Palestinian Journal
of Technology & Applied Sciences (PJTAS)

Palestinian Journal
of Technology & Applied Sciences (PJTAS)

Vision

Achieving leadership, excellence and innovation in the field of open learning, community service, and scientific research, in addition to reinforcing the University leading role in establishing a Palestinian society built on knowledge and science.

Mission

To prepare qualified graduates equipped with competencies that enable them to address the needs of their community, and compete in both local and regional labor markets. Furthermore, The University seeks to promote students' innovative contributions in scientific research and human and technical capacity-building, through providing them with educational and training programs in accordance with the best practices of open and blended learning approach, as well as through fostering an educational environment that promotes scientific research in accordance with the latest standards of quality and excellence. The University strives to implement its mission within a framework of knowledge exchange and cooperation with the community institutions and experts.

Core Values

To achieve the University's vision, mission and goals, the University strives to practice and promote the following core values:

- Leadership and excellence.
- Patriotism and nationalism.
- Democracy in education and equal opportunities.
- Academic and intellectual freedom.
- Commitment to regulations and bylaws.
- Partnership with the community
- Participative management.
- Enforcing the pioneer role of women.
- Integrity and Transparency.
- Competitiveness.

The Journal

The Palestinian Journal of Technology and Applied Sciences is an annual scientific refereed journal, issued by the Deanship of Graduate Studies and Scientific Research. The first issue of the Journal was published in January 2018 after obtaining an International Standard Serial Number (E- ISSN: 2521-411X), (P– ISSN: 2520-7431).

The journal publishes original research papers and studies conducted by researchers and faculty staff at QOU and by their counterparts at local and overseas universities, in accordance with their academic specializations. The Journal also publishes reviews, scientific reports and translated research papers, provided that these papers have not been published in any conference book or in any other journal.

The Journal comprises the following topics:

Information and Communication Technology, Physics, Chemistry, Biology, Mathematics, Statistics, Biotechnology, Bioinformatics, Agriculture Sciences, Geology, Ecology, Nanotechnology , Mechatronics, Internet of things , Artificial Intelligence and Big Data.

Palestinian Journal
of Technology & Applied Sciences (PJTAS)

Publication and Documentation Guidelines

First: Requirements of preparing the research:

The research must include the following:

1. A cover page which should include the title of the research stated in English and Arabic, including the name of researcher/researchers, his/her title, and email.
2. Two abstracts (English and Arabic) around (150-200 word). The abstract should include no more than 6 key words.
3. Graphs and diagrams should be placed within the text, serially numbered, and their titles, comments or remarks should be placed underneath.
4. Tables should be placed within the text, serially numbered and titles should be written above the tables, whereas comments or any remarks should be written underneath the tables

Second: Submission Guidelines:

1. The Researcher should submit a letter addressing the Head of Editorial Board in which he/she requests his paper to be published in the Journal, specifying the specialization of his/her paper.
2. The researcher should submit his research via email to the Deanship of Scientific research (tas@qou.edu) in Microsoft Word Format, taking into Consideration that the page layout should be two columns.

   (Check the attached digital form on the website of the Journal)
3. The researcher should submit a written pledge that the paper has not been published nor submitted for publishing in any other periodical, and that it is not a chapter or a part of a published book.
4. The researcher should submit a short Curriculum Vitae (CV) in which she/he includes full name, workplace, academic rank, specific specialization and contact information (phone and mobile number, and e-mail address).
5. Complete copy of the data collection tools (questionnaire or other) if not included in the paper itself or the Annexes.
6. No indication shall be given regarding the name or the identity of the researcher in the research paper, in order to ensure the confidentiality of the arbitration process.

Third- Publication Guidelines:

The editorial board of the journal stresses the importance of the full compliance with the publication guidelines, taking into note that research papers that do not meet the guidelines will not be considered, and they will be returned to the researchers for modification to comply with the publication guidelines.

1. Papers are accepted in English only, and the language used should be well constructed and sound.
2. The researcher must submit his/her research via email (tas@qou.edu)in Microsoft Word format, taking into consideration the following:
   - Font type should be Times New Roman, and the researcher should use bold font size 14 for head titles, bold font size 13 for subtitles, font size 12 for the rest of the text, and font size 11 for tables and diagrams.
   - the text should be single-spaced
   - Margins: Should be set to: 2cm top, 2.5 cm bottom, 1.5 cm left and right.
3. The paper should not exceed 25 (A4) pages or (7000) words including figures and graphics, tables, endnotes, and references, while annexes are inserted after the list of references, though annexes are not published but rather inserted only for the purpose of arbitration.
4. The research has to be characterized by originality, neutrality, and scientific value.
5. The research should not be published or submitted to be published in other journals, and the researcher has to submit a written acknowledgment that the research has never been published or sent for publication in other journals during the completion of the arbitration process. In addition, the main researcher must acknowledge that he/she had read the publication guidelines and he/she is fully abided by them.
6. The research should not be a chapter or part of an already published book.
7. Neither the research nor part of it should be published elsewhere, unless the researcher obtains a written acknowledgement from the Deanship of Scientific Research.
8. The Journal preserves the right to request the researcher to omit, delete, or rephrase any part of his/her paper to suit the publication policy. The Journal has also the right to make any changes on the form/ design of the research.
9. The research must include two research abstracts, one in Arabic and another in English of (150-200) words. The abstract must underline the objectives of the paper, statement of the problem, methodology, and the main conclusions. The researcher is also to provide no more than six keywords at the end of the abstract which enable an easy access in the database.

11. The researcher has to indicate if his research is part of a master thesis or a doctoral dissertation as he/she should clarify this in the cover page, possibly inserted in the footnote.
12. The research papers submitted to the Deanship of Scientific Research will not be returned to the researchers whether accepted or declined.
13. In case the research does not comply with the publication guidelines, the deanship will send a declining letter to the researcher.
14. Researchers must commit to pay the expenses of the arbitration process, in case of withdrawal during the final evaluation process and publication procedures.
15. The researchers will be notified of the results and final decision of the editorial board within a period ranging from three to six months starting from the date of submitting the research.

Four- Documentation:

1. Footnotes should be written at the end of the paper as follows; if the reference is a book, it is cited in the following order, name of the author, title of the book or paper, name of the translator if any or reviser, place of publication, publisher, edition, year of publishing, volume, and page number. If the reference is a journal, it should be cited as follows, author, paper title, journal title, journal volume, date of publication and page number.
2. References and resources should be arranged at the end of the paper in accordance to the alphabetical order starting with the surname of author, followed by the name of the author, title of the book or paper, place of publishing, edition, year of publication, and volume. The list should not include any reference which is not mentioned in the body of the paper.
   - In case the resource is with no specified edition, the researcher writes (N.A)
   - In case the publishing company is in not available, the researcher writes (N.P)
   - In case there is no author, the researcher writes (N.A)
   - In case the publishing date is missing , the researcher writes (N.D)
3. In case the researcher decides to use APA style for documenting resources in the text, references must be placed immediately after the quote in the following order, surname of the author, year of publication, page number.
4. Opaque terms or expressions are to be explained in endnotes. List of endnotes should be placed before the list of references and resources

   Note: for more information about using APA style for documenting please check the following link:

   http://journals.qou.edu/recources/pdf/apa.pdf

Five: Peer Review & Publication Process:

All research papers are forwarded to a group of experts in the field to review and assess the submitted papers according to the known scientific standards. The paper is accepted after the researcher carries out the modifications requested. Opinions expressed in the research paper solely belong to their authors not the journal. The submitted papers are subject to initial assessment by the editorial board to decide about the eligibility of the research and whether it meets the publication guidelines. The editorial board has the right to decide if the paper is ineligible without providing the researcher with any justification.

The peer review process is implemented as follows:

1. The editorial board reviews the eligibility of the submitted research papers and their compliance with the publication guidelines to decide their eligibility to the peer review process.
2. The eligible research papers are forwarded to two specialized Referees of a similar rank or higher than the researcher. Those Referees are chosen by the editorial board in a confidential approach, they are specialized instructors who work at universities and research centers in Palestine and abroad.
3. Each referee must submit a report indicating the eligibility of the research for publication.
4. In case the results of the two referees were different, the research is forwarded to a third referee to settle the result and consequently his decision is considered definite.
5. The researcher is notified by the result of the editorial board within a period ranging from three to six months starting from the date of submission. Prior to that, the researcher has to carry out the modifications in case there are any.
6. The researcher will receive a copy of the journal in which his/her paper was published, as for researchers from abroad, a copy of the Journal volume will be sent to the liaison university office in Jordan and the researcher in this case will pay the shipping cost from Jordan to his/her place of residency.

Six: Scientific Research Ethics:

The researcher must:

1. Commit to high professional and academic standards during the whole process of conducting research papers, from submitting the research proposal, conducting the research, collecting data, analyzing and discussing the results, and to eventually publishing the paper. All must be conducted with integrity, neutralism and without distortion.

2. Acknowledge the efforts of all those who participated in conducting the research such as colleagues and students and list their names in the list of authors, as well as acknowledging the financial and morale support utilized in conducting the research.
3. Commit to state references soundly, to avoid plagiarism in the research.
4. Commit to avoid conducting research papers that harm humans or environment. The researcher must obtain in advance an approval from the University or the institutions he/she works at, or from a committee for scientific research ethics if there is any, when conducting any experiments on humans or the environment.
5. Obtain a written acknowledgement from the individual/individuals who are referred to in the research, and clarify to them the consequences of listing them in the research. The researcher has also to maintain confidentiality and commit to state the results of his/her research in the form of statistical data analysis to ensure the confidentiality of the participating individuals.

Seven: Intellectual Property Rights:

1. The editorial board confirms its commitment to the intellectual property rights
2. Researchers also have to commit to the intellectual property rights.
3. The research copyrights and publication are owned by the Journal once the researcher is notified about the approval of the paper. The scientific materials published or approved for publishing in the Journal should not be republished unless a written acknowledgment is obtained by the Deanship of Scientific Research.
4. Research papers should not be published or republished unless a written acknowledgement is obtained from the Deanship of Scientific Research.
5. The researcher has the right to accredit the research to himself, and to place his name on all the copies, editions and volumes published.
6. The author has the right to request the accreditation of the published papers to himself.

Palestinian Journal
of Technology & Applied Sciences (PJTAS)
No. 6

Contents

The Effect of Irrigation with Fish and Dairy Farm Effluents on the Production of Date Palm cv. Medjool

تأثير الري بمياه المزارع السمكية ومياه مزارع الأبقار على إنتاج نخيل التمر صنف مجول

Alai Daoud Bitar
Associate Professor/Al-Quds Open University\ Palestine
abitar@qou.edu

علائي داوود البيطار
أستاذ مشارك / جامعة القدس المفتوحة / فلسطين

Hassan Ahmad Abu Qaoud
Professor/ An-Najah National University/Palestine
Hassan@najah.edu

حسان أحمد أبو قاعود
أستاذ دكتور / جامعة النجاح الوطنية / فلسطين

Husameddin Mustafa Isaid
Assistant Professor /Nakheel Palestine Company/ Palestine
hdsaeed@yahoo.com

حسام الدين مصطفى اسعيد
أستاذ مساعد/ شركة نخيل فلسطين / فلسطين

Ismail Hasan Daiq
Researcher /Valley Trading Company/ Palestine
Daiq.Ismail@gmail.com

إسماعيل حسن دعيق
باحث / شركة الوادي / فلسطين

Khalid Wajdi Aslan
Researcher /Al-Quds Open University/Palestine
Khalidaslan94@gmail.com

خالد وجدي أصلان
باحث / جامعة القدس المفتوحة / فلسطين

Ziad Hussein Al-Zuhoor
Researcher /Valley Trading Company/ Palestine
Ziad_Zhour@hotmail.com

زياد حسين الزهور
باحث / شركة الوادي / ف لسطين

Received: 23/ 11/ 2021, **Accepted:** 12/ 2/ 2022
DOI: 10.33977/2106-000-006-001
http://journals.qou.edu/index.php/PJTAS

تاريخ الاستلام: 23/ 11/ 2021م، **تاريخ القبول:** 12/ 2/ 2022م.
E - ISSN: 2521 - 411X
P - ISSN: 2520 - 7431

الملخص:

منذ العقد الأول من القرن الحادي والعشرين، أخذت أشجار النخيل تنتشر بسرعة عبر وادي الأردن في الضفة الغربية. وعلى الرغم من أن نخيل التمر قد يعيش في المناخات القاحلة، إلا أنه يحتاج إلى كمية كافية من الماء للوصول إلى كامل إنتاجه المحتمل. يُنظر إلى الموارد المائية المحدودة في الضفة الغربية على أنها أكبر عقبة أمام التنمية الزراعية.

أجريت هذه الدراسة خلال موسمي 2019 و2020، من أجل تقييم فعالية مصادر مائية بديلة لري أشجار النخيل. تم استخدام تصميم القطاعات الكاملة العشوائية (RCBD) في التجربة لتكرار المعاملات الثلاثة: مياه المزرعة (شاهد)، ومياه المزارع السمكية، ومياه مزارع الأبقار. تم إجراء التحاليل الميكروبية (خمائر، أعفان، E. coli) والكيميائية (درجة الحموضة، الملوحة، الكلور، الصوديوم، النيتروجين، الفسفور، والبوتاسيوم) لجميع أنواع مياه الري المستخدمة في التجربة.

أشارت النتائج إلى عدم وجود ميكروبات ضارة في معاملات مياه الري الثلاثة. كانت الملوحة أعلى في كل من مياه الأسماك والأبقار (6.85 و6.41 dS/m) على التوالي مقارنة بـ 2.89 dS/m في الشاهد.

أظهرت نتائج موسم 2019 عدم وجود فروق معنوية بين المعاملات في معظم الصفات المدروسة في التجربة، بينما أشارت نتائج موسم 2020 إلى تفوق معاملة الري بمياه المزارع السمكية مقارنة بالمعاملات الأخرى، حيث أعطت هذه المعاملة أعلى عدد من الأوراق (26.5 ورقة) لكل شجرة وأعلى إنتاجية للشجرة (67 كغم/ شجرة) وأعلى وزن ثمرة (18.78 غم).

كانت مياه المزارع السمكية ومياه مزارع الأبقار غنية بالنيتروجين والفسفور والبوتاسيوم والتي يمكن أن تعوض الاحتياجات السنوية من NPK لنخيل التمر. وبالتالي، يمكن استخدام مياه المزارع السمكية ومياه مزارع الأبقار كمصدر للري والأسمدة في زراعة نخيل التمر الحديثة .في المستقبل، هناك حاجة لأجراء المزيد من البحوث حول تأثير استخدام مصادر المياه المستخدمة على التربة والنبات

الكلمات المفتاحية: نخيل التمر، مجول، إعادة استخدام المياه، مياه المزرعة، مياه المزارع السمكية، مياه مزارع الأبقار.

Abstract:

Since the 2000s, date palm trees have been rapidly spreading across the Jordan valley in the West Bank. Although the date palm may survive in arid climates, it needs enough water to reach its full potential production. The West Bank limited water resources are seen as the most significant obstacle to agricultural development.

This study was carried out over two growing seasons, 2019 and 2020, in order to assess the efficacy of alternative irrigated water resources for the date palm trees. A randomized complete block design (RCBD) was used in the experiment to replicate the three treatments: farm water (control), fish farm effluent, and dairy farm effluent. Microbial analyses (yeasts, molds, and *E. coli*) and chemical analyses (pH, EC, Cl, Na N, P, and K) were performed for all types of irrigated water used in the experiment.

The results indicated no harmful microbial existence in the three irrigation water treatments. The EC was higher in both fish and dairy effluents (6.85 and 6.41 dS/m) respectively compared to 2.89 ds/m in the control. The results of 2019 season showed no significant differences among the treatments in most parameters studied in the experiment, while the results of the 2020 season indicated the superiority of irrigation with fish farm effluent compared to other treatments. This treatment gave the highest number of leaves (26.5 leaves) per tree, the highest yield per tree (67 kg/ tree) and the highest fruit weight 18.78 g. Fish and dairy farm effluents were rich in nitrogen (N), phosphorus (P), and potassium (K) that could compensate the annual requirements of NPK for the date palm tree. Thus, fish and dairy farm effluents could be used as a source of irrigation and fertilizer in modern date palm cultivation, however, long run investigation is needed to study the impact of using such water sources on soil and plant.

Keywords: Date palm, Medjool, reuse water, farm water, fish farm effluent, dairy farm effluent.

INTRODUCTION

The date palm (*Phoenix dactylifera L.*), represents a main resource for the development of the agricultural economy in Palestine and in other countries around the Mediterranean basin and the Middle East (Russo et al., 2017). In the previous ten years, date production has doubled, with global yields of almost 7.5 million tons in 2014 (Al Alawi et al., 2017).

Since the 2000s, date palm trees have been rapidly spreading across the Jordan valley in the West Bank. It offers promising opportunities for Palestinians to enhance their livelihoods and export profits (Abu-Qaoud, 2015). In Jericho and the Jordan Valley, the entire area of date palms is about 26,000 donum, with roughly 350,000 trees (Daiq, 2020). Although the date palm may survive in arid climates, it needs enough water to reach its full potential production (Abu-Qaoud, 2015). The West Bank's limited water resources are seen as the most significant obstacle to agricultural development (Mazahrih, 2012). The available water in the West Bank declined, and the water quality in many Jordan valley wells degraded. Similarly, the amount and quality of irrigation water in Jericho dropped and the majority of these wells aren't recharged on a regular basis.

In addition, the regional area including Palestine is expected to experience an average temperature rise of 1.4°C to 4°C over the next century, as well as a general decrease in precipitation of 25% regionally and up to 40% locally, a shift in rain seasons from winter and spring to autumn (East, 2019).

As a result of the intensive date palm agriculture in Jericho, as well as the influence of climate change on water availability and quality, including a rise in salinity in both irrigation water and soil, water availability and quality will be impaired (Russo et al., 2017). Therefore, another renewable water sources must be sought to fulfill the water shortage (Abu-Jaish, 2018). Several alternative renewable water sources have been proposed, including irrigation with fish waste and dairy farm effluent. In the last years, many large agricultural pools were constructed in Jericho and Jordan valley. Farmers and investors in agriculture have realized the importance of fish farming in agricultural pools used to irrigate date palm and other crops. Moreover, few farmers in Jericho and Jordan Valley adopted the integrated farming systems between aquaculture farms and plant production farms. These combinations may increase farm income and maximize the benefits from agricultural inputs and the conservation and the sustainable use of natural resources (Ahmed and Garnett, 2011). In 2020, the number of fish farm water pools in Jericho and Jordan valley was about 150 pools, while the number of dairy farm water pools was two pools (Daiq, 2020).

Reusing the drained water of fish farms is a critical approach to sustainable use of available water for irrigation. Drained water of fish farming as a good irrigation resource increased crop productivity and reduced the total costs of fertilizers (Abdelraouf and Hoballah, 2014). Furthermore, Abdelraouf (2017) reported the benefits of fish water effluent as a new water source to irrigate potato, soybean, and onion crops. This application did not require additional fertilizers and thus reduced the fertilizer costs. Khater (2006) utilized the outputs of fish farming as a new technique to minimum requirements of nutrients in growing vegetables, i.e., lettuce, cucumber, tomato, cabbage, etc. Drained water of fish culture was evidenced to be highly enriched with natural fertilizers due to fish excrete which contains high content of ammonia and urea, and organic matter (Elnwishy, 2008). Consequently, this promotes increased nitrogen and phosphorus in the water due to fish excretion and feed leftovers. Therefore, the fish farming effluent water could be noticeable as a source of irrigation and fertilizers in agriculture, especially in areas that suffer from water scarcity (Kaab Omeir et al., 2020).

Dairy farms produce large quantities of wastewater as a result of the thorough washing of milking and feeding parlours, such as wastewater distinguished by certain level milk components, such as casein, lipids, and lactose (Roufou et al., 2021). Applying t hese dairy farm effluents (DFE) to agricultural soils is a widely utilized disposal method due to the benefits these wastewaters bring to the soil and the accompanying improved crop productivity (Longhurst et al., 2000). Among the benefits of land application of DFE, the added organic matter, nutrients, and microbial

mass are known to improve soil aggregation, soil fertility, and water holding capacity (Barkle et al., 2000). Irrigation with DFE onto pasture was increasingly recognized as a means for biological treatment, and realized that DFE was a resource for its mineral content rather than a waste (Hawke and Summers, 2006). An experiment by Islam et al. (2017) studied the impacts of dairy farm's wastewater irrigation on maize's growth and yield attributes. The investigation showed that dairy farm wastewater contained different nutrients and organic matter, which optimistically contributed to the development and production of maize. The highest grain yield, was recorded under dairy farm wastewater irrigation with no fertilizer application. Similar results obtained by Islam (2015) who reported that dairy farm wastewater might be used as a source of irrigation and fertilizer for wheat production without any hazard to soil.

The importance of this study lies in its attempt to test other water sources to compensate the scarcity of water for irrigating date palms, such as fish and dairy farm waste water. There have been no research on the effect of fish and dairy farm effluents on the production of date palm. As a result, our research's goals were as follows:(1) evaluate the use of various water resources (fish and dairy farm effluents) in date palm irrigation in Jordan valley; and (2) examine the fruit quality and air fruit content under various water resources (fish and dairy farm effluents).

MATERIAL AND METHODS

Study site and climatic conditions: This study was conducted at Al-Wadi farm (Daiq Brothers) in Jericho/Palestine, and carried out by Al-Quds Open University Center for Agricultural Research-Jericho for two growing seasons, 2019 and 2020. The experimental site was located at 31°52′16″N latitude and 35°26′39″E longitude. The farm was located in the elevated plain at an altitude of 360 meters below the sea level (bsl). The area received rains about 166 mm annually (PMD, 2020). The climatic conditions at the site are characterized by hot dry summer and relatively warm winter, while the soil is characterized by having silt texture with high pH and high salts content (Dudeen et al., 2001).

Plant materials: The experiment was conducted on six-year-old Medjool date palm trees, spaced at 8×9 meters grown under a drip irrigation system. Date palm trees were selected for their uniformity in vigor and size.

Treatments, replicates, and experimental design: the plot size was one date palm tree. The experiment consisted of three treatments viz., farm water (control), fish farm effluent, and dairy farm effluent, in four replicates using randomized complete block design (RCBD) according to Little and Hills (1978). All trees in the experiment were subjected to the same usual horticultural practices.

Water sources: the Arab Development Society (ADS) provided the fish farm effluent for this experiment, while the dairy farm effluent was got from Al-Juneidi Dairy and Food Products Company which was established recently in Jericho as the biggest modern dairy farm.

Amount of water: Table 1. shows the amount and time of irrigation used in the experiment for all the types of irrigation water.

Table 1 :

Amount of monthly irrigation (m3 of water/tree of date palm cv. Medjool) schedule for all types of water used in the experiment during the two seasons of 2019 and 2020.

No.	Month	Amount of irrigation (m^3 of water/tree)	
		2019	2020
1	January	3.6	4
2	February	3.8	4
3	March	5.2	6
4	April	7.1	8
5	May	9.0	10
6	June	11.0	12
7	July	12.0	13
8	August	12.0	13
9	September	10.0	11
10	October	7.1	8
11	November	6.2	7
12	December	3.7	4
Total		90.7	100

Water analysis:

At the beginning of the experiment, water samples for all types of water sources were collected and sent to the Chemical and Biological Analysis Unit at An-Najah National University to test some microbial analysis (total yeasts count, total molds count and *E.coli*) and some chemical analysis (pH, EC, Cl, Na, N, P, and K) according to Fresenius et al.(1988).

Observations:

At harvest time, the number of leaves with thorn per each tree in the experiment was counted. The yield (kg/tree) was weighed. Average fruit weight (g) was measured by collecting 25 fruits randomly from each tree and weighed. The value obtained for 25 fruits was divided by 25 to obtain the average fruit weight and expressed in grams. Balah fruits refer to the yellow fruits (khalal stage) were computed by weighing the yellow fruits on the tree and expressed in kg /tree. Fruit with stalk refers to the fruit with calyx or cap, and fruit without stalk were those that their calyx or cap have detached from the fruit. Fruit with stalk was computed by weighing the fruit with stalk on the tree (yield /tree), whereas fruit without stalk was computed by weighing the fruit without a stalk (yield /tree). Skin separation (loose skin) refers to air content between the skin and the fleshy parts of the fruit. It categorized from 0-10%, 11-30%, 31-50% and>50% according to air content in the fruit. Air content (%) in the fruit were computed by weighing the fruit for each category (0-10%, 11-30%, 31-50% and >50%) on the tree (Daiq, 2020).

Statistically analysis:

All the data were statistically analyzed and the significant differences in the treatment means were separated according to the LSD test at a 5% level (SAS software, 1990).

RESULTS AND DISCUSSION

Microbial analysis of irrigation water: Yeasts and molds are two different forms of naturally occurring fungi present in the environment. Total Yeasts and Molds Counts (TYMC) are used to detect and quantify the amount of fungal growth on material.

In the present study, total yeasts count (3 CFU/100 ml) were found in dairy farm effluent, while no yeasts were present in both farm water and fish farm effluent (Table 2).

The highest total molds count (2.1 * 102 CFU/100 ml) were found in farm water, followed by fish farm effluent (1.9 * 102 CFU/100 ml), and the lowest (6 CFU/100 ml) were found in dairy farm effluent (Table 2). The higher molds count in the farm water treatment than the other treatments may be attributed to the presence of high content of Cl in both fish farm effluent (2220 mg/l) and dairy farm effluent (2000 mg/l) as compared to farm water (950 mg/l) (Table 3). It is well known that Cl is used extensively as a water disinfectant.

Regarding *Escherichia coli*, there is a human pathogen in irrigation water that can be transmitted to plants. Thus, consumption of fruits irrigated with *E. coli*-contaminated water can cause illness to humans. Fortunately, in the present study, the microbial analysis (Table 2) indicated that all types of water used in the experiment were free from *E. coli*. Based on the total number of *E. coli*, the standards of California recommend a 7-day median value of 2.2 CFU/100 ml, and a maximum value of 240 CFU/100 ml (Fewtrell and Bartram, 2001). On the contrary, the World Health Organization guidelines recommend 1000 CFU/100 ml (WHO, 2006).

Table 2:

Some microbial analyses of farm water, fish farm effluent, and dairy farm effluent used in the experiment.

No.	Test	Farm water	Fish farm effluent	Dairy farm effluent
1	Total yeasts count (CFU/100 ml)	Nil	Nil	3
2	Total molds count (CFU/100 ml)	$2.1 * 10^2$	$1.9 * 10^2$	6
3	E. coli (CFU/100 ml)	Nil	Nil	Nil

Chemical analysis of irrigation water: In the present study, the pH of the water was slightly alkaline, ranging from 7.43 in dairy farm effluent to 8.11 in farm water (Table 3). Generally, the pH of irrigation water used in most date-growing regions in the world ranges from pH 6.5 to 8.8. The pH outside of the normal range has the potential to cause an imbalance of nutrients (Bauder et al., 2011).

In this study, the EC of the water was very high in both fish farm effluent (6.85 dS/m) and dairy farm effluent (6.41 dS/m), while in farm water (2.89 dS/m) was accepted in date palm plantation (Table 3). The EC of irrigation water which was used in Jericho and Jordan valley ranges from 2.8 to 6.0 dS/m. Date palm is classified as the most salt-tolerant fruit crop (Fipps, 2003). However, the Food and Agriculture Organization (FAO) has established guidelines for agricultural water primarily based on salinity. These guidelines have been modified by Ayers and Westcot in 1985, and are being internationally accepted.

In the present study, Cl was very high in both fish farm effluent (2220 mg/l) and dairy farm effluent (2000 mg/l), while in farm water (950 mg/l) was high (Table 3). Although chloride is essential to plants in very low amounts, but it can cause toxicity to crops at high concentrations.

In addition, Na was very high in both fish farm effluent (1420 mg/l) and dairy farm effluent (1310 mg/l), while in farm water (481 mg/l) could be considered as a medium amount of Na in irrigation water (Table 3). Sodic water is not the same as saline water. The application of water with a high sodium content can reduce yield.

Consequently, in this study, the Cl and Na contents in both fish and dairy farm effluents were high to very high. Thus, the long term irrigation with such types of water may have a negative effect on soil and subsequently decreased growth and production of date palm.

In the present study, the total N was high in fish farm effluent (35 mg/l), followed by dairy farm effluent (15 mg/l). However, N was absent in farm water (Table 3).

The highest amount of total P (2.1 mg/l) was found in fish farm effluent, followed by dairy farm effluent (0.35 mg/l), while the lowest (0.04 mg/l) was found in farm water (Table 3).

Similarly, K was very high in both fish farm effluent (118 mg/l) and dairy farm effluent (110 mg/l), while the lowest (30 mg/l) was found in farm water (Table 3).

Because NPK were major nutrient elements in plant production and used extensively in date palm fertilization, the amount of NPK contents in 100 m3 of water/tree/year of all types of water used in the experiment was calculated by divided the values of NPK in Table 3 by 10. The converted values are presented in Table 4.

Fish and dairy farm effluents were rich in nutrient elements. The present study recorded the highest amounts of NPK in irrigation water with fish farm effluent, followed by dairy farm effluent. However, the lowest was recorded with farm water control (Table 3). These considerable amounts of NPK in fish and dairy farm effluents matched the yearly needs for date palm trees (AL-Rawi, 1998).

Table 3:

Some chemical analyses of farm water, fish farm effluent, and dairy farm effluent used in the experiment.

No.	Test	Farm water	Fish farm effluent	Dairy farm effluent
1	pH	8.11	7.88	7.43
2	EC (dS/m)	2.89	6.85	6.41
3	Cl (mg/l)	950	2220	2000
4	Na (mg/l)	481	1420	1310

No.	Test	Farm water	Fish farm effluent	Dairy farm effluent
5	Total N (mg/l)	0	35	15
6	Total P (mg/l)	.04	2.10	.35
7	K (mg/l)	30	118	110

Table 4:

Amount of NPK contents in 100 m3 of water/tree/year in all types of water used in the experiment and annual amount of nutrient requirement (kg/adult date palm tree).

Nutrient	Farm water (kg/100 m^3 of water)	Fish farm effluent (kg/100 m^3 of water)	Dairy farm effluent (kg/100 m^3 of water)	The standard amount of nutrient requirement (kg/tree/year)*
Total N	0	3.5	1.5	1.5–3.0
Total P	0.004	0.210	0.035	0.5
K	3.0	11.8	11.0	2-3

*Generally, 1.5–3.0 kg of nitrogen, 0.5 kg of phosphorus and 2.0–3.0 kg of potassium per tree yearly is recommended to maintain optimum growth of date palm tree (AL-Rawi, 1998).

Growth, yield, and quality:

In the season of 2019, there were no significant differences in the number of leaves with thorn per tree and yield (kg/tree). The highest significance fruit weight (16.40g) was recorded with farm water on par with dairy farm effluent (13.34g). However, in the 2020 season, the highest number of leaves with thorn per tree (26.5), yield (67 kg/tree), and fruit weight (18.78g) were recorded with fish farm effluent (Table 5).

Table 5.

show the effect of different quality of irrigated water on the number of leaves with thorn per tree, yield (kg/tree), and fruit weight (g) of date palm cv. Medjool, in 2019 and 2020.

Irrigation treatment	Number of leaves with thorn per tree		Yield (kgltree)		Fruit weight (g)	
	2019	2020	2019	2020	2019	2020
Farm water (control)	* 20.5 a	b 21.3	60.78 a	60.5 b	16.40 a	15.88 b
Fish farm effluent	21.5 a	26.5 a	59.66 a	67.0 a	12.28 b	18.78 a
Dairy farm effluent	19.3 a	23.5 b	50.04 a	61.8 b	13.34 ab	16.58 b

*Means followed by similar letters in each column are not significantly different according to the Fisher LSD at $p \leq .05$.

These results may be attributed to the fact that fish farm effluent containing high amounts of nutrients NPK (Table 3), which produced the highest number of leaves with thorn per tree (Elnwishy, 2008). This directly or indirectly has a positive effect on fruit weight and yield. Similar results were reported in other crops like tomato (Castro et al., 2006), potato (Abdelraouf and Hoballah, 2014), cucumber (Buang et al., 2018), maize and bean (Silva et al., 2018), and basil and purslane (Kaab Omeir et al., 2020) who reported improved in crop productivity and reduced in the total costs of fertilizers.

In both seasons, no significant differences were recorded among treatments for Balah weight (Table 6). Balah fruits, yellow fruits, or khalal

fruits of Medjoul variety (not like Barhi variety) were not palatable because of astringency taste (Alsmairat et al., 2013). Therefore, the less Balah weight per tree is the better. In this experiment, the weights of Balah fruits (kg/tree) in the whole experiment were relatively low and generally, these weights were accepted as postharvest losses in Medjoul production.

In the 2019 season, the highest significant fruit with a stalk (7.58 kg/tree) was recorded with fish farm effluent, which was on par with farm water (4.91 kg/tree), while in the 2020 season, no significant differences were recorded among the treatments (Table 6).

In the 2019 season, no significant differences were recorded among treatments for fruit without a stalk. In contrast, in the 2020 season, the highest significant fruit without a stalk (28.5 kg/tree) was recorded with fish farm effluent (Table 6).

It is well known that fruits with stalks are better than without stalks because of the fruits' good appearance, and the shelf life of the fruits will be extended because the harmful microbes cannot easily enter inside the fruits due to fewer opens on the surface of the fruits (Kader and Hussein, 2009). Unfortunately, in the whole experiment, the weights of fruits with a stalk were much less than those without a stalk (Table 6).

Table 6:

Effect of different quality of irrigated water on Balah fruits (kg/tree) and fruit with and without a stalk (kg/tree) of date palm cv. Medjool, in 2019 and 2020.

Irrigation treatment	Balah fruits (kg/tree)		Fruit with stalk (kg/tree)		Fruit without stalk (kg/tree)	
	2019	2020	2019	2020	2019	2020
Farm water (control)	1.49 a*	2.18 a	4.91 ab	6.90 a	25.23 a	24.50 b
Fish farm effluent	3.44a	2.43 a	7.58 a	7.90 a	24.40 a	28.50 a
Dairy farm effluent	2.03 a	2.18 a	3.60 b	7.50 a	19.39 a	25.50 b

*Means followed by similar letters in each column are not significantly different according to the Fisher LSD at $p \leq .05$.

Skin separation:

The less air content in date palm fruit means the better fruit quality. Fortunately, in the whole experiment, the weights of fruits in this category (from 0-10 %) were higher than the weights for the other categories (Table 7).

In the season of 2020, there were significant differences among the treatments in the category from 11-30 (%) which recorded the lowest significant skin separation with the farm water (5.05 kg/tree). Similarly, there were significant differences in the season of 2020 in the category from 31-50 (%) which recorded the lowest significant skin separation with the farm water (3.43 kg/tree). This treatment was on par with dairy farm effluent (4.10 kg/tree) (Table 7).

Moreover, there were significant differences among the treatments in the category > 50 (%) in the season of 2019 which recorded the lowest significant skin separation with the dairy farm effluent (0.95 kg/tree). This treatment was on par with farm water (1.56 kg/tree) (Table 7).

The phenomenon of skin separation occurs in fruits of many date palm cultivars (Gophen, 2014), it is crucial in Medjool (Lustig et al., 2014), and the price for fruit with skin separation is only one-half of similar fruit without it (Cohen and Glanser, 2015). Thus, improving the quality of the fruits is extremely important, and the elimination of skin separation in 'Medjool' fruits is a significant concern for date producers and the economy of this variety (Kader and Hussein, 2009).

Table 7:

Effect of different quality of irrigated water on the air content (%) in the fruit of date palm cv. Medjool, in 2019 and 2020.

Irrigation treatment	Air content (%) in fruit							
	0-10 (%)		11-30 (%)		31-50 (%)		> 50 (%)	
	2019	2020	2019	2020	2019	2020	2019	2020
Farm water (control)	22.38 a*	25.75 a	3.75 a	5.05 b	2.28 a	3.43 b	1.56 ab	1.68 a
Fish farm effluent	22.59 a	27.50 a	4.02 a	6.28 a	2.41 a	4.40 a	2.64 a	2.05 a
Dairy farm effluent	18.35 a	25.25 a	2.82 a	6.15 a	1.82 a	4.10 ab	0.95 b	1.71 a

*Means followed by similar letters in each column are not significantly different according to the Fisher LSD at $p \leq 0.05$.

Conclusions

It could be concluded from this study that the most effective treatment as an alternative source of irrigation water was fish farm effluent. No harmful microbial (*E. coli*) present in the three irrigation water treatments. The EC in both fish and dairy farm effluents was high, while the EC in farm water was accepted in date palm plantation. Fish and dairy farm effluents were rich in nitrogen, potassium, and phosphorus that could cover the annual requirements of NPK for the date palm tree. Thus, fish and dairy farm effluents could be used as a source of irrigation and fertilizer in modern date palm cultivation. However, long run investigation is needed to study the impact of using such water sources on soil and plant.

Acknowledgement

The authors would like to thank the Deanship of Scientific Research at Al-Quds Open University and the Palestinian Ministry of Higher Education and Scientific Research for the financial support of the present research project. Special thanks to the Al-Wadi farm (Daiq Brothers) for their great help in the accomplishment of this research in their farm. Great thanks for the Arab Development Society-Jericho for providing fish farm effluent and for the Al-Juneidi Dairy and Food Products Company for providing dairy farm effluent.

References:

- Abdelraouf, R.E. (2017). Reuse of fish farm drainage water in irrigation. In: The handbook of environmental chemistry. Springer, Berlin, Germany.
- Abdelraouf, R.E. & Hoballah, E.A. (2014). Impact of irrigation systems, fertigation rates and using drainage water of fish farms in irrigation of potato under arid regions conditions. International Journal of Scientific Research in Agricultural Sciences, 1(5), 67-79.
- Abu-Jaish, A.A. (2018). Determinants of the expansion of palm growing within the natural resources (land, water) available for the development of the agricultural sector in the Jordan Valley region. M.Sc. Thesis in Sustainable Rural Development, al-Quds University in Abu Dees, Jerusalem, Palestine.
- Abu-Qaoud, H. (2015). Date palm status and perspective in Palestine. In J.M. Al-Khayri, S.M. Jain, & D.V. Johnson. Date Palm Genetic Resources and Utilization, eds.: Volume 2: Asia and Europe, (Springer Press), p.423-439.
- Ahmed, N. & Garnett, S.T. (2011). Integrated rice-fish farming in Bangladesh: meeting the challenges of food security. Food Security, 3(1), 81-92.
- Al-Alawi, R.A., Al-Mashiqri, J.H., Al-Nadabi, J.S., Al-Shihi, B.I., & Baqi, Y. (2017). Date palm tree (*Phoenix dactylifera* L.): natural products and therapeutic options. Frontiers in Plant Science, 8, 845. https://doi.org/ 10.3389/fpls.2017.00845.
- Al-Rawi, A.A.H. (1998). Fertilization of date palm tree (*Phoenix dactylifera* L.) in Iraq. In: Proceedings of the 1st international conference on date palm, 8–9 Mar 1998, Al-Ain Hilton Hotel, Al-Ain, United Arab Emirates. pp 320–327.
- Alsmairat, N.G., El-Assi, N.M., Al-Abdallat, A.M. & Mehyar, G.F. (2013). Enhancement of edibility of" Barhi" and" Medjool" date palm cultivars at khalal mature stage. International Journal of Botany, 9(3), 123-132.
- Ayers, R.S. & Westcot, D.W. (1985). Water quality for agriculture. FAO Irrigation and drainage paper 29 Rev. 1. Food and Agricultural Organization. Rome, 1, 74.
- Barkle, G.F., Stenger, R., Singleton, P.L. & Painter, D.J. (2000). Effect of regular irrigation with dairy farm effluent on soil organic matter and soil microbial biomass. Australian Journal of Soil Research, 38(6),

1087-1097.

- Bauder, T.A., Waskom, R.M., Sutherland, P.L. & Davis, J.G. (2011). Irrigation water quality criteria (Doctoral dissertation, Colorado State University. Libraries).
- Buang, E.A., Nornasuha, Y., Nashriyah, M. & Khandaker, M.M. (2018). Effects of fish waste effluent on the growth, yield and quality of Cucumis sativus L. J. Agrobiotech., 9, 258–267.
- Castro, R.C., Azevedo, C.B. & Neto, F.B. (2006). Increasing cherry tomato yield using fish effluent as irrigation water in Northeast Brazil. Scientia Horticulturae, 110(1), 44-50.
- Cohen, Y. & Glanser, B.A.B. (2015). Date palm status and perspective in Israel. In J.M. Al-Khayri, S.M. Jain and D.V. Johnson. Date Palm Genetic Resources and Utilization, eds.: Volume 2: Asia and Europe, (Springer Press), p.265-298.
- Daiq, I. (2020). International date palm expert. Personal communication, Jericho, Palestine.
- Dudeen, B., Lacirignola, C., Montanarella, L., Steduto, P. & Zdruli, P. (2001). The soils of Palestine (The West Bank and Gaza Strip) current status and future perspectives. Soil Resources of Southern and Eastern Mediterranean Countries, 203-233.
- East, E.M. (2019). Climate change, water security, and national security for Jordan, Palestine, and Israel.
- Elnwishy, N.H. (2008). Effectiveness of fish culture implementation to improve irrigation water quality. International Journal of Agriculture and Biology, 10(5), 591-592.
- Fewtrell, L. & Bartram, J. (2001). Water quality: guidelines, standards & health. IWA publishing.
- Fipps, G. (2003). Irrigation water quality standards and salinity management strategies. Texas FARMER Collection.
- Fresenius, W., Quentin, K.E. & Schneider, W. (1988). Water analysis. A practical guide to physico-chemical, chemical and microbiological water examination and quality assurance.
- Gophen, M. (2014). Skin separation in date fruits. International Journal of Plant Research, 4(1), 11-16.
- Hawke, R.M. & Summers, S.A. (2006). Effects of land application of farm dairy effluent on soil properties: A literature review. New Zealand Journal of Agricultural Research, 49(3), 307-320.
- Islam, M.T. (2015). Effects of dairy farm's wastewater irrigation on wheat production and soil health. M.Sc. Thesis in Irrigation and Water Management, Bangladesh Agricultural University, Mymensingh, Bangladesh.
- Islam, M.S., Islam, M.T., Hossain, S.A., Adham, A.K.M. & Islam, D. (2017). Impacts of dairy farm's wastewater irrigation on growth and yield attributes of maize. Fundamental and Applied Agriculture, 2(2), 247-255.
- Kaab Omeir, M., Jafari, A., Shirmardi, M., & Roosta, H. (2020). Effects of irrigation with fish farm effluent on nutrient content of Basil and Purslane. Proceedings of the National Academy of Sciences, India Section B: Biological Sciences, 90(4), 825-831.
- Kader, A.A. & Hussein, A.M. (2009). Harvesting and postharvest handling of dates, ICARDA, Aleppo, Syria, 15p.
- Khater, E.G. (2006). Aquaponics: the integration of fish and vegetable culture in recirculating systems. M.Sc. Thesis, in Agric. Eng., Fac. Agric., Moshtohor, Benha Univ., Egypt.
- Little, T.M. & Hills, F.J. (1978). Agricultural Experimentation. John Wiley and Sons, Inc., New York, USA.
- Longhurst, R.D., Roberts, A.H.C. & O'Connor, M.B. (2000). Farm dairy effluent: A review of published data on chemical and physical characteristics in New Zealand. New Zealand Journal of Agricultural Research, 43(1), 7-14.
- Lustig, I., Bernstein, Z. & Gophen, M. (2014). Skin separation in Majhul fruits. International Journal of Plant Research, 4(1), 29-35.
- Mazahrih, N.T., Al-Zubi, Y., Ghnaim, H., Lababdeh, L., Ghananeem, M. & Ahmadeh, H.A. (2012). Determination actual evapotranspiration and crop coefficients of date palm trees (*Phoenix dactylifera* L.) in the Jordan valley. American-Eurasian J. Agric. & Environ. Sci., 12 (4), 434-443.
- Palestinian Meteorological Department (2020). Climatic data for Jericho. https://www.pmd.ps/en
- Roufou, S., Griffin, S., Katsini, L., Polańska, M., Van Impe, J.F. & Valdramidis, V.P. (2021). The (potential) impact of seasonality and climate change on the physicochemical and microbial properties of dairy waste and its management. Trends in Food Science & Technology, 116, 1-10.
- Russo, G., Dalla Ragione, I., Fantinati, L., Simonetti, C. & Lauteri, M. (2017, November). A socio-ecological analysis of the date palm (*Phoenix dactylifera* L.) system in Jericho Area, Palestine. In Euro-Mediterranean Conference for Environmental Integration (pp. 1425-1427). Springer, Cham.
- SAS Institute (1990). ''User's guide. Statistics. Ver. 6.4th ed. SAS Inst. Cary. N. C, 956.
- Silva, E.F.L., Botelho, H.A., Venceslau, A.D.F.A. & Magalhães, D.S. (2018). Fish farming effluent application in the development and growth of maize and bean plants. Científica, 46(1), 74-81.
- World Health Organization (2006). WHO guidelines for the safe use of wastewater excreta and greywater (Vol. 1). World Health Organization.

Forecasting COVID-19 Confirmed Cases Using Time Series Analysis

التنبؤ بعدد الحالات المؤكدة لكوفيد - 19 باستخدام تحليل السلاسل الزمنية

Akram Mohammed Radwan
Assistant Professor\ University College of Applied Sciences\ Palestine
aradwan@ucas.edu.ps

أكرم محمد رضوان
أستاذ مساعد/ الكلية الجامعية للعلوم التطبيقية / فلسطين

Received: 10/ 1/ 2022, **Accepted:** 14/ 11/ 2022
DOI: 10.33977/2106-000-006-002
http://journals.qou.edu/index.php/PJTAS

تاريخ الاستلام: 10/ 1/ 2022م، **تاريخ القبول:** 14/ 11/ 2022م.
E - ISSN: 2521 - 411X
P - ISSN: 2520 - 7431

Abstract:

The novel coronavirus (COVID-19) pandemic is a major global health threat that is spreading very fast around the world. In the current study, we present a new forecasting model to estimate the number of confirmed cases of COVID-19 in the next two weeks based on the previously confirmed cases recorded for 62 countries around the world. The cumulative cases of these countries represent about 95% of the total global up to the date of data gathering. Seven regression models have been used for three rounds of predictions based on the data collected between February 21, 2020 and December 29, 2020. A number of different time series features have generated using feature-engineering methods to convert a time series forecast into a supervised learning problem and then build regression models. The performance of the models was evaluated using root mean squared log error, root mean squared error, mean absolute error, mean absolute percentage error, coefficient of determination and running time. The findings show a good performance and can reduce the error about 72% with a high coefficient of R2 = 0.990. In particular, XGB and Random Forest models have demonstrated their efficiency over other models.

Keywords: COVID-19, predictive analytics, machine learning, regression, time series.

الملخص:

تعد جائحة كورونا (COVID-19) تهديدًا صحياً عالمياً رئيسياً انتشر بسرعة كبيرة في جميع أنحاء العالم. في الدراسة الحالية، قدمنا نموذجاً للتنبؤ لتقدير عدد الحالات المؤكدة لكوفيد- 19 في الأسبوعين المقبلين بناءً على أعداد الحالات المؤكدة مسبقاً التي سجلت في 62 دولة حول العالم. تمثل الحالات التراكمية لتلك الدول حوالي 95 % من الإجمالي العالمي حتى تاريخ جمع البيانات. تم استخدام سبع خوارزميات انحدار لثلاث جولات من التنبؤات بناءً على البيانات التي تم جمعها في الفترة ما بين 21 فبراير 2020 و29 ديسمبر 2020. تم استخراج عدد من ميزات السلاسل الزمنية باستخدام أساليب هندسة الميزات لتحويل التنبؤ بالسلاسل الزمنية إلى مسألة تعلم آلي خاضع للإشراف، ثم باستخدام الميزات الأكثر أهمية تم بناء نماذج الانحدار لعمل التنبؤ المطلوب. تم تقييم أداء النماذج باستخدام المقاييس التالية: جذر متوسط الخطأ اللوغاريتمي التربيعي، جذر متوسط الخطأ التربيعي، متوسط الخطأ المطلق، متوسط نسبة الخطأ المطلق، معامل التحديد (مربع معامل الارتباط) ووقت التنفيذ. أظهرت نتائج هذا البحث أن أداء نماذج الانحدار كانت جيدة، واستطاعت تقليل الخطأ بنسبة 72 % مع معامل تحديد عالٍ R2 وصل لـ 990. على وجه الخصوص، أظهرت كلّ من خوارزميات XGB و Random Forest كفاءة أعلى في الأداء مقارنة مع الخوارزميات الأخرى.

الكلمات المفتاحية: كوفيد- 19، التحليل التنبؤي، التعلم الآلي، الانحدار، السلاسل الزمنية.

1. Introduction

The new coronavirus disease (COVID-19) is an infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV2). This epidemic is spreading very quickly all over the world, affecting more than 200 countries. The World Health Organization (WHO) declared it as a global pandemic. According to WHO, 82,679,350 confirmed cases have been recorded and 1,872,994 deaths have been reported till the end of December 2020[(1)] globally. The spread of Covid-19 is very dangerous

requiring more strict policies, and plans that aid the healthcare service preparation, which have already been implemented in many countries around the world. Thus, it is very vital to forecast the confirmed cases in the upcoming days to support the prevention of outbreak of Covid-19 pandemic and to prepare against possible threats (Oliveira and Moral, 2021).

In the last year, numerous studies have addressed forecasting the number of confirmed cases of Covid-19. Various mathematical methods, time series models and machine learning (ML) techniques have been proposed to estimate the future trend of pandemic Covid-19 (Ahmad et al., 2020) and (Vytla et al., 2021). A few examples of these methods are Multiple Linear Regression (Rath et al., 2020), Bayesian Network, Auto-Regressive Integrated Moving Average (ARIMA)

(Hernandez-Matamoros et al., 2020), Deep learning via Long Short-Term Memory (LSTM) (Chowdhury et al., 2021), SEIR model (Feng et al., 2021), Adaptive Neuro-Fuzzy Inference System (ANFIS) (Chowdhury et al., 2021), and Simulation models (Hassanat et al., 2021).

Time series forecasting is a method to predict future values based on previously observed values using temporal features. This method has been studied widely in Covid-19 forecasting. Using historical time series data, we can forecast the number of daily new confirmed cases in the next days (Oliveira & Moral, 2021; Ahmad et al., 2020).

There exist a large number of evidences where regression algorithms have proven to give efficient predictions the Covid-19 prevalence in many countries (Ahmad et al., 2020; Rath et al., 2020). The prediction based on regression methods has many approaches.

Many researches aimed to predict the prevalence of Covid-19 in one country or union of territories (Chowdhury et al., 2021; Al-Qaness et al. (2020; Samson et al., 2020; Ribeiro et al., 2020), however, our study handles the estimation of the confirmed cases in the most affected countries worldwide.

The paper is structured as follows. Section 2 provides a literature review. Section 3 describes feature sets, which generated from time series data. In section 4, we present data and the models used in this study. Next Section 5 covers experimental results and performance evaluation. Finally, the conclusions are summarized in Section 5.

2. Literature Review

During the last two years, several research papers were published that tackled the applications of machine learning techniques for the prediction of Covid-19. In this section, we briefly review state-of-the-art that are relevant to our work.

Pandey et al. (2020) used linear and polynomial regression to predict the number of confirmed cases in India. They use data from January 30, 2022 to March 25. 2020 as the training data and predict the number of COVID-19 cases for next two weeks. The performance of the model was evaluated using RMSLE and achieved 1.75. Gu et al. (2020) applied cubic regression equations which use the number of days as the input variable to predict the confirmed cases in the whole of China except Hubei based on the existing data. Multiple Regression Analysis Models are suitable to fit the model and to predict the COVID-19 epidemic. Another research on the forecasting COVID-19 has been conducted by Rath et al. (2020). The authors used multiple linear regression to predict the next number daily active cases during the second week of August.

Gecili, Ziady and Szczesniak (2021) presented four different time series models for forecasting the numbers of confirmed cases, deaths and recoveries of COVID-19 for both the USA and Italy based on the daily reported data covered the period from February 22, 2020 until April 29, 2020. The performance of these models was evaluated using mean absolute error (MAE) and mean absolute percentage error (MAPE). The ARIMA model, is useful and powerful in time series analysis, was the most consistent across the other models and it had smaller prediction errors and narrower prediction intervals.

The study of Wang et al. (2022) proposed the ARIMA, SARIMA and Prophet models to predict daily new cases and cumulative confirmed cases in the USA, Brazil and India over the next 30 days based on the time series data from May 1, 2020 to November 30, 2021. The performance of different models was evaluated by using the root mean square error (RMSE), MAE and MAPE. The experimental results showed that the Prophet's model is more suitable for daily new cases of the USA with large fluctuations and has its unique advantages compared with ARIMA model, which is better for predicting of new cumulative cases in Brazil and India.

The authors in (Ribeiro et al., 2020) implemented six ML models to forecast with 1, 3 and 6 days ahead the COVID-19 cumulative confirmed cases of the most affected states of Brazil. They are ARIMA, ridge regression (RR), cubist (CUBIST), random forest (RF), SVR and stacking-ensemble learning (SEL) model. In the SEL approach, the CUBIST regression, RF, RIDGE, and SVR models are adopted as a base-

learners and Gaussian process (GP) as a meta-regressor. Their developed models can generate accurate prediction with a reasonable error.

Al-Qaness et al. (2020) combined ANFIS with an enhanced Flower Pollination Algorithm (FPA) and Salp Swarm Algorithm (SSA) to optimize the parameters of the model. The enhanced model is used to estimate the number of confirmed cases of COVID-19 in the upcoming ten days based on the previously confirmed cases recorded in China. Their approach showed better performance in terms of MAPE, Root Mean Squared Relative Error (RMSRE), coefficient of determination (R2), and computing time.

The COVID19 data is time-series data that compiles the number of confirmed cases where the cases are increasing over time until it arrives at a certain peak curve. Deep learning techniques such as LSTM can handle the nonlinearity and complexity of COVID-19 time-series data. The researchers in (Chowdhury et al, 2021) have used ANFIS and LSTM to predict the newly infected cases in Bangladesh. They showed that LSTM works better on a scenario based model for Bangladesh with MAPE= 4.51, RMSE= 6.55 and Correlation Coefficient= 0.75. Generally, LSTM is preferable in predicting the long term where ARIMA is for the short term (Kibria et al., 2022).

3. Features Engineering

Feature extraction is the most critical step in designing an algorithm in order to achieve good performance (Abuzir et al., 2021). We have used feature engineering to transform a time series raw data into a supervised learning dataset for machine learning algorithms. It is one of the most effective ways to improve predictive models' performance. The process takes in one or more existing columns of raw data and converts it into many columns of new features. Extracting useful information can help with time series data forecasting(2). From melting data, we can generate a number of various time series features that can be useful to predict future value based on these features.

A. Lag Features

When we try to predict the confirmed cases for a country, the previous day's cases are significant to make a prediction. In other words, the value at day t is affected by the value at day *t-1*. The past time series values are known as lags, so for *t-1* is lag_{t-1} for *t-2* is , and so on. We created lag features for three days.

B. Difference feature

This diff feature computes the difference between the confirmed cases in the previous day and the day before, i.e.,$D_{t-1}=X_{t-1}-X_{t-2}$. We created diff feature features for three days.

C. Rolling Window Features

The rolling window feature for time series calculates some statistical and aggregate functions based on past values. The size of the rolling window m is defined as the time frame, which in our case is the number of days. We created rolling window mean of size 3 at day t-1, denoted by M_{t-1}, would be mean (X_{t-1},X_{t-2} ,X_{t-3}).

D. Z-Scores scaling

The Z-score is linearly transformed data value having a mean of zero and a standard deviation of one. Z-score, or standard score, is used for standardizing scores on the same scale by dividing a score's deviation by the standard deviation in by the formula:

$$Z_t = \frac{X_t - mean(X)}{Std(X)} \quad (1)$$

where be a set of confirmed cases for a country and X_t is confirmed cases at day t. A z-score can be zero, positive or negative. A negative score indicates a value less than the mean, and a positive score indicates a value greater than the mean. The standardization of the time series for each country reduces the differences between the confirmed cases. In our model, we computed the z-score value at day t-1, denoted by Z_{t-1}

E. Rank

Rank feature gets the data frame by ascending order with a maximum rank value, and equal values have the same rank. In our model, we computed the rank value at day t, denoted by

F. Cumulative maximum

Cmax feature finds the cumulative maximum

valu $Cmax_{t-1} = \max\{X_1, X_2, \dots, X_{t-1}\}$ day t-1.

G. Entropy

The concept of entropy in information theory measures the amount of uncertainty of a random variable X. The entropy in terms of X, with po:

$$H(X) = -\sum_{i=1}^{n} p(X_i) log_2\ p(X_i) \qquad (2)$$

where $p(X_i)$ is simply the frequentist probability of a confirmed cases for a country at day *i*. When applies the entropy feature, all rows in a dataset with zero cases were deleted and the number of samples was reduced.

For this study, feature set is designed to include the following extracted features: lag_{t-1}, lag_{t-2}, lag_{t-3}, D_{t-1}, D_{t-2}, D_{t-3}, M_{t-1}, Z_{t-1}, $Rank_t$, $Cmax_{t-1}$ and $H(X_{t-1})$.

4. Experimental Design and Methods

This section presents the description of the used data, regression models with their parameter settings and the performance measures.

4.1. Dataset

The data includes time series data tracking the number of people affected by Coronavirus worldwide. The employed dataset contains data on Covid-19 including new daily-confirmed cases and it covers the period 21th February 2020 to 29th December 2020. Data is categorized by country named conforming to the WHO. It covers 62 countries around the world that are the most affected countries worldwide. The cumulative Covid-19 cases of these countries represent 95% of the total global up to the date of data collection. The data used in this work were collected from the repository of the John Hopkins University Center for Systems Science and Engineering (CSSE)[3]. We evaluated the performance of the presented method using three datasets of daily Covid-19 confirmed cases. The first one is called DS1; it starts from February 21 and continues until June 25, 2020. The second one is called DS2; it starts from June 26 to August 31, 2020 whereas, the third is called DS3; it starts from the first of September to December 29, 2020. The raw data consists of samples; each records daily-confirmed cases for 126 days in DS1, 67 days in DS2 and 120 days in DS3 for each country. Table 1 depicts a sample of dataset DS2.

When we try to fit a regression model for each country, we faced a problem due to having a little data entries (number of days in a dataset), which is small and not enough to get good results. To encounter this problem, we have used melting data, which converts wide-format data with several measurement columns into long-format with much more rows. In this case, each row becomes: Country, Day, Confirmed cases and we have 7874 samples (in DS1), 3960 samples (in DS2) and 7259 samples (in DS3) to train and test the models.

4.2. Regression Models

Regression models are statistical sets of processes that are used to estimate or predict the target or dependent variable based on one or more independent variables. It is widely used when both dependent and independent variables are linearly or non-linearly related, and the target variable has a set of continuous values. In this section, a brief of popular prediction algorithms are described, which are employed in the data analysis and experimental results.

1) Decision Tree (DT)

DT solves the regression problem by transforming the data into tree representation. Each internal node of the tree denotes an attribute or feature and each leaf node denotes a class label. While DT requires less effort for data preparation

during pre-processing, it often involves higher time to train the model.

Table 1.

Sample of the dataset DS2.

Country	June 26	June 27	..	Aug 30	Aug 31
Afghanistan	276	165	..	19	3
Algeria	240	283	..	365	348
Argentina	2886	2401	..	7187	9309
⋮	⋮	⋮	..	⋮	⋮
Italy	255	175	..	1365	996
Japan	107	92	..	605	438
Kazakhstan	569	0	..	111	77
South Korea	51	62	..	248	235
⋮	⋮	⋮	..	⋮	⋮
UK	1381	634	..	1752	1415
Ukraine	1121	957	..	2179	2202
US	45255	42705	..	35337	34156

2) Random Forest (RF)

RF is a bagging ensemble model that combines the prediction of multiple decision trees to create a more accurate final prediction. The final prediction is computed by taking the mean of the individual decision-tree predictions. RF is a fast and robust learning method able to deal with the randomness of the time series (Breiman, 2001).

3) Gradient Boosting Regression (GBR)

GBR is a type of ensemble where additional trees are added at each stage to compensate the shortcoming of the existing weak learners. These models are generally employed where features are too heterogeneous. Gradient Boosting model is more robust to outliers than boosting algorithm (Gumaei et al., 2021). In our model of Gradient Boosting Regressor we have used Huber loss function in loss function parameter.

4) Extreme Gradient Boosting (XGB)

XGB is a tree-based model. It stacks many trees, each new tree attempting to reduce the error of the preceding ensemble. The main goal is to develop a strong predictor by combining many weak predictors. XGB is one of the most powerful regression algorithms with high speed and performance (Chen, 2016). It runs more than ten times faster than existing popular solutions on a single machine. XGB is an efficient and scalable implementation of GBR. Moreover, it is feasible to train on large datasets. XGB can also be used for time series prediction.

5) Light Gradient Boosting Machine (LGBM)

LGBM is a gradient boosting framework based on a decision tree algorithm. LGBM has faster training speed with lower memory usage compared to XGB (Ke et al., 2017). Moreover, it can handle the large size of data and support GPU learning. Even though both XGB and LGBM models follow Gradient Boosting, XGB grows tree level-wise and LGBM grows tree leaf-wise.

6) Support Vector Regression (SVR)

This model works similarly to SVM (Support Vector Machine), but is adapted to handle regression. SVR uses kernel function to calculate the similarity between two data points when dealing with the non-linear problem. SVR involves two parameters that should be tuned for the model to perform well; the regularization parameter (referred to C) and the error sensitivity parameter (referred to ε) (Drucker et al, 1997).

7) Stacking-ensemble learning (SEL)

Stacking Generalization is an ensemble learning technique to combine multiple regression models (base- learners) via a meta-regressor. The individual regression models are trained based on the complete training set; then, the meta-regressor is fitted based on the outputs of the individual regression models in the ensemble (Ribeiro & Coelho, 2020). The main advantage of the SEL is that this approach can improve the accuracy and additionally reduce error variance. For this study, we trained a stacking-ensemble model using DT and RF as base-learners and as the LGBM meta-regressor.

This study aims to assess the ability of

the regression models with a selected feature set to forecast the confirmed Covid-19 cases by comparing their performances. Fine tuning predictive model hyperparameters is a crucial step to find the best fit parameters that improve accuracy of the forecasted results. The choice of inappropriate parameters' values may result in a poor performance. The parameters' setting for the models used in our study is listed in Table 2. For SEL method, no need to tune the parameters since this method is a combination of the best regressions.

Table 2.

Parameters' setting

Algorithm	Parameters Setting
DT	max_depth=5
RF	n_estimators=1000, n_jobs=-1, random_state=0
GBR	n_estimators =300, max_depth= 4, min_samples_split= 2, learning_rate= 0.01
XGB	learning_rate=0.1, base_score=0.5, max_depth=3, min_child_weight=2, n_estimators=300
LGBM	num_leaves=10, learning_rate=0.1, n_estimators=100, reg_lambda=0.30
SVR	C=5.0, ε =0.2

4.3. Evaluation Criteria

To check the performance of th seven models used in this study, we use the following statistical measures:

$$\text{RMSLE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\log(\hat{y}_i + 1) - (\log(y_i + 1))^2} \quad (3)$$

where N is the number of data observations, is the actual count and is the predicted count

- $$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \quad (4)$$

- $$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \quad (5)$$

$$\text{MAPE} = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (6)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2} \quad (7)$$

The lowest value of RMSLE, RMSE, MAE, and MAPE refers to the best model while the higher value R2 indicates better correlation for the model.

RMSLE is less sensitive to outliers than other metrics[4]. RMSLE is preferable when there is a wide range in the target variables and targets having exponential growth, such as population counts. Therefore, we can rely more heavily on this metric.

4.4. Research Methodology

In Figure 1, the procedure of COVID-19 forecasting has been shown. In the first step, the real data of COVID-19 are collected for the analysis. After collecting the raw data, we use melt method to change a data-frame from wide to long format and then obtain a time series of data to work with. After that we extract a time series feature set from melting data, as explained in Section 2, and it is used as inputs of each model. Then, the resulting data is divided into a training and test sets. We assign the last two weeks of data as a test set in order to evaluate regression models. The comparative analysis of seven regression models with optimal hyper-parameters are done and the best prediction model is identified based on the prediction results. The performance of the models will be verified by comparing the predicted data with real data via different statistical measures, including RMSE, MAE, MAPE, RMSLE and R2.

All experiments are implemented using python and its libraries such as scikit-learn, numpy and pandas. We have performed our experiments in Intel Core i7 CPU clocked at 2.00 GHz, 16 GB RAM (Raschka, Patterson & Nolet, 2020).

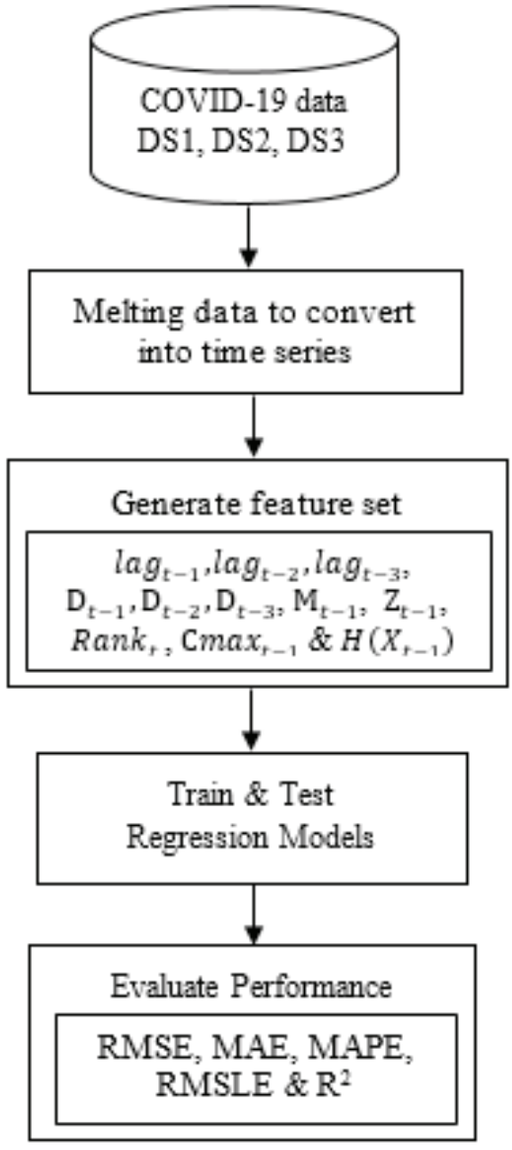


Figure.1

Proposed framework for forecasting

5. Results and Discussion

In this section, we described the performed experiments and discussed the obtained results. Our prediction results were obtained using seven regression models. Experiments were conducted over three rounds of forecasts where the first round was made for the two weeks from 12/06/2020 to 25/06/2020 based the data points available from 21/02/2020 to 11/06/2020 (training data for dataset DS1). The second round of the forecast was made for another two weeks from 18/08/2020 up end of Aug 2020 based on the actual data from 26/06/2020 to 17/08/2020 (training data for dataset DS2). Once more data became available, the third round was done for another two weeks from 16/12/2020 to 29/12/2020 based on the actual data from 01/09/2020 to 15/12/2020 (training data for dataset DS3). Raw melting data and extracted feature set, explained in Section 2, are used as inputs of each regression model.

The baseline prediction algorithm is used as the criterion by which all other regression models can be compared. The first row of experiments table results shows the baseline scores. Thus, if a model achieves a predictive score below the baseline, it is good. Whereas model with higher R^2 value indicates a good model.

Table 3 shows the predictive scores obtained by all regression models on three datasets, and the best performance is shaded with grey. Overall, all models except SVR demonstrated good performance. It can be concluded from Table 3, the RF and XGB outperformed the compared models in all measures. The results indicate the XGB is ranked first in terms of RMSLE and MAPE whereas the RF is ranked first in terms of RMSE and MAE. They are almost equal in term of R^2. However, the time computation (in seconds) undertaken by XGB is less than RF. Moreover, the scores produced by LGBM and SEL relatively well and they are ranked third after RF and XGB. Note that LGBM achieved optimal run time because it speeds up the training process and is almost seven times faster than XGB[(2)]. Thus, it is a much better technique of handling large datasets.

The value of R^2 indicates the correlation between the prediction obtained by the regression model and the actual Covid-19 confirmed cases. From Table 3, the high values of R^2, which are .961, .990 and .978 on DS1, DS2 and DS3 respectively, indicate the fitting goodness for predicting confirmed cases.

Table 4 demonstrates the average predictive scores achieved by baseline and the first ranked model in terms of RMSLE, RMSE, MAE and MAPE for three datasets. Based on these averages, we estimate error reductions, which are high in terms of RMSLE and MAPE; good in terms of RMSE and MAE.

Table 3.

Comparison results obtained by regression models on three datasets.

Dataset	Algorithm	RMSLE	RMSE	MAE	MAPE	R^2	Time (s)
DS1	Baseline	0.682	1947	483	78.96	0.831	--
	DT	0.366	1936	588	30.46	0.855	0.41
	RF	0.199	975	255	13.88	0.961	235.79
	GBR	0.275	2381	637	20.94	0.828	87.51
	XGB	0.159	1094	286	11.35	0.960	10.65
	LGBM	0.175	1367	342	12.56	0.938	1.55
	SVR	0.557	3159	825	91.68	0.694	21.40
	SEL	0.200	1104	301	14.32	0.956	209.79
DS2	Baseline	0.669	2587	737	43.26	0.925	--
	DT	0.421	1960	640	36.41	0.968	0.28
	RF	0.237	1014	324	16.30	0.989	123.99
	GBR	0.285	3246	871	22.62	0.918	48.86
	XGB	0.167	1093	331	11.48	0.990	6.24
	LGBM	0.232	1459	422	14.32	0.983	1.40
	SVR	0.659	4436	1193	94.39	0.819	4.93
	SEL	0.248	1194	356	16.64	0.986	124.23
DS3	Baseline	0.572	6295	1827	29.72	0.871	--
	DT	0.405	7353	2280	35.46	0.873	0.44
	RF	0.198	2846	931	14.13	0.978	254.23
	GBR	0.315	10915	2568	26.50	0.830	95.83
	XGB	0.158	3291	981	12.41	0.975	11.92
	LGBM	0.178	4616	1278	14.24	0.964	1.80
	SVR	0.690	13284	3323	102.77	0.668	18.97
	SEL	0.193	4419	1231	13.98	0.946	232.19

Table 4.

Error reduction.

Algorithm	RMSLE	RMSE	MAE	MAPE
Baseline	0.641	3610	1016	152
XGB	0.175			38.26
RF		1612	503	
Error reduction	72.7%	55.3%	50.5%	74.8%

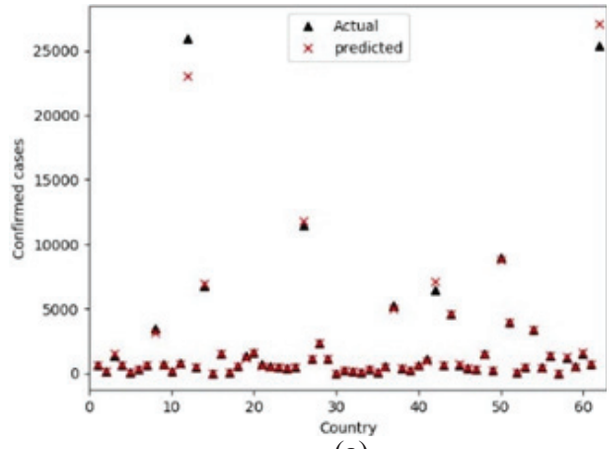
From the results, we can observe that the quality of our results is better than that appeared in recent previous studies. It can be easily seen that the RF model in our work outperformed the XGB and other models in Larabi-Marie-Sainte et al. (2022) for the four datasets (KSA, Brazil, Spain, and the US) in terms of RMSE and MAE. By comparing our estimates with those in Rguibi et al. (2022), we find that the prediction models in our study are better than ARIMA and LSTM models that used to predict the confirmed cases of Covid-19 in the upcoming two months in Morocco based on MAE and MAPE evaluation metrics. However, LSTM model showed lower prediction error values in term of RMSE and RMSLE. In such work, average values were RMSE = 795.3 and RMSLE = 0.00394. Another point to be highlighted is the comparison to the results obtained in the work of Wang et al. (2022), which tried to predict daily new cases in USA, Brazil and India over the next 30 days. We note that the prediction accuracies of our XGB and RF models are higher than that produced by SARIMA and Prophet models.

Since the best score was achieved by XGB, it was used to predict the Covid-19 daily-confirmed cases for the first day of the 2-weeks prediction. Figure 2 graphically compares the actual value with the predicted value of confirmed cases using XGB model for June 12, 2020, August 18, 2020 and December 16, 2020 respectively, where the x-axis represents the country, and the y-axis represents the corresponding daily-confirmed cases: actual (black) versus predicted (red). Figures 2 (a), (b) and (c) show that the dots representing the actual and predicted points are very close. However, we can observe that the XGB model gives some prediction values that are slightly far from the actual value, especially in DS1 and DS3, figure 2(a) and (c).

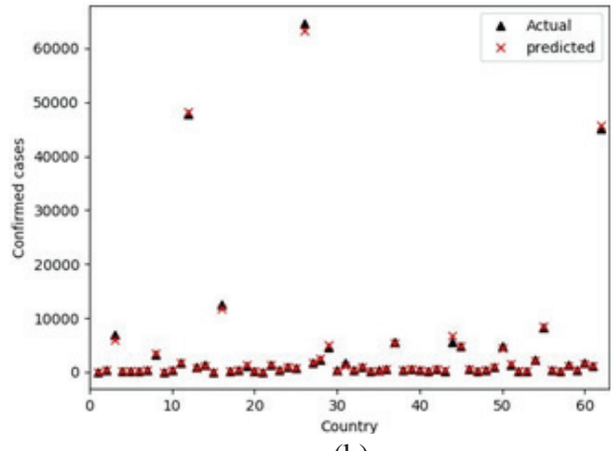
6. Conclusion

In this paper, we have conducted a three-round study of COVID-19 confirmed cases in the most affected countries worldwide. The most prominent techniques of regression models are used to analyze and predict the daily cases. The regression models are trained on a time series features, which are extracted through the original data. We have analyzed epidemic data made available by the CSSE within ten months to forecast the number of confirmed cases of COVID-19 for the next two weeks based on data available within enough time period before. The experimental results show that XGB and RF models produced good scores in terms of the five measures over three rounds and they may be appropriate for predicting the prevalence of COVID-19 in the future. Our analysis can help in understanding the trends of the pandemic outbreak.

There are some limitations in the forecasted numbers of COVID-19 cases. First, some countries have missing values for some days, so raw data record 0 values for these days and add missing cases for this day to next days. Therefore, the results for these countries are not accurate. Second, the prediction models rely on past behavior. Therefore, the existence of outliers and noise in the data make it hard to accurately predict the number of cases. Therefore, it is necessary to use noise filters to reduce noise›s effects. But this is missing in our study.



(a)



(b)

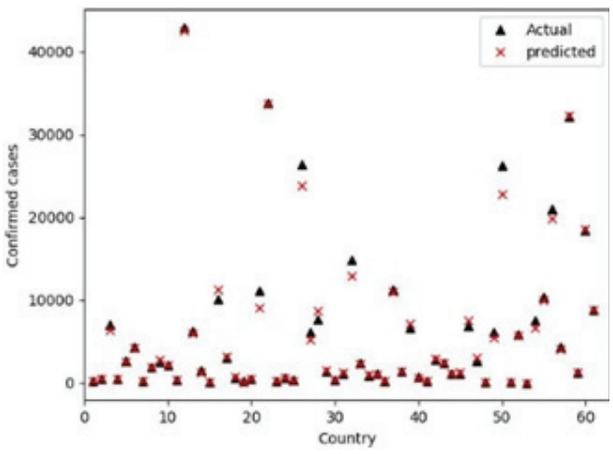


Figure 2.

Actual and predicted cases recorded for 62 countries in (a) June 12, 2020 (b) August 18, 2020 and (c) December 16, 2020

REFERENCES

- Abuzir, Y. S., Amro, I. Y., Tork, B., & Issa, M. (2021). A Prediction Model of Newly Admitted Students in the Level Exam Using Data Mining. Palestinian Journal of Technology and Applied Sciences (PJTAS), No 4, 23-35.
- Ahmad, A., Garhwal, S., Ray, S. K., Kumar, G., Malebary, S. J., & Barukab, O. M. (2021). The number of confirmed cases of covid-19 by using machine learning: Methods and challenges. Archives of Computational Methods in Engineering, 28(4), 2645-2653.
- Al-Qaness, M. A., Ewees, A. A., Fan, H., & Abd El Aziz, M. (2020). Optimization method for forecasting confirmed cases of COVID-19 in China. Journal of Clinical Medicine, 9(3), 674.
- Breiman, L. (2001). Random forests. machine learning, vol. 45, 5-32.
- Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).
- Chowdhury, A. A., Hasan, K. T., & Hoque, K. K. S. (2021). Analysis and prediction of COVID-19 pandemic in Bangladesh by using ANFIS and LSTM network. Cognitive Computation, 13(3), 761-770.
- Drucker, H., Burges, C. J., Kaufman, L., Smola, A., & Vapnik, V. (1996). Support vector regression machines. Advances in neural information processing systems, 9.
- Feng, S., Feng, Z., Ling, C., Chang, C., & Feng, Z. (2021). Prediction of the COVID-19 epidemic trends based on SEIR and AI models. PLoS One, 16(1), e0245101.
- Gecili, E., Ziady, A., & Szczesniak, R. D. (2021). Forecasting COVID-19 confirmed cases, deaths and recoveries: revisiting established time series modeling through novel applications for the USA and Italy. PloS one, 16(1), e0244173.
- Gu, C., Zhu, J., Sun, Y., Zhou, K., & Gu, J. (2020). The inflection point about COVID-19 may have passed. Science bulletin, 65(11), 865-867.
- Gumaei, A., Al-Rakhami, M., Al Rahhal, M. M., Albogamy, F. R., Al Maghayreh, E., & AlSalman, H. (2021). Prediction of COVID-19 confirmed cases using gradient boosting regression method. Comput Mater Continua, 66, 315-329.
- Hassanat, A. B., Mnasri, S., Aseeri, M. A., Alhazmi, K., Cheikhrouhou, O., Altarawneh, G., ... & Almoamari, H. (2021). A simulation model for forecasting covid-19 pandemic spread: Analytical results based on the current saudi covid-19 data. Sustainability, 13(9), 4888.
- Hernandez-Matamoros, A., Fujita, H., Hayashi, T., & Perez-Meana, H. (2020). Forecasting of COVID19 per regions using ARIMA models and polynomial functions. Applied soft computing, 96, 106610.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... & Liu, T. Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. Advances in neural information processing systems, 30.
- Kibria, H. B., Jyoti, O., & Matin, A. (2022). Forecasting the spread of the third wave of COVID-19 pandemic using time series analysis in Bangladesh. Informatics in Medicine Unlocked, 28, 100815.
- Larabi-Marie-Sainte, S., Alhalawani, S., Shaheen, S., Almustafa, K. M., Saba, T., Khan, F. N., & Rehman, A. (2022). Forecasting COVID19 parameters using time-series: KSA, USA, Spain, and Brazil comparative Case study. Heliyon, e09578.
- Oliveira, T. D. P., & Moral, R. D. A. (2021). Global short-term forecasting of COVID-19 cases. Scientific reports, 11(1), 1-9.
- Raschka, S., Patterson, J., & Nolet, C. (2020). Machine learning in python: Main developments and technology trends in data science, machine learning, and artificial intelligence. Information, 11(4), 193.
- Rath, S., Tripathy, A., & Tripathy, A. R. (2020). Prediction of new active cases of coronavirus disease (COVID-19) pandemic using multiple linear regression model. Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 14(5), 1467-1474.
- Ribeiro, M. H. D. M., & dos Santos Coelho, L. (2020). Ensemble approach based on bagging, boosting and stacking for short-term prediction in agribusiness time series. Applied Soft Computing, 86, 105837.
- Ribeiro, M. H. D. M., da Silva, R. G., Mariani, V. C., & dos Santos Coelho, L. (2020). Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil. Chaos, Solitons & Fractals, 135, 109853.
- Rguibi, M. A., Moussa, N., Madani, A., Aaroud, A., & Zine-Dine, K. (2022). Forecasting covid-19 transmission with arima and lstm techniques in

morocco. SN Computer Science, 3(2), 1-14.

- Samson, T. K., Ogunlaran, O. M., & Raimi, M. O. (2020). A Predictive Model for Confirmed Cases of COVID-19 in Nigeria. TK Samson, OM Ogunlaran, OM Raimi (2020), 1-10.
- Vytla, V., Ramakuri, S. K., Peddi, A., Srinivas, K. K., & Ragav, N. N. (2021, February). Mathematical models for predicting COVID-19 pandemic: a review. In Journal of Physics: Conference Series (Vol. 1797, No. 1, p. 012009). IOP Publishing.
- Wang, Y., Yan, Z., Wang, D., Yang, M., Li, Z., Gong, X., ... & Wang, Y. (2022). Prediction and analysis of COVID-19 daily new cases and cumulative cases: times series forecasting and machine learning models. BMC Infectious Diseases, 22(1), 1-12.

Endnote

1. https://covid19.who.int

2. https://machinelearningmastery.com/basic-feature-engineering-time-series-data-python

3. https://github.com/CSSEG ISand Data/COVID-19

4. https://medium.com/analytics-vidhya/root-mean-square-log-error-rmse-vs-rmlse-935c6cc1802a

5. https://neptune.ai/blog/xgboost-vs-lightgbm

Dynamic Energy-Efficient Routing Protocol for Wireless Sensor Networks

بروتوكول التوجيه الديناميكي الموفر للطاقة لشبكات الاستشعار اللاسلكية

Dalya Nouman Amro
Lecturer\ Al-Quds Open University\ Palestine
damro@qou.edu

Liana Jalal Al-Tamimi
Assistant Professor\ Palestine Polytechnic University\ Palestine
Liana_tamimi@ppu.edu

دالية نعمان عمرو
محاضر/ جامعة القدس المفتوحة/ فلسطين

ليانا جلال التميمي
أستاذ مساعد/ جامعة بوليتكنك فلسطين/ فلسطين

هذا البحث مستل من رسالة ماجستير.

Received: 15/ 1/ 2022, **Accepted:** 17/ 5/ 2022
DOI: 10.33977/2106-000-006-003
http://journals.qou.edu/index.php/PJTAS

تاريخ الاستلام: 15/ 1/ 2022م، **تاريخ القبول:** 17/ 5/ 2022م.
E - ISSN: 2521 - 411X
P - ISSN: 2520 - 7431

Abstract:

The devices in wireless networks operate on battery, so all routing protocols try to scale back the energy they consume while keeping acceptable performance for the tasks. For that, matching the routes to energy constraints to increase the network lifetime is what we are looking for, i.e, an efficient routing protocol.

Hence balancing the consumed energy among nodes within the networks is that the purpose of energy-efficient routing protocols. These protocols need to maximize the lifetime of the wireless network, and there are many algorithms to try to so like A-star algorithm or using mathematical formulation by changing the permittivity factor to a better value where we›ve high residual energy within the network and set it to a low value for the nodes that don›t have much energy left.

Our newly proposed protocol is trying to balance the vitality utilization to prolong the network lifetime using one between two protocols dynamically, which are DD and LEACH. The simulation results demonstrated that the newly proposed protocol is prolonging the WSNs lifetime.

Keywords: Wireless Sensor Network, Dynamic Energy-Efficient Routing, WSN Application, WSN Routing Protocols.

الملخص:

تعمل الأجهزة في الشبكات اللاسلكية على البطارية، لذا تحاول جميع بروتوكولات التوجيه تقليص الطاقة التي تستهلكها مع الحفاظ على أداء مقبول للمهام. لذلك، فإن مطابقة البروتوكولات مع قيود الطاقة لزيادة عمر الشبكة هو ما نبحث عنه؛ أي بروتوكول توجيه فعال.

ومن ثم فإن موازنة الطاقة المستهلكة بين العقد داخل الشبكات هي أن الغرض من بروتوكولات التوجيه الموفرة للطاقة. تحتاج هذه البروتوكولات إلى زيادة عمر الشبكة اللاسلكية إلى الحد الأقصى. وهناك العديد من الخوارزميات لمحاولة مثل خوارزمية A-star أو استخدام صياغة رياضية عن طريق تغيير عامل السماحية إلى قيمة أفضل حيث لدينا طاقة متبقية عالية داخل الشبكة وضبطها على قيمة منخفضة للعقد التي لا يتبقى لها الكثير من الطاقة.

يحاول بروتوكولنا المقترح حديثًا موازنة الاستخدام المرن لإطالة عمر الشبكة باستخدام واحد من بروتوكولين ديناميكيين، وهما DD و LEACH. أظهرت نتائج المحاكاة أن البروتوكول المقترح حديثًا يطيل عمر الشبكات اللاسلكية.

الكلمات المفتاحية: شبكة الاستشعار اللاسلكية، التوجيه الديناميكي الموفر للطاقة، تطبيق WSN، بروتوكولات التوجيه WSN

Introduction

As a day to day we communicate by transmitting data and sharing services. Network makes information and services obtainable to everyone on the network, notwithstanding the physical location of the resources or users.

Through wireless networks, devices and computers are connected by radio waves or the other wireless media, additionally , various wireless communication standards permit full mobility (Fahmy .(2016).

A Wireless Sensor Network (WSN) may be a group of specialized autonomous sensory and actuators according to Rawat, Singh, chaouchi, Bonnin.(2013).

With the potentiality of sensing, wireless communication and computations Guy (2006).

WSN may be a network of sensing devices connected together to the bottom station by wireless means (Springer) to watch, control physical or environmental conditions at diverse locations, and pass data to regulate the command to the specified actuators through the network(Springer).However, WSNs have their limitations that occur while using like cost, size, and limited power.

The proposed protocols and algorithms for WSNs have numerous problems to cope with. In addition, it is very complex to design routing protocols. The problems start with the energy dissipation, pocket loss rate, coverage and lifetime and they all are essential and need to be improved.

But the energy dissipation has a primary concern to the researchers due to the significance of the transmitted information at the WSN. For that, we need unique performance metrics to be optimized. Such a lot of algorithms had been proposed with trading off a few problems with the others related to the weighted importance to the problem they want to address according to Nabavi, Seyed Reza, Eraghi, Nafiseh Osati, Torkestani, Javad Akbari(2021), Engmann, Felicia, Katsriku, Ferdinand Apietu, Abdulai, Jamal-deen, Adu-manu, Kofi Sarpong (2020). Some of them took into consideration the strength efficient while others have got dead routes and nodes within the early stages of the WSN lifetime.

In our proposed protocol, we will address the energy dissipation by trying to solve the problem of limited power for the nodes.

The limited power problem causes issues within the consumption, which makes the network dies, before the time needed to transfer all the collected data. For this purpose, we try to prolong the network life time. That's why we proposed a replacement protocol that uses two different structures to function with two different protocols so as to save lots of the nodes energy.

The network uses the LEACH to optimize the network energy and use the advantages of the hierarchical characteristic.

The network is conserving the energy by operating the DD protocol on the lookup tables. Therefore, the nodes are on sleep mode till they›re required to work a replacement task of collecting information which is accomplished by LEACH.

The lookup table preserves the situation of the nodes with (X, Y) pairs and variety as a reputation to differentiate it from other nodes. Furthermore, it saves the initial energy (REC) for the network. The energy following every execution. Additionally, the efficient data aggregation (AGG) that utilizes the space and position of nodes to transmit data within the most effective way.

Those two parameters REC and AGG demonstrate the efficient usage of energy for our newly proposed protocol which extends the WSNs lifetime hence the facility source is that the most precious one for the WSNs.

Both the LEACH protocol and DD protocol are dynamically and alternatively utilized in our proposed protocol. The BS has the choice of what to work at a particular moment.

The simulation concludes that our dynamic protocol enhances the WSN lifetime with the share range of 20-25. (Ajit, Sunkara, Kumar (2013)) (Mardini, Khamayseh, AlZou'bi, Baniyassein. (2009).)

Related works

In the recent years, the importance of enhancing WSN has raised, and the researchers place a trial to propose brand-new algorithms to beat the disadvantages of the WSN. Since we have not had a tendency to use internet protocol (IP) address with WSN , this will cause a heavy load within the large areas. Additionally, it's exhausting for the restricted capabilities sensors to deal with the unpredictable topology changes particularly in mobile environments. The previous algorithms focused on prolonging the network lifetime ignoring the importance of quality of service (QoS) and that they have high energy consumption that makes the nodes to die resulting issues with the connectivity, and therefore the algorithms could have transmission errors, energy depletion additionally premature energy exhaustion because of heavy transmission-load. So we tend to think about enhancing the convergence and arrangement of the WSN to cope with the 4G and 5G networks that have a high bandwidth that make it expensive to apply widely (El-Esawy, Shimaa Gamal ,ElShennawy, Nada, Elfishawy, Nawal Ahmed(2018)). (Nabavi, Seyed Reza, Eraghi, Nafiseh Osati, Torkestani, Javad Akbari(2021)).

(Engmann, Felicia, Katsriku, Ferdinand Apietu, Abdulai, Jamal-deen, Adu-manu, Kofi Sarpong(2020)).

(Hassan, Ali Abdul-hussian, Shah, Wahidah MD, Habeb, Abdul-Hussien Hassan, Othman, Mohd Fairuz Iskandar ,AL-Mhiqani, Mohammed Nasser (2020) proposed an improved energy-efficient clustering protocol (IEECP) to prolong the time period of the WSN-based IoT. The planned protocol reduces and balances the energy consumption of nodes

by improving the clustering structure, where ver (IEECP) is appropriate for networks that need an extended lifetime. (Muzakkari, Bashir A, Mohamad, Mohamad A, Kadir, Mohd F. A, Mamat, Mustafa(2020)) presented an Energy Efficient and QoS-aware (EEQ) MAC protocol with a duty cycle scheme that adapts the node's duty cycle to the queue size and priority class of a packet to reduce the delay of high priority packets and support time bounded delivery of priority packets. This approach improves the energy efficiency and extending the lifespan of WSNs.

Bouzid, S.E., Serrestou Y, Raoof K, Omri, M.N. (2020) proposed a brand new routing protocol for WSN supported distributed Reinforcement Learning (RL). That ensures higher energy potency, postpones nodes death and isolation.

WSNs Energy-Efficient Routing Protocols

According to the networks, architecture routing protocols are generally classified as plane routing and level routing. Plane routing protocols like DD, SPIN and therefore the typical level routing protocols are LEACH, PAGASIS, and SPAN.

Therefore, in this thesis, we'll specialize in the LEACH and DD.

Irum, Nawaz, Murad, Saeed (2017), Nandy, Mitra. (2012), Salim. Ahmad, Osamy, Walid, Khedr. Amed M. (2014), Tandel (2016), Yulin, Xianguing. (2007). Pantazis, Nikolaos A. Nikolidakis, Stefanos A. Vergados, Dimitrios D. (2013). Hiremani, Nirmala, Basavaraju, T.G (2017). Elshakankiri, Maher N, Moustafa, Mohamed N, Dakroury, Yasser H.(2008).

Table 1:

Flat Protocols Characteristics.

Protocol	SPIN	DD
classification	flat	flat
mobility	supported	limited
Power management	limited	limited
Network life time	Good	Good
scalability	limited	limited
Resource awareness	Yes	Yes
Query based	Yes	Yes
Data delivery model	Event driven	Demand driven

Table 2:

Average Cluster Protocols Characteristics.

Protocol	LEACH	PEGA-SIS	SPAN
classification	clustering	clustering	clustering
mobility	Fixed base station	Fixed base station	Fixed base station
Power man-agement	maximum	maximum	limited
Network life time	Very good	Very good	good
scalability	Good	Good	limited
Resource awareness	Yes	Yes	Yes
Query based	No	No	No
Data delivery model	Cluster head	Chain based	continu-ously

Proposed Protocol

In this section, we discuss how the we created our dynamic protocol for the network, Ajit, Sunkara, Kumar (2013), Ganjali, McKeown(2005), Heidemanm, Estrin, Govindam, Intanagonwiwat (2003), Intanagonwiwat, (2002), Lingxi, Wenjun (2015), Xiao, Khatoon, Keshi (2013), Willig, Kaurl.

WSN Creation

The sensor nodes are positioned during a two-dimensional area of X*X meter, (N) of sensor nodes which are disseminated during a random manner with no mobility. Each node knows its location, the space to the BS, sink node and therefore the neighbor's location.

The network is partitioned into grids and each node belongs to a recognized location. The nodes took their position within the grids, memorized that position, the neighbor's position inside the grid they belong to and their assigned sink node

which will communicate and send the info to the BS as we will see in (Fig.1).

The elected sink nodes gather the collected data from the traditional nodes so as to pass it through the info mining algorithm to create the info pattern to be saved according to Zhn, Zhang, Yang, Wu (2016).

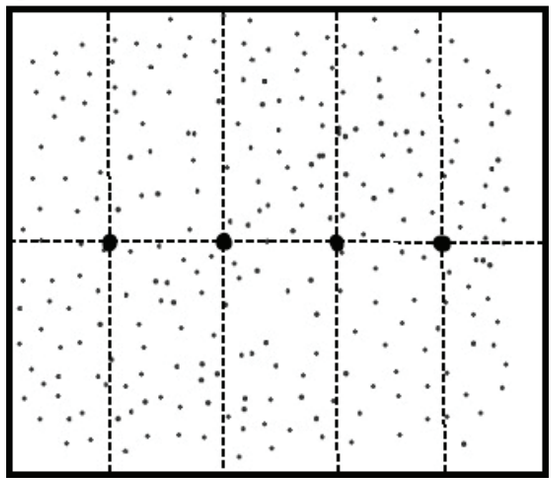


Figure 1: WSN Creation

WSN Route

The nodes choose the shortest path to send the info to the sink node, so it can send the collected data to the BS, which is seen in (Fig.2). The sink nodes communicate and send their locations with one another, and they know which one is that the closest to the BS.

The closer one is elected to speak with the BS on behalf of the entire network.

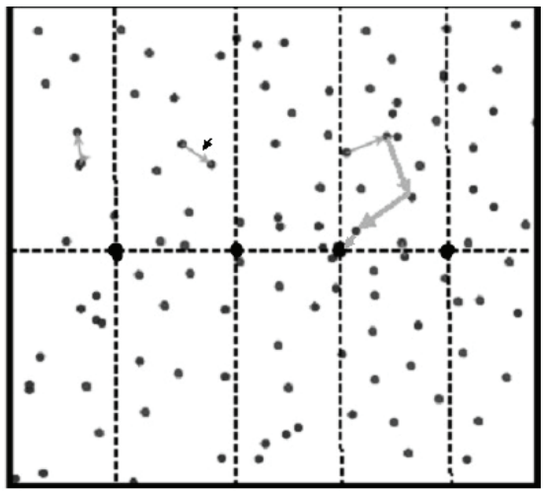


Figure 2: Best Route Finding

WSN Energy Efficient Protocol

The central server requests particular data from the BS then it searches for the data on the search for tables the use of DD protocol with VRS algorithm to look for the constructed patterns and the usage of the information mining algorithm with the decimal normalization. If the information is determined in the search for tables, the data is sent to the BS to the ultimate destination (central server) without looking the total network if not. The BS is sending the request to the sink node to operate LEACH protocol to experience and collect new information needed for the central server. The BS determines which sink node to send the statistics to, primarily based on the region which the BS knows and it has the data in it. In the network we have four sink nodes that send the sensed data (which is remarked with their location) from everyday nodes. Each node takes readings, which are arranged to be positioned in the lookup tables using data mining algorithm with decimal normalization to produce patterns of the data coming from the sink node to be dispatched and saved to the BS. After being divided into levels, we can search the area (X, Y) which has the readings the use of the precise range and location. The main content material in the search for tables are energy and location. The community grid structure has no impact on the constructed patterns. The tables are divided into corporations concurring to the records source location. LEACH protocol is used when we want to comprehend the total state of the network; hence, we want to collect data from all the nodes in a specific area belonging to one grid or many grids. Starting with the clustering, our protocol uses the clustering traits and the distance between nodes via solving the optimization hassle for the network. Thus, our protocol initializes a small world of sensors which are arranged randomly, and the adjacency matrix is developed by means of the usage of the distance formula.

D = [(X2 − X1)2 + (Y two − Y 1)2] (1/2) (1)

The distance formulation has been used to determine the superior number of chosen sink nodes with the aid of the usage of the clustering evaluation to share the resemblance of the centric communication function. Hence, to enforce it, the hierarchical clustering analysis is used to construct

a collection of sensor nodes from the lowest level to the absolute best with the consideration of the node's communication distance.

The cluster evaluation desires an appropriate quantity of sink nodes to be identified, which is scaled through the lifetime of the network. Sink node price and the premier power for the network. To attain that, we want a median node called key nodes linked without delay to the sink, which will supply the data from the sensor to sink. The data has to skip from (N-K)/K sensors. Where (N) is the whole network nodes number, while (K) demonstrating the range of key nodes that are linked with the BS. To restrict the increasing quantity of sink nodes we outline the community price as C=N.Cn+n.Cs and the fantastic variety of sinks will be computed by sink.

Table 3:

Nodes Communication Model.

Nodes	Nodes Task
Normal Node	Collect and sense data and send it to the sink
Sink Node	Communication between normal nodes and BS
BS	Request data and communicate with central server

Where (Cn) is the sensor node cost, whilst (Cs) is the BS cost, (Et) is the transmitted strength for communication per node pairs, (Er) is the obtained strength for communication per node pairs, (E0) power consumption while there is no verbal exchange between nodes. According to that, if we had two grids having the wished information the nearest one to the BS will do the verbal exchange process. In DD, The VRS has been used to ship the statistics to the region that wanted it, not the nodes that requested it. To achieve this, we will take care of the flooding problem that reasons overlapping and messages duplication. The computation for which protocol to be chosen, and the grid division is finished by way of the BS. In addition, BS will shop a look-up desk for the community conduct, so it doesn't have to do the computation all over again every time when the facts are needed from the nodes.

The BS of the network is the node with full energy, which is dividing the community into ten equal grids each and every sensor in these grids is gathering records in it till it is needed. Every node is aware of its area and distance to the BS, in LEACH protocol we have sink nodes that connect the regular nodes with the BS. Sink nodes speak at once with the BS. The packet is sent from the node to the sink node closer to the BS in the equal direction back and forth. The direction with the LEACH protocol is decided by using the optimization algorithm. The BS requires the wished data, By sending the request to the particular region that is aware of it can find the information required in it, and now not for the whole network which saves the energy for the unrelated nodes from recollecting and computation.

The VRS algorithm is operated and has a high-quality advantage. Those are not having a confined route discovery segment and an excessive tolerance for the rapid changes. If one node dies, the neighbors of this node are going to deal with the records request and communicate concurring to the time slots. The benefit of no longer having a restricted route path for the VRS is assisting with running the optimization algorithm used in LEACH to take care of the DD as well.

Table 4:

Transmitted Packets Contents.

Data Packet	Request Packet	Reply Packet	Error Packet
Packet Number	Packet Number	Packet Number	Packet Number
Source Location	Source Location	Source Location	Source Location
Destination Location			
Destination Location		Destination Location	
Destination Location			
Residual Energy	—	—	—
Protocol	Route Path	Route Path	Route Path
Data	Request	Reply	Error

When the network energy is suitable, the quantity of information is big (harvested from every sensor), and has a recognized location. We can use the LEACH protocol and use the optimization algorithm.

If the amount of statistics is small (in the look-up tables) we can use DD. The nodes gather the data and keep it until it is needed, the BS is going to ship a request according to the distance and place to a unique node and that is going to be the reply direction for the information to go returned to BS.

The optimization algorithm is used to measure the distance and the place for each node and VRS collects the information from nodes and shop them till needed. That's how the data is changed between the nodes and BS for the DD and LEACH. BS will take the statistics from the region for each node into consideration to supply the packet to the unique location, which wishes the information, which will give us so many advantages about power conservation due to the fact the nodes don't have to store the data tables for the routes.

The information packet arrival is primarily based on the nodes vicinity and the distance the usage of LEACH. Data series from all the cluster heads will be executed by way of CH, and it will check the redundancy of the information to recognize the passed off pattern, which is really useful for the sensor readings according to Dr. Arockiasamy, V.S.Anita (2011) and greve, gehlot, Sha (2012). This is focusing on energy conservation on WSN and with this way we make simply the cluster head communicates with the BS

S=U 1, U 2, ..Un

Computes the distance dj from Ui If (N (ui)) and distance dj from Ui Ci=Ui0, Ui1, ..Uin Do

Compute E(OiI) where I = zero to n

CHi is the cluster head of Ci if E(OiI) is excessive uni senses the records and passes the messages to CHi while(E(uiI)) >> (E(uij))

CHi passes the facts and aggregates it to the BS

The previous algorithm shows to the distance computation system for the sensor nodes S and that helps the nodes to follow the nearest CH then it computes the energy to determine which sensor is chosen to be the CH to communicate with the BS and aggregate the data.

The packets content material is the node location, distance between the node and the BS, the residual electricity and the direction is going to take it.

For the maintenance, if the BS sent a request whilst that node is useless and has no strength the network sends an error message and the closest neighbor node that has almost the equal location is going to deal with the request after sending the error message and for the new request or accepting the reply that is sent from the neighbor.

A network with (100) sensors are built, and those sensors are attending to collect data from their environment according to their location during a distributed way. However, this step is distributing the workload during a resilient way which will prolong the network lifetime.

Hence, the distributed way will control the quantity of communication between nodes and forbid heavy traffic from happening within the limited bandwidth of the wireless channels, which may be a source preserving. Because by this way, we control the power utilized in computation and communication process and to avoid any effects or unusual values in our information some normalization goes to be done on the data and decimal scaling is that the used technique by moving the decimal points.

To do so, distributed data algorithm is employed to cope with the nodes limitations

(power, computation, memory). The distributed algorithm goes to process the data locally then the results are aggregated. During this way, the energy for the communication process is reduced and at an equivalent time, the amount of messages during the process of data transfer to the central data is reduced.

The option of distributed data mining algorithm helps the WSNs to live longer and controls the massive data flow which will cause bottleneck and wastage of communication bandwidth because it‘s collecting an enormous data to analyze.

Initialize network

Stand by until receive request from server

IF {If data in BS}

search using DD in look up tables

else

Nodes take their positions

Compute cluster heads number

Compute the distance between nodes

Find the best route based on distance

Transmit data between nodes and the BS

Make patterns using data mining algorithm

Save in B using look up tables

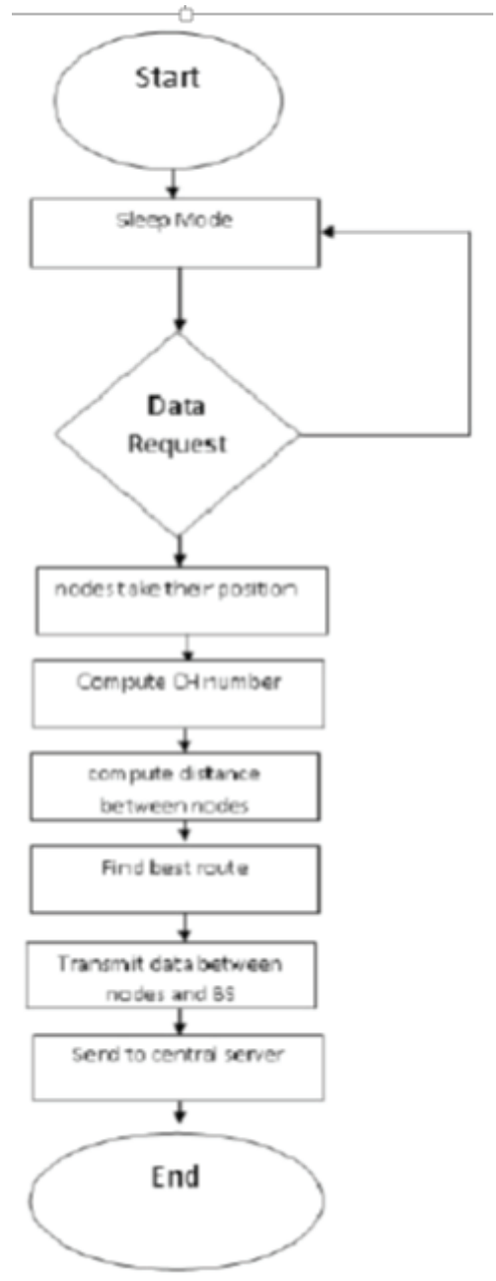
End if

Sending required data to server

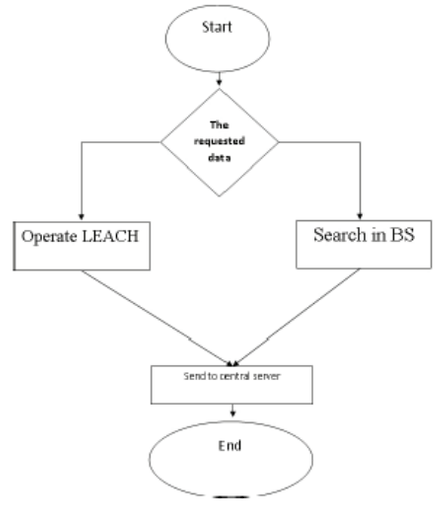
Go to step2

ENDIF

ENDIF



network initialization using LEACH



network flowchart

Assumptions

For this model, we made some assumptions as follows:

1. The BS is located at point (100,100) far from the sensors and it›s fixed.

2. The nodes are homogeneous and energy constrained.

3. There›s no mobility for sensor nodes.

4. The nodes can use power control to vary the quantity of power transmitted to reach the BS.

Our main concern is measuring the performance of WSN in terms of energy sufficiency. We are using metrics that have an influence on energy consumption. In WSN, the matrices are used to evaluate the performance of a WSN as an entire. These metrics are: [Tang, Tong. (2010), Verma, Sharma.(2013), Sunitha, Yadav(2014), Mardini, Khamayseh, A-lZou'bi, Baniyassein (2009).

1. ***End-to-End Delay***: it represents the time needed for a packet to be transmitted from source to destination.

2. ***Packet Delivery*** Rate: helps us to grasp the ratio of the delivered packets to their destination with the packets sent from the source.

3. ***Efficiency (Residual Energy):*** the energy of an isolated node is constant and independent of any changes occur within the system.

4. ***Network Throughput***: the quantity of data computed in any given period of time.

5. ***Packet Routing Load***: it represents the entire number of routing packets transmitted within the process of successful data transmission.

Simulation Environment

The proposed protocol has been tested to measure the metrics influence on the energy, to save lots of the energy and prolong the lifetime of the network. Within the following experiments, the nodes density has been fixed to 100 sensors and assumed there is no mobility within the network.

NS3

Our model simulation is completed by using NS3 simulator because of its advanced benefits over doing the real physical experiment within the world. Using simulation to visualize the connection between the parameters and seeing the interactions inside the network and therefore the way it deals with the obstacles and the normal operation within the communication process between the channels.

In the simulation, we had a dynamic environment to regulate our experiment and to create the network using mathematical equations and algorithm, which we›ve to adjust them if there›s a necessity to do so. Using dynamic environment like simulators helping to construct efficient implementation to the network model. Handling that is getting to help us understand the experiment and analysis easily. With the advantage of the low-cost experiment.

The used platforms within the simulators give us the graphs and therefore the results needed to deduce the new fact and helping to prove our model. The simulator helps you to verify and validate your model using its techniques.

In WSN, we›d like to use simulation to reinforce the model. Therefore, the simulator is constructing our model with accurate results. Network simulator version 3, is the tool that we are using to implement, which is an open source, discrete event network simulator. This has helped to model the actions of the dynamic compound system to offer a series of defined proceedings.

Results analysis and discussion

This section is related to discuss the results of the proposed protocol simulation compared to the opposite protocols.

Using five matrixes to get the energy consumption and therefore the QoS while not commerce off the two qualities

The obtained results show that the projected protocol has variations within the performance like considering message. Therefore the distance between nodes by emerging the benefits of the two protocols compared with.

This shows a significant enhancements over

the LEACH and VRS on their own and it provides us a sensible importance of the protocol.

The End-to-End Delay versus the Node Number

End-to-End delay means the packet time of arrival to their destination. This include the route discovery operation delay and therefore the packets transmission queue. The packets that are counted only have successfully reached the destination. We run the simulation to calculate the average delay for 100 (samples) nodes. Therefore, when the packet arrives to the destination we record the time of arrival to calculate the current delay. An equivalent operation is completed with every packet and that we increment the samples one by one then we will get the entire delay for the entire network.

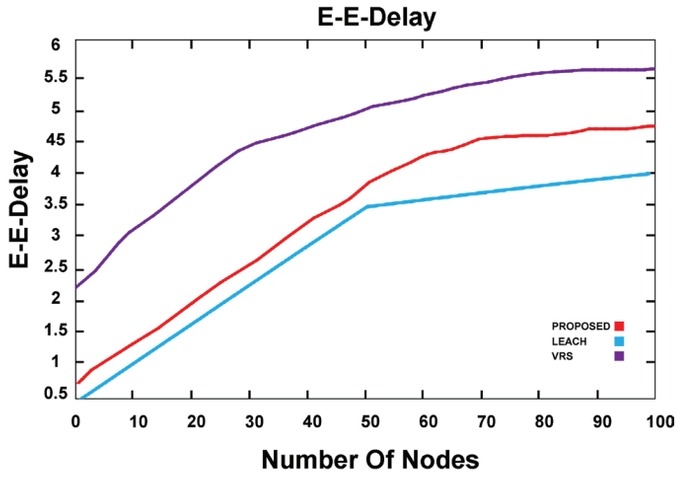


Figure 3:

End-to-End Delay.

Fig (3) presents the rate of the data arriving at their destination, it increases by the number of nodes participating in sending the required information. However, the rise is due to the new route discovery, which is completed by the nodes trying to compete to reach the wireless channel and to not be dropped or lost. Fig (3) represents the end-to-end delay for the proposed protocol gradually starts from 0.6 second for the transmission process for one node, which suggests that the network always has routes to the destination and is obtainable all the time. Then the proposed protocol starts stabilizing at node no. 70 tell the last node participating within the network.

Due to the new route discovery operated by LEACH, the delayed packets are increased to 1.3 second for node no.1 comparing with an equivalent node within the proposed model and starts stabilizing at node no. 50.

For the VRS it starts from 2.3 second for node no. 1 and increases steadily then stabilize at node no 80.

The proposed protocol gives us low rate of delay which means connected data without gaps or lost pieces of information which will make a deference in enhancing the performance. Stabilizing at 50 for the LEACH means gaps in the sent data and lost packets, on 80 for the VRS means overloading the transition bandwidth of data that may not be beneficial, hitting 70 with our proposed one means the needed data have been delivered before the nodes vanish.

Energy Consumption Versus the number

Energy consumption means what proportion energy the WSN consumes while operating. The main concern is that the efficiency of the WSN. Hence, presenting the energy consumption with the number of nodes, so as to the high energy consumption of the cluster head formation.

The energy consumption during a sensor is that the sum of the transmission energy consumption and therefore the received energy consumption multiplied with the time. Ec=(Et+Er)*t .

Where Ec is that the consumed energy, Et transmission energy consumption, Er received energy consumption, t is that the time.\

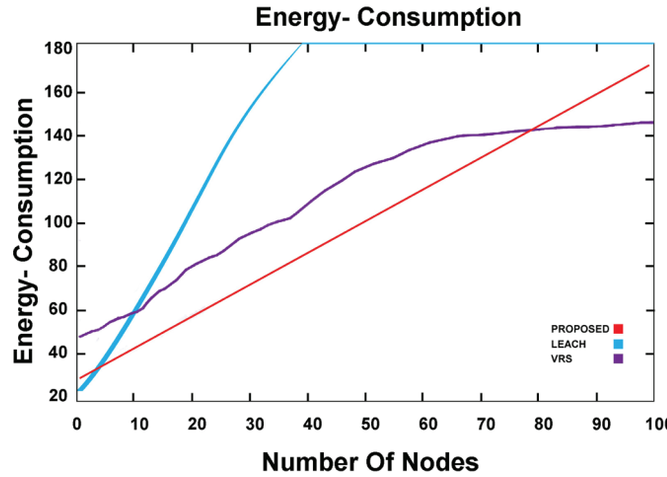


Figure 4:

Energy Consumption.

Fig (4) presents energy consumed by the activated nodes through the network while harvesting data to seek out the specified information which increases with the rise of the nodes number participating within the process.

Fig (4) shows that the energy consumption for the proposed protocol gradually starts with 20 for 10 nodes and reaches the maximum energy consumption 180 when using every node within the WSN. However, the LEACH reaches the maximum when using 40 nodes only which clearly shows that the proposed protocol is energy efficient more than LEACH. The proposed protocol is consuming the energy economically and does not reach the maximum until all the sensors are participating within the process.

For the VRS starts gradually from 30 for one node and the stabilized starting from node no.60.

The proposed protocol gives us high energy efficiency compared with the LEACH and VRS which means no lost data due to nodes vanish. This makes a deference in enhancing the performance by sending all the needed data before the power depletion.

The Packet Delivery Rate Versus Time

Packet delivery rate means the ratio of sent packets to the amount of received packets at the destination. We will compute it by dividing the successfully received packets to the entire sent packets. Packet delivery equals successfully received packets/ total packets.

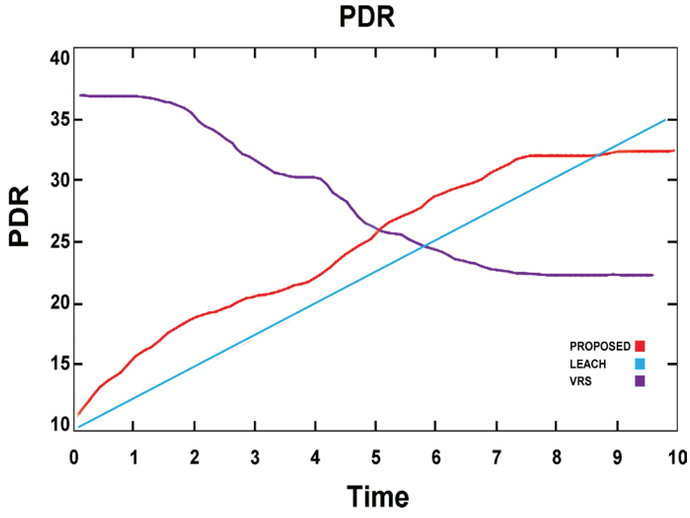


Figure 5:

Packet Delivery Rate.

Fig (5) shows that it increases with time. This indicates an honest performance for the WSN. Due to the effectively delivered packets to their destination without errors or being lost.

Fig (5) indicates that the packet delivery ratio for the proposed protocol is better than the packet delivery for the LEACH on its own. The packet delivery ratio for the proposed protocol gradually starts from 15 at the 1 second of the transmission process then stabilize at second 7. However, the LEACH start from 11 at the 1 second, which indicates that the proposed protocol is delivering the packet to the destination effectively.

For the VRS, it starts from 36 at the 1 second the decreases gradually then stabilize at second 6.

The proposed protocol gives us an efficient packet delivery and then stabilize after the 7 second that means no overhead and less lost packets compared to the LEACH which gradually increasing which may cause an overhead and the VRS which cause a lost packets because it gradually decreases.

The Throughput Versus the number of Nodes

Throughput is that the maximum infallible packets delivered to the destination during the communication process, and the way fast sending the packet through the communication channel is without having congestion.

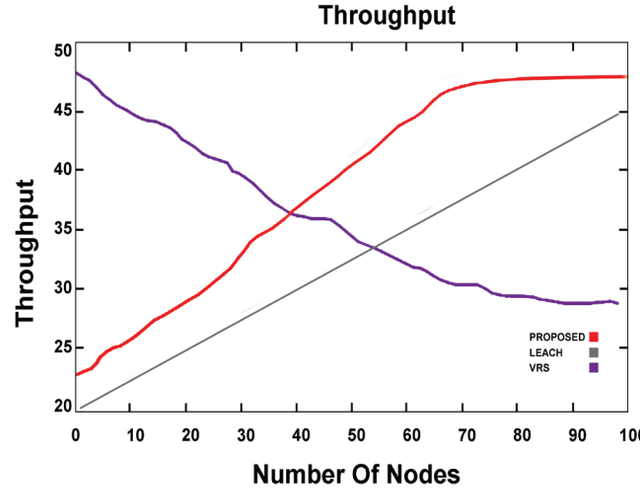


Figure 6:

Throughput.

Fig (6) presents that more nodes are delivering more packets. Which indicates an honest performance for the WSN.

Throughput for the proposed protocol starts gradually from 23 Mbps for the primary node then starts stabilizing when node no.60 is activated.

LEACH throughput depicts that it›d have some energy lost handling the congestion problem. Furthermore, the packets are delivered effectively to their destination through the communication process.

For VRS it decreases from 48 Mbps the stabilizes at the node no. 70.

The proposed protocol gives us less congestion which means more energy retention compared with the LEACH that will consume more energy to handle the congestion problem.

The VRS decreases because it is not competent to utilize the proposed protocol advantages.

The Packet Routing Load vs Simulation Time

The packet routing load expresses the entire number of transmitted packets over the successful transmission. During which the router chooses the simplest low-cost route to the destination, and during this case with low energy consumption.

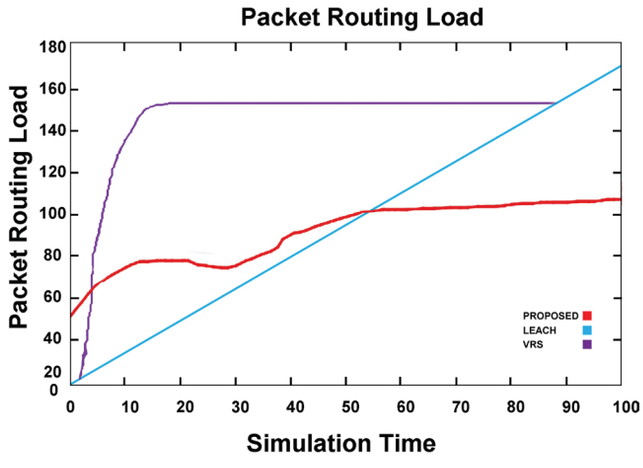


Figure 7:

Packet Routing Load.

Fig (7) shows that more packets are sent to their destinations at the primary few minutes and then, it stabilizes. This suggests that the proposed protocol saves the energy for the nodes that transmit data. Because it delivers 160 packets in less than 15 seconds then stabilize at this range of packet delivery Fig (7) represents LEACH routing load gradually increases which depicts that the increasing packet routing load with time consumes more energy from the nodes which reduce the network performance.

However, VRS starts from 50 packets then stabilizes at node no. 50 with 90 packets delivered.

Conclusion

In this paper, we discussed the dynamic proposed WSN routing protocol from the energy efficiency point of view to prolong the network lifetime and to scale back the amount of energy they consume while finding the route path from the source to destination. We used the hierarchical based (LEACH) is that the common protocol to reserve the energy and therefore the DD can find a solution for the routing problems like flooding using VRS algorithm. We will use this information to enhance the newly proposed protocol and have prolonged life for the network since the networks.

REFERENCES

- Dr. Arockiasamy, S., V.S. Anita Sofia. (2011). Generalized Framework for Energy Conservation in Wireless Sensor Network. International Journal of Computer Science Issues (IJCSI). 8(1). ISSN 1694-0814.
- Ajit, Pal, Sunkara, Kumar, Inodh, (2013). Assisted-Leach (A-Leach) Energy Efficient Routing Protocol for Wireless Sensor Networks. International Journal of Computer and Communication Engineering. 420-424. 10.7763/(IJCCE).2(218).
- Amro, Dalia, Dr. Atamimim, Liana.(2019). Dynamic energy-efficient routing protocol for wireless sensor. network. http://scholar.ppu.edu/bitstream/han thesis. pdf?sequence=1isAllowed=y
- Ganjali, Yashar, McKeown. Nick. (2005). Routing in a Highly Dynamic Topology. Second Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks.
- Greve, Robert, Gehlot, Jeguesh, Sha, Kewei. (2012). Multipath Routing Techniques in Wireless Sensor Networks: A Survey. Springer, DOI 10.1007/511277-0120723-2.
- Heidemanm, John, Estrin. Deborah, Govindam, Ramesh, Intanagonwiwat, Chakrmek (2003). Directed

Diffusion for Wireless Sensor Networking. IEEE/ ACM Transactions on networking, 11(1). 1663-6692.

- Intanagonwiwat, Chalermek, Govindan, Ramesh, Estrin, Deborah, Heidemann, John (2003). Directed Diffusion for Wireless Sensor Networking. IEEE/ ACM TRANSACTIONS ON NETWORKING, 11(1).
- Intanagonwiwat, Chalermek, (2002). Directed Diffusion: An Application-Specific and Data-Centric Communication Paradigm for Wireless Sensor Network. Doctor thesis, University of Southern California.
- Irum, Misbah, Nawaz, Mehmood, Murad. Maryam, Saeed. Nimrah (2017). Survey on Single Path and Multi Path Energy Efficient Routing Protocols for Wireless Sensor Networks. Journal of Computer and Communications, 5, 1-11, ISSN 2327-5227.
- Lingxi, Zhao, Wenjun, Wu (2015). Effective Local Dynamic Routing Strategy for Air Route networks. Chinese Journal of Aeronautics, 28(6):1709-1717.
- Mardini, W., Khamayseh. Y., AlZou'bi, A., Baniyassein. M., (2009). Improvement on LEACH Protocol of Wireless Sensor Network(V-LEACH). Research Gate, DOI: 10.1109.
- Markosyan, Mher, Mirfattahi, Mehdi, Mirnabibaboli, Miresmacil, (2011). Improving the Directed Diffusion in order to Reduce the Average of Energy Consumption in Wireless Sensor Network. The 5th International Conference on Sensor Technologies and Application, IARIA. ISBN 9781-1-61208-144-1.
- Nandy, Diya, Mitra, Rudranath, (2012). A Survey on Clustering Techniques for Wireless Sensor Network. International Journal of Research in Computer Science, 2(4), (pp. 51-57), ISSN 2249-8265.
- Salim, Ahmad, Osamy, Walid, Khedr. Amed M. (2014). IBLEACH: Intra-balanced LEACH Protocol for Wireless Sensors Networks. Springer. 20:1515-1525, DOI 10.1007/s11276-014-0691-4.
- Sunitha, Ch. Yadav, Lalita, (2014). Low Energy Adaptive Clustering Hierarch in Wireless Sensor Network (LEACH). International Journal of Computer Science and Information Technologies (IJCSIT), (5)3, ISSN:0975-9646.
- Tandel, Reshma, I. (2016). LEACH Protocol in Wireless Sensor Network: A Survey. International Journal of Computer Science and Information Technologies (IJCSIT), Vol 7.
- Tandel, Reshma, I. (2016). LEACH Protocol in Wireless Sensor Network: A Survey. International Journal of Computer Science and Information Technologies (IJCSIT), 7(4), P 1894-1896. ISSN 0975-9646 .
- Tang, Minghao, Tong, Mu. (2010). LEACH-B: An Improved LEACH protocol for wireless sensor network. IEEE, 978-1-4244-3709-2.
- Verma,Vandna, Sharma, Nishi (2013). Energy Efficient LEACH Protocol for Wireless Sensor Network. International Journal of Information and Network Security (IJINS), (l 2) 4, ISSN: 2089-3299.
- Willig, Andreas, Kaurl, Holger, Protocols and Architectures for Wireless Sensor Network. WILEY. com.
- Xiao, Mi, Khatoon, Shaheen, Keshi, Azhar Mahmood (2013). Data mining techniques for wireless sensor network: A survey. International Journal of Distributed Sensor Networks, https:doi.org 0.1155 2013 406316.
- Yulin, Song, Xianguing, Fan (2007). Improvement on LEACH protocol of Wireless Sensor Network. International conference on Sensor Technologies and Application, DOI 10.1109
- Zhn, Xiaoy, Zhang, Tiankui, Yang, Digcheng, Wu, Fahui, Xiao. Lin. (2016). Energy Effective Wireless Sensor Network Modeling Based on Complex Networks. Journal of sensors .Hindawi publishing corporation. ID 3831810.
- Fahmy (2016). Wireless Sensor Networks Concepts, Application, Experimentation and Analysis. Springer Ch (2).
- Guy (2006).Wireless Sensor Networks. SPIE, (6357,6357-1).
- https://www.spiedigitallibrary.org/conference-proceedings-of-spie
- Wireless Sensor Networks Principles, Design and Application. Springer CH (1- 2).
- Rawat, Singh, Chaouchi, Bonnin (2013).Wireless Sensor Networks: A Survey on Recent Developments and Potential Synergies. Journal of Supercomputing (10.1007/s11227-013-1021-9) Springer, Verlag.
- Nabavi, Seyed Reza, Eraghi, Nafiseh Osati, Torkestani, Javad Akbari (2021). WSN Routing using a multi objective greedy approach. Journal of Hindawi wireless communication and mobile computing,(6664669) https://doi.org/10.1155/2021/6664669.
- Engmann, Felicia, Katsriku, Ferdinand Apietu, Abdulai, Jamal-deen, Adu-manu, Kofi Sarpong(2020). Reducing the energy budget in WSN using time series models. Journal of Hindawi wireless communication and mobile computing,(8893064) https://doi.org/10.1155/2020/8893064.
- Hassan, Ali Abdul-hussian, Shah, Wahidah MD, Habeb, Abdul-Hussien Hassan, Othman, Mohd Fairuz Iskandar, AL-Mhiqani, Mohammed Nasser (2020). An Improved Energy-Efficient Clustering Protocol to Prolong the Lifetime of the WSN-Based IoT. Journal of IEEE Access(10.1109/ACCESS.2020.3035624) https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9247097
- Muzakkari, Bashir A, Mohamad, Mohamad A, Kadir, Mohd F. A, Mamat, Mustafa (2020). Queue and Priority-Aware Adaptive Duty Cycle Scheme for Energy Efficient Wireless Sensor Networks. Journal of IEEE Access(10.1109/ACCESS.2020.2968121)

https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=89638

- El-Esawy, Shimaa Gamal, ElShennawy, Nada, Elfishawy, Nawal Ahmed (2018)). An Improved Energy-Efficient Directed Diffusion Routing Protocol for Wireless Sensor Network. Journal of IEEE Access. https://ieeexplore.ieee.org/document/8636153
- Pantazis, Nikolaos A. Nikolidakis, Stefanos A., Vergados, Dimitrios D. (2013)). Energy-Efficient Routing Protocols in Wireless Sensor Networks: A Survey. Journal of IEEE Communications Surveys & Tutorials (10.1109/SURV.2012.062612.00084) https://ieeexplore.ieee.org/document/6248647
- Hiremani, Nirmala, Basavaraju, T.G (2017). Energy Efficient Routing Protocols, Classification and Comparison in Wireless Sensor Networks: A Survey. Journal of IEEE International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC)(10.1109/CTCEEC.2017.8455024) https://ieeexplore.ieee.org/document/8455024
- Elshakankiri, Maher N, Moustafa, Mohamed N, Dakroury, Yasser H.(2008). Energy Efficient Routing protocol For Wireless Sensor Networks. Journal of International Conference on Intelligent Sensors, Sensor Networks and Information Processing (10.1109/ISSNIP.2008.4762020) https://ieeexplore.ieee.org/document/4762020
- Bouzid, S.E., Serrestou Y, Raoof K ,Omri, M.N.(2020)
- Efficient Routing Protocol for Wireless Sensor Network based on Reinforcement Learning. Journal of International Conference on Advanced Technologies for Signal and Image Processing (ATSIP). (10.1109/ATSIP49331.2020.9231883). https://ieeexplore.ieee.org/document/9231883