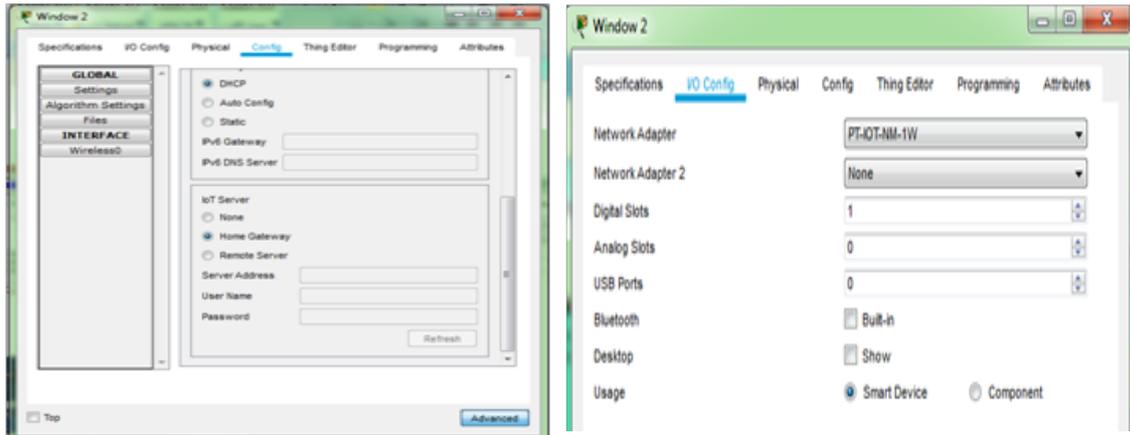
3. النتائج والمناقشة

1.3. النتائج المتحصل عليها من برنامج المحاكاة (Packet Tracer)

في البداية تم استخدام أداة Home Gateway كما هو موضح في الشكل (3) كأداة ربط بين أجهزة IoT. كما هو مبين بالشكل نلاحظ أن أجهزة IoT (Webcam, Door, Window, Light, Smartphone, etc) مرتبطة مع Home Gateway بطريقة لاسلكية وذلك بتغيير بطاقة الشبكة إلى بطاقة لاسلكية كما موضح في الشكل (4). ولتتمكن أجهزة IoT من التعرف على Home Gateway يتم ذلك من خلال تغيير خاصية IoT Server إلى Home Gateway كما مبين بالشكل (5).



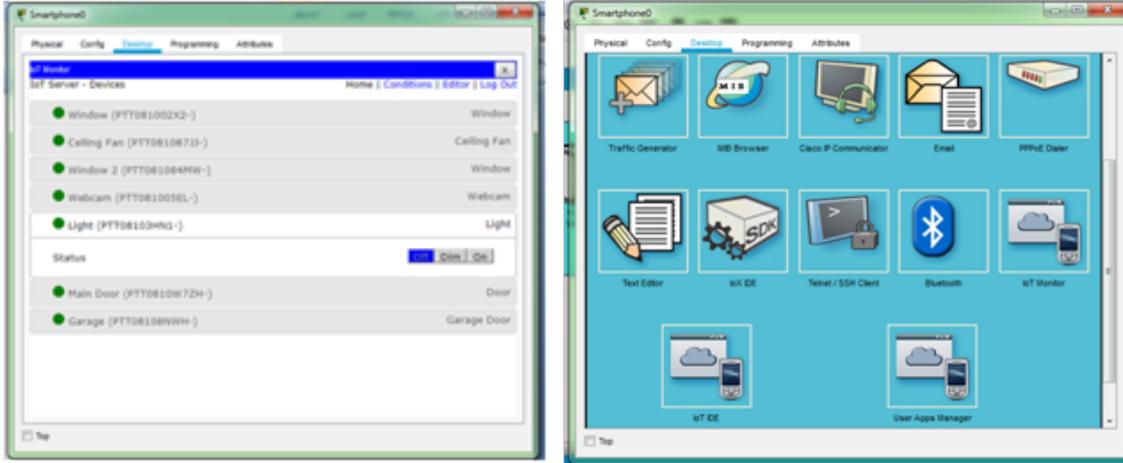
الشكل (3): أجهزة IoT مرتبطة مع Home Gateway



الشكل (5): تغيير خاصية IoT Server إلى Home Gateway

الشكل (4): تغيير بطاقة الشبكة إلى بطاقة لاسلكية

بالرجوع إلى الشكل رقم (3) فقد تم استخدام الهاتف الذكي Smart Phone فقد للتحكم في باقي أجهزة IoT. من خلال خاصية IoT Monitor في الهاتف الذكي كما موضح بالشكل (6)، فإن المستخدم يمكنه التحكم عن بعد في جميع أجهزة IoT المرتبطة مع جهاز Home Gateway من الهاتف الذكي. من خلال هذه الخاصية يمكن للهاتف الذكي من التحكم بحالات باقي أجهزة IoT على سبيل المثال يمكن التحكم في حالة جهاز Light من On إلى Off كما هو موضح بالشكل (7).



الشكل (7): تغيير حالات اجهزة IoT من خلال خاصية IoT Monitor

الشكل (6): خاصية IoT Monitor في جهاز Smartphone

2.3. النتائج المتحصل عليها من لوحة الاردوينو (Arduino board)

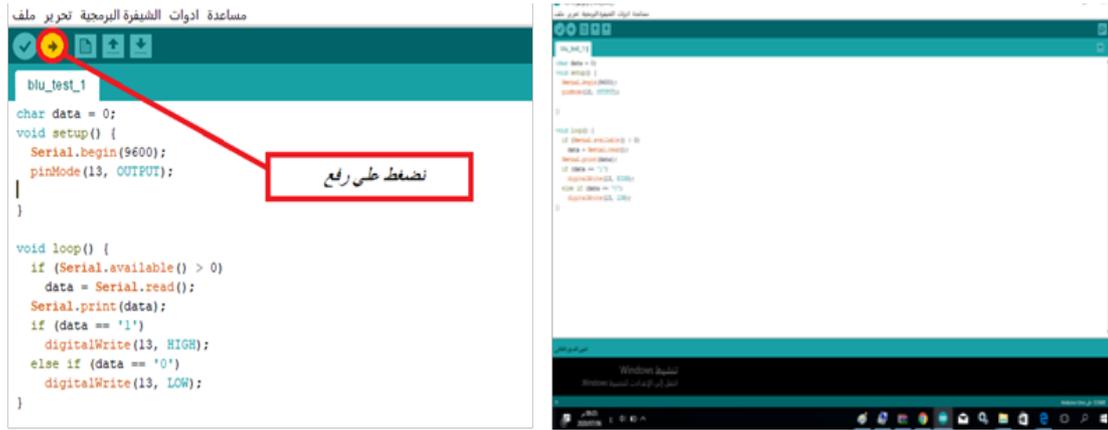
في البداية تم ربط لوحة الاردوينو بقطعة البلوتوث و باقي القطع المستخدمة و ذلك لجعل لوحة الاردوينو تنقل البيانات من و الي الحاسب. الشكل (8) يبين القطع المستخدمة مع لوحة الاردوينو وهي قطعة بلوتوث Bluetooth module CH-05 ، مصدر تغذية خارجي (cable) ، اسلاك توصيل Solderless jumper ، جهاز قياس فولتميتر، و مصابيح LED.



الشكل (8): القطع المستخدمة مع لوحة الاردوينو

بعد توصيل كامل القطع تم تعريف جهاز البلوتوث HC-05 المربوط مع الاردوينو مع جهاز الهاتف الذكي Smartphone. بعد إتمام عملية توصيل القطع، تم استخدام برنامج Arduino IDE لبرمجة لوحة الاردوينو بينما تم استخدام برنامج LED controller للاتصال بقطعة البلوتوث واستخدامه في واجهة التحكم.

برنامج Arduino IDE يمكنه التحكم في لوحة الاردوينو عن طريق لغة برمجة Arduino C كما هو موضح بالشكل (9). في البداية يتم كتابة النص البرمجي و من ثم يتم تحميله الي لوحة الاردوينو حيث يحتوي النص البرمجي علي آلية العمل التي سوف يتم تطبيقها علي مصابيح LED. بعد الانتهاء من كتابة الكود نقوم بالضغط علي زر تأكيد Verify الذي يقوم بالتحقق من صحة كتابة النصوص البرمجية قبل تحميلها. قبل عملية التحميل Upload يجب التأكد من أن كابل لوحة الاردوينو متصل بالكمبيوتر ثم يتم الضغط علي زر التحميل Upload وذلك حتى تتم عملية تحميل النص البرمجي علي لوحة الاردوينو. الشكل (10) يبين كيفية تحميل النص البرمجي.

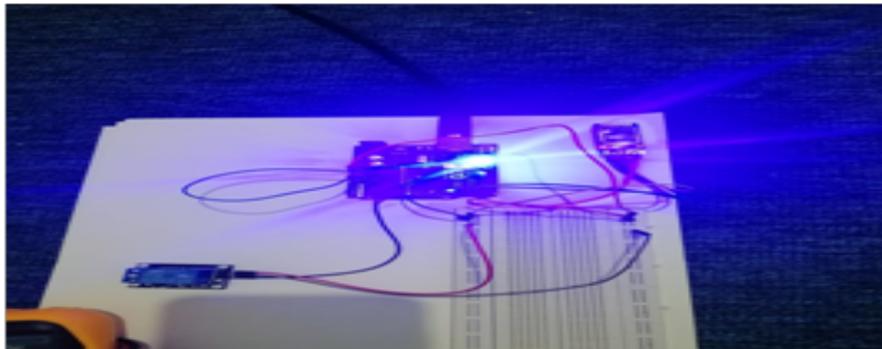


الشكل (10): كيفية تحميل النص البرمجي.

الشكل (9): برنامج Arduino IDE

خلال عملية التحميل يلاحظ أن مصابيح الـ LEDs الموجودة على اللوحة والتي معنون بجانبها TX- RX تضيئ بسرعة عالية ثم تقف قليلا و هذه الإضاءة تعني أن اللوحة بدأت تستقبل بيانات النص البرمجي الذي تمت كتابته و تحميله.

بعد ما أصبحت لوحة الاردوينو جاهزة وتم تحميل النص البرمجي عليها يتأتى دور برنامج LED Controller و الذي باستخدامه يمكن التحكم في تشغيل و إيقاف LED كما موضح في الشكل (11). LED Controller هو عبارة عن تطبيق يتم تحميله على الهاتف الذكي وربطه بلوحة الاردوينو المبرمجة عن طريق Bluetooth حيث يمكن من خلاله تشغيل وإيقاف إضاءة مصابيح LED. من هنا يتبين انه بإمكاننا ربط المصابيح الكهربائية أو أي أجهزة كهربائية أخرى عن طريق لوحة الاردوينو حيث يتم تشغيلها او إيقافها عن التشغيل



الشكل (11): التحكم في مصابيح LED باستخدام LED Controller

4. الخاتمة والتوصيات

إن بناء نظام شامل لإنشاء المنزل الذكي هي خطوة كبيرة في مجال تطوير هذا النوع من الأجهزة الذكية وتحتاج إلى تطوير دائم للتناغم مع التطور الهائل للتكنولوجيا. في هذه الدراسة البحثية تم تطبيق فكرة المنزل الذكي باستخدام برنامج المحاكاة Packet Tracer 7.3 و كذلك تم استخدام لوحة Arduino لتحويل جزء من برنامج المحاكاة الى تجارب عملية. حيث أثبتت النتائج أن المنزل الذكي يمكن تطبيقه ليس فقط باستخدام برامج المحاكاة Cisco Packet Tracer وإنما يمكن تطبيقه عمليا على الواقع وذلك باستخدام لوحة الاردوينو. نظرا لمحدودية القطع الملحقة في السوق فإن هذا العمل يحتاج إلى المزيد من التطوير في المستقبل وذلك لزيادة أمانة هذا النظام، الاتصال البعيد باستخدام الشبكة العالمية الانترنت ، إمكانية استخدام أدوات الاستشعار sensors ، وإمكانية الوصول والتحكم بدائرة الاردوينو عن طريق web browser.

المراجع

- [1]. Introduction of Internet of Things (2020) [access in February 2020] at <https://www.netacad.com>.



- [2]. IEEE computer society standard committee (2005) Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications.
- [3]. Anthony Bawa Maxwell Lewis Selby (2018) Design and Simulation of the Internet of Things for Accra Smart City, *Network and Complex Systems*, Vol. 8, PP, 17-30, 2018.
- [4]. Al-Emran, M., Malik, S. I., and Al-Kabi, M. N. (2020). A Survey of Internet of Things (IoT) in Education: Opportunities and Challenges. In *Toward Social Internet of Things (SIoT): Enabling Technologies, Architectures and Applications*. pp 197-209, Springer.
- [5]. Introduction of Internet of Everything (2019) [access in September 2019] at <https://www.netacad.com>.
- [6]. R. Tabeidi, S. Masaad, and B. Elshaikh (2019) Implementing Smart College Using CISCO Packet Tracer 7.2 Simulator, *Journal of Engineering Research and Application*, Vol 9, Issue. 4, pp 44-39.
- [7]. G. Alfarsi, J. Jabbar, R.Tawafak, and S.Malik (2020) Using Cisco Packet Tracer to simulate Smart Home, *International Journal of Engineering and Technical Research*.
- [8]. G.Ashok, P. Akram, M. Neelima, J. Nagasaikumar, and A.Vamshi (2020) Implementation Of Smart Home By Using Packet Tracer, *International Journal of Scientific & Technoloy Research*, Vol 9, Issue. 2, pp 678-685.
- [9]. Packet Tracer 7.3 [accessed in December 2019] at <https://www.netacad.com/courses/packet-tracer>
- [10]. Introduction to Packet Tracer (2019) [access in December 2019] at <https://www.netacad.com>.
- [11]. Introduction to Arduino Board [access in January 2020] at <https://www.arduino.cc>

The Virtual University: Trends and Challenges in Libya Elmergib University as Case Study

Muna Mohamed Ali Adrah¹ and Gharsa Ali Elmarash².

Computer Science Department, Faculty of Science, Elmergib University

¹mmadrah@elmergib.edu.ly, ²gharsa.ali@elmergib.edu.ly

ABSTRACT

E-Learning is playing a significant role in education to improve student's skills and teach them new ways for managing their knowledge and information. Recently few public universities in Libya have initiated limited attempts to use e-learning alongside traditional classroom. However, different obstacles are preventing successful adoption of this technology. It is obvious that effective ICT implementation should focus on understanding individual faculty needs. Then a thorough investigation is conducted in relation to the importance of implementing electronic education at Libyan universities, and including faculty members in every step of the planning and implementation of up-to-date technologies, as well as the challenges and solutions of ICT implementation. This paper focused on the overall scenario of virtual universities by narrative review of past researches.

The present research aimed to highlight challenges that hinder effective implementation of e-learning in Libya and recommend possible solutions to tackle them. A total of 70 respondents voluntarily participated in this research. They consisted of academic staff (N=8), professors in charge of e-learning (N=2), and undergraduate students (N=60). The method used to collect data is a survey instrument. Data was then analyzed and reported quantitatively and qualitatively. This provided in-depth understanding to the current status of e-learning in Libyan universities and highlighted major hindrances of its successful application. Based on this analysis, the study proffered many recommendations that should be considered in order to fully benefit from e-learning technologies and the possible of converting Elmergib university to virtual university.

Keywords:

Information
Communication
Technologies (ICTs),
e-learning,
virtual university,
e-university,
challenges of e-learning.

1- Introduction

The impact of communications and information technology has extended to the fields of work, where the tremendous and successive developments in the field of communications and information have created a new reality that will bring about major changes in various aspects of economic, social, military, space and other life, and thus many new specialties have been created, and then new areas of work. In addition to the decrease in demand for some specialties and other jobs.

In the recent years, we have seen exponential growth of Internet-based learning. The transition to online technologies in education provides the opportunities to use new learning methodology and more effective methods of teaching. The major perspectives of using e-learning can be generalized as follows: an opportunity for overcoming the limitations of traditional learning, such as large distance, time, budget or busy program; equal opportunities for getting education no matter where you live, how old you are, what your health and social status is; better quality and a variety of lecture materials; new consortia of educational institutions, where a lot of specialists work in collaboration,

use shared resources and the students get freedom to receive knowledge, skills and experience from other universities [1].

On the other hand, with the emergence of the Corona (COVID-19) pandemic globally at the present time, it has led us to resort to alternative methods to obtain high-quality education, and virtual education is one of the important means by which we can take advantage of modern technology to provide this type of education to students, as virtual education university students and others of students can acquire any kind of knowledge that they wish to obtain.

In Libya, there are significant differences in the perceptions of Libyan academics as to the potential use of the Internet for research purposes and the likely benefits from Internet access. To avoid potential problems in using the Internet for research purposes, a study by a Libyan researcher in 2012 [2] mentioned that there are a number of factors that could Keep staff in Libyan universities from using the internet for academic research purposed, and the study proposed several solutions to reduce the level of resistance displayed by some members of academic staff to the wide use of ICT tools and implementation of the national ICT policy , the most important of these factors: lack of access to specialized on line databases, lack of internet access, low speed of connection, quality of the information source, system availability, lack of skill in the English language, technical difficulties, high cost of Internet connections, lack of training, computer support and Internet skills, social factor and lack of encouragement and incentives from the educational institutions.

Today, many of high educational institutions have begun to use e-learning, so there is a need to know a new term called virtual university, which is an educational platform that simulates all university services through the Internet by creating a campus environment in accordance with criteria to be determined by academicians specialized in educational institutions [3]. Improving education systems in universities has become an essential concern for all nations, with a considerable amount of research being conducted in this area, particularly in terms of studying tools to facilitate the learning process and thereby improve the outcome and quality of national higher education systems. In the context of contemporary higher educational systems, Information and Communication Technology (ICT) is considered an essential tool to facilitate the acquisition of cognition and knowledge [4].

2- Paper Objectives

In this research, the main focus is on Libyan academic staff and student perspectives of what are the key challenges to the use of e-learning in higher education. And how successful the idea of converting the higher education system to a virtual education system by using e-learning. According to the best of the authors' knowledge, this area of research is unexplored in Libya because e-learning is a new experience in all Libyan universities. Hence, the present paper seeks to:

- identify the current status of e-learning application in Elmergib University.
- highlight challenges that hinder the effective integration of e-learning.
- recommend possible practical solutions that can be followed to address such barriers.
- The possibility of converting Elmergib University into a virtual university.

3- Related Works

In this part of the paper, as shown in Table 1 we give some experiences of Arabic countries to achievement and practical situations in which the concept of virtual university has been applied and works well. The papers recommended that is the issue of improving the quality of education can be addressed through the use of ICT. This approach will provide the opportunities for modernizing the education content, expanding the access to education, transferring the teaching sessions from traditional audiences to virtual ones and ensuring their flexibility, enhancing the learner's independent work and optimizing the teaching process. In order to realize the concept

of the application of distant technologies in education, it is important to make global changes in the education system as a whole and, first of all, the classic university model should be changed into the electronic university model.

Table 1: Related Works

Country name	References	The experiences
Saudi Arabia	[5]	The KAU university created a program that underwent rigorous course development and quality control to engage students more actively through asynchronous technologies- virtual classrooms for every face-to-face hour of every course- with synchronous components, using a learning management system developed in-house to integrate with all other university system.
Iraq	[6]	In this research, a simple prototype has been implemented to introduce some e-learning tools and technologies to the traditional education in the technical college of Mosul to let the students and faculty staffs get used on it. The prototype was implemented using Microsoft Windows 2003 server with other supporting servers and software.
Jordan	[7]	This study identified prevalent faculty perceptions toward the use of e-learning tools in higher education at one public Jordanian university, explored the level of integration of e-learning at the UJ, and examined the factors that are related to faculty's attitudes toward the use of e-learning tools, and probed what one academic institution can do to improve the utilization of e-learning at its campus.
Algeria	[8]	This research trying to assess the experience of E-learning site of the University of Ouargla and the extent of professors response and students to this modern method through the use of a tool for education, communication, knowledge sharing, publishing information and even conduct examinations on the line.
Syria	[9]	This study aims at exploring the impact of electronic services provided the Syrian Virtual University(SVU) on student satisfaction, the results indicated that the relationship between electronic services dimensions and student satisfaction at the SVU were positively significant except for the bulk SMS dimension.

4- Virtual University

In virtual university the term virtual is taken from computer science where ICT based experience transferred original experience in virtual reality, become popular in the 1990s [10].

"virtual university" emerged about 1995. Website of the International Council for Distance Education(ICDE) defines virtual institutions as "institutions which offer programs only by electronic means, and which are not open universities" .

Wikipedia offers more inclusive definitions: "A virtual university provides higher education programs through electronic media, typically the Internet".

Virtual universities are new types of education institutions that fully realize potential opportunities of ICT, offering educational programs with the application of only e-learning and distance learning technologies. Virtual universities do not have buildings, student dormitories, campus, cabinets for administrative staff, and lecture halls. It consists of collaborator, co-workers, trainers, educators, technicians, employed and interactively educated learners who live in remote distances, often beyond the national boundaries using modern ICT. This model has not been fully implemented yet. It faces serious difficulties, particularly the problem of virtual university accreditation, in other words, gaining public trust, obtaining the right to provide the diplomas and certificates and appropriate degree [11].

Generally there are some functions of virtual university. Firstly, create learning environment for collaboration in the development and use of emerging technologies that are needed to develop virtual education model, secondly, support services to students, which would include assessment of current skills and knowledge, advice regarding academic plans, quality, and record of learning, finally, improve the management in the organization, planning, design and delivery to learners of programs, curricula and courses that are pertinent to the human resource development [12].

4.1 Major differences between the traditional university and virtual university

in view of the promising hopes of virtual universities in development of education, the following is a comparison between virtual universities and traditional universities, with mentioning the advantages and disadvantages of each (Table 2) [13] [14] [15] [16] [17].

Table 2. Major differences between the traditional university and virtual university

Traditional university	Virtual university
A primary source of learning is the teacher.	A facilitator of educational resources is the teacher.
The teacher guides the student.	The student learns self.
Learning independently and not groups.	Learning collectively through student interaction.
The same thing is taught to all students.	Every student learns what he wants according to his circumstances.
Time and place limitations.	Time and space flexibility.
Being know to the teachers and the students and cultivation of social community.	There's no face to face interaction between teacher and student.
More expensive to provide.	Cost-effective for learners.
Immediate feedback.	Lack of immediate feedback in asynchronous e-learning.
Result by test mark.	Result by electronic work file
Educational environment is school .	Providing higher education opportunity for anyone in different major and prestigious

	universities and potentially available to global audience.
Subject matters are size limited & selected.	Variety of educational techniques in e-learning.
Educational interaction is direct, face to face.	There are technological limitations and when one is offline or the system fails, it cannot be used.
Supervising of the students is periodically face to face , limited in time.	Supervising of the students is difficult.
Motivating students.	Students are more motivated to learn(due to their self-centeredness).

4.2 Importance of Virtual University

The importance of virtual university can be derived in the following points:

- Promotes formal and informal learning.
- Enhance capacity of distance learning.
- Transfers the form of distance education system into new digitalized form.
- Provides educational opportunities to students who would otherwise not have them.
- Provides education to increasing number of students.
- Smash boundaries of culture, distance and nations.
- Rubs the blueprints of mortar institutes and stops degradation of efficiency de to building bricks.
- Boosts education capacity with limited costs.
- Give the chance to equip with knowledge to working adults.

4.3 Tools of virtual university

The virtual universities system is built to suit all user's needs so that they can acquire knowledge. The system often includes student-teacher communication tools, test tools, and other create the main tools bellow:

A. Content Management System(CMS):

The main content management of the university on internet is interested in the means of delivering the content of the courses [18].

B. Learning Management System (LMS):

This system often includes the management of courses, online discussions, courses and homework [19].

C. Human Resources Management System

strategies that contribute to the implement it is the human resources management of the University"s objectives [20].

D. Library Management System :

This system contains a large amount of scientific resources and data that the student needs during his study at the university[21].

E. financial Accounts Management System:

The purpose of this system is to manage the salaries of the employees of the university of teaching and administrative staff in addition to managing the tuition fees for students and all accounting matters related to the needs of the university equipment [22].

F. Exam Management System :

This involves the use of ICT in the presentation and processing of evaluation materials, the e-assessment includes exams with multiple choice questions constructed on machine grading this technology provides a more intelligent access to applications and users [23].

5- Research Methodology

This research sought to identify the main challenges in e-learning and convert the Libyan universities to virtual universities. Generally, responses were classified into either external or internal factors. The former is related to educational institutions or current circumstances in Libya. The latter, on the other hand, is associated to intrinsic features of users themselves. Table 3 depicts these two categories. The interesting result is that both academic staff and students highlighted similar challenges. This suggests that decision-makers, leaders, and e-learning administrators should exert great effort to address them.

Table 3: *External and Internal Factors*

	N of lecturers= 8	N of students= 60
External Challenges	%	%
Low internet bandwidth	22.97	62.5
Insufficient financial support	18.91	25
Inadequate training programs	54.05	62.5
Lack of technical support	20.27	25
Frequent electricity shortage	6.75	12.5
Internal Challenges	%	%
E-learning literacy	31.08	50
Lack of awareness, interest, and motivation	47.29	75

5.1 External Challenges

External challenges include technical implementation of e-learning. Looking at the results in Table 3, many lecturers (22.97%) and most students (62.5%) agreed that low internet bandwidth is one of the issues hindering successful application of e-learning in Libya. Users need a long time to upload or download, for example, pictorial learning materials such as videos, and graphs. It is noteworthy that students have identified this issue more than lecturers because they may not be able to afford.

Another hindrance towards e-learning application in accordance with lecturers (18.91%) and students (25%) is the shortage of funding. However, more attention should be given for the requirements of e-learning integration.

In addition, Table 3 shows that most lecturers (54.05%) and students (62.5%) pointed out that intensive training programs are required in order to enhance user skills towards computer and e-

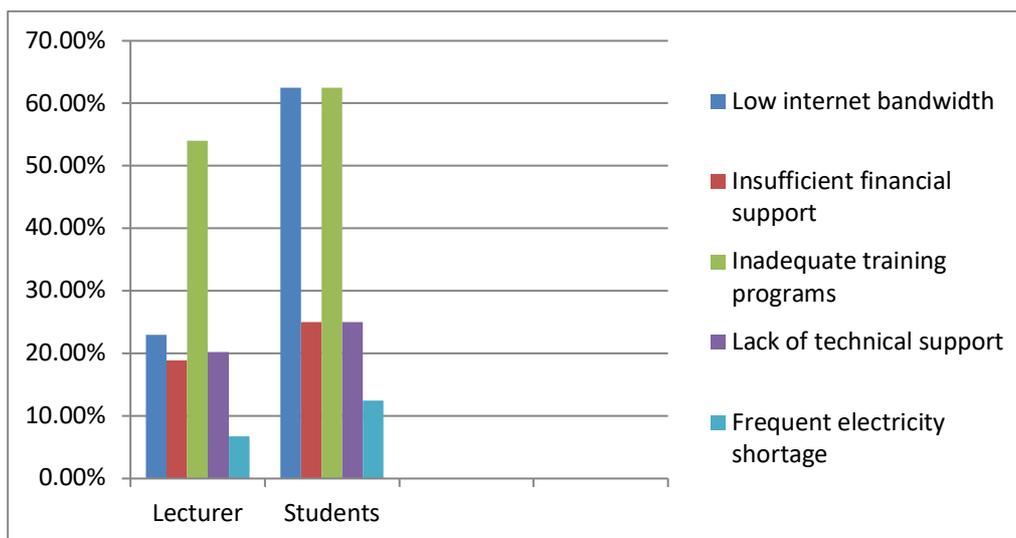
learning technologies. Undoubtedly, lack of individual skills can negatively affect their intention to use e-learning.

Successful implementation of any technology cannot be achieved without professional technical staff to set up and maintain it. As demonstrated in Table 3, instructors (20.27%) and students (25%) found that there is a lack of technical support to address issues that users may face. This means that preparation of professional staff should be a first step before launching a system (we mean here convert the traditional education system to virtual education system). As such, any failing in the system can be directly tackled to prevent user disruption.

Furthermore, results in Table 3 also demonstrated that few lecturers (6.75%) and students (12.5%) considered frequent electricity shortage as a factor that negatively influenced successful e-learning implementation in Libya. On the other hand, people use many alternative methods to get their minimum requirements of electricity. This may explain why only a few participants have mentioned this as a challenge that affects the application of e-learning.

Looking at Figure 1, we see that there are obstacles to activating the online educational system and transforming Libyan universities into virtual universities. But with the current state of students staying home due to the Corona pandemic, we used social media and a Google platform to give lectures. And this step is considered relatively successful in order to save time and not to miss the semester specifically for graduate students.

Figure 1: External Challenges



5.2 Internal Challenges

Internal challenges are more relevant to user willingness to move from traditional to more advanced teaching and learning approaches. Additionally, some users believe that e-learning means delivering learning content electronically by, for example, computer, iPhone, and mobile devices, whether offline or online. E-learning, however, entails a much wider concept than this superficial perspective. It includes exploiting the advantages of e-learning malleability to integrate pedagogical and learning theories [25].

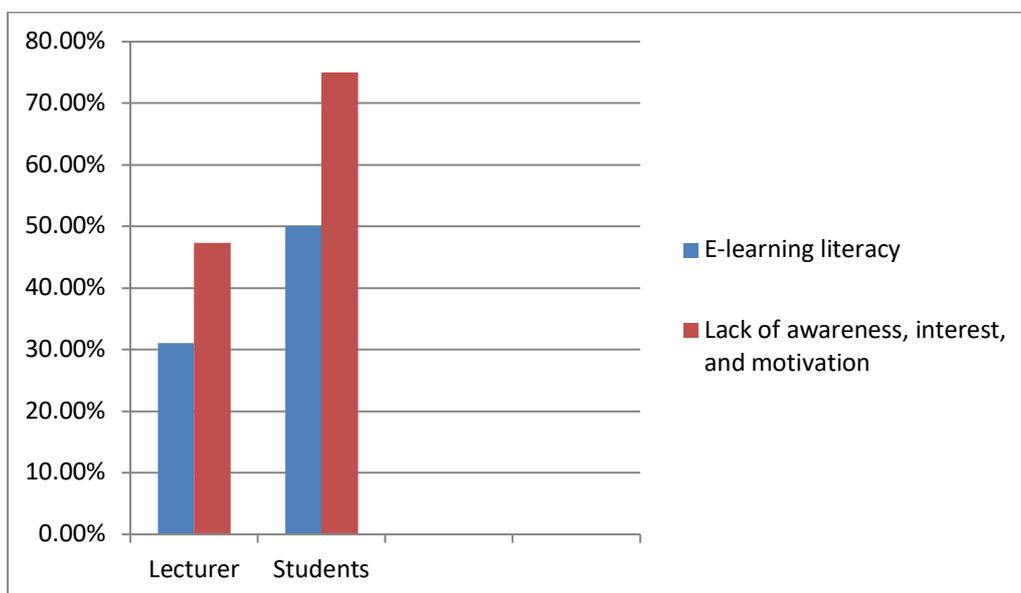
The results in Table 3 exhibit that many lecturers (31.08%) and half of students (50%) consider e-learning illiteracy to be hindering the effective integration of e-learning. Unsurprisingly, modern instructional technologies have recently been introduced to the Libyan Higher Education context. Therefore, older teaching members and those from humanitarian fields are unfamiliar with such technologies.

Table 3 also demonstrates that both staff (47.29%) and students (75%) pinpoint unwillingness, disinterest, and demotivation as other major challenges that hinder e-learning uptake. Some users resist any new experience. Respondents mentioned several aspects that could lead to such a negative attitude. For instance, the preparation of e-learning content requires more time than the traditional approach; e-learning will reduce teacher roles in educational practice; and the advantage of e-learning is limited to uploading electronic lectures only. Based on such perspectives, they have not found a difference between e-learning and other communication technologies such as email or even distributing written lectures in classrooms.

Accordingly, huge effort should be directed towards extending their academic understanding about the potential pedagogical impacts of e-learning on different learning aspects such as learner performance, satisfaction, and engagement. Academic staff should have an obligation to implement e-learning in order to fulfill the goals of all students more easily and enjoyably.

As shown in figure 2, a budget for establishing an integrated e-learning should be prioritized, as well as other necessary requirements. Also, a systematic strategy should be followed to implement e-learning. Since it is a new experience for Libyan universities.

Figure 2: Internal Challenges



6- Conclusions

The present research aimed to cast some light on major challenges that hinder effective application of use e-learning in virtual universities, and How effective is the transformation of Libyan higher education into a virtual education under realistic conditions. It was suggested to use e-learning methods and convert higher education in Libya into virtual education due to the Corona pandemic. Findings exhibited that the use of e-learning in Libya is still at a very early stage. In addition, it was illustrated that most public universities have started limited attempts to apply either custom or open source e-learning applications, but the majority of academic staff still do not have enough knowledge and skills to use it properly. This investigation confirmed the findings of existing studies about obstacles to e-learning integration in Libya. Causes that led Libya to be later to implement e-learning may encompass the previous isolation of staff members from the scientific world, the delay of introducing e-learning technologies to Higher Education, and the continuous national instability and insecurity.

Pertaining to the present investigation, five external and two internal challenges were demonstrated as a result of the quantitative and qualitative analysis. The main barriers comprised, but were not limited to, low internet bandwidth, insufficient financial support, inadequate training programs, lack of technical support, frequent electricity shortage, e-learning illiteracy, and lack of awareness, interest, and motivation toward e-learning technology. The key issues as identified by the majority of participants were lack of training programs. Based on the discussed barriers, many recommendations were provided. We hope that following them can lead to effective e-learning integration in Libyan universities. It is hoped that in the near future we will witness evolution in the e-learning sector as in other neighboring countries. And similar to the results of the questionnaire, it became clear that transforming the computer department into a virtual department requires the spread of the Internet among students and professors, increasing their knowledge and their use of electronic learning methods.

References

- [1] G. Georgieva, G. Todorov, and A. Smrikarov, "A model of a Virtual University some problems during its development," in Proceedings of the 4th international conference on Computer systems and technologies: e-Learning, 2003.
- [2] A. Elzawi and S. Wade, "Barriers to ICT adoption in quality of engineering research in Libya: how to bridge the digital divide?," ed: University of Huddersfield, 2012.
- [3] M. Kalochristianakis, S. Panagiotakis, and K. Vassilakis, "Design considerations for developing e-University solutions based on open source software," in 2012 International Conference on Telecommunications and Multimedia (TEMU), 2012, pp. 190-195.
- [4] A. Rose and S. Kadvekar, "ICT (information and communication technologies) adoption model for educational institutions," Journal of Commerce and Management Thought, vol. 6, pp. 558-570, 2015.
- [5] H. A. Al-Nuaim, "The use of virtual classrooms in e-learning: A case study in King Abdulaziz University, Saudi Arabia," E-Learning and Digital Media, vol. 9, pp. 211-222, 2012.
- [6] M. M. Shaker, A. Khidhir, and Y. A. Al-Rizzo, "Design and implementation of a virtual university prototype technology requirements and evaluation," Proc. CSIT, vol. 3, pp. 560-568, 2006.
- [7] M. Al-Shboul, "The Level of E-Learning Integration at the University of Jordan: Challenges and Opportunities," International Education Studies, vol. 6, pp. 93-113, 2013.
- [8] د. ع. الغني, "تقييم فعالية التعليم الافتراضي في الجامعة الجزائرية: دراسة حالة موقع التعليم الافتراضي and ر. خالد . ع. الغني, "تقييم فعالية التعليم الافتراضي في الجامعة الجزائرية: دراسة حالة موقع التعليم الافتراضي بجامعة ورقلة (www. elearn. univ-ouargla. dz)= Evaluate the Effectiveness of E-Learning in the Algerian University: Case Study of E-Learning Webe Site at the University of Ouargla (www. elearn. univ-ouargla. dz)," Algerian Review of Economic Development, vol. 339, pp. 1-12, 2015.
- [9] S. G. Al Azmeh, "The relationship between e-learning service and student satisfaction a case study at the Syrian Virtual University (SVU)," Business, Management and Education, vol. 17, pp. 49-71, 2019.
- [10] G. M. Farrell, The Development of Virtual Education: A Global Perspective. A Study of Current Trends in the Virtual Delivery of Education: ERIC, 1999.
- [11] M. H. Mammadova and H. A. Gasimov, "E-university: conceptual, technological and architectural approaches," Problems of Information Technology, pp. 51-62, 2017.
- [12] A. A. Sejzi, B. Aris, and N. Yahya, "The phenomenon of virtual university in new age: trends and changes," Procedia-Social and Behavioral Sciences, vol. 56, pp. 565-572, 2012.
- [13] R. H. Saloom, A. S. Elameer, and A. S. Jalal, "A New Style of an Academic Institutions in Iraq: e-University," in 2018 1st Annual International Conference on Information and Sciences (AiCIS), 2018, pp. 187-192.
- [14] M. Mirzakhani, H. Ashrafzadeh, and A. Ashrafzadeh, "The virtual university: Advantages and disadvantages," in 2010 4th International Conference on Distance Learning and Education, 2010, pp. 32-36.
- [15] A. A. Mahmood and Y. Wan, "The Disadvantage of ICT in Higher Education Students' Study and Life through Lanzhou University Postgraduate Students."

- [16] M. Fooladvand and M. H. Yarmohammadian, "A comparative study between virtual and traditional approaches in higher education in Iran," *Procedia-Social and Behavioral Sciences*, vol. 28, pp. 646-650, 2011.
- [17] D. Zhang, J. L. Zhao, L. Zhou, and J. F. Nunamaker Jr, "Can e-learning replace classroom learning?," *Communications of the ACM*, vol. 47, pp. 75-79, 2004.
- [18] T. El-Sakka, "Advances in Information Technology for the Holy Quran and Its Sciences (32519)," in 2013 Taibah University International Conference on, 2013.
- [19] A. I. Albarrak, H. A. Aboalsamh, and M. Abouzahra, "Evaluating learning management systems for University medical education," in 2010 International conference on education and management technology, 2010, pp. 672-677.
- [20] D. Ruppert, *Statistics and data analysis for financial engineering vol. 13*: Springer, 2011.
- [21] A. Calanducci, C. Cherubino, L. N. Ciuffo, M. Fargetta, and D. Scardaci, "A digital library management system for Grid," in 16th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE 2007), 2007, pp. 269-272.
- [22] C. Liu, L. Wei, R. Xu, and J. Jia, "Design of the financial accounting management system for universities," in *Proceedings of 2012 International Conference on Measurement, Information and Control*, 2012, pp. 318-321.
- [23] A. Ullah, H. Xiao, M. Lilley, and T. Barker, "Design, privacy and authentication of challenge questions in online examinations," in 2013 IEEE Conference on e-Learning, e-Management and e-Services, 2013, pp. 46-50.
- [24] M. M. A. Adrah, G. A. Elmarash, and E. E. Eljadi, "The possibility of using cloud computing technology in E-learning at Elmergib University," in *International Conference on Technical Sciences (ICST2019)*, 2019, p. 04.
- [25] A. Al-Azawei, P. Parslow, and K. Lundqvist, "Barriers and opportunities of e-learning implementation in Iraq: A case of public universities," *The International Review of Research in Open and Distributed Learning*, vol. 17, 2016.

Performance Evaluation of LSB Technique into visual Objects based on Steganography

Bashir M. Mujber^{1*}, Ali Ahmad Milad^{1,2}

¹mojbar_ca12@yahoo.com, ²alimilad@elmergib.edu.ly

¹Department of Computer, Faculty of Education, Elmergib University

²Faculty of Information Technology, Elmergib University

ABSTRACT

Steganography is an important field of research in recent years to embed a range of data, it is the science that hide information in cover medium without being accompanied by any effect or distortion in that medium. Nevertheless, most of the modern researches focus on hiding information in image according to its popularity. This paper studies the Least Significant Bit (LSB) (1-LSB, 2-LSB and 3-LSB) with one or RGB color based Steganography, The LSB algorithm has experimented on Bitmap 24 bits format as cover image to generate a stego images. The aim of this paper is to carry out various types of image steganography technique for purpose of identifying various principles of image steganography in terms of visual effectiveness and efficiency. However, the algorithm that has been chosen for this purpose is discussed in details in this paper. The visual effectivity of the stego were measured by comparing the histograms of the stego and cover images. In this study we used Mean Squared Error MSE calculation and discussed the implementation of this algorithm in detail. The results from experiments prove that algorithm is not affected by different visual characteristics of the cover images in so doing, the perceptual distortion to the cover image is nearly negligible and unlikely to be detected by simple visual inspection.

Keywords:

Steganography,
LSB,
RGB.

1- Introduction

Steganography means verbatim which covers the procedure of writing, and is studied as one of the most crucial communication arts [1]. It contains two words, the first word is Steganos means "coverage" and the other word graph in which means "writing" in Greek. The main idea of it is to conceal the communication procedure without using encryption algorithms that make the process of communication non-understandable except for those who have the right keys [2]. Steganography inside the image is the process of developing a hidden message within the same or given image so that nobody can know what the message is or should not be able to detect its presence.

The term Steganography means "cover writing" while Cryptography means "concealed writing" as showed in figure 1 and 2. Cryptography is the procedure of sending the message in distinctive forms so that only the involved people can demystify the message and then read it [3]. The message that is sent without encryption is called the plaintext whereas the enciphered message is labeled as ciphertext. The procedure for changing the plaintext to a ciphertext is called encryption while the reverse operation (i.e., alteration of encrypted text to a plaintext) is called decryption. Encryption protects the contents of the message by encrypted it during transmission of data from the sender to the receiver. However, when the receiver receives the message he decrypts it and ensures from its integrity Steganography works to conceal the message in plain view inside the data instead of encrypting it and does not require sending confidentially [4].

At the present time, steganography works on digital media as cover image and embedding digital media as secret message, the example for the used digital media are .wav, .gif, .bmp, .jpeg, .txt, .mp3, and .doc. Steganography is thought to be one of the most essential techniques to the future of the Internet in terms of privacy and security. The importance of steganography is highlighted because of the weakness in the encryption process and the desire to get the secrecy in the open systems. Lots of governments have made laws in an effort to decrease the strength of encryption systems or completely prevented them; this may create unfortunately weak and breakable encryption algorithms in the Internet community [5]. Hence, the of steganography appears more than ever significant where the hidden message inside another file can be detected and read only by the involved entities or individuals and no one has the ability to read the message even with the knowledge of its existence. However, encryption and steganography do not provide the desired privacy and confidentiality, but that can be accomplished by utilizing both technologies to provide acceptable limits of privacy and confidentiality of anyone connected to the open systems [5].

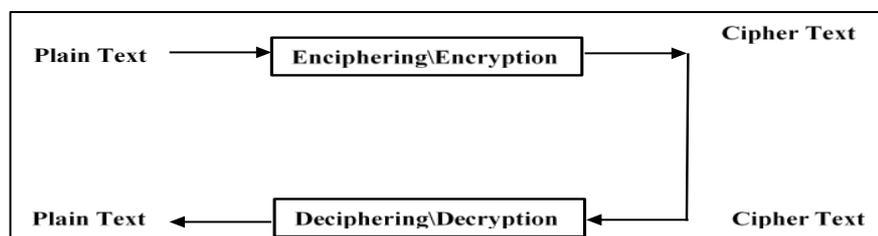


Figure 1:

Cryptography

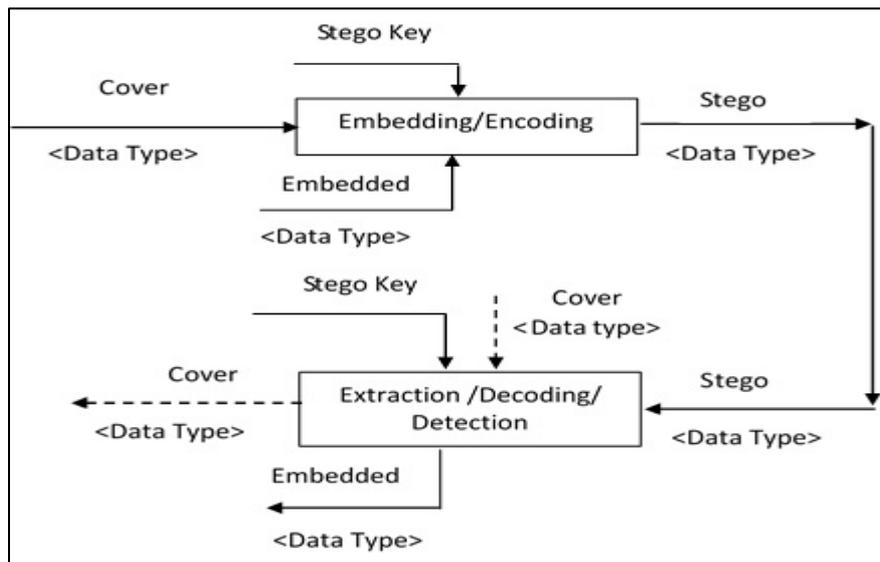


Figure 2: *Steganography*

There are many techniques which can be applied for concealing information within the multimedia objects, with audio files, image or video object cover depending on the object type given below. Here we will discuss various methods and techniques that are frequently applied in text, image, audio or video of steganography.

- Text: The procedure of hiding a text inside text can be accomplished by changing the text format or by changing certain properties in text elements such as letters. The goal of designing coding methods is to initiate changes that are unbreakable even if they contain noise [6], [7].

- Audio: The process of concealing information is computer-based audio steganography system, secret messages are embedded in digital sound [8]. Here a message is hidden in the mode of a simple change in the binary sequence of an audio file.
- Image: To embed specific messages or text inside an image that requires encrypting all message bits or inserting noisy areas which brings less attention so that we could hide our specific message within the region of huge natural colors variations. It is equally possible that the hidden message is randomly distributed to various sections of the image. There are numerous methods that are utilized for the reason of hiding information within the image [9].
- Video: The video file is a collection of concealed images with added voices to it, here the intended information is hidden in form of displayed images using image concealing techniques such as DCT [10].

There is a system applied to the colored image by using LSB method. The system compares the value for each ASCII character with palette location of the image if the value is equal to the palette value that is compensated in another location [11]. A system designed an efficient system to check the image as if to see it contains a secret message or not. Through the system has extracted the secret message, which may be in the form of text or image, and when it fails to extract the hidden message or keep it from traffic, the system destroys the hidden text [12]. Fridrich J. and Goljan M. In 2003 described steganography where they took large loads of gray scale images and added a few amplitude of the noise to the image pixels with certain specifications. Susceptibility noise distribution is arbitrary so the parties related to it have the possibility to hide the noise in conformity with the output noise of the devices themselves [13].

2- System Design

Least significant bit insertion (LSB): LSB is substitution method that uses specific k LSBs in each pixel to hide a secret message. It is thought as one of the easiest ways to conceal a secret image in a specific image. Nevertheless, it is not difficult to disclose a stego-image that is related by using the LSB insertion technique [14].

We have selected the image carefully for that it is complicated and of multi characteristics, it is expected here that it has an effect in power and capacity, all images have the same size 300*200 on two formats BMP, JPEG, as given in the following: We examined if the characteristics of the cover image has impact on the concealment efficiency of stego.

In figure 3 (A) Group_of_student: here the dominant color is the blonde color regarding the hair, but the color of the jackets is black which makes it also a dominant color in this picture. (B) Living_room_home_house, it is obvious that the dominant color is green color of the trees and the field which has a light green color. (C) Spring_sunshine_may the dominant color is pink color, although it is complicated to tell the exact color.

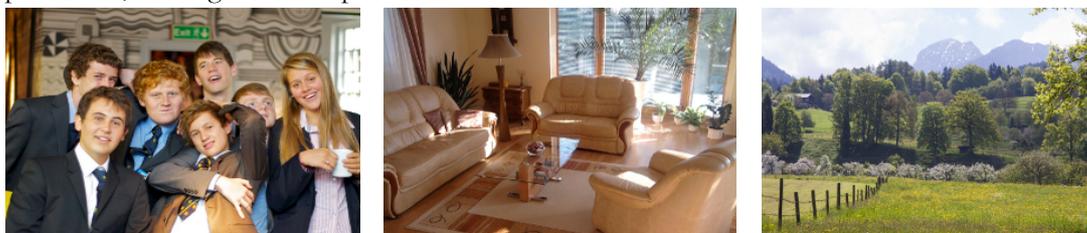


Figure 3: Cover Image Selection A, B and C

In this paper the text file was chosen in making our experiment and the reason for that is the objective which is, to use the maximum insertion capacity of every method which is not expected in few of them and also depends on formats and scenes of chosen image, that leads to use a

concealed message that is different in size for every experiment. We used a text message since it is not important for our study whether the message is text, voice or another image, it is just a data, a collection of 0's and 1's... Therefore, with regard to, 'Vocabulary as a reflection of life wisdom [15] we have chosen the below mentioned paragraph given in bold, as the concealed message for our experiment, and it is kept in a text file to be used with the suggested techniques. Emphasize the repetition of the same text to fill up the total capacity of the cover image which differs depending on the steganography technique used.

“Comprehend the environment in terms of discrete objects and events as a result we can say that the world consists of a multitude of uniquely defined objects and events They can be further organized into classes as groupings based on the criterion of similarity or shared characteristics The mental construction which comprises the criterion of similarity and which subsequently enables the classification of objects is called the concept In other words it stands for or represents a common set of attributes of an object or event” [15].

The advantage of using LSB steganography is not needed to a complex calculation for the purpose of the data hiding in the image as well as it is considered one of the most prevalent methods to hide data in the spatial domain of the image. These methods are simple to implement, however, over time the tool Internet offers the possibility to identify and extract the information in any bit plane. Therefore there is a need to make a program which permits us to withhold information in an image using proper to the purpose of the research, so we develop program to adjust some variation of parameters and getting the final result of the application as stego image file. Based on the inherent characteristics of the human eye and images in this paper.

LSB in Single Color

- **One LSB:** Basically it consists of changing the least important bit of the color bands (R, G or B) of an image cell array in consideration to enter the message utilizing the space of one bit per pixel to store the message. Besides the already mentioned low computational cost that is characteristic of LSBs algorithms, specifically in this regard we can quote high fidelity between the original image and the stego-image. The change of just one bit ensures a great difficulty to note the "naked eye" the contrast between them in Figure 4. Here, one bit stored in accordance to the method described in this Section.

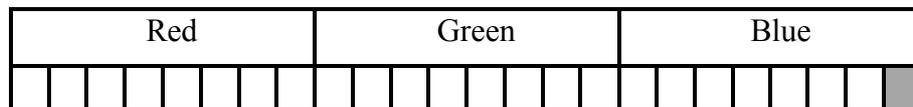


Figure 4: One bit in Blue Color

- **Two LSBs:** A slight variation of the one LSB, the only difference, as its name suggests, is to use two least significant bits of the color bands (R, G or B) separately in the image. The subtle difference is in the ability to the hide message stored twice, and the change in the figure is also a bit sharper, but still generally imperceptible to the "naked eye", as we can observe in Figure 5, where two bits of message stored in it.

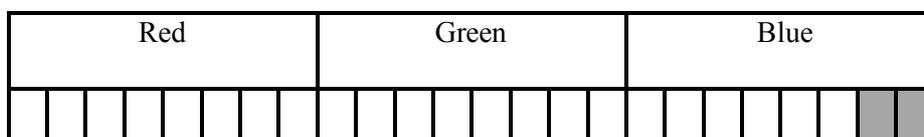


Figure 5: Two bits in Blue Color

- **Three LSBs:** In general, the three LSB bits boost the quantity of bits to be used. This causes a boost in storage capacity over a stego-image quite modified from the original. To illustrate, Figure 6 presents an example of inserted three bit in one color in pixel.

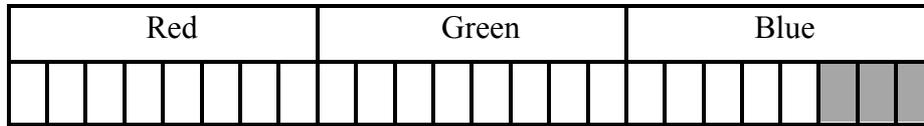


Figure 6: Three bits in Blue Color

- **LSB in Three Color (RGB):**

The least important bit to be modified is switched between bands cyclically. For example, a message to store only four bits in the image the algorithm record the first bit in the R band, the second bit in the G band, and B in the third quarter turn to band R in next pixel. Thus, a switchover is produced in order pixels to be modified to achieve thus imposing a difficulty more for steganalysis to detect and get the message correctly. Figure 7 shows the one, two and three bits of secret data using the cyclical LSB algorithm.

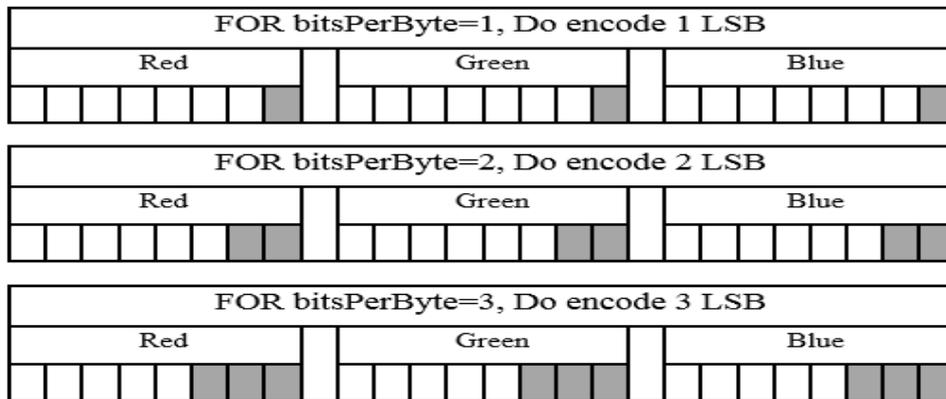


Figure 7: LSB Three Colors Embedded

The algorithm has been tested and validate using MATLAB and each allows the variation of Parameters such as number of bits or color component to modify. The techniques were developed initially in the shape of scripts, wherein each was implementing a method (LSB bit 1, etc.). Subsequently, aiming facilitating experimentation and obtaining the results, a graphical interface was created (GUI) using the guide Matlab. This interface provides a facility to choose the parameters of each experiment. In Figure 8 one can notice the GUI.

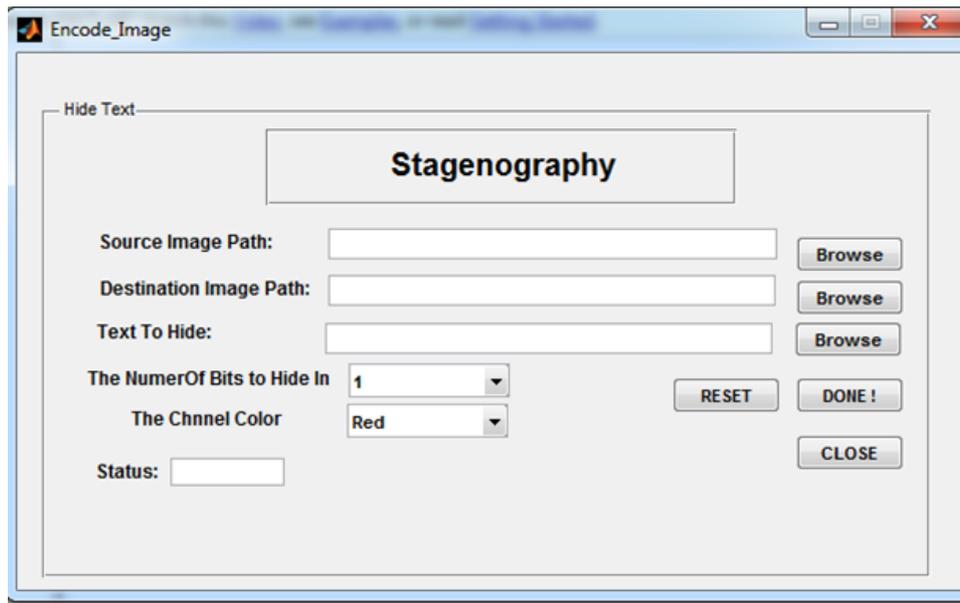


Figure 8: LSB Method Interface

With regard to the parameters of each algorithm, we varied the use of each one of them in regards to the number of LSB bits to be modified in the image in consideration to making it possible the analysis of the impact which this change does not only through use of the similarity indices, but also by our visual system. In choice of which color component (R, G or B) to modify we opted for, where possible, modify each of color separately and all of them because of the difficulty inherent to the human being to visualize this component. Therefore, we sought to make changes in the images less possibly noticeable to our eyes.

In this case, the quality criteria are either subjective (determined by the human eye), or based on image characteristics: shape and parameters of the brightness distribution, the width of the spatial spectrum, etc. Moreover, objective criteria used in assessing the quality of the images are criteria to get a computed image characteristic difference signal between two images: a real and some ideal, or it may be the original and transformed. They are called difference metric distortion. Using these criteria it allows to evaluate the quantitative changes of brightness levels of image distortion. When creating transformations (filtering, data compression, etc.) that is substantially the quality of the conversion means - algorithm or system. It is extremely necessary in the construction of algorithms and image processing systems and algorithms for evaluating quality.

The histogram is a significant statistical characteristic of data. In various image processing applications, the histogram is generally utilized as the basic characteristic to present the distribution of the intensity, color, and texture parameters of images. As a statistical feature, the histogram is equally not sensitive to translation and rotation of objects. Meanwhile, it is a standardized and compressed data storage type that can save much space. Because of these advantages and along with the same, the histogram is used mostly in image segmentation, registration, tracking, and especially in the image retrieval field that involves a large amount of data. The following formula calculates the histogram measurement where its parameters are k is the maximum pixel value in an image, m is the pixel value.

$$n_i = \sum_{i=1}^k m_i$$

The most popular distance tools for analysis of the level of distortions that are introduced into the cover image at the time to hide the information, therefore, MSE can be utilized to examine

the quality of the stego-images as well. MSE is the ratio of sum of the square of the differences in the pixel values between the corresponding pixels of the two images over total pixel number. MSE can be calculated if two images dimensions are equal. If two images are identical MSE value will be 0. Next formula shows how to calculate MSE value. X and Y are images with same dimensions. m and n are the dimensions of images [16].

$$MSE(X, Y) = \frac{1}{(m \cdot n)} \sum_{i=1}^m \sum_{j=1}^n [X(i, j) - Y(i, j)]^2$$

3- Results and Discussion

Here, the results of the histograms were collected for the stego images that have one color modulation in one form, and RGB colors stego images. The one least significant bit result is shown in the figures 9 and 10, two least significant bit result in the figures 11 and 12, and three least significant bit result in the figures 13 and 14.

One LSB:

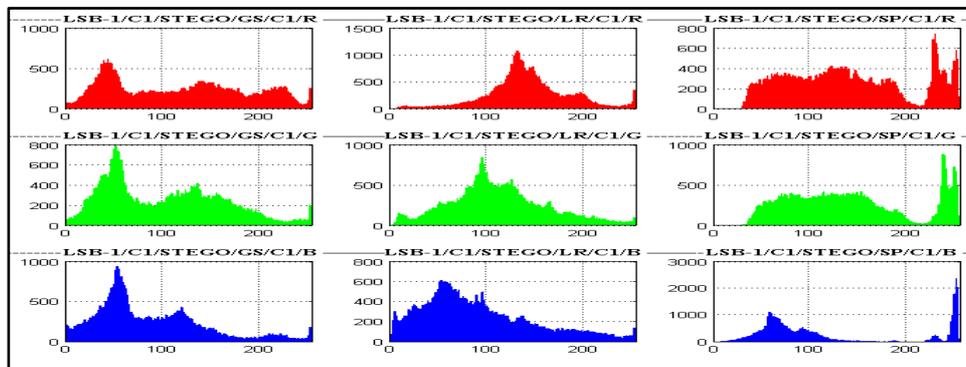


Figure 9: Histogram of One LSB One Color

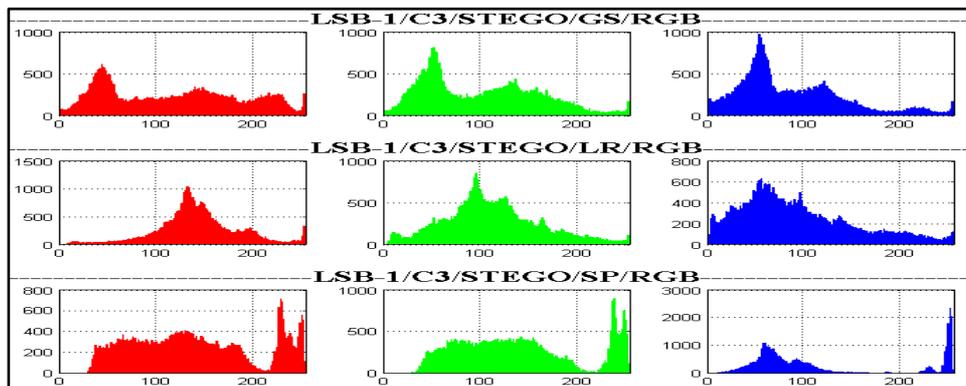


Figure 10: Histogram of One LSB RGB Color

Two LSBs:

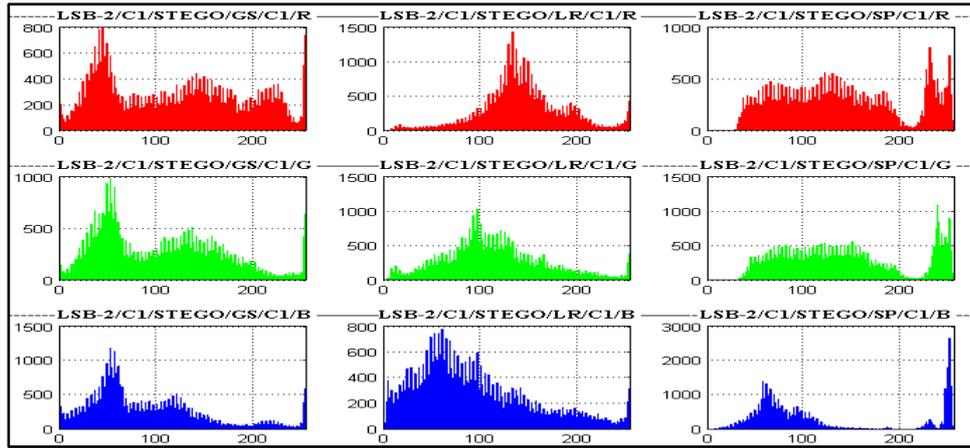


Figure 11: Histogram of Two LSB One Color

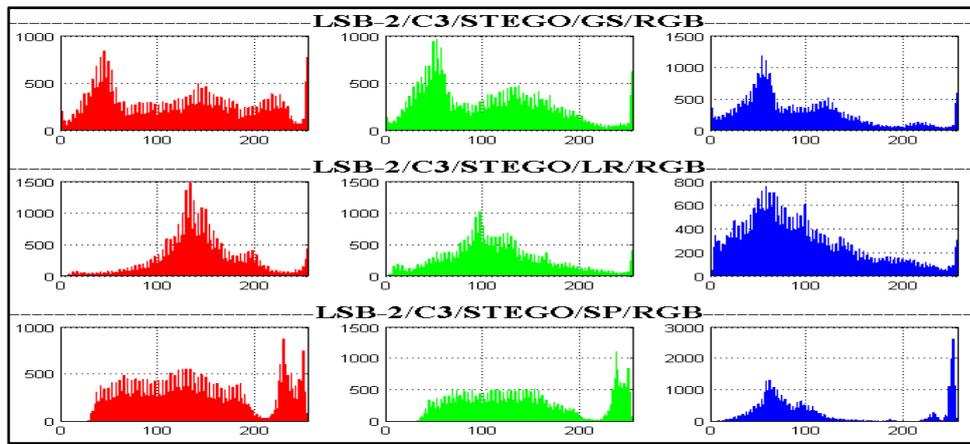


Figure 12: Histogram of Two LSB RGB color

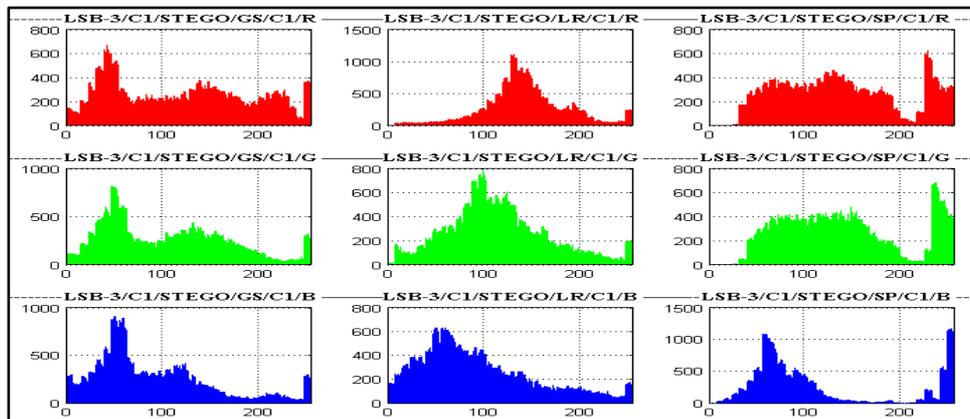


Figure 13: Histogram of Three LSB One Color

Three LSBs:

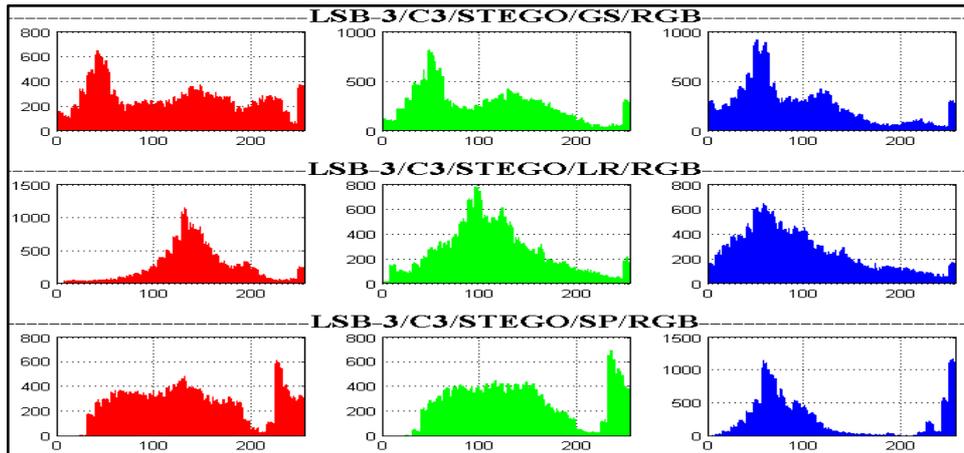


Figure 14: Histogram of Three LSB RGB color

In Table 1, we have presented the results of MSE parameter that have been obtained using LSB algorithm. Through this table, one can notice that the error values of a single color that is experimented on three images (as can be seen) are increasing with the increasing number of bits insertion, while, the increasing error values for all the stego-images are considered as slight which is between 0.0881% to 1.9801 in the average. However we found it imperceptible with the human eye. Nevertheless, the RGB error rate values as described in Table 2 are higher than the error rate for a single color and it is also considered as a small percentage and cannot be observed by the Human Visual System as well.

Table 1: MSE Results of LSB Method with One Color

LSB	Group_of_student			Living_room_home_house			Spring_sunshine_may		
	Red	Green	Blue	Red	Green	Blue	Red	Green	Blue
One	0.0903 %	0.0888 %	0.0898% %	0.0891 %	0.0896 %	0.0901 %	0.0886 %	0.0881 %	0.0884 %
Two	0.4447 %	0.4459 %	0.04437 %	0.4346 %	0.4458 %	0.4431 %	0.4454 %	0.4356 %	0.4345 %
Three	1.9801 %	1.9563 %	1.9532% %	1.9009 %	1.938% %	1.9334 %	1.8638 %	1.8985 %	1.7408 %

Table 2: MSE Results of LSB Method with RGB Color

LSB	Group_of_student	Living _room_home_house	Spring_sunshine_may
Bits	RGB	RGB	RGB
One	0.2686%	0.2679%	0.2650%
Two	1.3345%	1.3231%	1.3194%
Three	5.8868%	5.7983%	5.4970%

4- Conclusions

Steganography is an important tool in hiding the information either during transfer or in storage. It was discovered a long time ago as a cryptographic method. Since then, it has been used to hide large amounts of data securely and reliably to some extent. The technique which has been used in this research namely LSB, which is easy to use for concealments and extraction operations.

According to the results that were obtained in MSE analysis and by visual comparisons on histograms of the original and the stego-images, one can perceive the distinctive effects of different algorithms and also varying influence of the images with different visual characteristics. However, by looking to the results of MSE parameter that we have obtained, it is clear that the information hiding efficiency of LSB algorithm is not effected by different visual characteristics of the cover images which were in bitmap format if we use 1-LSB, 2-LSB and 3-LSB.

References

- [1]. N. F. Johnson, and S. Jajodia, "Exploring steganography: Seeing the unseen," Computer, vol. 31, no. 2, pp. 26-34, 1998. <https://ieeexplore.ieee.org/abstract/document/4655281/>
- [2]. N. Johnson, "Digital Watermarking and Steganography: Fundamentals and Techniques," ACM Digital Library, 2007. <https://dl.acm.org/doi/book/10.5555/1557373>.
- [3]. A. Milad, Z. Muda, Z. A. B. M. Noh, and M. A. Algaet, "Comparative study of performance in cryptography algorithms (Blowfish and Skipjack)," Journal of Computer Science, vol. 8, no. 7, pp. 91, 2012. <https://thescipub.com/abstract/jcssp.2012.1191.1197>
- [4]. S. M. Thampi, "Information hiding techniques: a tutorial review," arXiv preprint arXiv:0802.3746, 2008. <https://arxiv.org/abs/0802.3746>.
- [5]. Dunbar, "A detailed look at Steganographic Techniques and their use in an Open-Systems Environment," Sans Institute, vol. 1, 2002. <https://www.sans.org/webcasts/security-compliance-hypergrowth-startup-113570>.
- [6]. J. T. Brassil, S. Low, N. F. Maxemchuk, and L. O'Gorman, "Electronic marking and identification techniques to discourage document copying," IEEE Journal on Selected Areas in Communications, vol. 13, no. 8, pp. 1495-1504, 1995. <https://ieeexplore.ieee.org/abstract/document/464718/>.
- [7]. Singh, P. K. Singh, and K. Saroha, "A survey on text based steganography." Proceedings of the 3rd National Conference; INDIACOM-2009 pp. 332-335. 2009.
- [8]. N. F. Johnson, Z. Duric, and S. Jajodia, "Information Hiding: Steganography and Watermarking-Attacks and Countermeasures: Steganography and Watermarking: Attacks and Countermeasures." Springer Science & Business Media, 2001. www.wakp.nl.
- [9]. A. Singh, and S. J. Singh, "An overview of image steganography techniques," International Journal of Engineering and Computer Science, vol. 3, no. 07, 2014. <http://103.53.42.157/index.php/ijecs/article/view/1199>
- [10]. N.-D. Hoang, and Q.-L. Nguyen, "A novel method for asphalt pavement crack classification based on image processing and machine learning," Engineering with Computers, vol. 35, no. 2, pp. 487-498, 2019. <https://link.springer.com/article/10.1007/s00366-018-0611-9>.
- [11]. Y. Luo, J. Qin, X. Xiang, Y. Tan, Q. Liu, and L. Xiang, "Coverless real-time image information hiding based on image block matching and dense convolutional network," Journal of Real-Time Image Processing, vol. 17, no. 1, pp. 125-135, 2020. <https://link.springer.com/article/10.1007/s11554-019-00917-3>.
- [12]. S. A. Parah, J. A. Sheikh, J. A. Akhoun, N. A. Loan, and G. M. Bhat, "Information hiding in edges: A high capacity information hiding technique using hybrid edge detection," Multimedia Tools and Applications, vol. 77, no. 1, pp. 185-207, 2018. <https://link.springer.com/article/10.1007/s11042-016-4253-x>.
- [13]. A. Sharif, M. Mollaeefar, and M. Nazari, "A novel method for digital image steganography based on a new three-dimensional chaotic map," Multimedia Tools and Applications, vol. 76, no. 6, pp. 7849-7867, 2017. <https://link.springer.com/content/pdf/10.1007/s11042-016-3398-y.pdf>.
- [14]. K. Sahu, G. Swain, and E. S. Babu, "Digital image steganography using bit flipping," Cybernetics and Information Technologies, vol. 18, no. 1, pp. 69-80, 2018. <https://content.sciendo.com/view/journals/c`ait/18/1/article-p69.xml>.
- [15]. E. H. Rachmawanto, and C. A. Sari, "Secure image steganography algorithm based on dct with otp encryption," Journal of Applied Intelligent System, vol. 2, no. 1, pp. 1-11, 2017. <http://publikasi.dinus.ac.id/index.php/jais/article/view/1330>.
- [16]. Z. Wang, and A. C. Bovik, "Mean squared error: Love it or leave it? A new look at signal fidelity measures," IEEE signal processing magazine, vol. 26, no. 1, pp. 98-117, 2009. <https://ieeexplore.ieee.org/abstract/document/4775883/>.

Developing a mobile game app themed about Libyan culture using Unity engine

Khaled M. Ben Hamed^{1*}, Nahed F. Farah², Mohammed Binmiskeen³,
Mohamed Bashagha⁴

^{1*}k.benhamed@uot.edu.ly, ²n.farah@uot.edu.ly, ³m.binmiskeen@uot.edu.ly,

⁴m.bashagha@uot.edu.ly,

^{1,2,4} Department of Mobile Computing, Faculty of IT, University of Tripoli, Libya

³ Department Internet Technologies, Faculty of IT, University of Tripoli, Libya

ABSTRACT

This paper presents the design and implementation of an educational game App using Unity engine. The game aims to provide informative experience of Libyan traditions while keeping players entertained. Also, the game attempts to document Libyan fading traditions while being amusing and enjoyable. This game will be very first Libyan games to be launched into Google Play Store.

Keywords:

Mobile Computing,
Mobile Games,
Mobile App Development,
Unity Engine.

1. Introduction

With each passing day, the mobile game industry is witnessing alarming growth as the technology it runs on, evolves and matures[1]. In 2019, global mobile game revenue grew to \$68.5B and this number is expected to hit \$76.7B by the end of 2020[2]. This wide spread of mobile games provides an opportunity to spread information with it. Thus, making it one of the most efficient way for future education to help students to improve their technological, mathematical, scientific and engineering abilities[3].

Libya has a distinctive **history, traditions, and culture**[4]. A major problem facing Libya is the lack of cultural documentation and preservation of its historical treasures[5]. As games start to migrate more quickly and rapidly onto mobile devices, the culture and traditions of our ancestors face the risk of being forgotten and lost between the historical archives in museums forever. Of course, one mobile game app will not solve this problem. Nevertheless, it can only serve as a reminder of the impact these lost traditions have on our daily life style. In addition, to hopefully triggering the right people into taking care of this issue. This paper introduces the first mobile game themed about Libyan culture and traditions into the Google Play Store. The game integrates Iconic Libyan themes with game rules to create an educational and entertaining environment for the young youth and children to learn some of the oldest traditions, words and cultures, while also filling the space represented by Libya in the global game market.

2. Research, analysis and backgrounds

The very first stage of the project was to research and analyze the basic concepts and ideas that should be implemented within the game. This included researching successful mobile games apps, as well as different development processes and comparing them. The advantages and disadvantages of each game category should be taken into consideration. Additionally, software development tools benefit and constrains are very important to determine the approach of the development process. Also, the research included Libyan history, culture and traditions, while focusing on the

old city of Tripoli, and the traditional Libyan games, which are commonly played by the children in the streets of Libya. This allowed us to integrate a similar traditional vibe into the game.

From game comparisons research prospective, we choose the following games based on their popularity, success and creativity, these three attributes which we are looking forward to learn from and hopefully pass them into our game.

Super Mario Brothers [6]: it is the best-selling video game franchise of all time.

Created by: Shigeru Miyamoto

Considered an Ideal Model for good games design for many reasons such as:

Importance of the name: Miyamoto commented that if he had named Mario "Mr. Video", Mario likely would have "disappeared off the face of the Earth".

Importance of cultural reference: attributing Mario's nationality to his mustache.

Importance of colors: Due to the graphical limitations of arcade hardware at the time, Miyamoto clothed the character in red overalls and a blue shirt to contrast against each other and the background.

Importance of simplicity: red cap was added to avoid drawing the character's hairstyle. To make him appear human onscreen despite his small size.

Angry Birds: it became a very popular[7].

Created by Finnish company Rovio Entertainment

What makes it successful?

Original story idea, comical style, and simplicity.

Good advertising, easy to share, and low price.

Small psychological details (large impact).

Charming art style and cute characters.

From Libyan culture research prospective, before creating a game that represent a certain culture, we must have a deeper understanding of that culture, in this case Libyan games. Most commonly Libyan boys like to play games like *football*, *al-batch*, and *sab3a rashadat*. Libyan girls on the other hand play *niggi3a*, *fathy ya warda-sakri ya warda*, and *rope jumping*[8].

A Closer look at – **Battesh**. This game consists of small glass balls, also known as marbles, which come in different colors and variety of sizes. "Battesh" is one of the traditional Libyan games played by the children mostly around street alleys and schoolyards. Figure-1 shows the main character in our mobile game wearing traditional Libyan cloths and holding batch.



Figure 1: Batch

This traditional Libyan game seen less and less every day, as the word is shifting towards smart technology. Traditional Libyan food is one of the main aspects of culture[13], some of the most common Libyan food used in our mobile game are shown in Figure-2.

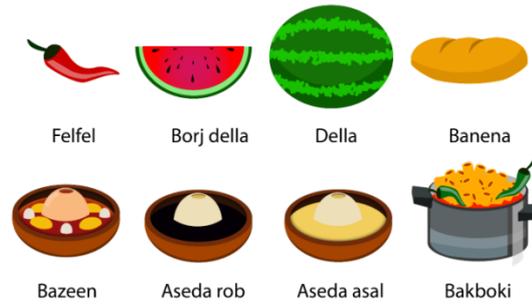


Figure 2: Food items design.

From implementation research prospective, there are several constrains to our project. All of which force us to make decisions and must be dealt with to succeed. These decisions are listed below:

- **Graphics 2d vs. 3d** [9]: It's not difficult to produce a simple 3d game, however, that will also defeat the objectives of this project. By doing 3d games, the quality of the work is decreased. 2d graphics allow greater focus on the design, the mechanics and the rules of the game, while also improving the quality of the work.
- **Vector vs. pixel** [10]: The pixel Art process is very difficult and time consuming as it deals with each and every pixel independently, forcing us to minimize the size of the picture and simplifying the details as much as possible, while trying to get our idea across to the user at the same time. Vector art clearly became the best choice.
- **Time constrains** are handled by efficient team work, organizational skills, planning ahead, making the right decisions and deciding priorities[11].
- **Budget constrains** limit the project team from the ability to obtain resources and might potentially limit the project scope.
- **Quality is restricted** by the specification of the product. The quality of the project is increased significantly by limiting the artwork themes, setting priorities and targeting specific devices and operating systems (in this case android devices).
- **Designers rules:** From the design prospective, the game integrates Iconic Libyan themes with game rules. there are a number of rules designers must follow to capture the attention and approval of a larger number of users. One of these rules is called “KISS” which stand for “keep it simple stupid”[12] by keeping the interface design simple, functional and understandable it covers a wider range of population around the world.

3. Requirements and planning

In Software engineering, a functional requirement defines a function of a system or its component. A non-functional requirement is a requirement that specifies criteria that can be used to judge the operation of a system[13]. As shown in Table 1, a numeric rating (in descending order of importance) to be assigned to all functional requirements of the project.

Table 1: priority scale

priority	1 very high	2 high	3 normal	4 low	5 optional
-----------------	--------------------	---------------	-----------------	--------------	-------------------

Let's start with the functional requirement, we set for the our mobile game:

- **Functional requirements**

Table 2 shows sample of game app functional requirements. Because games can be very demanding when it comes to functional requirements details, the functional requirements are split into three conceptual sections based on app system architecture.

- 1) **Game app section**, which is responsible for the app main menu and basic app navigations.
- 2) **Game world basics section**, it is an artificial universe, an imaginary place in which the events of the game occur. It represents the actual game play itself.
- 3) **The player basics section**, which may also be referred to as three basics, User, Player, Hero. **User**: the person who is using the game app. **Player**: is the user while in the state of interacting with the game world. Usually by playing in the role of a game character. In this particular case “Batcha”. **Hero**: is another name for the main character in this particular case, its (Batcha).

Table 2: Game app functional requirements

Game app basics			
ID	requirement	Testing method	priority
FRGB01	The system should allow user to become a player	Black Box	1
FRGB02	The system should allow easy access to game world	Black Box	1
FRGB03	The system should allow user to choose game level	Black Box	1
FRGB04	The system should allow user to choose language	Black Box	1
FRGB05	The system should allow user to choose language	Black Box	1
...

Once the Mobile game app functional requirements have been sorted and prioritized,

Next, we list some of the non-functional requirements:

- **Non-functional requirements**
 - **Response time:** The average response time between click and reaction must be less than 0.5 seconds.
 - **Required resources:** The game should be able to run with minimum of 512 MB of RAM. And shouldn't exceed 100 MB of total size.
 - **Platform:** The game must run on Android mobile devices.
 - **Frames rate:** The minimum frame rate should be more than 15 frames per second.

Use case diagrams

As a sample one of the use case diagrams, Figure-3 represents the interaction between the user and main menu.

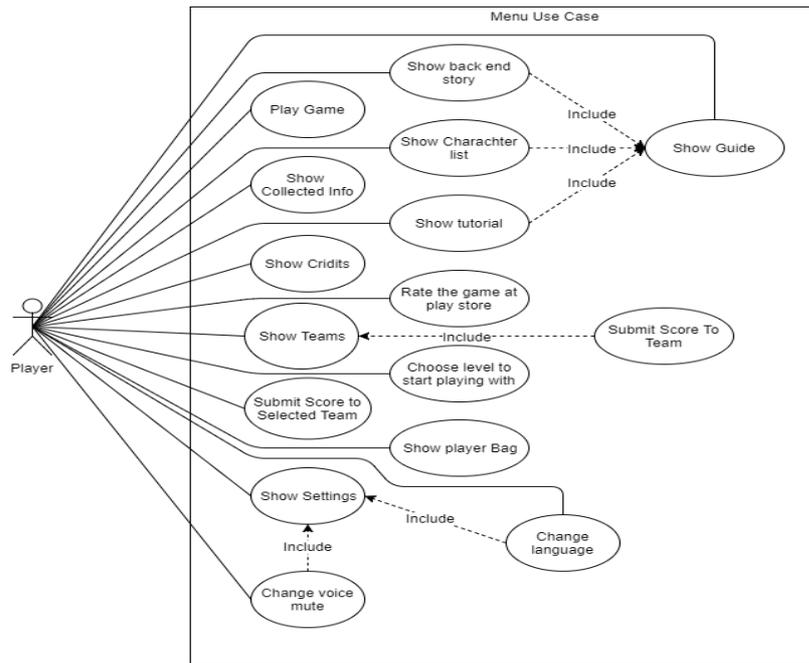


Figure 3: Main menu use case diagram
Activity diagram

Our mobile game play activity is shown in Figure-4, and it's showing the different activities that the user can perform in the game play menu.

- **Die:** refers to the event in which the player dies, from any cause,
- **Win the Game:** refers to the action of the player winning the game and this happens when the player goes to the end of the level without getting die.
- **Pause Menu:** refers to the action of pausing the game when clicking on the pause button.

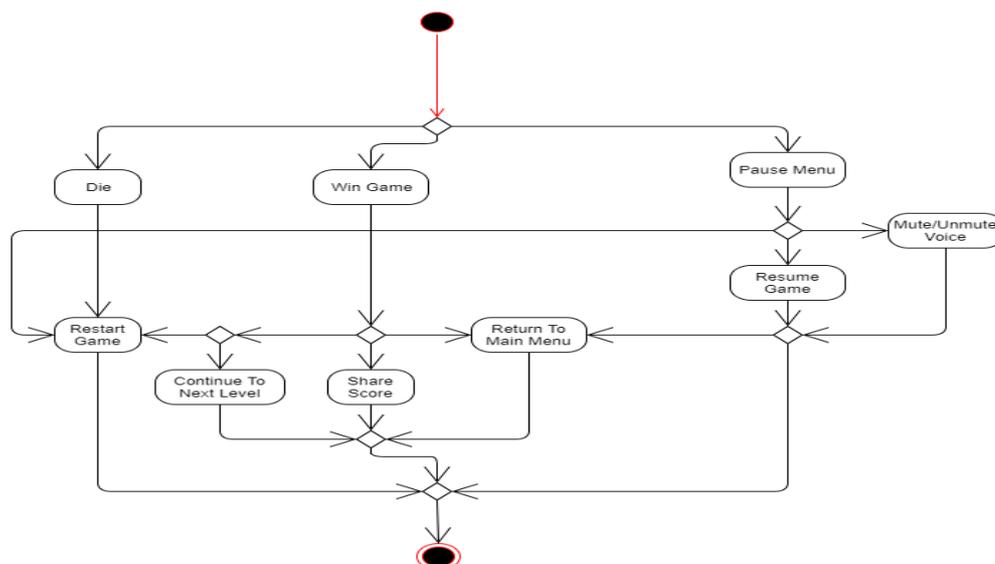


Figure 4: Game play activity diagram

4. Art and software design

Game design is split to art design and software design. First we must understand the relation.

4.1 Art design

The art design process is a very intense part of any game project and so it was for our mobile game too. Usually it's handled by a professional team, or by using premade assets which can be bought online. Thankfully we have managed to complete all the necessary elements of the game art in time. We made sure ALL elements are original in design to our mobile game only. The design of the five main characters of the our mobile game are shown in figure-5. Shown in figure-6, Azooz algayla – a famous evil character which is known through old Libyan stories, the old lady is known to kidnap kids who misbehave in the afternoon. Civilian and evil theme, in figure-7, are characters that are casually spread across game levels, their main function is to stop player from throwing batch randomly within the game, as they decrease the score with each hit they receive. Background scenes are a selection of blocks that represent the old city of Tripoli and some of its most important sights. Selected scenes of our mobile game are shown in figure-8.



Figure 5: Five main characters design



Figure 6: 3zooz algayla design

In figure-8, the Evil theme of the game, shows that the dark side of the game which the player is fighting against. The sky background image changes randomly at the start of each game session.



Figure 7: Civilians and evil theme design

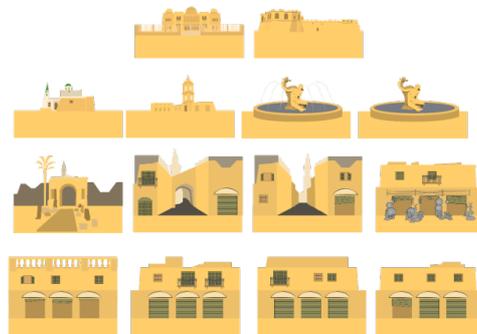


Figure 8: Background compatible blocks design

The figure-9 represent the variety of different animations that have been implemented into the hero character (Batcha) while jumping animation of the Our mobile game to represent different game states.



Figure 9: Hero jumping animation

Educational aspects are designed to help users to learn about Libya (the past, present and future...), culture and traditions, Libyan people and words (Libyan phrases), Geography (Map), and monuments and sightseeing.

4.2 Software design

As a data solution, Class serialization is chosen instead of relational databases so the data model implementation will be classes and class diagrams will be drawn to illustrate the data in them, but first there's explanation and figures of the Entity Relation Diagram.

An entity-relationship model (ERM) is an abstract conceptual representation of structured data. Entity-relationship modelling is a relational schema database modelling method, used in software engineering to produce a type of conceptual data model (or semantic data model) of a system, often a relational database, and its requirements in a top-down fashion.

The figure-10 illustrate the whole system Entity relationship diagram.

In the next section, we discuss the tools used in the implementations and testing conducted.

5. Implementation

In the implementation, we favoured Unity vs. libGDX. The problem we had with libGDX was that we were unable to display the Arabic letter on the screen[14].

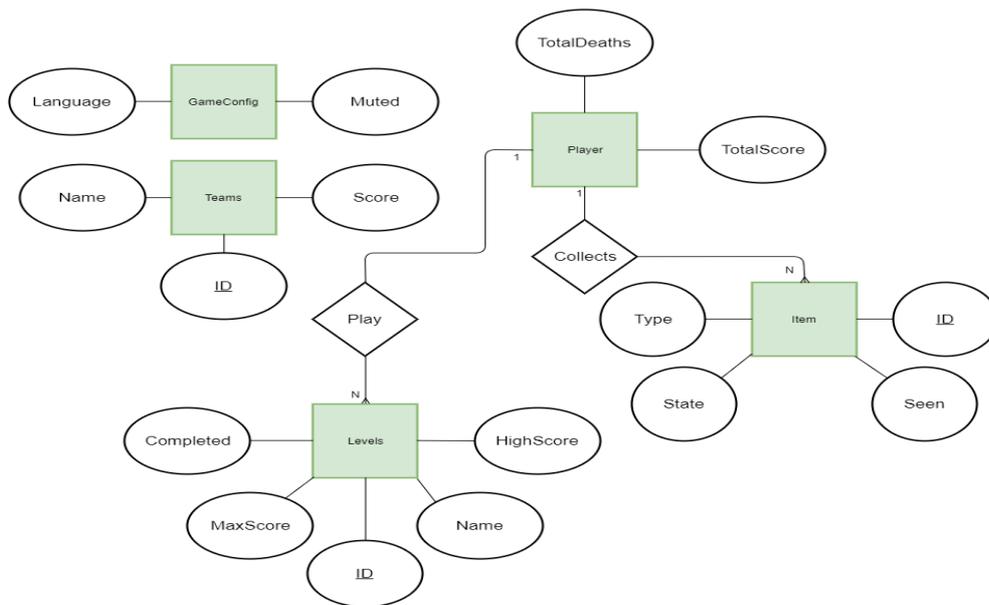


Figure 10: game ERD

5.1 Developments tools

C# programming language: We choose it for writing scripts in Unity, we choose it because its modern language packed with a lot of intensive features that will help us to get the job done in less time and in a more efficient way [15].

Localization Editor(LE): It gave us the ability to manage the variation of languages in our mobile game app and separate the texts of the UI from the code to make the editing process of the string more effective and less time consuming[16].

Spriter2UnityDX: We used it to translate the animations created by Spriter into language that can be understandable by Unity[17].

SpriterDotNet: It's scripts that gave us the ability to run the imported Spriter project in the .NET framework[18].

Firestore Realtime Database: We used it as a cloud-hosted database [19]. Our data is stored as JSON and synchronized in Realtime to every connected client.

5.2 Testing and debugging

We used white box testing and black box testing. White box testing is the testing when the developer should know the inner workings and start test based on that. We used Path Testing technique to calculate the number of possible paths and the complexity of the app which can be defined as:

$$V(G) = E - N + 2$$

Where, E – No. of edges and N – No. of Nodes

$$V(G) = P + 1$$

Where P = No. of predicate nodes (node that contains condition)

Example : Get Teams Data Tests

Step 1: Construction of graph with nodes and edges from the code, figure-11 Shows the Get Teams Data code with numbered paths, and figure-12 Shows all paths for this code.

Step 2: Identification of independent paths.

Path 1: 1 – 2 – 12 – 13

Path 2: 1 – 2 – 3 – 4 – 5 – 13

Path 3: 1 – 2 – 3 – 4 – 6 – 7 – 8 – 9 – 13

Path 4: 1 – 2 – 3 – 4 – 6 – 7 – 8 – 10 – 11 – 13

Step 3: Cyclomatic Complexity Calculation[20]

$$E = 15, N = 13, P = 3$$

$$V(G) = E - N + 2 \qquad V(G) = 15 - 13 + 2 = 4$$

$$V(G) = P + 1 \qquad V(G) = 4$$

Step 4:- Design of Test Cases

The Table 3 shows test cases for selected path we created.

Table 3: Path 2: 1 – 2 – 3 – 4 – 5 – 13

Inputs	Expected Results
FindObjectOfType<GameManager>().Connected == false && Application.internetReachability == NetworkReachability.NotReachable	Hide the please wait text and show internet error message and hide the error text after 3 Seconds

Black Box Testing is the method that let the tester test the system without knowing the specifics of the code itself, so it gives the ability for everyone to try it. Here we conducted tests on several of the parts of the system against the project requirements. Table 4 Shows a selected test case of change language of our mobile game.

Table 4: Test Case - 001

TC-001	
Purpose	Change the Language
Setup	Lunch the game
Steps	Select the settings menu Choose language
Expected Reponses	The language of the entire app should be changed
Result	Pass
Requirement covered	FRGB05

```

1 DatabaseReference reference = FirebaseDatabase.DefaultInstance.RootReference;
  DatabaseReference teamsRef = reference.Child("Database").Child("Teams");
  GameManager gm = FindObjectOfType<GameManager>();

  GameObject.Find("PleaseWait").GetComponent<Animator>().SetTrigger("Show");
2 gm.Get("https://www.google.com", () =>
3 {
4     Debug.Log("Completed");
5     if (!FindObjectOfType<GameManager>().Connected &&
6         Application.internetReachability == NetworkReachability.NotReachable)
7     {
8         GameObject errorText = GameObject.Find("ErrorText");
9         errorText.GetComponent<Text>().enabled = true;
10        errorText.GetComponent<LELocalize>().localized_string_key = "no_internet";
11        errorText.GetComponent<LELocalize>().UpdateLocString();
12        GameObject.Find("PleaseWait").GetComponent<Animator>().SetTrigger("Hide");
13        Invoke("HideErrorText", 3);
14    }
15    if (FindObjectOfType<GameManager>().Connected ||
16        Application.internetReachability != NetworkReachability.NotReachable)
17    {
18        teamsRef.GetValueAsync().ContinueWith((task1) =>
19        {
20            GameObject.Find("PleaseWait").GetComponent<Animator>().SetTrigger("Hide");
21            if (task1.IsFaulted)
22            {
23                // Handle the error...
24            }
25            else if (task1.IsCompleted)
26            {
27                DataSnapshot snapshot = task1.Result;
28                Debug.Log("Done: " + snapshot.Children);
29                Reafresh(snapshot.Children);
30            }
31        });
32    }
33 }, () =>
34 {
35     GameObject errorText = GameObject.Find("ErrorText");
36     errorText.GetComponent<Text>().enabled = true;
37     errorText.GetComponent<LELocalize>().localized_string_key = "no_internet";
38     errorText.GetComponent<LELocalize>().UpdateLocString();
39     GameObject.Find("PleaseWait").GetComponent<Animator>().SetTrigger("Hide");
40     Invoke("HideErrorText", 3);
41
42     Debug.Log("error");
43 });
44 }
45 }
    
```

Figure 11: Get teams data logic

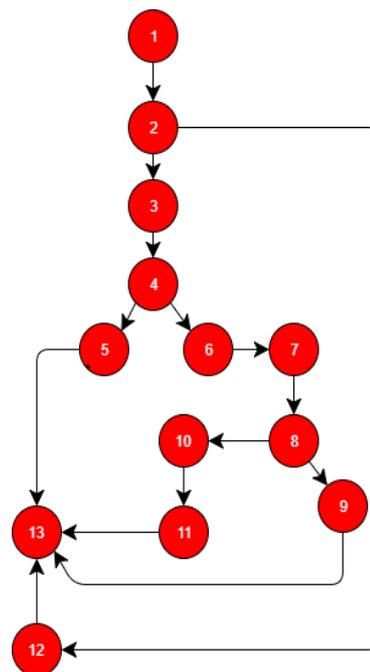


Figure 12: Get teams data logic graph paths

6. Conclusions:

In summary, we have created an educational Libyan game using Unity engine. From the design prospective, the game integrates Iconic Libyan themes with game rules. At different levels, various Libyan sights are displayed. Icons are presented in different shapes such as puzzles, tokens, or treasures for the player to collect. Also, obstacles, backgrounds and all game images have a sense of familiarity to Libyan life style and culture. From the implementation prospective, the game was developed using Unity game engine which allowed us to create a game for cross platforms, combined with the art work made using Adobe Illustrator to create vector arts and Spriter to create character animations for game play. Other programs and development tools are used as needed. Furthermore, we adhere to Agile approach in the game development process.

Future work is mainly perfecting the game design, focusing on the user feedback and experience, we'll continue to fix any issues with the game play. There are a few options that can be added to receive a decent income if the project is to be successful, however these options need to be studied sensibly and put into further careful consideration as they may hinder the overall performance and popularity of the game.

References

- [1]. K. Alomari, T. Soomro, K. Shaalan, "Mobile Gaming Trends and Revenue Models", International Conference of Industrial Engineering and Other Applications of Applied Intelligent Systems, pp. 671-683, August, 2016. DOI:10.1007/978-3-319-42007-3_58
- [2]. Market research firm Newzoo, "Global mobile market report", September, 2019. Access online on 10 September 2020 at Available: <https://venturebeat.com/2019/09/17/newzoo-global-game-revenues-will-hit-68-5-billion-in-2019/>
- [3]. C. Chang, G. Hwang, "Trends in digital game-based learning in the mobile era: a systematic review of journal publications from 2007 to 2016", International Journal of Mobile Learning and Organisation, vol.13 no.1, pp.68 – 90, 2019. DOI: 10.1504/IJMLO.2019.096468
- [4]. D. Mattingly, S. McLaren, E. Savage, Y. Al-Fasatwi, and K. Gadgood, "The Libyan Desert, natural resources and cultural heritage", Society for Libyan Studies, Cambridge University Press, vol. 37, No.6, pp. 165-167, March, 2015. DOI: <https://doi.org/10.1017/S0263718900004118>
- [5]. N. Brodie, "Why Is No One Talking about Libya's Cultural Destruction?", A journal of the American Schools of Oriental Research, vol. 78, no. 3, pp. 212-217, September, 2015. DOI: 10.5615/neareastarch.78.3.0212
- [6]. M. Garin, "Super Mario, the new silent clown: Video game parodies as transformative comedy tools", International Journal of Cultural Studies, vol. 18, no 3, pp. 305-309, May 2015. <https://doi.org/10.1177/1367877913513688>
- [7]. M. Böhmer, B. Hecht, J. Schöning, A. Krüger, G. Bauer, "Falling asleep with Angry Birds, Facebook and Kindle: A large scale study on mobile application usage", Proceedings of the 13th Conference on Human-Computer Interaction with Mobile Devices and Services, August, 2011. DOI: 10.1145/2037373.2037383
- [8]. Biuk, "Libyan Child Play," Access online on 10 September 2020 at Available: <http://www.tripolipost.com/articledetail.asp?c=17&i=9459>.
- [9]. Sinicki, "Learn Unity for Android Game Development A Guide to Game Design, Development, and Marketing", APress, 2017. DOI 10.1007/978-1-4842-2704-6_2
- [10]. R. Rosenbaum, C. Tominski, "published Pixels vs. Vectors: Presentation of Large Images on Mobile Devices", Conference: International Workshop on Mobile Computing (IMC), Packi Publishing Ltd, Jan, 2003. Access online on 10 September 2020 at <https://www.researchgate.net/publication/247160237>
- [11]. S. B. Kaleel and S. Harishankar. "Applying agile methodology in mobile software engineering: Android application development and its challenges", Technical report, Department of Computer Science, Ryerson University, 2013.



- [12]. K. Bjornard , K. principle, the Interaction Design Foundation. Access online on 10 September 2020 at Available: <https://www.interaction-design.org/literature/article/kiss-keep-it-simple-stupid-a-design-principle>.
- [13]. Tsui, O. Karam, B. Bernal, "Essentials of Software Engineering 4th", Jones and Bartlett Learning, 2018.
- [14]. M. Zechner, "libgdx", 2013. Access online on 10 September 2020 at Available: <https://libgdx.badlogicgames.com/>
- [15]. Troelsen, P. Japikse, " Pro C# 8 with .NET Core 3 Foundational Principles and Practices in Programming" , Apress, 2020. DOI:10.1007/978-1-4842-5756-2
- [16]. L. Editor. [Online]. Access online on 10 September 2020 at Available: <http://localizationeditor.com/>
- [17]. Spriter2UnityDX, Spriter2UnityDX. Access online on 10 September 2020 at Available: <https://github.com/Dharengo/Spriter2UnityDX>
- [18]. SpriterDotNet, Access online on 10 September 2020 at Available: <https://github.com/loodakrawa/SpriterDotNet>.
- [19]. M. Ohyver, J. Moniaga, I. Sungkawa, B. Subagyo, I. Chandra, "The Comparison Firebase Realtime Database and MySQL Database Performance using Wilcoxon Signed-Rank Test", the 4th international conference on computer science and computational intelligence, vol.157, , pp. 396-405, September, 2019. <https://doi.org/10.1016/j.procs.2019.08.231>
- [20]. Jay, J. Hale, R. Smith, D. Hale, "Cyclomatic Complexity and Lines of Code: Empirical Evidence of a Stable Linear Relationship", Journal of Software Engineering and Applications, vol.2, no. 3, pp. 137-143, January, 2009. DOI: 10.4236/jsea.2009.23020

Comparison of BER Performance of Cooperative Wireless Systems Based on D-OSTBC, D-EOSTBC and DF Relaying Protocol

Masoud Eddaghel^{1*}, Mohamed Nashnush², Adnan BenSaiid³, Alhossin Aljadai⁴

¹ masoud_edd@yahoo.com, ² mohammednash@alkafaa.net, ³ adnanbensaiid.adi@gmail.com,

⁴ a.aljadai1@unimail.derby.ac.uk

¹ Department of Communications & Microwaves, College of Electronic Technology-Bani Walid, Libya

^{2,3} Department of Communications, College of Electronic Technology-Tripoli, Libya

⁴ Engineering and Technology College, University of Derby, DE1 3HD, Derby, UK

ABSTRACT

The end-to-end bit error rate (BER) performance analysis of cooperative wireless communication networks based on distributed orthogonal space time block coding (D-OSTBC) (Alamouti) scheme, distributed extended orthogonal space time block coding (D-EOSTBC) and decode-and-forward (DF) relaying protocol, are addressed. There are two main relaying techniques which are amplify-and-forward (AF) and decode-and-forward (DF). Opposed to that in conventional multi-input multi-output (MIMO) systems, cooperative systems, the antennas are distributed among different terminals, which easy to establish cooperative MIMO. This paper will deal with three system models of wireless communications. The first model is conventional single-input single-output (SISO) system, which consists of one source and one destination. The second model is cooperative MIMO based on two relays and D-OSTBC (Alamouti) scheme. This system consists of one source, two relays and one destination. The third model is cooperative open loop MIMO based on four relays and D-EOSTBC technique. This system consists of one source, four relays and one destination. All these schemes are based on DF protocol and analyzed over flat fading channels. Also, QPSK digital modulation scheme is used for all schemes. MATLAB simulations confirm that the BER performance of cooperative SISO is the worst, whereas, the best BER performance is achieved by the system with four relays and based on D-EOSTBC techniques.

Keywords:

pair relays cooperation;
four relays cooperation;
D-OSTBC,
D-EOSTBC;
decode and
forward technique

1- Introduction

MIMO communications are applied in real-world wireless communications systems, such as standards, WiFi and LTE/LTE-advanced, which are used extensively for wireless LAN applications and conventional cellular communications, respectively. The version of WiFi that employs MIMO is defined by the IEEE 802.11n standard. IEEE 802.11n supports a wide variety of antenna configurations, ranging from SISO arrangements to $m \times n$ MIMO configurations with various combinations of STBC [1]. More specifically, as in [1], the antenna configurations of IEEE 802.11n is up to four antennas at either the transmitter, receiver, or both.

Cooperative communication is a new communication paradigm which generalizes conventional MIMO communications to much broader applications [2]. Conventional MIMO system may not always be practical to accommodate multiple antennas at the mobile stations, due to cost, size and other hardware limitations. Therefore, the distributed orthogonal space time block coding (D-OSTBC) techniques provide promising solutions to achieve cooperative MIMO systems requiring

reliable wireless communications at high rates. As in [3], an attempt to attain a better communication efficiency, the concept of cooperative MIMOs has been proposed for cellular systems. More specifically, a group of mobile nodes shares their antennas with other users to create a virtual antenna array to provide spatial diversity gain, this group of mobile nodes known as relays. Figure 1 demonstrates the concept of relay system and potential applications of various cooperative MIMO systems in a cellular network.

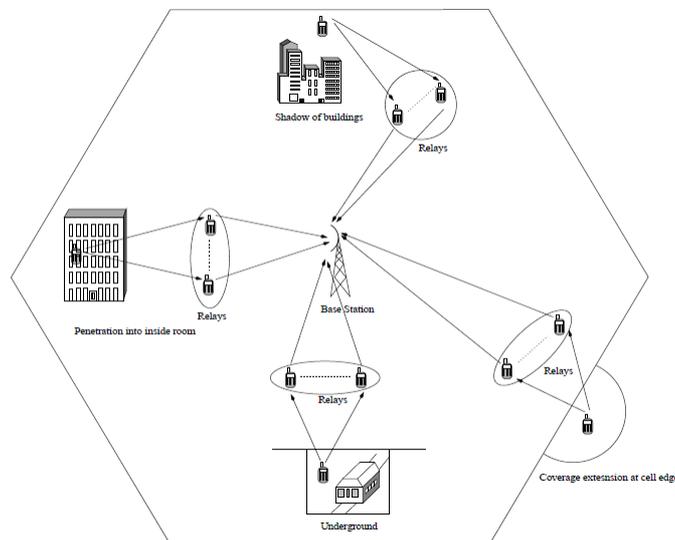


Figure 1: Cooperative MIMO systems in a cellular network [2].

For example, when a user is behind buildings or underground, as in Figure 1, direct communication with the base station (BS) becomes unreliable, due to severe shadow fading and path loss. In order to maintain reliable wireless communications, a group of neighbour users may form a virtual antenna array in order to fix a reliable connection to transmit the data between the source node and the BS. There are several types of relaying cooperation protocols such as Amplify-and-Forward (AF), Decode-and-Forward (DF) and Compress-and-Forward (CF). AF cooperation scheme, the relay nodes simply amplify the received signal waveforms, but they amplify the signal and noise jointly and hence are unable to improve the SNR. Whereas, in the DF cooperation strategy, the signals received at the relays are decoded and possibly re-encoded using different Forward Error Correction (FEC) codes, before being forwarded to the destination. Finally, the CF arrangement, the relay compresses, estimates or quantizes its observations without decoding the information. However, most routing protocols use a DF strategy at each relay node, where packets received by the relay are decoded to remove errors through error correction and retransmissions requested when errors are detected that cannot be corrected. Therefore, in this work, DF is exploited. However, DF strategy suffers from the error propagation problem because the relays do use a hard-decision operation on the received signal [4], [5].

The Focus of This Paper

In wireless digital communication system, the common measure of performance is a BER. Therefore, this paper focuses on measurements of BER performance of conventional SISO system, cooperative MIMO systems based on D-OSTBC (Alamouti) scheme and cooperative open loop MIMO systems based on D-EO-STBC technique. These systems exploit QPSK modulation scheme and DF relaying protocol. Moreover, the analysis is based on flat fading channels. In addition, the performance of these systems is compared with cooperative SISO systems and cooperative system with one relay and direct link between the source and destination. Finally, simple mathematical analysis of transmission and receive process of cooperative open loop MIMO wireless system with four relays over flat fading channel is presented.

Organization of the Paper

The rest of this paper is organized as follows: In Section 2, the system model of cooperative open loop MIMO systems based on D-EO-STBC techniques over flat fading channel is described. In Section 3, the Mathematical model of cooperative open loop MIMO systems based on four relays, D-EO-STBC techniques and DF relaying protocol, over flat fading channel. is explained. Simulation results are drawn in Section 4 and conclusions are given in Section 5.

Common Notations

Common notations that have been used in this paper are $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ which denote complex conjugate, transpose and Hermitian transpose operations, respectively.

Next, system model of cooperative open loop MIMO systems based on four relays, open loop D-EO-STBC techniques over flat fading channel will be presented.

2- System Model of Cooperative Open Loop MIMO System

The wireless cooperative communication system with one source node S , four relays nodes R , and one destination node D . In this system, each node has only one antenna as in Figure 2. The faded channel coefficient from the source node to the i^{th} relay and the faded channel coefficient from the i^{th} relay to the destination node are denoted as h_{SR_i} and h_{R_iD} , respectively. The system works under assumptions, the terminals operate in a half-duplex mode and the channel between any two terminals is quasi-static flat fading. Therefore, assume that h_{SR_i} and h_{R_iD} are independent complex Gaussian random variable with mean zero and variance one. The transmission process of the information from the source node to the destination node, they experience two phases. Phase one, the source node broadcast the information $\mathbf{s} = [s_1 \ s_2^*]^T$ to the relay nodes. Then, phase two relay nodes resend the information to destination node, which called relaying or cooperative phase.

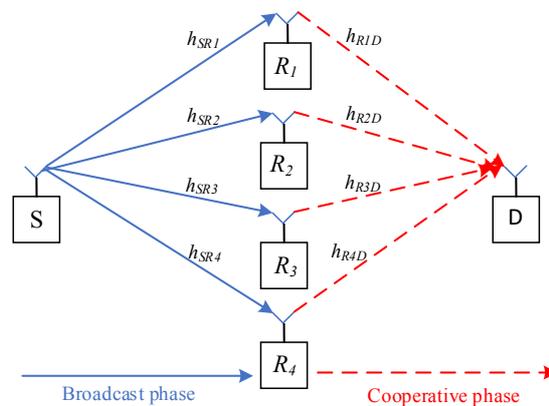


Figure 2: System model of cooperative open loop MIMO systems based on four relays, D-EO-STBC techniques and DF relaying protocol, over flat fading channel.

Next, mathematical model of cooperative open loop MIMO systems based on four relays, D-EO-STBC and DF relaying protocol, over flat fading channel will be explained.

3- Mathematical Model of Cooperative Open Loop MIMO System

Mathematical Model is obtained from the schematic diagram of cooperative open loop MIMO system based on four relays, D-EO-STBC technique, over flat fading channel. As in Figure 2., cooperation strategies involve two phases of transmission. Phase one, the source node broadcast the information $\mathbf{s} = [s_1 \ s_2^*]^T$ after modulation onto complex symbols to the relay nodes, R_1 , R_2 , R_3 and R_4 , where $(\cdot)^T$ denotes vector transpose. The received signal at the R_1 , R_2 , R_3 and R_4 , is corrupted by both the fading coefficient h_{SR_1} , h_{SR_2} , h_{SR_3} , h_{SR_4} and the noise n_{R_1} , n_{R_2} , n_{R_3} , n_{R_4} , respectively. The received signals in the first time slot: at i^{th} relay is:

$$r_{Ri} = h_{SRi} s + n_{Ri} \quad (1)$$

Then, the received signal is decoded at i^{th} relay by exploiting an STBC code matrix which it takes on the following form as in [6], [7]:

$$D_i = A_i r_{Ri} + B_i r_{Ri}^*, \quad (2)$$

where $\mathbf{s} = [s_1 \ s_2]^T$ is a set of symbols to be transmitted, and $\{A_i, B_i\}$ are fixed code matrices at relays of dimension number of data signal (\mathbf{s}) and number of transmission time slots (T). Then,

$$D_i = h_{SRi} A_i s + h_{SRi}^* B_i s^*. \quad (3)$$

In this system two relay pairs are used. These relay nodes are designed to use the following A_i and B_i matrices at the i^{th} relays:

$$A_1 = A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \& \quad B_1 = B_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad (4)$$

where, matrices A_i and B_i are at R_1, R_2, R_3 and R_4 , respectively. Then, the received signal at R_1 and R_2 during 1st time slot:

$$r_{Rij} = D_{ij}, \quad (5)$$

where, i and j denote number of relays and time slots, respectively. So, the received signal at R_1, R_2, R_3 and R_4 , during 1st time slot:

$$r_{Ri1} = h_{SRi} A_1 s + n_{Ri} = h_{SRi} s_j + n_{R1}, \quad i = 1, 2 \text{ and } j = 1, \quad (6)$$

$$r_{Ri1}^* = h_{SRi}^* B_1 s^* + n_{Ri} = h_{SRi} s_j + n_{Ri}^*, \quad i = 3, 4 \text{ and } j = 2. \quad (7)$$

Then, the estimated received signals $\widehat{s}_{11}, \widehat{s}_{12}, \widehat{s}_{23}$ and \widehat{s}_{24} at R_1, R_2, R_3 and R_4 , respectively, during 1st time slot, can be expressed as follows:

$$\widehat{s}_{11} = h_{SR1}^* r_{R11}, \quad \widehat{s}_{12} = h_{SR2}^* r_{R21}, \quad \widehat{s}_{23} = h_{SR3} r_{R31}^* \text{ and } \widehat{s}_{24} = h_{SR4} r_{R41}^* \quad (8)$$

Then, the received signal at R_1, R_2, R_3 and R_4 , during 2nd time slot:

$$r_{Ri2} = h_{SRi} A_1 s + n_{Ri} = h_{SRi} s_j + n_{Ri}, \quad i = 1, 2 \text{ and } j = 2, \quad (9)$$

$$r_{Ri2}^* = h_{SRi}^* B_1 s^* + n_{Ri} = -h_{SRi}^* s_j + n_{Ri}^*, \quad i = 3, 4 \text{ and } j = 1. \quad (10)$$

Then, the estimated received signals $\widehat{s}_{21}, \widehat{s}_{22}, \widehat{s}_{13}$ and \widehat{s}_{14} at R_1, R_2, R_3 and R_4 , respectively, during 2nd time slot, can be expressed as follows:

$$\widehat{s}_{21} = h_{SR1}^* r_{R12}, \quad \widehat{s}_{22} = h_{SR2}^* r_{R22}, \quad \widehat{s}_{13} = h_{SR3} r_{R32}^* \text{ and } \widehat{s}_{14} = h_{SR4} r_{R42}^*. \quad (11)$$

The matrix form of EO-STBC, which will be performed at relays, is

$$\mathbf{S} = \begin{bmatrix} s_1 & s_1 & s_2 & s_2 \\ s_2^* & s_2^* & -s_1^* & -s_1^* \end{bmatrix}. \quad (12)$$

EO-STBC is performed at relays by sending the decoded signal at each relay in time slots. In the first time slot R_1, R_2, R_3 and R_4 , send $\widehat{s}_{11}, \widehat{s}_{12}, \widehat{s}_{23}$ and \widehat{s}_{24} , respectively. Then, the received signal at the destination can be stated as follows:

$$r_{1D} = \widehat{s}_{11} h_{R1D} + \widehat{s}_{12} h_{R2D} + \widehat{s}_{23} h_{R3D} + \widehat{s}_{24} h_{R4D} + w_1. \quad (13)$$

In the second time slot R_1, R_2, R_3 and R_4 , send $-\widehat{s}_{21}^*, -\widehat{s}_{22}^*, \widehat{s}_{13}^*$, and \widehat{s}_{14}^* , respectively. Then, the received signal at the destination can be stated as follows:

$$r_{2D}^* = \widehat{s}_{13} h_{R3D}^* + \widehat{s}_{14} h_{R4D}^* - \widehat{s}_{21}^* h_{R1D}^* - \widehat{s}_{22}^* h_{R2D}^* + w_2^*. \quad (14)$$

Hence, the received signal at destination can be written in matrix form as follows:

$$\mathbf{r}_D = \mathbf{H}\mathbf{s} + \mathbf{w}_i, \quad (15)$$

where, $\mathbf{r}_D = [r_{1D} \ r_{2D}^*]^T$, $\mathbf{s} = [s_1 \ s_2^*]^T$, $\mathbf{w}_i = [w_1 \ w_2^*]^T$ and

$$\mathbf{H} = \begin{bmatrix} h_{R1D} + h_{R2D} & h_{R3D} + h_{R4D} \\ h_{R3D}^* + h_{R4D}^* & -h_{R1D}^* - h_{R2D}^* \end{bmatrix}. \quad (16)$$

The estimated received signal at destination can be represented as follows:

$$\widehat{\mathbf{s}} = \mathbf{H}^H \mathbf{r}_D, \quad (17)$$

where, $(\cdot)^H$ is the Hermitian transpose. By substituting (16) in (17), yields (18)

$$\widehat{\mathbf{s}} = \mathbf{H}^H \mathbf{H} \mathbf{s} + \mathbf{H}^H \mathbf{w}_i, \quad (18)$$

$$\begin{bmatrix} \widehat{s}_1 \\ \widehat{s}_2 \end{bmatrix} = \begin{bmatrix} \alpha + \beta & 0 \\ 0 & \alpha + \beta \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \mathbf{H}^H \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}, \quad (19)$$

where, $\alpha = \sum_{i=1}^4 |h_{RiD}|^2$ is the channel gain. $\beta = \beta_1 + \beta_2$, are the channel interferences, $\beta_1 = 2\text{Re}\{h_{R1D} h_{R2D}^*\}$ and $\beta_2 = 2\text{Re}\{h_{R3D} h_{R4D}^*\}$. Then, the estimated received signals are:

$$\hat{s}_1 = (\alpha + \beta) s_1 + (h_{R1D}^* + h_{R2D}^*) w_1 + (h_{R3D} + h_{R4D}) w_2^* \quad (20)$$

$$\hat{s}_2 = (\alpha + \beta) s_2 + (h_{R3D}^* + h_{R4D}^*) w_1 - (h_{R1D} + h_{R2D}) w_2^* \quad (21)$$

In next subsection, simulation results will be presented.

4- Simulation Results and Discussion

Comparison of simulation results of conventional SISO, cooperative SISO and cooperative MIMO systems

In this section, simulation results, by using MATLAB, are provided to demonstrate the comparison of BER performance conventional SISO system, cooperative SISO system, cooperative system consists of one source, one relay, one destination, and direct path between source and destination, cooperative MIMO based on D-OSTBC, and cooperative open loop MIMO based on D-EO-STBC. All these schemes are based on DF protocol and analysed over flat fading channels. Also, QPSK digital modulation schemes is exploited for all schemes. The perfect channel state information (CSI) is assumed to be perfectly available at relays and at the receiver side.

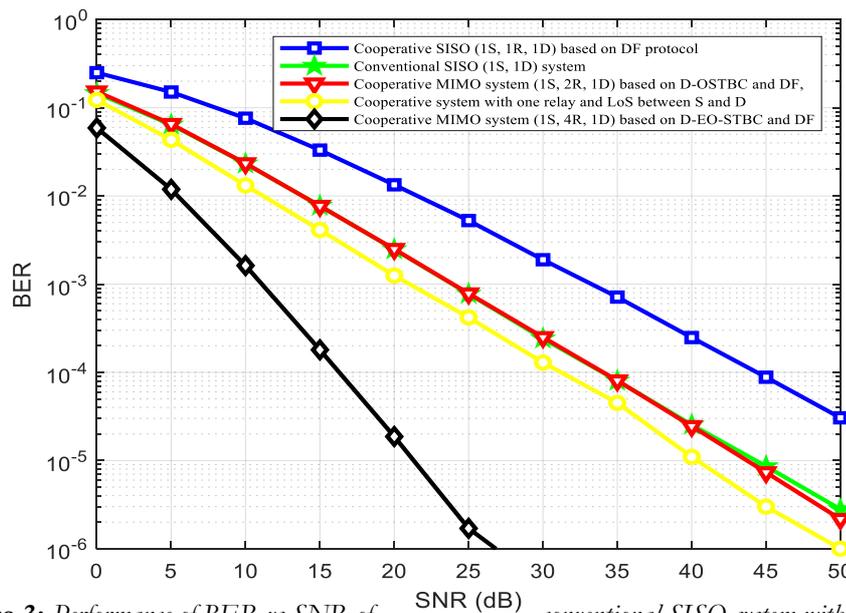


Figure 3: Performance of BER vs SNR of conventional SISO system with single transmit and receive antenna and cooperative SISO, cooperative MIMO (1S, 2R, 1D) systems based on D-OSTBC (Alamouti) scheme, and cooperative MIMO (1S, 4R, 1D) systems based on open loop D-EO-STBC scheme, over flat fading channel.

Figure 3 depicts the comparison of the BER performance of conventional SISO system, cooperative SISO system and cooperative system with one relay and direct link between the source and destination, cooperative MIMO systems based on D-OSTBC (Alamouti) and cooperative open loop MIMO system based on D-EO-STBC technique. All systems effected by flat fading channel. From the graph, it is clearly seen that the cooperative SISO scheme has the worst BER performance. Whereas, the performance of cooperative MIMO systems based on D-OSTBC (Alamouti) is almost identical to conventional SISO system. Also, cooperative open loop MIMO

system based on D-EO-STBC technique has the best BER performance. For example, at 10^{-5} BER reference, cooperative SISO needs approximately 47 dB, whereas cooperative system with one relay and direct link between the source and destination needs about 41 dB. However, cooperative open loop MIMO system based on D-EO-STBC needs just only 22 dB. This means that, the cooperative open loop MIMO system based on D-EO-STBC improves the BER performance by about 18 dB compared with cooperative system with one relay and direct link between the source and destination.

5- Conclusions

The end-to-end BER performance analysis of cooperative MIMO based on two relays and D-OSTBC scheme and cooperative open loop MIMO based on four relays and distributed extended orthogonal (D-EO-STBC) technique, over flat fading channels, were studied. QPSK digital modulation scheme was exploited. The results confirmed that the cooperative open loop MIMO scheme with four relays had the better BER performance. In contrast, the cooperative SISO has had the worst BER performance. Moreover, cooperative MIMO systems based on Alamouti scheme and conventional SISO system had almost identical performance. Cooperative closed loop MIMO system based on D-EO-STBC and DF relaying protocol, is a subject of our ongoing study.

6- Acknowledgment

The authors wish to thank The Collage of Electronic Technology-Bani Walid /Libya, The Collage of Electronic Technology-Tripoli /Libya and Engineering and Technology College-Derby University /UK for their cooperation.

References

- [1]. J. R. Hampton, Introduction to MIMO Communications. Cambridge University Press, 2014.
- [2]. W. S. K. Rylie, A. K. Sade and A. Kaminski, *Cooperative Communications and Networking*. Cambridge University Press, 2009. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.
- [3]. L. Hanzo, O. R. Alamri, M. El-Hajjar, N. Wu, Near-Capacity Multi-Functional MIMO Systems: Sphere-Packing, Iterative Detection and Cooperation, John Wiley & Sons, Ltd, 2009.
- [4]. A. Goldsmith, Wireless Communication. Cambridge University Press. 2005.
- [5]. Y.-W. Peter Hong, Wan-Jen Huang and C.-C. Jay Kuo, *Cooperative Communications and Networking*. Springer New York Dordrecht Heidelberg London, 2010.
- [6]. Erik G. Larsson, Petre Stoica, "Space-Time Block Coding for Wireless Communications", Cambridge University Press, 2008.
- [7]. T. Alotaibi and J. A. Chambers, "Extended orthogonal space-time block coding scheme for asynchronous cooperative relay networks over frequency-selective channels," *2010 IEEE 11th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Marrakech, 2010, pp. 1-5.

Classifying Various Bacteria Genera by Transfer Learning Model

Mohammed Elgaed, Khaled Elgdamsi and Ali Ganoun*

m.elgead@uot.edu.ly, k.elgdamsi@uot.edu.ly, a.ganoun@uot.edu.ly

Electrical and Electronic Engineering Department, Faculty of Engineering, University of Tripoli, Libya

ABSTRACT

Bacteria Classification using computer-aided methods makes the identification and recognition processes more automatic and thus greatly reduces the time needed for classification. In this paper, we explored an approach to automating the process of classifying bacteria with the use of deep Convolutional Neural Network (CNN). CNNs is one of deep machine learning methods that mimics the connectivity pattern between visual cortex neurons. It can extract hierarchical image feature representations based on multi-layer processing. The ‘transfer learning’ approach was used to retrain a famous convolutional neural network model with a dataset of 152 images of 7 different bacteria species. The retrained model has been able to recognize and classify all 7 different species of bacteria with very high accuracy.

Keywords:

bacteria classification,
deep learning,
convolution neural network,
transfer learning,
computer Vision

1- Introduction

The identification of bacterial genera and species is vital, as the biological knowledge of microorganisms is extremely significant in medicine, veterinary science, biochemistry or food industry. Although most of the microorganisms have positive impact on various areas of life, they can be a reason of many diseases (including the infectious ones). Therefore, automating the process of recognition can find application in different sectors such as in medical prevention, diagnosis and treatment. The recognition of microbiological samples is preceded by the culturing process. This process includes dedicated equipment and chemical agents to be used for staining the samples. It also follows stringent culturing procedures and safety protocols. As a result, we obtain samples that are studied for the purpose of discovering characteristic features and classifying the particular genera and species of bacteria.

Classical laboratory methods for bacteria identification and recognition require an expert knowledge and experience [1]. The use of recent artificial intelligence methods such as Convolutional Neural Network (CNN) to automatically identify and classify bacteria is vital for improved recognition accuracy.

Classification is an active research area in pattern recognition which has been used in many applications that involves a decision mechanism that assigns a new observation to a set of classes, based on a training dataset. Traditional methods largely rely on manually crafting descriptors and characteristics. On the contrary, this paper proposes a Deep Learning (DL) approach that can automatically extract features from two-dimensional bacteria images.

The CNN models are an extension of deep learning of artificial networks. They consist of Multi-hidden Layer Perceptron’s (MLPs) which involve multiple convolution, pooling, Rectified Linear Unit (ReLU), and fully-connected layers. Generally, the features maps of the previous layers are convolved with learnable weights on a convolutional layer and fed through the activation functions to form the output features maps. Each output map combines convolutions with multiple input features maps. For a specific output map, the input maps will be convolved with distinct kernels and the convolutional layer shares its weights in a same output map. Convolutional layers are

usually combined with pooling layers to reduce dimensions and increase speed. Although the pooling layer is desirable as it reduces the computational time, it may cause loss of some information. Generally, the pooling layer produces sub-sampled versions of the input maps [2].

2- System Model

Image-processing and pattern recognition techniques combined with various types of classifiers are often used as a useful tool for recognition of laboratory samples, occurring in the form of images. There are many methods for automatic recognition of bacteria species and strains such as statistical methods [3], artificial neural networks [4], or other machine learning classifiers [5]. Most of the mentioned methods are used to recognize very few species or strains (sometimes only one, e.g., tuberculosis). Moreover, in many cases, the algorithms based on the morphological features combined with some classification methods. Therefore, they are constrained.

This paper considers the AlexNet – a pre-trained Convolutional Neural Network (CNN) model – which trained on a subset of ImageNet database of around 1.2 million images of 1000 classes and tested on 150,000 images belonging to 1000 classes [6, 7]. The last three layers of the pre-trained AlexNet network are configured for 1000 classes and must be fine-tuned for the new classification problem. The fine-tuning done by extracting all layers except the last three from the pre-trained network. Then Transferring the layers to the new classification task by replacing the last three layers with a Fully Connected layer, a SoftMax layer, and a classification output layer. Figure 1 shows the AlexNet network architecture.

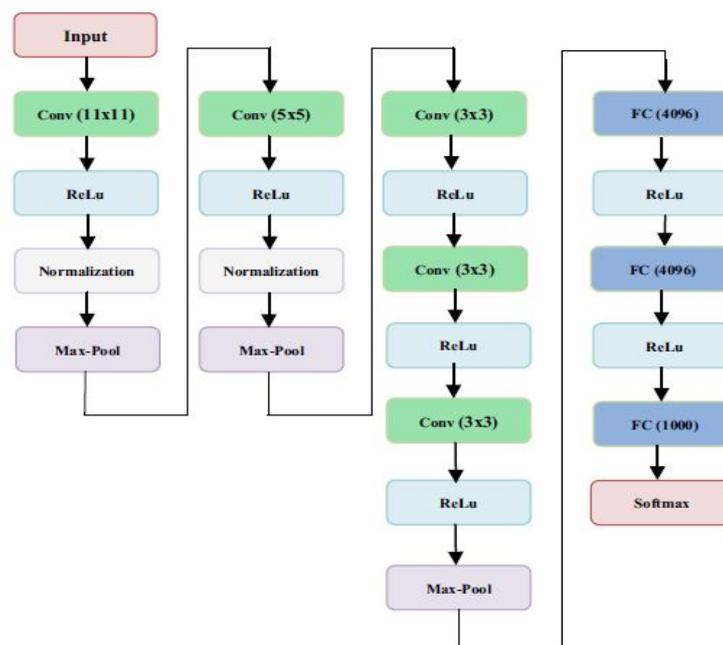


Figure 1: Structure of the AlexNet network architecture

Convolution & Pooling Layers

Convolution layer is the core building block of a convolutional neural network which uses convolution operation. The first layer is a convolution layer which is a feature extraction layer, it takes its input from the input layer or sampling layer. A pooling layer takes each feature map output from the convolutional layer and down-samples it, i.e., pooling layer summarizes a region of neurons in the convolution layer. Figure 2 shows the max-pooling technique

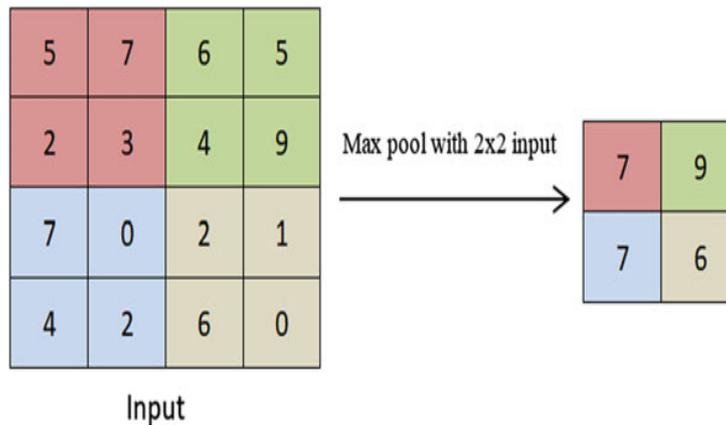


Figure 2: Max-Pooling Operation.

ReLU Layer

Choosing a proper activation function is crucial since it significantly increases the performance of a CNN for certain applications. ReLU is one of the most commonly used function, this activation function produces a graph which is zero when $x < 0$ and linear with slope 1 when $x > 0$ [6, 8]. The remaining layers of convolution and subsampling perform similar to those in previous layers. Only difference is that the extracted features become more abstract as the depth increases.

FullyConnected Layer

The process of convolution and pooling continues to detect the needed features. If there is a classification problem, then the detected features used in the spatial domain to obtain class score, i.e. the probabilities that these features represent each class [6].

Dropout Layer

Deep neural networks consist of multiple hidden layers used to learn more complicated features followed by fully connected layers for decision-making. A fully connected layer is connected to all features, and it is prone to overfitting. Overfitting refers to the problem when a model is trained and it works so well on training data that it negatively impacts the performance of the model on new data. In order to overcome the problem of overfitting, a dropout layer can be introduced in the model in which some neurons along with their connections are randomly dropped from the network during training. A reduced network is left; incoming and outgoing edges to a dropped-out node are also removed. Only the reduced network is trained on the data in that stage. The removed nodes are then reinserted into the network with their original weights. Dropout notably reduces overfitting and improves the generalization of the model [6, 9, 10].

3- Simulation Results

Data Set

The data images consist of 152 images with 7 different genera and species of bacteria. The data images are collected from Digital Image of Bacterial Species (DIBaS) dataset which contains 660 images with 33 different genera and species of bacteria [11]. Figure 3 shows the selected number of data images in each class and an image from each class.

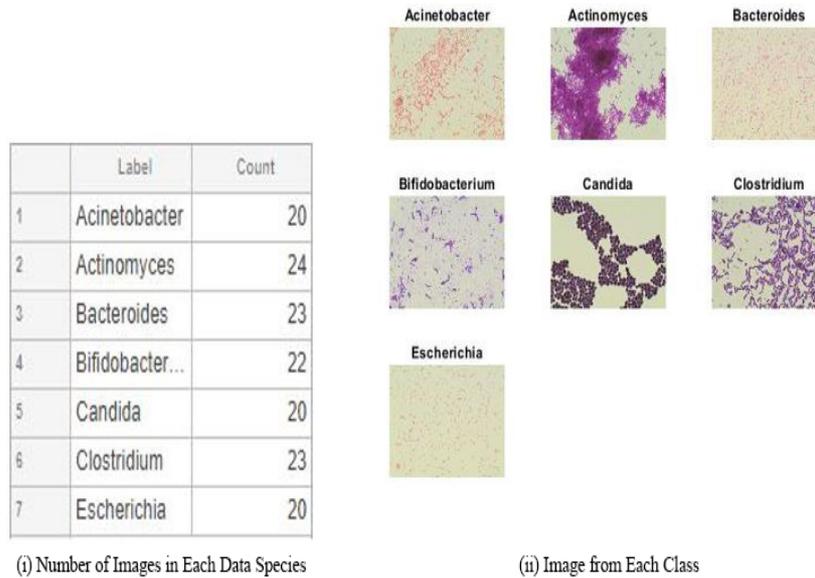


Figure 3: Data Set Images.

The data set divided randomly into 70% of the data for training and 30% for validation. The training data will be used to train the network and the validation data will be used to test the network, so that training data contains 106 images and the validation data contains 46 images. Figure 4 shows a histogram distribution of the different class labels in the training and validation sets.

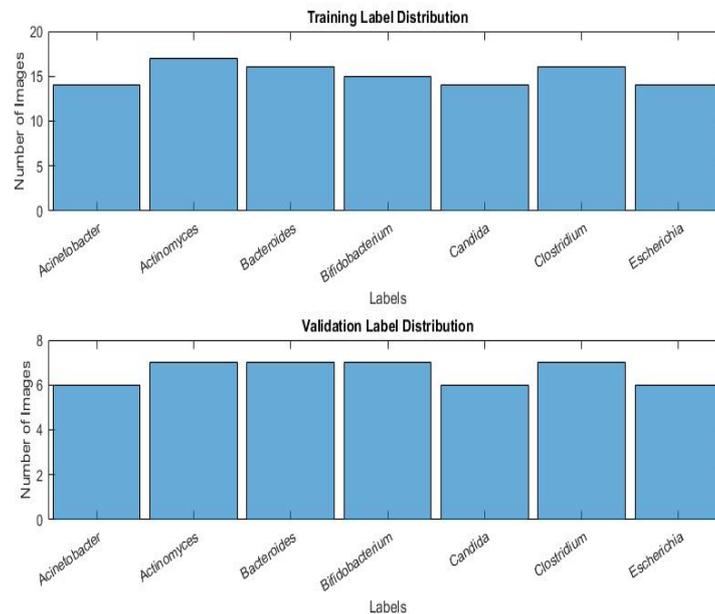


Figure 4: Distribution of Different Class Labels in the Training and Validation Data Sets.

Data Augmenter

An augmented image datastore has been used to automatically resize the training images. Additional augmentation operations have been applied on training images: randomly flip the training images along the vertical axis, and randomly translate them up to 30 pixels horizontally and vertically. Data augmentation helps prevent the network from overfitting and memorizing the exact details of the training images [7].

Network Training

The experiments were implemented with the following parameters: The maximum epoch is set to be 10 and the minibatch size is 15, wight optimization done with stochastic gradient descent algorithm with momentum.

From the training results shown in Figure 5 it can be noted that since the neural network is not trained enough (at the first training epochs), both accuracy and losses are poor. At some points, the training accuracy is significantly higher than the validation accuracy, which means at this point, the network is overfitting. In neural networks, the more training, the more improvement, both accuracy, and losses are improved after re-training, epoch after epoch, which leads to a good result after nine epochs. The training repeated several times to calculate average accuracy and time; Table 1 shows the re-training results.

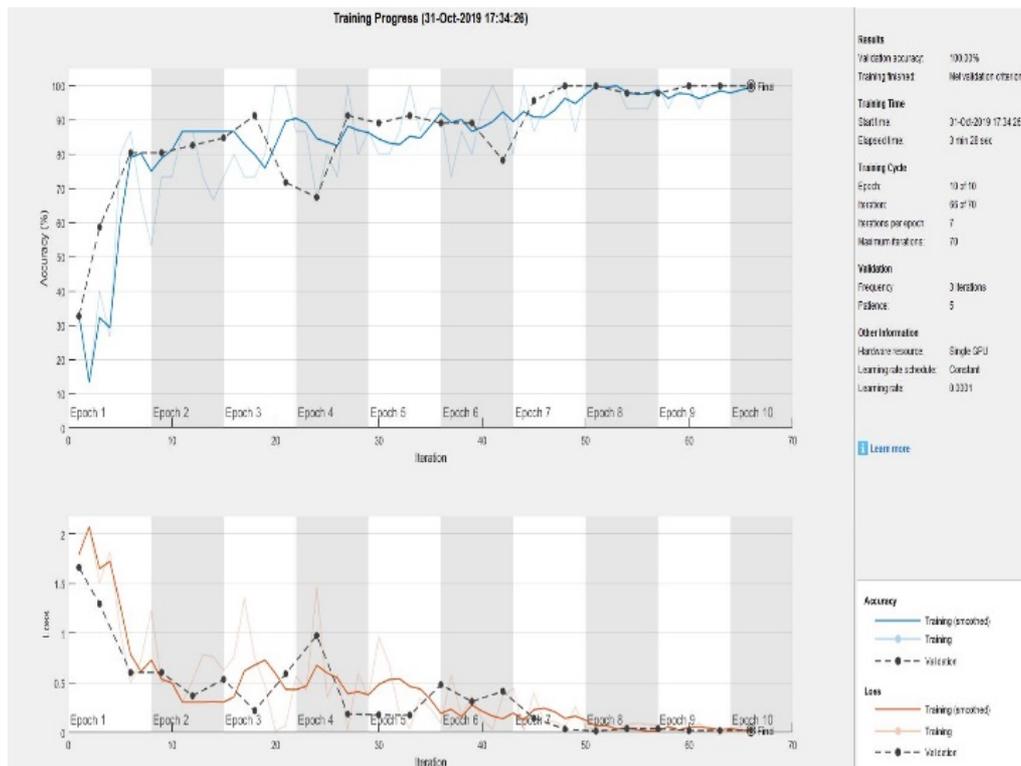


Figure 5: Training Progress.

In total, the training consists of 10 epochs and 70 iterations; each epoch consists of 7 iterations to complete the full data set. The entire network learning takes average time equal to 4 min & 27 sec and the average accuracy was 98.45%, which considered excellent and fast training compared to training a neural network from scratch.

Table1: Training progress results.

Training Progress	Completed Epochs	Mini Batch Size	Learning Rate	Accuracy	Time
1	10	15	1e-4	100%	3min&15sec
2	10	15	1e-4	100%	6min&33sec
3	7	15	1e-4	97.83%	2min&52sec

4	9	15	1e-4	95.65%	3min&32sec
5	10	15	1e-4	100%	6min&19sec
6	8	15	1e-4	95.65%	3min&28sec
7	10	15	1e-4	100%	7min&16sec
Average				98.45%	4min&27sec

Learned Features

Convolution neural networks use features to classify images. What the network learns during training is sometimes unclear, the layers towards the end of the network have larger receptive field sizes and learn larger features. The channels at the convolution layer 1 mostly contain edges and colors, which indicates that the filters at the convolution layer 1 are edge detectors and color filters, while the layers which are deeper into the network yield more detailed filters. Figure 6 shows some of the learned feature maps in the convolution layers.

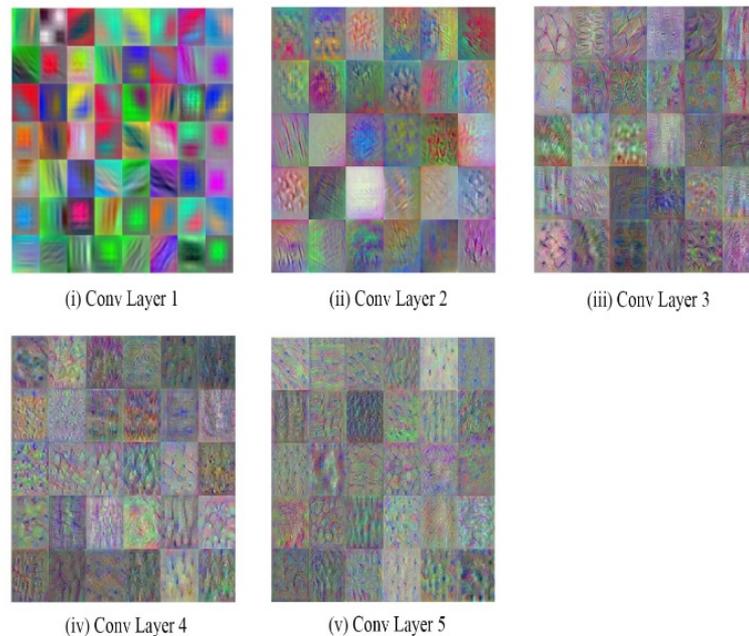


Figure 6: Convolution Layers Learned Features.

4- Classification Results

Figure 7 shows the confusion matrix of the network validation results. On the confusion matrix, the rows correspond to the predicted class (Output Class) and the columns correspond to the true class (Target Class). The diagonal cells correspond to observations that are correctly classified. The off-diagonal cells correspond to incorrectly classified observations. Both the number of observations and the percentage of the total number of observations are shown in each cell. Since the testing done with 100% accuracy model, it can be noticed that all the validation data images are classified correctly, and there's no overfitting during network classifying.

Sensitivity, Specificity, and Accuracy are the terms which are most commonly associated with a classification test and they statistically measure the performance of the test. In general, Sensitivity indicates, how well the test predicts one category and Specificity measures how well the test predicts the other category. The training has been repeated several times to measure the average sensitivity and specificity; Table 2 shows the sensitivity and the specificity of this model.

Validation Data Confusion Matrix

Output Class	Acinetobacter	Actinomyces	Bacteroides	Bifidobacterium	Candida	Clostridium	Escherichia	
Acinetobacter	6 13.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Actinomyces	0 0.0%	7 15.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Bacteroides	0 0.0%	0 0.0%	7 15.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Bifidobacterium	0 0.0%	0 0.0%	0 0.0%	7 15.2%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Candida	0 0.0%	0 0.0%	0 0.0%	0 0.0%	6 13.0%	0 0.0%	0 0.0%	100% 0.0%
Clostridium	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	7 15.2%	0 0.0%	100% 0.0%
Escherichia	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	6 13.0%	100% 0.0%
	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%
	Acinetobacter	Actinomyces	Bacteroides	Bifidobacterium	Candida	Clostridium	Escherichia	

Target Class

Figure 7: Confusion Matrix.

From the table and on average, the sensitivity tells us that 100% of the acinetobacter, bacteroides, candida, clostridium or escherichia images were correctly identified, and 93.87% of the actinomyces images were correctly identified, and 95.91% of the bifidobacterium images were correctly identified.

Table 2: Sensitivity and specificity of the model.

	Acinetobacter	Actinomyces	Bacteroides	Bifidobacterium	Candida	Clostridium	Escherichia
Avg. Sen.	100%	93.87%	100%	100%	100%	100%	100%
Avg. Spec.	100%	100%	100%	98.9%	100%	99.2%	100%

From the table and on average, the specificity tells us that 100% of the non-acinetobacter, non-actinomyces, non-bacteroides, non-candida or non-escherichia images were correctly identified, and 98.90% of the non-bifidobacterium images were correctly identified, and 99.62% of the non-clostridium images were correctly identified

5- CONCLUSION

Convolutional neural networks are biologically inspired networks that are used in computer vision for image classification and object detection. The most fundamental advantage of CNN is the automated feature extraction for a given mission. In general, when using such a network, a machine learning designer will need to think about the specific features for the task and assign each of them to an element of the input vector. On the other hand, a convolutional neural network appears to drive these features automatically. This becomes incredibly useful for tasks such as object recognition in which it is difficult for human-machine learning designers to come up with relevant features for the task.

A convolution neural network based approach has been applied for the classification of a variety of bacteria species was implemented In this paper. The proposed ConvNet architecture will automatically identify images of seven different bacteria genera and species with very high accuracy.

Also, instead of training a convolution neural network from scratch, a pre-trained model - AlexNet - was trained on a large dataset and tuned to the appropriate task. Transfer Learning is an effective important method of object recognition as it decreases training time and produces an improved overall performance.

References

- [1]. B. Zieliński et al, “Deep Learning Approach to Bacterial Colony Classification”, PLoS ONE 12(9), 2017.
- [2]. Yalcin and S. Razavi, “Plant Classification using Convolutional Neural Networks”, IEEE Fifth International Conference on Agro-Geoinformatics (Agro-Geoinformatics) 2016.
- [3]. S. Trattner et al, “Automatic identification of bacterial types using statistical imaging methods”, IEEE transactions on medical imaging. 2004; 23(7):807±820.
<https://doi.org/10.1109/TMI.2004.827481> PMID: 15250633.
- [4]. N. Blackburn et al, “ Rapid determination of bacterial abundance, biovolume, morphology, and growth by neural network-based image analysis. Applied and Environmental Microbiology”. 1998
- [5]. K. De Bruyne, et al, “Bacterial species identification from MALDI-TOF mass spectra through data analysis and machine learning. Systematic and applied microbiology, Systematic and Applied Microbiology 2011; 34(1):20±29.
<https://doi.org/10.1016/j.syapm.2010.11.003> PMID: 21295428
- [6]. M. Wani et al, “Advances in Deep Learning”, Studies in Big Data 57, Springer Nature Singapore Pte Ltd. 2019.
- [7]. A. Krizhevsky, et al,” ImageNet Classification with Deep Convolutional Neural Networks”. Communications of the ACM, 2017.
- [8]. A. Aggarwal, “Neural Networks and Deep Learning”, Springer International Publishing, 2018.
- [9]. N. Srivastava, et al, ”Dropout: A Simple Way to Prevent Neural Networks from Overfitting”, Journal of Machine Learning Research, 2014.
- [10]. L. Huang, W. Tu, “Novel neural network application for bacterial colony classification”, Theoretical Biology and Medical Modelling, 2018.
- [11]. Digital Image of Bacterial Species (DIBaS) dataset,
<http://misztal.edu.pl/software/databases/dibas/> (accessed on the 27th of September 2020).

An Extensive Study on Online and Mobile Ad Fraud

Hala Shaari^{1*}, Nuredin Ahmed²

¹ h.shaari@uot.edu.ly, ² nuredin@mtit.com.ly

¹ Software Engineering Department, Faculty of Information Technology, University of Tripoli, Libya

² Control Department, Technical Computer College, National Board for Technical and Vocational Education, Libya

ABSTRACT

The advertising ecosystem faces major threats from ad fraud caused by artificial display requests or clicks, created by malicious codes, bot-nets, and click-firms. Currently, there is a multibillion-dollar online advertisement market which generates the primary revenue for some of the internet's most successful websites. Unfortunately, the complexities of the advertisement ecosystem attract a considerable amount of cybercrime activity, which profits at the expense of advertisers. Web ad fraud has been extensively studied whereas fraud in mobile ads has received very little attention. Most of these studies have been carried out to identify fraudulent online and mobile ads clicks. However, the identification of individual fraudulent displays in mobile ads has yet to be explored. Additionally, other fraudulent activity aspects such as hacking ad-campaign accounts have rarely been addressed. The purpose of this study is to provide a comprehensive review of state-of-the-art ad fraud in web content as well as mobile apps. In this context, we will introduce a deeper understanding of vulnerabilities of online/mobile advertising ecosystems, the ad fraud's well-known attacks, their effective detection methods and prevention mechanisms.

Keywords:

Online Ad Fraud,
Mobile Ad Fraud,
Advertising Ecosystem,
Detection,
Prevention.

1- Introduction

World wide web today provides consumers with a versatile and easily available platform for listing and viewing advertising compared to more conventional media such as newspapers and printed booklets. Digital advertising has grown into one of the world's largest and most lucrative industries. It is one of the key methods to produce revenues from digital media (e.g., websites and mobile apps) by providing advertisements to customers. Online advertising is unfortunately a fractured industry with a highly inefficient business model which is vulnerable to fraud and abuse. Fraudsters have been able to leverage many vulnerabilities of the online advertisement model and have begun to manipulate the program to make a profit. Such fraudulent activities are popularly known as Advertisement Fraud (Ad Fraud), also known as Invalid Traffic.

Digital advertising is a cornerstone of finance that funds free internet content and services, and free smartphone applications. At a high level of online advertising, the basic concept is to sell space on web pages and apps for advertising. The mechanisms and infrastructure required for online ads are indeed highly diverse and complex. Since web and mobile ads use similar infrastructure, they face the same security issues related to them.

The ad ecosystem can be partitioned into three groups approximately: advertisers, publishers and intermediaries. Advertisers pay publishers to place a specified amount of creative content on websites and applications with embedded links such as text, display or video ads. Intermediaries (e.g. ad servers/networks and ad exchanges) are also used to enable communication between

advertisers and publishers. Intermediaries usually charge advertisers a fee for ad placements and/or ad clicks to publishers. What is instantly evident from this basic explanation is that revenues from publishers and intermediate platforms are directly related to the number of regular visitors to a website or app.

Therefore, publishers and intermediaries are strongly incentives to use whatever means are available to push user traffic to publisher pages. However, there is another form of fraud that does not involve the publishers. Dishonest advertisers aim to simulate demands on their competitor's ads to deplete the advertising budgets of their competitors. These methods for traffic generation have emerged, many of which are deeded as fraudulent by advertisers and intermediaries.

The aim of this research is to comprehensively review recent ad fraud activities in web content and also mobile apps. We intend to provide a better understanding of vulnerabilities of online/mobile advertising ecosystems and the well-known attacks. In Section 2, we first address the online advertising ecosystem model and mobile ecosystem model, and then discuss different revenue models for them. Online/Mobile workflows have also been introduced. The remainder of the paper is organized as follows. A comprehensive description of Ad Fraud types in Online/Mobile advertising provided in Section 3 and 4. Then, Ad Fraud existing prevention mechanisms presented in Section 5. Finally, we conclude the paper in Section 6.

2- Understanding Ad Ecosystem

This section provides an overview of the online digital advertising ecosystem and its main components for better understanding fraud activities, their major characteristics, and corresponding detection mechanisms, which will be described in the following Sections. Then, we describe a number of advertising types pertinent to our discussion. Both a business model and technical framework for delivering advertisements delivering advertisements to publisher websites and apps will introduce.

Main components

According to [1][2] Online ads primarily include four agents:

- Publisher is an organization that publishes content or provides a service through a website or a mobile app.
- Advertiser is an organization that pays the ad networks to show its ads.
- Users are any visitors to the website of the publisher who might be interested and then click on the advertiser.
- Ad networks/servers are companies that manage publishers and advertisers. They are able to buy and sell ad traffic (in the form of ad requests) internally as well as through other ad networks.

In additional to these components; Ad exchange is another prime agent to facilitate the purchase and selling of inventories in real time from various ad networks. An ad exchange serves as a broker to connect buyers and sellers to exchange information for them, enabling buyers and sellers to negotiate rates and deliver ads to end devices in real time.

Types of online advertising

There are many types of online advertising. For example, display-based advertising, search-based advertising, social media marketing, email advertising, chat advertising, classified advertising, affiliate marketing, and content marketing. Related advertising types to the purpose of this paper, have been described as follow:

- Display-based advertising. Display-based advertising visually transmits its promotional messages using text, logos, animations, videos, photographs or other graphics [3]. Display advertisers also

target users with different characteristics to increase the impact of the ads. Online advertisers (typically through their ad networks) Cookies, which are specific identifiers of particular computers, are often used to determine what advertisement to serve a particular user. Cookies will monitor when a user has left a page without purchasing something, so that the advertiser can later retarget ads to the user.

- Search engine marketing (SEM). Search engine marketing, or SEM, is designed to boost the visibility of a website in search engine results. Search engines have sponsored results as well as organic (natural or non-supported) results based on question from a web searcher. Search engines also use visual indicators to separate the sponsored results from organic results. Search engine marketing includes all of an advertiser's actions to make a website's listing more prominent for topical keywords [3].

- Mobile advertising. Advertisements are delivered through mobile devices such as smartphones, tablets or other smart devices (e.g., smart TVs). Mobile advertisements can take the form of static or rich media display ads, SMS (Short Message Service) or MMS (Multimedia Messaging Service) ads, mobile search ads, advertising within mobile websites, or ads within mobile applications or games. Mobile advertising is growing rapidly.

- Social Media advertising. Several social networks show advertisement driven business models. If we take for granted that a social network manager is a huge database, with vast amounts of qualitative data from its users, using those helps brands to micro-segment their promotional activities. Facebook is segmentation king. Its advertising network, Facebook Ads, enables the target audience of each ad to be delimited according to location, age, sex, languages and even interests and behaviors. Facebook is the social network with more data on its users. Facebook advertisement formats are displayed either on the sidebar on the platform's own right or on the user's own timeline (Web and mobile), as well as on the logout page [4]. Facebook advertisement formats are displayed either on the sidebar on the platform's own right or on the user's own timeline (Web and mobile), as well as on the logout page. And they seek to improve interaction, lead users to a website or exclusive deals, get more pages likes, download apps.

- E-mail advertising. E-mail marketing is an online marketing technique which uses email to submit commercial or advertisement information. This is a communication device for attracting new customers or keeping those already loyal to the brand. E-mail is currently the first Internet service to go along with social media. This volume of traffic includes legitimate e-mails and spam. The word "spam" applies to those messages we do not ask for and we do not want them or from an unknown sender, usually sent via mass mail. Although though spam can be used on other channels and apps, such as SMS on cell phones, the most important medium for this activity is e-mail.

Revenue Models

For online advertising or mobile application with advertisements so they can make money through those advertisements where revenue is typically determined by the amounts of impressions and/or clicks. The following revenue models [5] are generally used:

- Cost per mile (CPM) is that advertisers charge publishers with ad networks per thousand impressions. It is often referred to as the cost per thousand (CPT), because it calculates the cost per thousand views of the ads. This measurement is commonly available for advertisements for Android developers.

- Cost per click (CPC) is that the advertisers charge the users per click. It is used for the amount of times a website visitor or user clicks on a banner in an application. This measurement is also common because it can be carried out in a simple way.

- Cost per action (CPA) is that advertisers charge per specific action such as filling a form, signing up for an offer, completing a survey, or downloading software. This model seems beneficial to

advertisers since they only pay for concrete acts directly related to their advertising. Nevertheless, when it comes to complex behavior, implementation is not simple.

Online Ad Workflow

Online advertising is a very complicated process that starts when an internet user visits a web page and sends the web server an HTTP request. It will cause an Ad impression if there is any Ad banner on the requested web page, handled by Ad network/servers. When a user submits an HTTP request to the publisher web server for access to the site, e.g. a Web page containing one or more banners, the web server must contact its Ad network/server to request advertising information to be inserted into the site page. The Ad network/servers must contact server-side platform in most cases (or in real-time bidding scenarios) to prepare for bidding. Server-side platform is essentially an integrated programmatic technology platform that enables ad publishers to control their inventory of advertising spaces, maximize the selling of their online media estate, etc. The main purpose of the server-side platform is to allow publishers to link their inventory to multiple ad exchanges, demand side platforms and networks at once, so that publishers can optimize their benefit and monitor their inventory's selling price for different advertiser groups. This maximization could be overcome by the use of a data management platform that offers server-side platform data support. The server-side platform is now ready to send an auction to selected Ad exchanges after the requisite planning [6].

The details included in the bid request enables Ad Exchange and demand side platform to understand the meaning of the Ad banners to be offered to the audience. IAB OpenRTB specification [7] defines the actual format of the bid request for programmatic ad buying / selling. When an Ad exchange receives an Ad request from publishers, the request will be distributed as an Ad auction to all demand side platforms connected to the exchange. Demand side platform therefore provides broad access to inventory and vertical and lateral targeting, with the ability to serve advertising, bid on advertising in real time, monitor ads and maximize revenue. A demand-side platform will also rely on data management platforms for this purpose to provide efficient data support. Once the Ad exchange receives bid responses from all demand side platform within the time limit (typically less than 100ms from the bid), the one with the highest bid price will be chosen. The advertiser who wins the bid will transfer the details, such as the advertising URL along with the script code, to the server-side platform and then to the Ad network/server publisher. The web content server publisher and the ad network/server then react to the internet user, providing the client devices with the website and advertisements. The above procedure completes a single Ad transaction and this operation happens in real-time with less than 100ms delay in effect. And consumers are not experiencing frustrating latency.

Mobile Ad Workflow

When advertisers choose to promote a product(s), they contact an Ad network/server authority and supply the material to potential clients. The Ad network/server stores these contents and provides custom APIs to the publishers. If a publisher wants to generate revenue using online ads from their website or application, it signs up with an Ad network/server and accesses its APIs to allow mobile advertising. These APIs usually contain a UI component used to view the advertising and the publisher places the UI component in its GUI. When a user initially opens the website or the publisher's application, the Ad API loads the ad UI component with the content of the advertiser and this event is logged as an impression. Now if the user clicks on the Ad component (a click event) another request will be sent to the Ad network/server to redirect the user to suitable URLs [8]. Ad network/servers keep track of each such event. The advertiser must pay the Ad network/server on the basis of their sales model and the publisher earns a certain percentage of the Ad network/server payment. Among the revenue models, Pay-per-click and Pay-per-impression models are most popular.

3- Ad Fraud Types in Online Advertising

By definition, Ad fraud is synonymous with an activity in which views, clicks, acts or data events are misreported to criminally gain revenue, or for other purposes of deceit or malice. Ad-fraud activities aimed at generating revenue are more common, but noise generation and other non-revenue generating activities are also present today in the internet advertisement ecosystem. There are three main types of ad fraud: placement fraud, traffic fraud, and online advertisement action fraud at different rates, showed in figure 1. Reporting in each of these cases validates a visitor to be genuine, but is simply fraudulent. Such fraudulent tourists can be absolutely robotic, human or a mixture of both. This section includes a thorough analysis of the forms of Ad fraud. In addition, for each type of fraud, we will also review detection methods.

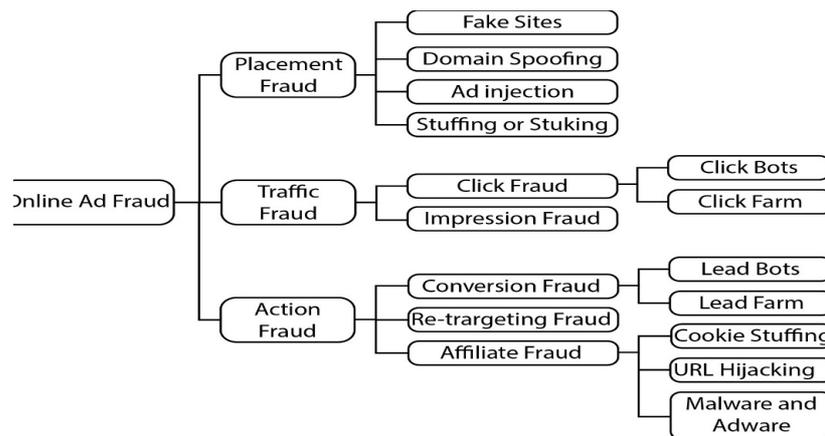


Figure 1: Online Ad Fraud Types

Placement Fraud

An Ad placement is always an iframe containing advertisements that contain messages, pictures or videos that are innovative material. Placement fraud is described as fraudulent acts or activities aimed at manipulating or changing the web pages of publishers or changing the web pages displayed on the devices of the user to increase impressions or clicks. Such fraudulent activities involve a variety of acts, ranging from simple keywords stuffing [1], misrepresenting Ad placement position so that a placement is positioned in the invisible frames and is never visible to the public [9] to Malvertising [10] which injects advertisement malware by attracting users to register and then redirecting traffic to malicious sites to generate inflated impressions. According to [6] placement frauds have been categorized into four groups, with each group focusing on one aspect of fraudulent actions, and review solutions to detect each type of placement fraud.

Stuffing or Stacking: Stuffing (whether keyword stuffing or pixel stuffing) is a way to show content that cannot be seen with bare eyes. It is widely used both for keyword stuffing [1] and for placement stuffing. In order to detect stuffing or stacking fraud, Double Verify Inc. [11] suggests many methods of detecting hidden or invisible ads in order to prevent stuffing or stacking fraud. One approach is to compare the ads with graphic images taken from html codes with this web page's snapshot. Any advertisement with the image which is not contained in the snapshot will be labelled invisible using image recognition technology [12]. Another approach is the geometrical analysis. The code snippet contained in the page must measure the position of the advertising, the position of the browser's viewable areas and the size of the open browser window to check whether the advertising is viewable.

Fake Sites: There are two distinct types of fake website fraud. One is to build websites with legal domain names but only contain Ad slots[3]. Then, by entering large Ad networks or Ad exchanges, fraudsters may get significant revenue from fake sites. Another way to deceive tourists is to copy

the content from well-known websites or register a similar domain name. There are two popular ways for detecting fake sites, namely blacklist lookups and recognizing fake sites using machine learning methods. Nearly every browser toolbar contains a blacklist to verify whether the site the user visits is fake or not [13]. Using data mining and machine learning methods to identify fake pages, Abbasi and Chen [14] proposed a classifier system.

Domain Spoofing: Web spoofing [6] is widely known on the Internet, where fraudsters create websites that mimic real websites to conduct fraud activities, for example, stealing identity information or account credentials. Since advertisers are willing to put their advertising on better quality sites and even offer higher rates, fraudsters spoof their domains in order to avoid being placed on the blacklists. Domain spoofing therefore refers to fraudulent activities which attempt to falsify the domain as if the traffic were from publishers in the whitelist. Domain spoofing is normally done in Ad networks via Malware & Tool-bars [15] or Ad Tag Misrepresentation [16].

Ad Injection and Malware: Ad injection and malware are more aggressive fraud activities that explicitly impact client web browsers to either change the ad [17] or view advertising on the current web page that have no Ad placement at all. One form of Ad injection comes from advertisement software, such as malicious adware, a program running on client computers to show unintended ads [18]. Additionally, ad injection can also be done via browser extension software which is often used to increase / enhance web browser functionality [19]. Thomas et al. [17] developed a detection method based on client-side DOM with Google websites to get Ad injection scripts from HTML pages. They search the customer's DOM first to classify suspect elements and fraudulent domains. They then filter scripts that are not affected by ad injection, by excluding usual programs such as browser toolbars and antivirus engines. Finally, they update the scripts manually depending on the quality of the scripts.

Traffic Fraud

The main objective of traffic fraud is to inflate the number of impressions produced by individual sites or placements by manipulating network traffic. However, for CPC-based campaigns, only the users' click action on the advertising displayed will result in a revenue, so, click fraud [20][21] is often widely seen and is one of the most prevalent fraudulent behaviours.

Impression Fraud: Impression fraud aims to increase the website traffic directly and thereby produce more impressions for auction. This form of fraud has the greatest effect on the CPM-based campaigns, since inflated impressions give advertisers little to no value for their advertising benefits [1]. In addition, it also affects campaigns focused on CPC and CPA, since most impression fraud cannot lead to click or conversion events, and therefore the click-through rate (CTR) will be reduced, because the CTR measurement denominator is the number of page views. In fact, fraud in impression is created through three approaches: hiring human labour to manually view pages, design different types of bots to generate impressions for auctions, and using expired domains to divert users to third-party pages. Since hiring human resources is seen as too costly to produce web traffic, whereas bot also has less intellectual capacity to mimic human trafficking, several hybrid methods seek to increase traffic on websites by incorporating really human behaviour and automated bot functions. For example, publishers gain impressions invisibly in pay-per-view (PPV) networks [22]. If a user views one publisher of this network and clicks anywhere on this page, an invisible frame with other publishers will be activated. The authors [22] implemented three countermeasures, i.e. filtering zero-sized viewports, blocking traffic from PPV networks using referral blacklists, and stopping running advertisement on publishers in blacklist.

Click Fraud: On any advertisement, a click event is a simple indication that a viewer is potentially interested in an advertisement and can thus become a customer. Tap through rate (CTR) is most also used to determine effectiveness at various rates, such as at the placement level, site level, or publisher level etc. Click fraud is probably the most prevalent fraud in the Ad ecosystem, mainly because campaigns focused on CPC dominate Ad networks. Fraudsters use various types of

approaches in a click fraud attack, either manually or with bots, to click on an advertisement. Two parties, publishers and advertisers can root a click fraud, with two motives, respectively, publisher click inflation or advertiser competition [23]. Publisher Click Inflation: Clicking events will carry immediate revenue from the publisher perspective, since publishers are paid on the basis of the percentage of advertising experiences that viewers click. Publishers therefore intuitively accept click fraud attacks, though they do not promote or engage in these activities. To do so, clicks are created either through the use of automated programs or human labours.

Advertiser Competition Clicks: Under the CPC revenue model, a small amount of advertisement budget is consumed per click. So, by making artificial clicks on the advertisements of the competitor, the advertising budget of the competitor may be depleted within a limited period of time. As a result, fraudulent advertiser advertising will have the benefit of targeting legitimate users with a higher chance of user clicking and resulting in a better conversion rate and pleasing the branding company. Some advertisers use a pacing rate control to defend click attacks which specifies daily or hourly advertisement spending cup for smooth budget delivery [24]. It will keep the whole campaign budget from running out in a limited time, whereas clicks of poor quality with no commercial value bring extra pressure to every campaign. Two widely used methods to produce fake clicks are click farms or click bots [25], Where the former is created by human viewers and the latter generated by computer programs.

Action Fraud

Action fraud aims to target relevant business activities of users, such as filling out an online form or survey, placing an online purchase order, or re-targeting important customers through previous acts or behaviours of users [14]. Since advertisers are focused on using cost-per-action (CPA) to determine their promotional costs versus revenue, action fraud affects the ad pricing, campaign preparation and several other major components of the Ad ecosystem directly.

Conversion Fraud: An Ad network conversion refers to one or a series of concrete business actions taken by site visitors that they convert to paid (or future paying) customers. Alternatively, a conversion may also be described as "agreed-upon user action" [26]. For instance, a simple conversion event may be to download a file or fill out a form, or complete an online purchase order. This fraud is also called spam conversion [1]. A conversion usually involves a variety of user activities, and generally a conversion event takes place minutes, hours, or even days after the initial Ad click. It should be noted that a conversion is usually tracked via the placed pixel on the branding sites (or the landing page), while usually a click is tracked on the sites of the publishers. The mapping is also achieved by matching details about the user cookies. A purchase is needed for most clicks via landing pages-based conversions, so users need to provide name, credit card and other important details. As this phase involves sophisticated interactions and financial engagement, there are very few fraudulent activities aimed at this kind of conversion. For conversions based on lead generation landing page users are only required to provide simple details or to take simple actions, such as filling in user name, household address, or downloading a file from the advertisers' site. All these acts can be carried out at minimal or nearly no financial cost. Most conversion fraud thus targets this sort of conversion. Conversion fraud also occurs across two forms of activity [27]: 1) Lead Bots: is a computer agent who fills out lead forms automatically with either randomly generated or partially correct information. 2) Lead Farm: In this situation, the fraudsters are able to employ people to manufacture conversions with lower labour costs from underdeveloped countries.

Re-targeting Fraud: It aims to target valuable customers with precision based on their previous Internet acts, Such as customer purchase history or site surfing history/customer activities [28]. This can be achieved by reviewing past transaction history or monitoring users ' cookies before visiting sites and showing interest in certain products to decide whether or not a user is interested in those products. Commonly, they use techniques like "cookie" or "pixel" as the snippet code. For the re-targeting of fraud, the fraudsters ' main goal is to mimic the particular activities of legitimate

consumers and make them behave like attractive users. This is typically done through the use of computer-generated agents, such as DeceptiBots [29], to mimic the thoughts and actions of a human being and believe that they are interested in a specific product or brand. As a consequence, the bots trick advertisers into believing bots are valuable potential customers, thus, put a higher price on the bots produced auctions/impressions.

Affiliate Fraud: Affiliate marketing is a kind of performance-based approach to marketing, where an associate (i.e. a business entity) collects the benefits from its marketing activities for each visitor or customer. Affiliates use various types of promotional strategies in affiliate marketing, including search engine optimization (SEO), e-mail marketing, or display advertising to attract visitors. Affiliate fraud refers to activities that mislead the framework for reporting commission / revenue not authorized by the affiliate. In most affiliate marketing, commission is only charged if a customer makes a transaction, so affiliate can claim commission only after a conversion has taken place. As a result, affiliate fraud specifically targets consumers who are already on the verge of making purchases [30]. An affiliate fraud [31] is widespread across the following three forms of approaches: 1) Malware and Adware, 2) Cookie Stuffing and 3) URL Hijacking.

4- Ad Fraud Types in Mobile Advertising

Although the literature includes a wide range of ad fraud research in web applications, relatively little attention has been given to such mobile fraud. A recent study [32] proposed a new taxonomy of Mobile fraud as showed in figure 2 that sums up nine different types of ad fraud, which is by far the largest number of types of ad fraud.

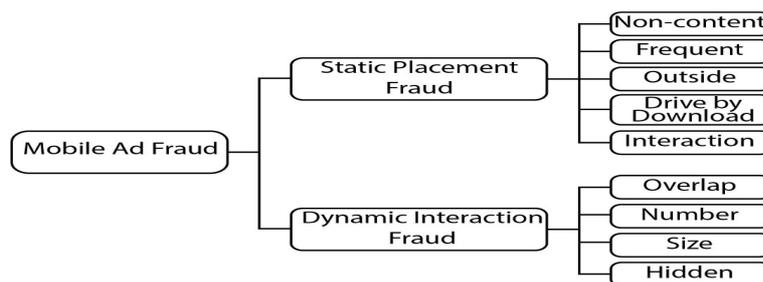


Figure 2: Mobile Ad Fraud Types

Static Placement frauds

Many cases of fraud are achieved simply by manipulating the shape and location of the ad view in a UI state. "Static" means that the identification of these frauds can be calculated by static data and takes place in a single UI state. "Placement" means fraudulent conduct exploits aspects of placement, e.g., scale, position, and number of ad views, etc. Four specific types of behaviour related to static placement frauds have been identified:

The Ad Hidden fraud: App developers can hide advertisements (e.g. under buttons) to give users the impression of an "ad-free app" that would hopefully improve user experience. However, these ads are not shown in compliance with the contract with advertisers paying for the advertising role [33][34].

The Ad Size fraud: While ad size advice given by ad networks is not compulsory, and there are no ad size requirements, the size ratio of the ad to the screen must be fair [33], allowing users to display advertisements normally[35]. By extending ad size to the limits, fraudulent activity can be implemented: with incredibly small ad sizes, app developers can have the feeling of an ad-free app,

but it can cheat advertisers; similarly, with an abnormally large ad size, there is a greater risk of attracting or pressuring users to click on the ad in an effort to close it.

The Ad Number fraud: Since advertising must be seen by users as pure extras in addition to the main content of the app, the number of ads must remain fair [34][36]. Unfortunately, developers also provide a large number of advertisements to increase the chance that user interests would be drawn, even if the user experience of the app is deteriorating and even seriously affects regular functionality when ad content exceeds valid product content.

The Ad Overlap fraud: To force users to trigger unwanted impressions and clicks, App developers can simply show ad views on top of app-relevant views [33][34][36]. By positioning ads in places covering consumer areas of interest in standard experiences with apps, App developers create irritating circumstances in which users will 'recognize' the ad.

Dynamic Interaction Frauds

Authors on [32] identified cases of fraud found that go beyond putting ad views on a single UI state, but it requires runtime behaviour, which can occur in an unpredictable scenario of application use. "Dynamic" means that such frauds are identified at runtime. "Interaction" means the fraudulent activity exploits scenarios of user interaction and may include several UI states.

The Interaction Ad fraud: Developers use interstitials in web programming (i.e. web pages shown before or after the intended content of a page) to view advertisements. Translated into mobile programming, when switching between UI states, some ad views are put. Nonetheless, manipulation can be achieved by putting interstitial ads on app load early or by exiting apps that could trick users into unintended clicks, as contact with the app / device is highly likely at this point [33][34][37].

The Drive-by downloads Ad fraud: Ads are intended to provide short advertising content created by marketers to attract the attention of users when they visit an external web page. When app developers are remunerated not by the number of clicks but by the number of users who ultimately become actual customers of the product / service advertised, there is a temptation of fraud. A common example of fraud includes causing unintended downloads (e.g. of advertised APKs) when you click on the ad view [34][37]. This activity also has a significant effect on user experience, and drive-by-downloads are not even readily cancelled in most cases.

The Outside Ad fraud: Advertisements are expected to appear on pages when users use the app. There are, however, illegal activities for ads while applications are running in the background or even outside of the user environment (e.g., Placed ad views on the home screen and covering device icons that users need to hit to launch new apps)[33][34][37][38]. In certain extreme situations, the advertisements appear spuriously and the user must identify them, because these advertisements can only be removed when the user recognizes and launches the application from which they originate.

The Frequent Ad fraud: App developers aim to increase the probability of ad impressions and clicks to gain more revenue. The number of UI states in the device limits this probability. Therefore, developers may enforce deceptive techniques by showing interstitial ads each time the user clicks on the core content of the application (e.g., even when the click is to show a menu in the same page)[33][37].

The Non-content Ad fraud: To boost ad impressions and trick users into accidental clicks, app developers may place advertisements on non-content-based pages such as thank you, error, login, or exit screens. Ads on these types of UI states may confuse a user in the assumption that the ads are actual app content [33].

A lot of research was proposed to classify ad frauds on the web. Such approaches can provide valuable tips for mobile community researchers and practitioners to devise promising methods to

detecting mobile ad fraud. Current mobile ad fraud studies have attempted to classify ad-fraud technologies where the fraudulent activities can be statically identified (the so-called static placement frauds). For examples, Pearce et al. [39] found that 49 percent of Android applications contain at least one advertisement library, and 46 percent of ad-supported applications are affected by these libraries. Shekhar et al. [40] suggested an approach called AdSplit to split applications from its advertisement libraries that could request sensitive privilege permissions. In addition, Liu et al. [41] examined Windows Phone's static placement abuse by examining app layouts. Furthermore, Crussell et al. [42] developed an automated method for the detection of click frauds. Shekhar et al. and Crussell et al. methodology is basically applied in three stages: (1) building HTTP request trees, (2) identifying ad request pages using machine learning, and (3) detecting clicks in HTTP request trees use heuristic rules. Unfortunately, the above-mentioned strategies are unable to classify the new fraudulent behaviors with the evolution of ad fraud, for example, they cannot be used to detect fraud involving dynamic interactions.

5- Ad Fraud Prevention Mechanism

To counter Ad fraud, current approaches typically depend on the following four types of mechanisms [16]:

- **Signature-based Prevention Mechanism:** This type of approach uses predefined features/patterns to identify malicious traffic or impression [43]. For example, research found that when execution of a client-side code is incompatible with established code execution models (such as JavaScript); the traffic is very likely not to be created by real human users, but by a bot [44]. Therefore, checking the execution of code environment, such as support for JavaScript or mouse event test [45], a large portion of fraudulent traffic can be filtered out. These behaviour analysis approaches for Clickbots have been studied extensively [46].
- **Anomaly-based Prevention Mechanism:** This approach uses statistical analysis and historical data to identify suspicious pages, websites or publishers whose traffic is deemed irregular compared to the general traffic of users. For example, as of April 2016, the average probability of click events in advertising display is about 0.17 percent, indicating that there are about 1.7 click events on every 1000 impressions on average [47]. A placement or publisher website that shows substantially higher clicking through rates would be considered anomaly and includes fraudulent activities that deserve further investigation [48].
- **Honeypot-based Prevention Mechanism:** To pinpoint fraudulent activities, Ad servers (such as advertisers) may purposely serve a range of carefully specified bluff ads to publishers where it is understood that bluff / honeypot ads are unidentifiable through individual users, and if bluff advertisements result in events, such as a click event, it will contradict the presumption and therefore involve fraud [49]. Traffic traffickers used such a honeypot strategy to analyse traffic for better Ad service [50].
- **Credential-based Prevention Mechanism:** The credibility of website publishers is strongly associated with possible fraud activities. To determine publisher's credentials, demand side platforms or advertisers may use reverse crawling to find the content of the web pages and test if their content is compatible with the impression-related tags when submitting an auction. However, one can also use the number of impressions produced by a publisher, and compare this value to trustworthy website rankings like Alexa or RageRank. A publisher with a much greater impression than its traffic rating would obviously suggest possible fraudulent activities.

6- Conclusion

Online advertising is now a popular form of business marketing and one of the reasons why free web content or mobile apps are available. The emergence of smart televisions and online content distribution services is expanding rapidly. Consequently, fraudsters are also exploiting the market to drain money from advertisers. This paper provides a detailed description of the current state-

of-the-art online/Mobile advertising fraud. However, more research needs to be addressed to all types of occurring ad fraud. As examples, in the context of online marketing campaigns, hacking ad-campaign accounts need to be studied. Additionally, to explore individual fraudulent displays in mobile advertising as well as dynamic interaction frauds that has been pointed out in this study.

References

- [1]. N. Daswani, C. Mysen, V. Rao, S. Weis, K. Gharachorloo, and S. Ghosemajumder, "online advertising fraud crimeware: understanding new attacks and defenses," 2008, vol. 40, no. 20, pp. 1–28.
- [2]. R. Oentaryo *et al.*, "Detecting Click Fraud in Online Advertising: A Data Mining Approach," *J. Mach. Learn. Res.*, vol. 15, no. 3, pp. 99–140, 2014, [Online]. Available: <http://jmlr.org/papers/v15/oentaryo14a.html>.
- [3]. D. Chaffey and F. Ellis-Chadwick, *Digital Marketing: Strategy, Implementation and Practice*, 1st ed. Pearson Education, 2012.
- [4]. T. Pineiro-Otero and X. Martinez Rolan, *understanding digital marketing basics and actions*. Springer-Verlag GmbH, 2016.
- [5]. N. Daswani, C. Mysen, V. Rao, S. Weis, K. Gharachorloo, and S. Ghosemajumder, "online advertising fraud," 2007.
- [6]. X. Zhu, H. Tao, Z. Wu, J. Cao, K. Kalish, and J. Kayne, *Fraud Prevention in Online Digital Advertising*, 1st ed. Springer Publishing Company, Incorporated, 2017.
- [7]. B. IA, "OpenRTB API Specification Version 2.5." Dec. 2016.
- [8]. G. Cho, J. Cho, Y. Song, D. Choi, and H. Kim, "Combating online fraud attacks in mobile-based advertising," *Eurasip J. Inf. Secur.*, vol. 2016, no. 1, pp. 1–9, Dec. 2016, doi: 10.1186/s13635-015-0027-7.
- [9]. B. Edelman, "Accountable? The Problems and Solutions of Online Ad Optimization," *IEEE Secur. Priv.*, vol. 12, no. 6, pp. 102–107, Nov. 2014, doi: 10.1109/MSP.2014.107.
- [10]. A. K. Sood and R. J. Enbody, "Malvertising exploiting web advertising," *Comput. Fraud & Security*, vol. 2011, no. 4, pp. 11–16, 2011.
- [11]. M. McLaughlin, R. K. Rosenfeld, L. M. Abu, and L. Simon, "system and method for identifying hidden content." Google Patents, 2015.
- [12]. Y. Zheng, B. Jeon, D. Xu, Q. M. Wu, and H. Zhang, "Image segmentation by generalized hierarchical fuzzy C-means algorithm," *J. Intell. Fuzzy Syst.*, vol. 28, no. 2, pp. 961–973, 2015.
- [13]. Y. Zhang, S. Egelman, L. Cranor, and J. Hong, "Phinding phish: Evaluating anti-phishing tools," 2007.
- [14]. A. Abbasi and H. Chen, "A comparison of tools for detecting fake websites," *Computer (Long Beach, Calif.)*, vol. 42, no. 10, pp. 78–86, 2009.
- [15]. innovation in magazine media 2015-2016, "advertising: digital advertising fraud." .
- [16]. B. Stone-Gross, R. Stevens, A. Zarras, R. Kemmerer, C. Kruegel, and G. Vigna, "Understanding fraudulent activities in online ad exchanges," in *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*, 2011, pp. 279–294.
- [17]. K. Thomas *et al.*, "Ad injection at scale: Assessing deceptive advertisement modifications," in *2015 IEEE Symposium on Security and Privacy*, 2015, pp. 151–167.
- [18]. S. Emily, "New adware apps bug facebook, google." 2011.
- [19]. X. Xing *et al.*, "Understanding malvertising through ad-injecting browser extensions," in *Proceedings of the 24th international conference on world wide web*, 2015, pp. 1286–1295.
- [20]. K. C. Wilbur and Y. Zhu, "Click fraud," *Mark. Sci.*, vol. 28, no. 2, pp. 293–308, 2009.
- [21]. Q. Zhang, T. Ristenpart, S. Savage, and G. M. Voelker, "Got traffic? An evaluation of click traffic providers," in *Proceedings of the 2011 Joint WICOW/AIRWeb Workshop on Web Quality*, 2011, pp. 19–26.
- [22]. K. Springborn and P. Barford, "Impression fraud in on-line advertising via pay-per-view networks," in *Presented as part of the 22nd USENIX Security Symposium (USENIX Security 13)*, 2013, pp. 211–226.
- [23]. N. Vratonjic, M. H. Manshaei, and J.-P. Hubaux, "advertising fraud," 2011.
- [24]. K.-C. Lee, A. Jalali, and A. Dasdan, "Real time bid optimization with smooth budget delivery in online advertising," in *Proceedings of the Seventh International Workshop on Data Mining for Online Advertising*, 2013, pp. 1–9.

- [25]. P. Pearce *et al.*, “Characterizing large-scale click fraud in zeroaccess,” in *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, 2014, pp. 141–152.
- [26]. B. Mungamuru and S. Weis, “Competition and fraud in online advertising markets,” in *International Conference on Financial Cryptography and Data Security*, 2008, pp. 187–191.
- [27]. Performics, “digital advertising fraud and abuse: strategies and recommendations for mitigation.” 2014.
- [28]. Z. Li, K. Zhang, Y. Xie, F. Yu, and X. Wang, “Knowing your enemy: understanding and detecting malicious web advertising,” in *Proceedings of the 2012 ACM conference on Computer and communications security*, 2012, pp. 674–686.
- [29]. P. Nowak, “Deceptibots: when machines go bad,” *New Sci.*, vol. 214, no. 2870, pp. 45–47, 2012.
- [30]. B. Edelman and W. Brandi, “Risk, information, and incentives in online affiliate marketing,” *J. Mark. Res.*, vol. 52, no. 1, pp. 1–12, 2015.
- [31]. P. Snyder and C. Kanich, “No Please, After You: Detecting Fraud in Affiliate Marketing Networks.” 2015.
- [32]. F. Dong *et al.*, “Frauddroid: Automated ad fraud detection for android apps,” in *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, 2018, pp. 257–268.
- [33]. Google, “google admob & adsense policies.” 2017.
- [34]. Wandoujia, “wandoujia (ali app) developer policy.” 2018.
- [35]. Firebase, “banner ads.” 2017.
- [36]. DoubleClick, “doubleclick ad exchange program policies.” 2017.
- [37]. C. C. S. Association, “mobile intelligent terminal malicious push information to determine the technical requirements.” 2017.
- [38]. H. Market, “huawei market app developer policy.” 2018.
- [39]. P. Pearce, A. P. Felt, G. Nunez, and D. Wagner, “addroid: privilege separation for applications and advertisers in android,” in *Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security*, 2012, pp. 71–72.
- [40]. S. Shekhar, M. Dietz, and D. S. Wallach, *Separating Smartphone advertising from applications*. Rice University, 2012.
- [41]. B. Liu, S. Nath, R. Govindan, and J. Liu, “ $\text{\$}\{\text{\$}\text{DECAF}\}\text{\$}$: Detecting and Characterizing Ad Fraud in Mobile Apps,” in *11th $\text{\$}\{\text{\$}\text{USENIX}\}\text{\$}$ Symposium on Networked Systems Design and Implementation ($\text{\$}\{\text{\$}\text{NSDI}\}\text{\$}$ 14)*, 2014, pp. 57–70.
- [42]. J. Crussell, R. Stevens, and H. Chen, “Madfraud: Investigating ad fraud in android applications,” in *Proceedings of the 12th annual international conference on Mobile systems, applications, and services*, 2014, pp. 123–134.
- [43]. V. Dave, S. Guha, and Y. Zhang, “Measuring and fingerprinting click-spam in ad networks,” in *Proceedings of the ACM SIGCOMM 2012 conference on Applications, technologies, architectures, and protocols for computer communication*, 2012, pp. 175–186.
- [44]. S. K. Adrian Neal, “quantifying online advertising fraud: ad-click bots vs humans,” 2015.
- [45]. H. Xu, D. Liu, A. Koehl, H. Wang, and A. Stavrou, “Click fraud detection on the advertiser side,” in *European Symposium on Research in Computer Security*, 2014, pp. 419–438.
- [46]. B. Miller, P. Pearce, C. Grier, C. Kreibich, and V. Paxson, “What’s clicking what? techniques and innovations of today’s clickbots,” in *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, 2011, pp. 164–183.
- [47]. C. Dave, “us, europe and worldwide display ad clickthrough rates statistics summary.” Apr. 2016.
- [48]. F. Yu, Y. Xie, and Q. Ke, “SBotMiner: Large Scale Search Bot Detection,” in *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, 2010, p. 421430, doi: 10.1145/1718487.1718540.
- [49]. H. Haddadi, “Fighting Online Click-Fraud Using Bluff Ads,” *SIGCOMM Comput. Commun. Rev.*, vol. 40, no. 2, pp. 21–25, Apr. 2010, doi: 10.1145/1764873.1764877.
- [50]. blog TT, “using honeypot banners to detect click fraud.” 2015.

Design of New Secret Key to Increase the Security of LSB Algorithm

Jalal M. Mehalhal^{1*}, Adema.bensaid², Mohammed F. Ighbeeshah¹

¹ Enjjalal@yahoo.com, ² A.bensaid@uot.edu.ly, ¹ abuabdualaziz2240@gmail.com

¹ High Institute-Alkhoms, Dept. of Computer Engineering, Alkhoms–Libya

² Department of Electronics, College of Science, Tripoli University, Libya

ABSTRACT

Nowadays, information security has become a big challenge for the world due to the rapid growth of Internet users day after day. Unauthorized access to confidential data can have serious implications such as financial loss, etc. One of the best techniques for secure communication is secret writing. Hiding data is very important nowadays as data travels over multiple insecure networks. To avoid this problem, encryption is used that hides data, but in some cases encryption cannot provide full security because the message is still available for encryption analysis. Encryption focuses on making the message unreadable to any unauthorized person who might intercept it. On the other hand, hiding information is a means of hiding the existence of a message to allow secure communication in a completely undetectable manner. Hide information and encryption are two different ways to hide data.

In this paper the researcher suggests how to hide the message using the least significant bit algorithm inside an image and encrypt it in a new way, by modifying the DES algorithm, the researcher generated subkeys from the DES algorithm and used them to specify the masking mechanism in the digital image.

Keywords:

cryptography,
LSB,
image,
pixel,
RBG color

1- Introduction

Information security science has become a subject of great interest by researchers and interested people who are trying to obtain new and updated solutions and technologies to ensure the protection of the information that is sent and received over the global network of information without any penetration or disclosure by intruders.

Therefore, it was necessary to keep abreast of the development of information security and the establishment of advanced technologies and means. Hence the science of concealing information and evolved by adopting the technique of concealment. Masking technology is one of the protections that make data sent and received invisible, by hiding specific messages within a specific cover. The goal of the concealment process is to not raise any point of doubt about the existence of hidden data, while the goal of the concealment analyst is to doubt all messages sent, and to check them to ensure that there are hidden data in them. The process in which a party tries to detect, read, change, or delete hidden information is called a decryption process.

Hence the need to find multiple means, for the purpose of communicating information and data in a correct and protected way, from the information of non-authorized parties to access this information, so the science of cryptography appeared as it is the science that is concerned with the methods of data protection and transmission in a wide field, and these methods depend on a secret key used to encrypt the data.

Although encryption is a good way to protect information, it is easy to discover and any intruder can manipulate, so the need for a more sophisticated, more confidential technology and information preservation, especially with the emergence and development of the global network of information.

The coverage system has been resorted to, because seeing the data in its encrypted form is sufficient to cause the intruder or the attacker to believe that there are important or sensitive data that lie in

the randomness or in the encrypted text, so he starts using anti-encryption techniques to try to get its content, and even if he is unable to achieve that, he might tamper with it, distort it, or use some available means to prevent it reaching its goal.

The major and major challenge that the information security field has faced is the emergence of computer networks and means of communication in order to store, enter and supply information internally within organizations and externally to and from remote host devices. So a new expression has been added to information security which is network security, which is defined as the correct protection of all components related to the computer network, including data, communication tools, and infrastructure.[1]

1.1 Image processing

Color image models With the color format, a digital image can record and provide more information than the gray scale format image does. Digital acquisition devices (such as scanners and digital cameras) can separate beams of light into three primary colors- red, blue, and green, through the assistance of spectrosopes and filters. In order to record the color information, we need at least three parameters (e. g. red, blue, and green) to represent a color. We use the color model to represent the color information of digital images. Since we need three parameters to represent a color, those color models must be with a three dimensional format. The models use some mathematical functions to represent a point position (in the three dimensional space) that is assigned to a color. Some color models (RGB, CMY, YIQ, HSI, l1_l2_l3, and L*a*b)[8].

1.2 Cryptography

Cryptography is synonymous to encryption. Encryption is a process of converting plaintext (data or message) into ciphertext (encrypted text). Since ancient times it has been a practice to try and send a message to your allies without it being picked up an adversary. It is used to provide the strategic leverage needed in order to prevent any unwanted person from gaining any intelligence that might jeopardize the entire mission, for example Julius Caesar used to send messages to his generals in a coded format and the generals would decipher it using a key that only they had access to hence giving birth to “Caesar cipher”. This has been an essential element in the art of war and a way of communicating war strategies even during World War I. These “keys” or ciphers usually follow a logical pattern or an algorithm but with the advent of the age of information technology, these methods started becoming obsolete as major algorithms that were earlier known to only an elite group became common knowledge and software’s enhanced the process of guessing and cracking the algorithm patterns both previously known or newly created.[6]

Cryptography acts as a shield for the data by keeping it safe from changes and pilferage and as a result has become a prerequisite when data needs to be transmitted via any public medium especially the internet or any other network. Cryptography can also be used as a tool for user verification, but can also be used for user authentication. Encryption is a process of converting plaintext into ciphertext. Decryption is a process where encrypted text or ciphertext is converted back into original message.

The following three cryptographic schemes are very popular typically used with some Cryptographic algorithm [7].

- a) Secret key (or symmetric) cryptography
- b) public-key (or asymmetric) cryptography and
- c) Hash functions.

Secret key cryptography is also known as symmetric key cryptography or private key cryptography. In this scheme both sender and receiver uses the same key to encrypt and decrypt data. The use of a secured channel in order to exchange the key in the process of symmetric encryption lowers its utility usefulness. So the main problem arrives when the keys are to be exchanged. If user wants to communicate with different people with separate confidentiality level he has to use different number of keys for each individual. If there is a group containing ‘N’ number of people. Who are using secret-key cryptography scheme, then it is mandatory to administer a number of keys equal to $N * (N-1) / 2$.

1.3 LSB Hidden algorithm

It is a widely used method, as it uses the field less important than a particular pixel to store information. It is the most used method, and requires that one or more broadcasts of the message to be hidden and replaced by broadcasts of less importance than the pictures be entered. The least important broadcast is the one with the lowest mathematical value ($2^0=1$) in that the most important broadcast is the one with the greatest mathematical value ($2^7 = 128$).

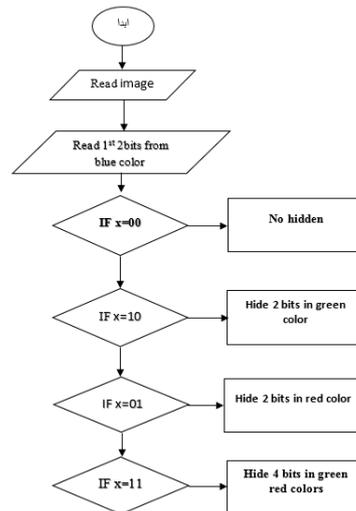


Figure 1: *LSB Algorithm*

2- Related work

In [2], The Sneha Arora suggested a technique to hide the text data into the color images using edge detection method. The alteration in edges cannot be distinguished well so edges can hide more data without losing quality of an image.

in [3] Khan Muhammad presents a novel approach for image steganography using Hue-Saturation-Intensity (HSI) color space based on Least Significant Bit (LSB). The proposed method transforms the image from RGB color space to Hue-Saturation-Intensity (HSI) color space and then embeds secret data inside the Intensity Plane (I-Plane) and transforms it back to RGB color model after embedding.

In [4] Muhammad proposes a way to improve Least Significant Bit (LSB) by randomly inserting message bits into an image to produce a more secure system .

In[5] Manjula K G use DES algorithm is utilized to image file encryption and decryption.

3- The Proposed Algorithm

In this section, the method proposed in this paper will be explained, which includes several operations, which is reading the text, image and secret key and incorporating text into the image based on the secret key, as well as incorporating data in several ways so that we seek to enter the largest amount of data into the image with the least distortions MSE calculation and image resolution evaluation with increased data size.

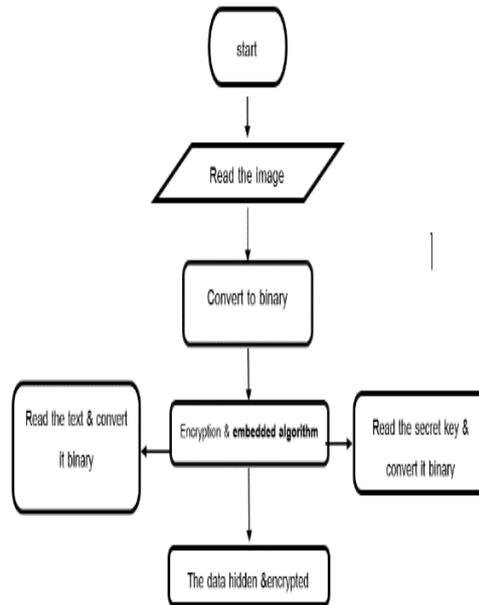


Figure 2: The Proposed Algorithm

3.1 The Secret Key Algorithm

In this paper, a secret key was proposed that encrypts the data, as it contributes to the distribution of data within the image in a way that makes it not possible for the third party to extract the text from the image in the event that he obtained the cached image.

Suppose that the value of the main secret key is 1001001101

The following step is to generate the subkeys where:

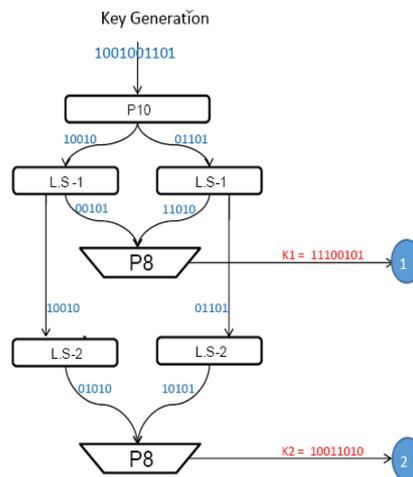


Figure 3: Generated two sub keys

p10: represents the demodulator bits from the main key

p8: represent the demodulator bits from the previous ten bits (the permutation operation) corresponding to the next equation

$$p8 = [6 \ 3 \ 7 \ 4 \ 8 \ 5 \ 10 \ 9]$$

LS_1 , LS_2 represent the shift to left

Whereon product the two sub keys represent the next

- 1- The first sub key to determine the value of address of the start pixel
- 2- The second sub key to determine the value of jump between the pixels

The next table to generate two sub keys

Table 1: *The function of each subkey*

The first sub key	The second sub key
the value of address of the start pixel	the value of jump between the pixels

It represents the address of the first pixel at which the masking process begins Table below The title of the first pixel at which the masking process begins Through the previous table, the subkey of the first pixel, the seventh row and the tenth column, was generated

Table 2: *select the address of the first subkey*

First subkey							
Row Number (for Image Matrix)				Column Number (for Image Matrix)			
1bit	2bit	3bit	4bit	5bit	6bit	7bit	8bit
1	1	1	0	0	1	0	1
Convert to decimal				Convert to decimal			
07				10			

The following algorithm explains the process of selecting the first pixel

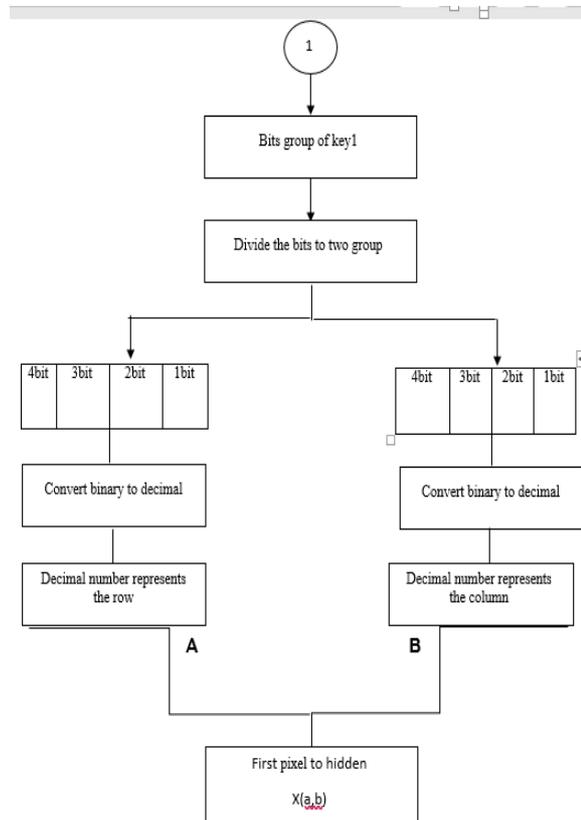


Figure 4: *The algorithm to select the first pixel*

2. Second sub key (jump value between each pixel)

The jumping technique was used and this technique allows to reduce interference in the image file. This technique selects the first two bits in the blue color for each pixel.

Table below The jump value for each pixel

Table 3: *Value of jump*

Value of key(binary)	00	01	10	11
Value of jump (decimal)	0	1	2	3

the table (3) shown the value of jump depending on the second sub key

Table 4: *Select the value of jump of secret subkey*

the second sub key								the value of first two bits from blue color
00		01		10		11		
1bit	2bit	3bit	4bit	5bit	6bit	7bit	8bit	The value of each bits of key
0	0	1	1	1	0	1	0	
0		3		1		1		The value of jump (decimal)

The following algorithm shows determine a jump value between each pixel

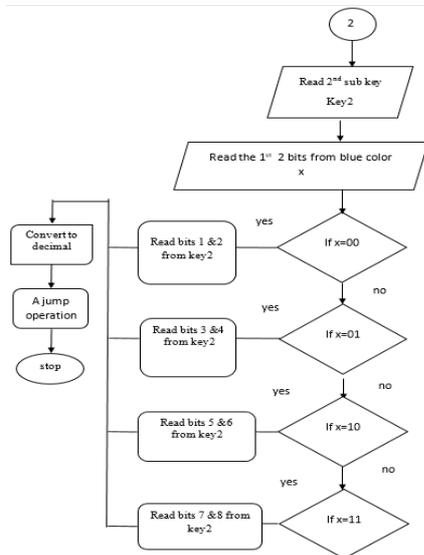


Figure 5: *The algorithm to determine a jump value*

When data is hidden using the key, the data inside the image will be distributed in such a way that the data extracted using the LSB algorithm is encrypted and the locations of the concealment cannot be predicted.

3.2 Evaluate the algorithm’s performance after using the encryption key

The image file is subject to many types of distortions during the stages that you may go through, such as storing, processing and compressing. These distortions affect the image quality, and there are several measures used to evaluate the image quality, such as PSNR,

RMSE, MSE, (Histogram) for each image, they are one of the most commonly used standards.

To evaluate the tool of each LSB algorithm and the proposed algorithm and to know which algorithms give better results, this is done by measuring the image quality. Results from each algorithm. MSE, (Histogram) for each image were used as measures of image quality.

MSE

It is the average squared difference between the original and the modified image

$$MSE = \frac{1}{M * N} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (1)$$

Histograms are a very useful tools used to analyze and compare significant changes in the frequency of appearance of the colors of the cover image with steganography images so as to be able to get a quick summary of the tonal range present in any given image.

3.3. Implementation environment

In this study, algorithms have been implemented using the MATLAB program, as the program has multiple uses such as work interferences, curved of direct value, building functions, establishing new functions, and the most important of them is that is easy to use.

MATLAB programmed for matrix solution , their interactions and their uses in creating programs for large operations. There is a section within the program to build programs for graphics and different types of shapes and results directly on the screen or document them within (M-files) files.

4- Results

In this part, after entered a set of images into the program and using it will be calculated (the Histogram) for each image before and after hiding, calculated text size (TEXT), key size (KEY), calculated the (MSE) standard using a secret key without a secret key, calculated the text and key in each image, and compiling The data is in a table for each image.

1. Image1:

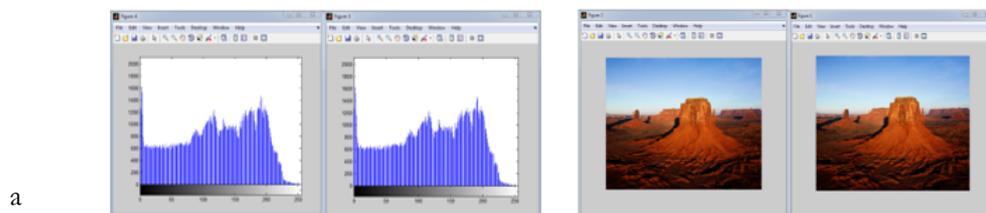


Figure 6(a,b) : The Image & Histogram before & after hide

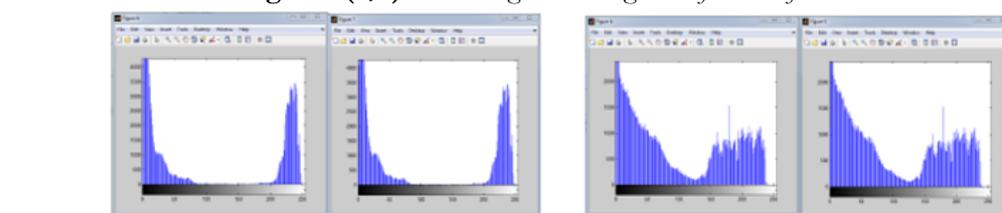


Figure 6(c,d) : The Image & Histogram before & after hide

- a- The image before& after hiding operation
- b- histogram of red color before& after hiding operation
- c- histogram of green color before& after hiding operation
- d- histogram of blue color before& after hiding operation

The following table calculate the MSE for two text , two keys & without key

Table 5: Calculate the MSE for Image1

Text	Size	key	size	Mse	Mse for Lsb without password
Text 1	488 bits	Key 1	248 bits	0.0011	0.000498453776
Text 1	488 bits	Key 2	80 bits	0.000086466	
Text 2	208 bits	Key 1	248 bits	0.0005776	0.0002746582031
Text 2	208 bits	Key 2	72 bits	0.00047302	

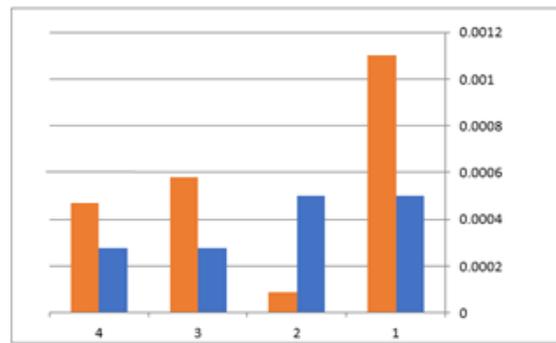


Figure 7 : the MSE for Image1

2. Image 2:

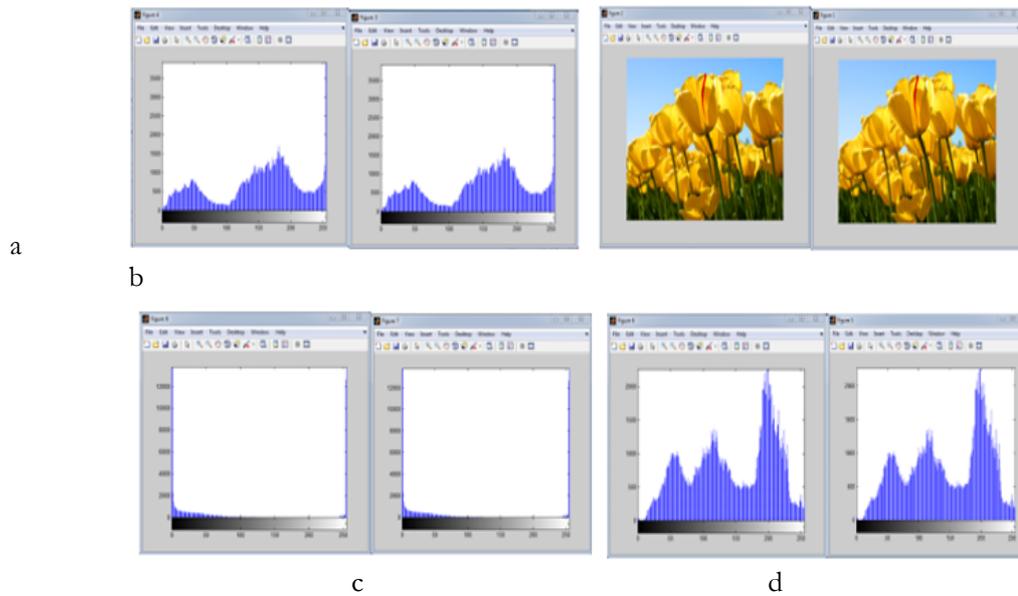


Figure 8: show the image & histogram

- e- The image before& after hiding operation
- f- histogram of red color before& after hiding operation
- g- histogram of green color before& after hiding operation
- h- histogram of blue color before& after hiding operation

The following table calculate the MSE for two text , two keys & without key:

Table 6: Calculate the MSE for Image2

Text	Size	key	size	Mse	Mse for Lsb without password
Text 1	488 bits	Key 1	248 bits	0.0015	0.0006256103516
Text 1	488 bits	Key 2	184 bits	0.0014	
Text 2	208 bits	Key 1	248 bits	0.00051371	0.0002136230469
Text 2	208 bits	Key 2	184 bits	0.00056458	

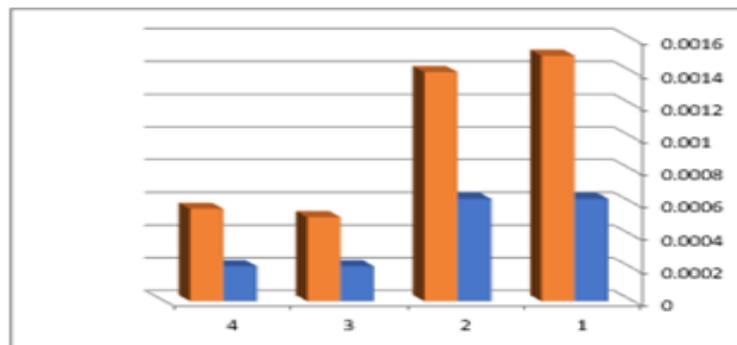


Figure 7 : the MSE for Image2

5- conclusion

in this method encrypted data was obtained Without using encryption technology, Using the secret key the data confidentiality service has been increased, No apparent distortions of the digital image, Having more than one jump and starting point depending on the secret key made it more difficult to extract pictures. As a first Recommendations Using the master key for other text hidden techniques such as (FMM) technology, As a second Recommendations This algorithm implements on another data such as video, audio, ..., As a third Recommendations increase the more possibilities for the number of jumps and the starting point by increasing the number of bits selected from the subkey, As a fourth Recommendations Use secret keys designed for encryption algorithms such as (RSA) to distribute data inside the image, and compare it with the proposed algorithm in this paper.

References

- [1]. K.Brindha, ,” Use of Symmetric Algorithm for Image Encryption”, International Journal of Innovative Research in Computer and Communication Engineering, Vol. 2: No. 5, May - 2014.
- [2]. Sneha Arora, ,” A Proposed Method for Image Steganography using Edge Detection " An International Journal of Engineering Sciences, Vol. 8, June - 2014.
- [3]. Khan Muhammad, " A Novel Image Stegano-graphic Approach for Hiding Text in Color Images using HSI Color Model”, International Journal of Computational Intelligence and Informatics, Vol. 3: No. 3, October - December 2013.
- [4]. Obaida Mohammad, ,” Hiding Data in Images Using New Random Technique”, International Journal of Computer Science Issues, Vol. 9, Issue 4, No 2, July 2012.



- [5]. Manjula K G," Color Image Encryption and Decryption Using DES Algorithm", International Research Journal of Engineering and Technology, Volume: 03 Issue: 07 | July-2016.
- [6]. V.Hemamalini ," A Survey on Elementary, Symmetric and Asymmetric Key Cryptographic Techniques ", International Journal of Computing Academic Research (IJCAR) ISSN 2305-9184, Volume 5, Number 1 (February 2016), pp.11-26.
- [7]. Unik Lokhande ," An Effective Way of using LSB Steganography in images along with Crypto-graphy ", International Journal of Computer Applications , Volume 88 – No.12, February 2014
- [8]. Muhmed F. Agbisha , " Encrypt The Text and Hide in the RGB Image using three Pointers of color" ,ICTS 2018.

Internet of Things for Libya Healthcare System: Challenges and Issues

Sarah Makhlu¹, Ali M. Abdulshahed^{2*}

¹ sarah.makhlouf@it.misuratau.edu.ly, ² a.abdulshahed@eng.misuratau.edu.ly

¹ Department of Networks and Communications, College of Information Technology, Misurata
University, Libya

² Department of Electrical and Electronic Engineering, College of Engineering, Misurata University,
Libya

ABSTRACT

This paper proposes a system based on the Internet of Things to Human Activity Recognition by monitoring and processing smartphone sensors remotely. A machine learning algorithm could be used to determine the activity done during the day. Meanwhile, it is expected to give a feedback during and after the activity is performed, using the machine learning method with remote visualization and intelligent decisions. All in all, it is hoped that the outcome of this proposed work will be beneficial for government officials, health policy makers, and healthcare providers in the area of the healthcare industry.

Keywords:

Internet of Things,
Artificial Intelligence,
Machine learning.

1- Introduction

The term "the Internet of Things" (IoT) is a concept reflecting a connected set of people, sensing objects, anytime, anyplace, any service, and any available network [1, 2]. This concept was, firstly, presented to the world by Kevin Ashton in 1999 [3]. According to Cisco's forecast, 500 billion smart devices will be connected by the year 2030 [4]. The recent advancements in technology and the availability of the Internet make it possible to connect various devices that can communicate with each other and share data. The Internet of Things is emerging communication concepts that will interconnect a variety of devices, sensors, actuators, smartphones, and people, where clearly identifiable and smart objects can communicate with each other in a defined environment to make autonomous decisions by analysing and processing the data collected from the environment.

With the exponential increase of human health data, a main challenge is to create suitable platforms for data analysis, where artificial Intelligence (AI) can be applied to various types of healthcare data. An artificial Intelligence (AI) system is a system that can take decisions, which would be considered intelligent if made by a human being. Artificial Intelligence is becoming more popular and particularly amenable to modelling complex systems, because it has demonstrated superior predictive ability compared to traditional methods [5]. Artificial Intelligence enabled IoT creates intelligent things that simulate smart behaviour and support in decision making with little or no human intervention.

IoT is expected to take more important roles in healthcare, human-machine interfacing, and biological studies. Chu, et al. [6] proposed "BodyNET"; a body area sensor network, which enables wireless transmission of multiple physiological signals to the digital format in a continuous and accurate fashion. The authors indicate that their system will require an efficient communications network to connect their layers, and this is yet to be built. Therefore, it must bridge the digital and physical worlds between individuals, objects and environments. Number of vehicles on the road have growth dramatically in the past few years. It leads to accidents that are the main cause of death in many countries around the world. In most of the cases of accidents, lack of providing quick first aid and timely medical service is the cause of death. Consequently, it is needed to develop an

intelligent framework for accident detection, avoiding other accidents, and provide a timely and reliable response. Kodali et al. [7] have proposed an automatic accident notification system using IoT technology. Message Queuing Telemetry Transport (MQTT) based on Wi-Fi network was used to send an alert through email service while a vehicle accident has detected and data received from accelerometer and ultrasonic sensors stored in an IoT platform. However, their system could not detect the exact location of the accident. As a result, rapid rescue and sending help is very difficult. Additionally, email communication is relatively slower than a text message service. Falls are the main cause of fatal and nonfatal injuries among older adults around the world. It has been reported that, falls cause over 37.3 million severe injuries and 646,000 deaths yearly, and hence has become a global public health issue [8]. It has become obvious that there is a need to develop an intelligent IoT system, which can monitor and detect elderly falls without human intervention. Santos et al. used a Convolutional Neural Network (CNN) to discriminate falls using three open datasets and against extant research based on acceleration measurements. Authors showed that the best results were achieved when using data augmentation during the training process. However, their model could be improved by considering different activities, e.g., standing, walking, jumping, and running and different fall types.

The messaging protocol is considered the main element in the Internet of things. These protocols play a big role to enable IoT all over life. One of IoT healthcare systems design is to choose the type of IoT messaging protocol. Chen and Kunz [9] studied IoT protocols under limited environments: MQTT, CoAP, DDS, and XMPP, using a healthcare case study based on a medical device. They examined a number of indicators such as latency and packet loss under progressively constrained environments. They found that both MQTT and DDS have zero packet losses in high latency environments. Although, the DDS has good performance in terms of latency, However, the DSS has higher bandwidth requirements.

To obtain the patient data and process it intelligently, different studies use different mechanisms such as RFID, 3G, 4G, 5G, and Wi-Fi technology. The 5G networks stand for fifth-generation mobile technology, which can outperform earlier generations of wireless communication technology. Al-Falahy et al. [10] have discussed the potential techniques for 5G networks, which lead to the upcoming challenges and prospects for future research. However, the trend in future networks has shown a different pattern from that of existing networks, because the main objective has changed from enabling users to connect wirelessly to the Internet to enabling massive numbers of users and devices to seamlessly connect by the internet of things technology.

Nowadays, Wi-Fi technology can be considered as a ubiquitous technology found in a wide range of consumer devices such as smartphones. However, in the past, Wi-Fi technology has not shown a significant existence in the IoT market, without any specification focused on IoT and its singularities. Taking into account the near future scenario for IoT communications; the new Wi-Fi 6 "IEEE 802.11ax" technology could bridge the gap by offering new amendment such as high data rate, less interference, and larger data channel widths [11].

Although Wi-Fi technology has enabled the rapid growth of personal wireless networks, the security protocols have shown weaknesses that can be exploited in order to gain unauthorized access to the network and the encrypted communication data. Leca [12] have gathered a database consisting of approximately 100000 unique access points gathered through war-driving in Romania. Results show that rural areas adopted WPA security directly when compared to urban areas that have made the transition from WEP to WPA and still use legacy equipment. The results also show that provider wireless access points on average offer better security than private networks. This work is also concerned with highlighting known vulnerabilities in Wi-Fi networks and their spread. The authors have gathered a database consisting of approximately 10,000 access points gathered through war-driving in Rabat, the capital of Morocco. They indicate that 77% of networks have used accepted Wi-Fi security algorithms. Furthermore, the network operators were aware of high interference that can occur on channel 6 and therefore, act to minimize interferences [13].

Due to the emerging technology, small wireless solutions which are coupled to IoT can make it possible to monitor patients remotely instead of visiting the traditional hospital. A variety of smart sensors which are attached to the body of a patient can be used to obtain health data securely, and the acquired data can be analysed (by means of Machine learning algorithms) and sent to the server using different network connections. All doctors, emergency centres, hospitals, and relatives can access from anywhere at any time, visualize the data, and take a decision accordingly. However, it has been noticed that in most of the above-mentioned studies and related discussion, a comprehensive platform is still missing in the IoT healthcare sector. Since, IoT is a new concept in Libya, there is a need to increase public awareness about its security and capabilities.

The aim of this work is to propose a system based on the Internet of Things to Human Activity Recognition by monitoring and processing smartphone sensors seamlessly and remotely.

2- The proposed system

With the continuous demand of an efficient healthcare system, a major challenge is to design suitable IoT model for an accurate recognition system. The IoT model proposed in this work for efficient deployment of IoT in the field of healthcare consists of four layers (sensor layer, network layer, internet layer, and service layer) as mentioned in [14]. The Human Activity Recognition system will be able to use the smartphone sensors to collect the patient data, which is integrated with the Internet of Things. A patient used built-in accelerometer sensor to obtain health related data. Machine learning algorithms will be applied on data to verify the conditions of the patient. If the value is out of the normal range, then an alert message is sent to a doctor, relatives and the emergency centre will take appropriate action accordingly. The whole scenario of the proposed system is shown in Figure 4 .

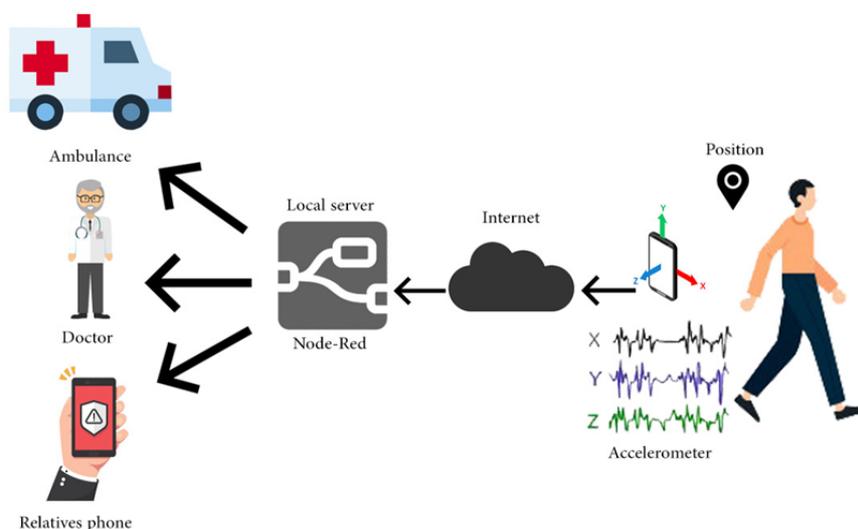


Figure 4. Human Activity Recognition -IoT system

In order to implement a practical IoT system for a healthcare application, the following tools can be used:

1. A smartphone with embedded accelerometer and Global Positioning System GPS; These embedded sensors allow the smartphone to be used as a useful tool to collect patient data.
2. Node-RED runs on a Node.js server; Node-RED is a programming tool for IoT software development created by IBM for “wiring together hardware devices, APIs and online services in new and interesting ways”.
3. A machine learning algorithm that runs on a Node.js server in order to Human Activity Recognition by monitoring and processing smartphone sensors seamlessly and remotely.

This system is proposed to be used by patients with chronic heart diseases, patients who have their health status in a non-critical condition but still need constant monitoring. The proposed system focuses on the daily routine, activity and physical therapy that each patient must have as part of their recuperation process .

The activity recognition can be implemented using a machine learning algorithm that uses the information gathered by the IoT system. The human activity recognition system is a specific application of pattern recognition and expert systems. This system works in two phases: a training phase and a recognition phase. Although both have similar steps, the training phase has a priori information of the activities done and the recognition phase uses the knowledge of the training phase to have accurate recognition. The learning or training phase is the first phase of any recognition algorithm. It has three main steps (see figure 5); Data collection, Feature extraction, and Learning process. On another hand, the recognition phase is considered as a final phase of an HAR system. It does the recognition of the activity performed using the training phase result. The recognition phase has three main steps (see figure 6); Data collection, Feature extraction, Recognition process.

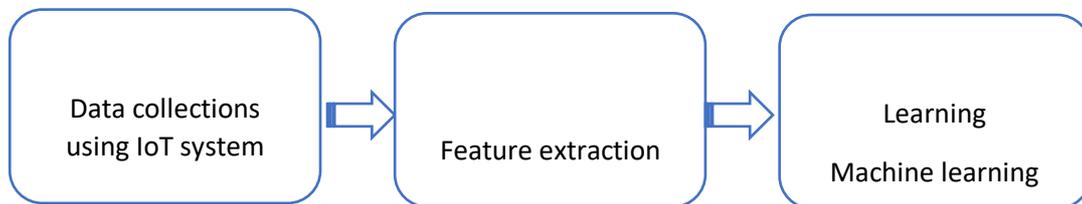


Figure 5. Learning process of the proposed system

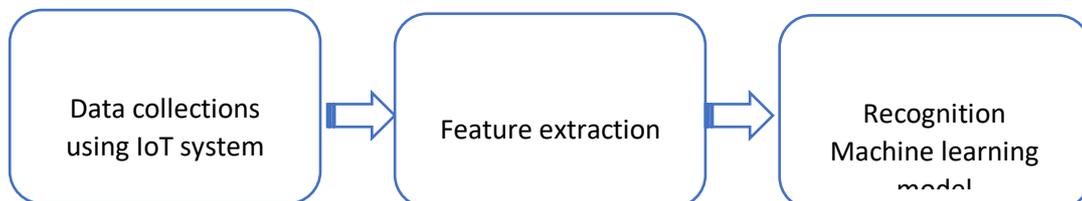


Figure 6. Recognition process of the proposed system

This proposed system aims to achieve a set of goals that have to do with the Internet of Smart Things in the field of health care through monitoring, analysis and making decisions in the right time, so that the followings can be considered: (1) Develop experimental platforms for interfacing smartphone equipped sensors using state-of-the art IoT technologies. (2) Observing the person's condition throughout the day and sending reports periodically and predicting any unfamiliar behaviour. (3) To increase public awareness on IoT wireless network security and to highlight existing vulnerabilities in Wi-Fi networks that could be avoided.

3- Discussion and future perspectives

Violence in Libya continues to have a huge impact on healthcare sector, with hospitals and other medical facilities bombed, shelled and patients at times denied prompt life-saving care or attacked while getting treatment. Moreover, there are so many people in Libya whose health may suffer because they do not have proper access to hospitals and health services. Furthermore, the accelerated spread of coronavirus has exposed and exacerbated many structural problems in the

healthcare systems. These problems can only be solved if there is a solution that is easily scalable and automated. Due to the emerging technology, small wireless solutions which are coupled to IoT can make it possible to monitor patients remotely instead of visiting the traditional hospital. A variety of smart sensors which are attached to the body of a patient can be used to obtain health data securely, and the acquired data can be analysed (by means of Machine learning algorithms) and sent to the server using different network connections. All doctors, emergency centres, hospitals, and relatives can access from anywhere at any time, visualize the data, and take a decision accordingly. Patients should be continuously monitored when they are out of the hospital. This should be possible by an efficient remote monitoring system that expands access to medical services. In this perspective, because most of the IoT sensing devices have low computing capability and small storage, the HTTP protocol, due to its complexity, cannot be considered in such an environment. The MQTT protocol, has some outstanding features that can change HTTP properties to increase requirements of IoT. Some of the key features of this protocol are low power consumption and overhead that make MQTT a suitable solution for the IoT implementation.

Personal wireless networks are one of the main components of IoT. Wi-Fi is used in various fields such as home healthcare, wearable sensor devices, mobile devices, and smart grids. In Libya, Wi-Fi is a favourite way to connect to the Internet. The falling costs and high availability of wireless equipment in Libya have enabled the growth in the number of wireless networks. Because of broadcast technology used, they are subject to security attacks. It is crucially important to investigate the risks and potential vulnerabilities of Wi-Fi networks to recommend users how to improve their security with less interference.

In most of the above-mentioned studies and related discussion, however, it has been noticed that a comprehensive platform is still missing in the IoT healthcare architectures to use smartphone sensors to sense and transmit important data related to a patient's health. In this work, it is important to conduct a study of Wi-Fi performance and security in Misurata (as a case study) in order to increase public awareness on IoT wireless network security. Moreover, we wonder if accurate recognition of activity can be obtained using the Internet of Things and artificial intelligence technology in a non-controlled environment.

4- Conclusions

In the near future, healthcare sector will be developed by IoT technology and Artificial intelligence tools. These emerging technologies will play a prominent role in vital signs monitoring system in hospitals and more importantly at homes. Remote patient monitoring gives great opportunities to not only increase serveries quality but also minimize serveries costs by diagnosing diseases and harmful situations. Nowadays, the healthcare services in Libya are costlier than ever before, while most of the patients are needed to enter the hospital for a certain period of time. Using AI-based IoT with a capability to seamlessly and remotely monitor patients can somehow overcome the healthcare challenges. These technologies, by acquiring the patient's real-time health data and transferring it to decision makers, will not only reduce the cost of healthcare services, but also enable the treatment of health issues before they become critical.

References

- [1]. S. R. Islam, et al., "The internet of things for health care: a comprehensive survey," IEEE access, vol. 3, pp. 678-708, 2015.
- [2]. G. Aceto, et al., "Industry 4.0 and health: Internet of things, big data, and cloud computing for healthcare 4.0," Journal of Industrial Information Integration, vol. 18, p. 100129, 2020.
- [3]. K. Ashton, "That 'internet of things' thing," RFID journal, vol. 22, pp. 97-114, 2009.
- [4]. A. Yastrebova, et al., "Future networks 2030: Architecture & requirements," in 2018 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2018, pp. 1-8.
- [5]. A. M. Abdulshahed, et al., "The application of ANFIS prediction models for thermal error compensation on CNC machine tools," Applied Soft Computing, vol. 27, pp. 158-168, 2015.
- [6]. B. Chu, et al., "Bring on the bodyNET," Nature News, vol. 549, p. 328, 2017.
- [7]. R. K. Kodali and S. Sahu, "MQTT based vehicle accident detection and alert system," in 2017 3rd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), 2017, pp. 186-189.
- [8]. G. L. Santos, et al., "Accelerometer-based human fall detection using convolutional neural networks," Sensors, vol. 19, p. 1644, 2019.
- [9]. Y. Chen and T. Kunz, "Performance evaluation of IoT protocols under a constrained wireless access network," in 2016 International Conference on Selected Topics in Mobile & Wireless Networking (MoWNeT), 2016, pp. 1-7.
- [10]. N. Al-Falahy and O. Y. Alani, "Technologies for 5G networks: Challenges and opportunities," IT Professional, vol. 19, pp. 12-20, 2017.
- [11]. V. Baños-Gonzalez, et al., "IEEE 802.11 ah: A technology to face the IoT challenge," Sensors, vol. 16, p. 1960, 2016.
- [12]. C. L. Leca, "Overview of Romania 802.11 wireless networks security," in 2017 9th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), 2017, pp. 1-4.
- [13]. A. Sebbar, et al., "An empirical study of wifi security and performance in morocco-driving in rabat," in 2016 International Conference on Electrical and Information Technologies (ICEIT), 2016, pp. 362-367.
- [14]. K. Ullah, et al., "Effective ways to use Internet of Things in the field of medical and smart health care," in 2016 International Conference on Intelligent Systems Engineering (ICISE), 2016, pp. 372-379.

Enhancement of VoIP Performance in MANET using Fuzzy Logic

Abdusalam Yahya¹, Ashraf Gasim Elsid Abdalla²

¹abdusalamabdullahyahya74@gmail.com, ²agea33@yahoo.com

¹Department of Computer Network, Collège of IT, Elmergib University, Libya

²Future university, Faculty of Postgraduates studies, faculty of Telecommunication and Space
Technology, Khartoum, Sudan

ABSTRACT

Voice over Internet Protocol (VoIP) application is a vital technology that is quickly growing in the Mobile ad hoc Network (MANET). Packet loss is a factor that can significantly affect the Quality of Service (QoS) for VoIP performance. Due to the dynamic nature of MANET, it is a challenging task to maintain the desired packet loss rate. This paper aims to enhance the performance of VoIP in the MANET using a fuzzy logic model. The input for the model is VoIP packet loss and the outputs are the optimal parameters of MANET (node number, pause time, maximum speed, and maximum connection). Network Simulator (NS2) was used to perform all simulations. Matlab was used to implement the proposed fuzzy model. Moreover, the performance of the model was evaluated using NS2, and the results show that our proposed fuzzy model offers a significant enhancement in terms of the VoIP packet loss rate (P.LR).

Keywords:

VoIP,
MANET,
Fuzzy Logic.

1- Introduction

Voice over Internet Protocol (VoIP) is one of the fastest rising Internet applications in recent years [1]. It replaces the old Public Switched Telephone Network (PSTN) with a converged network where data, voice, and video packets are carried together using a single IP-based network. The integration between old voice network (PSTN) and IP data network into one network offered low cost and improved network management. Originally networks were designed to provide services to non-real-time applications such as FTP, HTTP, and email services which can tolerate some delay, jitter, and packet loss. Therefore, they are not sensitive to low network performance.

The VoIP is a real-time application and it has special characteristics. It is unable to tolerate inefficient network performance. Thus, IP networks must afford exceptional services to carry voice packets [2]. VoIP is an application that employs Real-Time Transport Protocol (RTP) to transfer voice packets over IP based networks. RTP runs on top of the user datagram protocol (UDP), which is an unreliable delivery protocol. Regarding real-time communication, voice packets that delay in arrival at the destination are useless and equal to drop packets [3].

Public Switch Telephone Network (PSTN) has provided the requirements of businesses and consumers worldwide with telephone service for more than 100 years [4]. PSTN can accommodate all voice quality requirements, as it was originally designed for transmitting voice signal. In contrast, wireless networks are an IP-based network that was originally designed to carry only data packets. Moreover, due to the mobility of the MANET network, congestion occurs and introduces a high packet loss rate. As a result, carrying voice packets over the MANET network requires much concern. The major VoIP quality metrics in the IP networks are Packet Loss, Jitter, and End-to-End Delay [5] [6].

Fuzzy logic was invented by Dr. Lotfi Zadeh at the University of California at Berkeley in 1965. Fuzzy logic allows situations or problems to describe and processed in linguistic words such as “cold “or “hot” as an alternative to exact numeric values such as a degree or 40 degrees. The advantages of fuzzy logic such as a linguistic system led to adopting this technology in many fields such as electronic, engineering control, information processing, financial analysis, and much more [7].

Many researchers have proposed solutions for improving VoIP performance using different techniques. However, none of them has employed the VoIP packet loss rate as input for the fuzzy system. Therefore, the contribution of this paper is to build a new fuzzy logic system that considers a packet loss rate as the input. The proposed system is used to predict the best parameters of MANET to minimize the VoIP packet loss rate. The input for the model is VoIP packet loss rate and outputs are MANET parameters (number of nodes numbers, maximum speed, maximum connection, and pause time).

The rest of this paper was organized as follows: In section 2, the related work was reviewed. Section3 discussed and illustrated the proposed approach. Section 4 presented the evaluation process. Section 5 discussed the results. Finally, the conclusion was discussed in section 6.

2- Related Works

Using a Fuzzy logic-based approach for VoIP quality maintenance proposed in [8]. The authors employed the Resource reservation protocol (RSVP) to control the token bucket algorithm. They built the Fuzzy token system includes three variables (Bandwidth Rate, Buffer Size, and New Token bucket). Stability was used to determine the best method to get a new token bucket that does not cause any factors (Jitter – Delay – Packet loss). The result shows that whenever improving bandwidth rate and buffer size rate, a new token bucket at the same rate of improvement and vice versa [8].

The authors of [9] employed a multi-criteria decision-making scheme to evaluate and select the best network services. The decision-making algorithm is based on the fuzzy logic approach. The study includes five service providers in Turkey namely, ADSL, Fiber technology, Power Line Communication, and high-speed Internet access offered by Cable Television operators. Four user profiles were created from the target ISPs (profile1: video, profile2: VoIP, profile3: Streaming Media, profile4: Interactive). Their study concluded that the use of fuzzy logic led to a strong indication of the expert opinions on decision-making.

Sweta and Meenn in[10] proposed a decision support system to control two issues in buffer management (drop decision when buffer overflow occurs and scheduling decision when a transmission chance starts). They suggested a development to the Customs Service time scheduling by integrating it within a fuzzy-based buffer ranking mechanism. The study was conducted in Delay Tolerant Networks (DTNs). The results show that the proposed Fuzzy logic –based traffic differentiation technique reaches improved delivery performance over the existing traffic differentiation scheme for DTNs.

Demonstrating the use of fuzzy logic to control congestion in ATM and TCP/IP networks has been done In [11]. The researchers follow ATM Forum Traffic Management Specification version 4 to propose Fuzzy Explicit Rate Marking in ATM and TCP/IP best effort. They concluded that the proposed technique can offer important improvements in controlling congestion in computer networks.

In [12] the authors presented a dynamic QoS solution for enterprise networks. The proposed decision-making is based on 0-order TSK fuzzy model. This technique was employed to modify the priority level of a packet in real-time by adjusting the value of Differentiated Services Code Point (DSCP) in the Internet Protocol (IP) header. The developed system was tested in a real

network environment and the results show that the proposed system dynamically able to change data transmission priority for different situations to improve the quality of service.

Improving service quality and network performance were accomplished in[13]. The authors used the Sugeno fuzzy system and differential operation to optimize the output and input variables of the fuzzy controller. The results of the simulation show that the proposed algorithm is capable to control the queue more efficiently than classical techniques.

In [14] an Adaptive Neuro-Fuzzy Inference System proposed is proposed to identify congestion in the sink node. It enables the sink node to guess the congestion rate using local information like a participant, buffer occupancy, and traffic rate as input for the protocol. The results show that the proposed protocol can detect the congestion level correctly whenever there is a packet drop.

3- Implementation of Fuzzy logic

Fuzzy logic can be implemented in several techniques, one of them is the Fuzzy Inference System (FIS) [15]. A fuzzy inference system is used to control the connection between the input and output variables of a system. The crisp inputs in the fuzzy inference system are changed into fuzzy inputs by using a fuzzification interface. After fuzzification, the rule base is used. Defuzzification is used to convert the fuzzy value to the real-life value which is the output.

3.1 Traffic Generation

The first step to build our model is to create input data. According to previous studies, NS2 is the most appropriate tool to simulate the dynamic nature of MANET [16]. Therefore, to generate input data NS2 is used to build a MANET. VoIP traffic was modeled in ns2 by using Code Bit Rate (CBR) traffic. The Tool Command Language Script (TCLS) was used to write the TCL script. Table 1.1: illustrates simulation parameters designed to calculate the VoIP packet loss rate.

Table 1: *Simulation Parameters*

Parameters	Values
Simulation Area Size	500 X500 m2
Pause Time (P.T)	2-20
Maximum Connection (M.C)	2-20
Maximum Speed (M.S)	2-20
Node Number (Node. No)	5-40
Traffic Type	CBR (VoIP)
Simulation Time	300sec
Simulation runs	18

3.2 VoIP Packet Loss Rate

Packet loss arises once transmitted packets across a network never reach its destination [13]. The packet loss rate (P.L.R) is a significant parameter affecting the QoS of VoIP. However, it is expressed as a ratio and calculated using the following formula: =

$$(((\text{Number of sent packets} - \text{Number of received packets}) / \text{total sent packet}) * 100) [17]$$

Table 2: *Packet Loss ratio of VoIP*

Scenario.	MANET Scenarios Parameters				Performance
	Node	P.T (s)	M.P (M/S)	M.C	VoIP P.L.R (%)
1	5	2	2	2	0.01
2	8	4	4	4	0.02
3	10	5	5	5	0.08
4	12	6	6	6	0.19
5	14	7	7	7	0.1
6	16	8	8	8	0.55
7	18	9	9	9	4.5
8	20	10	10	10	6.5
9	22	11	11	11	7.45
10	24	12	12	12	11.5
11	26	13	13	13	14.99
12	28	14	14	14	20
13	30	15	151	5	26.5
14	32	16	16	16	39.16
15	34	17	17	17	39.5
16	36	18	18	18	40
17	38	19	19	19	51.14
18	40	20	20	20	52.06

3.3 Fuzzy System Construction

The second step is to build a fuzzy system using MATLAB R2016a. The input variable is the VoIP Packet loss rate and the output variables are optimum MANET parameters.

3.4 Membership Function

The membership function type for the input and outputs was Triangular membership that offered flexibility to represent the information. Table 3 shows the minimum and maximum value for the input and outputs membership functions.

Table 3: Range of Fuzzy system Membership functions

Fuzzy Parameters	Minimum	Maximum
Packet loss	0	40
Node number	2	20
Pause time (sec)	2	20
Max speed	2	20
Max connection	2	20

Figures 2 to 6 show a graphical representation of the membership functions.

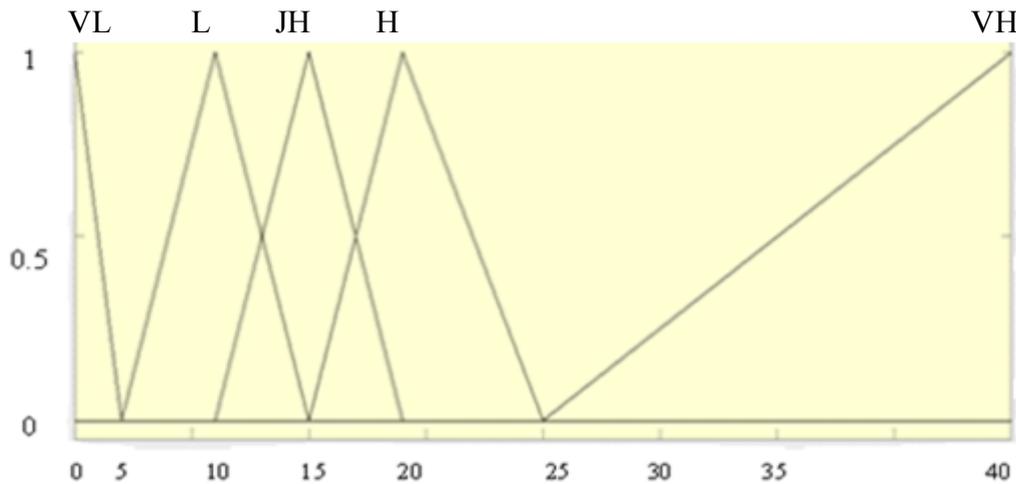


Figure 1: Membership Function of VoIP Packet loss Ratio

Figure.1 shows the triangular membership function of the VoIP packet loss ratio. The x-axis represents the packet loss rate with a range from 0 to 40, whereas the y-axis represents the degrees of membership in the [0,1] interval. The membership function consists of five fuzzy sets (Very low (VL), Low (L), Just High (JH), High (H), and Very High(H))

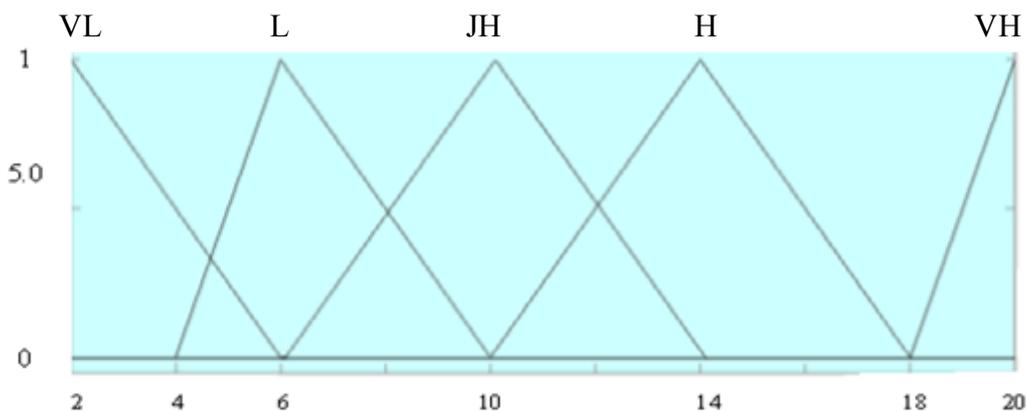


Figure2: Membership Function of MANET Number of nodes

Figure 2 displays the triangular membership function of the MANET node number. The x-axis represents the number of nodes with a range from 2 to 20, whereas the y-axis represents the degrees of membership in the [0,1] interval. The membership function involves five fuzzy sets (Very low (VL), Low (L), Just High (JH), High (H), and Very High(H))

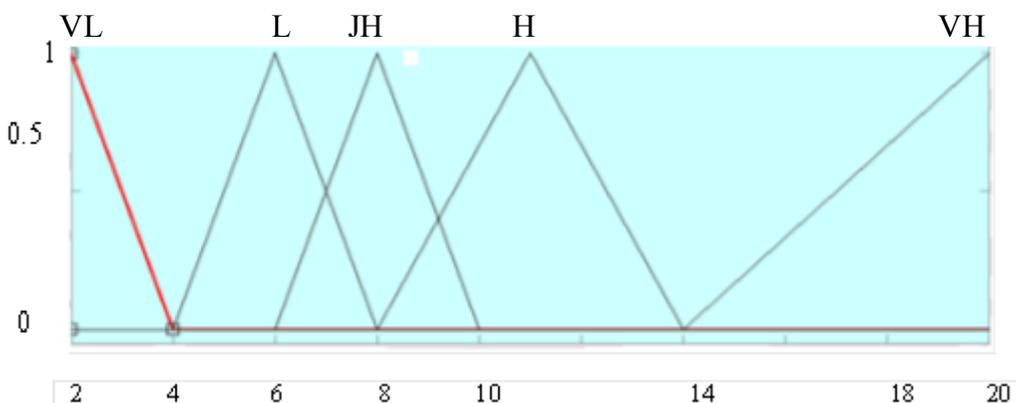


Figure 3: Membership Function of Maximum speed

Figure 3 illustrates the triangular membership function of maximum speed (M. S) of MANET nodes. The x-axis represents (M.S), whereas the y-axis represents the degrees of membership in the [0,1] interval. The membership function consists of five fuzzy sets (Very low (VL), Low (L), Just High (JH), High (H), and Very High(H))

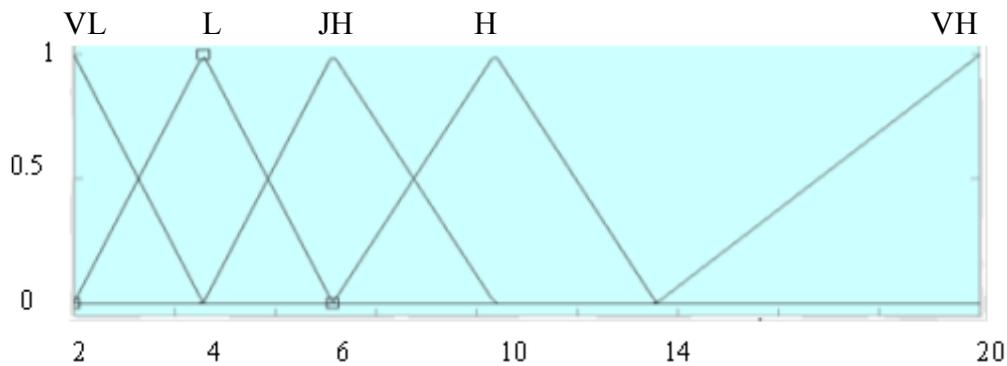


Figure 4: Membership Function of Maximum Connection

Figure 4 plates triangular membership functions of maximum connection (M.C) in between MANET nodes. The x-axis represents the M.C range from 2to 20, whereas the y-axis represents the degrees of membership in the [0,1] interval. The membership function contains five fuzzy sets (Very low (VL), Low (L), Just High (JH), High (H), and Very High(H))

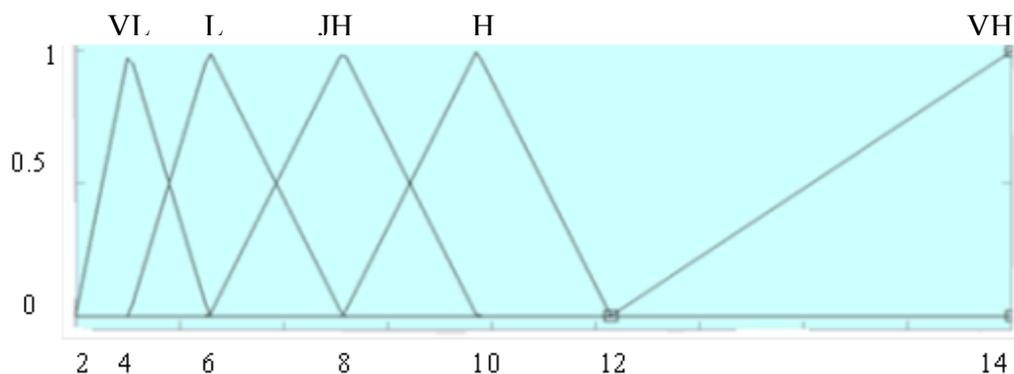


Figure 5: Membership function of Pause time

Figure 5 presents the triangular membership functions of Pause time (P.T) of the mobility of MANET nodes. The x-axis represents the (P.T) with a range from 2 to 20, whereas the y-axis represents the degrees of membership in the [0,1] interval. The membership function includes five fuzzy sets (Very low (VL), Low (L), Just High (JH), High (H), and Very High(H))

3.5 Fuzzy Rules

The fuzzy logic system decides according to control rules and linguistic terms. In the main, the inference system has two models Mamdani and Takagi-Sugeno models. In this study Mamdani approach is used because of its simple structure and straightforward design. The fuzzy rule consists of IF-THEN linguistic term and the output membership functions are worked by the fuzzy rule related between the input and the outputs. The suggested Fuzzy rules in this research are given in table 4.

Table 4: Rules used for the FIS knowledge base

IF	VoIP P.L.R	Then	Node No.	P. T	M.C	M.S
	Very Low (VL)		Very Low (VL)	Very Low (VL)	Very Low	Very Low
	Low (L)		Low (L)	Low (L)	Low (L)	Low (L)
	Just High (JH)		Just High (JH)	Just High (JH)	Just High	Just High
	High (H)		High (H)	High (H)	High (H)	High (H)
	Very High		Very High	Very High	Very High	Very High

4- Performance Evaluation

There are no performance standards available to evaluate the goodness or accuracy of the fuzzy Membership function[18]. therefore, the evaluation process is focused on the entire proposed fuzzy system. The following stages are performed to evaluate the model.

Packet loss values stated in table 2 are given as input to the model as shown in figure 6.

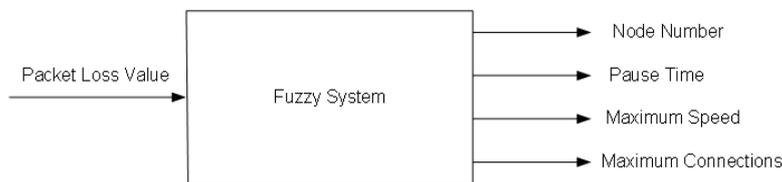


Figure 6: Proposed Fuzzy System

The output from the fuzzy system in response to the input is illustrated in table 5.

Table 5: Output of Proposed Fuzzy Mode

Scenario. No	VoIP	P.L.R	Nodes	M.S (m/s)	P.T (sec)	M.C
1	0.01		6	2	3	2
2	0.02		6	2	3	2
3	0.08		6	2	3	2
4	0.19		6	2	3	2
5	0.1		6	2	3	2
6	0.55		6	2	3	2
7	4.5		3	6	11	4
8	6.5		5	6	7	5
9	7.45		7	6	7	5
10	11.5		11	9	8	8
11	14.99		14	11	9	10
12	20		11	11	11	11
13	26.5		19	17	16	17
14	39.16		19	18	17	17
15	39.5		19	18	17	17
16	40		19	18	17	17
17	51.14		19	18	17	17
18	52.06		19	18	17	17

3 The model output values are used to set MANET scenario parameters as shown in figure 7



Figure 7: *Simulated Mobile ad-hoc Network*

4 The simulation is conducted using NS2 and VoIP packet loss ratio is calculated using AWK script language. Table 6 displays the result of run scenarios.

Table 6: *Performance of VoIP with Fuzzy model*

Scenario. No.	Nodes	M.S	P.T (sec)	M.C	VoIP P.L.R (%)
1	6	2	3	2	0.02
2	6	2	3	2	0.02
3	6	2	3	2	0.02
4	6	2	3	2	0.02
5	6	2	3	2	0.02
6	6	2	3	2	0.02
7	3	6	11	4	0.65
8	5	6	7	5	4
9	7	6	7	5	6
10	11	9	8	8	4.54
11	14	11	9	10	12.71
12	11	11	11	11	7.17
13	19	17	16	17	25.5
14	19	18	17	17	29.4
15	19	18	17	17	29.4
16	19	18	17	17	29.4
17	19	18	17	17	35
18	19	18	17	17	35

5- Results and Discussion

VoIP packet loss rate was calculated and compared with the proposed fuzzy model and without the model, as shown in table 7. The less VoIP packet loss rate was observed when using our model. This is due to applying fuzzy rules that produced the best MANET parameters. The result showed that the VoIP packet loss rate without fuzzy is between 0.01 % and 51 %. In contrast, when the fuzzy model is used the rate is between 0.02 % and 35% as shown in table 7.

Table 7: Comparison of VoIP Packet loss ratio

Scenario. No.	VoIP P.LR without Fuzzy (%)	VoIP P.LR with fuzzy (%)
1	0.01	0.02
2	0.02	0.02
3	0.08	0.02
4	0.19	0.02
5	0.1	0.02
6	0.55	0.02
7	4.5	0.65
8	6.5	4
9	7.45	6
10	11.5	4.54
11	14.99	12.71
12	20	7.17
13	26.5	25.5
14	39.16	29.4
15	39.5	29.4
16	40	29.4
17	51.14	35
18	52.06	35

6- Conclusion

In this paper, we employed a new fuzzy logic model to improve VoIP performance over MANET. The proposed system uses VoIP packet loss value to gain the ideal value of MANET parameters. The system is implemented in MATLAB R2016a.1. concerning testing the model, we performed simulation experiments using NS2. The results approved that we significantly minimize the level of VoIP packet loss rate in mobile ad hoc networks.

References

- [1]. Y.-B. Lin and I. Chlamtac, Wireless, and mobile network architectures. John Wiley & Sons, 2008.
- [2]. S. Na and S. Yoo, "Allowable propagation delay for VoIP calls of acceptable quality," in International Workshop on Advanced Internet Services and Applications, 2002, pp. 47-55: Springer.
- [3]. T.-K. Chua and D. C. J. I. N. Pheanis, "QoS evaluation of sender-based loss-recovery techniques for VoIP," vol. 20, no. 6, pp. 14-22, 2006.
- [4]. U. Varshney, A. Snow, M. McGivern, and C. J. C. o. t. A. Howard, "Voice over IP," vol. 45, no. 1, pp. 89-96, 2002.
- [5]. W. Wang, S. C. Liew, and V. O. J. I. t. o. v. t. Li, "Solutions to performance problems in VoIP over a 802.11 wireless LAN," vol. 54, no. 1, pp. 366-384, 2005.
- [6]. A. A. Yahya and M. A. Alhanish, "Study the Effect of OSPF and IS-IS Protocols Convergence on the VoIP Performance."
- [7]. Y. Tanaka, "An overview of fuzzy logic," in Proceedings of WESCON'93, 1993, pp. 446-450: IEEE.
- [8]. M. E. Ebrahim and H. A. Hefny, "Fuzzy Logic based Approach for VoIP Quality Maintaining," INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS, vol. 9, no. 1, pp. 537-542, 2018.
- [9]. S. Ballı and M. Toker, "A fuzzy multi-criteria decision analysis approach for the evaluation of the network service providers in Turkey," Intelligent Automation & Soft Computing, pp. 1-7, 2017.

- [10]. S. Jain and M. Chawla, "A fuzzy logic based buffer management scheme with traffic differentiation support for delay tolerant networks," *Telecommunication Systems*, vol. 68, no. 2, pp. 319-335, 2018.
- [11]. K. Kumar, K. Narayana, and B. Sangmitra, "Congestion control in high speed networks using fuzzy logic control," *International Journal of Mathematics and Soft Computing*, vol. 5, no. 1, pp. 45-55, 2015.
- [12]. J. Li, L. Yang, X. Fu, F. Chao, and Y. Qu, "Dynamic QoS solution for enterprise networks using TSK fuzzy interpolation," in *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2017, pp. 1-6: IEEE.
- [13]. N. K. Quoc, V. T. Tu, and N. T. Hai, "Some Improvements on Active Queue Management Mechanism Based on Adaptive Fuzzy Control," *EAI Endorsed Trans. Context-aware Syst. & Appl.*, vol. 2, no. 6, p. e4, 2015.
- [14]. T. Mekni, K. IbnTaarit, and M. Ksouri, "Adaptive neuro-fuzzy inference system congestion detection protocol," in *2018 International Conference on Advanced Systems and Electric Technologies (IC_ASET)*, 2018, pp. 363-3680: IEEE.
- [15]. A. Salama, R. Saatchi, and D. Burke, "Fuzzy Logic and Regression Approaches for Adaptive Sampling of Multimedia Traffic in Wireless Computer Networks," *Technologies*, vol. 6, no. 1, p. 24, 2018.
- [16]. T. P. Venkatesan, P. Rajakumar, and A. Pitchaikkannu, "Overview of Proactive Routing protocols in MANET," in *2014 Fourth International Conference on Communication Systems and Network Technologies*, 2014, pp. 173-177: IEEE.
- [17]. W. Mansouri, K. B. Ali, F. Zarai, and M. S. Obaidat, "Radio resource management for heterogeneous wireless networks: Schemes and simulation analysis," in *Modeling and Simulation of Computer Networks and Systems*: Elsevier, 2015, pp. 767-792.
- [18]. E. Kayacan, A. Sarabakha, S. Coupland, R. John, and M. A. J. E. A. o. A. I. Khanesar, "Type-2 fuzzy elliptic membership functions for modeling uncertainty," vol. 70, pp. 170-183, 2018

Comparison of BER Performance of SISO, MISO and MIMO Systems Based on OSTBC and EO-STBC Techniques

Usama I. Hamouda^{1*}, Mohammed J. Bendaw², T. M. Buzid³, Masoud M. Eddaghel⁴

¹ shafterosama6@gmail.com, ² mohammeddaw97@gmail.com, ³ t.buzid@gmx.de,

⁴ masoud_edd@yahoo.com

^{1,2} Department of Communications, College of Electronic Technology-Tripoli, Libya

³ Faculty of Sciences, Alghabel Algharbi University, Libya

⁴ Department of Communications & Microwaves, College of Electronic Technology-Bani Walid, Libya

ABSTRACT

Keywords:

SISO;
MISO;
MIMO;
OSTBC;
open and closed loop EO-STBC;
diversity order;
flat fading channel.

Comparison of bit error rate (BER) performance of single input single output (SISO), multiple input single output (MISO), and multiple input multiple output (MIMO) systems based on orthogonal space time block coding (OSTBC) and extended orthogonal space time block coding (EO-STBC) techniques, over flat fading channel, is addressed. The antenna configurations of MIMO system used in LTE and WiMAX, supports up to four antennas at either the transmitter, receiver, or both. Therefore, the comparison of BER performance of SISO, MISO and MIMO systems are a subject of this paper. The implementation of open loop SIMO and MIMO are based on OSTBC technique. Whereas, the implementation of closed loop SIMO and MIMO are based on EO-STBC technique. Losing the orthogonality decreasing the spatial diversity order. To achieve full spatial diversity order, closed loop (feedback) is used. The MATLAB simulation results confirm that, when the transmitted signals are fully orthogonal as in the system based on OSTBC (Alamouti) scheme, the system achieves full spatial diversity order which means that the best BER performance is reached. In contrast, when the transmitted signals are not fully orthogonal as in the system based on open loop EO-STBC scheme, the system can not reach full spatial diversity order which means that the worst BER performance is recorded. In the former system, feedback is exploited to achieve full spatial diversity order.

1- Introduction

The use of multiple antennas allows independent channels to be created in space which called diversity and is one of the most interesting and promising areas of recent innovation in wireless communications. Rayleigh fading wireless channels make a very large power penalty on the BER performance of the conventional single input single output (SISO) wireless system. To mitigate the effects of Rayleigh fading due to wireless channels, is to use diversity techniques. There are many types of achieving diversity in a wireless system. First method of achieving diversity is by using either two transmit antennas or two receive antennas with different polarization. Second method is frequency diversity, where it achieved by transmitting the same narrowband signal at different carrier frequencies, this technique requires additional transmit power to send the signal over multiple frequency bands. Third method, time diversity is achieved by transmitting the same signal at different times, the time difference should be greater than the channel coherence time. Time diversity does not require increased transmit power, but it does decrease the data rate. Fourth method is space diversity, this type uses multiple transmit antennas, receive antennas, or transmit

and receive antennas at both sides [1], [2]. In practical, the Long-Term Evolution (LTE), LTE-Advanced and Worldwide Interoperability for Microwave Access (WiMAX) based approaches utilize OFDM modulation and employ multiple input multiple output (MIMO) antenna technology. The antenna configurations of MIMO system used in LTE and WiMAX, supports up to four antennas at either the transmitter, receiver, or both [3]. Therefore, performance of MIMO system is a subject of this paper. Space diversity can be classified into three types. Firstly, receive diversity, which is implemented by single input multiple output (SIMO) system and based on maximum ratio combining (MRC) technique. Secondly, transmit diversity, which is implemented by multiple input single output (MISO) system and based on OSTBC (Alamouti) and open and closed EO-STBC schemes. Finally, transmit and receive diversity, which is implemented by MIMO system and also based on MRC, OSTBC (Alamouti) and open and closed EO-STBC schemes [1]. EO-STBC scheme consists of operating two OSTBC (two Alamouti code) schemes in parallel.

There are two common metrics that are used to characterize the amount of spatial diversity in a conventional SISO, SIMO, MISO and MIMO systems. These two common metrics are: spatial diversity order and spatial diversity gain [2]. Spatial diversity gain and spatial diversity order may appear fundamentally different; however, spatial diversity gain refers to the number of orthogonal signal replicas being combined which is equal to $M_T \times M_R$. Whereas, the spatial diversity order mentions to the slope of the BER versus average signal-to-noise ratio (SNR) curve. In another words, increasing the diversity order means increasing the number of orthogonal independent copies at high diversity order, the probability that at least one of the copies is not experiencing a deep fade is increases. As a result, improving the quality and reliability of received signal. For example, a MIMO system with M_T transmit antennas and M_R receive antennas theoretically offers $M_T \times M_R$ diversity gain, and hence a spatial diversity order is the slope of the BER versus SNR curve. Spatial diversity order is equal to diversity gain ($M_T \times M_R$) or less. However, the correlation between the individual M_T and M_R channel gains depends on the orthogonality of the transmitted signal and antenna spacing at the transmitter and the receiver. If the channel coefficients are fully correlated, then, the spatial diversity order of the channel is one. Likewise, if the transmitted signal is not fully orthogonal, then, the spatial diversity order is also not full. To achieve full diversity order, feedback is needed [4], [5], [6].

The Focus of This Paper

Comparing BER performance of conventional SISO, open MISO (2×1), open MISO (4×1), closed MISO (4×1), open MIMO (2×2), open MIMO (4×2), and closed MIMO (4×2) systems. All systems are exploiting QPSK modulation scheme and experiencing flat fading channels. Providing system and mathematical models for each system. Moreover, from each BER performance curve measuring the diversity order and comparing it for all proposed systems.

The Organization of the Rest of This Paper

The rest of this paper is organized as follows. In Section 2, the SISO wireless system over flat fading channel, including system and mathematical models, are presented. In Section 3, the MISO wireless system over flat fading channel, including system and mathematical models of open (2x1), open (4x1), and closed (4x1) are studied. In Section 4, the MIMO wireless system over flat fading channel, including system and mathematical models of open (2x2), open (4x2), and closed (4x2) are explained. In Section 5, diversity order and diversity gain are drawn, Finally, simulation results and conclusions are given in Section 6 and Section 7, respectively.

Next, system model of SISO wireless system over flat fading channel will be presented.

2- SISO Wireless System over Flat Fading Channel

System model of SISO (1×1) system

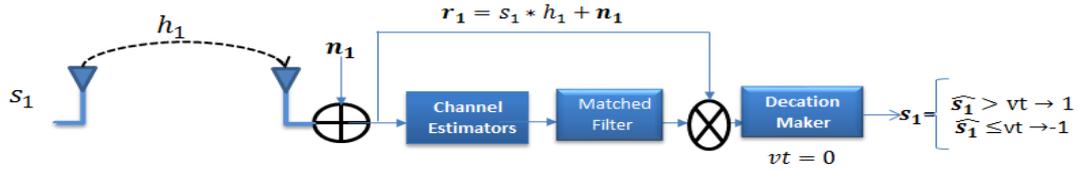


Fig. 1. Schematic system model of open loop SISO system (1Tx, 1Rx) over flat fading channel.

Mathematical model of SISO (1x1) system

The received signal can be represented as

$$r = Sh + n \quad (1)$$

Where S is the transmitted signal, n is the additive white Gaussian noise and h is the channel coefficient $h = \alpha \cdot e^{j\theta}$, then the channel state information is assumed to be known perfectly at the receiver. Then, the estimated received signal (\check{r}) before decision maker can be written as

$$\check{S} = H^H r \quad (2)$$

Then the estimated received signal \hat{S} after decision maker

$$\hat{S} = \begin{cases} \check{S} > V_t \rightarrow \hat{S} = 1 \\ \check{S} \leq V_t \rightarrow \hat{S} = -1 \end{cases} \quad (3)$$

System model of MISO wireless system over flat fading channel will be presented next.

3- MISO Wireless System over Flat Fading Channel

System model of open loop MISO (2x1) system based on OSTBC scheme

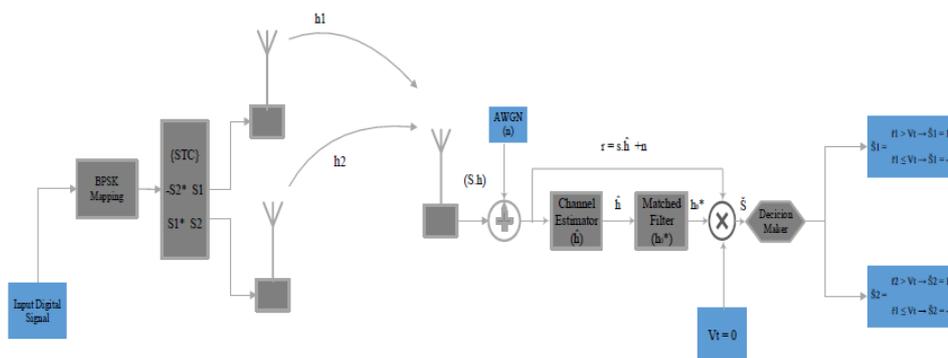


Fig. 2. Schematic system model of open loop MISO system (2Tx, 1Rx) based on OSTBC over flat fading channel.

Mathematical model of open loop MISO (2x1) system based on OSTBC scheme

In open loop MISO (2x1) system, the transmission based on OSTBC (Alamouti), which is done in two time slots, then, the received signal can be represented as:

-Time slot 1:

$$r_1 = S_1 h_1 + S_2 h_2 + n_1 \quad (4)$$

-Time slot 2:

$$r_2^* = S_1 h_2^* - S_2 h_1^* + n_2^* \quad (5)$$

The received signals in matrix form can be represented as:

$$\begin{bmatrix} r_1 \\ r_2^* \end{bmatrix} = \begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2^* \end{bmatrix} \quad (6)$$

Then, H Hermitian transpose operation can be written as:

$$H^H H = \begin{bmatrix} h_1^* & h_2 \\ h_2^* & -h_1^* \end{bmatrix} \begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} \quad (7)$$

Then, the estimated received signal can be written in matrix form as follows:

$$\begin{bmatrix} \check{S}_1 \\ \check{S}_2 \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} + \begin{bmatrix} h_1^* & h_2 \\ h_2^* & -h_1^* \end{bmatrix} \begin{bmatrix} n_1 \\ n_2^* \end{bmatrix} \quad (8)$$

The estimated received signals \check{S}_1, \check{S}_2 represented before decision maker as

$$\check{S}_1 = \alpha S_1 + h_1^* n_1 + h_2 n_2^* \quad (9)$$

$$\check{S}_2 = \alpha S_2 + h_2^* n_1 - h_1 n_2^* \quad (10)$$

Then the estimated received signal \hat{S}_1, \hat{S}_2 after decision maker

$$\hat{S}_1 = \begin{cases} \check{S}_1 > V_t \rightarrow \hat{S}_1 = 1 \\ \check{S}_1 \leq V_t \rightarrow \hat{S}_1 = -1 \end{cases} \quad \text{and} \quad \hat{S}_2 = \begin{cases} \check{S}_2 > V_t \rightarrow \hat{S}_2 = 1 \\ \check{S}_2 \leq V_t \rightarrow \hat{S}_2 = -1 \end{cases} \quad (11)$$

System model of open loop MISO (4×1) system based on EO-STBC

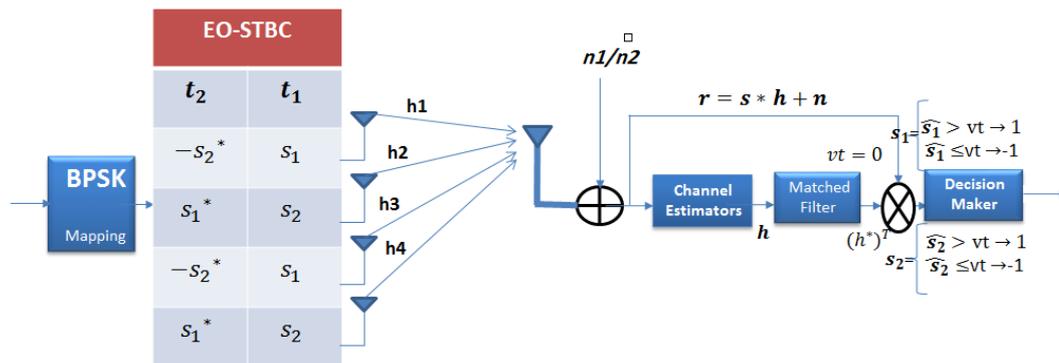


Fig. 3. Schematic system model of open loop MISO system (4Tx, 1Rx) based on EO-STBC over flat fading channel.

Mathematical model of open loop MISO (4×1) system based on EO-STBC scheme

In open loop MISO (4×1) system, the transmission based on the EO-STBC, which is done in two time slots, then, the received signal can be represented as:

-Time slot 1:

$$r_1 = S_1h_1 + S_1h_2 + S_2h_3 + S_2h_4 + n_1 \quad (12)$$

-Time slot 2:

$$r_2^* = S_1h_3^* + S_1h_4^* - S_2h_1^* - S_2h_2^* + n_2^* \quad (13)$$

The received signals r_1, r_2 at time slots 1 and 2, respectively can be represented as:

$$H^H H = \begin{bmatrix} h_1^* + h_2^* & h_3 + h_4 \\ h_3^* + h_4^* & -h_1 - h_2 \end{bmatrix} \begin{bmatrix} h_1 + h_2 & h_3 + h_4 \\ h_3^* + h_4^* & -h_1^* - h_2^* \end{bmatrix} = \begin{bmatrix} \alpha + \beta & 0 \\ 0 & \alpha + \beta \end{bmatrix} \quad (14)$$

Then, the estimated received signal before decision maker can be written in matrix form as follows:

$$\begin{bmatrix} \check{S}_1 \\ \check{S}_2 \end{bmatrix} = \begin{bmatrix} \alpha + \beta & 0 \\ 0 & \alpha + \beta \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} + \begin{bmatrix} h_1^* + h_2^* & h_3 + h_4 \\ h_3^* + h_4^* & -h_1 - h_2 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2^* \end{bmatrix} \quad (15)$$

The estimated received signals represented before decision maker as

$$\check{S}_1 = (\alpha + \beta)S_1 + (h_1^* + h_2^*)n_1 + (h_3 + h_4)n_2^* \quad (16)$$

$$\check{S}_2 = (\alpha + \beta)S_2 + (h_3^* + h_4^*)n_1 - (h_1 + h_2)n_2^* \quad (17)$$

Where:

$$\alpha = h_1 \cdot h_1^* + h_2 \cdot h_2^* + h_3 \cdot h_3^* + h_4 \cdot h_4^* \quad (18)$$

$$\beta = (h_1 \cdot h_2^* + h_2 \cdot h_1^* + h_3 \cdot h_4^* + h_4 \cdot h_3^*) \quad (19)$$

Then the estimated received signal \hat{S}_1, \hat{S}_2 after decision maker

$$\hat{S}_1 = \begin{cases} \check{S}_1 > V_t \rightarrow \hat{S}_1 = 1 \\ \check{S}_1 \leq V_t \rightarrow \hat{S}_1 = -1 \end{cases} \quad \text{and} \quad \hat{S}_2 = \begin{cases} \check{S}_2 > V_t \rightarrow \hat{S}_2 = 1 \\ \check{S}_2 \leq V_t \rightarrow \hat{S}_2 = -1 \end{cases} \quad (20)$$

System model of closed loop MISO (4x1) system based on EO-STBC scheme

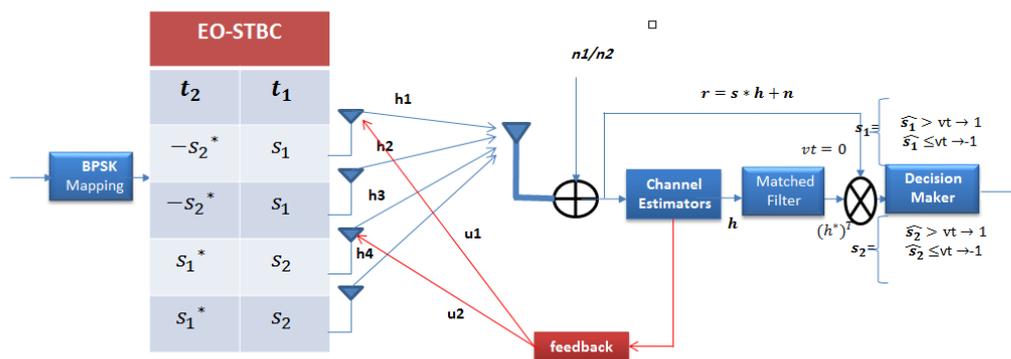


Fig. 4. Schematic system model of closed loop MISO system (4Tx, 1Rx) based on EO-STBC over flat fading channel.

Mathematical model of closed loop MISO (4x1) system based on EO-STBC scheme

In closed loop MISO (4×1) system, the transmission based on the EO-STBC, which is done in two time slots, then, the received signal can be represented as:

-Time slot 1:

$$r_1 = S_1h_1 + S_1h_2 + S_2h_3 + S_2h_4 + n_1 \quad (21)$$

-Time slot 2:

$$r_2^* = S_1h_3^* + S_1h_4^* - S_2h_1^* - S_2h_2^* + n_2^* \quad (22)$$

Then, the estimated received signals \check{S}_1, \check{S}_2 before decision maker in matrix form as follows:

$$\begin{bmatrix} \check{S}_1 \\ \check{S}_2 \end{bmatrix} = \begin{bmatrix} \alpha + \beta e^{j\theta} & 0 \\ 0 & \alpha + \beta e^{j\theta} \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} + \begin{bmatrix} h_1^* + h_2^* & h_3 + h_4 \\ h_3^* + h_4^* & -h_1 - h_2 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2^* \end{bmatrix} \quad (23)$$

The estimated received signals represented before decision maker as

$$\check{S}_1 = (\alpha + \beta e^{j\theta})S_1 + (h_1^* + h_2^*)n_1 + (h_3 + h_4)n_2^* \quad (24)$$

$$\check{S}_2 = (\alpha + \beta e^{j\theta})S_2 + (h_3^* + h_4^*)n_1 - (h_1 + h_2)n_2^* \quad (25)$$

Where:

$$\alpha = h_1 \cdot h_1^* + h_2 \cdot h_2^* + h_3 \cdot h_3^* + h_4 \cdot h_4^* \quad (26)$$

$$\beta = (h_1 \cdot h_2^* + h_2 \cdot h_1^* + h_3 \cdot h_4^* + h_4 \cdot h_3^*) \quad (27)$$

$$\theta = -\text{angle}(h_1 \cdot h_2^* + h_3 \cdot h_4^*) \quad (28)$$

As in [7], common phasor $e^{j\theta}$ is used to rotate the transmitted symbols from the first and third antennas and the rotation angle is existing in a range between 0 and 2π . It is apparent that this does not change the transmitted power. Since the phase rotation on the transmitted symbols is effectively equivalent to rotating the phases of the corresponding channel coefficients. Then, the estimated received signal \hat{S}_1, \hat{S}_2 after decision maker

$$\hat{S}_1 = \begin{cases} \check{S}_1 > V_t \rightarrow \hat{S}_1 = 1 \\ \check{S}_1 \leq V_t \rightarrow \hat{S}_1 = -1 \end{cases} \quad \text{and} \quad \hat{S}_2 = \begin{cases} \check{S}_2 > V_t \rightarrow \hat{S}_2 = 1 \\ \check{S}_2 \leq V_t \rightarrow \hat{S}_2 = -1 \end{cases} \quad (29)$$

Next, system model of MIMO wireless system over flat fading channel will be studied.

System model of open loop MIMO (2×2) system based on OSTBC scheme

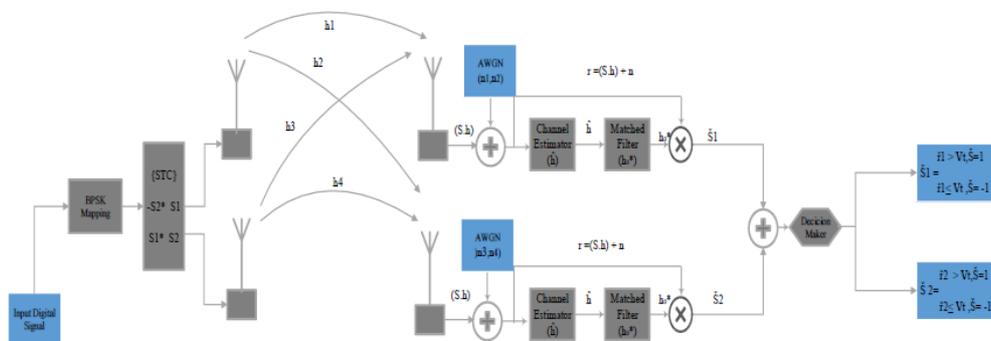


Fig. 5. Schematic system model of open loop MIMO system (2Tx, 2Rx) based on OSTBC over flat fading channel.

Mathematical model of open loop MIMO (2×2) system based on OSTBC scheme

In open loop MIMO (2×2) system, the transmission based on the OSTBC (Alamouti), which is done in two time slots, then, the received signal can be represented as:

-Time slot 1:

$$r_1 = S_1 h_1 + S_2 h_2 + n_1 \quad (30)$$

$$r_2^* = S_1 h_2^* - S_2 h_1^* + n_2^* \quad (31)$$

-Time slot 2:

$$r_3 = S_1 h_3 + S_2 h_4 + n_3 \quad (32)$$

$$r_4^* = S_1 h_4^* - S_2 h_3^* + n_4^* \quad (33)$$

The received signals in matrix form can be expressed as follows

$$\begin{bmatrix} r_1 \\ r_2^* \\ r_3 \\ r_4^* \end{bmatrix} = \begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \\ h_3 & h_4 \\ h_4^* & -h_3^* \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2^* \\ n_3 \\ n_4^* \end{bmatrix} \quad (34)$$

The part $H^H H$ yields to the following form of OSTBC, where α is the gain.

$$H^H H = \begin{bmatrix} h_1^* & h_2 & h_3^* & h_4 \\ h_2^* & -h_1 & h_4^* & -h_3 \end{bmatrix} \begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \\ h_3 & h_4 \\ h_4^* & -h_3^* \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} \quad (35)$$

The estimated received signals \check{S}_1, \check{S}_2 before decision maker

$$\begin{bmatrix} \check{S}_1 \\ \check{S}_2 \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} + \begin{bmatrix} h_1^* & h_2 & h_3^* & h_4 \\ h_2^* & -h_1 & h_4^* & -h_3 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2^* \\ n_3 \\ n_4^* \end{bmatrix} \quad (36)$$

Then,

$$\check{S}_1 = \alpha S_1 + h_1^* n_1 + h_2 n_2^* + h_3^* n_3 + h_4 n_4^* \quad (37)$$

$$\check{S}_2 = \alpha S_2 + h_2^* n_1 - h_1 n_2^* + h_4^* n_3 - h_3 n_4^* \quad (38)$$

Then the estimated received signal \hat{S}_1, \hat{S}_2 after decision maker

$$\hat{S}_1 = \begin{cases} \check{S}_1 > V_t \rightarrow \hat{S}_1 = 1 \\ \check{S}_1 \leq V_t \rightarrow \hat{S}_1 = -1 \end{cases} \quad \text{and} \quad \hat{S}_2 = \begin{cases} \check{S}_2 > V_t \rightarrow \hat{S}_2 = 1 \\ \check{S}_2 \leq V_t \rightarrow \hat{S}_2 = -1 \end{cases} \quad (39)$$

System model of open loop MIMO (4×2) system based on EO-STBC scheme

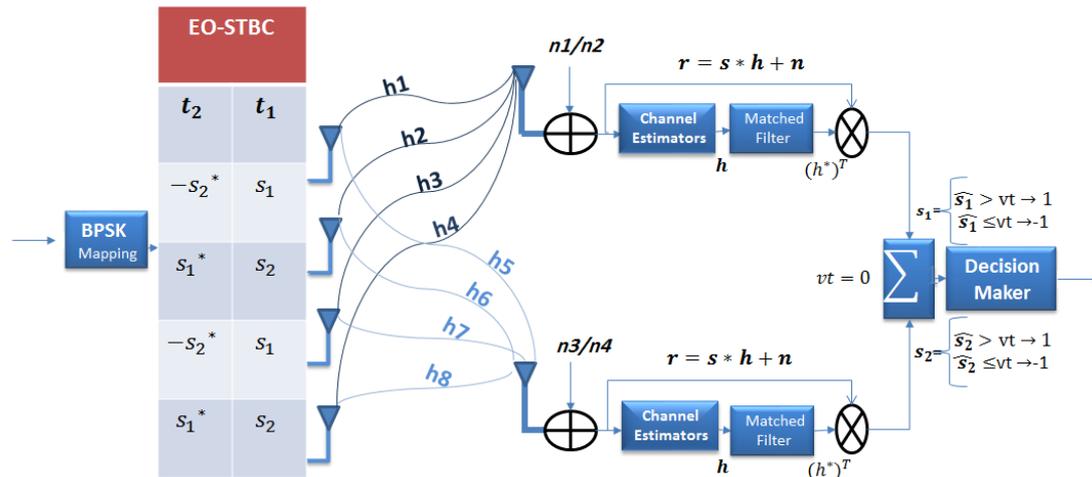


Fig. 6. Schematic system model of open loop MIMO system (4Tx, 2Rx) based on EO-STBC over flat fading channel.

Mathematical model of open loop MIMO (4x2) system based on EO-STBC scheme

In open loop MIMO (4x2) system, the transmission based on the EO-STBC, which is done in two time slots, then, the received signal can be represented as:

-Time slot 1:

$$r_1 = S_1 h_1 + S_1 h_2 + S_2 h_3 + S_2 h_4 + n_1 \quad (40)$$

$$r_2^* = S_1 h_3^* + S_1 h_4^* - S_2 h_1^* - S_2 h_2^* + n_2^* \quad (41)$$

-Time slot 2:

$$r_3 = S_1 h_5 + S_1 h_6 + S_2 h_7 + S_2 h_8 + n_3 \quad (42)$$

$$r_4^* = S_1 h_7^* + S_1 h_8^* - S_2 h_5^* - S_2 h_6^* + n_4^* \quad (43)$$

The received signals r_1, r_2, r_3, r_4 at time slots 1 and 2, respectively can be represented as

$$\begin{bmatrix} r_1 \\ r_2^* \\ r_3 \\ r_4^* \end{bmatrix} = \begin{bmatrix} h_1 + h_2 & h_3 + h_4 \\ h_3^* + h_4^* & -h_1^* - h_2^* \\ h_5 + h_6 & h_7 + h_8 \\ h_7^* + h_8^* & -h_5^* - h_6^* \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2^* \\ n_3 \\ n_4^* \end{bmatrix} \quad (44)$$

Then, H Hermitian transpose operation can be written as:

$$H^H \cdot H = \begin{bmatrix} h_1^* + h_2^* & h_3 + h_4 & h_5^* + h_6^* & h_7 + h_8 \\ h_3^* + h_4^* & -h_1 - h_2 & h_7^* + h_8^* & -h_5 - h_6 \\ h_5 + h_6 & h_7 + h_8 & h_7^* + h_8^* & -h_5^* - h_6^* \\ h_7^* + h_8^* & -h_5^* - h_6^* & -h_5^* - h_6^* & h_7 + h_8 \end{bmatrix} = \begin{bmatrix} \alpha + \beta & 0 \\ 0 & \alpha + \beta \end{bmatrix} \quad (45)$$

where α is the gain and β is the interference

$$\alpha = h_1 \cdot h_1^* + h_2 \cdot h_2^* + h_3 \cdot h_3^* + h_4 \cdot h_4^* + h_5 \cdot h_5^* + h_6 \cdot h_6^* + h_7 \cdot h_7^* + h_8 \cdot h_8^* \quad (46)$$

$$\beta = 2\text{real}(h_1 \cdot h_2^* + h_3 \cdot h_4^* + h_5 \cdot h_6^* + h_7 \cdot h_8^*) \quad (47)$$

The estimated received signals \check{S}_1, \check{S}_2 represented before decision maker as

$$\begin{bmatrix} \check{S}_1 \\ \check{S}_2 \end{bmatrix} = \begin{bmatrix} \alpha + \beta & 0 \\ 0 & \alpha + \beta \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} + \begin{bmatrix} h_1^* + h_2^* & h_3 + h_4 & h_5^* + h_6^* & h_7 + h_8 \\ h_3^* + h_4^* & -h_1 - h_2 & h_7^* + h_8^* & -h_5 - h_6 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \end{bmatrix} \quad (48)$$

$$\check{S}_1 = (\alpha + \beta)S_1 + (h_1^* + h_2^*)n_1 + (h_3 + h_4)n_2 + (h_5^* + h_6^*)n_3 + (h_7 + h_8)n_4 \quad (49)$$

$$\check{S}_2 = (\alpha + \beta)S_2 + (h_3^* + h_4^*)n_1 - (h_1 + h_2)n_2 + (h_7^* + h_8^*)n_3 - (h_5 + h_6)n_4 \quad (50)$$

Then, the estimated received signal \hat{S}_1, \hat{S}_2 , after decision maker

$$\hat{S}_1 = \begin{cases} \check{S}_1 > V_t \rightarrow \hat{S}_1 = 1 \\ \check{S}_1 \leq V_t \rightarrow \hat{S}_1 = -1 \end{cases} \quad \text{and} \quad \hat{S}_2 = \begin{cases} \check{S}_2 > V_t \rightarrow \hat{S}_2 = 1 \\ \check{S}_2 \leq V_t \rightarrow \hat{S}_2 = -1 \end{cases} \quad (51)$$

System model of closed loop MIMO (4x2) system based on EO-STBC scheme

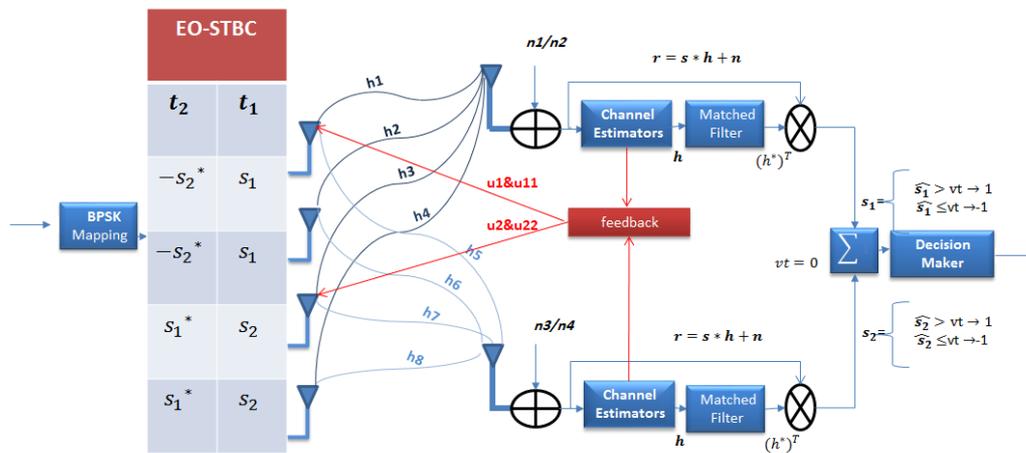


Fig. 7. Schematic system model of closed loop MIMO system (4Tx, 2Rx) based on EO-STBC over flat fading channel.

Mathematical model of closed loop MIMO (4x2) system based on EO-STBC scheme

In closed loop MIMO (4x2) system, the transmission based on the EO-STBC, which is done in two time slots, then, the received signal can be represented as:

-Time slot 1:

$$r_1 = S_1h_1 + S_1h_2 + S_2h_3 + S_2h_4 + n_1 \quad (52)$$

$$r_2^* = S_1h_3^* + S_1h_4^* - S_2h_1^* - S_2h_2^* + n_2^* \quad (53)$$

-Time slot 2:

$$r_3 = S_1h_5 + S_1h_6 + S_2h_7 + S_2h_8 + n_3 \quad (54)$$

$$r_4^* = S_1h_7^* + S_1h_8^* - S_2h_5^* - S_2h_6^* + n_4^* \quad (55)$$

The received signals r_1, r_2, r_3, r_4 at time slots 1 and 2, respectively can be represented as

$$\begin{bmatrix} r_1 \\ r_2^* \\ r_3 \\ r_4^* \end{bmatrix} = \begin{bmatrix} h_1 + h_2 & h_3 + h_4 \\ h_3^* + h_4^* & -h_1^* - h_2^* \\ h_5 + h_6 & h_7 + h_8 \\ h_7^* + h_8^* & -h_5^* - h_6^* \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2^* \\ n_3 \\ n_4^* \end{bmatrix}. \quad (56)$$

Then, H Hermitian transpose operation can be written as:

$$H^H \cdot H = \begin{bmatrix} h_1^* + h_2^* & h_3 + h_4 & h_5^* + h_6^* & h_7 + h_8 \\ h_3^* + h_4^* & -h_1 - h_2 & h_7^* + h_8^* & -h_5 - h_6 \\ h_5 + h_6 & h_7 + h_8 & h_7^* + h_8^* & -h_5^* - h_6^* \\ h_7^* + h_8^* & -h_5^* - h_6^* & -h_5^* - h_6^* & h_7 + h_8 \end{bmatrix} \begin{bmatrix} \alpha + \beta & 0 \\ 0 & \alpha + \beta \end{bmatrix} \quad (57)$$

Where, α is the gain and β is the interference.

$$\alpha = h_1 \cdot h_1^* + h_2 \cdot h_2^* + h_3 \cdot h_3^* + h_4 \cdot h_4^* + h_5 \cdot h_5^* + h_6 \cdot h_6^* + h_7 \cdot h_7^* + h_8 \cdot h_8^* \quad (58)$$

$$\beta = 2\text{real}(h_1 \cdot h_2^* + h_3 \cdot h_4^* + h_5 \cdot h_6^* + h_7 \cdot h_8^*) \quad (59)$$

The phasor $e^{j\theta}$ is used to rotate the transmitted symbols from the first and third antennas. The rotation angle is existing in a range between 0 and 2π . Then, the estimated received signals \check{S}_1, \check{S}_2 represented before decision maker as

$$\begin{bmatrix} \check{S}_1 \\ \check{S}_2 \end{bmatrix} = \begin{bmatrix} \alpha + \beta e^{j\theta} & 0 \\ 0 & \alpha + \beta e^{j\theta} \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} + \begin{bmatrix} h_1^* + h_2^* & h_3 + h_4 & h_5^* + h_6^* & h_7 + h_8 \\ h_3^* + h_4^* & -h_1 - h_2 & h_7^* + h_8^* & -h_5 - h_6 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2^* \\ n_3 \\ n_4^* \end{bmatrix} \quad (60)$$

$$\check{S}_1 = (\alpha + \beta e^{j\theta})S_1 + (h_1^* + h_2^*)n_1 + (h_3 + h_4)n_2^* + (h_5^* + h_6^*)n_3 + (h_7 + h_8)n_4^* \quad (61)$$

$$\check{S}_2 = (\alpha + \beta e^{j\theta})S_2 + (h_3^* + h_4^*)n_1 - (h_1 + h_2)n_2^* + (h_7^* + h_8^*)n_3 - (h_5 + h_6)n_4^* \quad (62)$$

Then, the estimated received signal \hat{S}_1, \hat{S}_2 , after decision maker

$$\hat{S}_1 = \begin{cases} \check{S}_1 > V_t \rightarrow \hat{S}_1 = 1 \\ \check{S}_1 \leq V_t \rightarrow \hat{S}_1 = -1 \end{cases} \quad \text{and} \quad \hat{S}_2 = \begin{cases} \check{S}_2 > V_t \rightarrow \hat{S}_2 = 1 \\ \check{S}_2 \leq V_t \rightarrow \hat{S}_2 = -1 \end{cases}. \quad (63)$$

Diversity order and diversity gain will be explained next.

4- Diversity Order and Diversity Gain

Diversity is important metric in random fading environments and it is a key factor influencing the relationship between the link quality such as average BER, and system performance such as throughput. The slope of the SNR versus BER, in the region of small BER, is often used to characterize the diversity order [8]. To quantifying the benefits of diversity is to use the slope of the curve obtained by plotting BER on a logarithmic scale versus average SNR in dB, when SNR gets large the curve is becoming linear. The resulting slope is defined as the real diversity gain of the system, which it denoted by G_d . the theoretical BER can express as in [4]:

$$P_b = \zeta [G_c(\overline{\text{SNR}})]^{-G_d}, \quad (64)$$

where ζ is a constant depends on the type of modulation scheme, G_c is a coding gain constant of the system, and the bar over SNR denotes the mean.

$$P_{b1} - P_{b2} = -G_d(\rho_1 - \rho_2), \quad (65)$$

where ρ_i represents SNRs and P_{bi} represents bit error probability.

Simulation results will be presented in next section.

5- Simulation Results of SISO, MISO, and MIMO Wireless Systems

Comparison the simulation results between SISO (1×1), open MISO (2×1), open MISO (4×1), closed MISO (4×1), open MIMO (2×2), open MIMO (4×2), and closed MIMO (4×2) systems

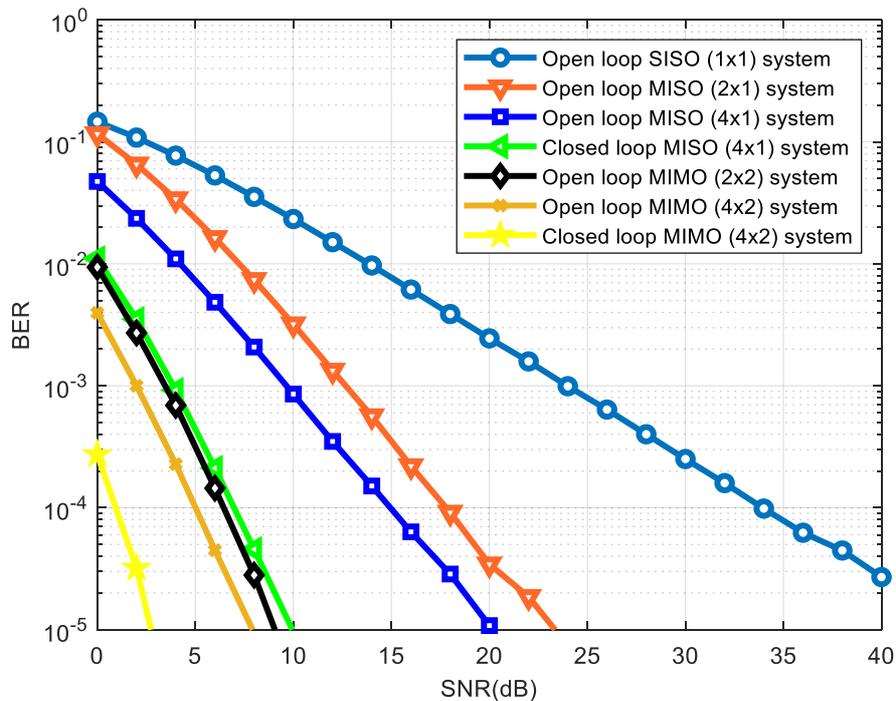


Fig. 8. Comparison the Simulation results between SISO (1×1) system, open loop MISO (2×1), open loop MISO (4×1), closed loop MISO (4×1) systems, MIMO (2×2), open loop MIMO (4×2), closed loop MIMO (4×2) systems with QPSK modulation scheme, over flat fading channel.

Figure 8 illustrate comparison of the simulation results between conventional SISO (1×1) system, open loop MISO (2×1), open loop MISO (4×1), closed loop MISO (4×1) systems, MIMO (2×2), open loop MIMO (4×2), closed loop MIMO (4×2) systems. All systems are experience flat fading channel and exploit QPSK modulation scheme. From the graph, it is clear that the conventional SISO (1×1) system has the worst BER performance because, according the equation (65), has the diversity order is equal just one. Whereas, closed loop MIMO (4×2) system has the best BER performance. Also, according equation (106), the diversity order equal to diversity gain which is equal eight. Moreover, open loop MISO (2×1) system based on OSTBC (Alamouti) scheme and open loop MISO (4×1) based on EO-STBC scheme achieve the same diversity order which is equal two, but with a 3 dB penalty in performance of MISO (4×1) system. Whereas, closed loop MISO (4×1) based on EO-STBC scheme has the same diversity order as open loop MIMO (2×2) without penalty which means the both curves are almost identical, the diversity order of both is equal four. In addition, open loop MISO (2×2) system based on OSTBC (Alamouti) scheme, closed loop MISO (4×1) based on EO-STBC scheme and open loop MIMO (4×2) based on EO-STBC scheme achieve the same diversity order which is equal four, but with improving by about 3 dB in performance of open loop MIMO (4×2) system. To conclude, MISO and MIMO systems based on OSTBC (Alamouti) have full diversity order and full rate. In contrast, open loop MISO and open loop MIMO systems based on EO-STBC scheme have approximately half diversity order

and half rate. Feedback has been used to improve just only spatial diversity gain. Therefore, closed loop MISO and closed loop MIMO systems based on EO-STBC scheme have full diversity gain and half rate.

6- Conclusions

In this paper, we have studied the comparison of the simulation results between conventional SISO (1×1) system, open loop MISO (2×1), open loop MISO (4×1), closed loop MISO (4×1) systems, open loop MIMO (2×2), open loop MIMO (4×2), closed loop MIMO (4×2) systems. The maximal diversity gain achievable is the diversity order of the channel $M_T \times M_R$. All systems are experience flat fading channel and they exploit QPSK modulation scheme. Conventional SISO (1×1) system had the worst BER performance with diversity order equal just one. Whereas, MISO and MIMO systems based on OSTBC (Alamouti) have had full diversity gain and full rate. Although system with feedback had overhead drawback, feedback had improved MISO and MIMO systems based on EO-STBC scheme to achieve full diversity gain. Cooperative MIMO system based on OSTBC (Alamouti), open and closed loop EO-STBC schemes, decode-and-forward (DF) and amplify-and-forward (AF) relaying protocols, is a subject of our ongoing study.

7- Acknowledgment

The authors wish to thank The Collage of Electronic Technology-Bani Walid / Libya, The Collage of Electronic Technology-Tripoli / Libya and Alghabel Algharbi University / Libya for their cooperation.

References

- [1]. Andrea Goldsmith, Wireless Communications. Cambridge University Press. 2005.
- [2]. E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, MIMO Wireless Communications. New York, NY, USA: Cambridge University Press, 2007.
- [3]. Rodger E. Ziemer, and William H. Tranter, Principles of communication: systems, modulation, and noise. Seventh edition, John Wiley & Sons, Inc. 2015.
- [4]. J. R. Hampton, Introduction to MIMO Communications. Cambridge University Press, 2014.
- [5]. H. Jafarkhani, Space-Time Coding: Theory and Practical. Cambridge University Press, 2005.
- [6]. Jeffrey G. Andrews, Arunabha Ghosh, Rias Muhamed, Fundamentals of WiMAX: understanding broadband wireless networking. Pearson Education, Inc. 2007.
- [7]. N. Eltayeb, S. Lambotharan, and J. Chambers, "A Phase Feedback Based Extended Space-Time Block Code for Enhancement of Diversity," in Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th, pp. 2296–2299, 2007.
- [8]. A. B. Gershman and N. D. Sidiropoulos, Space-Time Processing for MIMO Communications. John Wiley & Sons Ltd, 2005.



Track No. 6: Engineering Management and Quality

تقدير الكلفة الخفية لإنحراف خيوط الصوف الخالص المستخدم في صناعة السجاد (شركة الإنماء للصناعات الصوفية بني وليد)

علي احمد علي مصباح¹، عبدالرزاق فرج امحمد²، محمد صالحين قريميدة³، سهيلة التومي الشوشان⁴، عبدالسلام علي احمد دومه⁵
^{1,2} مركز البحوث الهندسية وتقنية المعلومات، بني وليد، ليبيا
شركة الإنماء للصناعات الصوفية ، بني وليد، ليبيا
منسق الجودة بقسم الإحصاء والاقتصاد القياسي، كلية الاقتصاد والعلوم السياسية، جامعة طرابلس، ليبيا
قسم الهندسة الميكانيكية والصناعية، جامعة بني وليد، ليبيا
resqe2010@gmail.com ² Aaammm198088@gmail.com ¹
³mohd.salaheen@gmail.com ⁴ souhaila.a.mhamed11@gmail.com
⁵abdussalam.a.ahmed@gmail.com

ABSTRACT

تعتبر صناعة الغزل والنسيج من اقدم الصناعات التي عرفها الانسان والتي تأثرت بالتطور التكنولوجي السريع الحاصل في تطور طرق ووسائل الانتاج والذي جاء تلبية لتطور رغبات واحتياجات المستهلك. انه من الطبيعي في حال التحدث عن رغبات واحتياجات متطورة واساليب وطرق انتاج حديثة ان تكون هناك وسائل وطرق قياس مواكبة لهذا التطور لتحديد مدى مطابقة المنتجات لما تم التعارف عليه باسم المواصفات القياسية.

Keywords:

هدفت هذه الدراسة الى تقدير التكلفة الخفية لعمليات انتاج خيوط الصوف الخالص المستخدمة في صناعة السجاد المنسوج ذات الرقم المتري 8/1 والمنتجة من قبل شركة الانماء للصناعات الصوفية - بني وليد، وذلك عن طريق اخذ عينات من خيوط الصوف ووزنها لغرض تحديد مطابقة رقم الخيط للمواصفات الليبية ومن ثم تطبيق نموذج تاجوشي لتقدير التكلفة الخفية الناتجة عن انحراف هذه العينات عن المواصفات القياسية. ولقد توصلت الدراسة الى ان هناك فقد كبير في كميات الصوف الخالص والمستخدم في انتاج خيوط الصوف نتيجة انحراف العينات عن القيمة المستهدفة في المواصفة وبناء على هذه النتائج فان الدراسة توصي بتركيز البحث في جميع مراحل الانتاج لتقدير التكاليف الكلية واستخدام التقنيات الحديثة لتطوير عمليات مراقبة الانتاج.

1- المقدمة

تعتبر صناعة الغزل والنسيج من أقدم الصناعات التي عرفها الانسان والتي تأثرت بالتطور التكنولوجي السريع الحاصل في تطور طرق ووسائل الانتاج والذي جاء تلبية لتطور رغبات واحتياجات المستهلك.

في ظل التطورات التي عرفتها المؤسسات المعاصرة في مختلف المجالات، اقتصادية، اجتماعية، او سياسية، كان من الضروري ان تسعى هذه المؤسسات الى مسايرة التقدم الفني و التكنولوجي في وسائل الانتاج المختلفة، ومحاولة تصحيح الانظمة الكلاسيكية المتبعة في جميع المستويات خاصة مع كبر حجم المشاريع وتنوع المنتجات وزيادة شدة المنافسة¹. كما أنه في حال التحدث عن رغبات واحتياجات متطورة واساليب وطرق إنتاج حديثة لابد أن تكون هناك وسائل وطرق قياس مواكبة لهذا التطور لتحديد مدى مطابقة المنتجات لما تم التعارف عليه باسم المواصفات القياسية، حيث تعد هذه المواصفات من قبل لجان تقوم بدراسة كل الموجودات والمتغيرات والتنبؤ بما هو آت ان امكن وتصدرها مراكز وهيئات معتمدة من الدولة حيث ان لكل دولة نظام من المواصفات خاص بها قد يتقاطع في بعض الاحيان مع مواصفات دول اخرى.

2- مشكلة الدراسة وهدفها.

تعاني شركة الانماء لصناعة السجاد من تكاليف لا يمكن حسابها وتحديدها في العمليات المحاسبية وهذه التكاليف هي عبارة عن تكاليف خفية ناتجة عن انحراف وزن الخيط عن القيمة المستهدفة في المواصفات والتي تقع ضمن حدود المواصفة.

3- الدراسات السابقة:

تكاليف الجودة وأثرها في تسويق المنتجات، (دراسة ميدانية على مصنع زينة للمنتجات الورقية) للباحث حسام الدين بابكر علي عبدالله تناول البحث تكاليف الجودة وأثرها في تسويق المنتجات، حيث اعتمد الباحث على تكاليف الجودة كمتغير مستقل وتسويق المنتجات كمتغير تابع. هدف البحث الي التعرف على تكاليف الجودة، ومعرفة العلاقة بين تكاليف الجودة وتسويق المنتجات، واثبات العلاقة بين جودة المنتجات وتسويقها من خلال الاهتمام بتكاليف الجودة. واعتمد البحث على المنهج التاريخي والاستنباطي والاستقرائي والتحليلي. توصل البحث الي عدة نتائج منها تركيز البرنامج التسويقي على جودة المنتج يعطي المنشأة ميزة تنافسية، وان المنتجات ذات الجودة العالية يسهل توزيعها، وان الاهتمام بتكاليف المنع والتقييم تساعد المنشأة على إيجاد قنوات جديدة لتوزيع المنتجات في أسواق جديدة، اون التركيز على تكاليف المنع والتقييم يزيد من رضا العملاء.

استخدام أنموذج تاجوشي في تقدير التكاليف الخفية للابتعاد عن مواصفات الجودة المستهدفة بالتطبيق على مصنع الغزل والنسيج في الموصل لـ مثنى فالح الزيدي خالص حسن الناصر، يهدف البحث بصورة أساسية إلى تحديد من هج محاسبي سليم لتقدير خسائر الجودة التي غالباً ما يتم تجاهلها وعدم تسجيلها في سجلات التكاليف، فضلاً عن بيان أهمية هذه التكاليف وتحديد مكوناتها. وقد تم وضع فرضية تتناسب مع من هجية التحليل في أن استخدام نموذج دالة خسارة الجودة يمكن أن يقدم منهجاً سليماً لإعطاء تقديرات معقولة لهذه التكاليف ويقدم معلومات مفيدة عن سبل تخفيضها. شمل الجانب العملي للبحث تقدير خسارة الجودة المتحققة لعينة من منتجات مصنع الغزل والنسيج في مدينة الموصل إذ قدمت النتائج التي توصل لها البحث دعماً لفرضية البحث، وعُرضت مجموعة من الاستنتاجات كان أبرزها أن أنموذج دالة خسارة الجودة يمكن أن يوفر مقاييس أداء لتقييم قدرة المصنع على اكتساب الفرصة لتخفيض الانحراف عن مواصفات الجودة المستهدفة.

واقع تكاليف الجودة في الشركات الصناعية الفلسطينية (دراسة تطبيقية) للباحث خالد سامي حمودة، هدفت هذه الدراسة إلى توضيح المفاهيم المتعلقة بالجودة وتكاليفها، وذلك من خلال التعرف على واقع تكاليف الجودة في شركات صناعة الأغذية الفلسطينية في قطاع غزة، من حيث مدى إدراك إدارة هذه الشركات بأهمية قياس تكاليف الجودة، ومدى تطبيق أنظمة تكاليف الجودة، ومدى اهتمامهم بأنواع ومسببات هذه التكاليف، ولتحقيق أهداف الدراسة تم استخدام الأسلوب الوصفي التحليلي لذلك تم تصميم استبانة وتوزيعها على مجتمع الدراسة البالغ عددها 38 شركة صناعة أغذية، وتم تحليل بيانات النتائج الأولية باستخدام برنامج الإحصائي (SPSS) واستخدام الاختبارات الإحصائية المناسبة. وخلصت الدراسة إلى عدة نتائج أهمها ما يلي: يوجد إدراك لدى إدارات الشركات الصناعية الفلسطينية لأهمية تكاليف الجودة بدرجة كبيرة بلغت %76.48 لجميع الأبعاد الأربعة لتكاليف الجودة، وتطبق إدارات الشركات الصناعية الفلسطينية نظام تكاليف الجودة بدرجة كبيرة بلغت %70.16 لجميع الأبعاد الأربعة لتكاليف الجودة، وتهتم إدارات الشركات الصناعية الفلسطينية بسياسات تخفيض تكاليف الجودة بدرجة كبيرة بلغت %71.20 لجميع الأبعاد الأربعة لتكاليف الجودة.

4- مفهوم الجودة.

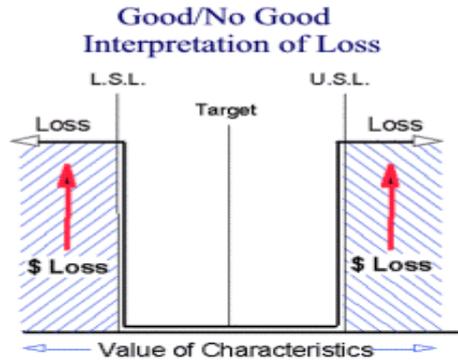
الجودة: هي مدى مطابقة المنتج للمواصفات القياسية، ويمكن تناول هذا التعريف من جانبين هما:

- جودة التصميم (Quality of Design) وهي ان جميع السلع والخدمات يتم انتاجها بدرجات او مستويات مختلفة من الجودة، وهذه الاختلافات في الدرجات أو مستويات الجودة متعمدة كما في صناعة السيارات.
- جودة المطابقة (Quality of conformance) تعني مدى مطابقة المنتج للمواصفات المطلوبة في التصميم وتتأثر جودة المطابقة بعدد من العوامل، بما في ذلك اختيار عمليات التصنيع، وتدريب القوى العاملة والإشراف عليها، والاختبارات، وأنشطة التفتيش المستخدمة، ومدى اتباع هذه الإجراءات، ودوافع القوى العاملة لتحقيق الجودة.

يمكن أن تُعرف الجودة على انها درجة الانسجام بين توقعات الزبون والمنتجات من السلع والخدمات وادراكهم لتلك التوقعات، والجودة تتناسب عكسياً مع التباين ويقصد بذلك أنه في حالة انخفاض التباين في الخصائص المهمة للمنتج، تزداد جودة المنتج.
(2)

5- مفهوم تكاليف الجودة

إن عدم مطابقة المنتجات للمواصفات المحددة يعد خسارة للمؤسسة لأن هذه المنتجات إما أن تكون تالفة أو تحتاج إلى إعادة تأهيل، أو أنها لا تلبي احتياجات المستهلك مما يترتب على كل ذلك تكاليف إضافية ترهق المؤسسة وتقلل من فرصها التنافسية، وفي واقع الامر؛ تكون المنتجات إما مطابقة تماما للقيمة المحددة في المواصفة (Target) أو تكون في ضمن النطاق المعروف بحدود المواصفة اي حديها العلوي والسفلي (Upper and lower Limits) وهذه المنتجات تعد مقبولة وفق اشتراطات المواصفة وانها لا تمثل خسارة للمؤسسة وهذا ما يعرف بالأسلوب التقليدي ويمكن توضيحه كما في الشكل الاتي:



يوضح المفهوم التقليدي

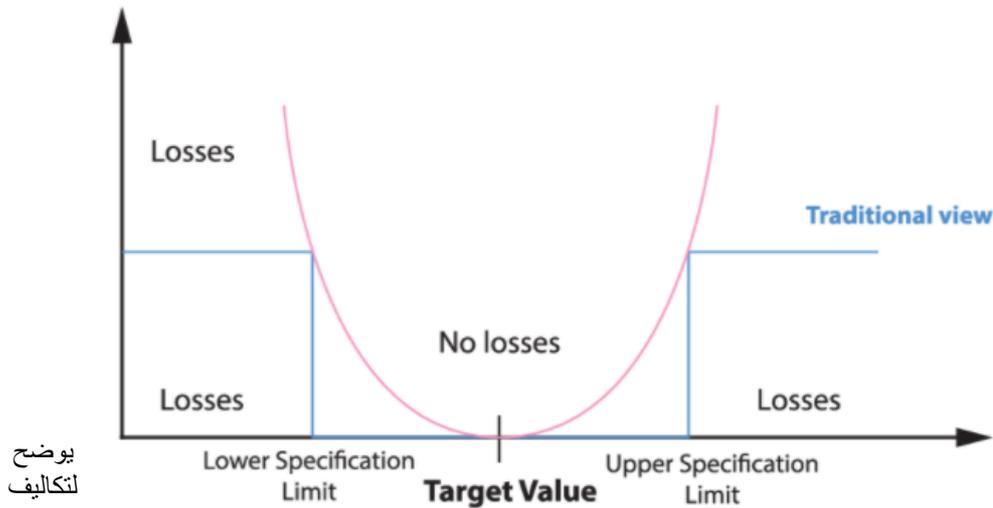
شكل 1

لتكاليف الجودة [3]

6- التكاليف الخفية للجودة.

تعرف تكاليف الجودة على انها مجموع التكاليف التي يتم انفاقها في المنظمة لضمان تقديم المنتجات من السلع والخدمات الى الزبون حسب متطلباته ورغباته. عادة ما تشمل هذه التكاليف تكاليف الاخفاقات التي تحدث نتيجة عدم الوفاء بمتطلبات الجودة على المستوى الداخلي للمنظمة او خارجها [4]

قام تاجوشي (Genichi Taguchi) بتقديم نموذج مطور لتقدير تكاليف الجودة حيث اعتبر ان اي مخالفة للهدف هو عبارة عن تكلفة اضافية تتناسب مع ابتعاد الخاصية عن الهدف والشكل الاتي يوضح هذا الانموذج.



يوضح
لتكاليف

شكل 2:
تاجوشي
الجودة [5]

7- تقدير التكاليف الخفية لجودة خيط الصوف الخالص رقم 8/1.

للوصول الى المنتج النهائي وهو السجاد المنسوج تمر المواد الخام بالعديد من العمليات وهي الفرز والغسيل والصبغة والغزل وغيرها، والنسيج والتجهيز النهائي، ويتم انتاج خيوط الصوف الخالص والمستخدم في انتاج السجاد المنسوج كإحدى العمليات التي تتم في قسم الغزل والذي تعتبر مخرجاته مواد اولية للقسم الذي يليه الا وهو النسيج، وتتم مراقبة جودة انتاج الخيوط عن طريق وزن كومة الخيوط البالغ طولها 100 متر ومن ثم حساب نمرة الخيط باستخدام المعادلة الاتية:

$$\text{wool sil number(metric)} \left(\frac{\text{meter}}{\text{gram}} \right) = \frac{\text{Silk length (meter)}}{\text{Silk weight (gram)}}$$

تشير المواصفة الليبية رقم (م م ق ل 454:2009) إلى هذه المعادلة واستخدامها في حساب رقم الخيط كالآتي:

$$\text{wool silk number(metric)} = \frac{100 (\text{meter})}{12.5(\text{gram})} = 8 \left(\frac{\text{meter}}{\text{gram}} \right)$$

كما تشير المواصفة الليبية رقم 2009:454 الى ان حدود المواصفة المسموح بها لرقم الخيط هي مساوية 0.5 % على جانبي القيمة المستهدفة والتي تساوي 8 متر/ جرام^[6]

الجدول رقم (1) يبين نتائج حساب رقم الخيط وهذه النتائج متحصل عليها من سجلات الشركة ولكي تكون قيمة التكاليف أكثر منطقية تم احتساب وزن هذه العينات اعتمادا على المعادلة السابقة، ولتلبية متطلبات دالة تاجوشي لحساب التكاليف الخفية تم حساب المتوسط الحسابي والانحراف المعياري للعينات، واستكمالا لمتطلبات دالة تاجوشي تم إنشاء الجدول رقم (2) والذي يتضمن حدي المواصفة الأدنى والأعلى والذين تم حسابهما وفقا للمعادلة الآتية:

$$\text{Weight (g)} = \frac{100 m}{8 \frac{m}{g}}$$

جدول رقم (1) يبين رقم الخيط بالنظام المتري ووزنه.

الوزن (جم)	رقم الخيط (متري)	العينة	الوزن (جم)	رقم الخيط (متري)	العينة
12.34	8.10	11	14.43 out	6.93	1
12.31	8.12	12	12.90	7.75	2
13.05	7.66	13	13.14	7.61	3
13.08	7.64	14	12.06	8.29	4
13.02	7.68	15	12.72	7.86	5
12.95	7.72	16	12.95	7.72	6
12.51	7.99	17	13.51	7.40	7
12.53	7.98	18	12.80	7.81	8
12.87	7.77	19	12.50	8.00	9
			12.69	7.88	10
12.77	الوزن (جم)	الوسط الحسابي Y (رقم الخيط بالنظام المتري)			
0.353	الوزن (جم)	الانحراف المعياري S ² (رقم الخيط بالنظام المتري)			

تم تطبيق المعادلة على جميع العينات لحساب وزن العينة باعتباره هو العنصر المتغير الوحيد وان رقم الخيط وطوله هما عبارة عن قيم ثابتة بحسب المواصفة (2009:453).^[7]

إن سجلات الشركة تشير الى ان تكلفة الكيلوجرام الواحد من خيوط الصوف الخالص (C) تساوي 6.5 دينار، وتم احتساب قيمة الثابت K باستخدام المعادلة الآتية:

$$K = \frac{C}{D}$$

جدول رقم (2) حدود المواصفة.

المواصفات / المتغيرات	رقم الخيط بالنظام المتري
الحد الاعلى للمواصفة (U)	13.125
الحد الادنى للمواصفة (L)	11.875
القيمة المستهدفة (T)	12.5
الحد الاعلى - القيمة المستهدفة (U-T)	0.652
القيمة المستهدفة - الحد الادنى (T-L)	0.652
الانحراف عن القيمة المستهدفة (D)	0.652
قيمة الخسارة بالدرهم /جم (C)	6.5
قيمة الثابت (K)	10.4

يتم احتساب التكاليف الخفية لمرحلة انتاج خيوط الصوف الخالص رقم (8/1) باستخدام معادلة تاجوشي كالآتي:

$$Loss(Y)_{avg} = K[S^2 + (Y - T)^2]$$

$$Loss(Y)_{avg} = 10.4[0.124 + (12.77 - 12.5)^2] = 4.42 \left(\frac{\text{د.م}}{\text{جم}}\right) = 2.04 \text{ دينار/كجم}$$

ان هذه القيمة تمثل الخسارة بالدينار لكل كيلوجرام من خيوط الصوف الخالص المنتج بالشركة، وبحسب السجلات فان الشركة تنتج ما يقارب 3000 كيلو جرام من خيوط الصوف شهرياً، ولاحتساب الخسارة السنوية اجريت المعادلة الحسابية الآتية:

الخسارة الكلية السنوية = متوسط الإنتاج الشهري من خيوط الصوف * عدد الأشهر بالسنة * خسارة الوحدة المحسوبة = الخسارة الكلية.

$$\text{الخسارة الكلية السنوية} = 3000(\text{كجم}) *$$

$$12(\text{شهر}) * 2.04 (\text{دينار/كجم}) = 73719.36 (\text{دينار}).$$

8- النتائج.

1. إن عدم مطابقة المنتجات للمواصفات المحددة يعد خسارة للمؤسسة لأن هذه المنتجات إما أن تكون تالفة أو تحتاج إلى إعادة تأهيل، أو أنها لا تلبى احتياجات المستهلك مما يترتب على كل ذلك تكاليف إضافية ترهق المؤسسة وتقلل من فرصها التنافسية.
2. تم تطبيق نموذج تاجوشي (Genichi taguchi) على خيوط الصوف الخالص المستخدمة في إنتاج السجاد.
3. أظهرت نتائج تطبيق نموذج تاجوشي أن الانحراف عن القيمة المستهدفة داخل حدود المواصفة يسبب كلف خفية وخسائر كبيرة في كميات الصوف المستخدم.
4. تتأثر جودة المنتج النهائي من السجاد بجودة سمك الخيط والذي يؤدي بانحرافه عن القيمة المستهدفة لانحراف جودة المنتج عن المعدلات القياسية للشركة.
5. تم احتساب التكاليف الخفية لمرحلة انتاج خيوط الصوف الخالص عن طريق معادلة تاجوشي وأظهرت النتائج خسائر قدرت ب (2.04) دينار لكل كيلو جرام صوف.
6. تم حساب الخسارة السنوية بسبب انحراف خيوط الصوف عن القيمة المستهدفة داخل المواصفة والتي بلغت (73719.36) دينار سنوياً.

9- التوصيات

لأجل تطوير شركة الانماء لإنتاج السجاد وضمان مكانتها التنافسية في السوق المحلي الذي اكتسبت المنتجات المشابهة فيه مكانة تنافسية فان الدراسة توصي بالاتي:

- 1- التأكيد على مطابقة المنتجات للمواصفات القياسية الليبية.
- 2- التأكيد على استخدام نظريات واساليب جديدة لاحتساب التكلفة الخفية.
- 3- التأكيد على تدريب العناصر البشرية وتأهيلها التأهيل المناسب في مجال مراقبة الجودة وحساب التكاليف.
- 4- اجراء الدراسات على جميع مراحل الانتاج ابتداء من الفرز والغسيل مرورا بالصباغة والغزل وانتهاء بالنسيج والتجهيز النهائي لاحتساب التكاليف الخفية لأجل تقاؤها .
- 5- العمل على تصميم برنامج حاسوبي يتضمن دالة تاجوشي لحساب التكاليف الخفية في جميع المراحل بشكل فوري ودقيق.

المراجع:

- [1]. حكوم حنان و روقي ابراهيم, الطرق الحديثة لحساب التكاليف , دراسة حالة , مؤسسة نفضال, صفحة(ب).
- [2]. Introduction to statistical quality control, Douglas C. Montgomery, sixth edition, page (06).
- [3]. <https://www.shmula.com/wp-content/bizpi.jpg>
- [4]. رعد هاشم جاسم, علاقة تكاليف الجودة بطريقة تاكوشي واثرها في تخفيض التكاليف, ورقة بحثية, مجلة الادارة والاقتصاد, العدد68, سنة 2008, الصفحة(06).
- [5]. <https://pdcahome.com/english/wp-content/loss-taguchi.png>
- [6]. المركز الوطني للمواصفات والمعايير القياسية, الخيوط (الغزول) الطبيعية والصناعية, م م ق ل (2009:454), الصفحة(03).
- [7]. المركز الوطني للمواصفات والمعايير القياسية, السجاد المنسوج اليا ووبرته من الصوف الخالص, م م ق ل (2009:453), الصفحة (03).

دراسة أثر استخدام استراتيجيات التنفيذ المختلفة للمشروعات الانشائية على أداء الاطراف الرئيسية في مشاريع الانشاءات باستخدام نموذج الأداء العام

*م. أسماء يوسف بن حميد¹، *د. وسام حسن بيت المال²

¹قسم الهندسة المدنية، كلية الهندسة، جامعة مصراته، ليبيا

²قسم الهندسة المدنية والبيئية، كلية الهندسة، جامعة ظفار، سلطنة عمان

²قسم الهندسة المدنية، كلية الهندسة، جامعة مصراته، ليبيا

¹asma.mse@eng.misuratau.edu.ly, ²wbeitelmal@du.edu.om

ABSTRACT

تهدف هذه الورقة البحثية إلى دراسة تأثير اختلاف استراتيجيات تنفيذ المشروعات الانشائية على أداء المسوقات (Drivers) (أصحاب المصلحة) أثناء مراحل المشروع، وعلى الرغم من وجود مجموعة من الخيارات لاستراتيجيات تنفيذ المشروعات، إلا أن العديد من مشروعات البناء محلياً لا تلبى توقعات مالكيها نظراً لاستخدامهم تقليدياً استراتيجيات واحدة لتنفيذ العقود الانشائية وهي: (التصميم، العطاء، التنفيذ). حيث أن مالكي المشروعات عادة ما يطمحون إلى نتائج مرضية للمشروع من ناحية الوقت والتكلفة ورضى مستخدمي المشروع وذلك بالتدخل والإشراف المباشر على كل أطراف للمشروع المختلفة. هذا البحث هو دراسة لتشجيع الخبراء المحليين للمساهمة في فتح الأفق على الاستراتيجيات المختلفة والمساهمة لتحسين مخرجات المشروعات الهندسية. حيث تم تكوين نموذج محاكاة باستخدام نموذج الأداء العام (General Performance Model) والمستخدم في العديد من البحوث العالمية المماثلة ذات العلاقة لدراسة القرارات الاستراتيجية ولتقييم التفاعلات بين العديد من المتغيرات الداخلة في تنفيذ المشروعات. تم عرض النموذج في ورشة عمل تضم ممارسين للمهنة لتحديد تأثير كل استراتيجية على أداء المسوقات. حيث اتضح ان هناك تأثير متفاوت للاستراتيجيات المختلفة على كل الأطراف المشاركة في المشروع الانشائي والهندسي (المسوقات) حيث وجد أن استراتيجية التصميم-العطاء-الانشاء لها تأثير إيجابي على أداء مدير المشروع. بينما استراتيجية التصميم-الانشاء تؤثر إيجابياً في أداء كل من المنفذ والمصمم/الاستشاري.

Keywords:

1. المقدمة

تنفيذ المشروع هو عملية شاملة بما في ذلك التعريف وتحديد النطاق والتصميم والبناء اللازمة لتنفيذ وإكمال منشأة بناء أو أي نوع آخر من المشروعات الهندسية [1]. حيث يعد تحديد استراتيجية (طريقة) تنفيذ المشروع أحد أهم القرارات التي يتخذها كل مالك يشرع في مشروع انشائي أو هندسي. والمقصود باستراتيجية التنفيذ هو تحديد شكل العلاقة بين الأطراف الرئيسية لتنفيذ المشروعات الهندسية [1]. تستخدم صناعة الإنشاءات أنواع مختلفة من استراتيجيات تنفيذ المشروعات الهندسية، حيث يتم اختيار إحدى هذه الطرق في كثير من الدول بشكل تقليدي محاولة لضمان الوقت، التكلفة، وجودة الأداء حسب طموحات المالك عادة. يجب أن يبدأ اختيار أفضل طريقة لأي مشروع بفهم جيد للخيارات المتاحة. كما يجب أن يكون لدى المالك إدراك لتأثير كل اختيار، ولذلك فمن المهم اختيار طريقة التنفيذ التي تلبى بشكل أفضل احتياجات المالك ومشروعه. هذا البحث هو دراسة لتشجيع الخبراء المحليين للمساهمة في فتح الأفق على الاستراتيجيات المختلفة والمساهمة لتحسين مخرجات المشروعات الهندسية. بسبب ندرة الدراسات السابقة في مجال تطبيق استراتيجيات تنفيذ المشروعات الهندسية المختلفة محلياً، تهدف هذه الدراسة إلى المساهمة في فهم والتوسع في هذا المجال بهدف تحسين أداء المشروعات وضمان قبول المخرجات من قبل المالك.

2. شرح الأدبيات المتعلقة باستراتيجيات التنفيذ لمشاريع الانشاءات

من المهم أن ينظر المالك في هذه المجالات الثلاثة (تنفيذ المشروع، طريقة التملك، شكل العقد) -والخيارات الموجودة داخل كل منها - عند اختيار طريقة تنفيذ المشروع [2].

يجب أن يبدأ اختيار أفضل استراتيجية لتنفيذ أي مشروع بفهم جيد للخيارات المتاحة. كما يجب أن يكون لدى المالك إدراك لتأثير كل اختيار، في جميع أنظمة التنفيذ، هناك دائماً ما لا يقل عن ثلاثة أطراف معنية: المالك، المصمم/الاستشاري والمنفذ [1]. من المهم اختيار طريقة التنفيذ التي تلي بشكل أفضل احتياجات المالك وظروف المشروع. ان اعتبارات المشروع لها تأثيرات أساسية على طريقة التنفيذ المختارة. وتشمل هذه الاعتبارات الميزانية الواقعية، والجدول الزمني ذو فترة أداء منطقية علمية، وعملية التصميم التي تضمن الجودة المطلوبة، وغيرها. فيما يلي، سنتطرق لبعض طرق التنفيذ المنتشرة الاستخدام عالمياً وهي خمسة طرق مرتبة حسب قدم استخدامها وانتشارها في مجال الإنشاءات كما يلي:

أ. التصميم-العطاء-الإنشاء: (DBB) Design-Bid-Build

المالك في هذه الاستراتيجية لديه عقود منفصلة مع كل طرف من أطراف التعاقد الأخرى (المصمم/الاستشاري والمنفذ) [1] في هذه الاستراتيجية، المالك يتأكد من مخرجات كل تفاصيل الأعمال لكل الأطراف بشكل منفصل لضمان تحقيق مخرجات المشروع.

ب. التصميم-الإنشاء: (DB) Design-Build

في هذه الاستراتيجية، المالك يتعاقد مع طرف واحد يمثل (المصمم والمنفذ) والذي سيقوم بكل من التصميم والإنشاء. المالك يشرف على الخطوات الرئيسية لضمان تحقيق الهدف/الأهداف الأساسية من إنشاء المشروع [1] توفر هذه الاستراتيجية بعض المزايا ومنها سرعة التنفيذ نتيجة تداخل بين مرحلتي التصميم والتنفيذ.

ج. مدير الإنشاءات والمقاول العام: (CMGC) Construction Manager / General Contractor

المالك في هذه الاستراتيجية لديه عقود منفصلة مع كل طرف من أطراف التعاقد الأخرى وهما (المصمم/الاستشاري والمنفذ). فالمصمم يدعم المالك في متابعة المنفذ لتحقيق أهداف إنشاء المشروع، كما يتم مشاركة المنفذ في القرارات ما قبل المباشرة في الإنشاء أي في مرحلة التصميم [1].

د. شراكة القطاعين العام والخاص: (PPP) Public-private partnership

الشراكة بين القطاعين العام والخاص هي عبارة عن ترتيب تعاوني بين كيان عام أو أكثر (عادة ما يكون المالك) وكيان آخر (عادة القطاع الخاص) لتصميم وبناء وتمويل المشروع. في بعض الأحيان تشغيل وصيانة المشروع لفترة محددة من الوقت نيابة عن المالك. هناك الكثير من الأنواع تقع تحت هذه الاستراتيجية لتنفيذ المشاريع منها: التصميم-الإنشاء-التشغيل، التصميم-الإنشاء-التشغيل-الصيانة، إنشاء-تشغيل-نقل ملكية. أهم الأسباب لاستخدام هذه الاستراتيجية هو توفير التغطية المالية من القطاع الخاص لدعم تنفيذ وتشغيل المشاريع [1].

هـ. استراتيجية المشروع المتكامل: (IPD) Integrated Project Delivery

في هذا النظام ينشئ المالك، والمصمم، والمنفذ عقداً واحداً. حيث يتم التواصل المباشر بين المالك، المصمم، والمنفذ، ويستمد المالك السيطرة أساساً من بناء ثقافة تنظيمية مشتركة تشجع على التحكم في الفريق. هذا النوع من الاستراتيجيات بدء استخدام ودراسة اثاره مؤخرًا على مخرجات المشروعات وخاصة في بناء المستشفيات وغيرها من المباني الخدمية [1].

3. المنهجية

تتضمن كلمة "الأداء" جميع جوانب عملية البناء. يعتبر الأداء كما هو مطبق على الأنشطة في الموقع أو الأنشطة المرتبطة به مصطلحاً واسعاً وشاملاً، ويشمل أربعة عناصر رئيسية، وهي: الإنتاجية والسلامة وحسن التوقيت والجودة. عند تطبيقه في تعريفه العام على الأنشطة في الموقع وخارج الموقع، فإنه ينطوي على جوانب إضافية. وصف أحد المؤلفين الأداء، في تعريف واسع، على أنه سبعة معايير أو عناصر يجب أن تركز الإدارة جهودها عليها وهي: الفعالية، الكفاءة، الجودة، الإنتاجية، الربحية، والابتكار [3].

الطبيعة الديناميكية لصناعة البناء اليوم تجبر المقاولين والمالكين على البحث عن استراتيجيات لتحسين الأداء. كذلك المنافسة المحلية والأجنبية الشديدة على عقود البناء، والتحالفات الجديدة بين المالكين والمقاولين. ففي ظل هذه الظروف، من الصعب التنبؤ بآثار استراتيجيات تنفيذ البناء الفردية أو المجمعة على أداء المشروع بشكل بسيط. لذلك، فإن إيجاد نموذج يساعد في فهم هذه الآثار هو مفتاح النجاح في مواجهة التحديات الجديدة التي تواجه مشروعات البناء [3].

حاول العديد من المؤلفين بناء نماذج مفاهيمية لعملية البناء من خلال شرح جوانب مختلفة من أداء المشروع. ارتبطت بعض هذه الجهود بإنتاجية البناء. تم أيضاً فحص نماذج الشبكات واستخدامها في تخطيط وجدولة ومراقبة المشروع. تم اقتراح نماذج المحاكاة واستخدامها كأدوات للتعامل مع حالات عدم اليقين في مشاريع البناء وخاصة الوقت والتكلفة.

تم استخدام تقنيات المحاكاة لدراسة عملية البناء على مستويات مختلفة. على مستوى المهمة، تم تطوير تقنيات النمذجة المحددة لتحليل عمليات البناء، وخاصة العمليات المعقدة التي تنطوي على دورات وعمليات متكررة. تم تنفيذ العديد من التطبيقات باستخدام مجموعة كبيرة ومتنوعة من منصات البرمجيات. معظم هذه التطبيقات موجهة نحو تحسين الإنتاجية في عمليات البناء حيث تكون مقاييس الأداء عادةً هي التكلفة والمدة والإنتاجية [4]، على مستوى المشروع، تم استخدام المحاكاة لتقدير وتقييم تأثير حالات عدم اليقين الموجودة في مشاريع البناء على نتائج مثل التكلفة والمدة. بشكل عام، تنقسم حالات عدم اليقين الموجودة في مشاريع البناء إلى عنصرين رئيسيين، وهما التباين في أداء المهمة، والتدخل من خارج المهمة الذي تحبط

تقدمها. يمكن للمحاكاة، بالنظر إلى صحة افتراضات المدخلات، أن تقدم تقديرات غير متحيزة لنتائج أداء المشروع مثل المدة والتكلفة [4]

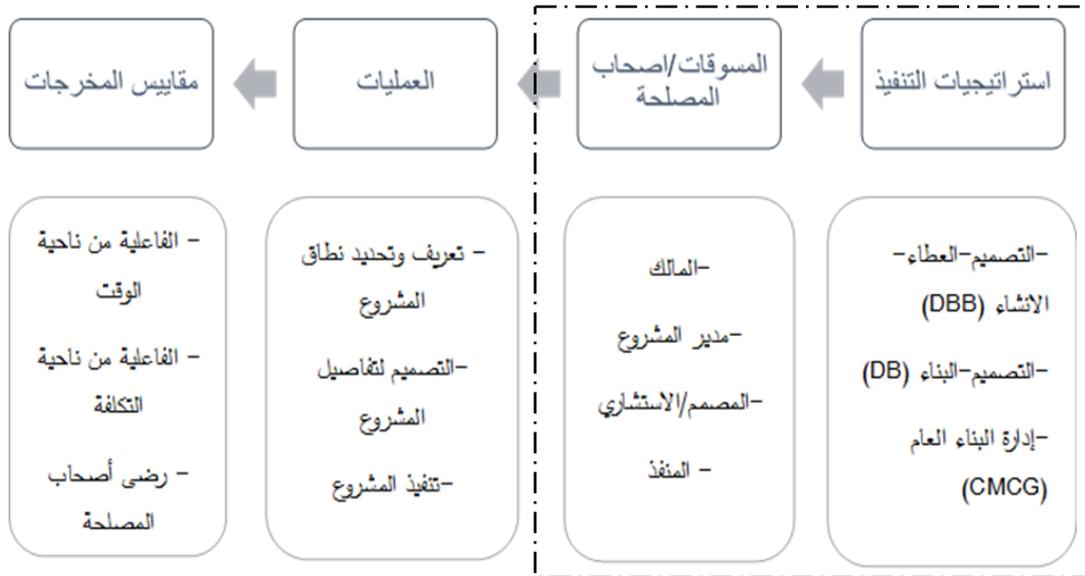
نماذج المحاكاة توفر قدرة غير محدودة تقريباً لنمذجة عمليات البناء مما يمنح المالك نظرة شاملة حول العوامل الأكثر أهمية وكيفية تفاعلها. يمكن تجربة استراتيجيات مختلفة من خلال هذه النماذج للتعرف على النتائج المتوقعة من تطبيقها واختيار أفضل البدائل. ويمكن تحليل الأنظمة الديناميكية في فترات زمنية قصيرة. ربما يكون العيب الرئيس لتقنيات المحاكاة هو أن تعقيد نماذج المحاكاة الحالية المطبقة على البناء يجعل من الصعب للغاية على المستخدمين فهم النماذج. وللأسباب نفسها، قد لا يدرك المستخدمون العيوب أو الآثار المترتبة على نتائج المحاكاة.

بالاعتماد على منهجية الأداء العام والتي اخترعت وطبقت من قبل (Project Team Risk/Reward) التابعة لمعهد صناعة البناء والتشييد (The Construction Industry Institute – CII)، تم تصميم النموذج الخاص بهذه الدراسة من قبل الباحث. الفقرة التالية تتضمن شرح مبسط للنموذج ومكوناته الرئيسية.

4. نموذج الأداء العام GPM

الفلسفة خلف هذا النموذج هو أنه نموذج مبسط للمتغيرات وتفاعلاتها التي تؤثر على أداء المشروع مع ضمان أن هذا النموذج يحتوي على المكونات الأساسية التي تجعله يمثل الواقع بما يتماشى مع متطلبات الدراسة. إن المكونات الأساسية هي: الاستراتيجيات، المسوقات، العمليات، مقاييس المخرجات. من خلال النموذج يمكن تمثيل التفاعل بين مكوناته الأساسية كالآتي: هناك تأثير مباشر للاستراتيجيات على المسوقات وهي التي تقود نقل تأثير الاستراتيجيات لباقي مكونات النموذج – الحالة تحت الدراسة. كما أن هناك تفاعل وتأثير بين المسوقات بالإضافة لتأثيرها المباشر على العمليات. أيضاً، هناك تفاعل بين عناصر العمليات كما للعمليات تأثير مباشر على مقاييس المخرجات ذات الاهتمام. وأخيراً، إذا كانت الدراسة مهمة بالعوامل الخارجية، فهناك تأثير مباشر منها على كل من المسوقات والعمليات [4]. لتفاصيل أكثر حول النموذج وتفاعلاته يمكن مراجعة [1,3,5] ولكن لأن دراستنا محلية ولن نقارن ظروف خارجية مختلفة، قمنا بعدم إدخال دراسة العوامل الخارجية في نطاق الدراسة.

نتيجة لاهتمام هذه الدراسة بمعرفة أثر تطبيق إحدى استراتيجيات التنفيذ على أداء الأطراف ذات العلاقة (أصحاب المصلحة) وبالتالي تأثيرهم على مقاييس المشروع المختارة، تم استخدام فلسفة نموذج الأداء العام لتمثيل هذه العلاقات المركبة وهي كما يلي: تم تمثيل تفاعلات استراتيجيات تنفيذ المشروع في مكون الاستراتيجيات. أما المسوقات فهي تضم أصحاب المصلحة وهم: المالك، مدير المشروع، المصمم، المنفذ. الذين سيترجمون علاقاتهم من خلال تفاعلاتهم لإتمام عمليات المشروع. هذه العمليات تم تليخيصها في ثلاث نقاط. وهي: تعريف وتحديد نطاق المشروع، تصميم تفاصيل المشروع، تنفيذ المشروع. بالإضافة تم تحديد ثلاث مقاييس مهمة لتحديد وفهم تأثير قرار تحديد استراتيجية التنفيذ. في هذه الورقة، سيتم التركيز على فهم العلاقة والتأثير بين الاستراتيجيات والمسوقات ومقاييس حساب هذه التأثيرات. الشكل رقم (1) يوضح كامل النموذج مع التركيز على الجزء الخاص بهذه الورقة.



شكل رقم (1): نموذج الأداء العام GPM

إن تسلسل التأثير في نموذج GPM يتمثل في: استراتيجية التنفيذ المختارة سوف تؤثر على دور وأداء المسوقات (أصحاب المصلحة)، كما تؤثر المسوقات (أصحاب المصلحة) على عمليات المشروع. وأخيراً، ستؤثر العمليات على أداء المشروع من حيث قياسات النتائج المحددة. علاوة على ذلك، هناك تأثيرات داخلية بين المسوقات والعمليات نفسها كما شرحناها مسبقاً.

استناداً إلى الدراسات السابقة، تتمثل إحدى الخطوات الرئيسية لمشاريع البناء الناجحة في اختيار المالك استراتيجية التنفيذ التي يرغبون في استخدامها. في هذه الدراسة، سيتم دراسة ثلاثة أنظمة تنفيذ وهي: التصميم – العطاء – الإنشاء (DBB)، مدير الإنشاءات والمقاول العام (CMGC)، التصميم – الإنشاء (DB). تم التركيز على هذه الاستراتيجيات لانتشار استخدامهم عالمياً ولعدم وجود ممارسين ذوي خبرة في الاستراتيجيات الأخرى وكذلك لهم خبرة في ظروف الإنشاءات في ليبيا. فيما يلي وصف لمميزات النموذج لهذه الدراسة:

أ. استراتيجيات (نظم) تنفيذ المشاريع

أحد أهم القرارات في بداية المشروع هو اختيار نظام تنفيذ المشروع. لذلك، يتعين على المالك تحديد نوع الهيكل التنظيمي لتطوير المشروع. يمثل النموذج GPM كل خيار من خيارات هذه القرارات. تشير الاستراتيجيات إلى مستوى مشاركة المالك، مدير المشروع، المصمم، والمنفذ خلال مراحل التعريف والتصميم والبناء. يقوم هذا النموذج GPM بتقييم ثلاثة هياكل أساسية لأنظمة التنفيذ:

- **التصميم – العطاء – الإنشاء (DBB)** المالك في هذه الاستراتيجية لديه عقود منفصلة مع كل من أطراف التعاقد الأخرى (المصمم والمنفذ). المالك يتأكد من مخرجات كل تفاصيل الأعمال لكل الأطراف لضمان تحقيق مخرجات المشروع.
- **التصميم – الإنشاء (DB)** المالك يتعاقد مع طرف واحد يمثل (المصمم والمنفذ) والذي سيقوم بكل من التصميم والإنشاء. المالك يشرف على الخطوات الرئيسية لضمان تحقيق الهدف/الأهداف الأساسية من إنشاء المشروع.
- **مدير الإنشاءات والمقاول العام (CMGC)** المالك في هذه الاستراتيجية لديه عقود منفصلة مع كل من أطراف التعاقد الأخرى (المصمم والمنفذ). المصمم يدعم المالك في متابعة المنفذ لتحقيق أهداف إنشاء المشروع. كما يتم مشاركة المنفذ في القرارات ما قبل المباشرة في الإنشاء.

ب. الأسواق (أصحاب المصلحة)

في نموذجنا، استخدمنا الأسواق باعتبارهم الأطراف / أصحاب المصلحة الرئيسيين الذين سيكونون جزءاً من أنشطة مشروع البناء. يحدد أصحاب المصلحة هؤلاء استناداً إلى مؤسسة إدارة المشاريع (PMI) "فرد أو مجموعة أو منظمة، قد تؤثر على أو تتأثر أو تتأثر بنفسها لتتأثر بقرار أو نشاط أو نتيجة لمشروع" اعتماداً على مراجعة الأدبيات، تستخدم الدراسة الحالية أصحاب المصلحة الأربعة التالية [4]

- **المالك:** هو الكيان الذي يقوم ببدء المشروع، تمويله والتعاقد عليه والاستفادة من مخرجاته.
- **مدير المشروع:** هو المسؤول عن تنسيق وإدارة جميع أنشطة المشروع، حيث يقوم بمراجعة جميع المستندات ورسومات التصميم المقدمة من قبل الاستشاريين والموافقة على عمليات الدفع، إلخ.
- **المصمم/الاستشاري:** لديه مسؤوليات تشمل التصميم والجدولة والمناقشة والإشراف على جميع أعمال البناء من قبل المقاولين.
- **المنفذ:** عليه مسؤوليات تتضمن جدولاً لنشاطات أعمال البناء وتنفيذ جميع الأعمال الميدانية.

5. تجميع البيانات

حاولنا التواصل مع العديد من الجهات ذات العلاقة بالمشاريع الهندسية، ولكن للظروف التي تمر بها البلاد، وجدنا صعوبة في تحفيز عدد أكبر من ذوي الاختصاص في هذه الدراسة رغم اقتناعهم بأهميتها ورغبتهم السابقة في المشاركة. تم الاكتفاء بالحد الأدنى لهذا النوع من الدراسات وورش العمل لتجميع البيانات المتخصصة وهو عدد ثلاث ممارسين للمهنة ذو خبرة [6].

فالمهندسين الثلاثة ذوي خبرة في موضوع البحث وهو تصميم وإدارة العقود ولهم الخصائص الموضحة في الجدول رقم (1). الخطوة الأولى هي السؤال عن شمولية النموذج واحتوائه على كل العناصر المهمة ذات العلاقة والضرورية للإجابة على السؤال البحثي موضوع الدراسة. حيث قمنا في الورشة الأولى وبعد شرح الفكرة وطرح السؤال البحثي بعرض النموذج وسؤال المشاركين عن مدى شمولية النموذج على العناصر الأساسية والضرورية لتمثيل مشكلة الدراسة. تم الموافقة على النموذج وانتقلنا لمرحلة التقييم.

جدول رقم (1) خصائص المشاركين في ورشة العمل

المهنة	سنوات الخبرة	الوظيفة قبل ورشة العمل	المشاريع ذات الاهتمام والتي شارك في إدارة عقودها
1	30	المدير التنفيذي لشركة خاصة	مشاريع الطرق والجسور (الدائري الثالث بطرابلس)
2	20	مدير مشروع المطار -وزارة المواصلات	مشاريع الطرق والمطارات
3	20	مدير شركة مقاولات خاصة	مشاريع الطرق والمباني الإدارية

تم تنفيذ ورشة العمل الأولى لتجميع تأثير الاستراتيجيات على المسوقات. في هذه الورشة تم عرض نموذج الأداء العام والذي يشمل على 51 تفاعل/تقييم بشكل كامل وتم التركيز في هذه الورشة على تقييم التأثير المباشر بين الاستراتيجيات والمسوقات والتي تحتوي على 12 تقييم فقط.

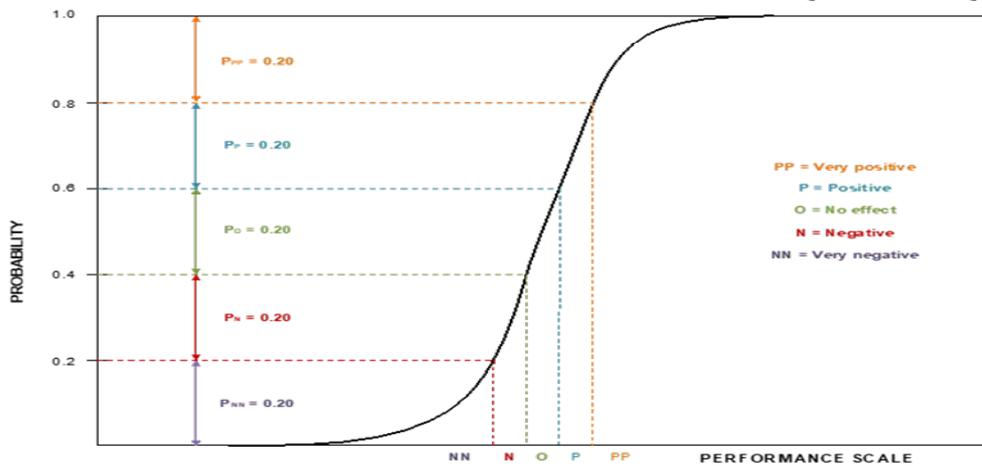
في هذه الورشة، تم مناقشة وتقييم التفاعلات في حدود ثلاث ساعات. أغلب الوقت تم استغلاله في مناقشة النموذج والتأكد من فهم طبيعة التفاعلات والتعريفات المقترحة لكل المكونات. أيضاً، تمت عملية مناقشة مقياس التقييم للتأكد من فهم أبعاده وأسلوب استخدامه لتقديم التقييم المطلوب بشكل متجانس بين المشاركين. حيث تم تقديم اساس التقييم كما يلي:

المؤسسة تُمكن تحقيق المسوقات (أصحاب المصلحة) أقصى إمكاناتها بما لا يقل عن 40% ولا يزيد عن 60% من حالات التعاقد (بين 4 و 6 عقود من أصل 10)

وفقاً للأساس السابق، تم طلب تقييم أثر الاستراتيجيات المختلفة للتنفيذ على مستوى أداء المسوقات حسب ما سيرد لاحقاً مع استخدام مقياس التقييم التالي:

- تأثير إيجابي عالي (PP): التأثير الإيجابي العالي يحسن من تحقيق المسوقات لأقصى إمكاناتها بما لا يقل عن ثمانية عقود من أصل 10 عقود.
- تأثير إيجابي (P): التأثير الإيجابي يحسن من تحقيق المسوقات لأقصى إمكاناتها بما لا يقل عن ستة عقود من أصل 10 عقود.
- بدون تأثير (O): لا تأثير على أداء المسوقات. المسوقات تحقق أقصى إمكاناتها بما لا يقل عن أربعة عقود ولا يزيد عن ستة عقود من أصل 10 عقود.
- تأثير سلبي (N): التأثير السلبي يقلل من تحقيق المسوقات لأقصى إمكاناتها بما لا يقل عن عقدين ولا يزيد عن أربعة عقود من أصل 10 عقود.
- تأثير سلبي عالي (NN): التأثير السلبي العالي يقلل من تحقيق المسوقات لأقصى إمكاناتها بما لا يزيد عن عقدين من أصل 10 عقود.

الشكل التالي يوضح المقياس الموضح أعلاه.



شكل رقم (2) التمثيل البياني لمقياس التقييم للاستراتيجيات

الجدول النهائي للتقييمات لهذه الورشة كان كما يلي:

جدول رقم (2) : تقييمات ورشة العمل الأولى

المسوقات				البيانات	البيانات الاستراتيجية
المصمم/ الاستشاري	المنفذ	مدير المشروع	المالك		
O	O	PP	N	التصميم – العطاء-الانشاء (DBB)	
P	P	O	P	مدير الانشاءات والمقاول العام (CMGC)	
PP	PP	P	N	التصميم – الانشاء (DB)	

6. الاستنتاجات والتوصيات

يتضح من النتائج المتحصل عليها من ورشة العمل أن اختلاف الاستراتيجيات المتبعة في مشروعات الانشاء تؤثر تأثير كبير على أداء المسوقات خلال مراحل المشروع وهم الأطراف الأساسية في نجاح أو فشل أي مشروع انشائي أو هندسي. لهذا نوصي الشركات المعنية بصناعة الانشاءات بزيادة الاهتمام بأساليب الإدارة واتباع الاستراتيجيات الحديثة لما لها من دور كبير في تحسين أداء الشركة.

من خلال النتائج، وجدنا أن أفضل استراتيجية تنفيذ لتحسين أداء مدير المشروع هي استراتيجية: التصميم-العطاء-الانشاء. فمن خلال هذه الاستراتيجية، كل المستندات المطلوبة والتي تدعم أداء مدير المشروع مثل التصاميم والمواصفات تكون متكاملة وجاهزة قبل البدء في الانشاء. بينما نجد أن استراتيجية: التصميم-الانشاء تعتبر الأفضل لكل من المنفذ والمصمم/الاستشاري حيث أن هذه الاستراتيجية تسرع من عملية البدء وتعطي مرونة لهما من حيث التعديل في التصميم واختيار أساليب التنفيذ حسب الحاجة لاحقاً بعد بدء العمل. أما استراتيجية إدارة البناء العام فهي تخلق ظروف إيجابية عند أغلب الأطراف وهم المالك والمنفذ، والمصمم/الاستشاري لأنها تتيح فرصة للتعاون بين هذه الأطراف في وقت مبكر من عمر المشروع وهو ما يدفعنا للتأكيد على ضرورة الاتجاه في هذا المسار وتطبيق استراتيجيات أخرى ذات طبيعة منفتحة وتكاملية بين المسوقات (الأطراف المهمة في المشروع) في وقت مبكر من عمر المشاريع الهندسية والانشائية. من هذه الاستراتيجيات المقترحة: استراتيجية المشروع المتكامل (IPD).

أما من ناحية التأثيرات السلبية، نجد أن (DBB) و (DB) لهما تأثيرات سلبية على أداء المالك. من خلال النقاش، تم التأكيد على أن المالك يتأثر سلباً عند استخدام استراتيجية (DBB) نتيجة طول فترة الإعداد للتصاميم والمستندات الخاصة بالمشروع والعطاء. وكذلك، يتأثر المالك سلباً من استراتيجية (DB) نتيجة فقدانه التأثير المباشر على تفاصيل المشروع حيث يبدأ المشروع قبل اكتمال التصاميم عادة. وهذا مؤشر إضافي للحاجة لاختيار الاستراتيجيات الحديثة والتكاملية خاصة للتغلب على هذه الآثار السلبية على أداء الأطراف المهمة في نجاح أو فشل المشاريع الهندسية والانشائية.

وختاماً ونظراً للظروف التي تمر بها البلاد، واجه البحث العديد من الصعوبات منها قلة الخبرات المحلية في مجال استراتيجيات التنفيذ واقتصار الدراسة على المنطقة الغربية فقط وبالتحديد طرابلس ومصراته. كما نوصي لأي دراسة مستقبلية ان يتم التوسع في تنفيذ ورش العمل لتضم العديد من المدن الليبية وكذلك الاستعانة بخبراء في ظروف مشابهة ليبييا للحصول على نتائج أعم وأدق.

المراجع

- [1]. A. Mesa, K. R. Molenaar, L. F. Alarcón, "Exploring performance of the integrated project delivery process on complex building projects", *Int. J. Proj. Manag.*, 2016, 16, 1089-1101.
- [2]. Venegas, P. & Alarcón, L.F. (1997) Selecting long term strategies for construction firms. *Journal of Construction Engineering and Management*, **123**, 388–389
- [3]. L. F. Alarcon-Cardenas D. B. Ashley, "Project Performance Modeling: A Methodology For Evaluating Project Execution Strategies", 415.
- [4]. L. F. Alarcón D. B. Ashley, "Project management decision making using cross-impact analysis", *Int. J. Proj. Manag.* 1998, 16, 145-152.
- [5]. W. H. Beitelmal, K. R. Molenaar, A. Javernick-Will, O. Smadi, "Strategies to Enhance Implementation of Infrastructure Asset Management in Developing Countries", *Transp. Res. Rec. J. Transp. Res. Board*, 2017, 48–39, 1, 2646, doi: 10.3141/2646-05.
- [6]. R.A.Krueger, M.A.Casey,(2015) "Focus Groups A Practical Guide For Applied Research",



Track No. 7: Mechanical & Industrial Engineering

Effects of a Curved Surface of the Blade on the Performance Characteristics of Axial Flow Rotors

Ali R. kwedikha

ali.kwedikha@zu.edu.ly

Department of Mechanical Eng., College of Engineering, University of Zawia, Libya

ABSTRACT

The benefit of using a Computational Fluid Dynamics (CFD) technique is to predict the effects associated with non-radial stacking (NRS). Three-dimensional (3D) flow field for NRS blades of an axial turbomachine is studied. The non stacking line of the fan is provided by a curved surface by means of enlarging blade airfoil only at the base of the fan blade. The effectiveness of the NRS on the axial fan rotor of low-aspect-ratio was studied in two separate case studies a standard blade (SB) and a curved blade (CB).

In developing a complete structured hexahedral mesh for the entire computational domain, comparative studies of CB and SB were conducted at the design and off-design flow rates. The structured mesh technique minimizes cell counts, cell skewness, and enables cost-effective CFD investigation.

The results are presented in the form of local radial velocity, local ideal and total pressure rise of the outlet as well as the static pressure on the blade suction side. Specifically, it is pointed out that in the design point the CB rotor exhibits the highest efficiency for the most part of the entire span, whereas SB exhibits the lowest efficiency along the entire span. While in the off-design point, the CB rotor exhibits the highest efficiency at the blade hub only.

Keywords:

Three-dimensional turbomachinery flow, axial fan, non-radial stacking, structure meshing.

1. Introduction

Improving the axial fan blade shape, by means of the stacking line technique, has become the main variable design. The blade shape designs, provided by stacking line performance, are often offered to increase efficiency and to reduce losses in axial turbomachines. Distortion of stacking line using sweep, dihedral, skew and/or any other forms have become nowadays a matter of requiring need in the design of turbomachinery blades. The form of the stacking line, which is only a radial line, is modified to take on new shapes by changing the blade profiles in definite directions. Some of these modifications are called a sweep, dihedral and skew (the combination of sweep and dihedral) depending on a certain movement direction of the blade profile.

The blade can be modified by sweep and/or dihedral if the blade sections of a datum blade of radial stacking line are displaced parallel to and/or normal to the chord, respectively [1].

There are four main types of sweep in the blade which are backward, forward, positive, and negative sweep. A blade is swept forward or backward at a given radius if the blade sections of a radially stacked datum blade are shifted parallel to their chord in such a way that a blade section under consideration is upstream or downstream of the neighbouring blade section at lower radius,

respectively, sweep is said to be positive or negative near an endwall when a blade section under consideration is upstream or downstream of the adjacent inboard section, respectively [2].

In this work, two axial fan blades are designed and analyzed. The results are also discussed and interpreted. The first blade is called the Standard blade (SB) which is to be easy to manufacture, and the second one has the same characteristics of the first one, but it was modified with an additional curved surface at the base of the blade, this blade is called Curved blade (CB).

The curved blade is prepared by means of increasing the blade chord length and enlarging blade airfoil at the base the fan blade, this curved blade is defined to be as a part of positive backward sweep profile at the hub as shown in **Figure 1**.

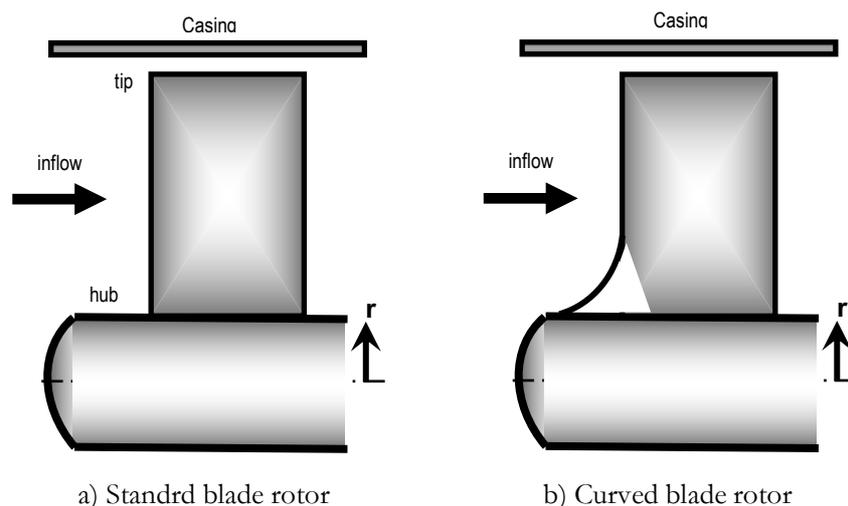


Figure 1: Schematic drawing for CB rotor.

The use of properly curved surface attached to the base of the fan blade is considered to prevent a total pressure loss and to provide a smooth transition of the flow stream through the blades of the fan. As an example, compressors have generally benefited from the use of backward sweep at the hub and/or forward sweep at the blade tip [1].

In general, Govardhan et al. [3], mentioned that, backward swept blades show minor efficiency improvement relative to the unswept rotor but at a considerable loss in stall margin. The peak efficiency of the backward swept rotor is often found to be much closer to the stall line to be of practical use.

The aim of this study is to understand the nature of three-dimensional flow when the blade has a curved surface on the hub in a low speed axial fan. The performance of the CB blade type is compared to the SB blade type and analyzed as well. Geometry construction

Curved blade geometry

To improve the performance of the axial fan, we made a modification to the design of the blade itself similarly to the designs of the sweep. The procedure was in such a way that modifying the base of the fan blade. Simply by magnifying and enlarging the blade airfoil section and increasing the chord line at the blade hub, then reducing this amplification towards the blade tip, without displacing the blade sections from the trailing edge and without changing the type of blade sections. The blade section used is profile of C4 (10%) along the entire span [4] as shown in **Figure 2**. However, there is a lack of information and knowledge on the curved surface applied in the blade design and the reference [4] serves as a preliminary reference for the data.

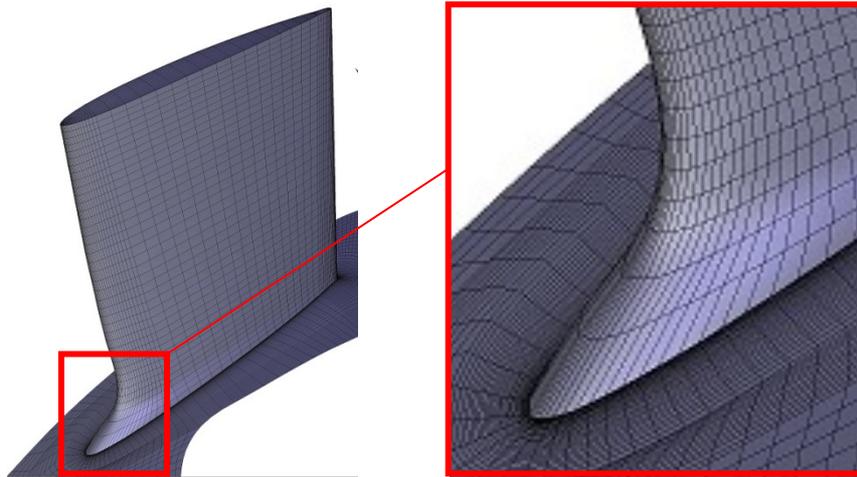


Figure 2: 3D view of the curved surface of the CB blade geometry.

The Reynolds number (Re) is approx. 1.07×10^6 based on the characteristic length as the chord line at the blade tip (0.372 m), the circumferential speed of the blade tip (43.56 m/s), and the kinematical viscosity of air at 20 °C ($1.516 \times 10^{-5} \text{ m}^2/\text{s}$).

Rotating cascades

The rotor blades are arranged so that the blades are assembled in an annular cascade, which is surrounded by a cylindrical casing, as shown in **Figure 3**. Another choice to use instead of the annular cascade is a fixed linear cascade of blade rotors in a linear cascade, which is an easy technique to design and it is not used in this study, because the radial effects such as the radial pressure gradients, the radial distribution and the secondary flows will not be developed as in the annular cascade. An important interest in the rotation cascades is their ability to simulate the effect of radial forces on the flow through the blade passages.

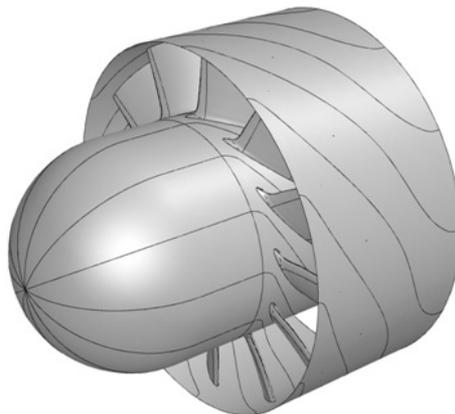


Figure 3: The CB rotor in cascade form.

Because the mesh periodicity for one domain is one of the most important constraints on the organized structured meshes, it has been taken into consideration more seriously. In cascade investigations this means that the flow periodicity adjustment becomes more difficult.

2. CFD technique

A commercially available ANSYS-17.2 finite volume code is used. The examined geometry in this work has 12 blades surrounded by a casing of diameter 2000 mm. Initially, a three-dimensional volumetric domain was constructed around one blade. The domain is divided into three parts. From the inlet to the exit of the domain, the first part begins with the inlet until the end of the hemisphere; the second part begins at the end of the first part until the front of the blade hub, and the last one contains the rest of the domain including the blade, hub and the outlet. The tip clearance (ν) is a 5% span with a Hub-to-tip ratio (τ) of 0.6. A typical computational domain for the CB rotor is shown in **Figure 4**.

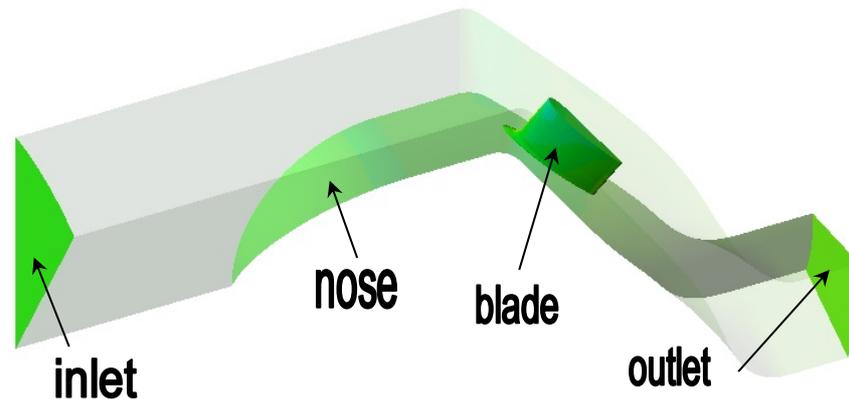


Figure 4: Computational domain for CB rotor.

A grid is said to be structured if all the vertices of the internal cell belong to the same number of cells of the control volume in which the domain is divided. The grid can be mapped to a rectangle (in 2D) or a block (in 3D), while unstructured grids divide the domain into simple elements, generally triangular (in 2D) or tetrahedral (in 3D) without implicit connectivity [5].

Because good mesh quality is crucial for good CFD analysis, it is important to construct a high-quality grid, especially where detailed flow are needed and at the interesting zones of the domain, such as in the vicinity of blade surfaces. The multi-block structure method allows the use of specific local structural grids in different places. Ali and Tucker [6] found that the multi-block structured meshes offer better computational efficiency than the unstructured meshes which, on the other hand, are more flexible for complex geometries.

At beginning the domain was divided into multiple separate volumes. By using an adequate hexahedral meshing each volume was meshed separately. Only H-Grid and C-Grid meshes were used for organized 2D meshes.

The majority of the number of cells in the domain is condensed around the vicinity of the blade. The multi-block structure method is used in this work; the domain was divided to 39 blocks and meshed to 304306 hexahedral cells. The equiangle skewness was 69.41% far from the blade, with restriction of periodic domain. A domain of structured grid is shown in **Figure 5**.

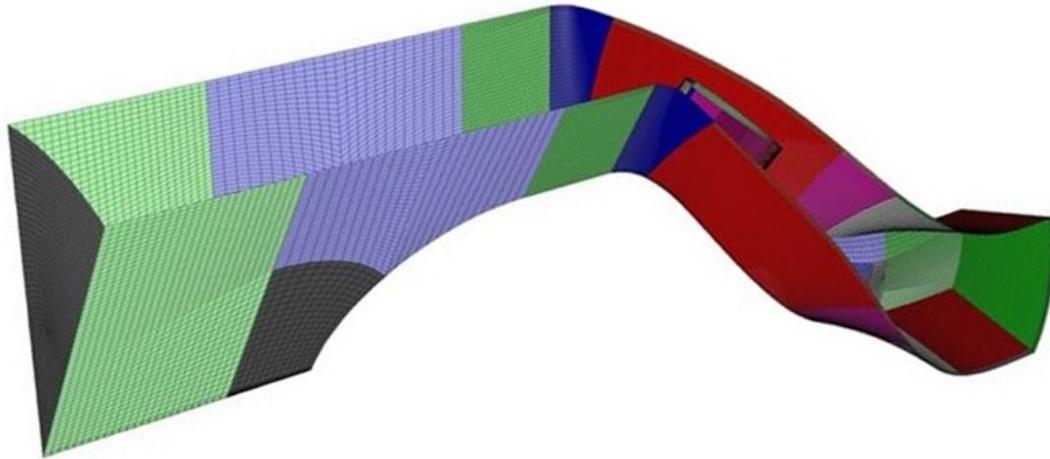


Figure 5: The splitted computational domain of structured grid.

(The casing is hidden for clarity)

3. Boundary conditions

For the design flow rate, the velocity of 9.2 [m/s] is used at the domain inlet, while 7.5 [m/s] is used for the off-design flow rate, the outflow is used as an outlet and the turbulence model used is k-ε model with enhanced wall treatment. The y^+ values are mostly between 30 and 100 in the interesting regions such as at the vicinity of the blade. Some values of y^+ are larger than 300 on grids where the areas are not interesting and located far from the blade. Typical computational run required about 6000 iterations.

4. Results and Discussion

The results provided to be on outlet planes mounted at the axial direction very close to the blade. Also, other investigations as well, located at 20% and 90% measured from the hub in the spanwise direction. Some parameters and symbols must be mentioned here and clarified, such as (SS) and (PS). They represent the blade suction side and pressure side, respectively.

The fraction of span (σ) is the radial distance from the hub divided by blade span height, the reference velocity (u_{ref}) or (u_t) in [m/s] is represented as the blade tip speed = ($d_t \pi n$) where, d_t is the blade tip diameter and n is the rotor speed (in revolutions per second), accordingly the definition of the radial local flow coefficient (ϕ_r) is [4]

$$\phi_r = v_r / u_t \dots\dots\dots(1)$$

Where, (v_r) is the radial velocity. Figure 6 shows, in the range of the rotor hub up to 65% of span, the averaged radial velocity of CB is almost stable while, the averaged radial velocity of SB is reduced in this range, and then the radial velocity for rotors have increased to 90% of span, after that the averaged radial velocities are reduced at the blade tip. The velocities are quite similar for both rotors from 80 % of span to the casing. According to continuity, this yield to rearranging the values of radial velocity at the blade tip between SS and PS, so that pitch-wise averaged radial velocities has been the same for both blades.

Lower than the 80 % of span, some different reorganization in radial fluid for SB and CB blades are appearing with more radially inward flow for CB blade, as observed by Beiler [7].

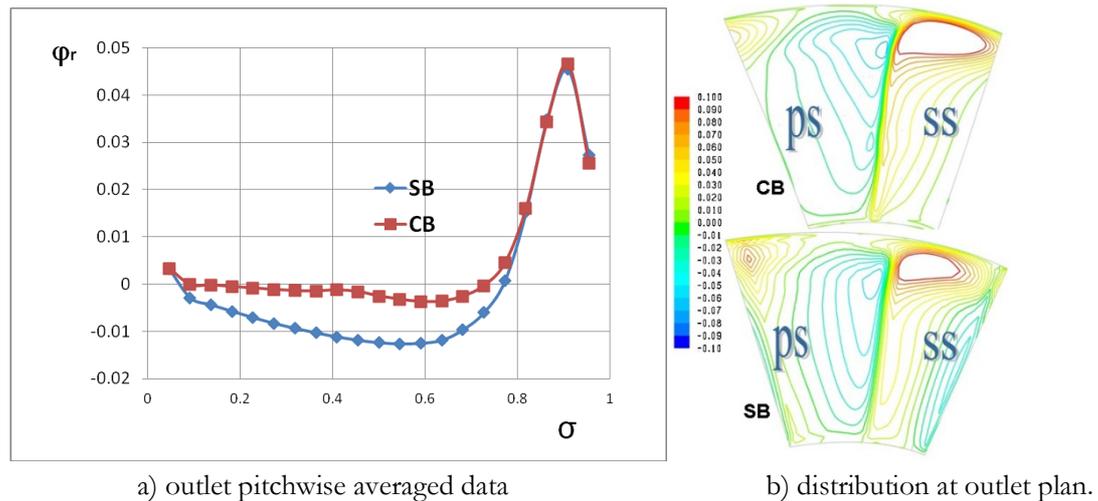


Figure 6: Outlet radial velocity.

The definition of the local total pressure rise coefficient (ψ) [4] is

$$\psi = \Delta p_t / (\rho u_{ref}^2 / 2) \dots\dots\dots(2)$$

Where, ρ is the air density and Δp_t is the pitchwise mass-averaged local total pressure rise. A mass averaged quantity is obtained by integrating the scalar time mass flow divided by total mass flow over the region. The outlet plan is divided to 21 slide regions.

The CB rotor recognized increased total pressure rise compared to SB at the dominant blade spanwise. The contrary noticed by Beiler and Carolus [8]. CB rotor tends to increase the total pressure rise, especially in the vicinity of the trailing edge. Figure 7 indicates, SB and CB rotors perform lower total pressure rise at the hub and tip.

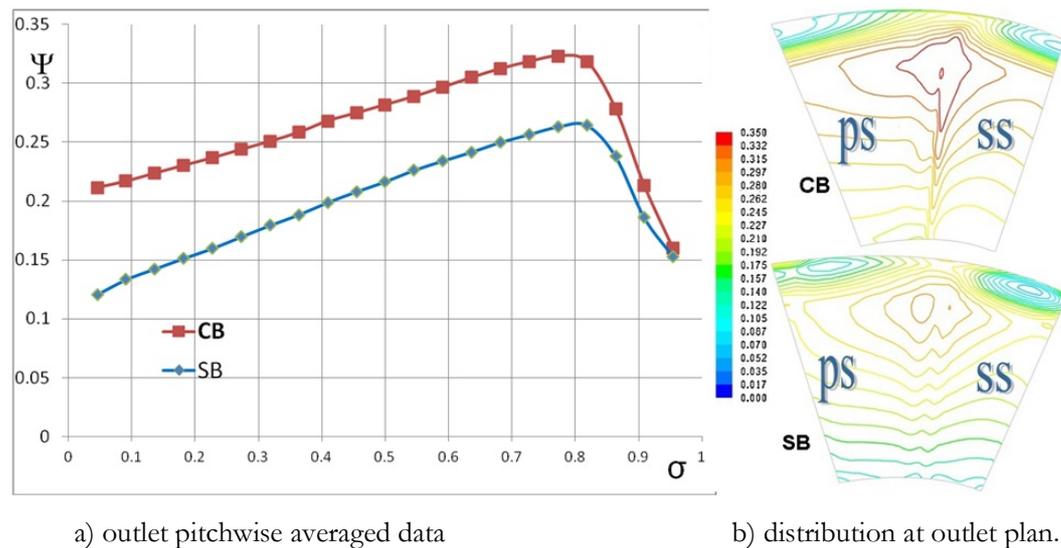


Figure 7: Local total pressure rise.

In general for both rotors, there is less total pressure rise at SS close to blade tip compared to lower radii. Figure 8 shows that the blades having curved surfaces cause increasing the total pressure on PS's at lower and higher radii.

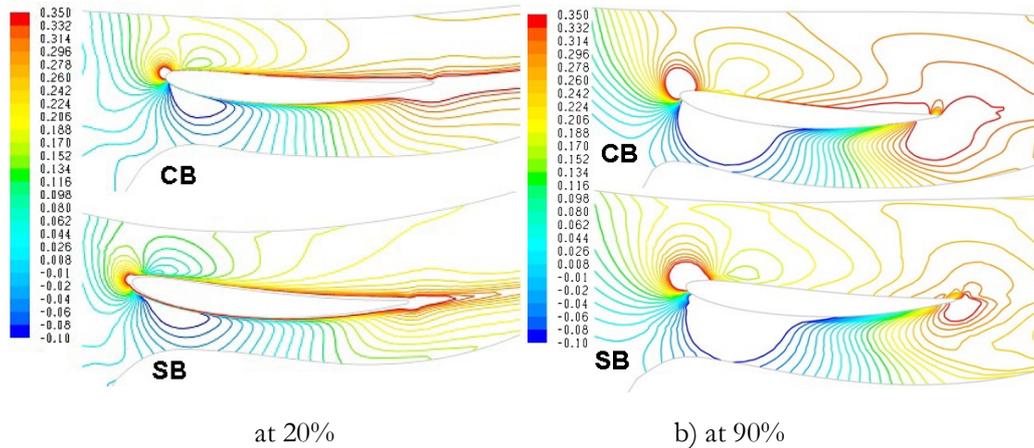


Figure 8: Distribution of local total pressure rise at 20% and 90% of span.

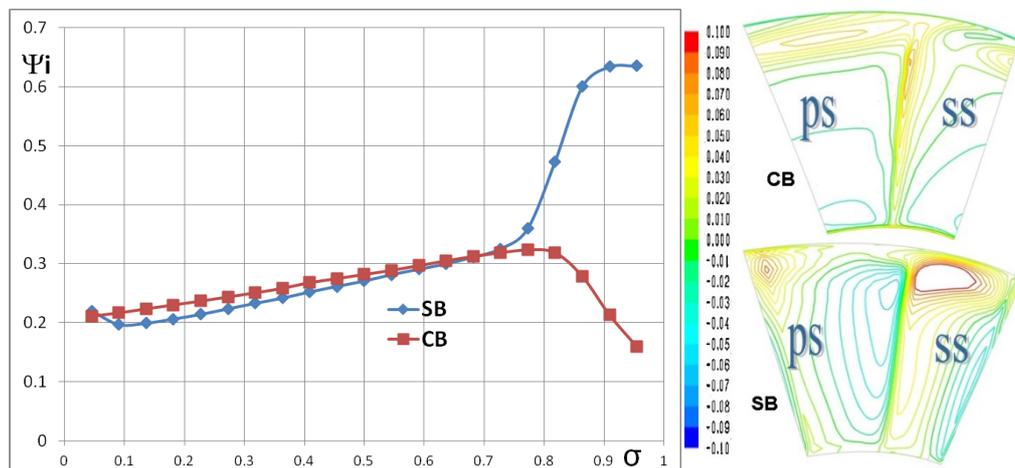
Dealing to the equation of the Euler, the total pressure rise of the inviscid flow is called the ideal total pressure rise. The definition of ideal total pressure rise coefficient (ψ_i) [4] is

$$\psi_i = \Delta p_{ti} / (\rho u_{ref}^2 / 2) \dots\dots\dots(3)$$

$$\text{and, } \Delta p_{ti} = \rho r \omega v_{t2} \dots\dots\dots(4)$$

Where, v_{t2} is the tangential pitchwise mass-averaged tangential velocity, r is radial coordinate and ω is rotor angular speed.

Figure 9 shows, CB rotor performs increased a little ideal total pressure rise compared to SB up to 75 % spanwise. The SB rotor increases the ideal total pressure rise after 75 % spanwise especially close the blade tip SS. However, due to the non-radial blade stacking, the Euler work at the tip is rapidly reduced; this is also being noticed by Clemen et al. [9]. Such an effect may cause a significant increasing local efficiency in this range of blade spanwise.



a) outlet pitchwise averaged data

b) distribution at outlet plan.

Figure 9: Ideal local total pressure rise

In general, the fluid in the boundary layer of SS has a propensity to move radially outward. Due to CB blade, the isobars in the tip region are disorganized in a coarse tendency than for SB blade. In CB blade, the local radial outward flow is located in a higher location region on the blade trailing

edge. Similar results have been illustrated in Yamaguchi et al. [10]. **Figure 10** illustrates the mechanism of CB blade to redirect the outer radial flow on the blade suction side.

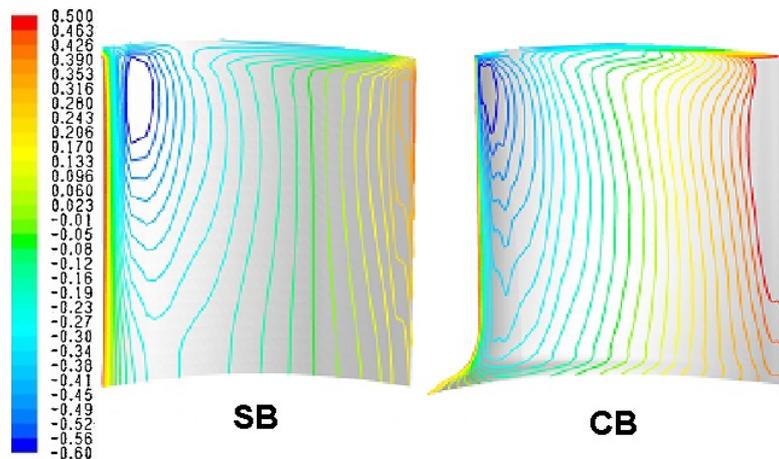
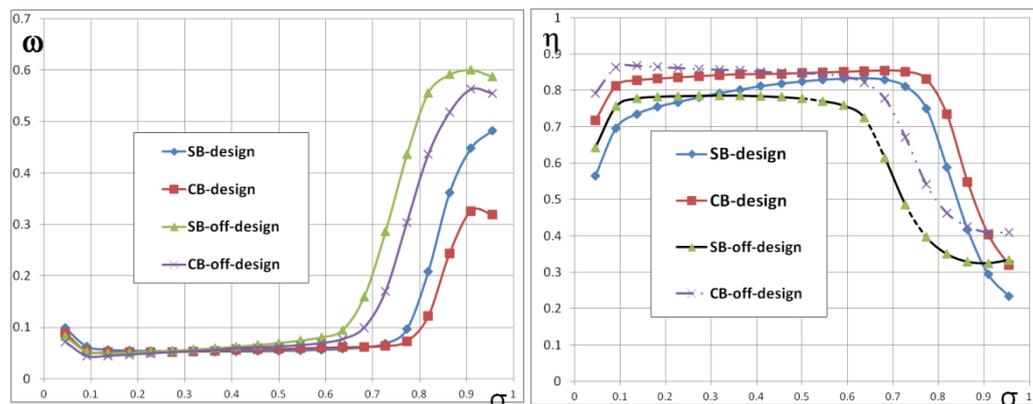


Figure 10: Distribution of static pressure coefficient on the SS.

The definition of the total pressure loss coefficient is expressed as $(\omega) = (\psi_i - \psi)$ and the local total efficiency is expressed as $(\eta) = (\psi / \psi_i)$ [4]. Figure 11a shows the two rotors at the design flow rate having the same pressure loss coefficient at the hub and midspan until 75% of span, while the values of pressure loss at the blade tip are more increased for the SB rather than for CB rotor, as observed by Halder and Samad [12]. Whereas, at the off-design flow rate, the two rotors are having the same pressure losses up to 60% of blade spanwise. Toward the blade tip, a similar tendency of total pressure loss occurs, but with higher values for the off-design flow rate than for the design flow rate. In the axial fan blade shape design, Seo et al. [11] mentioned that, the efficiency parameter is the major objective for the design by using the stacking line technique. Figure 11b shows the local total efficiency profiles at the design and off-design flow rate along the span. The CB rotor at off-design flow rate exhibits the highest efficiency at the hub, whereas CB rotor at design flow rate exhibits the highest efficiency at the tip.



a) total pressure loss coefficient

b) total efficiency

Figure 11: Local spanwise distributions.

5. Conclusions

The purpose of this work is to examine the effects of the Non-radial stacking line by means of an applied curved surface to the rotor of an axial fan at the design and off-design flow rates. Comparative CFD studies have been carried out on the CB and SB rotors, without geometrical correction of the elemental blade cascades of the blade cross sections.

CB rotor obviously exhibits the highest local total efficiency in the design flow rate at the blade tip and in the off-design flow rate at the blade hub. While SB rotor exhibits the lowest local total efficiency in the design flow rate at the blade hub and in the off-design flow rate at the blade tip

References

- [1]. C. Clemen, and U. Stark, "Compressor Blades with Sweep and Dihedral: a Parameter Study," European Conference Turbomachinery Fluid Dynamics and Thermodynamics, Proceedings, 5th Prague, Czech Republic, , pp. 151-161. 2003
- [2]. Corsini, and F. Rispoli, "Using Sweep to Extend the Stall-Free Operational Range in Axial Fan Rotors," J Power and Energy, Vol. 218, pp. 129-139, Proc. Instn Mech. Engrs, Part A, 2004
- [3]. M. Govardhan, O.G. Krishna Kumar, N. Sitaram, "Investigations on Low Speed Axial Compressor with Forward and Backward Sweep" Journal of Thermal Science 16(2):121-133, 2007
https://www.researchgate.net/publication/226046716_Investigations_on_low_speed_axial_compressor_with_forward_and_backward_sweep.
- [4]. J. Vad, A. R. A. Kwedikha, Cs. Horváth, M. Balczó, M. M. Lohász and T. Régert, "Combined Aerodynamic Effects of Controlled Vortex Design and Forward Blade Skew in Axial Flow Rotors," Journal of Power and Energy, Proceedings of the Institution of Mechanical Engineers – Part A , Vol. 221, pp. 1011-1023. 2007
- [5]. Ivo Wenneker, "Computation of flows using unstructured staggered grids" Dissertation at Delft University of Technology. Delft, the Netherlands, 2002.
<https://repository.tudelft.nl/islandora/object/uuid%3A5abe2984-0cdf-4af9-bd92-188c29a7e5c4>
- [6]. Z. Ali, and P.G. Tucker, "Multiblock Structured Mesh Generation for Turbomachinery Flows," In: 22nd International Meshing Roundtable, pp. 165-182 Orlando, Florida, 13-10-2013 https://link.springer.com/chapter/10.1007/978-3-319-02335-9_10
- [7]. M. G. Beiler, "Untersuchung der dreidimensionalen Strömung durch Axialventilatoren mit gekrümmten Schaufeln". Doctoral Dissertation, Nr. 298, Universität-GH-Siegen, VDI Verlag Düsseldorf, Reihe 7: Strömungstechnik, 1996
- [8]. M.G. Beiler, and T.H. Carolus, "Computation and Measurement of the Flow in Axial Flow Fans With Skewed Blades," J Turbomachinery, Vol. 121, pp. 59-66. ASME. 1999.
<https://asmedigitalcollection.asme.org/turbomachinery/article-abstract/121/1/59/438320/Computation-and-Measurement-of-the-Flow-in-Axial?redirectedFrom=fulltext>
- [9]. C. Clemen, V. Gümmer, M. Goller, H. Rohkamm, U. Stark, and H. Saathoff, "Tip-aerodynamics of forward-swept rotor blades in a highly-loaded single-stage axial-flow low-speed compressor," 10th International Symposium on Transport Phenomena and Dynamics of Rotating Machinery, Honolulu, (CD-ROM Proceedings), Paper No. 027, ISROMAC, 10. 2004
- [10]. N. Yamaguchi, T. Tominaga, S. Hattori, and T. Mitsuhashi, "Secondary-Loss Reduction by Forward-Skewing of Axial Compressor Rotor Blading," Proceedings Yokohama International Gas Turbine Congress, pp. II.61-II.68. Yokohama, Japan, 1991.
<https://www.semanticscholar.org/paper/IGTC-8-Secondary-Loss-Reduction-by-Forward-Skewing-Yamaguchi-Tominaga/a9ff6fe325cb4bd808f8daec86c3b192f5ecf8e2>
- [11]. S. J. Seo, S. M. Choi, and K. Y. Kim, "Design of an Axial Flow Fan with Shape Optimization," 8th Asian International Fluid Machinery Conference, China, 2005,
<https://www.researchgate.net>
- [12]. Paresh, S. Abdus, "Marine Energy Turbine Performance: Effect of Blade Sweep" 5th International Conference on Advances in Energy Research, Mumbai, India. ICAER, 15-17, December, 2015. <https://www.sciencedirect.com/>

Dimensions Effect of the Rotating Fluid Zone on the Results when CFD Modeling of Friction Stir Welding

Esam Hamza *

emh6141@yahoo.com

Authority of Natural Science Research and Technology, Tripoli, Libya

ABSTRACT

This study aims to investigate the influence of the shape and size of rotating fluid region on the results of computational fluid dynamic (CFD) model of friction stir welding (FSW). Accordingly, 3D time-dependent CFD based model was used to simulate friction stir welding of aluminum alloy AA2014-T6. A rotating fluid region was proposed to represent the Thermo-mechanical affected zone (TMAZ) in the model where the temperature distribution within the materials being welded together has been critically analysed. By using different shapes and dimensions for the fluid region, a number of numerical experiments has been carried out. The results revealed that the thermal profile for the circular zone indicates the best agreement among the other with the experimental thermal profile. Additionally, a semi-empirical equation was developed to calculate the maximum temperature based on the dimensions of TMAZ.

Keywords:

Friction Stir Welding (FSW),

Computational Fluid Dynamics (CFD),

Thermo-mechanical Affected Zone (TMAZ).

1- Introduction

Friction Stir Welding (FSW) was developed by Wayne Thomas et al at The Welding Institute (TWI) in The United Kingdom in 1991[1]. This welding process overcomes many of the problems associated with traditional joining techniques. FSW is applied in different field of industry due to its advantages such as high quality, the welding temperature does not exceed the melting point and shielding gas is not required. Furthermore, it's an effective alternative to weld dissimilar materials [2].

As a solid-state welding process, Friction Stir Welding (FSW) is carried out using a non-consumable, rotating and translating tool that is brought at the interface between the two work pieces to be welded. The heat generation leads to increase the temperature of the surfaces being welded in the near-tool region, which makes them soft. Then, the tool movement and its high mechanical pressure melds and joins both materials together.

For the purpose of understanding the physical phenomena associated with FSW particularly those are represented in the thermomechanical interrelations, several computational fluid dynamics (CFD) tool-based studies have been conducted. For simplification purpose, assumptions and different modelling techniques have been adopted particularly those related to defining the fluid region which represents thermo-mechanical affected zone (TMAZ) that taking place in the workpiece around the welding tool. However, the change in the proposed shape and size of the fluid region may affect the model accuracy.

The earlier works to CFD modelling of FSW have been mostly carried out in two-dimensions [3, 4]. However, P. A. Colegrove [5] have published a critical review of modelling of FSW which revealed that one of the most important issues is the development and validation of robust 3-

dimensional model. After that, P. A. Colegrove and Shercliff [6] described the application of (CFD) code in modelling of friction stir welding. In their work, CFD package FLUENT 15.0 was used to build a 3-dimensional steady-state model which includes five regions. Firstly, a rotating zone represents the fluid region nearby the tool and moves at its rotational speed. Secondly, a slightly wider strip surrounds the rotating zone and extending to the length of the workpiece which represents the material flow at the tool travelling speed. The rest of the workpiece in both sides is modelled as solid aluminium regions to simulate the thermal response far away from the deformation zone. Finally, to calculate the heat losses to the tool and backing plate both of them were defined as solid regions. The results have demonstrated that the CFD package, Fluent can be used to analyse the heat and material flow in FSW. However, the study has not shown the effect of the change in rotating zone dimensions as the size of the deformation zone in the model was larger than that recorded in the experimental work.

The 3-dimensional CFD model developed by S.Z Aljoaba et al [7] consisted two regions. Firstly, the tool which is assumed to be a solid body. Secondly, the workpiece which is assumed to be a fluid region. Although, the model has been validated with experimental data the rotating zone concept was not considered and temperature at the sides of the workpiece were kept fixed at room temperature.

The work published by Atharifar et al [8] used Fluent to solve the governing equations in the computational domain. In order to simplify the calculation time, a cylindrical computation zone was assumed with a diameter twice the diameter of the tool shoulder and rotates at the tool speed. This cylindrical fluid region is small compared to the actual size of the workpiece. The numerical results were found in a good agreement with the experimental data. However, the study has not given any information about the effect of the shape and size of the computation domain on the results.

Z. Zhang et al [9] have published a study includes an investigation of the effect of shoulder size on the temperature distributions and the material deformations in FSW. The workpiece was represented in the model used by a rounded zone with 80 mm diameter. The welding speed has been applied in the inflow region while the tool was considered to be a rigid body.

A 3-D transient CFD model was developed by ZHENZHEN YU et al [10] where the tool rotational motion was defined by giving the speed to a rotating zone which surrounds the tool. The authors used this model to investigate the influence of the pin thread on the thermal distribution and the material flow behaviour.

In the work carried out by S. Kang et al [2] the rotation affected zone RAZ concept was used in the computational domain of the model where the momentum and energy equations were solved using the CFD program Fluent. In the model used, RAZ represents the thermo-mechanically affected Zone TMAZ in real FSW process where the material plastic deformation occurs. The size of RAZ has determined in terms of dimensional variables a and b where a parametric study was carried out for various ranges of a and b to exactly define the RAZ. Consequently, the thermal analysis has shown when the value of a is 1mm and b is 2-3mm the temperature results have given the best agreement with the reference data. However, the study has been conducted in steady state condition and did not include the effect of the change in RAZ shape. Furthermore, the fluid zone proposed in the model developed is still big which includes the inflow and outflow region as well as the RAZ.

In advanced step to improve the transient CFD modelling of FSW, E. Hamza et al [11] has rotated a circular conical zone around the tool. The translational motion of the tool has been simulated by using the dynamic mesh while sliding mesh technique was employed to mimic the rotational

movement. As the model was validated with experimental data, influence of the tool pin size on the temperature distribution has been investigated.

The majority of the published works presented above have not focused on the effect of the proposed rotational zone on the modelling results particularly in the transient analysis conditions. Hence, the current study is an extension of the works conducted by S. Kang and E. Hamza [2, 11], where different shapes and sizes for TMAZ are proposed and their effects on the thermal cycle of FSW will be studied. Consequently, a dimensionless formula will be developed to describe the relationship between the temperature and TMAZ dimensions. The model that will be used was constructed using the CFD commercial program Fluent 16.0.

2- Numerical Modelling

The 3-D CFD transient model used in this study was developed and validated by E. Hamza [11] where the computational domain that is shown in Figure 1 consists of two regions. The first region represents the thermo-mechanical affected zone (TMAZ) around the tool. This zone is treated as a fluid region which has a radial and conical shape, where its dimensions are based on design variables a and b as displayed in Figure 2 [2]. The tool pin has a conical triangular profile with a length of 4.7mm, where the pin side lengths at the root and the tip of the pin are 5.19mm and 3.11mm respectively. The shoulder has a constant diameter of 12mm. The dimensions of the plates to be welded are 300mm x 100mm x 5mm.

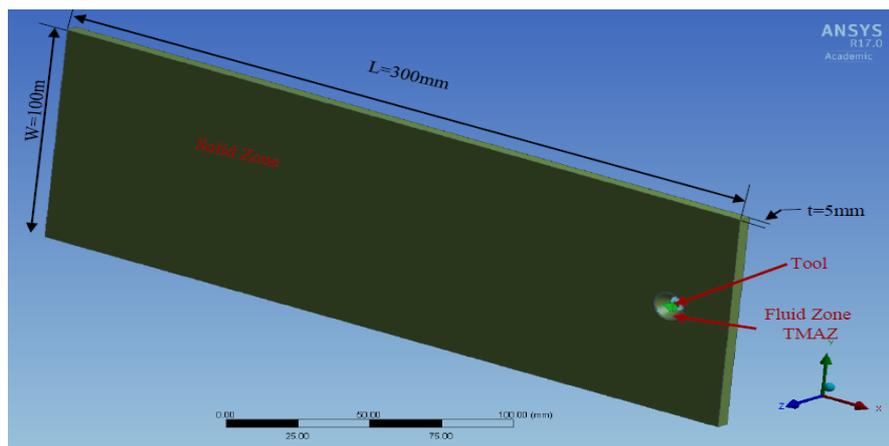


Figure 1: Model Geometry

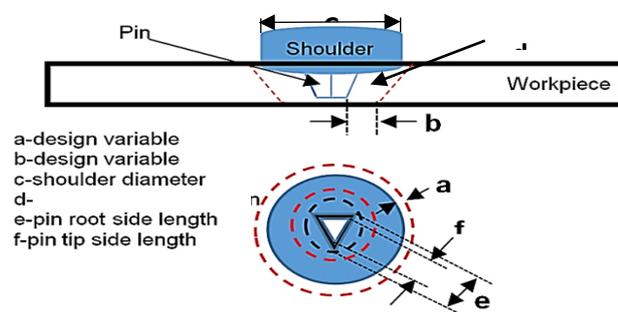


Figure 2: Details of the Computational Domain

Figure 3 shows the meshing of the computational domain where a new methodology was considered in which the translation motion of the tool has been specified using dynamic mesh technique, and considering the path followed by the tool as a solid region rather than fluid. As the temperature and velocity are steeply gradient in the (TMAZ), it has been finely meshed to capture

this level of gradient whereas the rotation of the tool has been specified using sliding mesh technique. Due to configurations, tetrahedral elements were used for both the fluid and the solid regions. With a view to simulate FSW, the boundary conditions have been defined where the tool was specified with rotational and travel speeds of 1000rpm 7.73mm/s, respectively. User Defined Function (UDF) subroutine has been used in applying of the travel speed. Thermally, to define the process of heat generation, a heat flux was applied to the tool surface after it had been calculated by the analytical heat generation equation given as GADAKH [13]

$$Q = \frac{2}{3} \pi \mu \omega P (R_{Shoulder}^3 + \frac{9}{4\pi} R_{prob}^2 H_{prob}) \quad (1)$$

where, μ is the friction coefficient taken as 0.5, ω is the rotational speed of the tool, P is the plunging pressure which was kept at 90Mpa. During the heat transfer analysis, the convection heat transfer coefficient from the top and side surfaces of the workpiece are 25W/m² °C, and whereas the bottom surface is supported by a backing plate, the coefficient value is considered to be 200W/m² °C. Kang [2].

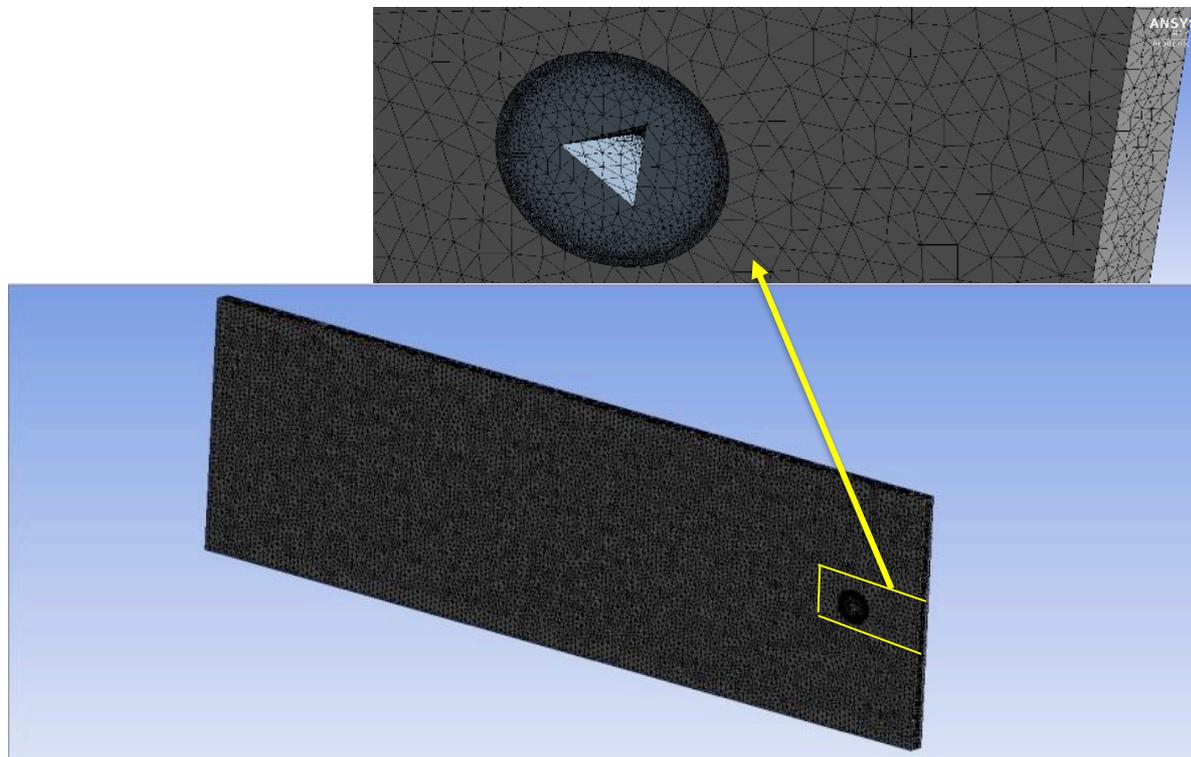


Figure 3: *Meshing of the Computational Domain*

Due to the resultant heat generation from the rotational speed of the tool, the material behaves as non-Newtonian viscoplastic fluid in the near tool region, with laminar flow conditions. The fluid flow is governed by mass and momentum conservation equations, whereas, in order to specify the conductive and convective heat transfer rates, the energy conservation equation is also employed. These equations can be found in many published literatures such as by Arora [12]. By using the Finite Volume Method (FVM), these governing equations for 3D transient heat transfer and fluid flow are discretized and solved iteratively during 6 seconds which are the welding time. More information about the model such as the material properties of the aluminium alloy AA2014-T6 are provided in the literature [2, 11, 14].

3- Investigating the effect of fluid zone (TMAZ) shape on the welding thermal cycle

In order to investigate the effect of the fluid region (TMAZ) shape in the current model, circular conical zone was proposed at the beginning with design variables $a = 1\text{mm}$ and $b = 2.5\text{mm}$ [2]. Then, in terms of a and b , the rear half of the zone has been changed by adding certain values for a and b where the region has got semi elliptical shapes. The reason behind that is just to examine the effectiveness of zone shape at the rear half in which the fluid zone shape and size are expected to be bigger than in the circular one as there is more hot metal accumulation in the trailing side than in the leading side. Figure 4 illustrates the fluid zone with different shapes. For conducting the study, six experiments have been designed with different shapes around a tool with a triangle pin as shown in table 1. As the model used in the current study has already been validated [11] and to study the effect of shape and dimensions of the fluid zone (TMAZ), the thermal cycles were monitored in all of proposed numerical experiments at a point with a distance 4mm from the weld centre, and a depth of 2mm from the top surface of the plates.

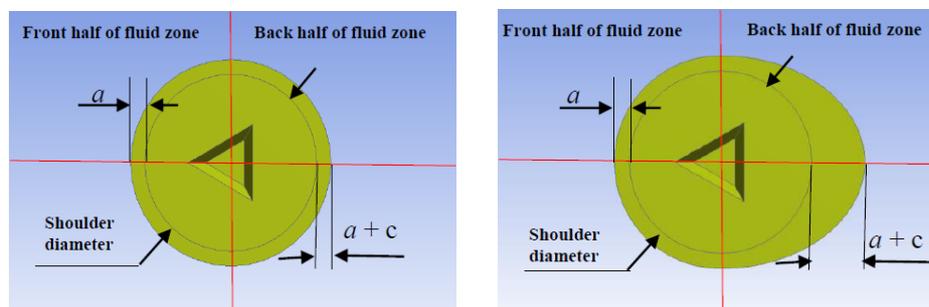


Figure 4: conical circular and conical semi elliptical fluid zones.

Table 1: different shapes of fluid zones

Experiment	Zone shape	Design variables(mm)	$c(\text{mm})$	$a+c(\text{mm})$	$b+c(\text{mm})$
Ex.1	circular	$a=1, b=2.5$	0	1	2.5
Ex.2	semi elliptical	$a=1, b=2.5$	1	2	3.5
Ex.3	semi elliptical	$a=1, b=2.5$	1.5	2.5	4
Ex.4	semi elliptical	$a=1, b=2.5$	2	3	4.5
Ex.5	semi elliptical	$a=1, b=2.5$	2.5	3.5	5
Ex.6	semi elliptical	$a=1, b=2.5$	3	4	5.5

4- Investigating the effect of fluid zone (TMAZ) dimensions on the welding thermal cycle

The circular conical shape of the fluid zone (TMAZ) has been chosen to carry out this analysis whereas several values for design variables a and b were proposed. Based on the proposed values of a and b , fifteen different sizes of TMAZ were used in the numerical experiments as shown in table 2. In order to monitor the thermal cycle, the time dependent temperature profile was recorded for all proposed numerical experiments at a point with a distance 4mm from the weld centre, and a depth of 2mm from the top surface of the plates. As well known, the maximum temperature is one of the significant thermal features that has an important role in defining the final microstructure of the welded joint and its properties as well which in role estimate the final weld quality. Therefore, the maximum global temperature $T_{g, max}$ was monitored and recorded for each value of a and b . After conducting the set of experiments that have been designed based on

different fluid zone dimensions and using a tool with the triangular pin, a relationship between the temperature ratio, design variables and the shoulder radius is expressed using multiple variable regression analysis.

Table 2: *different sizes of fluid zone*

Experiments No.	Value of a (mm)	Value of b (mm)
Ex.1	0.5	0.5
Ex.2	0.5	1.5
Ex.3	0.5	2.5
Ex.4	0.5	3.5
Ex.5	0.5	4.5
Ex.6	1	0.5
Ex.7	1	1.5
Ex.8	1	2.5
Ex.9	1	3.5
Ex.10	1	4.5
Ex.11	1.5	0.5
Ex.12	1.5	1.5
Ex.13	1.5	2.5
Ex.14	1.5	3.5
Ex.15	1.5	4.5

5- Results and discussion

As the numerical experiments shown in table 1 have been conducted for different shapes of TMAZ, what can be seen in Figure 5 is the thermal cycles for those simulated welding processes. The thermal profile of TMAZ with $a=1\text{mm}$ and $b=2.5\text{mm}$ (the circular conical zone) indicates the best agreement among the other with the experimental thermal profile. By considering the maximum temperature of the best CFD thermal profile, one can notify that there is no much difference compared to the maximum temperature of the other calculated thermal profiles. However, the cooling rate becomes faster when using the semi elliptical shape of TMAZ rather than the circular one. It would be said that the heat dissipation between the fluid zone and the solid zone is enhanced as long as the heat transfer area is increased due to increase in the fluid zone dimensions. For more explanation, the static temperature distribution has been depicted in Figure 6. It can be seen that the temperature is higher on the advancing side than on the retreating side, which is due to the opposite direction of the material flow on the advancing region to that of the tool motion.

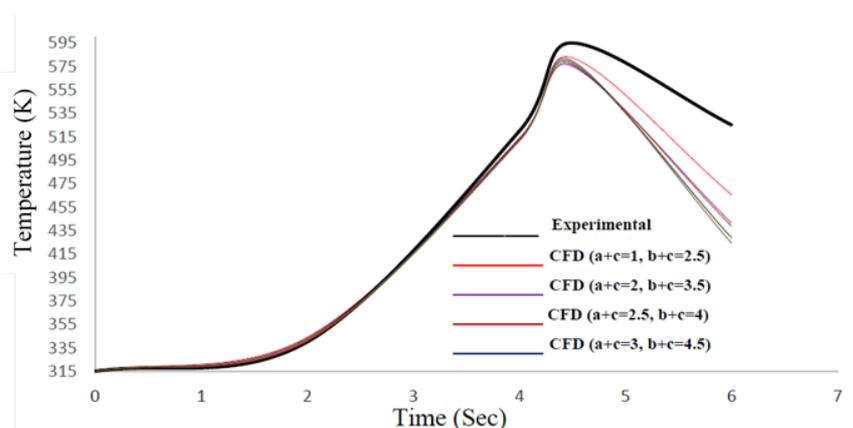


Figure 5: comparison between different thermal cycles for different TMAZ shapes corresponding measured thermal cycle

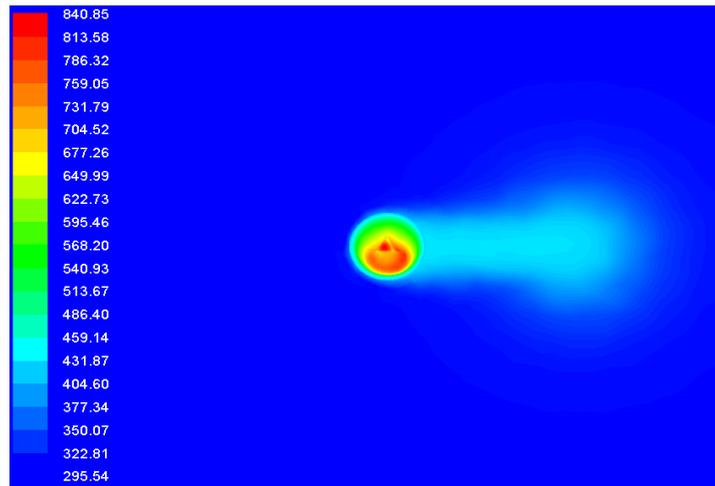


Figure 6: Temperature distribution within the flow domain at the end of the operational cycle

Based on the results recorded in section 3, the circular conical shape of TMAZ has been considered to conduct the analysis in the current section. Figures 7, 8 and 9 exhibit the comparison between different thermal cycles for different values of a and b . Looking at these figures and considering the temperature trend with time, maximum temperatures and cooling rates, it can be found that some values of the maximum temperature exceed the corresponding measured values. Those values are not physically accepted in terms of fact of the maximum temperature should not exceeds the experimental value for the same due to ignoring the part of heat generation from material deformation. Accordingly, the thermal profile of $b=2.5\text{mm}$ (blue

curve) for the three values of a has in general the best agreement with the experimental data among the other profiles in the same set of experiments. However, by drawing an analogy between the best three profiles; the one with $a=1\text{mm}$ and $b=2.5\text{mm}$ depicts the optimum behaviour based on the three mentioned thermal features. Consequently, this is an approval in a transient state for the work done by Kang et al. [2] who has used a steady state CFD model in finding the optimum values for a and b which are 1mm and 2.5mm respectively.

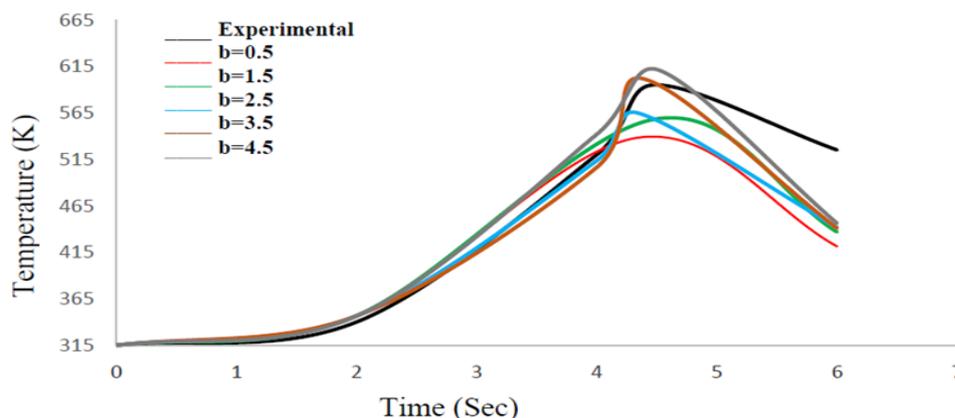


Figure 7: a comparison between different thermal cycles for different values of b when $a=0.5$

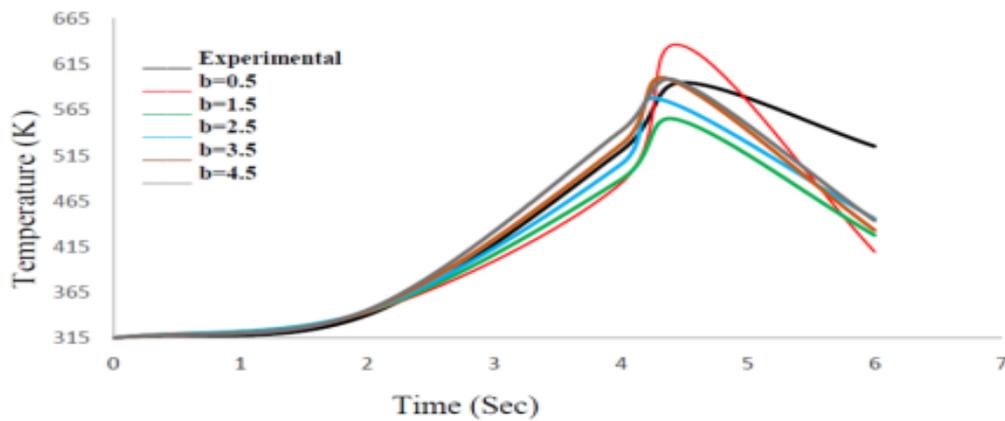


Figure 8: a comparison between different thermal cycles for different values of b when $a=1$

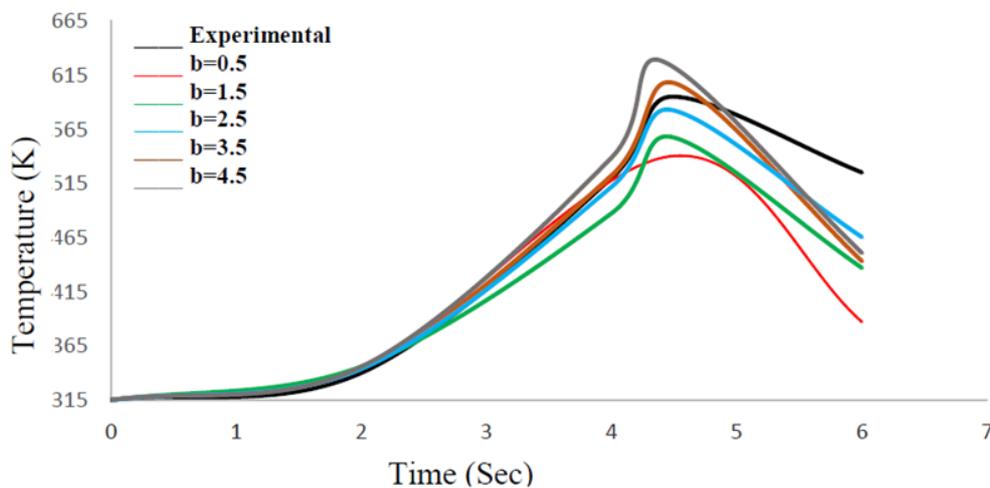


Figure 9: a comparison between different thermal cycles for different values of b when $a=1.5$

Table 3 presents the recorded values of $T_{g, max}$ against the values of a and b in the fifteen numerical experiments. Noticeably, a proportional relationship between the values of b and $T_{g, max}$ is taking place, where the increasing in b is met by an increasing in $T_{g, max}$. Beside the absence of heat amount due to the plastic deformation, the quantity of fluid mass increases according to the increase in b which leads to less volume change compared to a [2]. Therefore, the heat resistance becomes higher as the outer fluid particles experiences decreasing in their kinetic energy which allow the temperature to be risen up.

Table 3: the computed values of $T_{g, max}$ according to different values of a and b

$a=0.5$ (mm)		$a=1$ (mm)		$a=1.5$	
b (mm)	$T_{g, max}$	b (mm)	$T_{g, max}$	b (mm)	$T_{g, max}$
0.5	735	0.5	720	0.5	726
1.5	743	1.5	734	1.5	734
2.5	760	2.5	780	2.5	760
3.5	790	3.5	799	3.5	777
4.5	796	4.5	799	4.5	798

Based on the data from CFD experiments with the aid of multiple variable regression analysis, a dimensionless formula can be developed to describe the relationship between the temperature ratio $T_{g, max} / (T_{g, max} - T_{in})$ and geometrical variables a , b , and the shoulder radius R_s .

$$\frac{T_{g, max}}{T_{g, max} - T_{in}} = \frac{1.66 \left(\frac{a}{R_s}\right)^{0.0044}}{\left(\frac{b}{R_s}\right)^{0.0311}} \quad (2)$$

where T_{in} is the initial temperature which is equal to 315°K. Then, Equation (2) can be modified to find a direct expression to calculate $T_{g, max}$ as following:

$$T_{g, max} = - \left[\frac{\frac{1.66 \left(\frac{a}{R_s}\right)^{0.0044}}{\left(\frac{b}{R_s}\right)^{0.0311}}}{1 - \frac{1.66 \left(\frac{a}{R_s}\right)^{0.0044}}{\left(\frac{b}{R_s}\right)^{0.0311}}} \right] \times T_{in} \quad (3)$$

It is observed that b has a greater effect on $T_{g, max}$ than a . In order to authorise the utility of equation (3), the difference between the $T_{g, max}$ values from equation (3) and the CFD $T_{g, max}$ values, both of them were plotted in Figure 10. Obviously, within the $\pm 5\%$ error bounded of the developed equation it would be seen that 100% of the data lies in this range. Therefore, Equation 3 can be used to determine the maximum temperature with reasonable accuracy [15] particularly; the current study would be experimentally and analytically difficult so far. Last but not least, the above expressed equation is valid for $R_s=6\text{mm}$.

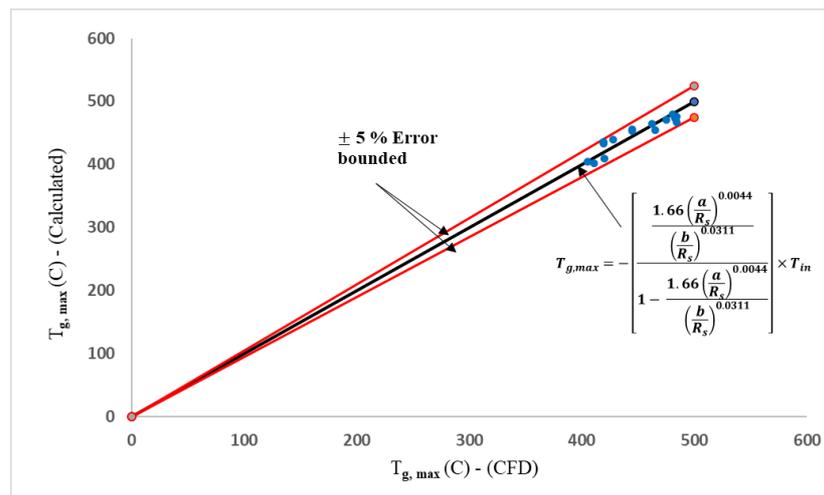


Figure 10: a comparison between the maximum temperature computed by the developed equation and that CFD estimated

6- Conclusion

In this work, heat transfer analysis is carried out using the CFD commercial program 'Fluent 17.0'. In order to achieve the objectives defined for the current study, a number of numerical experiments were carried out. When examining the effect of TMAZ shape on the welding thermal features, the thermal profile of TMAZ with $a=1\text{mm}$ and $b=2.5\text{mm}$ (the circular conical zone) indicates the best agreement among the other with the experimental thermal profile. Additionally, the cooling rate becomes faster when using the semi elliptical shape of TMAZ rather than the circular one.

Similarly, when analysing the effect of TMAZ dimensions on the thermal features, the thermal profile of $b=2.5\text{mm}$ for the three values of a has in general the best agreement with the experimental data among the other profiles in the same set of experiments. In spite of that, and between the best three profiles, the one with $a=1\text{mm}$ and $b=2.5\text{mm}$ depicts the optimum behaviour. Moreover, the developed equation based on multiple variable regression analysis has revealed that $T_{g, max}$ is greater influenced by b than a and the usefulness of this formula has been authorised where within ± 5 error bounded of the developed equation; 100% of the data lies in this range. Finally, this work can be extended by considering other operational and geometrical variables.

References

- [1]. W. Thomas, et al, "Friction stir butt welding Int Patent App PCT/GB92/02203, and GB Patent App 9125978.8," *US Patent*, vol. 5, pp. 317-460, December 1991.
- [2]. S.W. Kang, et al, "A study on heat-flow analysis of friction stir welding on a rotation affected zone," *Journal of mechanical science and technology*, vol. 28, no.9, pp. 3873-3883, 2014. <https://doi.org/10.1007/s12206-014-0851-6>.
- [3]. X. Deng, et al, "Two-dimensional finite element simulation of material flow in the friction stir welding process," *Journal of manufacturing processes*, vol. 6, no. 2, pp. 125-133, 2004. [https://doi.org/10.1016/S1526-6125\(04\)](https://doi.org/10.1016/S1526-6125(04)).
- [4]. T. Seidel, et al, "Two-dimensional friction stir welding process model based on fluid mechanics," *Science and technology of welding and joining*, vol. 8, no. 3, pp. 175-183, 2003. [https://doi.org/ DOI 10.1179/136217103225010952](https://doi.org/DOI%2010.1179/136217103225010952).
- [5]. P. A. Colegrove, "Modelling of friction stir welding," *Doctoral Thesis*, Cambridge University, 2004. <https://doi.org/10.17863/CAM.14008>.
- [6]. P. A. Colegrove, et al, "3-Dimensional CFD modelling of flow round a threaded friction stir welding tool profile," *Journal of Materials Processing Technology*, vol. 169, no. 2, pp. 320-327, 2014. <https://doi.org/10.1016/j.jmatprotec>
- [7]. S. Aljoaba, et al, "Modeling of friction stir processing using 3D CFD analysis," *International journal of material forming*, vol. 2, no. 1, pp. 315-318, 2009. <https://doi.org/10.1007/s12289-009-0662-y>.
- [8]. Atharifar, et al, "Numerical and experimental investigations on the loads carried by the tool during friction stir welding," *Journal of Materials engineering and Performance*, vol. 18, no. 4, pp. 339-350, JMEPEG, 2009. <https://doi.org/10.1007/s11665-008-9298-1>.
- [9]. Z. Zhang, et al, "Effect of shoulder size on the temperature rise and the material deformation in friction stir welding," *The International Journal of Advanced Manufacturing Technology*, vol. 45, no. 9, pp. 889-895, Int J Adv Manuf Technol, 2009. <https://doi.org/10.1007/s00170-009-2034-7>.
- [10]. Yu. Zhenzhen, et al, "Transient heat and material flow modeling of friction stir processing of magnesium alloy using threaded tool," *Metallurgical and Materials Transactions A*, vol. 43, no. 2, pp. 724-737, 2012. <https://doi.org/10.1007/s11661-011-0862-1>.
- [11]. E. Hamza, et al, "Computational Fluid Dynamics based Transient Thermal Analysis of Friction Stir Welding," *6th International and 43rd National Conference on Fluid Mechanics and Fluid Power*, December, 2016. <http://eprints.hud.ac.uk/id/eprint/29395/1/Esam.pdf>.
- [12]. A.Arora, et al, "Strains and strain rates during friction stir welding," *Scripta Materialia*, vol. 61, no. 9, pp. 863-866, 2009. <https://doi.org/10.1016/j.scriptamat.2009.07.01>.



-
- [13]. V. Gadakh, et al, “Analytical modeling of the friction stir welding process using different pin profiles,” *Welding Journal*, vol. 94, no. 4, pp. 115-124, 2015.
 - [14]. M. Mehta, et al, “Numerical modelling of friction stir welding using the tools with polygonal pins,” *Defence Technology*, vol. 11, no. 3, pp. 229-236, 2015.
 - [15]. T. Asim, et al, “CFD based investigations for the design of severe service control valves used in energy,” *Energy Conversion and Management*, vol. 153, pp. 288-303, 2017.
<https://doi.org/10.1016/j.enconman.2017.10.012>.

Risk-Based Maintenance (RBM) Approach for Identifying the Optimum Time of Whole Shutdown (SD) for Gas Liquid Recovery Unit; Processing Columns as a Case Study

Abdelnaser Elwerfalli^{1*}, Salih Alsadaie²

¹ gq7qg@yahoo.com, ² salsadaie@hotmail.com

¹ Department of Mechanical Engineering, College of Mechanical Engineering Technology
Benghazi – Libya

² Department of Chemical Engineering, University of Sirte – Libya

ABSTRACT

Pieces of equipment for any a processing unit, which run continuously under severe conditions usually exposed to undesirable failures such as a corrosion, leakage and other reasons due to over pressure and fluctuated temperatures. These failures could be resulted in huge risk of the unit. Therefore, these pieces of equipment should be subjected to risk-based maintenance approach in order to conduct planned shutdown event without taking the recommended periods of the original equipment manufacturers into account, which may not be represented the optimum solution of planned shutdown scheduling in the long-term due to operating conditions that differ from a unit to another. The purpose of this work is to determine optimum interval of planned shutdown for unit based on processing columns to avoid an unplanned estimation of shutdown, mitigate risk, extend critical equipment life, maximise uptime, decrease maintenance cost, reduce production losses and improve reliability of system. The results of risk-based maintenance application in gas liquid recovery unit demonstrated that interval of planned shutdown could be increased based on the risk assessment related to processing columns.

Keywords:

Planned Shutdown (SD),
Risk-Based Maintenance (RBM),
Failure Analysis, and
processing columns.

1- Introduction

Static equipment pieces in Gas Liquid Recovery Unit (GLRU) consider a complex in terms of process, operation and maintenance, especially which run continuously under severe operating conditions. Processing columns is a part of equipment that can be tended to deteriorate over time due to aging, corrosion, fatigue and fluctuating temperatures and pressures. Therefore, unit these pieces cannot be maintained or inspected during the normal operation of a unit unless facilities of unit are total shutdown to execute Planned Shutdown (SD) activities: inspections, modifications, installations, replacements and repairs [1]. Duffuaa and Ben Daya [2] defined also SD as an entire shutdown of the unit during a certain time period to carry out SD activities associated with inspection, repair, overhaul, modification, and replacement in accordance with Scope of Work (CWo).

Ghosh and Rao [3] proposed optimisation of the maintenance intervals using the reliability based on cost/benefit ratio. Megow et al. [4] stated that the SD interval of a large unit can be repeated more than one year. Rusin and Wojaczek [5] presented optimizing maintenance intervals of power machines by taking the risk into account. Obiajunwa [6] reported that interval of SD for petrochemical and refinery plant conducted every two years and power plant executed every four

years. Hameed and Khan [7] also presented a framework to estimate the risk-based shutdown interval to extend intervals between shutdowns for a processing unit. Swart [8] reported that historically, intervals of SD identified without any real strategy associated with operating process. Swart [8] stated that the estimation of the SD interval is either indiscriminate or has become as a redundant.

The constant identification of SD interval means that SD activity was not based on the residual life of the equipment that can increase risk due to fixed-interval. Consequently, this study is designed to estimate SD interval of a processing unit based on the processing columns, while ensuring that the overall cost is kept to a minimum. RBM provides an efficient way to select the most critical processing columns based on the RBM approach to manage assets in comparison to the individual equipment strategy and achieve better results with less operating expenses.

2- Risk-Based Maintenance (RBM) of Processing Columns

Processing columns comprise seven columns that can be classified into C-701 to C-707. These columns are spread across many units of gas plant (Gas Liquid Recovery Unit, Gas Treating Unit, Gas Drying Unit, Cryogenic Column and Fractionation Unit).

The GLRU experts demonstrated that most of inspected columns that were exposed corrosion due to fluctuating feed temperatures resulting from errors in the operating process and sometimes specification of natural gas.

According to the failures records of seven columns in the GLRU, it was found many trays and caps in these columns, which were prone to corrosion due to the presence of water vapour in natural gas resulting from fluctuated temperatures as shown in Figure.1. These can be resulted to a lot of consequences:

- Reduce the ability of gas to flow in the flowlines and process systems.
- Water vapour causes corrosion in trays parts.
- At low temperature, water vapour forms hydrates - complicated molecules of hydrocarbon liquid and water, causing blockage of lines.

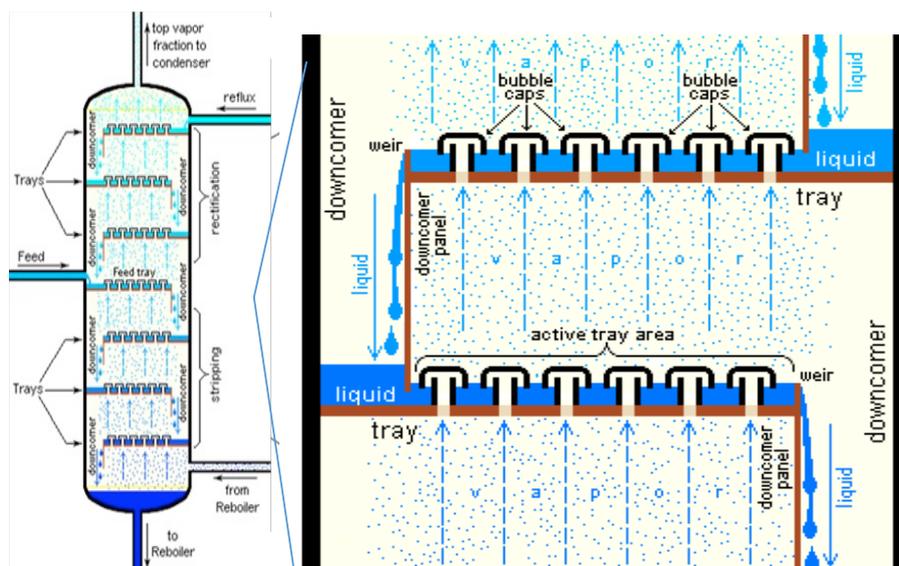


Figure 1: The prone trays to failure in GLRU columns

These parts are one of the outstanding challenging problems in the natural gas industries caused by complex operating conditions that constitute the highest risk. Therefore, it is necessary to

highlight the RBM approach to mitigate consequences on the functional performance of the unit resulting from over pressure and fluctuating temperatures. RBM approach has played an important role in the decision-making process due to complex processes that required a higher reliability of their facilities [9].

The proposed risk analysis matrix is applied according to RBM approach to identify the most critical pieces of processing columns that represent the highest risk on the unit facilities. These critical pieces of processing columns of a unit can be distributed the risks' matrix (5 x 5) according to Probability of Failure (PoF) and Consequences of Failure (CoF) on GLRU in terms of Environment Damage (ED), Production Losses (PL) and Asset Damage (AD). The estimated risk of system can be identified by using Equations (1) and (2), below:

$$\text{Estimated Risk} = \text{Probability of Failure} \times \sum \text{Economic Consequence of Failure} \quad (1)$$

$$ER = [1 - \prod_{i=1}^n Ri(t)] \times [\sum CoF_{ED} + CoF_{PL} + CoF_{AD}] \quad (2)$$

This paper focuses on seven pieces of columns located in GLRU. Those columns distribute on the risk matrix (5x5) to estimate risk ranking as shown in Figure 2. Two of seven columns are rated in the low risk zone, two pieces of columns are classified in the moderate risk zone and three columns (C-706 and C-707A/B) are rated in the high risk zone.

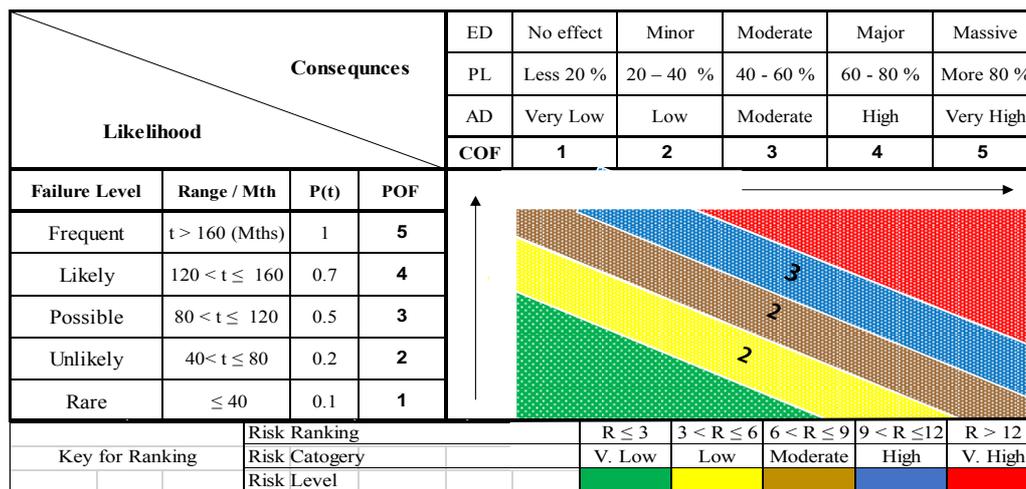


Figure 2: Risk matrix for seven processing columns

As a result, four pieces were rated in the low and moderate risk zone respectively, two pieces of columns are classified in the moderate risk zone. These columns are not required SD and must be excluded from the current SD. However, these pieces should be taken into account in the next cycle of SD according to priorities. Three of seven columns were classified in the high risk zone. These items must be involved to a scheduled SD to be determined the reliability and unreliability function using the appropriate probability distributions.

3- Failure Analysis

This case is designed to focus on the outcome of risk matrix (the located in the high risk zone – blue area). Three of seven pieces of columns that were classified in the high risk zone. Based on the shape parameter (β) for each one, it was found that the Weibull model was the most appropriate in the estimate of Probability of Failure $F(t)$ due to its inherent flexibility of modelling the behaviour for other distributions.

3.1 The Weibull Distribution

The Weibull distribution can mimic the behaviour of other distributions, such as normal for ($\beta = 3.5$) and exponential for ($\beta = 1$) distributions. A decreasing failure rate ($\beta < 1$) corresponds to an early life failure or infant mortality.

A constant failure rate ($\beta = 1$) suggests that units are failing from random events. An increasing failure rate ($\beta > 1$) suggests that wear out is occurring and that parts are more likely to fail over time [3]. The shape parameter β estimated by the failure data provides an insight into the failure processes of unit. This includes reliable operation for certain durations and when the device enters into the wear out zone. All SD activities are based on this assumption so that action can be taken before any failures occur. It is necessary to express the probability of failure for unit as a function of time for RBM interval estimation. In this work, the Weibull model with the parameters β and η is used to model the time dependent reliability of the static equipment involved in the unit. Ebeling [10] stated that the reliability of unit following the Weibull distribution is defined as:

$$Ri(t) = \left[e^{-\left(\frac{t}{\eta}\right)^{\beta_1}} \right] \times \left[e^{-\left(\frac{t}{\eta}\right)^{\beta_2}} \right] \times \left[e^{-\left(\frac{t}{\eta}\right)^{\beta_3}} \right] \quad (3)$$

$$F(t) = 1 - \prod_{i=1}^n Ri(t) \quad (4)$$

3.2 The Reliability Block Diagram (RBD)

The RBD of three pieces C-706 and C-707A/B of GLRU are connected with each other as series configuration as shown in Figure 3.

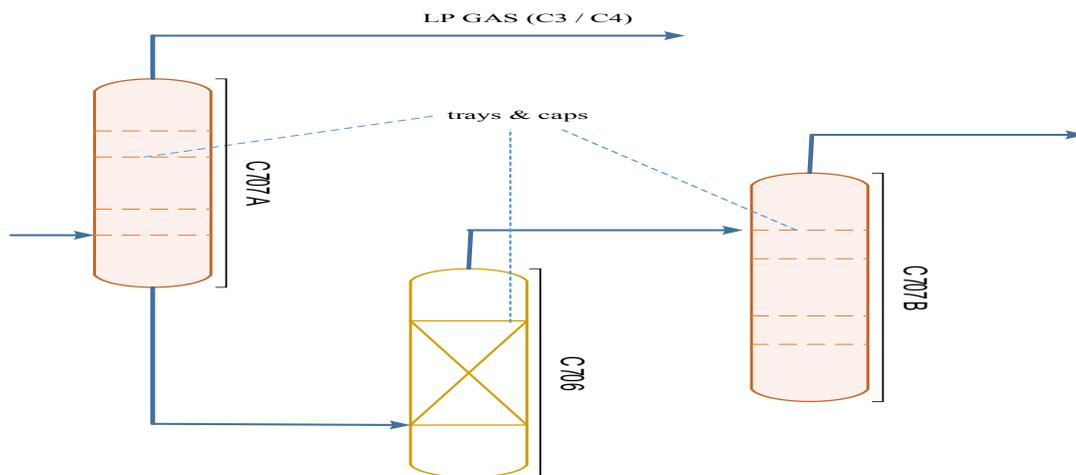


Figure 3: RBD for selected columns of Gas Liquid Recovery Unit (GLRU)

The Weibull distribution is used for modelling the SD scheduling of Splitter Column C-706A/B and two Debutanizer Column C-707 A/B columns based on shape and scale parameters as shown in Table 1.

Table 1: Shape and scale parameters of C-706 and C-707 A/B

Equipment Code	C-706	C-707A	C-707B
Shape Parameter β	5.80	5.6	5.5
Scale Parameter η (hr)	83430	95880	97335

4- Results and Discussions

Based on a tolerable risk that is adopted based on expert judgment of safety criteria of GLRU, the estimated risk of C-706 and C-707 A/B associated with melting trays and caps, and disappearing bolts due to overpressure and increase feed temperatures (fluctuating temperatures). Table 2 shows that an estimated risk is consistent with a tolerable risk at 81000 operational hour and 47.4% of F(t) according to the following constraint as shown in Figure 4:

$$\text{Estimated Risk (ER)} \leq \text{Tolerable Risk (TR)} \quad (5)$$

$$250\$/h \leq ER \leq 500\$/h$$

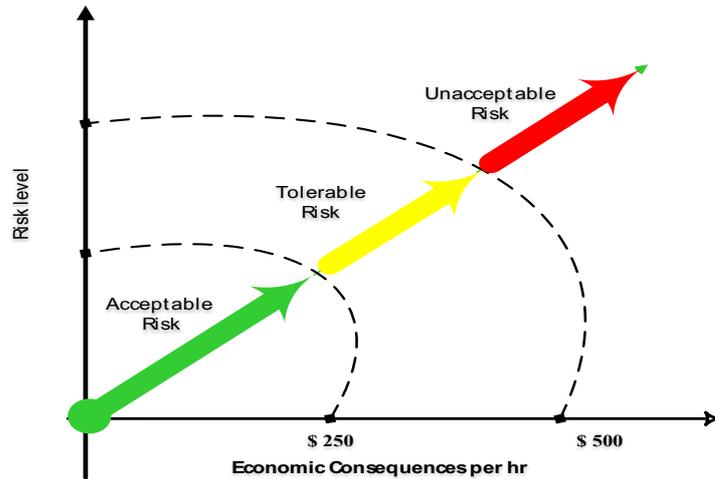


Figure 4. Tolerable risk of GLRU

$$ER = F(t) \times \sum [PL(\$) + AD(\$) + ED(\$)] \quad (6)$$

$$ER = [0.474] \times [\$85,420,000] = 500\$/hr$$

Table 2: Scenario of R(t), F(t), and ER results of C-706 and C-707A/B

SD (hr)	50000	60000	70000	75000	81000	85000	90000	95000	100000
R(t)	0.795	0.713	0.625	0.565	0.526	0.438	0.371	0.250	0.185
F(t)	0.205	0.287	0.375	0.435	0.474	0.562	0.629	0.750	0.815
ER	350.25	408.6	457.6	495.5	500	564.77	597	674.3	696

Based on processing columns for GLRU, the ER is achieved 500 \$/hr at 81000hrs. This means that ER can be acceptable with comparing to the assumed TR. Therefore, 81000 hours has become an indicator to start whole SD of plant and to execute maintenance events representing an inspection, repairs, modifications and improvement for critical equipment pieces that cannot be maintained during the normal operation of plant and to avoid an unexpected risk, which may be occurred in the plant assets, environment damage and production losses resulting from melting trays and caps, and disappearing bolts due to overpressure and increase feed temperatures. However, implementation of SD event is a feasible means to avoid these consequences, which justifies the widespread adoption of SD in real maintenance of processing columns. This means that, it is prudent to make total shutdown of the plant every nine years and three months based on the most critical processing columns.

5- Conclusions

The purpose of the work was to identify the optimum time of SD (operational period) at GLRU that worked under fluctuated temperatures and pressures. The goal was increased interval between SDs periods to increase reliability, availability and reduce production losses and maintenance cost. To address this concern RBM approach was applied as a new techniques to identify the most critical equipment in GLRU.

All of approximately sixteen pieces of processing columns were identified and distributed on the risk (5x5) matrix. Two out of seven pieces were selected as the most critical columns in the GLRU based on PoF and CoF.

The proposed RBM was generated well-maintained and well-regulated for critical equipment pieces at less hazardous agents. Therefore, this approach could be implemented in any an industrial environment run continuously under severe operational conditions, however, proper attention must be identified according to the most critical pieces of equipment which cannot be inspected or maintained during the normal operation of a plant.

References

- [1]. Elwerfalli, A., Khan, M. and Munive, J. 'Developing Turnaround Maintenance (TAM) Model to Optimise TAM scheduling for Gas Plants Based on Critical Static Equipment, International Journal of Industrial Engineering and Operations Management (IJIEOM), 2018.
- [2]. Duffuaa, S. and Ben Daya, M. "Turnaround maintenance in petrochemical industry: practices and suggested improvements", Journal of Quality in Maintenance Engineering, Vol. 10 Issue: 3, pp.184-190, 2004.
- [3]. Ghosh, D. and Rao, S. 'Maintenance optimization using probabilistic cost-benefit analysis', Journal of Loss Prevention in the Process Industries, (22), pp. 403–407, 2009.
- [4]. Megow, N., Möhring, H. and Schulz, J. 'Decision support and optimization in shutdown and turnaround scheduling', INFORMS Journal on Computing, 23(2), pp. 189-204, 2011.
- [5]. Rusin, A. and Wojaczek, A. 'Optimization of power machines maintenance intervals taking the risk into consideration'. Eksploatacja i Niezawodnosc, e-Maintenance and Reliability, 14(1), pp.72-76, 2012.
- [6]. Obiajunwa, C. 'A framework for the evaluation of turnaround maintenance projects'. Journal of Quality in Maintenance Engineering, 18(4), pp.368-383, 2012.
- [7]. Hameed, A. and Khan, F. A framework to estimate the risk-based shutdown interval for a process plant, Journal of Loss Prevention in the Process Industries, vol. 32, pp. 18-29, 2014.
- [8]. Swart, P. An Asset Investment Decision Framework to Prioritise Shutdown Maintenance Tasks, MSc thesis, University of Stellenbosch, 2015.
- [9]. American Petroleum Institute, Risk Based Inspection; 2nd ed., Washington, DC, USA: API publication 581, 2008.
- [10]. Ebeling, E. An introduction to reliability and maintainability engineering. New York, N.Y., USA: McGraw-Hill, 1997.

Frequency Response Function (FRF) Technique for the Diagnosis of Suspension System

Moamar Hamed^{1*}, Mohammed Elrawem², Feng Gu³

¹ moamar.ehmeid@gamil.com, ² mselrawemi@elmergib.edu.ly, ³ f.gui@hud.ac.uk

¹ Department of Mechanical Eng; , College of Engineering, Elmergib University, Libya

³ Centre for Efficiency and Performance Engineering, University of Huddersfield, UK

ABSTRACT

Some of the common faults associated with suspension components are damaged or leaking shock absorbers, spring weakness, wearing down of the pivot and bushing. To investigate these problems, a seven degree-of-freedom (7-DOF) model has been developed, for a full vehicle, using MATLAB. In the simulation, the suspension faults have been considered via the damage caused to the shock absorbers (dampers) and the faults were seeded by reducing the damper coefficient by 25%, 50% and 80%.

Frequency Response Function (FRF) technique was used to develop conditioned monitoring tools for suspension faults and detects the level of damping coefficients. To validate the model and evaluate the FRF technique stated in this study, experimental investigation was carried out on 4-post-test rig at the University of Huddersfield in order to measure the FRFs of the suspension for different damping setting.

The results demonstrate that, the shape of the frequency response depends on the damping coefficient (C), since low damping coefficients lead to good isolation properties of the vehicle mass in the mid- to high frequency range, but also leads to high amplitudes of the acceleration in the range of the natural frequency of the vehicle body (sprung mass). Therefore, FRFs analyses provide an effective monitoring of the suspension and to detect the level of change in damping coefficients.

Keywords:

suspension system,
Frequency response
Function (FRF),
damping coefficients,
Suspension Faults.

1- Introduction

During the last ten years, a variety of FFT-based two channel digital spectrum analysers have become commercially available which can make frequency response measurements[1]. The Frequency Response Function is a fundamental measurement that isolates the inherent dynamic properties of a mechanical structure [2]. The FRF describes the input-output relationship between two points on a structure as a function of frequency. An FRF is a measure of how much acceleration response a structure has at an output Degree Of Freedom (DOF), per unit of excitation force at an input DOF. The (FRF) is, in general, a complex valued function or waveform defined over a frequency range and includes both a real and imaginary component that can be resolved into magnitude and phase. Therefore, the process of identifying parameters from this type of measurement is commonly called curve fitting, or parameter estimation. Several experimental modal parameters such as natural frequencies, mode shapes and associated damping ratios can be extracted from a set of FRF measurements. Frequency Response Function Model can be Considered as a linear system as represented by the diagram in Figure 0 .

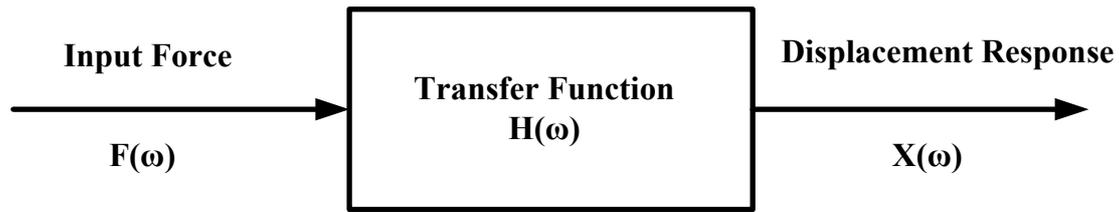


Figure 0: *Frequency Response Function Model [3]*

Where: $F(\omega)$ is the input force as a function of the angular frequency ω . $H(\omega)$ is the transfer function. $X(\omega)$ is the displacement response function. Each function is a complex function, which may also be represented in terms of magnitude and phase. Each function is thus a spectral function. There are numerous types of spectral functions. For simplicity, consider each to be a Fourier transform [3].

The relationship in Figure 0 can be represented by the following equations

$$X(\omega) = H(\omega) \cdot F(\omega) \quad (1)$$

$$H(\omega) = \frac{X(\omega)}{F(\omega)} \quad (2)$$

Similar transfer functions can be developed for the velocity and acceleration responses.

2- Frequency Response Function for single degree of freedom (SDOF) system

For simplicity a single-degree-of-freedom system subjected to a force excitation has been considered [3] as shown in Figure 2 .

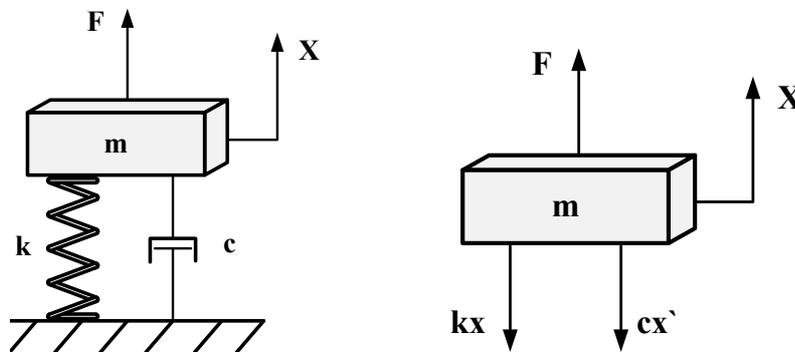


Figure 2: *Single-degree-of-freedom system and free-body diagram [3]*

Where m = mass, c = viscous damping coefficient, k = stiffness, x = absolute displacement of the mass, F = applied force.

Summation of forces in the vertical direction

$$\Sigma F = m\ddot{x} \quad (2)$$

$$m\ddot{x} + c\dot{x} + kx = F \quad (3)$$

$$\ddot{x} + (c/m)\dot{x} + (k/m)x = F/m \quad (4)$$

By convention,

$$(c/m) = 2\xi f_n \quad (5)$$

$$(k/m) = f_n^2 \quad (6)$$

Where f_n is the natural frequency in (radians/sec) and ξ is the damping ratio

Substituting the convention terms into equation (5)

$$\ddot{x} + 2\xi f_n \dot{x} + f_n^2 x = f_n^2 F/k \quad (7)$$

The Fourier transform of each side of equation (9) may be taken to derive the steady-state transfer function for the absolute response displacement, as shown in Reference 1. After many steps, the resulting transfer function is

$$\frac{X(\omega)}{F(\omega)} = \left[\frac{1}{k} \right] \left[\frac{f_n^2}{f_n^2 - f^2 + j(2\xi f f_n)} \right] \quad (8)$$

This transfer function, which represents displacement over force, is sometimes called the receptance function [4]. The transfer function can be represented in terms of magnitude and phase angle ϕ as following:

$$\frac{X(\omega)}{F(\omega)} = \left[\frac{1}{k} \right] \left[\frac{f_n^2}{\sqrt{(f_n^2 - f^2)^2 + (2\xi f f_n)^2}} \right] \quad (9)$$

$$\frac{X(\omega)}{F(\omega)} = \left[\frac{1/m}{\sqrt{(f_n^2 - f^2)^2 + (2\xi f f_n)^2}} \right] \quad (10)$$

$$\phi = \arctan \left[\frac{2\xi f f_n}{f_n^2 - f^2} \right] \quad (11)$$

The mobility function (Velocity / Force) can be considered as

$$\frac{V(\omega)}{F(\omega)} = \left[\frac{1}{k} \right] \left[\frac{j f f_n^2}{f_n^2 - f^2 + j(2\xi f f_n)} \right] \quad (12)$$

$$\left| \frac{V(\omega)}{F(\omega)} \right| = \left[\frac{I}{k} \right] \left[\frac{f f_n^2}{\sqrt{(f_n^2 - f^2)^2 + (2\xi f f_n)^2}} \right] \quad (13)$$

$$\left| \frac{V(\omega)}{F(\omega)} \right| = \left[\frac{I}{m} \right] \left[\frac{f}{\sqrt{(f_n^2 - f^2)^2 + (2\xi f f_n)^2}} \right] \quad (14)$$

$$\theta = \arctan \left[\frac{-f_n^2 + f^2}{2\xi f_n} \right] \quad (15)$$

The acceleration function (Acceleration / Force) can be presented as

$$\frac{A(\omega)}{F(\omega)} = \left[\frac{I}{k} \right] \left[\frac{-f^2 f_n^2}{f_n^2 - f^2 + j(2\xi f f_n)} \right] \quad (16)$$

$$\left| \frac{A(\omega)}{F(\omega)} \right| = \left[\frac{I}{k} \right] \left[\frac{-f^2 f_n^2}{\sqrt{(f_n^2 - f^2)^2 + (2\xi f f_n)^2}} \right] \quad (17)$$

$$\left| \frac{A(\omega)}{F(\omega)} \right| = \left[\frac{I}{m} \right] \left[\frac{-f^2}{\sqrt{(f_n^2 - f^2)^2 + (2\xi f f_n)^2}} \right] \quad (18)$$

$$\alpha = -\pi + \arctan \left[\frac{2\xi f_n}{f_n^2 - f^2} \right] \quad (19)$$

The basic assumption for single-mode approximations is that in the vicinity of a resonance, the response is due primarily to that single mode. The resonant frequency can be estimated from the frequency response data (illustrated in Figure 3) by observing the frequency at which any of the following trends occur:

- The magnitude of the frequency response is a maximum.
- The response lags the input by 90° phase.

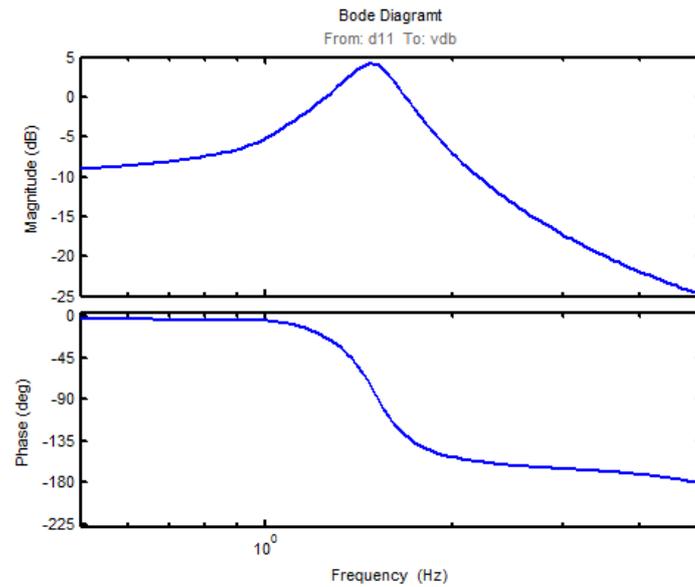


Figure 3: Bode Diagram magnitude and phase versus frequency for SDOF system

It can be noted that the height of the resonant peak is a function of damping. Damping serves to minimize or reduce the amplitude at the natural frequency and is often regarded as a measure of the frictional energy that is defined by the molecular characteristics of the structure. The level of damping on a system also has an effect on the magnitude of the amplitude at the resonant frequency of a system [5]. This is shown in figure 4. It was discussed earlier that the height of the resonant peak is a function of damping. The damping factor can be estimated by the half power method or other related mathematical or graphical method. In the half-power method, the damping is estimated by determining the sharpness of the resonant peak. It can be shown from Figure 4 that damping can be related to the width of the peak between the half-power points: points below and above the resonant peak at which the response magnitude is .7071 times the resonant magnitude.

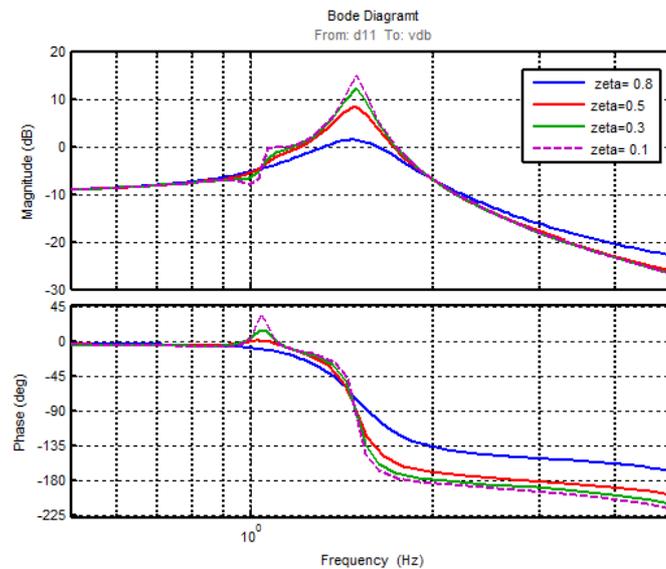


Figure 4: Bode diagram for different damping ratio (ζ)

3- Experimental investigation for FRF measurements

Early detections of abnormal events in automotive suspension systems can reduce the damage caused to the vehicle in driving situations, in addition to improving passenger comfort and security. The performance of a vehicle is often downgraded due to the appearance of faults with the

suspension. The common faults associated with suspension components are damaged or leaking shock absorbers, spring weakness, and damage to the main support member assembly. These faults can cause reduced operating efficiency and cause potential damage to the vehicle.

In order to evaluate suspension condition monitoring (CM) techniques stated in this study and also to measure the FRFs of the suspension for different damping setting, experimental work was carried out on 4-post-test rig at the University of Huddersfield. During the tests, the car is induced with the fault of damaged shock absorber, by replacing the front passive shock absorbers with adjustable shocks from SPAX Company; these shock absorbers can offer different damping force, and also more safer to perform the tests. The test was conducted to measure the vibration response of the car body and the vibration of the platform (shaker), the vibration signals were collected for baseline conditions and for different damping coefficients of suspension. For this purpose, the two front shock absorbers of the car were replaced with two adjustable shock absorbers from SPAX Company; these shock absorbers can offer different damping force.

This test consists of three main parts a commercial car, a measurement system (vibration sensors, amplifiers, data acquisition and portable laptop) and the 4-post simulator system (shaker). Figure 5 is an illustrative photograph of the 4-post-test rig and the tested car used in this study. This section comprises the experimental facilities and test procedures.

In practice, the models produced by modal testing often have poor quality due to factors inherent in the measurement. One of the sources of a lack of precision in modal testing is the errors caused by mechanical devices such as accelerometers, suspension springs and stingers.



Figure 5: the tested car (Vauxhall Zafira)

4- Experimental Facilities

4.1 The tested car specification

In this study a Vauxhall ZAFIRA car, 4 cylinder engines, with a front wheel drive and tyre model LT245/75R16E has been used, the specifications of the tested car are shown in the table 1. To evaluate the condition monitoring techniques stated in this study and also to measure the FRFs of the suspension for different damping setting, the two front shock absorbers of the car were replaced with adjustable shock absorbers (struts from SPAX Company), these shock absorbers can offer different damping forces and all adjustments can made without any dismantling of components, making it simpler, faster and easier to obtain ultimate set-up for the suspension. The SPAX shock absorbers have 28 stages of damping force adjustment. This allows a wide range of adjustment to suit different driving style and road conditions. Struts shock absorbers (adjuster at the top) are adjustable in the rebound.

To obtain the perfect set up for the car and passengers, the manufacturer recommended initially setting all the shock absorber to fully soft (anti-clockwise), and then by adjusting up in 4 click

increments, the conditions of the shock absorber will be changed from soft to hard. Figure 6 shows the SPAX shock absorber.

Table 1: *specification of the tested car*

Body style	MPV 5-Doors	Height	1634 mm
Driven Wheels	Front (FWD)	Wheel Base	2694 mm
Length	4317 mm	Weight	1448 kg
Width	1742 mm		



Figure 6: *Adjustable shock absorber from SPAX*

4.2 Measurement instrumentation

4.2.1 Accelerometers

Accelerometers are inertial measurement devices that convert mechanical motion into a voltage signal. The signal is proportional to the vibration's acceleration using the piezoelectric principle. The piezo-electric accelerometer is the most commonly used transducer for vibration measurement.

This sensor consists of a piezoelectric crystal and small mass normally enclosed in a protective metal case. When the accelerometer is subjected to vibration, the mass exerts a varying force on the piezoelectric crystal, which is directly proportional to the vibratory acceleration. The charge produced by the piezoelectric crystal is proportional to the varying vibratory force. The charge output is measured in Pico-coulombs per g (pC/g) where g is the gravitational acceleration. Some sensors have an internal charge amplifier, while others have an external charge amplifier. The charge amplifier converts the charge output of the crystal to a proportional voltage output in mV/g.

In this research, a piezoelectric accelerometer (model CA-YD-104T) is used to measure the vibration of the test car. These accelerometers are used on most vibrating surfaces because their mass does not significantly affect the movement of the surface, and possess a wide enough frequency range appropriate for measuring automobile vibration (0.5Hz to over 7 kHz). They are environmentally robust enough to withstand the conditions existing on cars and are of an adequate sensitivity. However, they do have an upper temperature limit of about 120°C due to their inbuilt microelectronics. The accelerometers are connected to an external charge amplifier YE5857A-1 providing power to the accelerometers and amplifying the charge signals. The amplified output signals of this device are sent to the data acquisition system for recording and analysis. Figure 7

shows photos of both the accelerometer and the preamplifier. A detailed specification of the accelerometer used in this work is summarized in table 2.



Figure 7: *accelerometer and amplifier*

Table 2: specification of the accelerometer

Vibration sensors	Sensor 1	Sensor 2
Model	CA-YD-104T	CA-YD-104T
Serial number	07143	07213
Sensitivity	3.243 pc/ms^{-2}	3.573 pc/ms^{-2}
Frequency range	0.5 to 7000 Hz	0.5 to 7000 Hz
Temperature	-20 to 120 °C	-20 to 120 °C

4.2.2 Data acquisition system

The Model YE7600 has 4 channels each channel has its own measurement converter A/D (analog/digital), which can be set to the sampling frequency in the range 96 kHz. Converters working in parallel are synchronized by the measure, giving the possibility of the simultaneous acquisition on these 4 channels. This type of data acquisition is adequate for monitoring vibration signal as its high sampling frequency which is sufficient for keeping the key information of original analog signal during the analog-to-digital conversion. This data acquisition is easily connected to the portable laptop by USB wire for recording and analysing the data.



Figure 8 *Data acquisition equipment (model YE6261B)*

4.3 Test procedures

Testing was carried out on the tested car. The vibration signals were collected using two accelerometers (CA-YD-104T) mounted on the car body and the base of the platform (shaker), as illustrated in Figure 5. The two identical accelerometers were used to measure FRFs. One of the main considerations regarding the accelerometers is that of the accessibility of installing them. To measure the vertical acceleration of the car body, the sensors mounted on the car include: (1) a vibration sensor mounted on the upper mounting point of the front left shock absorber where the mounting surface is flat and where the sensor can be mounted easily by gluing. This is common for many condition monitoring applications which place sensors in close proximity to vibration sources in order to acquire good vibration response characteristics. Then the same sensor 1 was moved and mounted in the same position on the upper mounting point of the front right shock absorber, figure 9 (a) shows the position of the front sensors. For measuring the vibration in the rear car body, the same sensor 1 also moved and fixed on the end of each rear corner of the car body, where the sensor can be close to the rear shock absorbers, figure 9 (b) illustrated the position of the rear sensors . Sensor 2 is represented to measure the acceleration in vertical direction, generated by the platform .The accelerometer catching on the platform was made in rigid assembly by sticking with Superglue as shown in the Figure 9 (c).



Figure 9 (a) Sensor position on front wheels (b) sensor position on rear wheels and (c) sensor position on the shaker base.

The most fundamental aim of the test was to obtain the acceleration (vibration) response of the suspension system to calculate the FRFs of the vehicle. For this purpose, multiple measurements were made when the oscillatory motion was generated by oscillating each platform separately. Determination of dynamic response has been made for the front left, the front right, the rear left and the rear right wheel respectively. The vehicle is exposed to a linear (sine sweep) frequency with the amplitude of 5 mm and a frequency range of 0.5 to 8Hz was sufficient to include the vehicle body characteristics, the frequency was increased by the rate of 0.03Hz/sec. The signals were collected and measured simultaneously and the sampling rate of data acquisition was 3000Hz. For each measurement, 765000 data points were collected and recorded in 255 seconds.

Two different experimental have been conducted, Test 1 was to measure the FRFs of the vehicle under normal operating conditions, meaning the suspension was running without any induced faults and the tyre pressure of each wheel was applied to a normal pressure of 2.2bar before starting the tests. Each test was repeated three times for each corner to ensure that the data consists of the interested characteristics of the tested vehicle.

Test 3 was performed to measure the FRFs of the vehicle for different damping coefficients, the aim of this test is to evaluate the on-road tests, for this purpose the front shock absorbers of the vehicle were replaced with adjustable shock absorbers (struts from SPAX Company). This test was conducted in two different cases, in the test (2) the damping setting of the front left shock absorber was changed in four cases as presented in Table 3. In the test (3), the damping setting of the front

right shock absorber was changed in four cases as shown in the table 4. Each test has repeated three times and after each test, the data acquired were viewed and imported to MATLAB files for further processing.

Table 3: *Experimental set up*

Test (2.1)	Damping conditions		Test (2.2)	Damping conditions	
	Front left	Front right		Front left	Front right
A	Fully soft	Fully soft	A	Fully soft	Fully soft
B	2-pteps hard	Fully soft	B	Fully soft	2-pteps hard
C	4-pteps hard	Fully soft	C	Fully soft	4-pteps hard
D	Fully hard	Fully soft	D	Fully soft	Fully hard

Fully soft that mean the shock absorber is in the baseline conditions, 2-steps hard, 4-steps hard and fully hard refer to decreases in the damping coefficients as a result of the faults in the damping system, which may induced to the system as a result of one or more of the following factors: worn seals, a reduction in the oil volume due to leakages, broken mounts and extruded or worn bushings section should extend, not repeat, the background to the article already dealt with in the Introduction and lay the foundation for further work.

5- Results and Discussion

It is well known in general that each machine will produce its own spectrum of frequencies. If the baseline vibration of a healthy machine is compared to the signal of a similar machine running under similar conditions, any increase over the baseline at any forcing frequency can indicate the presence of abnormalities in the machine. The baseline vibration signals (waveform) were recorded by the accelerometers of 765000 data points for 225 seconds. Baseline raw data of the acceleration response and FRF corresponding to the vehicle body as output sensor 1 and the front left platform as input sensor 2 are shown in Figures10. It can be noted that, in swept frequency signal, the frequency is proportional to variations of time.

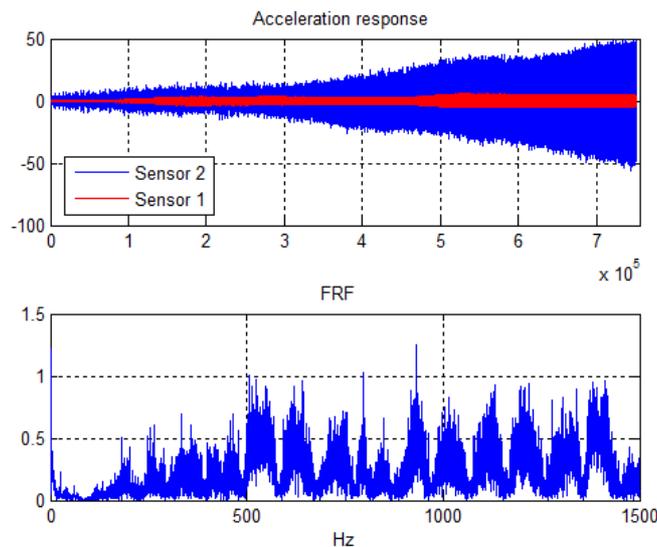


Figure 10: *Acceleration response and FRF corresponding to the vehicle body sensor 1 and the platform sensor 2*

To evaluate the proposed conditioned monitoring method for suspension faults, which has been developed in this research, to evaluate the on-road tests and to detect the level of damping

coefficients, the Frequency Response Function (FRF) technique was used. The FRF is a fundamental measurement that isolates the inherent dynamic properties of a mechanical structure and also describes the input-output relationship between two points on a structure as a function of frequency.

Figure 11 shows the amplitude-frequency characteristic curves for the baseline conditions of the suspension in the four corners of the vehicle front left (FL), front right (FR), rear left (RL) and rear right (RR), the road tests have shown some differences in the energy transmitted to the car body at each corner. In this figure, the four corners of the vehicle can show a notable difference in the vibration signals. It can be noted that, the damping was higher in the front left corner than the front right corner, and the rear corners also produce some differences. This can be identifying the reason behind the differences in the energy transmitted to the vehicle in the on-road tests, which cause of some disagreements between the simulation and the experimental results.

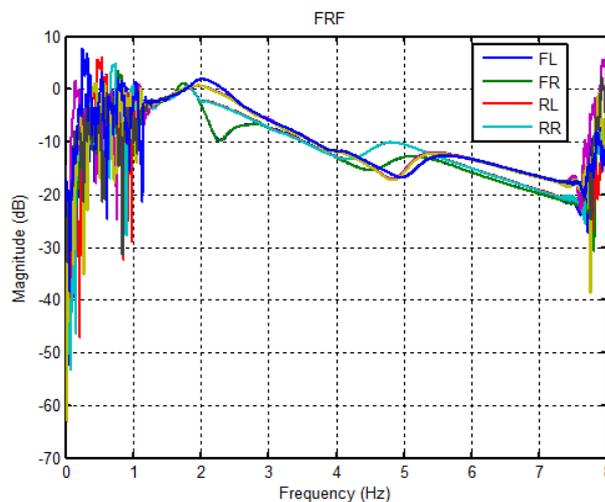


Figure 11: FRFs for the four corner of the vehicle

Figure 12 shows the acceleration ratio of the vehicle body (sprung mass) to the platform input with 4 damping conditions (fully soft, 2-steps hard, 4-steps hard and fully hard) as presented in the table 4. Theoretically, when the level of this ratio is around unity (0 dB) with no phase difference, the force between the wheel and the platform is said to be the static load and the tyre deflection is constant.

The plots reveal that increasing the damping coefficient reduces the value of the response magnitude for excitation frequencies of the vehicle body natural frequency. From this, it can be concluded that in the frequency range close to the natural frequency of the vehicle body, a high damping is required. However, lower damping is required to provide better vibration isolation in the mid- to high frequency range.

The same can be said for the figure 11, which presents the FRFs of the vehicle body and the platform for the front right corner of the vehicle. However, the differences in this side are slightly less than the differences in the left side.

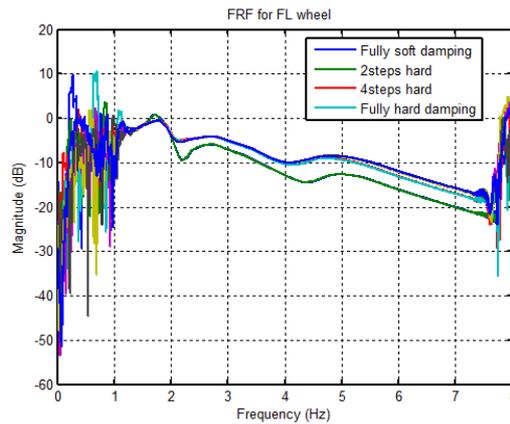


Figure 12: FRFs for the front left wheel with different damping coefficients

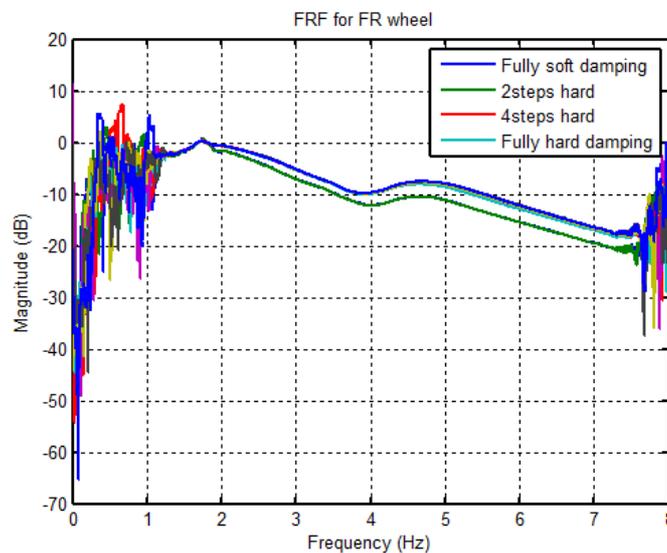


Figure 13: FRFs for the front right wheel with different damping coefficients

6- Conclusion

System being similar to those for the linear viscous damping, the isolation ability between the bounce mode and the wheel-hop mode is worse for the non-linear damping. This means the cubic damping might cause the sprung mass to have high level of response for the intermediate excitation frequencies.

One can say that with the higher force compressing the tyre, more grips will be obtained. The tyre is vibrating when either the ratio is not equal to unity or unity but with phase difference. If the tyre stiffness is too soft this will cause much vibration in the deflection of the tyre due to the undulating surface input. This can lead the wheel to vibrate against the stiffness of the tyre. This might result in poor quality of road holding, since to obtain good road holding the tyre deflection.

References

- [1]. M. H. Richardson and D. L. Formenti, "Parameter estimation from frequency response measurements using rational fraction polynomials," in Proceedings of the International Modal Analysis Conference, 1982, pp. 167–182.
- [2]. Herlufsen, "Modal Analysis using Multi-reference and Multiple-Input Multiple-Output Techniques," Brüel Kjær Appl. Note, 2004.



-
- [3]. T. Irvine, “an Introduction to Frequency Response Functions,” Rapp. Coll. Eng. Comput. Sci., 2000.
 - [4]. Packard, “The fundamentals of modal testing,” application note 243-3, 1997.
 - [5]. S. Cornelius and G. B. Paresh, “Practical Machinery Vibration Analysis and Predictive Maintenance,” 2004.

SCARA Robot Links Length Optimization by Using MATLAB and Verification with SimMechanics and Solidworks

Ph.D. Mustafa Mashali^{1*}, Miftah Addeif², Mohamed Embarak³

¹ mhmarshali@elmergib.edu.ly, ² moftaheddeif@gmail.com, ³ mohamed.embarak197@gmail.com
^{1,2,3} Department of Mechanical Engineering, College of Engineering, Elmergib University, Libya

ABSTRACT

This paper aims for optimizing links length that consumed the minimum energy, for a customized SCARA robot. Nine link length combinations are tested and simulated. This research is a part of a project of designing a robotic arm for a packing task. Kinematic and dynamic studies are performed for a 2R robotic arm. The results of kinematic study which are angular displacement, angular velocity and angular acceleration for each joint are determined and exported to the dynamic study to obtain the torque and power consumed. The dynamic study is performed with the aid of MATLAB code, MATLAB/SimMechanics and Solidworks are used to simulate and analyze the dynamic of the robotic arm. The energy consumed for each link length combination using the three methods is calculated.

Keywords:

Solidworks Motion Study,
MATLAB/SimMechanics,
SCARA Robot,
Link length optimization.

1- Introduction

In light of technological development and the increased demand for a huge amount of products of all kinds, it has become very important to increase reliance on robots in factories. This is because of their speed and accuracy that are essential in industrial sectors. Using robots can protect the laborers from multiple hazards such as toxic chemicals, high temperature and physical pollution. In addition, they lead to reduce the density of workers and maintain safe distances among the operators that can significantly mitigate the risk of spreading epidemics and infectious diseases such as the Corona pandemic (COVID-19).

This research is a part of a project that aims of designing a robotic arm for packing Tuna-cans at Al-Wafa Tuna factory at Al-Khums Libya. This paper focuses on finding the optimal links length of robotic arm for packing process. Figure 1 shows the packing unit of Tuna-cans in the factory. The packing unit consists of a table with a conveyer belt and six workers at six substations. The robotic arm is attached to the one of the substations. The designed robotic arm is SCARA (Selective Compliant Assembly Robot Arm) configuration. That is because SCARA robot is characterized by speed and accuracy that are required for assembly and packing process [1].

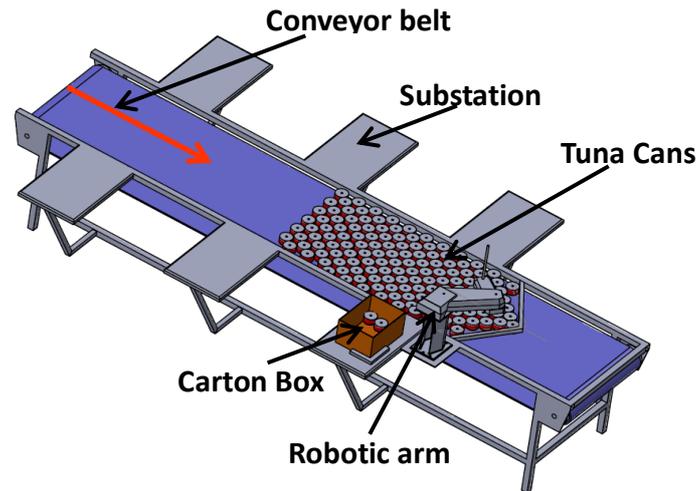


Figure 1: SCARA robot in the packing process

This arm was designed by Japanese engineer Hiroshi Makino in 1979 [1, 2]. A lot of studies has been carried out for this arm. A problem of verifying the kinematic model and the technical specifications of the four degrees of freedom SCARA robot by applying the D-H method was addressed by Fang et.al. The kinematic model was created in a MATLAB environment and simulated to monitor the movement of each joint in the robot to verify the validity of the model[3]. Alshamasin et.al presented the design and development of the four-axis SCARA robot system for drilling tasks. They worked on developing and simplifying the mathematical model. They developed it by deriving depending on the forward and inverse kinematic equations from the D-H (Denavit–Hartenber) model. The Solid Dynamic design program and the MATLAB Simulink program were used for mathematical analysis to confirm the results. The results are compatible with both programs[4]. AVCHAT et.al demonstrated the design of a SCARA three degree of freedom robot with working area dimensions of 1200 x 500mm. This arm can maintain torque of 60 N-m, at the end of the clutch. Ansys software was used in simulating the design[5]. Elaikh et.al presented the kinematic analysis by giving an angle at each joint and analyzing vibration using the Lagrange method. This is to obtain the frequency equation and analyze the effects of vibration loading on the dynamic stability of the robotic arm. The vibration analysis was for the possibility of improving the robotics system performance of the type of SCARA robots[6]. Mariappan et.al presented four-degree-freedom SCARA robot. It designed for drilling tasks and developed using Solidworks software. The CAD files were converted into a block diagram in a MATLAB/SimMechanics environment in the second generation technology. A simulation was conducted to monitor performance and dynamic variables. The results indicated that the structure variables affect the dynamic variables[7].

Fernini presented the design and development of the robot system using Solidworks and modeling using MATLAB/Simulink. Analysis was also performed for this robot to study the dynamic behavior when moving in a straight line by comparing the elbow movement up or down, which showed in the results that there is no difference in the energy exerted[8]. Gouasmi et.al presented two degrees of freedom robot design for pick and place tasks. Solidworks software was used to design and develop the robot. MATLAB/Simulink software was used to simulate and verify the results. This was done by giving an arc path for movement, and testing the movement of the elbow up and down. By reviewing the results, it was found that the movement of the elbow down requires the least torque for the robot joints[9]. Subhashini presented method to find the optimal link lengths of SCARA robot while moving in a rectangular path to remove the edges of the rectangular pieces. He also addressed the topic of finding the optimal link lengths of the robot when using the circular path by setting different lengths and analyzing the energy used in each case. The lengths that give the least energy possible were determined[10, 11].

In this paper, the problem of choosing the optimal links length for straight line trajectory is presented. This problem is for a modified SCARA robot arm. The joint configuration of the robot is (R-R-P) where R is a revolute joint and P is a prismatic joint. Solidworks motion and MATLAB/SimMechanics software is used for dynamic study. The straight-line path is generated. Torque, power and energy consumption are measured and analyzed at each joint to obtain the optimal links length in which the overall energy consumption is minimized.

1 Kinematic and Dynamic Model

As it can be seen from Figure 1, the designed robotic arm has a complicated geometry. The aim of this research is to determine the optimal length for the first and second links that can be carried out without using the actual geometry. Therefore, a simple arm of 2R is enough for performing the experiments. Figure 2 shows 3D model of the simple arm created in Solidworks.

Kinematic model gives a consideration to the motion of the arm neglecting forces that cause this motion. The kinematic model can be determined using D-H method. To define the D-H parameters for the simple arm, first, the frames have to be attached to the arm as shown in Figure 2A. Second, The D-H parameters can be determined according to the methodology explained in [12]. D-H parameters are illustrated in Table 1.

1.1 Forward Kinematic Model

The forward kinematic model calculates the end-effector position and orientation of a robot from joint angles and link lengths. Equation (1) is the general equation for determining the position and orientation of the end-effector for an arm with N joints. In this case $N = 2$.

$${}^0_N T = {}^0_1 T {}^1_2 T \dots \dots {}^{N-1}_N T \quad (1)$$

Where N is the number of joints

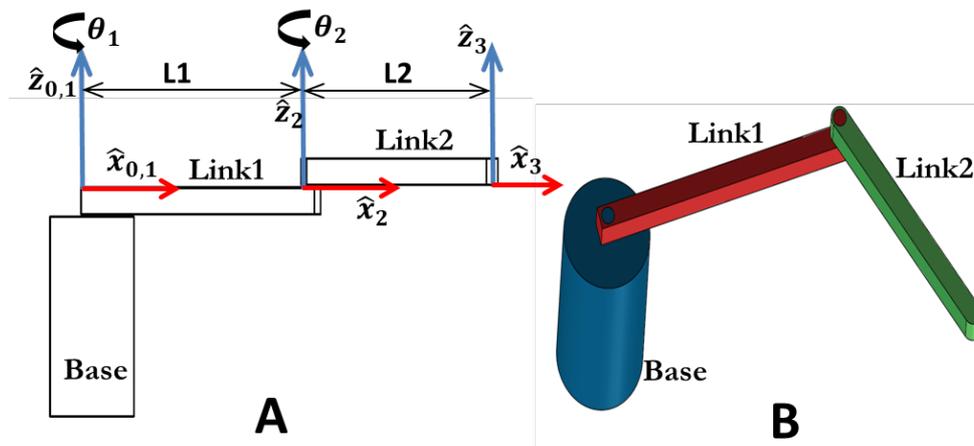


Figure 2: demonstrate 2-D and 3-D model of the RR simple arm and frame assignment.

Table 1: D-H parameters.

i	α_{i-1}	a_{i-1}	d_i	θ_i
1	0	0	0	θ_1
2	0	l_1	0	θ_2
3	0	l_2	0	0

$${}^0_1T = \begin{bmatrix} C_1 & -S_1 & 0 & 0 \\ S_1 & C_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; {}^0_2T = \begin{bmatrix} C_{12} & -S_{12} & 0 & l_1 C_1 \\ S_{12} & C_{12} & 0 & l_1 S_1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

$${}^0_3T = \begin{bmatrix} C_{12} & -S_{12} & 0 & l_1 C_1 + l_2 C_{12} \\ S_{12} & C_{12} & 0 & l_1 S_1 + l_2 S_{12} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where C_1 is $\cos(\theta_1)$, S_1 is $\sin(\theta_1)$, C_{12} is $\cos(\theta_1 + \theta_2)$, S_{12} is $\sin(\theta_1 + \theta_2)$, l_1 is the length of the first link, l_2 is the length of the second link.

1.2 Jacobian Matrix

The Jacobian matrix ($J(\theta)$) expresses the relationship between joints velocity ($\dot{\theta}$) relative to linear velocity (V) and angular velocity (ω) of end-effector. In the case of simple arm, there is no orientation so the equation will be as follow:

$$V = J(\theta)\dot{\theta} \quad (3)$$

The Jacobian matrix in simple arm case is as follow:

$$J(\theta) = \begin{bmatrix} -l_1 S_1 - l_2 S_{12} & -l_2 S_{12} & 0 \\ l_1 C_1 + l_2 C_{12} & l_2 C_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

1.3 Dynamic Model

The dynamic model studies the behaviour of moving objects and the forces acting on them.

The equations of motion are derived using Newton Euler's method as it is explained in[12]. This is to determine the torque at each joint (τ_1, τ_2).

$$\tau_1 = \left(I_1 + I_2 - l_1 S_2 - \frac{1}{2} l_2 + l_1 C_2 + \frac{m_1}{4} l_1^2 + \frac{1}{2} l_2 + \frac{l_1 C_2}{2} \right) \ddot{\theta}_1 + I_2 \ddot{\theta}_2 + (l_1 C_2 - m_2 l_1 C_2 S_2 + l_1 l_2 m_2 S_2) \dot{\theta}_1^2 + \left(\frac{l_1 l_2 m_2 S_2}{2} \right) (\dot{\theta}_1 + \dot{\theta}_2)^2 \quad (5)$$

$$\tau_2 = \left(I_2 + \frac{l_2}{2} + \frac{l_1 C_2}{2} \right) \ddot{\theta}_1 + \left(I_2 + \frac{l_2}{2} \right) \ddot{\theta}_2 + (l_1 l_2 m_2 S_2) \dot{\theta}_1^2 \quad (6)$$

$$m_i = \rho * h^2 * l_i \quad (7)$$

where I_i is the moment of inertia for the link i , m_i is the mass of the link i (kg), ρ is the density of the material (kg/m^3).The power consumed by the joint motors is determined as follow:

$$P_j = \dot{\theta}_j * \tau_j \quad (8)$$

$$E = \sum_{k=1}^{n-1} \frac{P_j(k+1) + P_j(k)}{2} * dt \quad (9)$$

where: P_j is power consumed in the joint j , $n = \frac{t}{dt}$; t is the complete period of time, dt : is the time interval, E is the energy consumption which is calculated using numerical integration (trapezoidal method) of the power versus time curve.

2 Problem formulation

To minimize the energy consumed, the robot's end-effector has to move in a straight line trajectory to pick and place Tuna cans. As shown in Figure 3, the robot moves from picking position, in which the robot arm is shown in red color, to the placing position in which the robot arm are shown in blue color. The picking position is fixed by cans' guides. This is to eliminate the cost of vision system for detecting the cans position. So in order to design the robot to do this type of tasks, you have to give a consideration in some of the robot's mechanical properties, and one of them is links length. Selecting the appropriate links length leads to the minimum energy consumption due to motors motion. This also leads to smooth motion of the end-effector.

3 Problem Solving Methodology

3.1 Data Used in Analysis

As mentioned before, the end-effector travels from the picking to the placing position in a straight line trajectory. As shown in Figure 4, the picking position is at a distance of 570 mm along X-axis and a distance of 200 mm along Y-axis. The placing positions are varying relative to the carton box. Therefore, the center of the carton box is chosen as the placing position which is at a distance of -130 mm along X-axis -minus sign indicates the opposite direction-and a distance of 600 mm along Y-axis. The travelling distance of the end-effector is 807 mm. The time in which the end-effector finishes the path is $t = 5$ seconds. The path is divided to points every $dt = 0.1$ second. Therefore, the path has 50 points equally spaced. The Cartesian velocities of the end effector (V_x, V_y) can be calculated as follows:

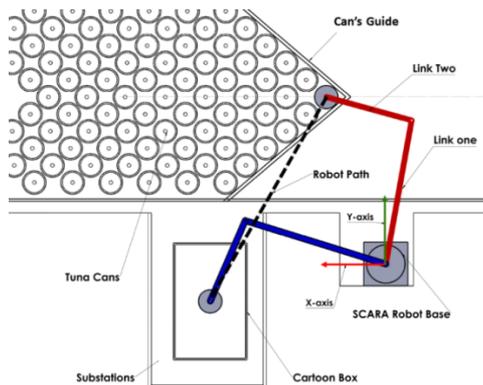


Figure 3: SCARA robot with important dimensions.

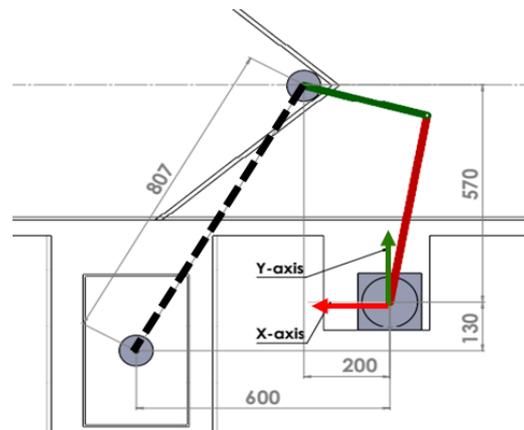


Figure 4: The robot path during the packing processes of the Tuna cans.

$$V_x = \frac{Px_i - Px_{i-1}}{dt} \quad (10)$$

$$V_y = \frac{Py_i - Py_{i-1}}{dt} \quad (11)$$

where i is the point number.

Table 2 displays the Cartesian velocities of the end-effector throughout the entire path. Using Jacobian matrix ($J(\theta)$) and Cartesian velocity (V mm/sec) the joint velocity ($\dot{\theta}$ rad/sec) can be determined using Equation (3). As the joint velocity is evaluated, the joint displacement (θ rad) and joint acceleration $\ddot{\theta}$ rad/second² can be calculated as follow:

$$\theta_i = \dot{\theta}_i * dt \tag{12}$$

$$\ddot{\theta}_i = \frac{\dot{\theta}_i - \dot{\theta}_{i-1}}{dt} \tag{13}$$

where i is the point number.

The joint displacement (θ), joint velocity ($\dot{\theta}$), and joint acceleration ($\ddot{\theta}$) are used as inputs to the torque equations

(5),(6).

Table 2: The position and velocity of the end-effector.

Point	Time t	P_x	P_y	V_x	V_y
1	0	570	200	-139.97	79.99
2	0.1	556	208	-138.26	80.24
3	0.2	542	216	-138.40	80.34
4	0.3	528	224	-138.49	80.42
5	0.4	514	232	-138.57	80.50
6	0.5	500	240	-138.64	80.59
⋮	⋮	⋮	⋮	⋮	⋮
47	4.6	-88	576	-135.72	81.96
48	4.7	-102	584	-135.72	82.1
49	4.8	-116	592	-135.73	82.26
50	4.9	-130	600	-135.76	82.43

The links in the simple arm has square cross section area with 20mm for each side. These dimensions are determined according to the maximum bending moment in the stress equation using Solidworks. Links lengths vary from (200mm to 600mm) with a step of 50 mm. The total link lengths are $l_1 + l_2 = 800\text{mm}$. There are 9 links length combinations. The total link length is chosen in the way to prevent the arm from getting into a singularity configuration or near it. If the arm is in singularity or near it, the Jacobian matrix has no inverse which makes the arm to move with a high velocity.

3.2 Dynamic Model Analysis Methods

In this paper, three methods are used to solve the dynamic model. These methods are: MATLAB program, MATLAB/SimMechanics and Solidworks motion.

3.2.1 MATLAB Program

MATLAB code is written to program the simple arm equations of motion to get joints torque in each link length combination and to get the joint displacement θ , velocity $\dot{\theta}$, and acceleration $\ddot{\theta}$. These are the inputs of the MATLAB/SimMechanics and Solidworks motion models. The steps used in MATLAB program to solve the dynamic model and visualize the results are illustrated below.

Set equation of motion inputs.

Set time;

Generate straight line trajectory.

Do for all link length combination

Defining l_1 and l_2

Get the joint velocities from Jacobian matrix and Cartesian end-effector velocities.

Get joint displacement and joint acceleration from joint velocity $(\theta, \dot{\theta}, \ddot{\theta})$;

Determine the Torque in each joint motor using torque equations (τ) ;

Determine the Power consumed in each joint using power equation (P) ;

Display torque and power data Vs time;

Determine the Energy consumed in each joint using energy equation (E) ;

End loop

Check for link length combination that consumes less Energy.

End program

3.2.2 MATLAB/SimMechanics Model

As it was mentioned before, the actual designed robotic arm is simplified to a simple 2R robotic arm. The simple robotic arm is designed in Solidworks software. In order to get the MATLAB/SimMechanics model, the CAD model is imported from Solidworks software as an (xml) file to MATLAB/ SimMechanics. This generated a block diagram and 3D model in Simulink environment as shown in Figure 5. All 9 link length combinations are designed in Solidworks and imported to MATLAB/ SimMechanics.

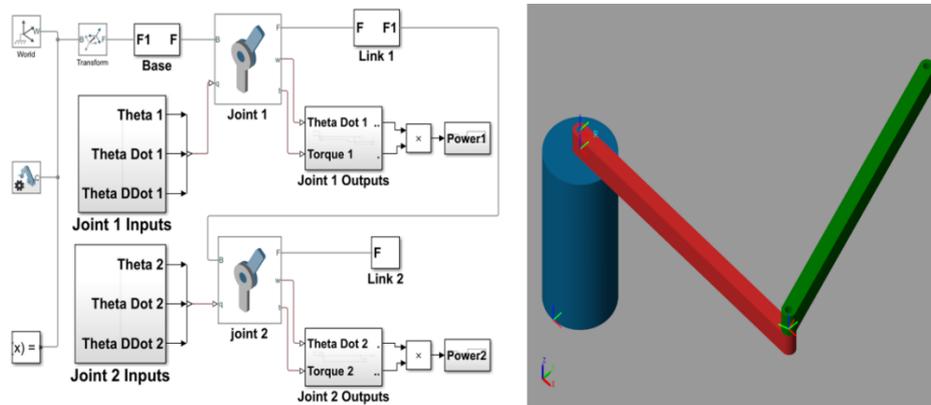


Figure 5: The MATLAB/SimMechanics block diagram and 3D model of the simple arm in the Mechanic Explorer.

3.2.3 Solidworks motion study

Solidwork motion is an add-on to the solidworks software. Solidworks motion is used to study and analyze the simple arm motion. Motion study for the 9 simple arm CAD models are carried out. These studies are generated by attaching a rotary motor for each joint. Data point file of joint displacement θ and time was imported to the joints motors. The outputs of the motion study are kinematic and dynamics results such as: Displacement, velocity, acceleration, forces, and power.

Also, it provides animation for the motion study as shown in Figure 6. The results were exported from graphs as a (CSV) file for the 9 links length combinations.

4 Results and Discussion

In this work, three methods of solving the Kinematics and dynamic models are presented: MATLAB, MATLAB/SimMechanics and Solidworks motion. These three methods were used to calculate joint velocity $\dot{\theta}$, torque τ and power consumption P for each joint. It means that, 18 different result have been analyzed for each links length combination.

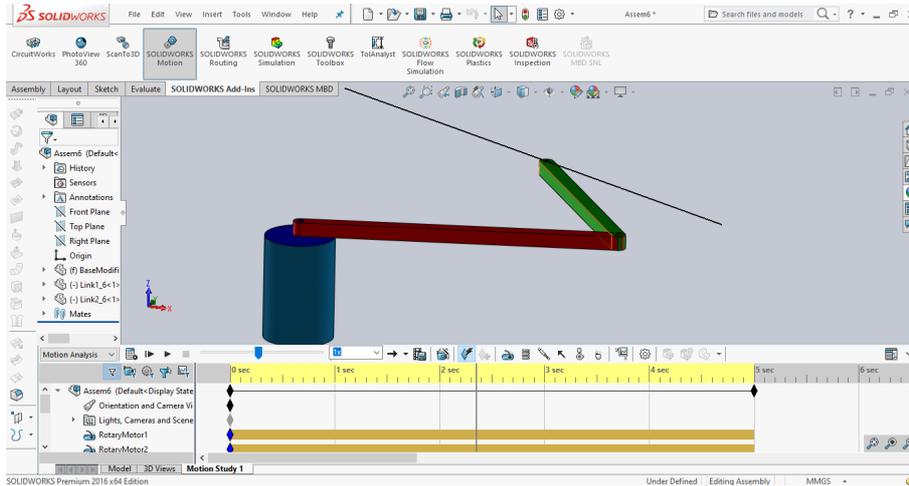


Figure 6: Simple Arm in Solidworks Motion study set.

Due to space limitation, the only sixth links length combination results will be demonstrated. Using MATLAB program, the joint angular velocity $\dot{\theta}$ was calculated for each joint. Then Joint displacement θ was determined by multiplying $\dot{\theta}$ with time. Figure 7 shows the simple robotic arm movements from the picking position at $t = 0$ second to the placing position at $t = 5$ seconds

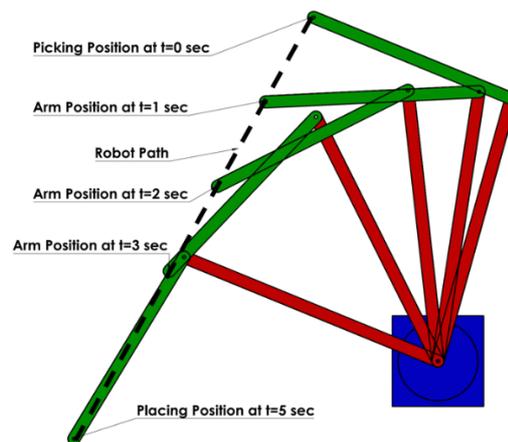


Figure 7: Time sequence of the simple robotic arm.

Figure 8 shows the joint velocity for both of the simple arm joints at the sixth links length combination. It shows that, there is a very small difference between the two joint velocities. For joint one the difference disappears as time progress. The MATLAB/SimMechanics joint angular velocity is the same as MATLAB program. Figure 9 shows the total power consumed by the two simple arm joints' motors in the sixth links length combination. The total power consumed in the two joints was computed by integrating the joint torque with the joint angular velocity as it was stated in equation (8). Figure 10 demonstrates the total power consumed in each link length

combination for whole period of time. To obtain the total power for each link length combination, the power consumed was numerically integrated for each method results. Three power consumed values corresponding to the three methods were obtained for each link length combination. It shows clearly that the sixth link length combination has the minimum power consumption. It means that the optimal links length are $l_1 = 450 \text{ mm}$ and $l_2 = 350 \text{ mm}$.

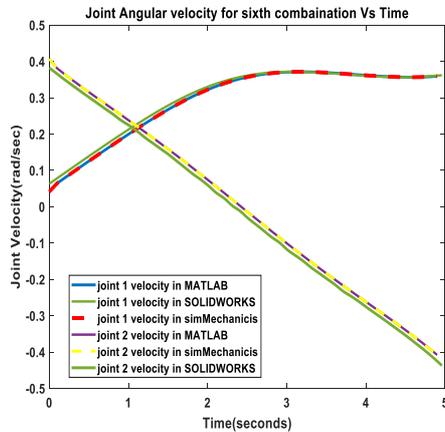


Figure 8: Joint velocity Vs Time

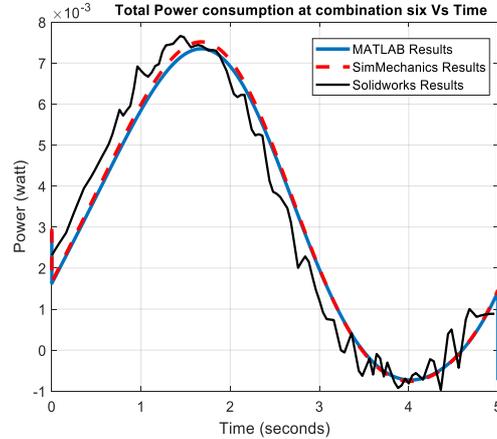


Figure 9: Power consumption by each joint for the sixth combination

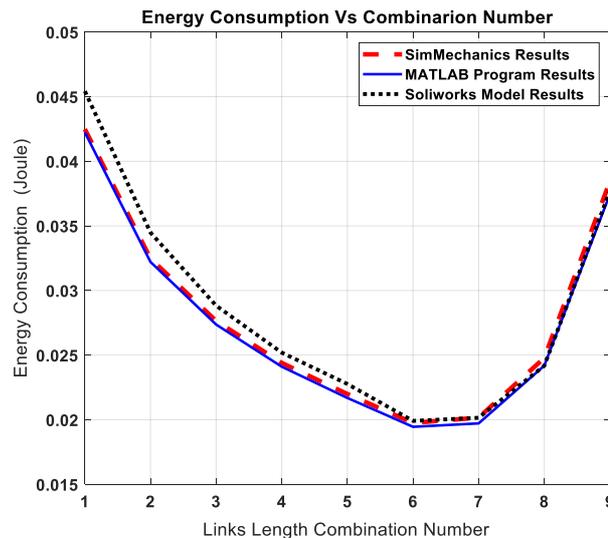


Figure 10: Energy consumption at each combination

5 Conclusion

In this paper, the optimal links length that give a minimum power consumption are determined for customized SCARA robot. Nine link length combinations are tested and simulated. This research is a part of a project of designing a robotic arm for packing task in tuna factory. The robotic arm end effector has to travel mainly in straight line for about 800 mm to pick up and place the Tuna-Cans. The kinematic and dynamic models are built for the 2R robotic arm. The results of kinematic model which are angular displacement θ and angular velocity $\dot{\theta}$ and angular acceleration $\ddot{\theta}$ for each joint are determined and exported to the dynamic model to obtain the torque and power consumed. Three method which are MATLAB program, and MATLAB/SimMechanics and Solidworks are used to simulate and analyze the dynamic of the robotic arm. The power consumed for each link length combination using the three methods is calculated. From the results, it can be concluded that the sixth link length combination has the minimum power consumption.

References

- [1] M. Shariatee, A. Akbarzadeh, A. Mousavi, and S. Alimardani, "Design of an economical SCARA robot for industrial applications," in *2014 Second RSI/ISM International Conference on Robotics and Mechatronics (ICRoM)*, 2014, pp. 534-539.
- [2] S. Y. Nof, *Handbook of industrial robotics*: John Wiley & Sons, 1999.
- [3] J. Fang and W. Li, "Four degrees of freedom SCARA robot kinematics modeling and simulation analysis," *International Journal of Computer, Consumer and Control*, vol. 2, pp. 20-27, 2013.
- [4] M. S. Alshamasin, F. Ionescu, and R. T. Al-Kasasbeh, "Kinematic modeling and simulation of a scara robot by using solid dynamics and verification by matlab/simulink," *European Journal of Scientific Research*, vol. 37, pp. 388-405, 2009.
- [5] V. AVCHAT, R. KALE, J. KODAM, and R. KOKATE, "DESIGN AND ANALYSIS OF 3-AXIS SCARA ROBOT TO SUSTAIN 60 Nm TORQUE BY NUTRUNNER," 2019.
- [6] T. E. Elaikh, H. J. Abed, K. M. Abed, and S. M. Swadi, "Vibration and kinematic analysis of scara robot structure," *Diyala Journal of Engineering Sciences*, vol. 6, pp. 127-143, 2013.
- [7] S. M. Mariappan and A. Veerabathiran, "Modelling and simulation of multi spindle drilling redundant SCARA robot using SolidWorks and MATLAB/SimMechanics," *Revista Facultad de Ingeniería Universidad de Antioquia*, pp. 63-72, 2016.
- [8] B. Fernini, "<Dynamic_Behavior_of_a_SCARA_Robot_by_using_N-E_met.pdf>."
- [9] M. Gouasmi, M. Ouali, B. Fernini, and M. h. Meghatria, "Kinematic modelling and simulation of a 2-R robot using solidworks and verification by MATLAB/Simulink," *International Journal of Advanced Robotic Systems*, vol. 9, p. 245, 2012.
- [10] P. Subhashini and K. Chinmayi, "Parametric Optimization of Link Lengths of A SCARA Robot for Deburring Of Rectangular Paths," *Materials Today: Proceedings*, vol. 5, pp. 11498-11506, 2018.
- [11] P. Subhashini, N. Raju, and G. V. Rao, "Parametric optimization of link lengths of a SCARA robot for deburring of circular paths," in *CAD/CAM, Robotics and Factories of the Future*, ed: Springer, 2016, pp. 127-135.
- [12] J. Craig, "Introduction to Robotics," 2018.

تخفيض الاحمال الكهربائية باستبدال السخانات الكهربائية بأنظمة تسخين المياه بالطاقة الشمسية

*د. عياد المخزوم محمد الواعر¹، أ. علي المبروك النانض²
1قسم تقنيات الطاقات المتجددة، المعهد العالي للعلوم والتقنية، ترونة، ليبيا
2قسم الهندسة الكهربائية، المعهد العالي للعلوم والتقنية، ترونة، ليبيا
*ayss732001@gmail.com, alimbruok1975@gmail.com

ABSTRACT

تواجه ليبيا في الوقت الراهن طلباً متزايداً على الطاقة فقد بلغ الحمل الأقصى للشبكة العامة إلى 6000 ميجاوات وتعتمد فيه محطات توليد الطاقة الكهربائية اعتماداً كلياً على استخدام أنواع الوقود المختلفة المشتقة من النفط في حين تمتلك أهم المصادر الطبيعية والتمثلة في شدة الإشعاع الشمسي بساعات سطوح مرتفعة طيلة أيام السنة والتي يمكن استغلالها كمصدر بديل للطاقة وذلك بتوظيف التقنيات المتاحة والتي تقوم بتحويل ضوء الشمس إلى طاقة كهربائية أو حرارية لسد العجز الحاصل في الطلب على الطاقة والمساعدة في تخفيض الاحمال الكهربائية التي تعاني منها كافة شرائح المستهلكين.

Keywords:

يهدف هذا البحث إلى إدخال بعض أنواع أنظمة تسخين المياه بالطاقة الشمسية في القطاع المنزلي، من أجل تقليل الحمل على الشبكة العامة للكهرباء، ومدى جدوى استخدام هذه الأنظمة في التطبيقات المنزلية عن طريق تسخين المياه بواسطة المجمعات الشمسية المتمثلة في مجمعات الألواح الشمسية المسطحة ومجمعات الأنابيب الزجاجية المفرغة بدلاً عن السخانات الكهربائية والتي تُعد من أكبر الأجهزة المنزلية استهلاكاً للطاقة. وتعرض هذه الورقة أيضاً نبذة للتعريف بهذه التقنيات وكيفية تحويل الطاقة ومميزاتها وطرق اختيار الأنسب منها والقوانين التي تحدد الكفاءة والأداء الحراري والمعادلات التي تحسب الكميات التي توفرها هذه الأجهزة من طاقة حرارية يمكن توظيفها والاعتماد عليها من ضمن مشروعات ترشيد استهلاك الطاقة واستغلالها بشكل أمثل على الصعيد المحلي للمساعدة في حل أزمة الطلب المتزايد على الكهرباء وإفاء الشركة العامة للكهرباء بتزويد كافة المستهلكين على مدار الساعة باحتياجاتهم من الطاقة الكهربائية. حيث أظهرت النتائج حسابياً أن الاستغناء عن 400000 سخان كهربائي يمكن أن يوفر 720 ميجاوات وهو ما يعادل توليد محطة كاملة. وبينت الدراسة حسابياً احتياجات الأسرة من الماء الساخن يومياً وسنوياً ومقدار كميات الطاقة اللازمة لتسخين هذه المياه عند استخدام السخان الكهربائي وكذلك بعد استبداله بمنظومة المجمع الشمسي. وتوضح أنه بعد حساب تكلفة إجمالي الاستهلاك الكهربائي للسخان الكهربائي وفقاً لسعر تعريف الاستهلاك (المنخفضة) للكيلو وات ساعة هو 334.88 دينار ليبي/سنة فإنه يمكن تحديد فترة استرداد ثمن المجمع الشمسي (**Pay-back period**) في حال الاعتماد عليه واستبداله بالسخان الكهربائي خلال فترة لا تتعدى العامين.

1. المقدمة

تعتبر مصادر الطاقة الشمسية في ليبيا هي الأعلى من بين دول العالم، لكونها تقع ضمن منطقة الحزام الشمسي، والتي تتمتع بشدة سطوح شمسي خلال أغلب أيام السنة والتي تتراوح فيها قيمة شدة الإشعاع الشمسي من 1900 كيلو وات ساعة/متر مربع/سنة في المناطق الساحلية وترتفع هذه القيمة لتصل إلى أكثر من 2800 كيلو وات ساعة/متر مربع/سنة في المناطق الداخلية الصحراوية، وكما هو معروف أن تكلفة وحدة إنتاج الطاقة تنخفض كلما ارتفعت قيمة شدة الإشعاع الشمسي الساقط على الأرض [1].

بسبب ارتفاع أسعار الوقود والكهرباء والتلوث البيئي الكبير الناتج عن حرق الوقود أصبح من الضروري البحث عن مصادر جديدة للطاقة تكون اقتصادية وذات أسعار معقولة ولا تؤدي إلى إلحاق الضرر بالبيئة، فكان السخان الشمسي بمختلف تصاميمه التقنية هو الحل المناسب لجزء كبير من هذه المشكلة ونتيجة لذلك انتشرت أجهزة تسخين المياه بالطاقة الشمسية بشكل واسع في الأسواق حيث لاقت أجهزة تسخين المياه بالطاقة الشمسية انتشاراً واسع النطاق وإقبالاً من المواطنين خاصة في الدول الغير منتجة للنفط لتكون إحدى استراتيجيات الطاقة الوطنية للاعتماد عليها بشكل ملحوظ لحل أزمة الطاقة [2]. فمن المهم إعطاء ولو فكرة مفصلة عن أهم أنواع أجهزة تسخين المياه بالطاقة الشمسية المتوفرة في الأسواق والتنويه عن الاختلاف فيما بينها والطرق المختلفة لكيفية عملها. في هذه الدراسة سنتعرف على الأنواع المختلفة لأجهزة تسخين المياه بالطاقة الشمسية شائعة الاستخدام والتي تتميز بسهولة التركيب والصيانة.

إن القيمة اليومية للحمل الأقصى للشبكة العامة قد تصل إلى 6000 ميجاوات ومع هذه القيمة المبالغ في استهلاكها في أغلب الأوقات مازالت قطاعات الدولة والمستهلكين الكبار منهم والصغار يعانون من انقطاع الطاقة الكهربائية التي وصلت في بعض

الاحيان إلى عدة ساعات في اليوم، في الوقت الذي يزداد فيه الطلب على الطاقة الكهربائية نتيجة للنمو السكاني يزداد الاستهلاك الغير مقنن من قبل شريحة المستهلكين بشكل عام نتيجة للتوسع السكاني العشوائي والزيادة في الامدادات الكهربائية الغير قانونية ولهذا فإن شركة الكهرباء لم تعد تستطيع الايفاء بتزويد المستهلكين على مدار الساعة بالطاقة الكهربائية المطلوبة على مدار العام وخاصة في فصلي الشتاء والصيف وهي الاوقات التي تزداد فيها معدلات الاستهلاك في الاستخدام المبالغ فيه والمتعلق بأجهزة التسخين والتدفئة والتبريد والغير خاضعة لأبسط مشروعات ترشيد استهلاك الطاقة في المجال الكهربائي.

إن تسخين المياه باستخدام السخانات الكهربائية تعتبر من التقنيات عالية الكفاءة ولكنها تعد من الأجهزة التي تأخذ معدلات استهلاك عالية من الإنتاج الكلي للطاقة الكهربائية وان عدم وعي المستهلك وقلة اهتمامه بإرشادات شركة الكهرباء في كيفية التشغيل الأمثل للسخانات الكهربائية بما يكفل الحد من الاحمال الكهربائية المفروضة يزيد من تفاقم مشكلة العجز في الإيفاء بإمدادات الطاقة الكهربائية، لذلك جاءت الحاجة للتركيز والاهتمام على امكانية استخدام نظم تسخين المياه بالطاقة الشمسية المتاحة في الاسواق والمناسبة للاستخدام المحلي ودورها في الحد من الانتاج الزائد والاستهلاك المفرط للطاقة الكهربائية، الامر الذي جعلنا نفكر في تقديم هذه الدراسة وهي استبدال هذا النوع من السخانات الكهربائية بغيرها من الأنظمة والتي تعتمد على الشمس كمصدر للطاقة.

2. منهجية البحث

عرض منظومات الطاقة الشمسية الحرارية والمتمثلة في نوعين منها (مجمعات الألواح المسطحة ومجمعات الانابيب المفرغة) وتوجيه وإرشاد شرائح المستهلكين الكبار والصغار اللذين يعتمدون على نظم توليد الطاقة الحرارية مثل السخانات الكهربائية، والنظر في كيفية استخدام تقنيات الطاقة الشمسية وبالتحديد المجمعات الشمسية الحرارية شائعة الاستعمال في الاسواق المحلية كي تنوب او تعمل بدلاً من السخانات الكهربائية التي تستهلك كما هائلاً من الطاقة الكهربائية، ايضاً تحديد نوع ومواصفات منظوماتها لاستخدامها الاستخدام الأمثل وصولاً لتخفيف الاحمال الكهربائية ترشيداً لاستهلاك الطاقة الكهربائية.

الطريقة العلمية والتي تم اختيارها في هذه الورقة هي إمكانية إثبات الجدوى الاقتصادية وذلك بمقارنة المنظومات التي تعتمد على الشمس كمصدر للطاقة مع غيرها من المنظومات الأخرى وتخصص دراستنا هذه على منظومة السخان الكهربائي التقليدي وستكون المقارنة من حيث الكفاءة وتوفير كميات الطاقة المطلوبة وحاجة المستهلك وعلى الرغم من أن الطاقة الشمسية هي طاقة مجانية إلا ان تكلفة منظوماتها والتي تحول الطاقة الشمسية إلى طاقة حرارية تكون باهضة الثمن، فسيتم تحديد الحسابات الخاصة بالطاقة المطلوبة لتسخين المياه بواسطة مجمعات الطاقة الشمسية وكذلك الحسابات التي تحدد استهلاك السخان الكهربائي وتكاليف الاستهلاك بوحدة (ك. و. س) اعتماداً على التعريف المحددة من قبل الشركة العامة للكهرباء ايضاً سيتم حساب الفترة الزمنية لإعادة المبالغ المصروفة (**Pay-back Period**) المتعلقة بالسخان الشمسي.

3. تحويل الطاقة الشمسية وتخزينها

التحويل الحراري للطاقة الشمسية يعتمد اساساً على تحويل الإشعاع الشمسي المباشر إلى طاقة حرارية وذلك باستخدام مختلف تقنيات المجمعات الشمسية والمواد الحرارية المستعملة في هذا المجال. فإذا تعرض جسم أسود (داكن) معزول إلى الإشعاع الشمسي المباشر فإنه يمتصه وبالتالي يمكن الاستفادة من هذه الحرارة في نقلها وتخزينها لاستخدامها في التدفئة والتسخين. إن تقنيات المجمعات الشمسية هي الأكثر شيوعاً في مجال تحويل الإشعاع الشمسية إلى طاقة حرارية، فعلى الرغم من أن الطاقة الشمسية هي الأكثر استثماراً مقارنة بالأنواع الأخرى المختلفة من الطاقات المتجددة، إلا أنها تظل مرهونة بتوفير كمية وكثافة عالية من الأشعة الشمسية الساقطة طيلة فترة السطوح الشمسي لذلك فمن الضروري التركيز على تطوير تقنيات التخزين الحراري المصاحبة لتقنيات التحويل الحراري للطاقة الشمسية ورفع كفاءتها لتعظيم الاستفادة منها أثناء فترات غروب الشمس [3].

ومن الجدير بالذكر ان معظم المناطق في العالم غنية بالإشعاع الشمسي ذو الكثافة العالية مما يتيح لها فرصة أكبر لاستخدام تقنيات المجمعات الشمسية بمختلف انواعها وهو من ضمن الأسباب الذي يدفعنا لاستبدالها بالسخانات الكهربائية التي تستهلك مقداراً كبيراً من الطاقة الكهربائية وذلك لأجل حل مشكلة أزمة الطاقة على الصعيد المحلي.

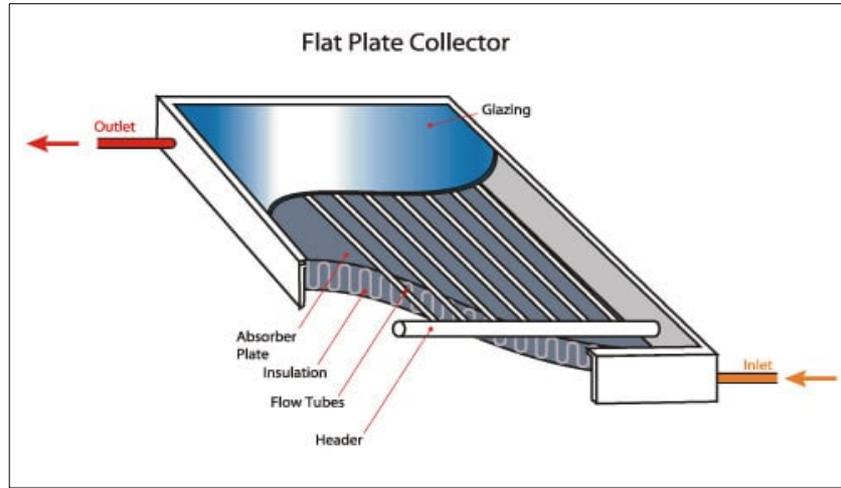
4. أهم أنواع المجمعات الشمسية المستعملة محلياً

تعد السخانات الشمسية من أهم الوسائل الحديثة والنظيفة لاستغلال الطاقة الشمسية لأغراض تسخين المياه في المنازل والتجمعات السكنية، وقد شهدت تطوراً كبيراً في السنوات القليلة الماضية. كما أنها تعتبر طريقة فعالة ورخيصة الثمن لتزويد المنازل بالمياه الساخنة الضرورية للاستعمالات اليومية، حيث تستطيع العمل في مختلف الطقوس والعوامل الجوية، وتسخن المياه بسرعة وكفاءة عالية حتى في الأيام الباردة وقودها هو أشعة الشمس المجانية وهنا تكمن اقتصاداتها. في هذه الورقة سيتم التركيز وعرض نوعين من هذه التقنيات الأكثر شيوعاً واستعمالاً [5]، والمتمثلة في مجمعات الألواح الشمسية المسطحة ومجمعات الأنابيب الزجاجية المفرغة.

1.4 مجمع اللوح الشمسي المسطح

مجمع اللوح الشمسي المسطح كما هو موضح في الشكل رقم (1) هو وحدة امتصاص وتسخين تقوم بتحويل الطاقة الحرارية الشمسية المتمثلة في الأشعة الساقطة عليه إلى حرارة عبر المياه التي تتحرك داخل الانابيب النحاسية بمنظومة الامتصاص

المعزولة لتجنب الفقد الحراري قدر الامكان، ترتفع درجة حرارة المياه داخل هذه الانابيب مع مرور الزمن ويتم تخزينها في سخانات ذوات ساعات مختلفة تكون معزولة بمواد عزل عالية الكفاءة للحفاظ على حرارة المياه ليتم استخدامها في اوقات مختلفة [5].



شكل رقم (1): المجمع الشمس المسطح

يصنع السطح الماص من مواد معدنية لها امتصاصية عالية تكون مطلية بمادة انتقائية داكنة لرفع معدل الامتصاصية، لأن المواد الداكنة لها خاصية امتصاص أشعة الشمس بمعدلات عالية من الأشعة الشمسية، ولكن مازالت تعاني من عيوب ضياع الحرارة منها بواسطة الاشعاع حيث يصل معدل الفقد الحراري الى 89% تقريباً، أي أن السطوح الممتصة للحرارة تكون فاقدة لها عن طريق الاشعاع بشكل ملحوظ وبالتالي يكون معدل الطاقة المستفاد قليل، فالسطوح الماصة غالباً ما تصنع من الألمنيوم او النحاس او الحديد المجلفن ويفضل عادة الألمنيوم او الحديد المجلفن لأنها رخيصة الثمن، حيث يُطلى السطح الماص للحرارة بمواد انتقائية عالية الامتصاصية ومنخفضة الإشعاعية مثل أكاسيد الكروم والكوبالت. من ضمن المكونات ايضاً الانابيب المعدنية والتي تمتلك معدلات امتصاص عالية لنقل الحرارة التي يتم امتصاصها بواسطة الانابيب الى السائل، ولذا يجب ان تكون هذه الانابيب ذات موصلية حرارية مرتفعة حتى تنقل كل الحرارة المستفاد، وتصنع هذه الانابيب من معادن مختلفة ويفضل استخدام النحاس لموصلته العالية. من مكونات المجمع الشمسي المسطح ايضاً الغطاء الزجاجي والذي تكون وظيفته السماح للأشعة الشمسية بالوصول الى الانابيب الحرارية الماصة، وتقليل الفقد الحراري من خلال القسم العلوي كما انه يمنع الهواء والماء من التسرب الى داخل المجمع، ومن أهم الانواع التي يمكن استعمالها انواع الزجاج ذو النفاذية العالية.

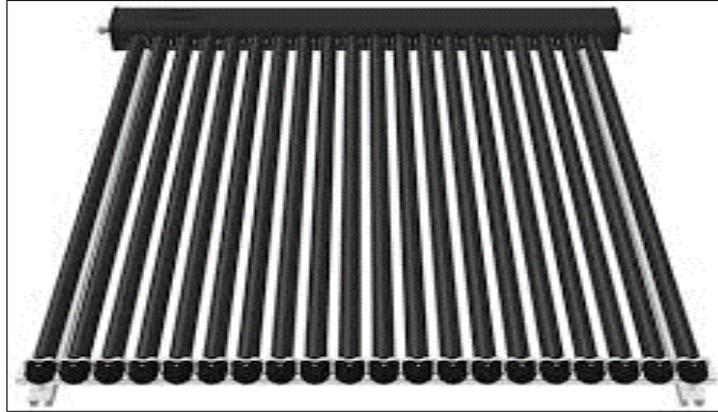
ان ارتفاع درجة الحرارة داخل السخانات بالمقارنة مع درجة حرارة الجو المحيط يزيد من معدل الفقد الحراري بالتوصيل وذلك عن طريق جسم السخان، وكذلك الفقد الحراري من الغطاء الزجاجي بالحمل والاشعاع وبالتالي يوصي استعمال مواد العزال الاقتصادية التي تمتلك خاصية الحد من ضياع الحرارة [5].

2.4 مجمع الأنابيب الزجاجية المفرغة

ظهرت هذه التقنية منذ فترة زمنية قريبة، وهي تعتمد على ما يعرف باسم الأنابيب الزجاجية المفرغة كما هو موضح في الشكل رقم (2) حيث أن هذا النوع من التقنية يمتص الأشعة الشمسية بمعدلات كفاءة عالية وتحولها إلى طاقة حرارية يتم استغلالها في عمليات تسخين المياه. يتألف كل أنبوب من هذه الأنابيب المفرغة من أنبوبين من الزجاج، متداخلين مع بعضهما البعض ويصنعان من زجاج البوروسيليكيت الذي يتميز بالمتانة ومقاومته للكسر، يكون الأنبوب الخارجي شفافاً حيث يسمح لأشعة الشمس الساقطة بالنفاذ من خلاله بقليل من الانعكاس بينما يُطلى الأنبوب الداخلي بطبقة سوداء عادة ما تكون من النيكل والكروم فيعمل على امتصاص أشعة الشمس الساقطة عليه بنسبة عالية تصل إلى 98% تقريباً [6].

يتم تسخين الأنبوبين المتداخلين إلى درجات حرارة عالية لتفريغ الهواء الموجود بينهما، ليتم بعد ذلك فوراً تثبيت نهايتي الأنبوبين عن طريق عملية الصهر لينتج من عملية التفريغ هذه تكون منطقة عزل بين الأنبوبين وهذا ما يمنح هذه الأنابيب كفاءة عالية لامتناس أشعة الشمس، ويعمل الفراغ المتكون بين الأنبوبين بمنع تسرب الحرارة التي اكتسبتها المياه المارة خلال الأنبوب الداخلي، ويعمل هذا الفراغ ايضاً على إيقاف الفقد الحراري من عمليتي التوصيل والحمل الحراريين، ونتيجة لذلك تتجاوز درجة حرارة الأنبوب الداخلي 150 درجة مئوية ويظل الأنبوب الخارجي بارداً. إن عملية التفريغ تمتلك خواص عزل جيدة تسمح بوجود فرق عالي في درجات الحرارة بين الأنبوبين الداخلي والخارجي، وهذا ما يجعل الأنابيب المفرغة تعمل بمقدار عالي من الكفاءة حتى في المناخ البارد بينما تكون كفاءة المجمعات الشمسية المسطحة في مثل هذه الأجواء منخفضة بسبب فقدانها للطاقة الحرارية [7]. من جهة اخرى هناك نوع من مجمعات الأنابيب الزجاجية المفرغة التي تستعمل بداخلها الأنبوب الحراري المعدني في نقل الحرارة فهو أنبوب مجوف من الداخل ومفرغ من الهواء تماماً وتوجد به كمية

صغيرة من سائل العمل يستقر في الجزء السفلي منه والذي يتبخر عند تلقيه الحرارة من الشمس عبر الانبوب الزجاجي المفرغ فيرتفع البخار حتى يصل إلى رأس الانبوب (المكثف) المثبت في المشعب الذي ينقل الحرارة من البخار إلى سخان فيبرد البخار ويتكثف ويسيل ويعود مرة أخرى إلى أسفل أنبوب الحرارة بواسطة الجاذبية وهكذا تتكرر العملية [6،7،8].



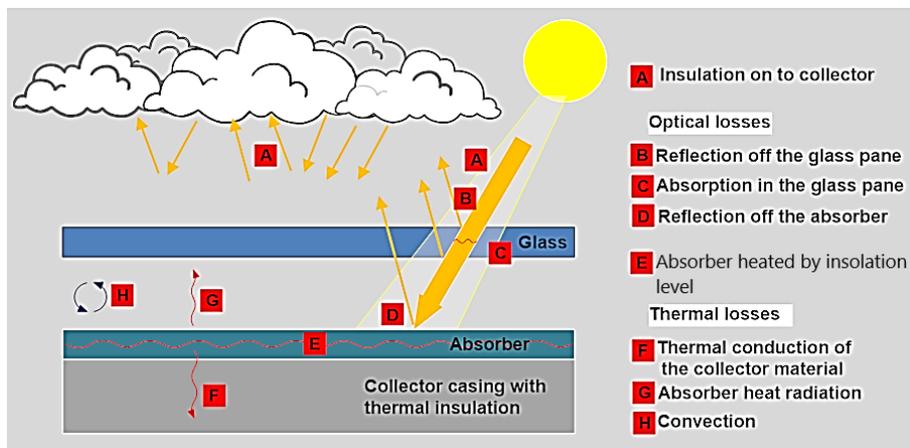
شكل 2 شكل رقم (2): مجمع الانابيب الحرارية المفرغة

3.4 مساحة المجمعات الشمسية

في المجمعات الشمسية المسطحة ينبغي أن تتناسب سعة السخان الذي يحوي الماء المراد تسخينه مع مساحة المجمع الشمسي، فعلى سبيل المثال في الدول العربية نحتاج إلى مجمع شمسي مساحته تتراوح من 1.2 متر مربع إلى 1.5 متر مربع لكل سخان ذو سعة 100 لتر. وتزيد مساحة هذا المجمع في مناطق أوروبا الشمالية، حيث نحتاج إلى 4 متر مربع لكل سخان ذو سعة الـ 100 لتر، حيث توفر هذه المساحة في حدود 90% مما يتم طلبه من المياه الساخنة في فصل الصيف و60% تقريباً في فصل الشتاء. فعلى سبيل المثال عند اقتناء هذا النوع من المجمعات الشمسية فإن سخان المياه ذو سعة 300 لتر يجب ان ألا تنقص مساحته عن 3.6 متر مربع عند تركيبه في إحدى دول الجزيرة العربية أو 4.5 متر مربع عندما يتم تركيبه بدول المغرب العربي. ومن الجدير بالذكر انه عند اقتناء هذه الأنواع من الأنظمة يجب التركيز ان تكون مساحة المجمع متناسبة مع سعة السخان وذلك للوصول إلى كفاءة وجودة التسخين المطلوبة. وفيما يتعلق بتقنية سخانات ذات الأنابيب المفرغة فإن سخان المياه ذو سعة 100 لتر يحتاج إلى مساحة المجمع الذي يحتوي على 10 انابيب، والسخان ذو سعة 300 لتر يحتاج إلى مساحة المجمع الذي يحتوي على 30 أنبوب مفرغ. وكما هو الحال في تقنية سخانات الألواح المسطحة يجب التركيز على عدد انابيب الذي يحتويه المجمع الشمسي وما يناسبه من السعات المختلفة لسخانات المياه عند أخذ القرار لاستعمال هذا النوع من التقنيات [5].

4.4 تحديد كفاءة المجمعات الحرارية الشمسية.

تعتمد كفاءة المجمع الشمسي على قدرته على امتصاص الحرارة والممانعة عن فقد الحرارة بمجرد امتصاصها. حيث يوضح الشكل رقم (3) مبادئ تدفقات الطاقة في المجمع الشمسي.



شكل رقم (3): مبادئ تدفقات الطاقة في المجمع الشمسي [9]

لحساب كفاءة المجمع الشمسي يمكن استخدام المعادلة رقم (1) والبارامترات الموجودة في دليل البيانات الخاصة بكل مجمع [9].

$$\eta_c = \eta_0 - a_1 \frac{(T_m - T_a)}{G} - a_2 \frac{(T_m - T_a)^2}{G} \dots \dots \dots (1)$$

حيث:

η_0 : أقصى قدر من الكفاءة إذا لم يكن هناك فقدان للحرارة [%]

a_1 : معامل فقدان الحرارة الأول [W/ (°C·m²)]

a_2 : معامل فقدان الحرارة الثاني [W/ (°C·m²)]

يجب تحديد هذه البارامترات وفقاً للمعيار المناسب ومن ثم توفير المعلومات الأساسية لتحديد الكفاءة.

η_c : كفاءة المجمع [%]

G: إجمالي الإشعاع الساقط على سطح المجمع [W/m²]

T_m : متوسط درجة حرارة سائل المجمع [°C]

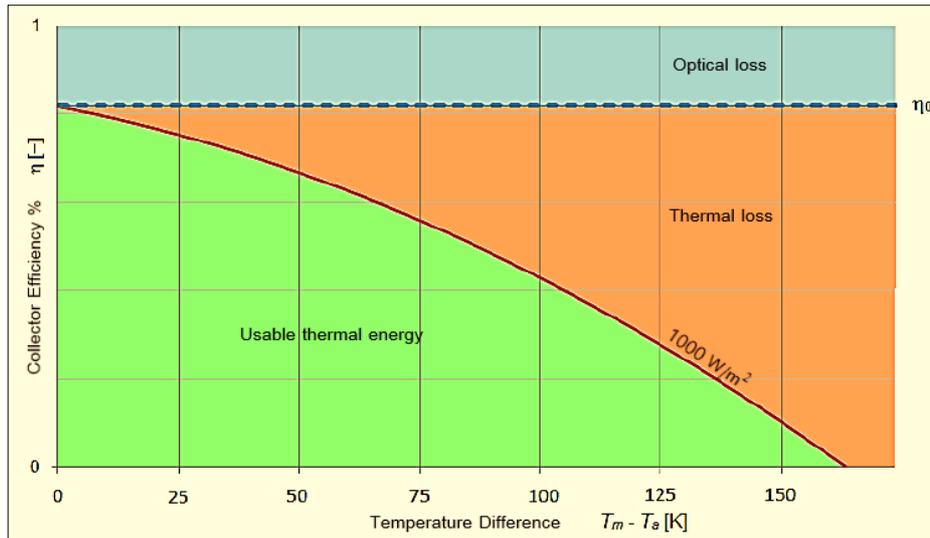
T_a : درجة حرارة الهواء المحيط [°C]

الشكل رقم (4) مثال للتعبير عن الكفاءة من المعادلة رقم (2) حيث (الخط الأحمر) يوضح الخسائر والطاقة المفيدة لمستوى إشعاع معين (على سبيل المثال 1000 وات/متر مربع). يتم تمثيل الكفاءة كدالة للفرق في درجات الحرارة بين متوسط درجة حرارة المجمع T_m ودرجة الحرارة المحيطة T_a . (اللون الأزرق) يشير إلى النسبة بين الخسائر البصرية (**Optical Loss**) و(اللون البرتقالي) يمثل فقدان الحرارة (**Thermal Loss**)، بينما (اللون الأخضر) هي الطاقة المفيدة (**Usable Loss**)، والتي يمكن مقارنتها بالمجموع الكلي للطاقة من الإشعاع (يساوي 100٪ على المحور الثانوي)، ويمكن حساب خرج طاقة المجمع باستخدام المعادلة رقم (2)

$$P_c = A_c \cdot \eta_c G \dots \dots \dots (2)$$

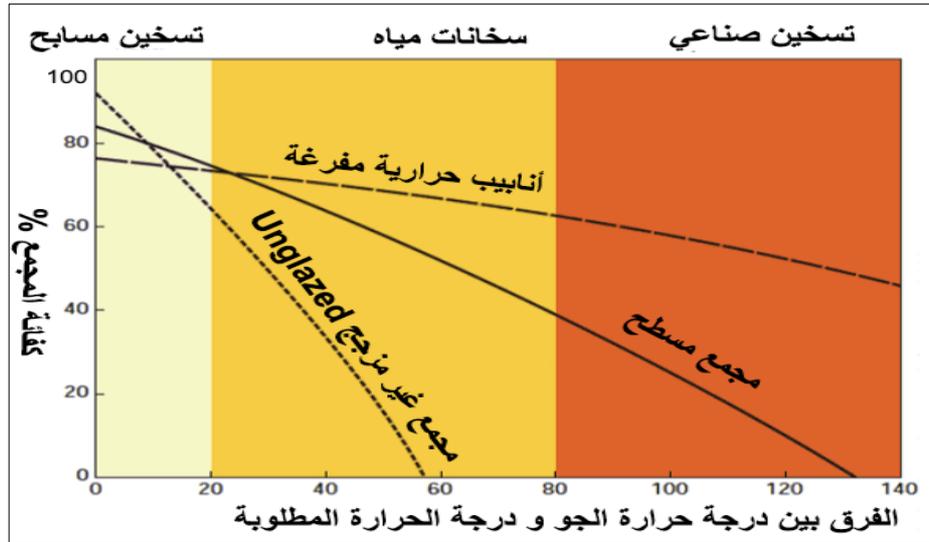
حيث أن: P_c : إنتاج الطاقة من المجمع [W]

A_c : مساحة المجمع [m²]



شكل رقم (4): التعبير عن الكفاءة [9,4]

الشكل رقم (5) يوضح أن أداء سخان ذو الأنابيب المفرغة أكثر كفاءة من أداء السخان ذو الألواح المسطحة حيث أن الأنابيب المفرغة تلائم المناطق الشمالية والتي يكون مناخها بارداً وذلك لأنها تتميز بكفاءة عالية رغم انخفاض درجة حرارة الجو، أما في المناطق العربية فأن الفرق بين النوعين من هذه التقنيات ليس كبيراً [9].



شكل رقم (5): الكفاءة الحرارية للمجمعات الشمسية [4]

5. الطاقة المطلوبة لتسخين الماء

لتحديد او معرفة كمية الطاقة المطلوبة لتسخين ماء السخان يمكن حسابها باستخدام المعادلة رقم (3)، [11,10,8]

$$Q = m c_p \Delta T \dots \dots \dots (3)$$

$$\Delta T = (t_2 - t_1)$$

$$Q = m c_p (t_2 - t_1)$$

حيث أن: Q هي كمية الطاقة المطلوبة بالكيلو جول (KJ)

m ، كتلة او كمية الماء المطلوب تسخينه (Kg)

C_p ، الحرارة النوعية للماء °C 4.186 KJ/Kg

، ΔT فرق درجتي الحرارة قبل وبعد التسخين

، (t₂ - t₁) درجات حرارة الماء قبل وبعد التسخين على التوالي

اما لتحديد او معرفة كمية الطاقة الكهربائية الاجمالية التي يستهلكها السخان الكهربائي خلال وحدة الزمن فيمكن حسابها باستخدام المعادلة رقم (4)

$$P = w * t \dots \dots \dots (4)$$

حيث أن: P هي الطاقة المستهلكة بالكيلو وات ساعة (Kw.hr)

w ، القدرة الكهربائية للسخان الكهربائي (Kw)

t ، زمن التشغيل (hr)

6. تجميع البيانات وتحليلها

تعتمد هذه الورقة على البيانات المتاحة من تقارير الشركة العامة للكهرباء كمعدلات التوليد للوفاء بالطلب على الطاقة الكهربائية والاستهلاك النوعي للسخانات الكهربائية طيلة العام حيث بينت الشركة العامة للكهرباء وعبر وسائل عديدة وفي برامج توعوية لحث المستهلكين لترشيد استهلاك الطاقة الكهربائية بفصل التيار الكهربائي عن السخانات في بعض الأوقات لأن السخان الكهربائي هو أكثر الأجهزة المنزلية استهلاكاً للطاقة حيث يتراوح استهلاك الطاقة لهذا النوع من سخانات المياه من 1000 وات ساعة إلى 3000 وات ساعة.

ومن خلال تقارير معدلات النمو السكاني التي بينت أن عدد الاسر في مدينة طرابلس يقدر بحوالي 200000 أسرة [12]، وان متوسط استعمال عدد السخانات الكهربائية لكل أسرة (منزل) هو اثنان سخان كهربائي أي ان اجمالي عدد السخان الكهربائية المستعملة في مدينة طرابلس وتستهلك التيار الكهربائي هو 400000 سخان وفقاً للقيمة التقديرية لتعريف استهلاك الكهرباء لمختلف شرائح المستهلكين لكل (ك. و. س) المبينة في الجدول رقم (1)

جدول رقم (1): القيمة التقديرية لشرائح المستهلكين/ (ك.و.س) [13]

القيمة (درهم / ك.و.س)	تصنيف الاستهلاك
20	المنزلي
30	زراعي صغار
32	زراعي كبار
42	صناعي خفيف
31	صناعي ثقيل
68	تجاري

معدلات توليد الطاقة الكهربائية للشركة العامة للكهرباء يدل على ان هناك انتاج كافي مقارنة بالنمو السكاني في كل عام وتشير البيانات أن احتياج الطاقة لشرائح المستهلكين في ليبيا يرتفع بشكل مطرد، ويوضح الجدول رقم (2) مقدار معدلات الحمل الأقصى على الشبكة العامة وارتفاعها بشكل تدريجي مع مرور الوقت.

جدول رقم (2): معدلات أقصى حمل على الشبكة العامة لكل سنة [14]

السنة	الحمل الأقصى ميغاوات	السنة	الحمل الأقصى ميغاوات
1970	151	2005	3857
1980	795	2006	4005
1985	1243	2007	4420
1990	1595	2008	4756
1995	1976	2009	5282
2000	2630	2010	5759
2001	3081	2011	5515
2002	3081	2012	5981
2003	3341	2015	6000
2004	3612	2019	6000

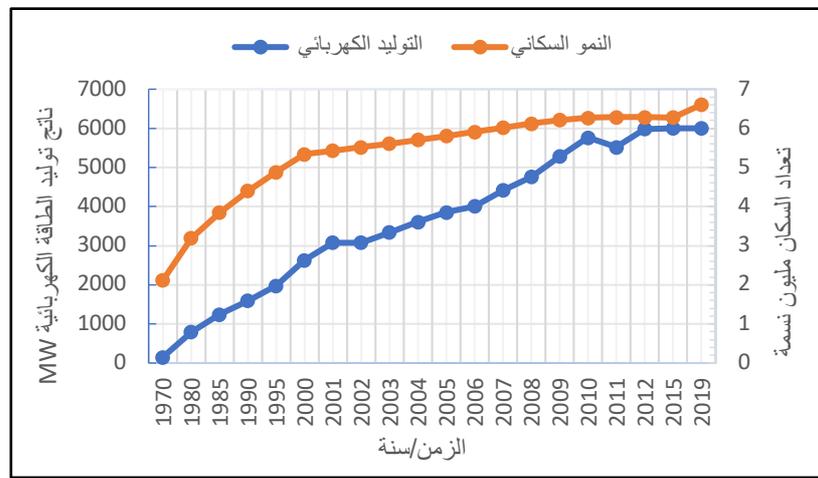
الجدول رقم (3) يوضح معدلات النمو السكاني منذ عام 1970 وحتى عام 2019 وبالمقارنة مع الجدول رقم (2) نلاحظ انه في عام 1970 كان مقدار أقصى حمل على الشبكة العامة في حدود 151 ميغاوات وكان عدد السكان في ذلك الوقت حوالي 2.114 مليون نسمة بينما ارتفع معدل أقصى حمل على الشبكة العامة بنسبة نمو بلغت 3873.5% ليصل إلى 6000 ميغاوات للأعوام 2015، 2016-2019 مع زيادة في نسبة معدل النمو السكاني حيث بلغ 212.5% ليصل إلى 6.606 مليون نسمة كما هو مبين بالعلاقة البيانية في الشكل رقم (6) الذي يوضح معدلات الأحمال القصوى والنمو السكاني.

جدول رقم (3): معدل النمو السكاني في ليبيا لكل سنة [14]

السنة	عدد السكان مليون نسمة	السنة	عدد السكان مليون نسمة
1970	2.114	2005	5.802
1980	3.191	2006	5.907
1985	3.841	2007	6.018
1990	4.398	2008	6.123

6.209	2009	4.878	1995
6.266	2010	5.337	2000
6.289	2011	5.428	2001
6.283	2012	5.518	2002
6.278	2015	5.609	2003
6.606	2019	5.703	2004

إن معدلات استهلاك الطاقة الكهربائية بدون أدوات التسخين والتدفئة الكهربائية التقليدية هي معدلات طبيعية ويتم تغذية الطلب على الطاقة بدون عجز (عدم اللجوء لطرح الأحمال) ولكن مع دخول فصل الشتاء وانخفاض درجات الحرارة يبدأ الاستهلاك المفرط للطاقة الكهربائية عن طريق تشغيل (الدفايات وسخانات الماء الكهربائية) والاعتماد عليها بشكل أساسي الأمر الذي يعود سلباً في تغذية الطلب على الطاقة خلال الذروة الشتوية.



شكل رقم (6): معدلات أقصى حمل على الشبكة العامة والنمو السكاني

7. النتائج 1.7 السخان الشمسي

إن الاحتياج للماء الساخن للاستخدام المكثف يكون بين شهري سبتمبر وأبريل لفترة زمنية مدتها 8 أشهر ومتوسط الشعاع الشمسي في اليوم تقريباً 8 ساعات. فإذا كانت أسرة مكونة من خمسة أفراد وحاجة كل فرد 50 لتر/يوم من الماء أي ان الاحتياج الكلي سيكون 250 لتر/يوم أي ان الاحتياج الكلي لكمية الحرارة لتسخين كمية الماء خلال اليوم يمكن حسابه باستخدام المعادلة رقم (3)

$$2250 \times 4.186 \times 8 (50 - 20) = 251160 \text{ KJ/day}$$

وكمية الحرارة اللازمة لمدة 8 أشهر

$$251160 \times 8 \times 30 = 60278400 \text{ KJ}$$

2.7 السخان الكهربائي

في حالة السخان الكهربائي كل الطاقة الكهربائية تتحول إلى طاقة حرارية ولهذا تعتبر كفاءته 100% ولحساب الطاقة المطلوبة خلال السنة:

$$Q = 60278400/3600 = 16744 \text{ Kw.hr}$$

سعر الكيلو وات ساعة 0.02 دينار لبيي فان التكلفة الكلية واللازمة لقيمة هذا الاستهلاك من الكهرباء تكون

$$\text{Total cost} = 16744 \times 0.02 = 334.88 \text{ LD/year}$$

3.7 إعادة المبالغ المصروفة

إذا كان سعر السخان الشمسي ذو جودة عالية ويوفر مقدار الطاقة المطلوبة هو 500 دينار لبيي وكما هو معروف في حساب زمن إعادة المبالغ المصروفة (**Pay-back period**) هو حاصل قسمة السعر الكلي للمنظومة (السخان الشمسي) على قيمة التوفير. ونظراً لأن كفاءة السخان الشمسي هي 80% من كمية الطاقة المطلوبة واللازمة للتسخين، فان إعادة المبلغ يكون على النحو التالي:

$$\text{Pay_back period} = 500/334.88 \times 0.8 = 1.8 \cong 2 \text{ years}$$

سنتان كمدة تقريبية هي الفترة الزمنية اللازمة لاسترجاع القيمة المدفوعة لشراء السخان الشمسي الذي تم تركيبه بدلاً من الصيانة.

ولتحديد الاستهلاك الكلي لسخان الماء الكهربائي باعتباره أحد أكثر الأجهزة المنزلية استهلاكاً للطاقة يمكن استعمال المعادلة رقم (4) أخذين بنظر الاعتبار الفترة الزمنية لاستعمال السخان خلال العام هي 8 أشهر ومتوسط ساعات التشغيل في اليوم 5 ساعات تقريباً وأن متوسط قدرته هي 1.5 كيلو وات.

$$E = 1.5 \times 5 = 7.5 \text{ kW. hr} \quad \text{معدل الاستهلاك اليومي للسخان الواحد}$$

$$E = 7.5 \times 30 \times 8 = 800 \text{ kW. hr} \quad \text{معدل الاستهلاك السنوي للسخان الواحد}$$

$$E = 1800 \times 400000 = 720 \times 10^6 \text{ kW. hr} \quad \text{وأن معدل استهلاك 400000 سخان كهربائي سيكون}$$

$$E = 720 \text{ MW}$$

وهو المعدل الكلي للاستهلاك السنوي خلال فترة استعمال الماء الساخن لمدة 8 أشهر.

8. الاستنتاجات

تطبيقات الطاقة المتجددة في ليبيا يعتبر محدوداً جداً نظراً لاعتماد الدولة على المصدر الوحيد للطاقة وهو الوقود الأحفوري (النفط) لتوليد الطاقة التي تحتاجها شرائح المستهلكين ومع زيادة معدلات النمو السكاني يزداد الطلب على الطاقة كما هو واضح في الشكل رقم (6) وان قيمة الأحمال والتي تصل ذروتها إلى 6000 ميجاوات واحياناً إلى اقل بكثير من ذلك نتيجة لأسباب عدة منها الإفراط الغير مقنن في الاستهلاك من قبل المواطنين والذي يفتقر إلى ابسط تدابير مشروعات ترشيد استهلاك الطاقة الكهربائية علاوة على ذلك عدم التزام المستهلكين بدفع وصادق قيم فواتير استهلاك الكهرباء رغم رخص قيمة التعريفات للكيلو وات ساعة، أيضاً من ضمن الأسباب الرئيسية في عدم ثبات قيم التوليد تهالك المحطات وعدم الالتزام بأجراء الصيانة الدورية مما أدى إلى انخفاض في كفاءة الأداء لمعظم محطات التوليد وخروجها عن الخدمة في أوقات الذروة نتيجة لضغط الطلب في الظروف المناخية السيئة أيضاً الأوضاع الراهنة وعدم استقرار الدولة ساهم في عرقلة امدادات الوقود لتزويد المحطات وغيب العدالة في التوزيع المتساوي لساعات طرح الأحمال مما أدى إلى الانهيار المستمر لدوائر الشبكة وحدوث الإخلام التام (BLACKOUT).

إن معدلات التوليد الحالية حتماً ستكون كافية إذا ما رافقها الاستهلاك الأمثل للطاقة وتحقيق الجباية الكاملة لقيم الاستهلاك من مختلف شرائح المستهلكين علاوة على أن عدد سكان ليبيا يعتبر قليلاً مقابل ما يتم توليده من طاقة كهربائية لسد احتياجات المواطنين إذا ما قورن بعدد سكان ومعدلات التوليد مع دول الجوار على الأقل.

ولهذا فإن الاعتماد على الطاقة البديلة كمصدر أخر للطاقة سيحقق نتائج إيجابية كما أظهرته هذه الدراسة من نتائج سواء على صعيد تخفيض الاحمال الكهربائية لسد العجز في أزمة توليد الكهرباء وكذلك المحافظة على البيئة من التلوث نتيجة لتخفيض الانبعاثات الناتجة من عوادم المحطات التي تعمل بالوقود المشتق من النفط.

إن الاستثمار في توظيف تقنيات المجمعات الشمسية ذات موثوقية ويمكن الحصول على درجات حرارة عالية اعتماداً على كفاءة المجمع الشمسي وشدة الإشعاع الشمسي عند استبدالها بالسخان الكهربائي على الأقل في الاستخدام السكني سيوفر كميات عالية من الطاقة الكهربائية تعادل توليد محطات متكاملة إذا ما التزمت به الدولة كمشروع وطني استثماري وأن عوائد هذا الاستثمار ستكون واعدة وفترة استرجاع راس المال ستكون قريبة جداً.

إن إدخال 200000 سخان مياه شمسي بدلاً من 400000 سخان كهربائي لمائتي ألف أسرة (منزل) يؤدي إلى توفير استهلاك في الطاقة الكهربائية بحوالي مقدار 720 ميجاوات.

وأن فترة استرداد ثمن المجمع الشمسي (Pay-back period) في حال استبدالها بالسخان الكهربائي ستكون خلال فترة لا تتعدى العامين.

في اعتقادنا أن هذا البحث يفتح المجال لإجراء دراسات أكثر شمولية فهو يقدم دراسة علمية معرفية أولية عن أهم التقنيات الشمسية شائعة الاستعمال والمناخية في مجال تسخين المياه بالطاقة الشمسية ويركز على كيفية الاستفادة من هذه المنظومات على الصعيد المحلي كبديل عن منظومات السخانات الكهربائية ومقدار الوفر الذي تحققه للمساهمة في حل أزمة الطاقة كما هو واضح في النتائج والحسابات.

9. التوصيات

تقدم هذه الورقة نظرة عامة عن بعض أجهزة تسخين المياه بالطاقة الشمسية وتعرض النتائج التي تم الحصول عليها بعد استبدالها بالسخانات الكهربائية المنزلية لترشيد استهلاك الطاقة الكهربائية والمساهمة في خفض ساعات طرح الأحمال على المواطنين ولتحقيق ذلك بالشكل الأمثل يُوصى بما يلي:

- 1- ضرورة التركيز وإنجاز المزيد من البحوث التطبيقية لدراسة إمكانية تطبيق هذه المنظومات الشمسية التي تُستعمل في توفير الماء الساخن على كافة الأصعدة وتأثيرها على تنمية الاقتصاد نظراً لما تحققه من وفر في المال والمحافظة على البيئة.
- 2- التأكيد على ضرورة ربط المراكز البحثية والجامعات ومؤسسات التعليم العالي التقنية والصناعية بحيث توجه الخطط البحثية بما يطرح العديد من الدراسات في هذا المجال.
- 3- تشجيع المواطنين عبر وسائل الإعلام المختلفة لحثهم على استعمال مثل هذه التقنيات لما لها من فوائد واعدة كإحدى أزمات الطاقة وتخفيض فاتورة الكهرباء والمساهمة في الحفاظ على البيئة.
- 4- تصميم وتنفيذ المباني العامة والخاصة لإمكانية استيعاب منظومات الطاقة الشمسية بمختلف أنواعها لضمان سهولة تركيبها وثبوتها.
- 5- أن تتبنى الدولة توطيد مشاريع الطاقات البديلة كمشروع وطني يساهم في توفير القدر المطلوب من الطاقة وحل الأزمات الناتجة من محطات توليد الطاقة التقليدية.

المراجع

- [1]. نوري محمد الكشويو & مصطفى الطاهر العائب، دراسة أداء مجمعات شمسية كهر وحرارية تحت الظروف المناخية لمدينة طرابلس باستخدام برنامج المحاكاة (TRNSYS) *المجلة الدولية المحكّمة للعلوم الهندسية وتقنية المعلومات*، المجلد 3، العدد 2، يونيو (2017) (8-13)
- [2]. Kalogirou, S. Solar Energy Engineering: Processes and Systems. In *Solar Energy Engineering*. Elsevier Inc. (2009)
- [3]. A. A. M. SAYIGH. Solar Energy Engineering, 2nd Edition. (2013) *Handbook*.
- [4]. Laughton, C., & Jackson, F. Solar domestic water heating. (2010). The Earthscan expert handbook for planning, design and installation.
- [5]. Kalogirou, S. a. Solar thermal collectors and applications. *Progress in Energy and Combustion Science*, (2004). 30(3), pp 231–295.
- [6]. Mahjour, F. Vacuum tube liquid-vapor (heat-pipe) collectors. *Proceedings of the Solar Conference*, (2004). pp 1–6.
- [7]. Vinubhai, T. S., *A Review: Solar Water Heating Systems*. April (2014).
- [8]. Ayompe, L. M., & Duffy, A. Thermal performance analysis of a solar water heating system with heat pipe evacuated tube collector using data from a field trial. *Solar Energy*, (2013). 90, pp 17–28.
- [9]. Trier, D. Solar district heating guidelines. (2012). 7.1-3, pp 1–15.
- [10]. Weiss, W. ,Conversion of solar radiation energy into other energy forms. *Institute for Sustainable Technologies (AEE INTEC) AUSTRIA*, 1–77.
- [11]. Alwaer, A., & Gryzagoridis, J. Desalination system using evacuated tube heat pipe solar collector with improved insulation of the geyser March 2019. *Proceedings of the 27th International Conference on the Domestic Use of Energy, DUE 2019, March*, pp 140–146.
- [12]. [http://bsc.ly/\(Bureau-of-Statistics-and-Census-Libya\)](http://bsc.ly/(Bureau-of-Statistics-and-Census-Libya))
- [13]. [WWW.GECOLLY\(General-Electricity-Company-of-Libya\)](http://WWW.GECOLLY(General-Electricity-Company-of-Libya))
- [14]. https://ar.wikipedia.org/wiki/قطاع_الكهرباء_في_ليبيا

Study the Performance and Behavior of the Hybrid System

Hamza Ahmed¹

¹ Hamzaahmado@yahoo.com

¹ Department of Marine and offshore Engineering, College of Engineering, Tripoli University, Libya

ABSTRACT

Keywords:

Renewable energy,
the hybrid system alkaline
electrolysis,
PEM fuel cell,
Study the performance.

The hybrid power generation system is considered one of the cleanest technologies to generate energy, environmentally friendly and does not emit any pollutants. The hybrid system is a mixture of renewable energy sources, as it consists of the two most important sources of renewable energy, solar or wind energy and hydrogen energy. The hydrogen gas production by alkaline electrolysis cell is one of the cleanest methods, the fuel cell (PEM) type is a clean device that generates energy by using hydrogen gas as a main fuel and is considered an environmentally friendly device that only produces pure water and heat. Preparing an accurate mathematical model is an important process to study the behavior of this system and how it works. The performance of system has been studied by operating it with different loads of current density, the modeling process showed satisfactory results. The operating process of system at different temperatures showed good results and clear improvements in system behavior, the operating power of the alkaline electrolysis cell decreases with increasing temperature, because the temperature effects on the reverse and activation voltage when the process of breaking the covalent bond to split the water molecule, the power of fuel cell is increases with increase the temperature. The pressure improves the performance of the both cells but not significantly and with small values, the voltage and power of the both cells is increases with increasing the value of the operating pressure, but this increase is considered very small. However, the system must not be operated at temperatures higher than (95°C) in order to preserve the manufacturing materials of the both cells. Also, keeping the amount of water in the alkaline electrolysis cell from reaching the steam stage, and keep the humidity of electrolyte membrane and the operating life of the fuel cell.

1- Introduction

Energy in the world today has become an important priority, as energy is the main axis of the world. The world cannot be imagined without any type of energy, for a simple example the sunlight is considered a type of energy. The impossible thing is live on planet Earth without the light energy of the sun and impossible to do any simple work without consuming some amount of energy. Energy is divided into two main types, traditional (non renewable) and renewable energy. Oil, gas and coal are traditional energy sources these sources have been used for several years as a main source of energy, and they have proven extremely dangerous and seriously harmful to the components of the world, such as humans, animals and plants because these sources of energy emit high amounts of pollution, fumes and harmful emissions and when the harm and danger became clear and threatening the world, and researchers began to notice a large amounts of

pollution and huge statistics. In this way, it became necessary to go to sources of energy production safer and more secure. Renewable energy sources have achieved very satisfactory results in terms of safety and environmental protection, and have proven their complete environmental friendliness and protection to the world. Scientific research has proven over the years the effectiveness of renewable energy in protecting the world as it does not emit any pollutant, harmful gases or fumes during the process of operation, production and consumption. Renewable energy types include solar energy, dams, wind and hydrogen energy but these sources are not without disadvantage like other sources due the interruption of these sources from time to time, such as solar energy when clouds are formed. Also, some countries do not have these sources, such as wind energy, as some countries have little to no winds, but some renewable resources may be available at all times, through some operational and design techniques, for example, hydrogen energy is considered a smooth energy and easy to use.

The hydrogen element is a good carrier of electrical energy cause atom of hydrogen carries two electrons, the molecular hydrogen element is not available in nature, hydrogen can be produced using several processes, including thermal and electrical, but thermal processes result in polluting carbon compounds, this part has been excluded in energy production processes such as conventional energy [5], but electrical processes such as electrolysis of the water element do not produce pollutants or gases, only the water is splitting to primary compounds which are hydrogen and oxygen gas by breaking the covalent bond by electric current. The splitting process is done by the alkaline electrolysis cell then the produced gas is collected and stored in the storage cylinder. The hydrogen gas produced by the electrolysis process is used to generate electrical energy by the hydrogen fuel cell, an integrated device is designed to produce hydrogen and use it to produce energy called direct coupling advanced alkaline electrolysis and fuel cells system where the design was very satisfactory and the operational results were homogeneous and close between the units of the system to work as one unit for the production and generation of power [6].

In this paper, the system will be operated on different operating parameters and monitor the system performance and operational behavior in terms of the energy produced.

2- Previous Literature

The researchers proposed the design of an integrated system that combines solar energy and hydrogen energy to generate electrical energy, also achieving the cost for this system. As this system consists of solar panels (photoelectric), battery, alkaline electrolysis, hydrogen storage system and proton exchange membrane fuel cell. The aim of this study was to design and implement this system, as it is considered an independent power system for homes, in order to overcome the need to connect to the electrical network. Also, the effect of the pressure and the operating temperature of the system was studied, the Nernst voltage increases with increasing the pressure of the cell, as for the electrical efficiency increases with increasing pressure. As for the temperature has a negative effect on the Nernst voltage, as the Nernst voltage decreases with increasing temperature, as for the current density, increasing the current density increases the losses in the cell [12].

The clean, environmentally friendly device has been designed that integrates renewable energy sources, this system relies on hydrogen energy as a basic fuel which is producing by alkaline electrolysis cell when the electric current is passing from wind or solar energy, and then generate the electrical energy by injecting the hydrogen produced into the proton exchange membrane fuel cell, this technology is considered useful because of the problems faced by renewable energy sources such as solar energy and wind, including the weather and the lack of availability of these sources. The simulations result show that, the voltage of alkaline electrolysis is higher than the fuel cell, the electrolysis cell does not need to inject more water, only the water generated from the fuel cell and is injected to electrolysis. In this research paper, this system will be developed by studying the effect of temperature and operating pressure of the system [6].

3- Materials and Methods

3.1 The Hybrid System

Renewable energy sources are widely available in terms of scientific research, technical development and in energy generation processes. Solar and wind energy are the most important of these sources, as these sources are more abundant and have achieved clear development and technical progress during the last few years. However, these sources are exposed to some operational problems and difficulties, such as bad weather. Thus, electricity generation is variable in operational times and inability to meet the required load during a period of low available resources, it is appropriate to overcome such difficulties and problems the generated energy was stored in the battery. However, batteries have problems and disadvantage, including short battery life, problems related to storage depth, and also lose their storage content quickly [7].

Therefore, it is better to search for other ways in renewable energy sources to store the generated energy, the technique of combining renewable energy sources is considered one of the most successful methods, such as combining solar energy with hydrogen energy.

The Hydrogen molecule does not exist in nature independently, but combines with other elements, the hydrogen gas can be produced by clean technologies such as electrolysis process. Hydrogen gas is produced by the alkaline electrolysis cell and utilized by storing it in cylinders and then injecting it into the hydrogen fuel cell to produce energy according to the required power at the appropriate time. This technology is called a hybrid system and was designed in the previous research paper [6].

The hybrid system consists of two basic units, which are the gas production unit (alkaline electrolysis cell) and the electrical power production unit (fuel cell) as shown in Figure .1.

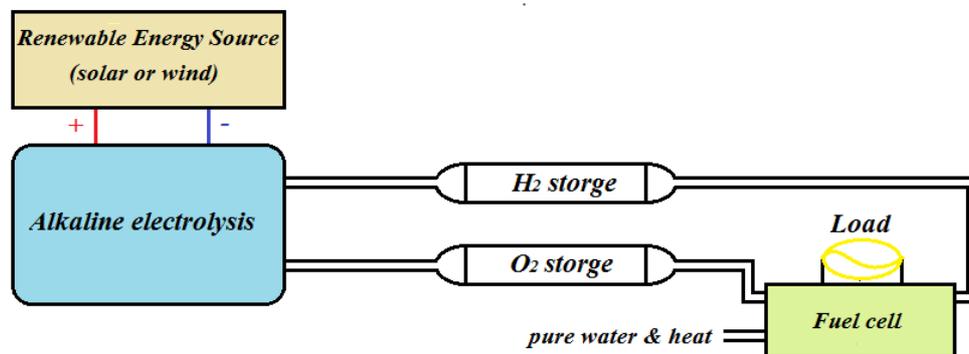


Figure 1: The layout of hybrid system.

The alkaline electrolysis cell requires an electrical current which is generated by a renewable energy source (solar or wind energy) to produce hydrogen gas, then the hydrogen gas that was produced is pumped into the fuel cell to produce the electrical energy at the required power and at the right time.

3.2 Alkaline Electrolysis Cell

Alkaline electrolysis cell consists of several metal chips connected together, these chips are divided into two important parts, the cathode and the anode. The two electrodes are immersed in a conducting electrolyte solution as shown in Figure .2. The electrodes of the alkaline electrolysis cell are connected to a direct current (DC). Electrons pass from the negative electrode through the cathode, where hydrogen atoms form on the cathode as a result of the consumption of hydrogen ions (protons) and the hydroxide ions move towards the anode of the cell. The two compartments are separated by a diaphragm, the produced gas is received by gas collector [8].

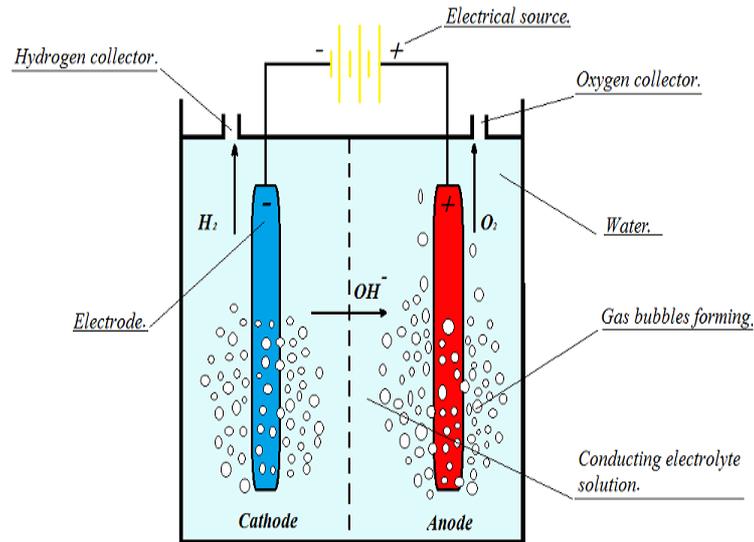


Figure 2: Alkaline electrolysis cell.

3.2. PEM Fuel Cell

Fuel cell is a device that converts the chemical energy in a chemical reaction into electrical energy through a set of reactions which is occur inside the fuel cell, hydrogen gas is the main fuel for the fuel cell. The fuel cell consists of two main electrodes, the cathode and the anode, and the electrolytes are between them. The Catalysts are placed between the electrolyte and the electrodes, hydrogen gas is pumped at the anode and oxygen gas at the cathode as showing in figure .3.

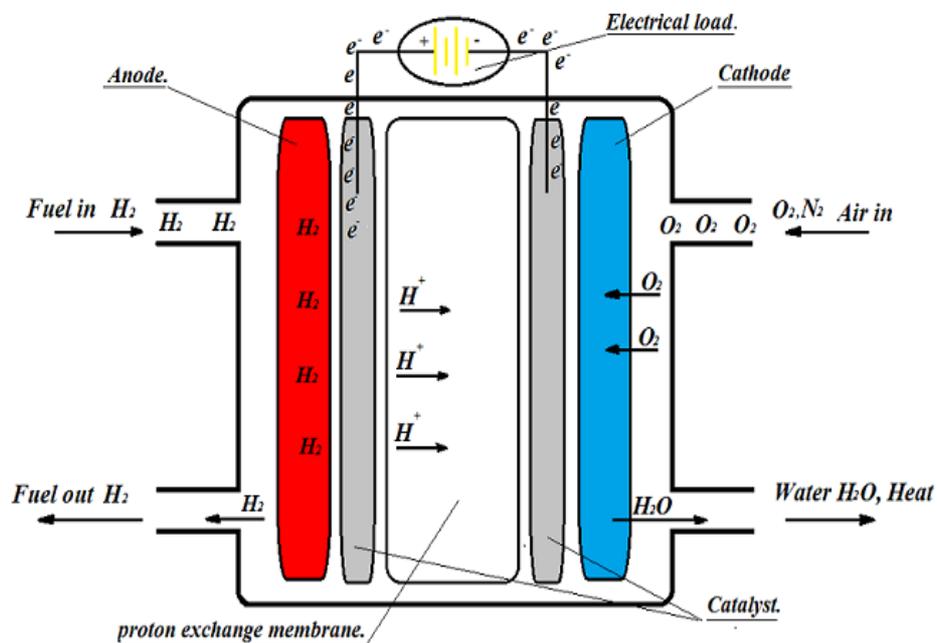


Figure 3: PEM fuel cell

The ions pass through the electrolyte to cross into the cathode, and does not allow electrons to pass through it. Thus, electrons pass to the external electrical wire (electrical load), after the occurrence of the processes and chemical reactions, the cell outputs are pure water and heat [9].

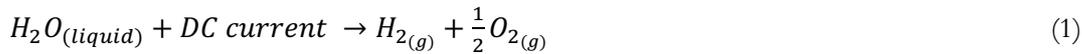
4- The Mathematical Model of System

A system is an assembly unit of components with each other, the system's function is to make a custom function or process under a specific behaviour, the system is described by the mathematical

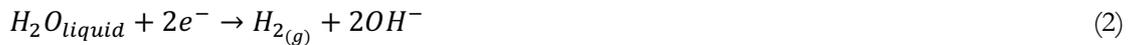
model known as the mathematical concept. Mathematical model is developed by mathematical modelling. The modelling is defined as the writing of equations to describe the mode of a system. The mathematical model is used to predict the behaviour of the system and its mechanism of work so the mathematical model must be more accurate. Preparing an accurate mathematical model of the system is done by using a computer simulation program to perform difficult mathematical operations and many calculations. The computer simulation program allows the system to run with many parameters and different values [1].

4.1 Alkaline electrolysis model description

The description of the system dynamically and mathematically depends on preparing an integrated mathematical model for the alkaline electrolysis cell. The model also describes the processes that occur within the system, including the thermodynamic and electrochemical sides. The general electrochemical reaction of electrolysis [10]:



1. Hydrogen evolution reaction (HER)



2. Oxygen evolution reaction (OER)



The reversible voltage It can be said that the maximum possible useful work of the electrolysis (reversible work) [2], and can be calculated by following formula [11]:

$$V_{rev} = V_{rev}(T) + \frac{R \cdot T}{z \cdot F} \ln \frac{(P - P_{KOH})^{1.5} \cdot P_w}{P_{KOH}} \quad (4)$$

$$V_{rev}(T) = 1.5184 - 1.5421e^{-3} \cdot T + 9.523e^{-5} \cdot T \cdot \ln T + 9.84e^{-8} \cdot T^2 \quad (5)$$

Where P_w is the vapour pressure of a purified water, P_{KOH} the vapour pressure of the electrolyte solution in the (atm) unit, R gas constant $8.314 \text{ (J/mol} \cdot \text{K)}$, F Faraday constant 96485 coulombs , T the temperature of electrolysis cell (K), z the number of electrons in reaction. The voltage of alkaline water electrolysis cell can be expressed in the following relationship [11]:

$$V_{cell} = V_{rev} + V_{act} + V_{ohm} \quad (6)$$

The activation voltage required for the electrochemical reaction, it is also a measure of the extent of the electrode activity of the electrolysis cell [11], Can be calculated by the following formula:

$$V_{act} = 2.3026 \frac{R \cdot T}{z \cdot F \cdot \alpha_{a/c}} \log \left(\frac{i}{i_o} \right) \quad (7)$$

Where the $\alpha_{a/c}$ is the charge transfer coefficient of the electrodes, i_o the exchange current density (A/cm^2).

The ohmic voltage of the alkaline electrolysis cell occur as a result of the ohmic losses due to the flow of electrons on the surface of the electrodes and conductors of the alkaline electrolysis cell, and V_{ohm} is given by:

$$V_{ohm} = r \cdot i \quad (8)$$

$$r = \frac{\delta}{\sigma_\varepsilon} \quad (9)$$

$$\sigma_\varepsilon = \sigma_o(1 - \varepsilon)^{1.5} \quad (10)$$

δ The electrolyte thickness (cm), σ_o the conductivity of KOH solution (S/cm), σ_ε the electrical conductivity in the presence of bubbles and ε the void fraction of the electrolyte.

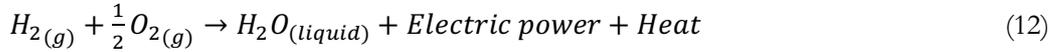
The electrolysis power can be calculated by the following law, which is multiplying the electric current by the electrolysis voltage.

$$Power_{electrolysis} = \left(\frac{i_{FC}}{A} \right) \cdot V_{cell} \quad (11)$$

4.2 PEM Fuel Cell Model Description

Fuel cell modeling is one of the most important processes for obtaining a suitable design for the cell, as developers work to improve the process of designing and manufacturing a fuel cell by preparing an accurate and robust mathematical model that predicts the performance of the fuel cell under a wide range of operating conditions. The model also finds and solves fuel cell problems. The mathematical model illustrates the most important sides of the cell, the thermodynamic and the electrochemical side by showing the chemical equations and electrical laws of the cell [4].

1. The general electrochemical reaction of PEM fuel cell:



2. The anode side:



3. The cathode side:



The electrical model of fuel cell describes the electrical behavior, such as electric current and potential difference, where the cell voltage can be calculated by the following law:

$$V_{Cell} = E - V_{act} - V_{ohm} - V_{conc} \quad (15)$$

Where E the reversible voltage of cell is 1.229 V at standard operating conditions 1 bar and 25°C, and V_{act} the activation voltage of the cell that is occurs due the activation energy when a chemical reaction is start. It can be calculated using the following law [3]:

$$E = 1.229 - 0.85 \cdot 10^{-3} (T - 298.15) + 4.3085 \cdot 10^{-5} \cdot T \cdot \ln \left[\frac{P_{H_2} \cdot P_{O_2}^{0.5}}{P_{H_2O}} \right] \quad (16)$$

Where T the temperature of cell (K), P_{H_2} pressure of fuel, P_{O_2} pressure of oxidant, P_{H_2O} pressure one mole of water.

$$V_{act} = -\xi_1 + \zeta_2 \cdot T + \xi_3 \cdot T \cdot \ln(C_{O_2}) + \zeta_4 \cdot T \cdot \ln(i_{FC}) \quad (17)$$

Where ξ_i are parametric coefficients, i_{FC} is the cell current (A) C_{O_2} the concentration of oxygen gas (mol/cm³) can be found in the following equation [3]:

$$C_{O_2} = \frac{P_{O_2}}{(5.08 \cdot 10^{-5} \cdot \exp(\frac{-498}{T}))} \quad (18)$$

The ohmic voltage of the cell occurs as a result of ohmic losses due to the passage of electrons through the electrical wires and conductors between the two electrodes, the ohmic voltage value is estimated by the following law:

$$V_{ohm} = i_{FC} (R_m + R_c) \quad (19)$$

Where R_c is the constant part of cell's resistance, and R_m depends on the temperature and on a parameter Ψ which represents membrane hydration level (takes the value 14 if the membrane is fully hydrated and 23 if the membrane is over saturated), and R_m is given by:

$$R_m = \frac{\rho_m \cdot l}{A} \quad (20)$$

Where l is the membrane thickness, A the cell active area (cm^2), ρ_m specific membrane resistance ($\Omega \cdot \text{cm}$) and estimated by the following law [3]:

$$\rho_m = \frac{(181.6 \cdot [1 + 0.03 \cdot (\frac{i_{FC}}{A})] + 0.062 \cdot (\frac{T}{303})^2 \cdot (\frac{i_{FC}}{A})^{2.5})}{([\Psi - 0.364 - 3 \cdot (\frac{i_{FC}}{A})] \cdot \exp(\frac{4.18 \cdot (T - 303)}{T}))} \quad (21)$$

The concentration voltage of the cell occurs as a result of the excess reactive concentration near the catalyst surfaces, where B is parameter that depends on the type of fuel cell, and can be calculated by using following law:

$$V_{conc} = -B \cdot \ln(1 - \frac{i}{i_L}) \quad (22)$$

The cell power can be calculated by the following law, which is multiplying the electric current by the cell voltage.

$$Power_{fuel\ cell} = (\frac{i_{FC}}{A}) \cdot V_{cell} \quad (23)$$

Table 1: Setting parameters of alkaline electrolysis [6].

Constant parameter	Symbol	Value
Area of electrodes	A	475 cm^2
Faraday's constant	F	96485 C/mole
Number of cells	n_c	36
Number of electrons	z	2
Gas constant	R	8.314 J/mole*K
exchange current density	i_o	30 A/ cm^2

Table 2: Setting parameters of PEM fuel cell [3].

Parameter	Value	Parameter	Value
T	343 K	B	0.016
A	50.6 cm^2	ξ_1	-0.948
L	178 μm	ξ_2	$2.86 \cdot 10^{-3} + 2 \cdot 10^{-4} \ln(A) + 4.3810 \cdot 10^{-5} \ln(C_{H_2})$
P_{H_2}	1 atm	ξ_3	$7.6 \cdot 10^{-5}$
P_{O_2}	1 atm	ξ_4	$-1.93 \cdot 10^{-4}$

P_{H_2O}	1 atm	Ψ	23
R_c	0.0003 Ω	i_L	1.5 $\frac{A}{cm^2}$

5- Results and Discussion

5.1 The Model Validation

In this paper, a hybrid system for clean energy generation was modeled, as this system includes two basic units, a hydrogen gas generation unit (alkaline electrolysis cell) and a power generation unit (fuel cell). The system was operated on different loads of current density and different parameters. The aim of this study was to design an accurate model for this system, and then operate it in a correct and smooth way. Also, monitor system behavior when operating loads change and note the results of both units.

The results were very satisfactory as they were presented in the previous research paper, due to the same approach followed [6]. The system was developed in this research paper by applying different equations for the fuel cell. The results obtained were compared with research papers to ensure the validity of the results that prove the validity of the modeling and compensation [11] [3].

However, in this paper the focus was on other operational parameters such as temperature and pressure, which is very important to clearly predict the behavior of the system. The change in the performance of the both cells was very clear as a result of the change of the operational parameters, as the operating temperature had a significant effect on the cell voltage and power. As for the operational pressure does not have a significant effect on cell performance.

5.2 (I-V) Characteristics of Units

The curve of current and voltage (I-V) characteristics is one of the most important results that must be obtained to evaluate the model being designed, the cells voltage is one of the most important parameters, which is affected by several operational factors such as temperature, operating pressure and current density. The voltage results of both cells are different in terms of values as showing in figure .4, as the value of the current density increases, the voltage value of the alkaline electrolysis cell increases and the voltage of the fuel cell decreases due to the difference in the nature of the chemical reaction in both cells, where the chemical reaction in the alkaline electrolysis cell need a large amount of energy to break the covalent bond between the hydrogen and oxygen atom.

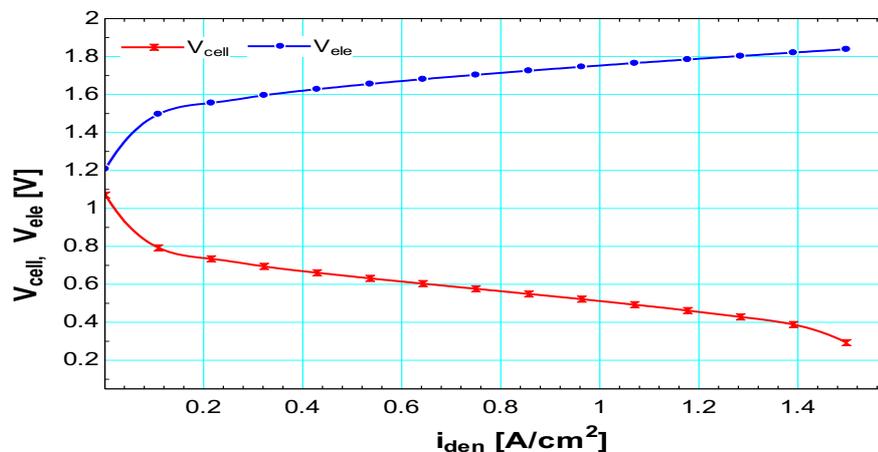


Figure 4: The system voltage changes with the change in the current density.

5.3 The Operating Temperature Effect on Electrolysis Cell Parameters

- The Power and Voltage of Electrolysis Cell

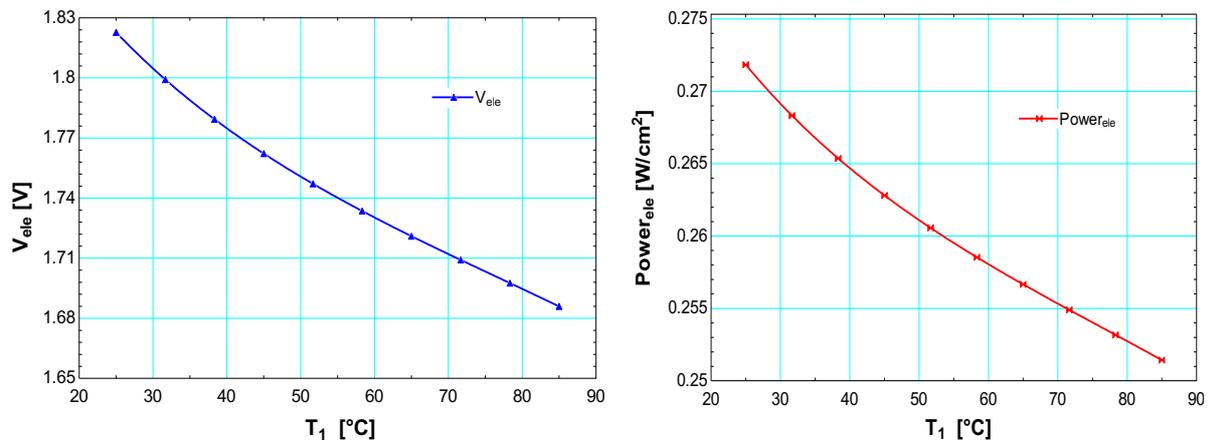


Figure 5: The temperature effect on the voltage and power of alkaline electrolysis.

The effect of temperature on the power and voltage of the fuel cell is very clear, the alkaline electrolysis voltage decreases when the operating temperature increases as showing in figure.5, that the voltage of the alkaline electrolysis cell is a function of temperature such as the reverse voltage and the activation voltage as noted in equation 4, 5, 6. The temperature change affects the water molecules and the covalent bond of their atoms. Thus, the process of break the covalent bond of water molecules will be easy.

5.4 The Operating Pressure Effect on Electrolysis Cell Performance

- The Power and Voltage of Electrolysis Cell

As showing in Figure. 6, the value of voltage and the power of the alkaline electrolysis cell increases with the increase in the operating pressure, but this performance improves only slightly and by small values. However, it is better not to operate the cell with more pressure than the allowable values, in order to preserve the electrolysis mechanism and operation life.

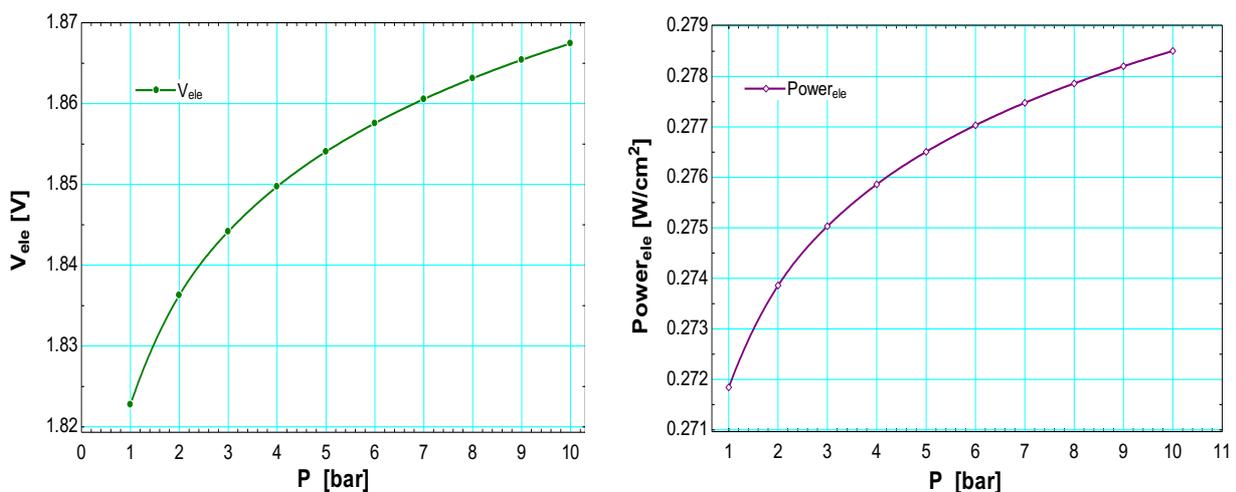


Figure 6: The pressure effect on the voltage and power of alkaline electrolysis.

5.5 The Operating Temperature Effect on Fuel Cell Parameters

- The Power and Voltage of Fuel Cell

The performance of the fuel cell is clearly affected by the change in the operating temperature, unlike the effect in alkaline electrolysis cell. The curve in Figure 7 illustrates the relationship between the change of operating temperature and cell performance (voltage and power), the cell performance increases with increasing cell temperature, and the performance improves significantly, it is possible to improve the performance of the cell by increasing the temperature, but the cell cannot be operated at the temperature (95 °C) or more to run the fuel cell in a safe state, that to keep the humidity of electrolyte membrane and the operating life of the fuel cell.

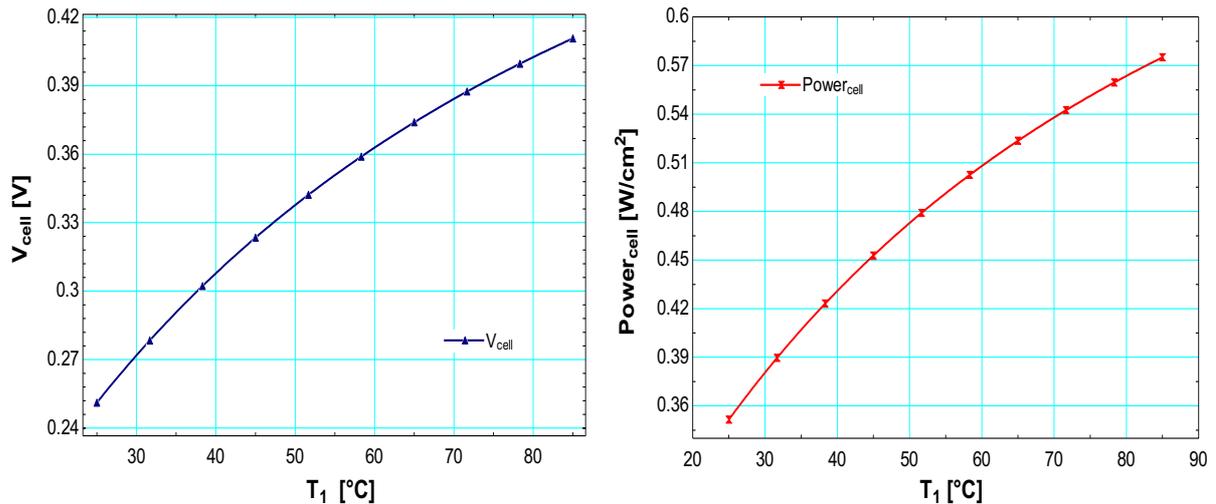


Figure 7: The temperature effect on the voltage and power of fuel cell.

5.6 The Operating Pressure Effect on Fuel Cell Parameters

- The Power and Voltage of Fuel Cell

The operating pressure of the fuel cell is one of the parameters that improve the performance of the cell but not significantly as shown in figure 8, unlike the effect of temperature.

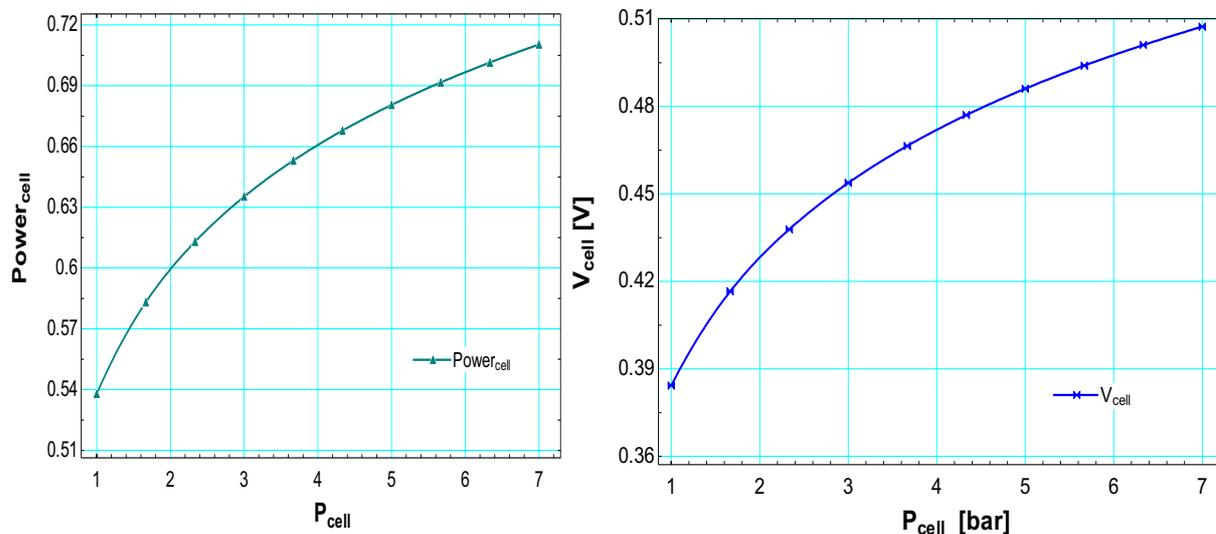


Figure 8: The pressure effect on the voltage and power of fuel cell

6- Conclusions

The renewable energy sources are very clean sources and do not produce any harmful carbon emissions or polluting fumes. However, some of these clean sources face difficulties and problems during power generation and transmission processes, it was necessary to focus on integrating two sources of renewable energy to overcome these difficulties. Solar and hydrogen energy are among the best renewable energy sources, a device consisting of two basic units has been designed as the alkaline electrolysis cell and the fuel cell called the hybrid system.

In this paper, a mathematical model for this system is designed and then the system is operated on different loads of current density. Also, making improvements to the system's performance and studying its operational behavior in terms of the generated voltage and the power produced by operating the system at different temperatures and pressures.

The results showed that the voltage of the alkaline electrolysis cell is $V_{ele} = 1.82 V$ at the current density $i_{ele} = 0.14 \frac{A}{cm^2}$, and the voltage of the fuel cell is $V_{fuel\ cell} = 0.38 V$ at the current density $i_{fuel\ cell} = 1.4 \frac{A}{cm^2}$. The process of operating the system at different temperatures showed good results and clear improvements in system behavior, the operating power of the alkaline electrolysis cell decreases with increasing temperature, because the temperature effects on the reverse and activation voltage when the process of breaking the covalent bond to split the water molecule occurs during the reaction, as the power of fuel cell is increases with increase in the operating temperature, and the cell performance improves significantly. However, the system must not be operated at temperatures higher than (95°C) in order to preserve the manufacturing materials of the both cells. Also, keeping the amount of water in the alkaline electrolysis cell from reaching the steam stage, and keep the humidity of electrolyte membrane and the operating life of the fuel cell. The operating pressure improves the performance of the both cells but not significantly and with small values.

Acknowledgment

The authors acknowledge Tripoli University, Tripoli, Libya for the support of this work.

References

- [1]. A. Patil and S. More, "Mathematical Modeling of Physical System", IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE), Volume 9, Issue 3 Ver. II (May – Jun. 2014), PP 57-64.
- [2]. O. Ulleberg, "Modeling of advanced alkaline electrolyzers: a systemsimulation approach", International Journal of Hydrogen Energy 28 (2003) 21–33.
- [3]. W. Saeed and G. Warkozek, "Modeling and Analysis of Renewable PEM Fuel Cell System", Energy Procedia 74 (2015) 87 – 101.
- [4]. Z. Ural and M. Gencoglu, "Mathematical Models of PEM Fuel Cells", 5th International Ege Energy Symposium and Exhibition (IEESE-5) 27-30 June 2010.
- [5]. C. Kwasi-Effah, I. Obonor and A. Aisien, "A Review on Electrolytic Method of Hydrogen Production from Water", American Journal of Renewable and Sustainable Energy Vol. 1, No. 2, 2015, pp. 51-57.
- [6]. H. Ahmed, A. Musa, "Design of Direct Coupling Advanced Alkaline Electrolysis and Fuel Cells System", 2nd Conference for Engineering Sciences and Technology -CEST2 29-31 October 2019 - Sabratha –Libya.
- [7]. K. Sobotka, "A wind power fuel cell hybrid system study", Master thesis, RES | the School for Renewable Energy Science, University of Iceland & the University of Akureyri, Akureyri, February 2009.



- [8]. D. Santos, C. Sequeira and J. Figueiredo, “Hydrogen production by alkaline water electrolysis”, *Quim. Nova*, Vol. 36, No. 8, 1176-1193, 2013.
- [9]. M. Gencoglu and Z. Ural, “Design of a PEM fuel cell system for residential application”, *international journal of hydrogen energy* 34 (2009) 5242-5248.
- [10]. M. Rashid, M. Al Mesfer, H. Naseem and M. Danish, “Hydrogen Production by Water Electrolysis: A Review of Alkaline Water Electrolysis, PEM Water Electrolysis and High Temperature Water Electrolysis”, *International Journal of Engineering and Advanced Technology (IJEAT)* ISSN: 2249 – 8958, Volume-4 Issue-3, February 2015.
- [11]. J. Koponen, “Review of water electrolysis technologies and design of renewable hydrogen production systems”, Master’s Thesis 2015, Lappeenranta University of Technology.
- [12]. H. Khater, A. Abdelraouf and M. Beshr, “Optimum Alkaline Electrolyzer-Proton Exchange Membrane Fuel Cell Coupling in a Residential Solar Stand-Alone Power System”, *International Scholarly Research Network, ISRN Renewable Energy* Volume 2011, Article ID 953434, 13 page.