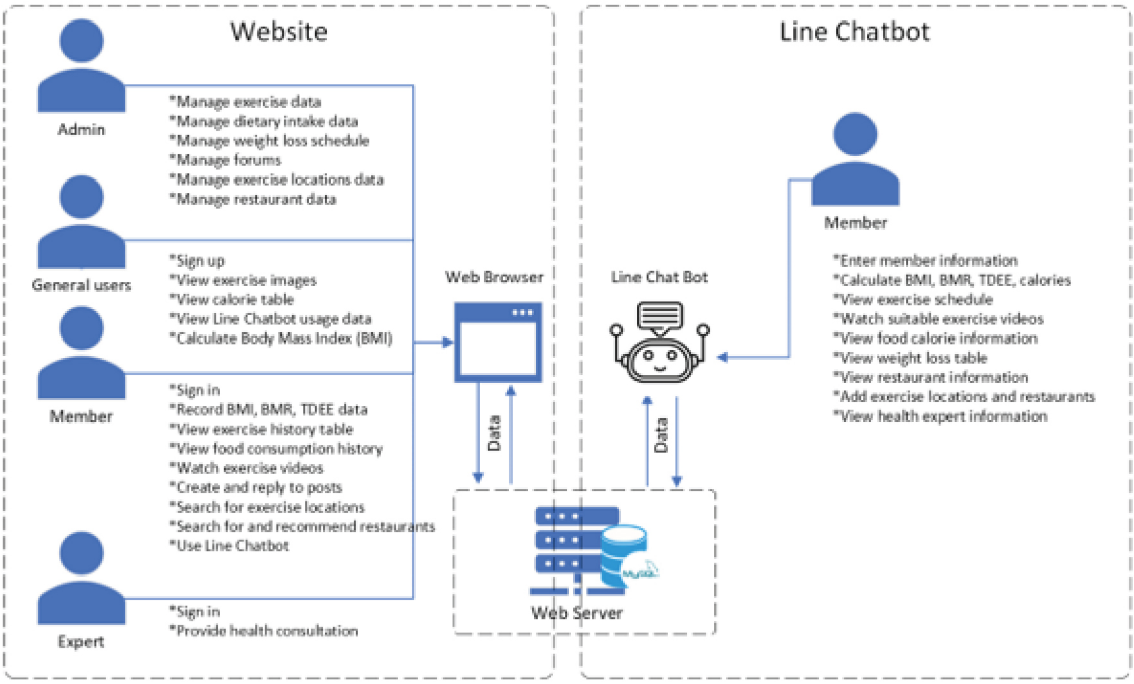
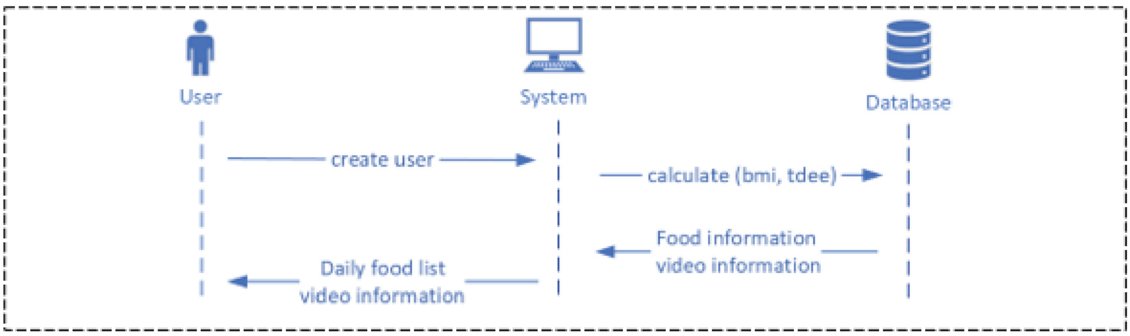
SSSTJ VOL.11 NO.2 July 2024
ISSN 2351-0889

SSSTJ

Suan Sunandha Science and Technology Journal
Volume 11, Number 2, July 2024





Available online at https://li02.tci-thaijo.org/index.php/ssstj

Suan Sunandha Science and Technology Journal

General Information

The Suan Sunandha Science and Technology Journal (SSSTJ) is a double-blind peer reviewed scientific journal published twice a year (January and July) by the Faculty of Science and Technology, Suan Sunandha Rajabhat University. Submissions of manuscripts should be sent to the Editor of the SSSTJ by online system: https://li02.tci-thaijo.org/index.php/ssstj The manuscript will be taken that all contributing authors attest that manuscripts and material submitted to the SSSTJ are original and have not been published or submitted elsewhere and the authors concede to the open-access distribution of the manuscript, including all contents contained therein.

Aim and scope:

Suan Sunandha Science and Technology Journal (SSSTJ) is an international academic journal that gains foothold at Suan Sunandha Rajabhat University, Thailand and opens to scientific communications in Southeast Asia, Asia and worldwide. It aims to contribute significant articles in science and technology research. Published papers are focused on state-of-the-art science and technology. The committee of the journal and association will review submitted papers. The authors may include researchers, managers, operators, students, teachers and developers.

Following areas are considered for publication:

1.Biology
2.Biotechnology
3.EnvironmentalScienceandTechnology
4.FoodScienceandTechnology
5.Microbiology
6.Applied Science
7.Computer Science and Information Technology
8.Other related fields

Available online at https://li02.tci-thaijo.org/index.php/ssstj

Suan Sunandha Science and Technology Journal

Editorial Team

Available online at https://li02.tci-thaijo.org/index.php/ssstj

Suan Sunandha Science and Technology Journal

Available online at https://li02.tci-thaijo.org/index.php/ssstj

Suan Sunandha Science and Technology Journal

SSSTJ Authors Guide

Types of paper:

The journal publishes research articles and academic articles. Full length articles are typically 4 - 10 manuscript pages (including tables, figures, equations and formulas). The articles may cover a variety of topics which are related to science or technology. Articles must use previously published works to support the discussion of a particular subject or topic.

Plagiarism:

Submission of an manuscript implies that the work described has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, without the written consent of the copyright-holder.

Articles and any other material published in the SSSTJ the opinions of the author(s) and should not be construed to reflect the opinions of the editor(s) or the publisher. Manuscripts that do not meet the general criteria or standards for publication in the SSSTJ will be immediately returned to the authors, without detailed review. At their discretion, editors may request from the corresponding author a statement describing specific original contributions made by each co-author.

Manuscript submission:

For submission of all manuscripts, follow the instructions of the online submission system. Hard copy submissions cannot be accepted. Before submission, prepare and keep ready all information on the manuscript (title, full name and affiliation of all authors, corresponding author email, abstract, name of all files to be submitted).The author submitting the manuscript will be corresponding author. Manuscripts are accepted only in American English. All submitted articles should report original, previously unpublished research results, experimental or theoretical, and will be peer-reviewed.

Role of the funding agency:

You are requested to identify who provided financial support for conducting the research and/or preparation of the article and to briefly describe the role of the sponsor(s), if any, in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. If the funding source(s) had no such involvement, then this should be stated.

Suan Sunandha Science and Technology Journal

Journal access:

Open access:

- Articles are freely available to both subscribers and the wider public.
- Articles are made available to subscribers at https://li02.tci-thaijo.org/index.php/ssstj

Regardless of how you choose to publish your article, the journal will apply the same peer review criteria and acceptance standards.

Peer review:

All published journal papers are refereed by the international competent researchers and scientists. Therefore, a full double - blind international refereeing process is used in which 1) Papers are sent to reviewers for their peer review process. 2) The reviewers' recommendations determine whether a paper will be accepted / accepted subject to change / subject to resubmission with significant changes / rejected. For papers which require changes, the same reviewers will be used to ensure that the quality of the revised paper is acceptable. All papers are refereed, and the Editor-in-Chief reserves the right to refuse any typescript, whether on invitation or otherwise, and to make suggestions and/or modifications before publication.

Additional information:

Manuscripts must be submitted by the senior author, who must accept responsibility on behalf of all authors for all ethical requirements. The author submitting the manuscript will be listed as the corresponding author in the published version of each accepted manuscript. There are no submission fees or page charges.

Revised manuscripts should be returned including revision notes. The revision notes should address the issues raised in the referee report and clearly state per page (indicate paragraph and line) which changes have been made. Additional materials may be requested at the discretion of the editor. Authors requested to revise their submission to SSSTJ will be given 4 weeks in which to submit the revised submission. Revised submissions received after 4 weeks from the time at which the revision was requested will be considered as new submissions.

References:

There are no strict requirements on reference formatting at submission. But the authors are expected to follow the American Psychological Association (APA 6th edition) style. Where applicable, author(s) name(s), journal title/book title, chapter title/article title, year of publication, volume number/book chapter and the pagination must be present. Use of DOI is highly encouraged.

Article structure:

Subdivision - numbered sections

Divide your article into clearly defined and numbered sections. Subsections should be numbered 1.1, 1.2, etc. (the abstract is not included in section numbering). Use this numbering also for internal cross-referencing:

do not just refer to 'the text'. Any subsection may be given a brief heading. Each heading should appear on its own separate line.

Article formatting:

Full-length articles published in the SSSTJ should include a title, the author's name, the author's address, the corresponding author email, Abstract, Introduction, Materials and Methods, Results and Discussion, Acknowledgements and References.

- **Font:** Times New Roman (headings 11pt and normal text 10pt)
- **Title:** Concise and informative. This should briefly summarize the scope of the article. Avoid abbreviations and formulas where possible. The title should be centered at the top of the page. e.g., Notes on the Biology of *Varanus semiremix*
- **Authors name and affiliation:** The full names referenced by numerical superscripts with affiliation and addresses of all authors, and the full address of the corresponding author. The name of the corresponding author along with phone, fax and e-mail information. e.g.

Author's Name
Author's affiliation
Author's address
Corresponding author email

- **Corresponding author:** Clearly indicate who will handle correspondence at all stages of refereeing and publication, also post-publication. Ensure that the e-mail address is given and that contact details are kept up to date by the corresponding author.
- **Abstract:** The abstract should contain brief information on purpose, methods, results and conclusion (without subheadings). It indicates significant data, and point out major findings and conclusions. Provide 200 – 250 words in length.
- **Keywords:** 3 – 6 keywords should be listed.
- **Introduction**: The introduction should provide sufficient background knowledge and information to allow the reader to understand and evaluate the value of the study. It must also supply a rationale for the study. Cite references to provide the most salient background rather than an exhaustive review of the topic.
- **Materials and Methods:** Specific details about materials, instrumentation and experimental protocols should be given here. This section should contain sufficient details so that others are able to reproduce the experiment(s). The source of special equipment or chemicals should also be given with the name and location of manufacturers, e.g. (Merck, Germany) or (Sigma, USA). Techniques previously published or standardized can be simplified by literature citations. The statistical procedures used should be explained. Primary headings for this section are in bold, indented, with numbering. The text is run from a new line under the heading with an indentation.
- **Results and Discussions:** The results should be presented with clarity and precision. The discussions should provide an interpretation of the data and relate them to existing knowledge.

Emphasis should be given to important new findings and new hypotheses should be described clearly. This section may be subdivided and headed as in the Materials and Methods section. At the end of this section, the authors should state the conclusion in a few sentences.

- **Conclusion:** The main conclusions of the study may be presented in a short conclusion section, which may stand alone or form a subsection of a Results and Discussions section.
- **Acknowledgment:** The acknowledgments should be stated in separate section at the end of the article before the references.
- **References:** Please make sure that every reference cited in the text is also present in the reference list and vice versa.
- **Appendices (optional):** If there is more than one appendix, they should be identified as A, B, etc. Formulae and equations in appendices should be given separate numbering: Eq. (A.1), Eq. (A.2), etc.; in a subsequent appendix, Eq. (B.1) and so on. Similarly for tables and figures: Table A.1; Figure. A.1, etc.

- **Tables and Figures:**

Tables: Please submit tables as editable text and not as image. Tables must be embedded into text. Headings should be in bold letters, and aligned in the text left. Vertical lines are not used. A table should not exceed one page when printed. Use lower case letters in superscripts a, b, c ... for special remarks.

Figure legends: It should be typed in numerical order. Begin each legend with a title and include sufficient description so that the figure is understandable without reading the text of the manuscript.

Figures: Each figure must be embedded into text. It should be prepared using applications capable of generating high resolution TIFF, JPEG, PNG.

- Color or grayscale photographs: minimum of 300 dpi
- Bitmapped line drawings: minimum of 1000 dpi
- Combination bitmapped line (color or grayscale photographs: minimum of 500 dpi)
- Headings should be in bold letters, and aligned in the text Centre

Checklist before submission:

The following list will be useful during the final checking of an article prior to sending it to the journal for review.

Ensure that the following contents are present:

One author has been designated as the corresponding author with contact details:

- E-mail address
- Affiliation
- Full postal address
- Corresponding author email

All necessary files have been uploaded, and contain:

- Keywords
- All figure captions

Available online at https://li02.tci-thaijo.org/index.php/ssstj

Suan Sunandha Science and Technology Journal

- All tables (including title, description)
- Manuscript has been 'spell-checked' (Spellchecker tool of MS word) and 'grammar-checked'.
- All references mentioned in the reference list are cited in the text, and vice versa.

Additional information: For additional information regarding format, content, submissions, or authoring guidelines, please contact the editor (e-mail: ssstj.contact@gmail.com).

SSSTJ has both online and printing versions.
ISSN 2351-0889 (Print)
e-ISSN 2539-5742 (Online)
Published by: Faculty of Science and Technology, Suan Sunandha Rajabhat University

SUAN SUNANDHA SCIENCE AND TECHNOLOGY JOURNAL
Suan Sunandha Rajabhat University, Faculty of Science and Technology
1 U-thong Nok Road, Dusit, Bangkok 10300 THAILAND

CONTENTS

July 2024, Volume 11, No.2

Available online at https://li02.tci-thaijo.org/index.php/ssstj

Suan Sunandha Science and Technology Journal

A Model for Converting Data into NoSQL Data Warehouse for Developing a Real-time Financial Data Warehouse System

Klaokanlaya Silachan[1], Sanya Kuankid[1,2], Thanin Muangpool[1,2*]

[1]Faculty of Science and Technology, Nakhon Pathom Rajabhat University, Malaiman Road, Muang, Nakhon Pathom 73000, Thailand
[2]Creating Innovation for Sustainable Development Research Center, Nakhon Pathom Rajabhat University, Nakhon Pathom, Thailand
***Corresponding author e-mail: signal@npru.ac.th**

Received: *5 February 2024* / **Revised:** *8 May 2024* / **Accepted:** *12 June 2024*

Abstract
This research introduces a novel model, the Financial Data Warehouses API (FDW-API), developed using PHP, Node.js, and Express.js. The model is designed to transform banking credit dataset information into a data warehouse format using a Non-Only SQL (NoSQL) database, stored in JSON format. Three types of databases were employed: MongoDB Node, MongoDB Serverless, and Cassandra. The study includes a comparative analysis of the data retrieval speed from all three databases. The model's applicability was tested in a real-time credit approval web application, demonstrating its effectiveness in transforming and storing data. Testing involved loading datasets ranging from 200, 300, 400, 500, 600, 800, and 1000 entries. Results indicate that the MongoDB serverless database outperformed others in terms of efficiency. Additionally, the FDW-API model streamlines data transformation and storage, facilitating real-time analysis and decision-making for financial institutions and data-driven businesses. Its flexibility integrates seamlessly with existing systems, reducing development time and costs, while its scalability accommodates growing data volumes and evolving business needs, providing a valuable tool for strategic insights and competitive advantage.

Keywords: API, NoSQL, MongoDB node, MongoDB serverless, Cassandra

1. Introduction

The incorporation of database systems into information systems yields benefits for management and facilitates the identification of relationships within business operations. Predominantly, individuals favor the utilization of Online Transactional Processing (OLTP) for data processing, a principle aimed at data storage that reduces redundancy and upholds data accuracy, thus mitigating errors resulting from data editing. Retrieval of available data from various sources is possible (Songsiri & Tamee, 2022). However, OLTP is not conducive to decision support systems requiring specificity. This is due to the necessity for large databases containing more extensive data than usual, segmented into smaller tables based on design principles, rendering it incapable of supporting the query format essential for decision-making support. This includes historical data retrieval to predict potential future trends based on developed models (Harvy, Matitaputty, Girsang, Michael, & Isa, 2019).

The management of large databases experiencing daily data volume increments challenges conventional techniques, posing a risk of errors. The method of segmenting the database into parts enhances its suitability for use. With a database system capable of supporting extensive data entry and simultaneous processing, Online Analytical Processing (OLAP) facilitates immediate service delivery (Wang, Li, Xu, Wang, & Wang, 2021). Consequently, a data warehouse system storing substantial data differs in structure from general databases, optimizing efficiency in data retrieval for end-users. The process of constructing an organization's data warehouse typically involves storing long-term data for at least 5-10 years, utilized in analyzing business operational processes (Garani, Chernov, Savvas, & Butakova, 2019).

The transformation of data from general databases into Online Transactional Processing (OLTP) format for integration into a data warehouse is achieved through the Extract, Transform, and Load (ETL) process (Singsanit, 2021). This process facilitates the extraction of data from various operational sources to execute Data Migration in alignment with the objectives of the software. Furthermore, the development of tools for data transformation through the ETL process encompasses two predominant modalities: tool utilization and code generation-based programming, typically implemented in batch or offline formats (Yulianto, 2019). The intricate interchangeability inherent in the ETL process may encounter challenges if not appropriately designed, when considering data volume (Barahama & Wardani, 2021).

Addressing challenges in loading and exporting data from large-scale operational databases involves prolonged loading times and necessitates a waiting period before data becomes usable (Nizzad & Irshad, 2021). Conversely, if a real-time data warehouse is operational, capable of instantaneously receiving data from OLTP systems, the data is promptly transmitted from the primary database to the data warehouse. This enhances processing efficiency for concurrent retrieval and analysis, requiring the development of a system to perform real-time data transformation through code scripting into real-time data software, as opposed to batch or offline data transformation. This approach contributes to expediting the loading process, enabling instantaneous data analysis. Moreover, selecting an appropriate database structure is crucial for efficient and swift data loading, facilitating subsequent data analysis and report generation. Generally, structured data is stored in data warehouses, characterized by a meticulous data schema. If data modifications are needed, all structured data must be updated, which is a time-consuming and resource-intensive process (Hassan et al., 2022).

Consequently, non-relational database models have been employed in the development of various systems to accommodate the anticipated surge in data volume. Researchers have introduced non-relational database models as a novel approach with potential enhancements for contemporary data warehouses. Currently, system development incorporates the use of Application Programming Interfaces (APIs) as interfaces for application services or various modules, contributing to accelerated development processes (Jose & Abraham, 2020; Oditisi, Bicevska, Bicevskis, & Karnitis, 2018; Petricioli, Humski, & Vrdoljak, 2021).

In this research endeavor, the proposed approach involves the development of a system model in the form of a web API based on the principles of Node.js and Express JS. The objective is to transform data into the non-relational data warehouse structure in the JSON document format. A performance comparison is conducted to evaluate the efficiency and speed of data retrieval from databases, encompassing MongoDB nodes, MongoDB serverless, and Cassandra (Boonhao, 2020; Bouaziz, Nabli, & Gargouri, 2019; Chauhan, 2019). This comparative analysis aims to identify the most time-efficient and optimal data loading strategy from the data warehouse, providing insights for the subsequent development of a real-time web application credit approval system.

2. Objectives

2.1 To advance the procedures for transforming data from API format to a non-relational database structure.

2.2 To compare and evaluate the results against procedures involving a relational database, measuring the system's impact on data access and efficiency.

3. Research Methodology

The research methodology encompasses five distinct stages, outlined as follows:

3.1 Data structuring for real-time credit systems

This phase involves an in-depth exploration of relevant information. The researcher focused on understanding the development methods and characteristics of data transformation. Due to the complexity of the real-time credit application system, importing data involves intricate personal data, which makes precise measurement challenging. Therefore, a preliminary assessment of data sets was conducted, and the chosen method and non-relational database structure that exhibit optimal responsiveness were implemented. The research utilized a newly generated OLTP data set constructed from the financial dataset structure obtained from the PKDD CUP 99 data source (Alfred & Kazakov, 2006, Al-Mamory & Jassim, 2013; Tavallaee, Bagheri, Lu, & Ghorbani, 2009). This dataset was subsequently refined to align with the credit data structure of banking institutions. A representative data structure is illustrated in Figure 1.

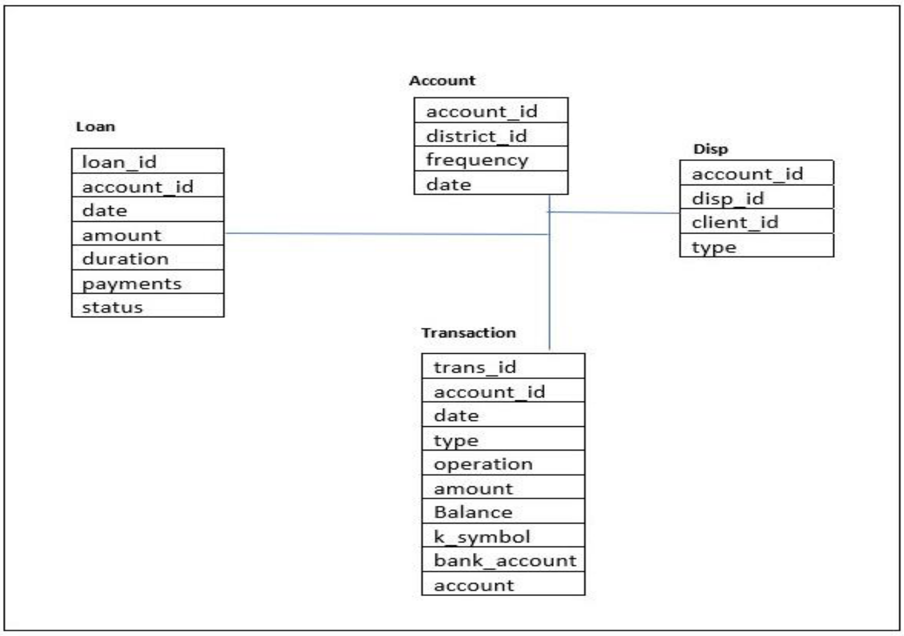


Figure 1. Dataset structure.

3.2 System analysis

The system analysis phase is a critical step where the studied system is examined to define its scope. This entails conducting a comprehensive review of the system's structure and the methodologies employed in its development and testing. Specifically, the analysis focused on the development and testing of three data transformation processes designed to import data into a non-relational database. These processes included MongoDB nodes, MongoDB serverless, and Cassandra, as depicted in Figure 2. The primary objective was to evaluate the efficiency of these processes in terms of response time for data transformation into an optimized database structure.

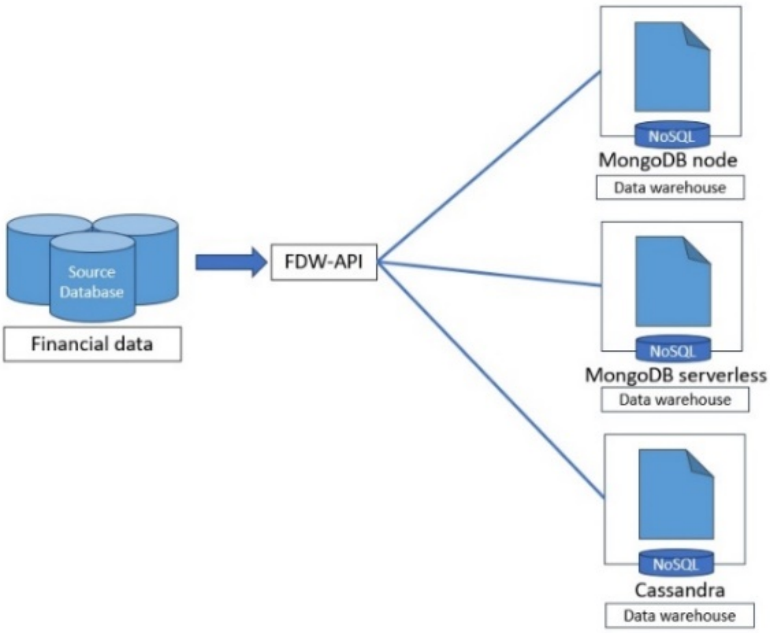


Figure 2. Diagram of data transformation into non-relational database using API.

3.3 System design

In system design, the approach is divided into two parts based on usage scenarios:

3.3.1 Design for data warehouse

In this segment of the system design, the focus was on the transformation of data to the data warehouse. The process involved converting data from the analytical online processing of the credit dataset through FDW-API to MongoDB nodes, MongoDB serverless, and Cassandra. Notably, Cassandra was served as the non-relational database with a distinct data structure compared to MongoDB nodes. The design adopted a hybrid approach, blending table-based data storage with key-value pairs. This allows for the conversion and utilization of JSON documents for storage in the Cassandra table structure, as illustrated in the representative data structure shown in Figure 3.

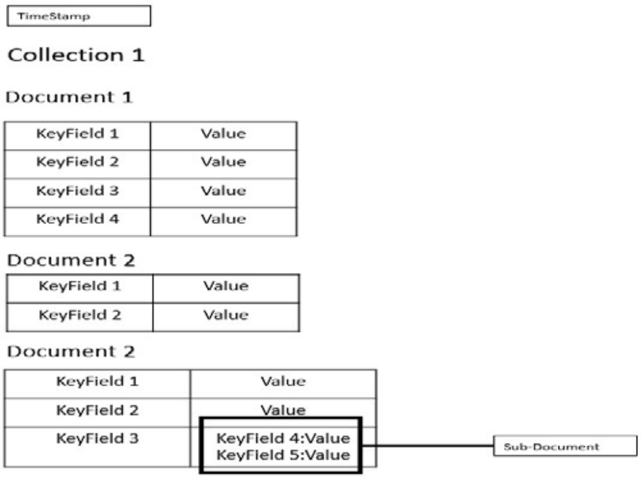


Figure 3. Structure of the document-oriented database.

3.3.2 System design

In the system design pertaining to the API-based model, the overall process for data transformation was outlined to identify the most efficient performance for immediate integration into the real-time system. The following steps elucidate the FDW API (Financial Data Warehouses API) model:

Step 1: FDW API transformation process

Utilizes Python Flask for the conversion of data from .asc (1) format to JSON documents conforming to the JSON Schema. Subsequently, the transformed data is stored in three distinct databases, namely MongoDB nodes, MongoDB serverless, and Cassandra.

Step 2: Node.js and Express.js API creation

- Reads ASCII files with semicolon delimitation from the dataset, assigning field names based on the file headers.
- Detects the data type and maximum length of the data in each field.
- Identifies the central table that has the highest number of foreign key connections to other tables.
- Embeds the data in JSON document format using the identified primary table.
- Displays the results as JSON documents through node.js + express.js, employing specified endpoints in the form of GET requests:
    - / warehouse/ json/ <table_name> for retrieval from MongoDB nodes, MongoDB serverless, and Cassandra.
    - /warehouse/mysql/<table_name> for retrieval from MySQL with full table join.
- Conducts load testing to evaluate and compare the performance of the designed API-based model.

3.4 System development

The system development phase focused on the implementation of the data transformation and management processes into the data warehouse. The FDW API (Financial Data Warehouses API) was developed using PHP, Node.js, Express.js, and in the form of a RestAPI. This API was designed to interact with non-relational databases, including MongoDB, MongoDB Serverless, and Cassandra. Additionally, for the relational database segment handling financial credit dataset, MariaDB was employed.

3.5 System testing and evaluation

This phase involves the comprehensive testing and evaluation of the system's functionality. The system demonstrates its ability to process, transform, and extract data from the data warehouse for accurate and timely presentation. Comparative speed assessments for data transformation have been conducted across non-relational database formats, including MongoDB, MongoDB Serverless, and Cassandra. The test results have been analyzed to identify the optimal transformation time, forming the basis for establishing a non-relational database for real-time credit data warehousing development.

4. Results

4.1 System development results

From the FDW API methodology outlined in Section 3.2.2, the designed process was executed to convert the dataset into a document-oriented data store format known as JSON. The details conform to the structure of JSON documents, facilitating data storage in a non-relational database format characterized by document properties akin to transaction timestamps, as depicted in Figure 4.

Figure 4. Example of converting JSON document structure for data storage in a JSON-oriented database.

4.2 Performance evaluation results

In this section, we present the performance evaluation results, specifically focusing on the processing speed comparison between non-relational database formats. The comparison encompasses MongoDB, MongoDB serverless, Cassandra, and is benchmarked against the average processing time. These results correspond to the second objective, which aims to compare and evaluate the results against procedures involving a relational database, measuring the system's impact on data access and efficiency.

The software components include MongoDB, MongoDB serverless, and Cassandra for database management. The hardware utilized for the evaluation comprises an Intel Core i7 8700 processor (6 cores / 12 threads), 16GB of RAM, Western Digital Black M.2 Solid State Drive, 10/100/1000 Gbps Ethernet, and the operating system employed is Microsoft Windows 11.

The system was developed using the PHP, Node.js, and Express.js programming languages in the form of RESTful APIs. The experimentation involved the Financial Dataset to assess and compare the response times of the proposed models and the non-relational database.

The results are analyzed by comparing the response times against the quantity of data entries used in the testing phase. The evaluation considers both the presented model and the non-relational database, as depicted in Figure 5.

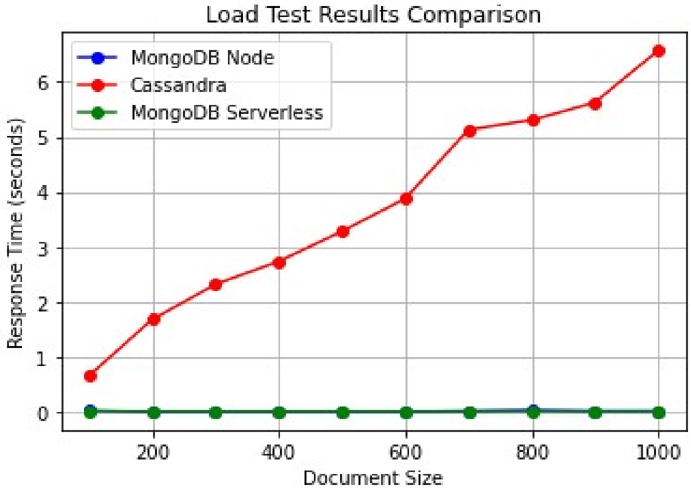


Figure 5. Response time evaluation results.

The response times of the MongoDB Node and MongoDB Serverless databases exhibit closely aligned values, prompting a dedicated comparison as shown in Figure 6.

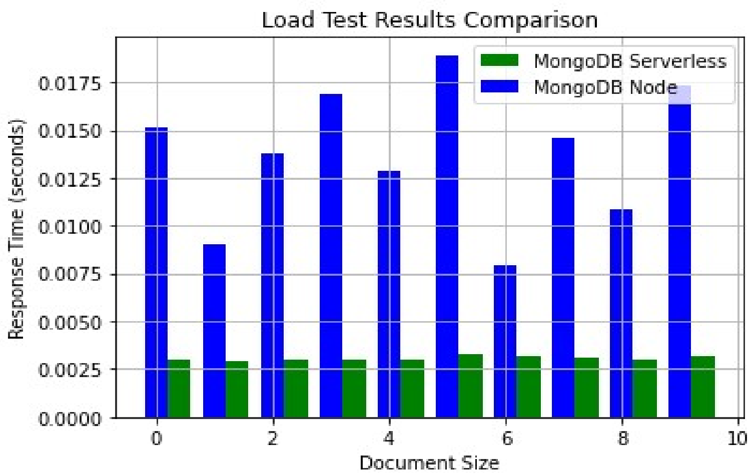


Figure 6. Comparative analysis of response times between MongoDB Node and MongoDB Serverless databases.

The average load times resulting from the assessment of the initial 100 dataset entries, followed by increments of 200, 300, 400, 500, 600, 800, and concluding with the last set of 1000, are tabulated below in Table 1.

Table 1. Average load times for testing datasets.

Number of Record Loaded	MongoDB Node (seconds)	Cassandra (seconds)	MongoDB Serverless (seconds)
100	0.015157	0.707098	0.003006
200	0.009023	1.311107	0.002936
300	0.013725	1.881718	0.003028
400	0.016828	2.448808	0.002962
500	0.012839	3.087998	0.002991
600	0.018908	3.733704	0.003286
700	0.007931	4.332826	0.003177
800	0.014600	5.038835	0.003118
900	0.010808	5.391285	0.003006
1000	0.017298	5.989958	0.003162

The experimentation results clearly demonstrate the successful development of the FWD-API model for establishing a non-relational data warehouse. Moreover, upon comparing the response time outcomes, it was evident that the MongoDB serverless operates more efficiently within the system, outperforming both MongoDB and Cassandra. This superiority is attributed to its faster data loading capabilities, resulting in enhanced performance, stability, and faster overall processing when compared to alternative models.

4.3 Development of real-time web-enabled credit data warehouse system

The outcome of the performance evaluation in developing a real-time banking credit web application, utilizing the FDW-API model to transform credit approval requests into the MongoDB serverless data warehouse, is presented. This section reflects the development and testing phase, showcasing the system's functionality in real-world application scenarios. When a credit applicant submits the request through the data entry form, the information is stored in the online analytical processing database. Upon approval, detailed data regarding the credit applicant and the approval transaction are stored in two segments: the MariaDB database and the MongoDB serverless data warehouse. This design enables further analysis and reporting based on the structure outlined in Figure 7.

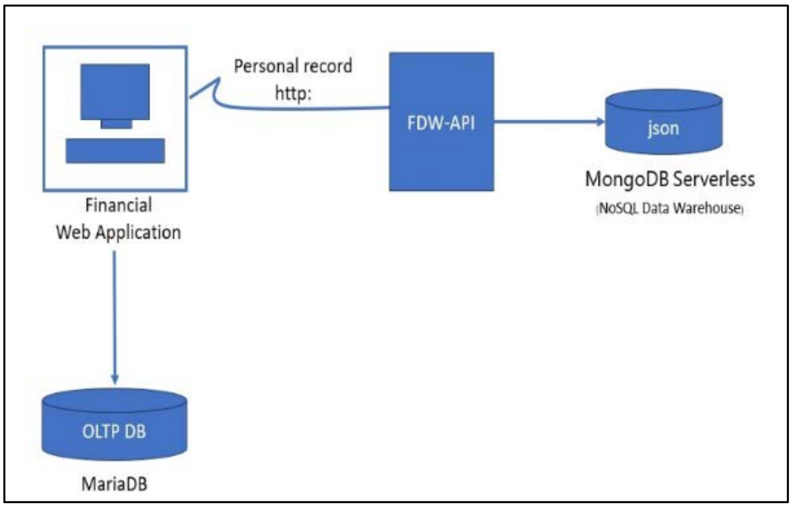


Figure 7. System development structure.

5. Discussion

This research endeavors to develop a real-time credit data warehouse system, employing API principles to manage data integration into a non-relational database structure. The chosen document-oriented JSON structure, known as FDW-API, was implemented using PHP, Node.js, Express.js, and utilized across three database formats: MongoDB, MongoDB serverless, Cassandra. The objective was to identify the most responsive format for real-time credit application systems.

Upon testing, it was observed that the MongoDB serverless data warehouse exhibited the best overall performance. This conclusion stems from its superior data retrieval speed and efficient operational outcomes. These findings align with the research conducted by Jaratsantijit, Y. (Jaratsantijit, 2022), which explored data storage strategies in NoSQL databases, specifically comparing the performance of relational databases with NoSQL databases for information systems. The choice of MongoDB as the NoSQL database for this research is substantiated by its optimal responsiveness and efficiency in handling real-time credit application systems.

It is recommended that future research further explores the scalability and performance of the FDW-API model across diverse industry sectors. Investigating the implementation of the model in sectors beyond finance could provide valuable insights into its adaptability and efficacy in different contexts. Furthermore, guidelines should be developed for agencies and stakeholders interested in implementing the FDW-API model. These guidelines could outline best practices for data integration, system architecture design, and performance optimization to maximize the model's benefits. Additionally, educational resources and training programs could be developed to support organizations in effectively leveraging the model for real-time data analysis and decision-making.

6. Conclusion

This study focuses on the design and development of a procedural model for transforming data from a credit database transaction set (OLTP) to a non-relational data warehouse in the form of a document-oriented model, specifically leveraging the MongoDB database. The study aims to accommodate both data storage and retrieval needs. Through performance testing involving 100 initial datasets, followed by increments of 200, 300, 400, 500, 600, 800, and finally 1000 entries, the procedural model and the data warehouse, based on the non-relational MongoDB serverless architecture, demonstrated efficient response times, indicating rapid data processing. This suggests that the developed model is well-suited for application as a data warehouse.

Acknowledgements

This research was made possible by the kind support of the Research and Development Institute of Nakhon Pathom Rajabhat University.

Conflict of Interest

No conflict of interest.

ORCID

Thanin Muangpool
https:// orcid.org/0000-0002-1088-5725
Sanya Kuankid
https:// orcid.org/0000-0003-0918-1402

Publication Ethic

Submitted manuscripts must not have been previously published by or be under review by another print or online journal or source

References

Alfred, R., & Kazakov, D. (2006). Pattern-based transformation approach to relational domain learning using dynamic aggregation for relational attributes. *Proceedings of the 2006 International Conference on Data Mining* (pp. 118-124). Las Vegas, Nevada.

Al-Mamory, S., & Jassim, F. S. (2013). Evaluation of different data mining algorithms with KDD CUP 99 data set. *Journal of University of Babylon*, *21*.

Barahama, A. D., & Wardani, R. (2021). Utilization extract, transform, load for developing data warehouse in education using Pentaho Data Integration. *Journal of Physics: Conference Series*, *2111*. doi:10.1088/1742-6596/2111/1/012030

Boonhao, P. (2020). NewSQL databases and usability trends. *Mahidol R2R e-Journal*, *8*(2), 38-52.

Bouaziz, S., Nabli, A., & Gargouri, F. (2019). Design a data warehouse schema from document-oriented database. *Procedia Computer Science*, *159*, 221-230. doi:10.1016/j.procs.2019.09.177

Chauhan, A. (2019). A review on various aspects of MongoDB databases. *International Journal of Engineering Research & Technology (IJERT)*, *8*(05), 90-92.

Garani, G., Chernov, A., Savvas, I., & Butakova, M. (2019). A data warehouse approach for business intelligence. *Proceedings of the 2019 IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)* (pp. 70-75). Napoli, Italy.

Harvy, I., Matitaputty, G. A., Girsang, A. S., Michael, S., & Isa, S. M. (2019). The use of book store GIS data warehouse in implementing the analysis of most book selling. *Proceedings of the 7th International Conference on Cyber and IT Service Management (CITSM)* (pp. 1-5). Jakarta, Indonesia.

Hassan, C. A. U., Hammad, M., Uddin, M., Iqbal, J., Sahi, J., Hussian, S., & Ullah, S. S. (2022). Optimizing the performance of data warehouse by query cache mechanism. *IEEE Access*, *10*, 13472-13480. doi:10.1109/ACCESS.2022.3148131

Jaratsantijit, Y. (2022). *Comparative study of query performance between relational database and NoSQL database for information system: A case study of the asset database of information technology service center.* Chiang Mai: Chiang Mai University.

Jose, B., & Abraham, S. (2020). Performance analysis of NoSQL and relational databases with MongoDB and MySQL. *Materials Today: Proceedings*, *24*(3), 2036-2043. doi:10.1016/j.matpr.2020.03.634

Nizzad, A. R. M., & Irshad, M. B. M. (2021). Data warehouse implementation: Cost effective approach for small businesses. *Journal of Information Systems & Information Technology (JISIT)*, *6*(2), 62-71.

Oditisi, I., Bicevska, Z., Bicevskis, J., & Karnitis, G. (2018). Implementation of NoSQL-based data warehouses. *Baltic J. Modern Computing*, *6*(1), 45-55.

Petricioli, L., Humski, L., & Vrdoljak, B. (2021). The challenges of NoSQL data warehousing. *Proceedings of International Conference on E-business Technologies* (pp. 44-48). Serbia.

Singsanit, K. (2021). The development of executive information system for managing research in university by data integration techniques on ontology on business intelligence. *Journal of Buddhist Education and Research: JBER*, *7*(1), 157-174.

Songsiri, K., & Tamee, K. (2022). Development of data warehouse for financial report in Faculty of Science, Naresuan University. *Journal of Applied Informatics and Technology*, *4*(2), 99-113. doi:10.14456/jait.2022.8

Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani A. A. (2009). A detailed analysis of the KDD CUP 99 data set. *Proceedings of the 2009 IEEE Symposium on Computational Intelligence in Security and Defense Applications (CISDA 2009)* (pp. 1-6). Ottawa, Canada.

Wang, D., Li, Q., Xu, C., Wang, P., & Wang, Z. (2021). Research of data warehouse for science and technology management system. *Proceedings of the International Conference on Service Science (ICSS)* (pp. 65-69). Xi'an, China.

Yulianto, A. A. (2019). Extract transform load (ETL) process in distributed database academic data warehouse. *Journal on Computer Science and Information Technologies*, *4*(2), 64-71. doi:10.11591/APTIKOM.J.CSIT.36

Available online at https://li02.tci-thaijo.org/index.php/ssstj

Suan Sunandha Science and Technology Journal

Electrospun Cellulose Acetate/Maleic Acid Functionalized Gold Nanoparticles as Pb(II) Ions Colorimeter Senser Strip

Pimolpun Niamlang[1*], Chularat Iamsamai[1], Piyanut Muangtong[1], Piyachat Chuysinuan[2]

[1]Department of Materials Engineering, Faculty of Engineering, Rajamangala University of Technology Rattanakosin, Phutthamonthon, Nakhonpathom 73170, Thailand

[2]Laboratory of Organic Synthesis, Chulabhorn Research Institute, Bangkok 10210, Thailand

***Corresponding author e-mail: pimolpun.kam@rmutr.ac.th**

Received: *10 February 2024* / **Revised:** *14 May 2024* / **Accepted:** *28 June 2024*

Abstract

This study investigates the synthesis of gold nanoparticles and surface modifications to enable their utilization as a colorimetric sensor strip on electrospun nanofibrous. Additionally, it utilizes cellulose acetate electrospun fiber preparations as substrates for loading the gold nanoparticles. In order to produce gold nanoparticles, citrate reduction was employed, followed by modification with maleic acid (MA-GNPs). MA-GNPs were significantly more selective for Pb^{2+} than for other ions (Co^{2+}, Cu^{2+}, Hg^{2+}, Ni^{+}), according to the results. Following the introduction of Pb^{2+}, the solutions exhibited a color change from red to blue or purple, which was attributed to the aggregation of nanoparticles, as determined by UV-Vis spectrometry. An optical band indicative of GNPs and MA-GNPs was detected at an estimated wavelength of 520 nm. Approximately 600 nanometers after the addition of Pb^{2+}, the intensity of the solution progressively decreased to 520 nanometers, and a new band emerged. To fabricate cellulose acetate nanofibrous via electrospinning, the impact of solvent ratios, polymer concentrations, and process conditions were investigated. The ideal parameters for the fabrication of the substrate were as follows: 15% w/w polymer solution, 15 kV electrospinning voltage, 15 cm needle-to-collector distance, and 48 hours of collecting time. In order to examine the color change of the prepared strip, it is observed that the strip transforms from red to blue upon exposure to Pb^{2+} ions of varying concentrations, just as it does in solution form. An examination of water samples indicated that the strip remained colorless when tested with DI water. However, testing with wastewater sourced from a battery facility identified a transformation of the strip from red to purplish blue. This is equivalent to an estimated lead concentration of 40 parts per million. The outcomes showcased the capacity of the discolored strips to identify the preliminary concentration of lead ions, suggesting potential for future advancements in this area.

Keywords: Gold nanoparticle, Lead detection, Electrospinning process, Colorimetric sensor, Sensor strip

1. Introduction

The pollution of harmful substances in soil and water is currently a serious topic being addressed by a number of parties. This issue has a profound impact on the environment, ecosystems, public health, and human livelihoods. Rapid industrialization, technological advancement, and economic growth have resulted in the emission of pollutants beyond permissible limits, resulting in major and difficult environmental challenges. One of the consequences is heavy metal leakage and poisoning of water supplies (Abdullah, Yusof, Lau, Jaafar, & Ismail, 2019; Feng et al., 2022).

Heavy metals are elements with a specific gravity greater than four, and they are primarily members of the transition metals group. They accumulate in soil and water because they are persistent and incapable of breaking down naturally (Feng et al., 2022). Heavy metals are crucial in numerous industries, including plastics, PVC, paint, and agricultural products such as fertilizer and insecticides (Chen & Wang, 2004; Fewtrell, Kaufmann, & Prüss-Üstün, 2003; Han, Zou, Zhang, Shi, & Yang, 2006). In everyday living, people are at danger of swallowing heavy metals through food or water that contains these compounds. Communities living near factories may contribute to inappropriate waste disposal methods, resulting in soil and water contamination. This contamination has a negative influence on aquatic life, compromising the delicate balance of ecosystems required to maintain life. The government recognizes the importance of this issue and has put in place guidelines for managing industrial wastewater outflow. For example, in compliance with the Factory Act of 1992, the Ministry of Industry issued regulations in 2017 to regulate the emission of heavy metals from industries. The laws establish allowable limits for elements like zinc and copper in wastewater (World Health Organization, 2004).

Traditional techniques such as Atomic Absorption Spectroscopy (AAS) and Inductive Coupled Plasma-Mass Spectroscopy (ICP-MS) have been used to monitor heavy metal ions (Ghaedi, Ahmadi, & Shokrollahi, 2007; Milne, Landing, Bizimis, & Morton, 2010). These approaches have limitations, such as high costs, complex operations, and high professional requirements. In comparison, the colorimetric method for heavy metal ions has gained growing attention due to its many advantages, particularly with regard to prompt and on-site detection. Specifically, its portability allows for field analysis, rapid detection using the naked eye without complex instruments, and ease of use for non-experts, suggesting significant potential for future applications.

A colorimetric sensor system for environmental and medical applications has been developed in response to these concerns. The sensor device, which employs gold nanoparticles, provides an affordable and user-friendly alternative to standard approaches. Furthermore, the use of nanomaterials, such as nanofibers and nanowebs electrospun (Huang, Zhang, Kotaki, & Ramakrishna, 2003) improves the sensitivity and biocompatibility of the sensing platform (Chansuwan, 2022; Hammami, Alabdallah, Al jomaa, & Kamoun, 2021; Ratnarathorn, Chailapakul, & Dungchai, 2015; Sartore, Barbaglio, Borgese, & Bontempi, 2011).

This study aims to investigate the properties of maleic acid-modified gold nanoparticles/electrospun cellulose acetate nanofiber mats as a sensor platform for identifying color shifts related to concentrations of heavy metal ions.

2. Materials and Methods

2.1 Materials

$HAuCl_4$ was purchased from Alfa Aeser, United States. Tri-sodium citrate was purchased from Ajax Finechem, Auckland. Maleic acid from TCI, Japan. Cellulose acetate, N, N-Dimethylacetamide (DMAc), and acetone were purchased from Sigma-Aldrich. All metal nitrate salts were AR grade from QRëC®, New Zealand.

2.2 Preparation of electrospun cellulose acetate nanofiber mats

As a supporting platform or test strips, cellulose acetate was formed into nanofiber mats. A 15% w/v cellulose acetate solution was made by dissolving it in a 2:1 solvent mixture of acetone and N, N-Dimethylacetamide (DMAc) (Zhang et al., 2019). The solution was stirred for 10 hours at room temperature using a magnetic stir bar until the solution became clear. The obtained solution was then electrospun into nanofibers using an electrospinning device with the electric voltage adjusted sequentially to 10, 12.5, and 15 kV. The needle tip to collector distance was 20 cm, and the solution was allowed to spin for 15 minutes. SEM (Scanning Electron Microscopy) was used to evaluate the resultant nanofibers for morphological properties, and Image J was used to measure size. The nanofiber mats were then formed for 2 hours and their thickness was measured with a micrometer. Finally, the nanofiber mats were dried in an air oven at room temperature (25°C) to eliminate any residual solvent.

2.3 Preparation of Maleic Acid/Gold Nanoparticles (MA-GNPs)

The preparation of gold nanoparticles using the method developed by Nalin Ratnarathorn and colleagues (Ratnarathorn et al., 2015) involves a modified procedure, which can be summarized as follows: A volume of 100 ml of tetrachloroauric acid solution with a concentration of 0.01% w/v is placed in a 250-ml amber glass bottle. The solution is then heated until boiling. Subsequently, 3.5 ml of sodium citrate solution with a concentration of 1% w/v is added rapidly, and the mixture is vigorously stirred with a stirrer bar for 15 minutes. Continuous stirring without additional heating is maintained for another 30 minutes. The solution is then allowed to cool to room temperature and stored in a refrigerator at 4°C. The light absorption characteristics are analyzed using a UV-Vis spectrophotometer.

2.4 Measurement of heavy metal ions

2.4.1 Measurement using MA-GNPs solution

A 1000 ppm solution of heavy metal ions, including Co^{2+}, Cu^{2+}, Hg^{2+}, Ni^{+}, and Pb^{2+} ions, was prepared. This solution was then diluted to quantities ranging from 10 to 100 ppm. To evaluate the detection capability of heavy metal ions, the diluted heavy metal ion solution was dropped into the MA-GNPs solution in quantities of 5, 10, 20, 30, 40, and 50 µl.

2.4.2 Measurement using MA-GNPs on cellulose acetate nanofiber mats

The cellulose acetate nanofiber mats were impregnated with MA-GNPs by dropping the MA-GNPs solution onto the nanofiber mats and allowing them to dry at room temperature. The detection of heavy metal ions was then tested by dropping 10-100 ppm heavy metal ion solutions onto the MA-GNPs-impregnated nanofiber mats. The results were recorded using a high-resolution camera for color comparison.

3. Results and Discussion

3.1 The condition of electrospining

The size of electrospun cellulose acetate fibers was measured using different electrostatic voltages (10, 15, and 20 kV). The following parameters were maintained: a cellulose acetate concentration of 15% w/v in a solvent mixture of acetone and N, N-dimethylacetamide (DMAc) in a 2:1 ratio, a needle tip-to-collector distance of 15 cm, and a spinning period of 2 hours. Subsequently, the fiber diameters were examined using scanning electron microscopy (SEM), with the results shown in Figure 1.

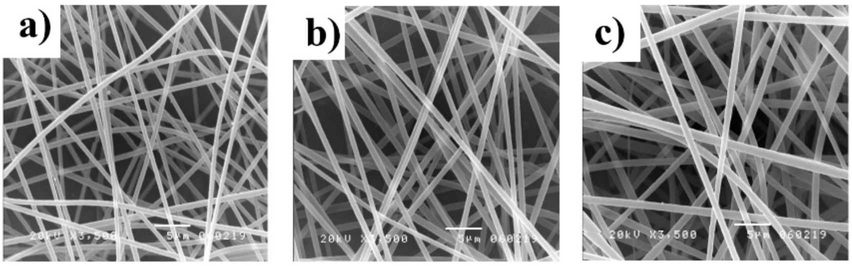


Figure 1. SEM images of cellulose acetate fibermat using electrospinning at different electric potential: (a) 10 kV, (b) 15 kV, and (c) 20 kV, respectively.

The optimal conditions for cellulose acetate fiber mat production are unquestionably 15 kV electrospinning voltage. The observed fiber diameters from SEM images reveal average diameters of 0.666, 0.620 and 0.795 nm and average variations of 0.0814, 0.0468, and 0.1237 for electrospinning voltages of 10, 15, and 20 kV, respectively. Additionally, 15 kV fibers showed the lowest average deviation, indicating more consistent and stable fiber diameters than 10 and 20 kV fibers. This means that a 15 kV electrospinning voltage results in more uniform and stable fiber diameters, showing its suitability for the electrospinning process, which will be studied further.

3.2 Concentration of maleic acid-modified gold (MA-GNP)

A gold nanoparticle solution with a volume of 0.4 milliliters was used to test the synthesis of gold nanoparticles (GNPs) treated with maleic acid. In order to modify the surface, an acid-modifying solution was mixed in a volume of 0.20 milliliters at various concentrations (5, 10, 15, and 20 millimolar). Figure 2(a) shows how the results were analyzed with a UV-Vis Spectrophotometer (Hammami et al., 2021).

The greatest absorption value of nanoparticles made by the citrate reduction technique was observed at 520 nm in wavelength. When utilizing acid concentrations of 10 mM or above, the absorption at 520 nm started to decrease after the surface was modified. This phenomenon can be attributed to the hydrogen bonding that occurs between the particles (Ding et al., 2010). Figure 2(b), which shows the relationship between the absorbance value at 520 nm and the concentration of maleic acid-modifying solution, illustrates this result.

These MA-GNP were selected for further heavy metal ion detection testing since the results also show that a 10 mM of maleic acid concentration is optimum for surface modification.

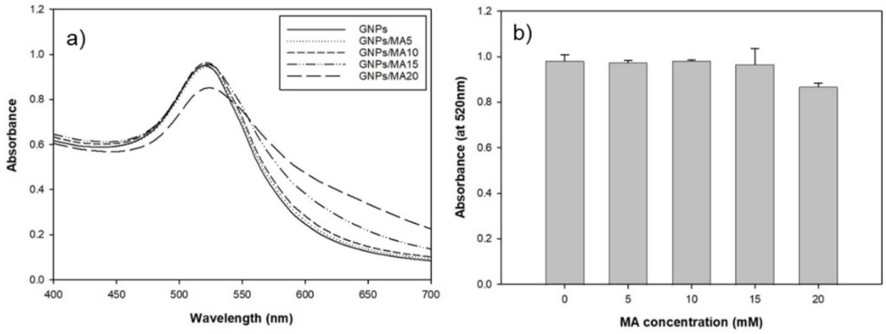


Figure 2. (a) Absorption wavelength range of 400-700 nm of gold nanoparticle with concentration of maleic acid solution from 0 - 20 mM (b) Relationship between acid concentration affecting the light absorption of gold nanoparticles at 520 nm.

4.3 Detection of heavy metal ions

4.3.1 Preliminary detection using MA-GNPs solution

The preliminary heavy metal ion detection tests involved five types of metal ions: Co^{2+}, Cu^{2+}, Hg^{2+}, Ni^{+}, and Pb^{2+}. A volume of 0.4 ml of metal ion solution with a concentration of 50 ppm was added to the MA-GNPs solution with a volume of 0.6 ml. After 15 minutes, the spectral analysis using UV-VIS Spectrophotometer revealed a clear change in the absorption spectra, as depicted in Figure 3.

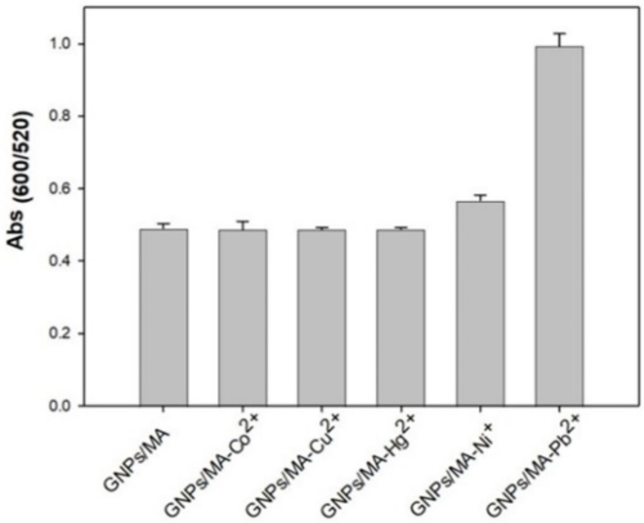


Figure 3. Detection of Co^{2+}, Cu^{2+}, Hg^{2+}, Ni^{+}, and Pb^{2+} ions using MA-GNPs with an acid concentration of 10 mM.

The absorption ratio (A600/520) demonstrated distinct changes after 15 minutes when adding ions such as Co^{2+}, Cu^{2+}, Hg^{2+}, and Ni^{+}. This indicates the high specificity of the detection method for lead ions. It was observed that the absorption ratio increased when the concentration of ions increased**.** The results suggest that MA-GNPs can effectively detect lead ions with high sensitivity and selectivity.

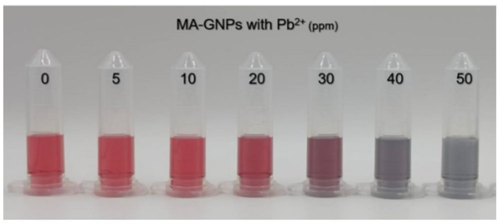


Figure 4. Color change of MA-GNP particles after adding Pb^{2+} ions of different concentrations.

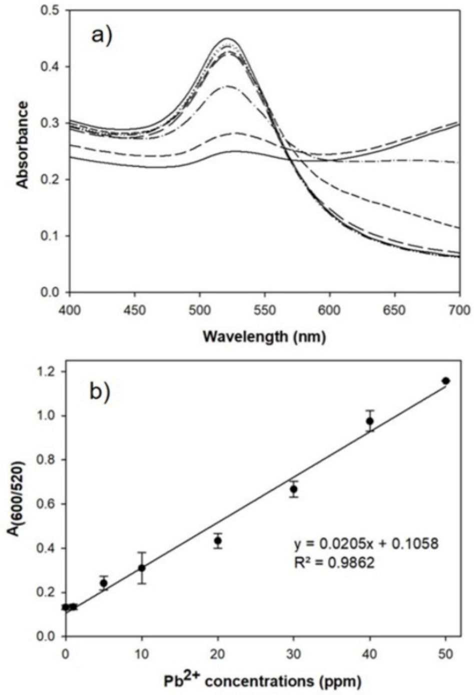


Figure 5. (a) Absorption Spectrums by MA-GNP after the addition of varying concentrations of Pb(II) ions. (b) Linear correlation between the absorbance ratio A600/520 and the concentration of Pb(II) ions in the range of 0-50 ppm.

The investigation of MA-GNPs nanoparticles for detecting lead ions at varied concentrations revealed distinct color intensity levels and change in the absorption spectra. As the concentration of lead ions increases, the color response of the solution shifts from red to blue, as can be observed in Figure 4. The graph indicates the changes in absorption of light values caused by varying lead ion concentrations.

Figure 5(a) indicates the shift in light absorption at 520 nanometers, which reduces, and at 600 nanometers, which increases. This suggests that the ability of MA-GNPs to aggerating lead ions corresponds to their concentration (Li et al., 2014). In addition, Figure 5(b) shows a linear relationship between the absorption ratio (A600/520) and the concentration of lead ions ranging from 10 to 100 ppm (R^2 = 0.9862).

4.3.2 Detection using cellulose acetate fiber mats loaded with MA-GNPs

Lead ions were also detected using cellulose acetate fiber mats coated with MA-GNPs. The mats were made by dropping a 100-µl solution of MA-GNPs into 10x10 mm cellulose acetate fiber mats and air-drying them. Color changes were detected after the electrospun fiber mats were subjected to Pb(II) ion solutions of different concentrations (0, 10, 20, 30, 40, and 50 ppm). The results are shown in Figure 6.

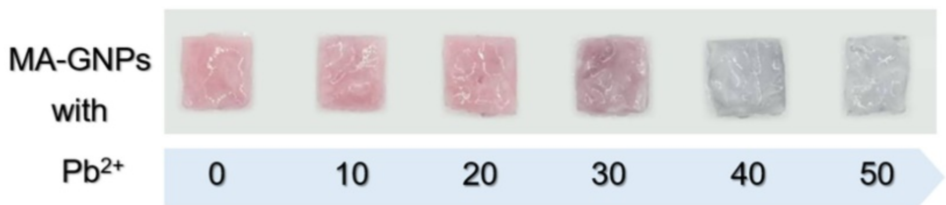


Figure 6. Color change of cellulose acetate fiber mats loaded with MA-GNPs when exposed to citrate ion solutions of different concentrations (0, 10, 20, 30, 40, and 50 ppm).

The fast detection of lead ions by these membranes was observed by the naked eyes and digital camera. The images showed MA-GNPs/cellulose acetate fiber mats with incubated at 0, 10, 20, 30, 40, and 50 ppm. These results confirmed that the color changes from red to purple-blue in cellulose acetate fiber mats loaded with MA-GNPs were the result of the interaction between Pb^{2+} and MA-GNPs in nanofibers.

4.3.3 Detection from sample water

Pb(II) ions were detected using cellulose acetate fiber mats loaded with MA-GNPs in two types of water samples: DI and wastewater from a battery manufacturing company. A 100-micro-liter MA-GNPs solution was dropped onto 10x10 mm cellulose acetate fiber mats and air-dried. The fiber mats were then exposed to the water samples and color changes were noticed. The outcomes are depicted in Figure 7.

Figure 7. exhibits the color change of cellulose acetate fiber mats loaded with MA-GNPs after exposure to various water samples.

The fiber mats changed color from red to purple-blue when exposed to wastewater from the battery manufacturing company, according to the test results. According to the concentration analysis, this color shift was connected with a Pb(II) ions concentration of around 40 ppm. These findings show the new test strips' potential for preliminary detection of citrate ions and suggest further development for broader applications.

4. Conclusions

In summary, electrospun tailored cellulose acetate fiber mats were successfully loaded with selectively manufactured and surface-modified gold nanoparticles (MA-GNPs) of approximately 20 nm. These MA-GNPs changed color from red to purple when exposed to lead ion concentrations ranging from 0 to 100 ppm. The color shift correlated with lead ion concentration in a proportional manner, with an absorbance ratio (A600/520) demonstrating a linear relationship in the 10 to 100 ppm range (R^2 = 0.9862). Furthermore, when exposed to lead ion-containing wastewater, these fiber mats had a similar color-changing response, demonstrating their use as preliminary lead ion detection tools in environmental monitoring applications. This points to intriguing future directions for the development and application of these color-changing strips in a variety of scenarios.

Acknowledgements

Authors wish to express their gratitude to National Research Council of Thailand (NRCT) for their financial support. We appreciate the ongoing collaboration with the Research and Development Institute at Rajamangala University of Technology Rattanakosin, the Organic Chemistry Synthesis Laboratory at Chulabhorn Research Institute, and the Department of Materials Engineering at Rajamangala University of Technology Rattanakosin. Their collective support, resources, and expertise have been instrumental in the success of our research.

ORCID

Pimolpun Niamlang
https://orcid.org/0000-0003-1680-5316

References

Abdullah, N., Yusof, N., Lau, W. J., Jaafar, J., & Ismail, A. F. (2019). Recent trends of heavy metal removal from water/wastewater by membrane technologies. *Journal of Industrial and Engineering Chemistry*, *76*, 17-38. doi:10.1016/j.jiec.2019.03.029

Chansuwan, W. (2022). The design of chemosensors for naked-eye detection of mercury (II) ion. *KKU Science Journal*, *42*(4), 748-760. Retrieved from https://ph01.tci-thaijo.org/index.php/KKUSciJ/article/view/249317

Chen, J. P., & Wang, L. (2004) Characterization of metal adsorption kinetic properties in batch and fixed-bed reactors. *Chemosphere*, *54*, 397-404. doi:10.1016/S0045-6535(03)00714-8

Ding, N., Cao, Q., Zhao, H., Yang, Y., Zeng, L., He, Y., … Wang, G. (2010). Colorimetric assay for determination of lead (II) based on its incorporation into gold nanoparticles during their synthesis. *Sensors*, *10*, 11144-11155. doi:10.3390/s101211144

Feng, X., Long, R., Wang, L., Liu, C., Bai, Z., & Liu, X. (2022). A review on heavy metal ions adsorption from water by layered double hydroxide and its composites. *Separation and Purification Technology*, *284*, 120099. doi:10.1016/j.seppur.2021.120099

Fewtrell, L., Kaufmann, R., & Prüss-Üstün, A. (2003). *Lead: Assessing the environmental burden of disease at national and local levels* (pp. 5-21). Geneva: World Health Organization.

Ghaedi, M., Ahmadi, F., & Shokrollahi, A. (2007). Simultaneous preconcentration and determination of copper, nickel, cobalt and lead ions content by flame atomic absorption spectrometry. *Journal of Hazardous Materials*, *142*, 272-278. doi:10.1016/j.jhazmat.2006.08.012

Hammami, I., Alabdallah, N. M., Al jomaa, A., & Kamoun, M. (2021). Gold nanoparticles: Synthesis properties and applications. *Journal of King Saud University - Science*, *33*(7), 101560. doi:10.1016/j.jksus.2021.101560.

Han, R., Zou, W., Zhang, Z., Shi, J., & Yang, J. (2006). Removal of copper(II) and lead(II) from aqueous solution by manganese oxide coated sand: I. Characterization and kinetic study. *Journal of Hazardous Materials*, *137*, 384-395. doi:10.1016/j.jhazmat.2006.02.021

Huang, Z. M., Zhang, Y. Z., Kotaki, M., & Ramakrishna, S. (2003). A review on polymer nanofibers by electrospinning and their applications in nanocomposites. *Composites Science and Technology*, *63*(15), 2223-2253. doi:10.1016/S0266-3538(03)00178-7

Li, Y., Ding, B., Sun, G., Ke, T., Chen, J., Al-Deyab, S. S., & Yu, J. (2014). Solid-phase pink-to-purple chromatic strips utilizing gold probes and nanofibrous membranes combined system for lead (II) assaying. *Sensors and Actuators B: Chemical*, *204*, 673-681. doi:10.1016/j.snb.2014.08.048

Milne, A., Landing, W., Bizimis, M., & Morton, P. (2010). Determination of Mn, Fe, Co, Ni, Cu, Zn, Cd and Pb in seawater using high resolution magnetic sector inductively coupled mass spectrometry (HR-ICP-MS). *Analytica Chimica Acta*, *665*, 200-207. doi:10.1016/j.aca.2010.03.027

Ratnarathorn, N., Chailapakul, O., & Dungchai, W. (2015). Highly sensitive colorimetric detection of lead using maleic acid functionalized gold nanoparticles. *Talanta*, *132*, 613-618. doi:10.1016/j.talanta.2014.10.024

Sartore, L., Barbaglio, M., Borgese, L., & Bontempi, E. (2011). Polymer-grafted QCM chemical sensor and application to heavy metal ions real time detection. *Sensors and Actuators B: Chemical*, *155*, 538-544. doi:10.1016/j.snb.2011.01.003

World Health Organization. (2004). Guidelines for drinking – Water quality (3rd ed.). Geneva: WHO.

Zhang, N., Wang, X., Ma, C., Qiao, Y., Zhang, H., Shi, C., … Zhong, J. (2019). Electrospun nanofibrous cellulose acetate/curcumin membranes for fast detection of Pb ions. *Journal of Nanoscience and Nanotechnology*, *19*, 670-674. doi:10.1166/jnn.2019.15893

Available online at https://li02.tci-thaijo.org/index.php/ssstj

Suan Sunandha Science and Technology Journal

Nutritional Value, Bioactive Compounds, and Antioxidant Activity of Phak-liang (*Gnetum gnemon* Linn. var. *tenerum* Markgr.) in Surat Thani Province, Thailand

Marisa Intawongse[1], Araya Pranprawit[2*]

[1]Program of Chemistry, Faculty of Science and Technology, Suratthani Rajabhat University,
Khun Ta-Lay, Muang, Surat Thani 84100, Thailand
[2]Program of Community Public Health, Faculty of Science and Technology, Suratthani Rajabhat University,
Khun Ta-Lay, Muang, Surat Thani 84100, Thailand
*Corresponding author e-mail: araya.pra@sru.ac.th

Received: *3 April 2024* / **Revised:** *24 June 2024* / **Accepted:** *9 July 2024*

Abstract

The nutritional value, bioactive compounds, and antioxidant activity of Phak-liang (*Gnetum gnemon* Linn. var. *tenerum* Markgr.) were investigated. The indigenous vegetables were collected from three districts; Muang, Bannasarn, and Khirirat Nikom in Surat Thani Province. The nutritional value of the samples was carried out according to the methods of AOAC. The moisture contents ranged from 84.72-87.52%, and the ash contents ranged from 0.78-1.07% fresh weight. The percentage of fiber, protein, fat, and carbohydrate in the leafy vegetables were in the range of 3.06-3.81, 3.44-4.68, 0.65-1.08 and 7.49-9.55 fresh weight, respectively. The samples were analyzed by ICP-OES for their mineral determinations. The amounts found in the samples, in descending order were: K > Ca > Mg > Mn > Na> Fe > Zn > Cu. Considering the amounts of minerals found in the vegetable and their recommended daily values by the Bureau of Nutrition (Ministry of Public Health, Thailand), the daily value (DV) percentage of Mn in the samples were relatively high (28.45-96.65%), whereas K, Mg, Ca, Fe, Cu and Zn contributed from 1.29–8.42% of the daily values. The vegetable had negligible amounts of Na (0.03 - 0.07% of the DV). The antioxidant capacities were analyzed using Ferric Reducing Antioxidant Power (FRAP) and 2,2-diphenyl-1-picrylhydrazyl (DPPH) assay. The results showed that antioxidant activity increased by the concentration of the ethanol extracts. The FRAP values of the extracts ranged from 89.01-160.70 µM $FeSO_4$/100 g dry weight, and the IC_{50} values calculated for DPPH radical scavenging activities of the extracts ranged from 54.37-60.57 mg/mL. The results suggested that Phak-liang is a potential source of natural phytochemicals and could be used for health promotion in the community.

Keywords: Nutritional value, Bioactive compounds, Antioxidant activity, *Gnetum gnemon* Linn. var. *tenerum* Markgr.

1. Introduction

Considerable attention has been focused on investigating the significant role of various edible plants recognized for their potential in mitigating chronic disease, particularly those abundant in antioxidants (Alam, Saqib, & Ashraf, 2017; Boeing et al., 2012; Pranprawit, 2019; Pranprawit, Heyes, Molan, & Kruger, 2015). A comprehensive review on the associations between the consumption of vegetables and fruits and the risk of several chronic diseases have been documented (Boeing et al., 2012). The findings of this review highlight that increasing the intake of vegetables and fruits from a range of botanical families, which provide sources of nutrients, dietary fiber and phytochemicals reduces the risk of hypertension, coronary heart disease, stroke, certain eye disease, dementia and osteoporosis. The presence of a wide range of bioactive compounds in fruits and vegetables, such as ascorbic acid, tocopherols, phenolic acids, flavonoids, carotenoids, and anthocyanin, has been suggested as the underlying cause for their beneficial effects. These substances possess antioxidant activities and play a significant role in promoting health and preventing chronic diseases, thereby enhancing the functionality of these natural foods (Alfa & Arroo, 2019; Paran, Novack, Engelhard, & Hazan-Halevy, 2009; Zehiroglu & Sarikaya, 2019).

Phak-liang (*Gnetum gnemon* Linn. var. *tenerum* Markgr.) is an indigenous leafy vegetable in the southern region of Thailand. It has gained popularity as a consumable vegetable, with concurrent cultivation observed alongside rubber plantations, oil palm plantations, or fruit orchards. It is widely consumed in various forms, including raw consumption as leafy greens or as an ingredient in culinary preparations. Notably, the leafy vegetable stands out for its exceptional nutritional profile, exhibiting high levels of essential nutrients and bioactive compounds of considerable importance (Kongkachuichai, Charoensiri, Yakoh, Kringkasemsee, & Insung, 2015). Phak-liang vegetable possesses components including β-carotene, lutein, polyphenol, vitamin C, and vitamin E, which exhibit antioxidant properties that may contribute to anti-aging mechanisms and strengthen the immune system (Kongkachuichai et al., 2015). Nevertheless, the comprehensive investigation of the nutritional composition, bioactive compounds, and antioxidant properties of the leafy vegetable remains scarce. Thus, the objectives of this study were to determine the nutritional value of Phak-liang (*Gnetum gnemon* Linn. var. *tenerum* Markgr.) collected from Surat Thani Province, southern Thailand and to assess the bioactive compounds and antioxidant activities inherent in the vegetable extracts. This study endeavors to provide valuable insights into the nutritional composition of the indigenous leafy vegetable, thereby contributing to the existing body of knowledge in this field.

2. Materials and Methods

2.1 Sample collection and preparation

The vegetable samples were randomly collected from 9 different locations (Table 1) in the 3 districts i.e., Muang (S1, S2 and S3), Bannasarn (S4, S5 and S6), and Khirirat Nikom (S7, S8 and S9), Surat Thani Province, southern Thailand in December 2021 to February 2022. Four kilograms of the vegetable from each sampling location were taken to the laboratory in a cool box. The fresh vegetables were washed with tap water to remove dust or contaminant particles and rinsed with deionized water. The edible portion was separated and the deionized water was dried off at room temperature. The samples were dried in an oven at 45°C for 48 hours, ground to fine powder with a grinder. All homogenized samples were kept at 4°C until analysis.

Table 1. Geographical coordinates and locations of the sampling sites.

Sampling site	Geographical coordinate	Area of collection
S1	9°00'57.9"N 99°21'56.0"E	Village No.10, Khun-Ta-Lay Sub-district, Muang District
S2	9°00'32.7"N 99°20'04.4"E	Village No.8, Khun-Ta-Lay Sub-district, Muang District
S3	9°04'18.0"N 99°23'19.4"E	Village No.8, Ma-Kham-Tia Sub-district, Muang District
S4	8°56'24.5"N 99°21'45.4"E	Village No.1, Tung-Tao-Mai Sub-district, Bannasarn District
S5	8°55'51.9"N 99°22'03.5"E	Village No.5, Tung-Tao Sub-district, Bannasarn District
S6	8°58'23.9"N 99°24'09.9"E	Village No.4, Tung-Tao-Mai Sub-district, Bannasarn District
S7	9°03'02.8"N 99°00'38.0"E	Village No.4, Ban-Yang Sub-district, Khirirat Nikom District
S8	9°00'17.0"N 98°55'43.7"E	Village No.6, Tha-Kanhon Sub-district, Khirirat Nikom District
S9	8°59'47.6"N 98°56'37.1"E	Village No.5, Yan-Yao Sub-district, Khirirat Nikom District

2.2 Proximate composition and nutritional analysis

Moisture content was determined by oven drying at a temperature of 100 ± 5°C to a constant weight according to the AOAC method 950.46 (Latimer, 2016). Total ash content of the sample was determined by furnace incineration at 600°C followed by the AOAC method 942.05 (Latimer, 2016). Total fiber content was carried out by digesting the samples with an acid and a base solution, then drying and combusting the residue according to the AOAC method 978.10 (Latimer, 2016) Crude protein content was determined using a carbon/nitrogen determinator according to the combustion method of AOAC 990.03 (Latimer, 2016). The protein content was calculated from total nitrogen content multiplied by 6.25. Crude fats were extracted with hexane solvent in an automated fat extraction system (S306MK Gerhardt, Germany) according to AOAC method 922.06 (Latimer, 2016). Total carbohydrates were calculated by subtracting the sum of percentage of moisture, protein, fat, and ash from 100%. Macroelements and microelements were analyzed by an inductively coupled plasma – optical emission spectrometer (ICP-OES, Prodigy 7, Teledyne Leeman Labs) after the samples were acid digested according to EPA Method 3052 (United States Environmental Protection Agency, 1996) using a microwave system.

2.3 Analysis of bioactive compounds and antioxidant activity

2.3.1 Preparation of crude extract of Phak-liang

Samples of dried Phak-liang were extracted using 95% ethanol. The preparation was done by soaking dried Phak-liang powder in 95% ethanol solvent in a ratio of 1g of powder to 5 mL of solvent and periodically shaking for 24 hours at room temperature. The extracts were then filtered twice through filter paper Whatman#1, and were evaporated using a rotary evaporator at 45°C, then stored in a refrigerator at 4°C for further analysis.

2.3.2 Total phenolic content (TPC)

Total phenolics were measured using the Folin-Ciocalteu procedure (modified from the previous method of Singleton, Orthofer, and Lamuela-Raventós (1999) and Waterhouse (2002). Briefly, 200 µL of extract was added to 200 µL of Folin-Ciocalteu phenol reagent and 800 µL of Milli-Q water, and allowed to react for 5 min. Subsequently, 2 mL of 7.5% (w/v) sodium carbonate solution and 1.6 mL of Milli-Q water were added and incubated for 30 min at room temperature. Absorbance of extracts were spectrophotometrically measured at 765 nm. A standard curve was prepared using different concentration (0-1,000 µg/mL) of gallic acid solution. Total phenolics was calculated as mg gallic acid equivalent (GAE) per 100g of dry weight sample.

2.3.3 Ferric reducing antioxidant power (FRAP) assay

The FRAP assay was carried out according to the previous method of Benzie and Strain (1996) with slight modification. FRAP reagent consisted of 300 mmol/L acetate buffer (pH 3.6), 10 mmol/L TPTZ (2,4,6-tripyridyl-s-triazine) in 40 mmol/L hydrochloric acid, and 20 mmol/L ferric chloride (III) solution at the ratio of 10:1:1 (v/v/v), respectively. The sample extract (200 µL) was added to 1.8 mL of FRAP reagent and 2 mL of Milli-Q water, mixed well and incubated at 37°C for 30 min, then the absorbance was measured spectrophotometrically at 593 nm. A standard curve was prepared using various concentrations of $FeSO_4.7H_2O$. The antioxidant activity was calculated based on the ability of extracts to reduce ferric (III) iron to ferrous (II) iron, and results were expressed as micromole ferrous ion (II) equivalent per g of dry weight sample.

2.3.4 Scavenging of diphenyl-picrylhydrazyl (DPPH) radicals

The ability of the extract to scavenge DPPH-radical was determined using the previous method of Prior, Wu, & Schaich (2005) with slight modification. Briefly, 1 mL of sample extract was added to 2 mL of 0.2 mM DPPH in 95% ethanol and 1 mL of Milli-Q water, mixed well, and incubated in the dark at room temperature for 30 min to allow the reaction to progress to completion before reading the absorbance at 517 nm. The absorbance reading was compared to the standard curve of 0-100 µg/mL of ascorbic acid. The scavenging activity of samples (% inhibition) at various concentrations was calculated and linear regression was performed to determine the concentration of samples that can scavenge DPPH-radical by 50% (IC_{50}).

2.4 Statistical data analysis

Descriptive statistics, including mean, standard deviation, and percentage were analyzed. The data were subjected to Analysis of Variance (ANOVA) to evaluate the variability, and Pearson's correlation coefficient (r) was used to determine linear relationships, with a significance level set at $P < 0.05$.

3. Results

3.1 Nutritional analysis

The vegetables were harvested at approximately 2-3 years of age, having been cultivated in sandy loam soil, and were subjected to the application of a balanced fertilizer formula consisting of either 15-15-15 or 16-16-16. The nutritional composition of Phak-liang vegetable was carried out on fresh weight basis and has shown in Table

2. The moisture contents of the vegetable were in the range of 84.72 - 87.52%, while the ash contents were ranged from 0.78 - 1.07%. The crude fiber, protein, fat, and carbohydrate contents were found between 3.06 - 3.81, 3.44 - 4.68, 0.65 - 1.08, and 7.49 - 9.55% fresh weight, respectively.

Table 2. Nutritional composition of Phak-liang vegetable (*Gnetum gnemon*) (%FW, n = 3).

Sampling site	Moisture	Ash	Fiber	Protein	Fat	Carbohydrate
S1	87.52 ± 0.40	0.85 ± 0.00$_2$	3.30 ± 0.05	3.44 ± 0.02	0.70 ± 0.00$_3$	7.49 ± 0.41
S2	86.90 ± 0.33	0.78 ± 0.00$_3$	3.66 ± 0.03	3.77 ± 0.01	0.68 ± 0.01	7.87 ± 0.33
S3	86.30 ± 0.25	0.92 ± 0.01	3.18 ± 0.03	4.06 ± 0.01	0.76 ± 0.00$_4$	7.96 ± 0.26
Mean±SD	86.91 ± 0.60	0.85 ± 0.06	3.38 ± 0.22	3.76 ± 0.27	0.71 ± 0.04	7.77 ± 0.36
S4	85.39 ± 0.19	1.07 ± 0.01	3.08 ± 0.07	4.68 ± 0.02	1.08 ± 0.01	7.78 ± 0.21
S5	84.76 ± 0.36	0.88 ± 0.01	3.06 ± 0.05	3.81 ± 0.01	1.00 ± 0.01	9.55 ± 0.36
S6	84.72 ± 0.92	0.96 ± 0.01	3.24 ± 0.07	4.52 ± 0.01	0.93 ± 0.01	8.85 ± 0.92
Mean±SD	84.96 ± 0.60	0.97 ± 0.08	3.12 ± 0.10	4.34 ± 0.40	1.00 ± 0.06	8.73 ± 0.92
S7	86.69 ± 0.34	0.79 ± 0.00$_2$	3.18 ± 0.08	3.63 ± 0.01	0.65 ± 0.00$_1$	8.24 ± 0.34
S8	85.79 ± 0.76	0.83 ± 0.00$_2$	3.81 ± 0.09	4.13 ± 0.02	0.75 ± 0.00$_4$	8.50 ± 0.77
S9	85.91 ± 0.24	1.00 ± 0.00$_3$	3.34 ± 0.05	4.24 ± 0.02	0.66 ± 0.00$_3$	8.19 ± 0.22
Mean±SD	86.13 ± 0.60	0.87 ± 0.10	3.44 ± 0.29	4.00 ± 0.28	0.69 ± 0.05	8.31 ± 0.46

3.2 Macroelements and microelements in the vegetable

Macroelements (Na, K, Mg, and Ca) and microelements (Fe, Cu, Zn, and Mn) were measured by an ICP-OES. The mean and standard deviation values of the elements in the vegetable are summarized in Table 3. The results showed that the amounts of macroelements, in descending order, were K, Ca, Mg, and Na with the range of 354.75-558.02, 34.37-64.18, 27.50-45.33, and 0.65-1.35 mg/100g fresh weight, respectively. In addition, the concentrations of microelements, in descending order, were Mn, Fe, Zn, and Cu, ranging from 1.76-5.18, 0.58-0.93, 0.39-0.52, and 0.05-0.08 mg/100g, respectively.

Table 3. Macroelements and microelements present in Phak-liang vegetable (*Gnetum gnemon*) and percentage daily value.

Sampling Site	Mean ± SD (mg/100g FW, n = 3)							
	Na	K	Mg	Ca	Fe	Cu	Zn	Mn
S1	1.35 ± 0.06	421.91 ± 11.27	43.57 ± 2.07	45.56 ± 1.57	0.58 ± 0.03	0.05 ± 0.01	0.43 ± 0.03	2.00 ± 0.15
S2	1.04 ± 0.04	354.75 ± 11.45	40.28 ± 1.37	59.39 ± 4.04	0.68 ± 0.03	0.08 ± 0.01	0.52 ± 0.03	2.86 ± 0.09
S3	0.98 ± 0.03	440.12 ± 20.05	40.85 ± 1.64	59.92 ± 3.67	0.85 ± 0.06	0.07 ± 0.01	0.46 ± 0.03	3.05 ± 0.08
Mean±SD	1.12 ± 0.18	405.59 ± 41.00	41.57 ± 2.13	54.95 ± 7.60	0.70 ± 0.12	0.07 ± 0.01	0.47 ± 0.05	2.64 ± 0.49
S4	0.73 ± 0.06	558.02 ± 19.44	42.09 ± 1.75	43.32 ± 2.60	0.71 ± 0.07	0.05 ± 0.01	0.47 ± 0.03	1.93 ± 0.08
S5	1.09 ± 0.03	398.44 ± 12.72	36.95 ± 1.92	64.18 ± 3.67	0.93 ± 0.05	0.06 ± 0.01	0.43 ± 0.03	3.70 ± 0.20
S6	0.65 ± 0.06	504.49 ± 15.17	40.00 ± 2.53	42.59 ± 2.97	0.65 ± 0.05	0.07 ± 0.01	0.39 ± 0.02	1.92 ± 0.08
Mean±SD	0.82 ± 0.21	486.98 ± 71.69	39.68 ± 2.88	50.03 ± 10.95	0.76 ± 0.14	0.06 ± 0.01	0.43 ± 0.04	2.52 ± 0.89
S7	0.98 ± 0.05	442.49 ± 17.51	27.50 ± 2.35	34.37 ± 2.82	0.58 ± 0.02	0.05 ± 0.01	0.44 ± 0.05	1.76 ± 0.11
S8	0.90 ± 0.05	475.10 ± 14.33	43.94 ± 2.82	58.59 ± 2.64	0.58 ± 0.03	0.05 ± 0.01	0.48 ± 0.04	1.92 ± 0.10
S9	1.05 ± 0.04	505.96 ± 18.37	45.33 ± 2.66	54.46 ± 2.55	0.70 ± 0.03	0.06 ± 0.01	0.49 ± 0.05	5.18 ± 0.43
Mean±SD	0.97 ± 0.07	474.51 ± 31.11	38.92 ± 8.88	49.14 ± 11.46	0.62 ± 0.06	0.06 ± 0.01	0.47 ± 0.04	2.95 ± 1.68
Recommended daily value, mg	910	2925	280	900	15	1.5	10	2.5
mg per serving*	0.26 - 0.61	148.06 - 246.21	10.67 - 20.25	13.40 - 29.32	0.23 - 0.42	0.02 - 0.04	0.16 - 0.24	0.71 - 2.42
Daily value, %**	0.03 - 0.07	5.06 - 8.42	3.81 - 7.23	1.49 - 3.26	1.56 - 2.79	1.29 - 2.41	1.59 - 2.37	28.45-96.65

* 1 serving size = 43 g fresh weight, based on the food exchange list that provides 25 kcal (Kongkachuichai et al., 2015)

** Daily value (%) = mineral amounts in the vegetable sample / recommended daily value x 100

3.3 Total phenolic content and antioxidant activity

3.3.1 Total phenolic content (TPC)

For the total phenolic content (TPC), each Phak-liang vegetable was extracted using 95% ethanol. The results are shown in Table 4. The TPC of ethanoic extract of Phak-liang vegetable collected from Muang district (S1-S3) ranged from 121.26 to 149.53 mg GAE/100 g dried weight, while these values were between 73.21 to

102.51 mg GAE/100 g dried weight in Phak-liang vegetable collected from Bannasarn district (S4-S6) and ranged from 109.15-127.79 mg GAE/100 g dried weight in Phak-liang vegetable collected from Khirirat Nikom district (S7-S9).

3.3.2 Ferric reducing antioxidant power (FRAP)

The antioxidant activity was determined using two commonly used methods: ferric reducing antioxidant power (FRAP), and scavenging of diphenyl-picrylhydrazyl (DPPH) radicals. In this study, ethanoic extract of Phak-liang vegetable collected from 3 different sites- Muang (S1-S3), Bannasarn (S4-S6), and Khirirat Nikom (S7-S9) in Surat Thani province had the reducing ability, which was measured by FRAP, in the range of 128.81-132.46, 89.01-101.59 and 131.02-160.70 µM $FeSO_4$ per 100 g dried weight, respectively (Table 4).

3.3.3 Scavenging of diphenyl-picrylhydrazyl (DPPH) radicals

The ability of ethanoic extract of Phak-liang vegetable to scavenge DPPH-radicals increases with the concentration of the extract. The vegetable collected from Muang district (S1-S3) at the concentration of 100 mg/mL possess the ability to scavenge DPPH-radicals (Table 4) between 74.86% and 75.47% inhibition, whereas Phak-liang from Bannasarn (S4-S6) and Khirirat Nikom (S7-S9) exhibited inhibition in the range of 65.20% to 68.52%, and 71.30% to 72.71%, respectively. When IC_{50} was calculated to assess the concentration of the extract required to scavenge 50% of the initial DPPH radicals. The lower the IC50 value, the more potent is the extract at scavenging DPPH and a higher antioxidant activity. It can be found that ethanoic extract of Phak-liang from Muang district (S1-S3), Bannasarn (S4-S6), and Khirirat Nikom (S7-S9) contained IC_{50} values in the range of 54.368 ± 0.911, 60.573 ± 4.293 and 57.569 ± 1.702 mg/mL, respectively. However, the scavenging ability of the extracts across all areas collected was significantly lower ($P< 0.05$) when compared to the standard ascorbic acid (0.022 ± 0.000 mg/mL).

Table 4. Total phenolic content, ferric reducing antioxidant power, and scavenging of DPPH-radical in ethanolic extract of Phak-liang vegetable (n = 3).

Sampling site	Total phenolic content (mg GAE/100g DW)	FRAP value (µM $FeSO_4$/100g DW)	DPPH (% inhibition)*	DPPH IC_{50} (mg/mL)
S1	126.19 ± 3.57	128.81 ± 4.96	75.47	53.315
S2	121.26 ± 1.67	132.46 ± 5.50	74.86	54.907
S3	149.53 ± 4.50	129.47 ± 7.67	75.33	54.880
Mean±SD	132.33 ± 15.10	130.24 ± 1.94	75.22 ± 0.32	54.368 ± 0.911**
S4	73.21 ± 1.56	89.01 ± 4.35	65.20	60.774
S5	102.51 ± 1.99	101.59 ± 0.31	68.52	64.762
S6	90.29 ± 2.14	93.03 ± 2.61	67.85	56.183
Mean±SD	88.67 ± 14.72	94.55 ± 6.43	67.19 ± 1.76	60.573 ± 4.293**
S7	127.79 ± 2.70	160.70 ± 2.80	72.10	59.096
S8	109.15 ± 1.33	135.56 ± 5.33	71.30	55.735
S9	121.33 ± 2.55	131.02 ± 1.86	72.71	57.877
Mean±SD	119.42 ± 9.47	142.43 ± 15.99	72.04 ± 0.71	57.569 ± 1.702**
Standard : Ascorbic acid				0.022 ± 0.000

* Indicate scavenging activity of DPPH-radical of Phak-liang at the concentration of 100 mg/mL

**Indicate statistically significant differences from standard (ascorbic acid) at $p<0.05$

4. Discussion

Proximate composition and nutritional analyses including moisture, ash content, fiber, crude protein, crude fat, and carbohydrates, were carried out in this study. With the exception of fat content, these results were in agreement with a previous finding indicated that the moisture content, protein, fat and carbohydrate of Phak-liang vegetable collected from 6 Provinces in southern Thailand were 85.66, 4.23, 0.58, and 8.51% fresh weight, respectively (Kongkachuichai et al., 2015). Additionally, both macroelements (Na, K, Mg, and Ca) and microelements (Fe, Cu, Zn, and Mn) were examined. This revealed an alignment between the amounts of Mg and Ca in the present study and those of the examined vegetables collected from six provinces in the southern region of Thailand. Nevertheless, the concentrations of K and Fe in the vegetable were found to be comparatively higher in relation to the results obtained from the previous findings (Kongkachuichai et al., 2015). This could be attributed to variations in the sources of the vegetables, which come from different cultivation areas with diverse soil conditions and agricultural management practices, such as the application of dissimilar chemical or organic fertilizers.

Research on mineral content in Phak-liang vegetable is relatively limited. Therefore, additional findings were presented by referencing studies conducted on other types of green leafy vegetables. A study reported the amounts of mineral in 15 leafy *Amaranthus* species., which indicated that Na concentrations were in consistent with the findings of this research. However, the concentrations of Ca, K, Mg, Cu, Fe and Zn present in the vegetables were relatively higher compared to this study, whereas Mn exhibited lower levels (Jiménez-Aguilar & Grusak, 2017). In addition, a study conducted on minerals in leafy vegetables of *Amaranthus blitum* reported that the levels of Mn were in agreement with the findings of this research, while the levels of Ca, Mg, Cu, Fe, and Zn were higher compared to the results obtained from this study. Conversely, the levels of K were found to be lower than those observed in this research (Sarker & Oba, 2020).

When considering the levels of essential and trace minerals present in the vegetable, a comparison was made with the recommended daily intake of minerals, which is expressed as a percentage daily value (Bureau of Nutrition, 2020) shown in Table 3. It was found that Mn in the samples had a high percentage daily value (28.45-96.65%), whereas the six minerals (K, Mg, Ca, Fe, Cu, and Zn) in the samples had relatively low percentage daily value (1.29-8.42%). In the case of Na, the percentage of the recommended daily intake is remarkably low (0.03-0.07%). Therefore, Phak-liang vegetable is considered a significant source of essential and trace minerals, especially for Mn.

Phak-liang vegetables are commonly found in local areas of the southern Thailand, however previous studies on bioactive compounds and antioxidant activity in Phak-liang vegetables are relatively limited. The study of Pichairat and Mahea (2014) investigated the phenolic content and antioxidant capacity of *Gnetum gnemon* Linn. var. *tenerum* Markgr. collected from Yantakhao district, Sikao district, Palian district and Nayong district in Trang province. The average total phenolic content in Phak-liang vegetables was 180.95 ± 2.84 mg GAE/100 g fresh weight, and the vegetables also contained antioxidant activity of 26.87 ± 0.40 and 17.08 ± 0.51 mg GAE/100 g fresh weight as measured by DPPH and FRAP assay, respectively. For the study with regards to indigenous vegetables in Surat Thani province, Yakoh, Weschasat, & Suwannachote (2017, 2018) analyzed the phenolic content and antioxidant activity by FRAP assay of Phak-liang vegetables collected from Muang district. In these two studies, total phenolic content of Phak-liang vegetables was compared between summer and rainy season, and found that in rainy season, Phak-liang vegetables had the TPC value of 279.23 ± 47.70 mg GAE/100 g, which exhibited the antioxidant activity (FRAP assay) in the average of 883.80 ± 89.6 µmol TE/100g. While in summer season, the TPC value found in Phak-liang was increased to 584.73±53.67 mg GAE/100 g and exhibited antioxidant activity of 430.83±203.20 µmol TE/100 g.

The total phenolic content of Phak-liang vegetables found in the study of Yakoh et al. (2018) is relatively higher than the results of TPC (132.33 ± 15.10 mg GAE/100 g in Muang district) found in this study, which collected the samples in rainy season between December to February. The reason may be due to the difference in the preparation and extraction methods between the study. This study used the hot-air oven, which may cause to loss of some phenolic compounds during the drying process, and leading to lower TPC values than the study of Yakoh et al. (2018) which used the freeze drier. However, it should be noted in this study that among various indigenous vegetables, *Gnetum gnemon* Linn. var. *tenerum* Markgr. (Phak-liang) contained higher total phenolic compounds than *Musa acuminate* Colla (Hua-plee) (23.5 mg GAE/100 g) *Momordica charantia* Linn. (Mara-kee-nok) (69.5 mg GAE/100 g) found in Surat Thani province (Pranprawit, 2019), *Lasia spinosa* Thw. (Phak-naam) (50.14 mg GAE/100 g) and *Antidesma ghaesembills* Gaerth. (Mao-khai-pla) (64.03 mg GAE/100 g) in Trang province (Pichairat & Mahea, 2014). Furthermore, Phak-liang vegetable also had higher total phenolic compounds than some vegetables that are commonly consumed in daily life, such as cauliflower (81 mg GAE/100 g) cabbage, round (46 mg GAE/100 g) cucumber (17 mg GAE/100 g) and carrot (16 mg GAE/100 g) (Isabelle et al., 2010).

The results of this study are consistent with previous reports (Charoenteeraboon et al., 2019; Pakdee, Poowanna, & Prathumtet, 2021; Pranprawit et al., 2015; Thonginla, Wanwimolruk, & Chuaybamroong, 2014) in finding a significant linear positive relationship between total phenolic content and antioxidant activity in various plants and vegetables. These findings clearly demonstrated that the phenolic compound such as flavonoids and phenolic acid etc., which commonly found in green leafy vegetables including Phak-liang could contribute to their total antioxidant ability and health benefits. The study of Kato, Tokunaga, and Sakan (2009) found an abundance of stilbenoids (namely resveratrol), saponins, flavonoids, and tannins in seeds of *G. gnemon* collected from Indonesia. Moreover, young leaves of *G. gnemon* are an excellent source of polyphenols (253.45±5.68 mg GAE/100 g) (Kongkachuichai et al., 2015). In this study, there was a moderately high correlation between total phenolics and FRAP values ($r = 0.748$, $p < 0.01$), and a high positive correlation between total phenolics and DPPH values ($r = 0.915$, $p < 0.01$), suggesting that the phenolic compounds are likely to be the major contributors to the antioxidant activity, moreover, the phenolic concentration in Phak-liang extracts is more likely to exert a radical scavenging ability rather than the reducing ability. The results of this study suggested that Phak-liang vegetables, which is an indigenous plant in Southern Thailand, are a good source of phenolic compounds and effective antioxidants. Therefore, regular intake should be encouraged among consumers for health promotion in local community.

5. Conclusion

This research examined the leaf extracts of *Gnetum gnemon* Linn. var. *tenerum* Markgr. collected from 3 districts; Muang, Bannasarn, and Khirirat Nikom in Surat Thani province for their nutritional values, bioactive compounds and antioxidant activity. In general, Phak-liang vegetable contain considerable amounts of crude fiber (3.06 - 3.81% FW), protein (3.44 - 4.68% FW) and carbohydrate (7.49 - 9.55% FW). it is also considered a significant source of essential and trace minerals, such as K, Ca, Mg, especially for Mn which contributed from 28.45-96.65% DV recommended by the Bureau of Nutrition (Ministry of Public Health, Thailand). Furthermore, the ethanolic extract of Phak-liang vegetable showed high amounts of phenolic compounds (73.21-149.53 mg GAE/100 g DW) and antioxidant activity which analyzed by FRAP (89.01-160.70 µM $FeSO_4$/100 g DW) and DPPH assay (54.37-60.57 mg/mL), and the phenolic content found in the extract of Phak-liang vegetable were positively related to their antioxidant activity.

The results of this study regarding the nutritional value, bioactive compound and antioxidant efficacy of this indigenous vegetable can be used for further development as a dietary supplement, and will also be useful for consumption recommendation in communities for the prevention and treatment of chronic diseases.

Acknowledgement

The Office of Thailand Science Research and Innovation and Suratthani Rajabhat University are acknowledged for financial support of this research.

References

Alam, F., Saqib, Q. N., & Ashraf, M. (2017). *Gaultheria trichophylla* (Royle): a source of minerals and biologically active molecules, its antioxidant and anti-lipoxygenase activities. *BMC Complementary and Alternative Medicine*, *17*, 1-9. doi:10.1186/s12906-016-1511-4

Alfa, H. H., & Arroo, R. R. J. (2019). Over 3 decades of research on dietary flavonoid antioxidants and cancer prevention: What have we achieved? *Phytochemistry Reviews*, *18*, 989-1004. doi:10.1007/s11101-019-09632-0

Benzie, I. F. F., & Strain, J. J. (1996). The ferric reducing ability of plasma (FRAP) as a measure of "antioxidant power": the FRAP assay. *Analytical Biochemistry*, *239*(1), 70-76. doi:10.1006/abio.1996.0292

Boeing, H., Bechthold, A., Bub, A., Ellinger, S., Haller, D., Kroke, A., … Watzl, B. (2012). Critical review: Vegetables and fruit in the prevention of chronic diseases. *European Journal of Nutrition*, *51*, 637-663. doi:10.1007/s00394-012-0380-y

Bureau of Nutrition. (2020). *Dietary reference intake for Thais 2020*. Bangkok: Department of Health, Ministry of Public Health.

Charoenteeraboon, J., Ngawhirunpat, T., Chaidedgumjorn, A., Asavapichayont, P., Mahadlek, J., & Tuntarawongsa, S. (2019). Total phenolic content and antioxidant activities of different parts of Hom Thong banana extract and extraction related factors. *Thai Bulletin of Pharmaceutical Sciences*, *14*(2), 47-60.

Isabelle, M., Lee, B. L., Lim, M. T., Koh, W. P., Huang, D., & Ong, C. N. (2010). Antioxidant activity and profiles of common vegetables in Singapore. *Food Chemistry*, *120*, 993-1003.

Jiménez-Aguilar, D. M., & Grusak, M. A. (2017). Minerals, vitamin C, phenolics, flavonoids and antioxidant activity of *Amaranthus* leafy vegetables. *Journal of Food Composition and Analysis*, *58*, 33-39. doi:10.1016/j.jfca.2017.01.005

Kato, E., Tokunaga, Y., & Sakan, F. (2009). Stilbenoids isolated from the seeds of melinjo (*Gnetum gnemon* L.) and their biological activity. *Journal of Agricultural and Food Chemistry*, *57*(6), 2544-2549. doi:10.1021/jf803077p

Kongkachuichai, R., Charoensiri, R., Yakoh, K., Kringkasemsee, A., & Insung, P. (2015). Nutrients value and antioxidant content of indigenous vegetables from Southern Thailand. *Food Chemistry*, *173*, 838-846. doi:10.1016/j.foodchem.2014.10.123

Latimer, G. W. (2016). *Official methods of analysis of AOAC International* (20th ed.). Maryland, USA: AOAC International.

Pakdee, N., Poowanna, R., & Prathumtet, J. (2021). Comparison of antioxidant activity and amount of total phenolic compound from watermelon seeds. *Huachiew Chalermprakiet Science and Technology Journal*, *7*(2), 52-60.

Paran, E., Novack, V., Engelhard, Y. N., & Hazan-Halevy, I. (2009). The effects of natural antioxidants from tomato extract in treated but uncontrolled hypertensive patients. *Cardiovascular Drugs and Therapy*, *23*, 145-151. doi:10.1007/s10557-008-6155-2

Pichairat, D., & Mahea, N. (2014). Effect of blanching on the phenolic content and antioxidant capacities of some Southern Thai indigenous vegetables. *Rajamangala University of Technology Srivijaya Research*, *6*(2), 36-46.

Pranprawit, A. (2019). Local vegetables traditionally used for lowering hyperglycemia in Surat Thani Province, Thailand. *International Journal of Applied Pharmaceutics*, *11*(S3), 3-6. doi:10.22159/ijap.2019.v11s3.M0008

Pranprawit, A., Heyes, J. A., Molan, A. L., & Kruger, M. C. (2015). Antioxidant activity and inhibitory potential of blueberry extracts against key enzymes relevant for hyperglycemia. *Journal of Food Biochemistry*, *39*(1), 109-118. doi:10.1111/jfbc.12094

Prior, R. L., Wu, X., & Schaich, K. (2005). Standardized methods for the determination of antioxidant capacity and phenolics in foods and dietary supplements. *Journal of Agricultural and Food Chemistry*, *53*(10), 4290-4302. doi:10.1021/jf0502698

Sarker, U., & Oba, S. (2020). Nutrients, minerals, pigments, phytochemicals, and radical scavenging activity in *Amaranthus blitum* leafy vegetables. *Scientific Reports, 10,* 3868. doi:10.1038/s41598-020-59848-w

Singleton, V. L., Orthofer, R., & Lamuela-Raventós, R. M. (1999). Analysis of total phenols and other oxidation substrates and antioxidants by means of Folin-Ciocalteu reagent. *Methods in Enzymology*, *299*, 152-178. doi:10.1016/S0076-6879(99)99017-1

Thonginla, V., Wanwimolruk, C., & Chuaybamroong, P. (2014). Correlation between antioxidant capacities of fruits analyzed with DMPD and phenolic contents, vitamin C, vitamin E and beta-carotene. *Burapha Science Journal*, *19*(2), 93-104.

United States Environmental Protection Agency. (1996). *Method 3052: Microwave assisted acid digestion of siliceous and organically based matrices*. Retrieved from https://www.epa.gov/sites/default/files/2015-12/documents/3052.pdf

Waterhouse, A. L. (2002). Determination of total phenolics. *Current Protocols in Food Analytical Chemistry*. New York: John Wiley & Sons.

Yakoh, K., Weschasat, T., & Suwannachote, P. (2017). Antioxidant activity of indigenous vegetables in Suratthani Province. *Proceeding of National & International Conference: Research 4.0 innovation and development SSRU's 80th anniversary* (pp. 233-243). Bangkok, Thailand.

Yakoh, K., Weschasat, T., & Suwannachote, P. (2018). Phenolic compound of 10 indigenous vegetables from Surat Thani Province. *Journal of Thai Traditional and Alternative Medicine, 16*(2), 185-194.

Zehiroglu, C., & Sarikaya, S. B. O. (2019). The importance of antioxidants and place in today's scientific and technological studies. *Journal of Food Science and Technology*, *56*, 4757-4774. doi:10.1007/s13197-019-03952-x

Available online at https://li02.tci-thaijo.org/index.php/ssstj

Suan Sunandha Science and Technology Journal
©2024 Faculty of Science and Technology, Suan Sunandha Rajabhat University

Developing an Innovative Health Information Service System: The Potential of Chatbot Technology

Ittikorn Klinthaisong, Narawit Subinnam, Attapol Kunlerd*, Chonlada Muangsri, Sakchan Luangmaneerote

Depertment of Computer Technology, Faculty of Agriculture and Technology,
Rajamangala University of Technology Isan Surin Campus, 145 Moo 15, Nok Mueang, Mueang, Surin 3200, Thailand
***Corresponding author e-mail: Attapol.ku@rmuti.ac.th**

Received: *2 May 2024* / **Revised:** *28 June 2024* / **Accepted:** *7 July 2024*

Abstract

This study is a research and development with 3 main objectives: (1) to design and to develop a health management system; (2) to evaluate the efficiency and acceptability of technology; and (3) to compare the average Body Mass Index (BMI) before and after utilizing the developed health management system guided by the principle of the wellness plan. The study was divided into 3 phases: (1) study phase that conducts semi-structured interview with a group of students and healthcare professionals, along with literature review; (2) development phase that the health management system used an adapted waterfall model, dividing the system into two parts: website and Line Chatbot; and (3) experimentation and evaluation phase which the data were collected by using questionnaires that were used and analyzed by using descriptive statistics. This study found that the developed health management system was evaluated at a high level by 3 experts. Additionally, the system acceptability was evaluated at a high level by sample group of 33 individuals. The result of research showed that the developed system was able to be used as a tool for primary health management in the areas of diet planning and exercise including to the ability to track and analyze your own health.

Keywords: Health informatics, Healthcare, Digital health, Line Chatbot, Web development

1. Introduction

Healthcare involves the maintenance and improvement of both physical and mental aspects of well-being to sufficiently meet individual needs and functions. In addition to the treatment of diseases and illnesses, healthcare also focuses on disease prevention and creating a conducive environment for the body and mind. The foundation of good health begins with health literacy, which refers to the individual's ability or potential to understand health information and make appropriate decisions about their health. Promotion of health literacy enhances individual's knowledge and understanding of health information and the benefits of self-care. For example, exercising not only enhances the body but also reduces the risk of Alzheimer's disease (Department of Disease Control, 2021a). Overall, good health is primarily attributed to both physical and mental well-being. Mental health refers to the ability to handle one's own problems appropriately without causing trouble for others. Meanwhile, physical health is related to aspects such as Body Mass Index (BMI), cholesterol levels, blood pressure, and health examination (Faculty of Dentistry, 2019). Generally, understanding physical health is fundamental knowledge that people of all ages can acquire. However, promoting the education and programs regarding habits, diets, healthy exercise, weight control, as well as attitudes, enhances individual health behaviors (Phokhwang, Sarakshetrin, Ngamwongwiwat, & Pramnoi, 2023). In medical aspect, several indicators are used in primary healthcare management, including BMI, Basal Metabolic Rate (BMR), and Total Daily Energy Expenditure (TDEE). BMI serves as a standard indicator to evaluate physical condition by assessing the balance between body weight and height (Department of Disease Control, 2021b). BMR represents the minimum number of daily calories needed to maintain homeostasis. Meanwhile, TDEE refers to the total daily energy expenditure, including both daily energy requirement and physical activity energy. TDEE can be calculated by adding the BMR value to the number of calories expended through physical activity within a 24-hour period (Lalam, Chaimai, & Fukfon, 2022). These indicators enable individuals to plan their appropriate diet and exercise routines to achieve and control their desired weight.

The education in university level is one of the most important periods for personal physical and mental health development. Healthcare during this period is paramount, as it profoundly impacts learning effectiveness and ability to manage daily stressors and problems (Ruanphet, Khamanek, Ngomsangad, Hiranwitchayakun, & Suwannawong, 2023). In high-ranking educational institutions, students often face considerable academic pressure. Inadequate stress management potentially harm student health. Therefore, appropriate physical and mental health management is essential for students to achieve effective performance and prevent physical health-related issues. Generally, universities promote the student exercise activities in many aspects, such as providing gym and sport facilities, and offer courses related to healthcare (RMUTI Surin, 2022). On the other hand, proper diet and exercise education are crucial for students. Lack of healthcare knowledge and time management is commonly cited as major problems among students during their university education. Additionally, many students do not prioritize their health properly, which leads to various diseases such as obesity, heart diseases, and diabetes, consequently impacting their study or work performance (Kimsungnoen, Setthawong, Boonya, Yantaporn, & Dechjob, 2023). However, a previous study has shown that developing health literacy has been an effective way to help students acquire knowledge and develop positive attitudes, leading to behavioral changes (Khamngoen et al., 2023).

Therefore, this research introduces a healthcare management method based on the concept of wellness plan (Department of Health, 2022) and health informatics. This healthcare management system combines information technology and clinical knowledge to improve and solves problems in the healthcare system (Nelson, 2020) by using basic health indicators, including BMI, BMR, and TDEE. This system presents personalized diet and exercise plans to encourage individuals to access their health information anytime and anywhere. This results in increased communication with medical and fitness professionals. Additionally, this system can be used as a tool for in-depth collection and analysis of personal health information and for planning health promotion policies. This developed healthcare management system utilizes 2 technologies, consisting of a health information management website and LINE Chatbot. Users record their basic health information (e.g. age, sex, weight, height, and daily activity) on the health information management website. The LINE Chatbot enables users to inquire about diet and exercise information, restaurant recommendation and fitness locations in the study area (Surin Municipality).

2. Research Objectives

2.1 To design and develop a health management system

2.2 To evaluate the efficiency and acceptability of the developed health management system technology

2.3 To compare BMI before and after using the developed health management system

3. Materials and Methods

3.1 Scope of study

The samples for this study were selected using purposive sampling techniques (Wadeejaroen, Lertnaisat, & Teekasap, 2017). The samples were divided into 2 groups: Group 1 consisted of 3 experts (including 1 professional nurse, 1 trainer, and 1 information specialist); and Group 2 consisted of 30 students (Year 1-4 in academic year 2022) from the Department of Computer Technology, School of Agriculture and Technology, Rajamangala University of Technology, Surin campus, Thailand. Additionally, the selected individuals were those who were interested in healthcare in terms of diet and exercise and attended the university gym at least once.

Moreover, the scope of content and period, this research presented health care in terms of diet and exercise plans that are appropriate for each individual's behavior by collecting information on restaurants and exercise venues in the Surin Municipality area. The research was conducted between March 2021 - April 2022.

3.2 Design and development of health management system

This study was divided into 3 phases:(1) study and research phase; (2) platform design and development phase; and (3) monitoring and evaluation phase. The details of each phase are described below:

Phase 1: Study and research

The needs of users were gathered in this phase. Data were collected from various sources, including: (1) research and healthcare theories in the areas of wellness plan, diet, and exercise (Department of Health, 2022). (2) open-ended and closed-ended questionnaires administered to the selected sample population; and (3) unstructured interview (Wadeejaroen et al., 2017) with 2 health care professionals (a professional nurse and a trainer).

Phase 2: Platform design and development

This phase comprised of 2 steps as follows:

Step 1 which focused on design of health management system consisted of 2 parts. Part 1 focused on the development of the health management website using a client-server architecture system (Nyabuto, Mony, & Mbugua, 2023). Users were divided into 4 levels: (1) system administrator, (2) general user, (3) member, and (4) expert. Users were able to access the system via a web browser. The functionality was analyzed by the development team based on data from phase 1 to ensure the consistency with health theories of diet and exercise, and to meet the need of users as much as possible. Part 2 focused on health management using LINE Chatbot. This served as an interface between system and member users, providing real-time information to users at any time. Both systems utilized the same database in order to reduce data storage conflicts. The system architecture is shown in Figure 1. In order to use the system, users were required to register in the system by filling in basic information (name, surname, sex, weight, and height), the system then calculated BMI and displays preliminary health information and proportional assessment (underweight, normal weight, overweight, or obese). Then the users uploaded their daily activity information, and the system displayed the calories content attributed to such activities. Additionally, the system displayed recommended foods, with information on the calories that users required to meet their needs. Users watched videos of exercises which focused on specific parts of the body as shown in Figure 2.

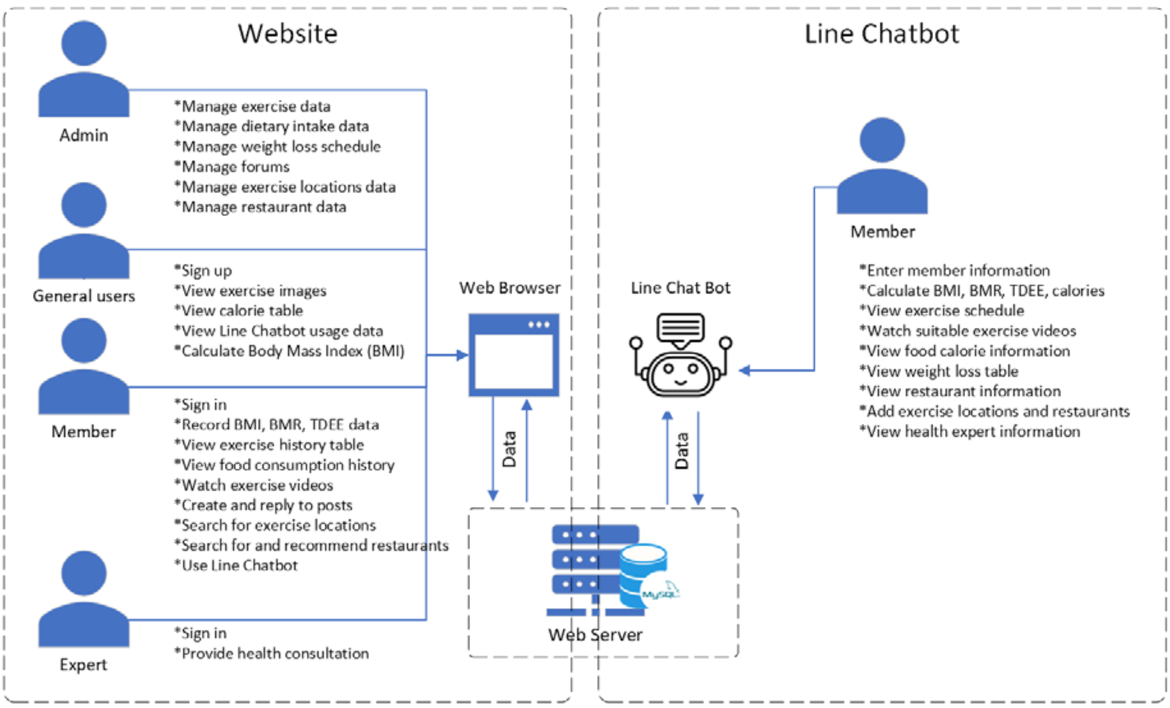


Figure 1. Architecture system of health management system.

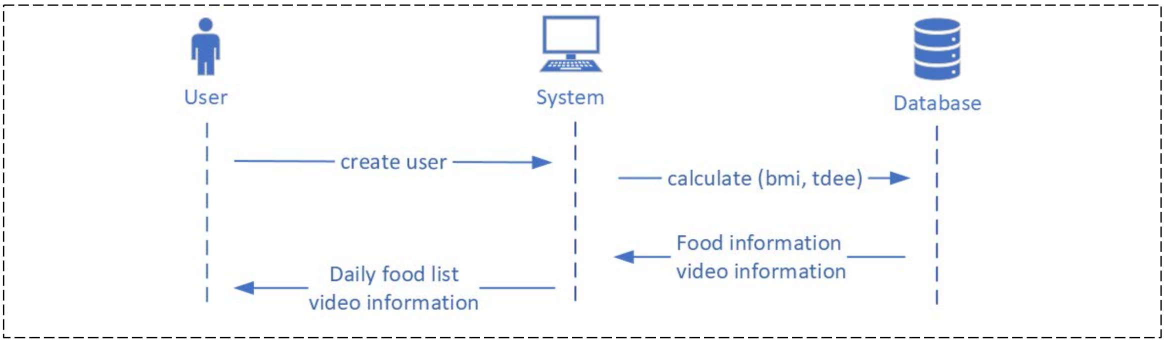


Figure 2. The flowchart of system process of health management system.

Step 2 which focused on the system development consisted of 2 parts. Part 1 involved the development of health management system website. The Front-End technology used HTML, JavaScript, JQuery, CSS, Ajax [Reference]. The Back-End technology utilized PHP and SQL. Additionally, MySQL was used as database, and it was managed by phpMyAdmin (Lamsal, 2020). The server model development involved server simulation by using the XAMPP program (Apache Friends, 2024). Subsequently, the developed system was installed on a server using the Ubuntu server as the operating system (Canonical Ltd., 2024). In addition, Part 2 focused on the development of Line Chatbot by using Dialog flow (Google, 2024). The overview of technologies which were used in the system development is illustrated in Figure 3.

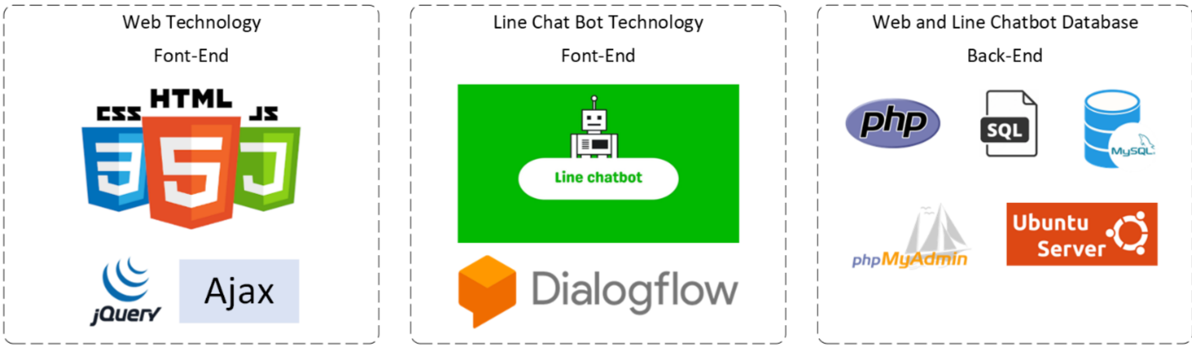


Figure 3. The overview of technologies used in the system development.

The system testing was conducted by 3 information system development professionals, including 2 developers and an independent tester (Assawamekin, 2015, p. 265). The Black-Box Testing technique, performed at a unit-level, was used to examine whether the function of system in Figure 2 met the expectation (Assawamekin, 2015, p. 282). The development team addressed any incomplete functionality. Afterwards, the software was evaluated again by performing regression testing (Assawamekin, 2015, p. 290) in order to determine if the previous change resulted in proper results. This procedure was performed until no system fault was detected. Subsequently, a user manual document was made to clarify how to use the system for different sorts of users (4 categories as shown in Figure 1) so that each category was able to comprehend and operate the developed system. This manual was generated in two formats: 1) document and 2) video explaining how to use the system.

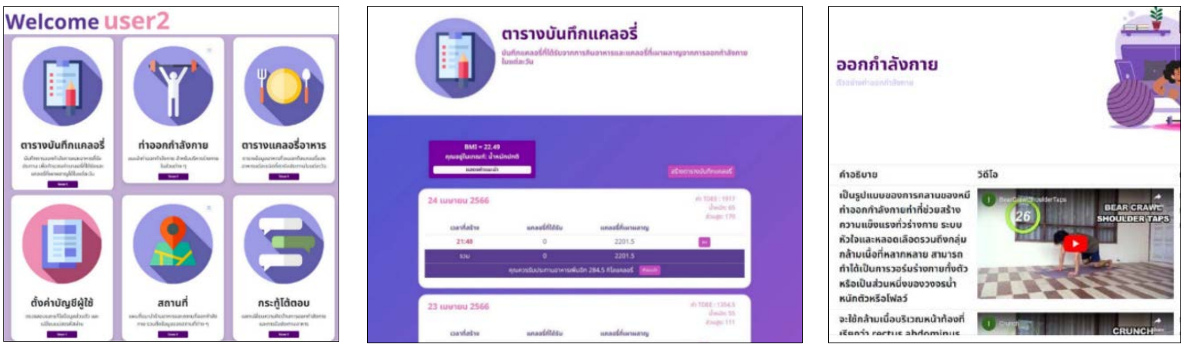
Phase 3: Follow-up testing and evaluation

In this phase, the system was integrated into practical usage. The research team used online training to train the first sample group, whereas classroom learning method was applied to train the second group (Phiwhlueng, 2024), with the research team acting as lecturers. The 2 training sessions included a workshop where participants applied their knowledge of the system and were allowed to ask the research team questions as they arose. Furthermore, the research team then facilitated the sample group to utilize the system for one month. Users had an access to the online platform, and contact information of the development team for further information, or spoke directly with the research team. When the system usage testing period ended, the following step was assessment. In this step, the research team collected data from the system users by asking them to assess the system after first providing clarification on the questionnaire. The performance evaluation was partitioned by 3 information specialists and assessed according to the framework for success in the development of improved systems based on the model of DeLone and McLean (DeLone & McLean, 2003). Furthermore, the acceptability of the developed system was examined by 2 groups of 33 participants. Group 1 evaluated satisfaction with administrator and expert functions, whereas group 2 evaluated satisfaction with general user, member, and LINE Chatbot features. The questionnaire was divided into 3 parts: Part 1: general information of the respondents; Part 2: Likert scale questions (Almohtadi & Aldarabah, 2021), where opinions were divided into 5 levels (Rating Scale), including 1 = very dissatisfied, 2 = dissatisfied, 3 = neither dissatisfied nor satisfied, 4 = satisfied and 5 = very satisfied and; Part 3: Suggestions. Furthermore, quantitative data (Wadeejaroen et al., 2017, p. 307) from Part 1 and 2 of the questionnaires were collected to assess and interpret the results by implementing descriptive statistics (Wadeejaroen et al., 2017, p. 307), including mean and standard deviation (SD) (Wadeejaroen et al., 2017, p. 315). The following was interpretation of the mean into satisfaction level: 4.50-5.00 the highest level, 3.50-4.49 high level, 2.50-3.49 moderate level, 1.50-2.49 low level, 1.00-1.49 the lowest level.

4. Results

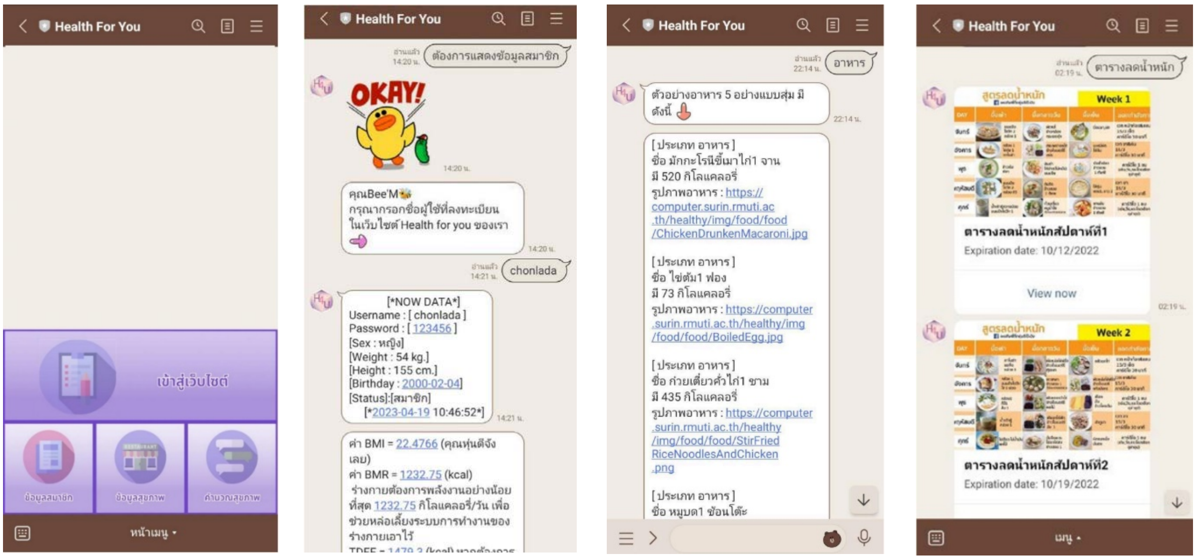
The findings of this study were split into 3 categories according to the objectives: 1) the outcome of design and development of health management system; 2) the outcome of assessment of technology acceptability and system efficiency; and 3) the outcome of sample groups' weight comparisons before and after utilizing the health management system. Details of the results were as follows:

1) The outcome of design and development of health management system was separated into 2 parts. Part 1 was the website for managing health information, where users were divided into 4 different groups: level 1 was administrator, level 2 was general user, level 3 was member, and level 4 was expert. The system provided daily activities and diet recommendations based on user weight, height, and suitable calorie content. In accordance with the user's demand, it also suggested workout routines for managing particular body areas. In order to compare weight fluctuations over the course of a month, users were required to input their weight every day in order to be recommended with restaurants with fitness centers and provide routes to those spots by the system. From the usage of a client/server architecture, the website was constructed and accessible via a web browser. The results of research are presented in Figure 4. Part 2 involved the LINE application, which linked data from the same database as the website, hosted LINE Chatbot for managing health information. The Chatbot made suggestion on a diet which adapted to the user's daily activities. The Chatbot as well determined calorie content based on the body's energy metabolism. It also provided information on exercise videos and the locations of eateries and fitness centers, depending on user preferences. Membership applications for accessing information were processed through the website. The outcomes of the LINE Chatbot's development are demonstrated in Figure 5.



a) Member homepage | b) Calorie log page | c) Exercise video page

Figure 4. Development results of a health information management website.



a) Chatbot homepage b) User information c) Food recommendations d) Weight loss schedule

Figure 5. Development results of LINE Chatbot.

2) The performance evaluation by 3 experts. The results revealed that the experts gave their opinion to the system efficiency in term of information quality at the highest level ($\bar{x}$ = 4.67, S.D. = 0.00). The experts as well rated the overall system performance at high level ($\bar{x}$ = 4.33, S.D. = 0.71) as shown in Table 1. Moreover, the evaluation of technology acceptance by 33 individual samples which were divided into 2 groups, resulted that the evaluation of the ease of use in group 1 (3 participants) rated an issue 3: ease of learning at the highest level ($\bar{x}$ = 5.00, S.D. = 0.00). The evaluation result of recognition of benefits found that the participants rated an issue 2: Ability to provide valuable information for decision-making at the highest level ($\bar{x}$ = 4.67, S.D. = 0.71). Meanwhile, the evaluation of the ease of use by group 2 (30 participants), indicated that group 2 rated an issue 3: ease of learning at high level ($\bar{x}$ = 4.27, S.D. = 0.71). As well as, the evaluation result of recognition of benefits found that the participants rated an issue 2: Ability to continuously track and record the weight loss results of users at high level ($\bar{x}$ = 4.33, S.D. = 0.71) as shown in Table 3.

Table 1. The results of the health management system efficiency evaluation

Topics	$\bar{X}$	S.D.	Results
1. system quality	3.67	0.00	high level
2. information quality	4.67	0.00	highest level
3. service quality	4.00	0.00	high level
4. perceived usefulness	4.33	0.71	high level
5. Overall	4.33	0.71	high level

Table 2. The result of the evaluation of technology acceptance by the sample group 1 with 3 participants.

Topics	$\bar{X}$	S.D.	Results
Field 1: Ease of use			
1. user friendly	4.33	0.00	high level
2. system flexibility	3.67	0.71	high level
3. ease of learning	5.00	0.00	highest level
Average	4.33	0.24	high level
Field 2: Recognition of benefits			
1. Ability to manage data and create clear and concise report	4.33	0.00	high level
2. Ability to provide valuable information for decision-making	4.67	0.71	highest level
3. Ability to provide effective support and services.	4.33	0.71	high level
Average	4.44	0.47	high level

Table 3. The result of the evaluation of technology acceptance by the sample group 2 with 30 participants.

Topics	$\bar{X}$	S.D.	Results
Field 1: Ease of use			
1. user friendly	4.03	1.41	high level
2. system flexibility	4.13	0.00	high level
3. ease of learning	4.27	0.71	high level
Average	4.14	0.71	high level
Field 2: Recognition of benefits			
1. Recommendation of appropriate weight loss plans based on user information (diet and exercise)	4.03	0.71	high level
2. Able to continuously track and record the weight loss results of users	4.33	0.71	high level
3. Promote and support user in receiving treatment and improving health along with providing recommendations	4.23	0.00	high level
Average	4.20	0.47	high level

3) The average body mass index comparison findings from sample group 2's health management system before and after was found that the average body mass index increased by 20 cases and decreased by 10 cases which are demonstrated in Figure 6.

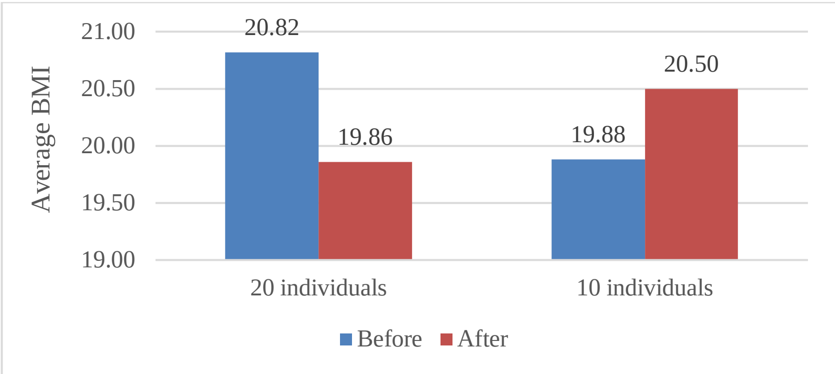


Figure 6. Compare BMI before and after system usage.

5. Discussion

5.1 Principles of software engineering play a crucial role in the development of health management systems. The LINE Chatbot and website are 2 technologies that can be utilized. As a result, users are likely to access the system quickly and simply without the necessity of installing it on their computer. The system, which connects and shares information with others, is continuously used. This is consistent with the research of Naemi et al. (2024), who has created a website to document medical data for individuals with lazy eye illness. According to the operational result, the convenience of users' assessment scores has been at a good level. The website also provides features for storing and retrieving patient information to help in managing additional challenges during treatment. Ability of the developed LINE Chatbot to respond to queries and offer health-related information to users' around-the-clock stimulates their interest and improves communication effectiveness. This is consistent with the research findings of Ulrich et al. (2024), whereby they have created a Chatbot to address and resolve patients' problems on several fronts. The Chatbot has received favorable feedback from users, according to the development findings. It could have lessened feelings of depression and anxiety. The developed Chatbot is automatically accessible and flexible. Additionally, this draws the conclusion that Chatbot may be a useful and successful instrument for altering behavior.

5.2 The result of efficiency evaluation by 3 experts showed an overall opinion at a high level ($\bar{X}$= 4.33, S.D.=0.71). The average opinions on information quality, perceived usefulness, service quality, and system quality were ranked from highest to lowest, respectively. Furthermore, the evaluation results of technology acceptance from 2 sample groups found that Group 1 (3 individuals who assessed the system administrator and expert functions) rated the advantages of technology at a high level ($\bar{X}$= 4.44, S.D.=0.47). The ease of use was also rated at a high level ($\bar{X}$= 4.33, S.D.=0.24). In addition, Group 2 (30 individuals who assessed the member, LINE Chatbot, and general user functions) rated the advantages of technology at a high level ($\bar{X}$= 4.20, S.D.=0.47). the ease of use was rated at a high level ($\bar{X}$= 4.14, S.D.=0.71). It was evident from the assessment findings of 2 sample groups that the perceived advantages element showed a high level of result. This was a result of the system's development methodology based on the idea of a “Wellness Plan”, which was a personalized healthcare plan based on each individual's requirements and objectives. Basic user data, including age, gender, height, weight, and activities, were linked to and analyzed in a way that contributed to each individual's own diet and exercise plan. This enabled the users to receive guidance and information that was relevant and valuable for promoting health care on a personal basis. This facilitated decision-making as well, allowing data to be utilized to alter an individual's behavior. This is in line with the research of Taptaemtut (2023). In the research, a lifestyle plan based on health literacy has been recommended for individuals of working age with metabolic syndrome, characterized by a malfunctioning metabolic system resulting in abdominal obesity. The operations result has shown that there has been a significant difference ($P<0.05$) in the sample group's health literacy scores before and after they participated in the program. This has suggested that applying medical recommendations to the people of working age could have enhanced clinical results as well as provide useful health information.

5.3 When comparing the average Body Mass Index (BMI) before and after using the health management system, it was found that among the sample group, 20 participants who used the system more than 80% of the time showed a decrease in average BMI. Conversely, 10 participants who used the system less than 80% of the time showed an increase in BMI. These results were arisen from the consistency in system usage that the users were facilitated by the ease of usage. The users inputted health information via the LINE Chatbot, which caused the data collection to be operated easily and obtained instant answers. As a result, there was motivation to consistently adopt the system's advice. This is consistent with the study conducted by Chenglai, Prampate, and Ponsung (2021), who have created a weight loss program that has enabled participants to share information with one another. As a result, the sample group has been motivated that has led to consistency, which subsequently resulted in more successful weight loss.

6. Conclusions

The result of this research indicated that LINE Chatbot and websites were capable of collecting basic user data (sex, age, height, weight, and daily activities) and then analyzing it to provide personalized exercise and diet regimens based on each person unique body requirements. The developed system had performance evaluation result and technological acceptability at high level. And the results of the comparison of the BMI of the sample group found that there were 20 cases with a reduced BMI, which showed that the developed system help increase efficiency in health management. This research was helpful in the social context of organizations which were seeking to support staff members' fundamental health since the system was simple to use, offers information that was actionable, and was easy to follow. Executives have access to summary data on the general health-related behavior of personnel of agencies, which were utilized for subsequent policy development. The research suggestions were that the data recording into the system was still recorded by the users. In some cases, there were errors or data were not recorded continuously, affecting the diet and exercise plan. Therefore, future research should add methods to record data into the system automatically, add a system to notify users continuously, which may result in better efficiency of the system.

Conflict of Interest

The authors declared that there are no conflicts of interest.

References

Almohtadi, R. M., & Aldarabah, I. T. (2021). University students' attitudes toward the formal integration of Facebook in their education: Investigation guided by Rogers' attributes of innovation. *World Journal of Education*, *11*(1), 20-28.

Apache Friends. (2024). *What is XAMPP*. Retrieved from www.apachefriends.org/index.html

Assawamekin, N. (2015). *Fundamentals of software engineering*. Bangkok: SE-Education.

Canonical Ltd. (2024). *Ubuntu Server documentation*. Retrieved from https://ubuntu.com/server/docs

Chenglai, N., Prampate, C., & Ponsung, S. (2021). Development of health literacy model of overweight students using the artificial intelligence activities and social network in Trang Province. *Journal of Sakon Nakhon Hospital*, *24*(1), 11-22.

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean Model of Information Systems Success: A ten-year update. *Journal of Management Information Systems*, *19*(4), 9-30.

Department of Disease Control. (2021a). *Guidebook for creating health knowledge in disease prevention and health hazard control*. Nonthaburi: Department of Disease Control, Office of Risk Communication and Health Behavior Development.

Department of Disease Control. (2021b). *Know your numbers & know your risks*. Nonthaburi: Ministry of Public Health [in Thai].

Department of Health. (2022). *Wellness plan*. Nonthaburi: Bureau of Elderly Health, Department of Health, Ministry of Public Health [in Thai].

Faculty of Dentistry. (2019). *Self-care manual for good health and happiness*. Bangkok: Chulalongkorn University Press [in Thai].

Google. (2024). *Dialogflow*. Retrieved from https://cloud.google.com/dialogflow/docs

Khamngoen, R., Phetrot, T., Jampairoh, N., Liamthong, P., Towangjorn, R., Paiklai, V., & Nuakalasin, S. (2023). The effectiveness of the health behavior change program based on PBRI's model and 7 colors ball tool on knowledge and health behavior of nursing students with overweight. *Journal of Health Research and Innovation*, *6*(2), 1-15.

Kimsungnoen, N., Setthawong, K., Boonya, P., Yantaporn, C., & Dechjob, N. (2023). Health promoting behaviors of students at Rangsit University. *APHEIT Journal of Nursing and Health*, *5*(2), 1-15.

Lalam, P., Chaimai, A., & Fukfon, K. (2022). Metabolic syndrome in older persons: Impacts and nursing roles. *Journal of Nursing and Health Research*, *23*(2), 13-27.

Lamsal, K. (2020). *Designing and developing a dynamic website using PHP*. Bachelor Degrees in Information Technology, Oulu University of Applied Sciences.

Naemi, R., Akbarian, M., Ebrahimi, M., Shahmoradi, L., Masoomian, B., & Rezayi, S. (2024). Design and evaluation of a web-based electronic health record for amblyopia. *Frontiers in Medicine*, *11*. doi:10.3389/fmed.2024.1322821

Nelson, R. (2020). *Health informatics: An interprofessional approach*. Elsevier Health Sciences.

Nyabuto, G. M., Mony, V., & Mbugua, S. (2023). Architectural review of client-server models. *International Journal of Scientific Research & Engineering Trends*, *10*(1), 139-143.

Phiwhlueng, A. (2024). The development of training course curriculum on classroom action research for strengthening the potential of teaching Thai language of teaching professional students. *Journal of Mani Chettha Ram Wat Chommani*, *7*(1), 80-95.

Phokhwang, C., Sarakshetrin, A., Ngamwongwiwat, B., & Pramnoi, T. (2023). Attitude toward body image, nutrition and exercise knowledge for weight control, and weight control behavior among nursing students from two educational institutions: A pilot study. *DRIRDI Research for Community Service Journals*, *9*(2), 58-73.

RMUTI Surin. (2022). *Bachelor of Science Program in Computer Science*. Retrieved from https://computer.surin.rmuti.ac.th

Ruanphet, K., Khamanek, S., Ngomsangad, Y., Hiranwitchayakun, A., & Suwannawong, S. (2023). Factors affecting happiness among students attending Western University, Buriram Campus. *Regional Health Promotion Center 9 Journal*, *17*(3), 797-810.

Taptaemtut, M. (2023). A health literacy lifestyle program for working age people with metabolic syndrome, Lamae District, Chumphon Province. *The Journal of Boromarjonani College of Nursing Suphanburi*, *6*(1), 145-159.

Ulrich, S., Gantenbein, A. R., Zuber, V., Von Wyl, A., Kowatsch, T., & Künzli, H. (2024). Development and evaluation of a smartphone-based Chatbot coach to facilitate a balanced lifestyle in individuals with headaches (BalanceUP app): Randomized controlled trial. *Journal of Medical Internet Research*, *26*.

Wadeejaroen, W., Lertnaisat, R., & Teekasap, S. (2017). *Research methods: From theoretical concepts into practice*. Bangkok: SE-Education [in Thai].

Available online at https://li02.tci-thaijo.org/index.php/ssstj

Suan Sunandha Science and Technology Journal

The Efficacy of Clustering Algorithms for Young '*Nam-Hom*' Coconut Gene Expression Data in Unveiling the Specific Genes Determining the Flavor: A Comparative Analysis of K-means and Fuzzy C-means

Kairung Hengpraprohm[1], Supoj Hengpraprohm[1*], Kriengkrai Meethaworn[2]

[1]Programm in Data Science,
[2]Programm in Agricultural innovation and management,
Faculty of Science and Technology, Nakhon Pathom Rajabhat University. 85 Malaiman Rd., Mueang, Nakhon Pathom, 73000, Thailand.
***Corresponding author e-mail: supojn@yahoo.com**

Received: *31 March 2024* / **Revised:** *20 May 2024* / **Accepted:** *12 July 2024*

Abstract
This study explores the application of K-means and Fuzzy C-means clustering techniques to analyze gene expression data related to the flavor of young '*Nam-Hom*' coconuts. By comparing these clustering methods, the research aims to identify gene clusters that significantly influence the aromatic and off-flavor profiles of young '*Nam-Hom*' coconuts stored at different temperatures (4°C and 25°C). Specifically, our findings highlight clusters involved in lipid metabolism and cold stress response which are crucial for developing desirable and undesirable flavors, such as LOX1 and ADH2 genes. The study advances our understanding of coconut genetics demonstrates the utility of clustering techniques in agricultural genomics, offering valuable pathways for future genetic enhancement and storage optimization strategies aimed at improving coconut aroma.

Keywords: Young '*Nam-Hom*' coconut, Gene expression data, Clustering algorithms

1. Introduction

Young '*Nam-Hom*' coconut (*Cocos nucifera* L.) is a crop of paramount economic importance and a symbol of agricultural heritage in tropical countries including, Thailand. It is a globally popular fruit because of its refreshing, high nutritional value, and tender-soft jelly-like endosperm (kernel) (Yong, Ge, Ng, & Tan, 2009). Moreover, it plays a significant role in the international market, and Thailand is the leading exporter of young '*Nam-Hom*' coconut. The main markets are the United States and China. In 2023, the exported value was more than 640,810 tons, 430 million USD (Office of Agricultural Economics, 2024). Despite the global challenges posed by the COVID-19 pandemic in 2020, the export of young '*Nam-Hom*' coconuts from Thailand experienced a notable increase of 30%. The demand for coconuts, both in fresh forms and for processing into various products, remains high. Specifically, a significant domestic and international demand for '*Nam-Hom*' coconut varieties contributes to an export value exceeding one billion baht annually, with an upward trend observed.

However, the thriving export market for Thai young '*Nam-Hom*' coconuts encounters several production and preservation challenges, including cracking of fruits (Meethaworn, 2021; Pakcharoen, Meethaworn, & Mohpraman, 2012), browning reaction after peeled (Mohpraman & Siriphanich, 2012), aromaloss (Siriphanich et al., 2011), and the development of off-flavor during long-term storage at low temperatures (Meethaworn, Imsabai, Zhang, Chen, & Siriphanich, 2022; Meethaworn, Luckanatinwong, Zhang, Chen, & Siriphanich, 2019). These issues, particularly off-flavor-flavor during cold storage, necessary for the minimum four-week shipping period to international destinations, represent a critical challenge. Coconuts are stored at temperatures between 2 and 4 degrees Celsius to maintain freshness during transport. However, this can lead to a phenomenon known as chilling injury, resulting in the emergence of rancid oil-like odors in the coconut kernel and coconut water of the young '*Nam-Hom*' coconuts.

Previous studies have indicated that only a subset of genes expressed at different temperatures are implicated in the developed off-flavor-flavors, suggesting a complex genetic basis for this trait. Therefore, this research aims to explore the gene expressions (traits) influencing the flavor in young '*Nam-Hom*' coconuts, employing machine learning techniques such as data clustering with K-means and Fuzzy C-means algorithms. By identifying informative features (genes) through clustering similar gene expression data, this study intends to reveal the groups of genes associated with the development of flavor in young '*Nam-Hom*' coconuts. The goal is to use this information to improve or limit the occurrence of off-flavor-flavor during storage, thereby enhancing the quality of coconuts for export.

2. Backgrounds

2.1 K-means clustering

K-means clustering (Sinaga and Yang, 2020) is an essential unsupervised learning technique utilized across various disciplines, including bioinformatics, for effectively organizing large datasets into meaningful clusters. This algorithm partitions a dataset into k distinct clusters based on the similarity of data points, making it particularly useful for identifying patterns and groupings in complex data sets such as gene expression profiles.

Workflow Summary:

(1) *Initialization*: Begin by selecting k initial centroids randomly or through a more deliberate method, where k is the desired number of clusters.

(2) *Assignment*: Assign each data point in the dataset to the nearest centroid, typically determined by the shortest Euclidean distance, thereby forming preliminary clusters.

(3) *Update*: Recalculate the position of each centroid to be the mean of the data points assigned to its cluster, effectively moving it to the cluster center.

(4) *Iteration*: Repeat the assignment and update steps iteratively until the centroids stabilize (no longer move significantly), signaling the algorithm has converged on a solution.

K-means aims to minimize within-cluster variance, the sum of squared distances between data points and their respective cluster centroid, achieving compact, homogenous clusters.

2.2 Fuzzy C-Means clustering

Fuzzy C-Means (FCM) clustering (Bezdek et al., 1984) extends beyond the traditional hardly partitioning methods, like K-means, by allowing data points to belong to multiple clusters with varying degrees of membership. This soft clustering technique is valuable, particularly in fields: such as bioinformatics, where the biological phenomena under study often exhibit inherent ambiguities and overlapping characteristics.

Workflow Summary:

(1) *Initialization*: Select the number of clusters, k, and initialize the cluster centers randomly or based on a heuristic. Unlike K-means, FCM starts with an assumption about the cluster centers that will be refined.

(2) *Membership Assignment*: Assign each data point a degree of membership for each cluster. This degree ranges from 0 (no membership) to 1 (full membership), based on the distance to the cluster centers, allowing data points to be partially in more than one cluster.

(3) *Centroid Update*: Update the cluster centers to reflect the calculated memberships, with each center being the weighted mean of all points, where weights are the membership degrees. This step accounts for the fuzziness of cluster boundaries by considering the degree to which points belong to clusters.

(4) *Iteration*: Iterate the membership assignment and centroid update steps until the cluster centers stabilize within a predetermined tolerance or until a maximum number of iterations is reached. Convergence is achieved when changes in the degrees of membership between two consecutive iterations are negligible.

The primary goal of FCM is to minimize an objective function that represents the distance between data points and cluster centers, weighed by the membership degrees. This minimization leads to the formation of clusters that best capture the underlying structure in the data, considering the fuzziness of natural groupings.

2.3 The Silhouette Score

The Silhouette Score (Shahapure & Nicholas, 2020) is a metric used to assess the quality of clusters created by a clustering algorithm. Introduced by Peter J. Rousseeuw in 1987, this measure evaluates how similar an object is to its cluster compared to the other clusters. The value of the Silhouette Score ranges from -1 to 1, where a high score indicates that the object is well-matched to its cluster and poorly matched to neighboring clusters, thus providing a clear indication of the appropriateness of the cluster assignments.

Calculation of the Silhouette Score: The Silhouette Score for each sample in the dataset is calculated using the formula $s=(b-a)/max(a,b)$, where:

a is the average distance from the sample to all other points in the same cluster.

b is the smallest average distance from the sample to points in a different cluster, minimized over clusters.

Interpretation:

A score approaching 1, signifies that the sample is far from the neighboring clusters.

A score approaching 0, indicates that the sample is on or very close to the decision boundary between two neighboring clusters.

A score approaching -1, means the sample has been assigned to the wrong cluster.

In the context of gene expression data analysis, the Silhouette Score facilitates the evaluation of clustering algorithms like K-means and Fuzzy C-means, offering a quantitative measure to compare the performance of different models. By utilizing this metric, researchers can more confidently discern the underlying patterns in gene expression, leading to more informed conclusions about genetic influences on traits of interest, such as the flavor of coconuts.

2.4 Related research

Alagukumar, S. and Lawrance, R. conducted a comparative study on the efficiency of two association rule mining techniques, FP-Growth and Apriori, applied to different microarray datasets for analyzing gene expression levels. Their results showcased FP-Growth's superior performance in terms of computational speed, underscoring the potential of sophisticated data mining techniques in unraveling complex gene expression patterns (Alagukumar & Lawrance, 2015).

Saensuk et al. studied a unique Thai dwarf green coconut variety known for its distinct "pandan-like" aroma, attributed to the compound 2-acetyl-1-pyrroline (2AP). Their transcriptomic analysis sought to identify the genes responsible for 2AP biosynthesis, providing insights into the genetic basis of flavor profiles in coconuts (Saensuk et al., 2016).

Zhu et al. presented a new approach to data clustering named subspace clustering guided unsupervised feature selection (SCUFS). This technique focuses on selecting representative data features that maintain the integrity of data subgroups, demonstrating that SCUFS can outperform traditional unsupervised feature selection methods in clustering efficacy (Zhu, Zhu, Hu, Zhang, & Zuo., 2017).

Meethaworn et al. explored the effects of low-temperature storage on the developed off-flavor in young coconuts, attributing the phenomenon to the lipoxygenase (LOX) pathway. Their research, which varied storage conditions, indicated that lower temperatures exacerbate the production of undesirable odors, pointing to specific genes involved in lipid oxidation (Meethaworn et al., 2019).

Hengpraprohm et al. investigated the identification of molecular markers indicating the developmental stages of ovarian maturation in black tiger shrimp using microarray data. By applying data mining techniques, they have shown that they can find the genetic information crucial for reproductive system functions, highlighting the application of genomic analyses in improving aquaculture breeding practices (Hengpraprohm, Jungjit, Hengpraprohm, & Thammasiri, 2019).

These highlighted studies demonstrate a broad spectrum of approaches and technologies applied in genomics and bioinformatics, ranging from association rule mining to advanced clustering and feature selection techniques. The collective insights from these research efforts deepen our understanding of genetic mechanisms behind important traits offer practical avenues for agricultural innovation and food production enhancement.

3. Materials and Methods

This study employed K-means and Fuzzy C-means clustering techniques to investigate the gene expression groups related to the flavor of young '*Nam-Hom*' coconuts. The methodology is outlined in the following steps:

3.1 Data collection

This step involved collecting and selecting data to ensure its suitability for model development. The research utilized secondary data representing gene expression levels associated with scent development in young coconuts. The dataset used in the experiment was from Meethaworn et al. (2019). This data was collected from coconuts stored at two different temperatures, 4°C and 25°C, for 15 days. The gene expression data, in the form of RNA sequencing technique, was converted into numerical values indicating the expression levels of genes, which were then organized into a Microsoft Excel format as illustrated in Figure 1.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Gene_ID	Length	Annotation	Name	swissprot_description	GO_Term	Con0d_A	Con0d_B	Con0d_C	RT3d_A	RT3d_B	RT3d_C	C3d_A	C3d_B	C3d_C	RT6d_A	RT6d_B	RT6d_(
2	TRINITY_DN40438_	1474	PREDICTED: replication fact	RFC2	Replication factor C sul	GO:0003689(DN	39.41	42.96	44.37	54.11	43.05	39.65	38.94	37.06	33.05	42.08	42.74	37.89
3	TRINITY_DN39538_	3275	PREDICTED: ubiquitin-conju	UBC5A	Ubiquitin-conjugating e	GO:0004842(ubi	65.85	62.55	51.68	73.61	103.78	110.93	53.36	53.79	69.76	90.94	87.44	81.99
4	TRINITY_DN44538_	505	-	-	-	-	15.20	19.08	25.52	15.04	22.05	21.01	16.11	9.20	9.80	18.18	21.72	12.41
5	TRINITY_DN46286_	680	hypothetical protein B456_00	-	-	-	12.79	15.94	17.74	16.78	11.20	14.03	31.55	40.05	17.93	10.85	14.87	25.50
6	TRINITY_DN45959_	2434	PREDICTED: LOW QUALI	RS31	Serine/arginine-rich spli	GO:0000166(nuc	40.94	37.60	33.30	102.87	100.34	106.14	50.88	53.49	47.28	79.59	75.35	56.13
7	TRINITY_DN36412_	2209	PREDICTED: E3 ubiquitin-pr	AIP2	E3 ubiquitin-protein liga	GO:0000209(pro	11.73	11.69	11.56	18.31	13.43	14.53	10.53	9.52	12.68	9.85	9.87	9.78
8	TRINITY_DN42104_	702	Actin 7 isoform 1 [Theobroma	ACT7	Actin-7 OS=Arabidops	GO:0005200(stru	409.89	404.76	416.09	682.65	453.72	470.62	409.35	396.94	453.68	655.62	772.94	627.85
9	TRINITY_DN46053_	3950	PREDICTED: sister-chromati	SCC3	Sister-chromatid cohesi	GO:0005515(pro	5.68	6.21	6.14	8.69	8.15	8.78	4.88	4.37	7.43	15.13	15.12	12.84
10	TRINITY_DN45484_	214	-	-	-	-	0	0	1.46	3.85	0	0	5.97	9.05	4.35	0	0	2.73
11	TRINITY_DN41248_	2876	PREDICTED: LOW QUALI	CWF19L2	CWF19-like protein 2	GO:0003824(cat	4.78	4.31	4.87	3.61	8.55	8.86	7.71	7.98	8.56	8.64	8.19	8
12	TRINITY_DN3178_	236	-	-	-	-	11.44	0	0	0	6.55	0	0	0	0	0	0	0
13	TRINITY_DN44608_	279	-	-	-	-	4.02	1.82	5.14	3.44	1.48	3.05	0	2.06	3.44	9.05	10.92	3.72
14	TRINITY_DN42271_	1790	PREDICTED: pectinesterase-	PME2.1	Pectinesterase 2.1 OS=	GO:0030599(pec	22.60	21.39	22.90	0.47	1.52	2.70	2.84	1.28	2.32	1.98	1.44	0.63
15	TRINITY_DN50003_	238	-	MT2345	Uncharacterized Na(+)	-	0	0	0	0	0.80	0	11.21	7.34	0	0	3.29	0
16	TRINITY_DN46260_	312	PREDICTED: uncharacterize	-	-	-	31.03	31.44	28.77	33.55	29.66	30.03	17.39	22.48	25.49	64.97	65.72	67.90
17	TRINITY_DN45669_	204	-	-	-	-	0	1.58	1.70	0	0	0	0	0	1.26	0	0	0
18	TRINITY_DN46256_	222	-	-	-	-	0	0	0	0	1.97	0	0	0	0.97	2.76	1.35	3.67
19	TRINITY_DN41479_	2193	PREDICTED: BTB/POZ dom	At3g05675	BTB/POZ domain-cont	GO:0005634(nuc	46.21	44.18	37.89	44.40	43.27	40.29	30.21	17.92	30.95	65.29	81.12	72.67
20	TRINITY_DN45789_	355	-	-	-	-	2.86	0.90	0	0	3.10	4.29	5.90	7.22	14.30	18.66	12.50	7.60
21	TRINITY_DN41011_	605	PREDICTED: long chain acyl	LACS4	Long chain acyl-CoA s	GO:0004467(lon	57.37	56.18	56.24	43.94	37.21	40.25	60.29	48.78	54.94	23.35	32.29	27.11
22	TRINITY_DN43916_	2344	PREDICTED: mitogen-activat	MPK12	Mitogen-activated prot	GO:0004672(pro	29.71	35.17	34.05	22.85	31.87	30.72	23.25	15.65	23.38	31.86	27.08	22.93
23	TRINITY_DN43789_	2596	PREDICTED: THO complex	THO1	THO complex subunit 1	GO:0000347(TH	7.04	8.15	7.85	14.86	15.89	15.74	8.49	9.47	8.76	20.55	20.48	21.22
24	TRINITY_DN44784_	2832	PREDICTED: LOW QUALI	-	-	GO:0008150(bio	19.79	19.18	20.30	17.78	15.16	13.30	12.32	10.76	10.97	19.59	18.44	15.96

Figure 1. Table of gene expression data of young '*Nam-Hom*' coconut stored at 25°C and 4°C for 15 days using RNA sequencing technique.

3.2 Data preparation

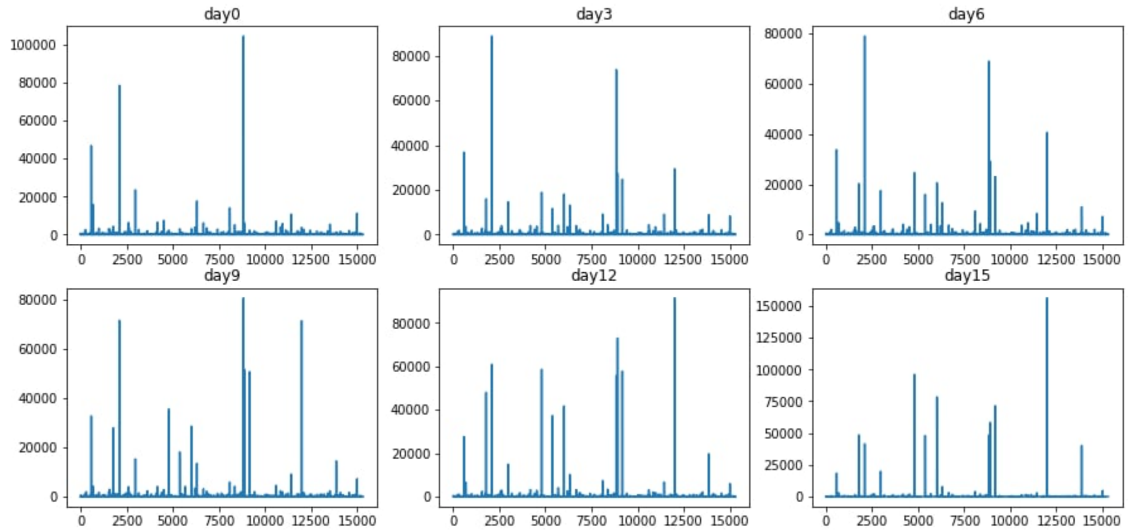
In this phase, the collected data underwent preparation to make it suitable for analysis using Python. It involved formatting the data into a .csv file format, as demonstrated in Figure 2. The structured data consisted of rows and columns where each row represented one of the total 15,000 genes, and the columns included 12 specific data points, detailed as follows:

Gene: The gene identifier.
con0d: Gene expression data of young '*Nam-Hom*' at the beginning of storage.
RT3d: Gene expression data of young '*Nam-Hom*' stored at 25°C for 3 days.
C3d: Gene expression data of young '*Nam-Hom*' stored at 4°C for 3 days.
RT6d: Gene expression data of young '*Nam-Hom*' stored at 25°C for 6 days.
C6d: Gene expression data of young '*Nam-Hom*' stored at 4°C for 6 days.
RT9d: Gene expression data of young '*Nam-Hom*' stored at 25°C for 9 days.
C9d: Gene expression data of young '*Nam-Hom*' stored at 4°C for 9 days.
RT12d: Gene expression data of young '*Nam-Hom*' stored at 25°C for 12 days.
C12d: Gene expression data of young '*Nam-Hom*' stored at 4°C for 12 days.
RT15d: Gene expression data of young '*Nam-Hom*' stored at 25°C for 15 days.
C15d: Gene expression data of young '*Nam-Hom*' stored at 4°C for 15 days.

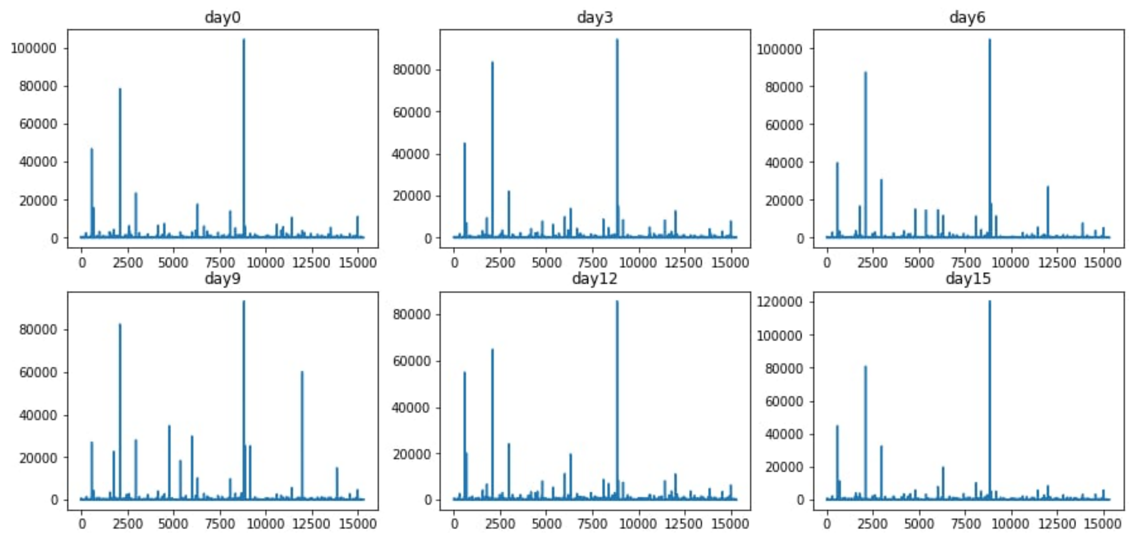
Gene	con0d	RT3d	C3d	RT6d	C6d	RT9d	C9d	RT12d	C12d	RT15d	C315d
1	42.247	45.603	36.35	40.903	30.173	28.393	30.51	27.973	21.313	30.6	10.803
2	60.027	96.107	58.97	86.79	83.197	51.13	21.857	107.153	23.637	98.267	13.453
3	19.933	19.367	11.703	17.437	14.397	13.477	7.687	35.233	11.497	34.91	5.187
4	15.49	14.003	29.843	17.073	17.03	22.173	15.74	15.147	16.633	13.63	40.24
5	37.28	103.117	50.55	70.357	49.21	44.857	34.137	75.08	26.117	60.88	15.397
6	11.66	15.423	10.91	9.833	12.143	8.923	7.247	10.747	6.123	8.97	2.547
7	410.247	535.663	419.99	685.47	523.627	602.957	412.917	525.333	341.9	581.967	299.923
8	6.01	8.54	5.56	14.363	7.407	12.96	5.953	15.727	4.697	18.843	2.423
9	0.487	1.283	6.457	0.91	6.843	3.78	1.36	4.283	2.28	11.923	0.283
10	4.653	7.007	8.083	8.277	7.65	9.42	5.8	10.62	3.833	10.757	2.193
11	3.813	2.183	0	0	0	0	0	0	0	0	0
12	3.66	2.657	1.833	7.897	4.01	8.22	2.233	1.05	2.167	6.697	6.57
13	22.297	1.563	2.147	1.35	1.82	2.44	2.69	0.223	1.717	0.607	2.5
14	0	0.267	6.183	1.097	0.62	0.293	0.49	0	0	0	0

Figure 2. The prepared gene expression data for the experiment.

This organized data allowed for the comparison of gene expression at different temperatures and time point of storage, aiming to identify variations and patterns in gene expression related to the flavor of young '*Nam-Hom*' coconuts. Figure 3 illustrates the comparative gene expression data of young '*Nam-Hom*' coconuts stored at different temperatures over 15 days. Sub-figure 3(a) presents the comparison data of all 15,000 genes' expression profiles at 4 degrees Celsius across various days. Conversely, Subfigure 3(b) details the gene expression profiles at 25 degrees Celsius.



(a) Coconut gene expression with 4°C



(b) Coconut gene expression with 25°C

Figure 3. Comparative gene expression data of young '*Nam-Hom*' coconuts stored at 4°C and 25°C for 15 days.

By structuring the data in this manner, it became feasible to apply clustering techniques to analyze the gene expressions systematically. This preparation was crucial for the subsequent model construction phase, where the data's structure and organization played a significant role in the effectiveness of the clustering algorithms.

3.3 Model construction

The study aimed to identify significant gene characteristics influencing the scent of coconuts through data mining principles. The K-means and Fuzzy C-means clustering techniques were applied to group the data and compare the performance of these models. The similarity between data clusters and the consistency of model data were evaluated using the Silhouette Score, which ranges from -1 to 1. A score closer to 1 indicates well-suited data grouping, whereas a score near -1 suggests inappropriate clustering. The conceptual framework is depicted in Figure 4.

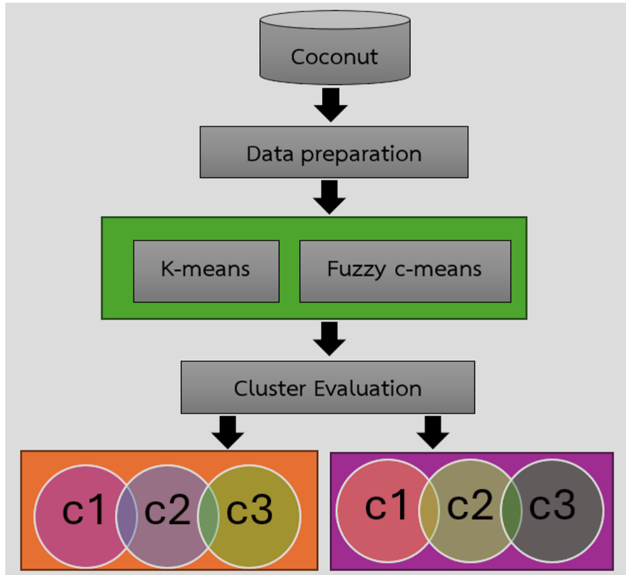


Figure 4. The conceptual framework for gene characteristic selection.

This research began with preparing the young '*Nam-Hom*' coconut data for clustering significant genes, divided into two sets based on storage temperatures at 25°C and 4°C. The K-means and Fuzzy C-means techniques were then used for data clustering. The efficiency of both methods was assessed to determine the best-performing model based on the Silhouette Score. The chosen model's data was then utilized to identify or describe genes impacting the developed flavor in young '*Nam-Hom*' coconuts during storage at both temperatures. The model development was conducted using Python programming, facilitating a thorough comparison and selection process for the genes associated with young '*Nam-Hom*' coconut flavor traits.

4. Results

The comparative analysis of gene clustering related to flavor development in young coconuts highlighted the efficacy of K-means and Fuzzy C-means clustering techniques. According to Table 1, the optimal number of clusters (K) was determined to be 3 based on the analysis depicted in Figure 5, which presents the distortion score elbow chart. The 'elbow' point in the chart is applied to determine the optimal number of clusters. At k=3, there is a noticeable elbow in the chart, indicating a significant reduction in the distortion score with a lesser rate of decrease beyond this point. This suggests that increasing the number of clusters beyond three leads to diminishing returns of clustering compactness and separation improvements. Therefore, k=3 is considered the optimal choice for achieving the most meaningful and distinct clustering of gene expression data without unnecessary complexity.

The K-means clustering showed superior performance achieving a Silhouette Score of 0.996, slightly higher than the Fuzzy C-means, which scored 0.995. This slight difference underscores the precision of both clustering methods in segregating gene expression data effectively.

Table 1. The Silhouette score comparison of K-Means and Fuzzy C-means clustering.

Method	K	Silhouette Score
K-means	3	0.996*
Fuzzy-C-means	3	0.995

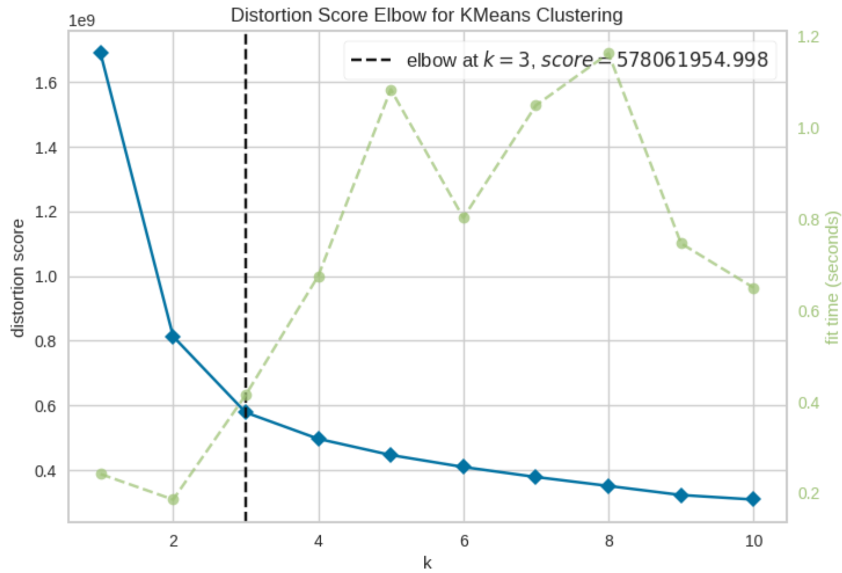


Figure 5. The distortion score elbow for K-Means clustering.

K-means Clustering Findings:

Cluster 1 primarily associated genes with lipid degradation processes and stress response signaling mechanisms during storage. Interestingly, this cluster did not encompass genes directly linked to the generated characteristic '*Nam-Hom*' coconut flavor or off-flavor.

Cluster 2 was identified to contain genes involved with both lipid degradation and transformation processes, indicative of energy utilization for cell nourishment, especially notable during seed germination. Gene expression in this cluster was predominantly observed at 25°C, suggesting that storage at this temperature facilitates energy-related metabolic activities more than at 4°C.

Cluster 3 stood out by including genes related to lipid degradation and the related genes to both '*Nam-Hom*' aroma and off-flavor in young '*Nam-Hom* coconuts. This cluster was the only one encompassing genes like Badh2, which controls the synthesis of 2-AP (Saensuk et al., 2016), a compound linked to the aroma of '*Nam-Hom*' coconuts. The expression of Badh2 was notably higher at 25°C, indicating that lower storage temperatures could diminish the coconut's aromatic quality. Moreover, genes associated with off-flavor development, such as LOX1 and ADH2, showed higher expression levels at 4°C, aligning with the observation that off-flavors are more developed at lower storage temperatures (Meethaworn et al., 2019; Meethaworn et al., 2022).

Furthermore, the exclusive presence of lipase genes in Cluster 3, involved in breaking down free fatty acids from triglycerides, signifies their role in the subsequent oxidation by the LOX pathway, leading to off-flavor generation. The expression of these genes was evident at both 4°C and 25°C, underscoring their potential involvement in scent modulation irrespective of the storage temperature.

The K-means clustering findings delineate three distinct clusters with unique gene expressions linked to the different lipid metabolism aspects and flavor profile modulation. Noted, Cluster 3 includes genes critical for the characteristic synthesis of '*Nam-Hom*' aromas and off-flavors, showing differential expression patterns across the two temperatures studied. Moreover, analysis of the results obtained from the Fuzzy C-means clustering corroborates these findings, displaying similar cluster characteristics and supporting the consistency of the gene grouping trends. This alignment between the two clustering methods underscores the robustness of our clustering approach in capturing the essential aspects of gene expression that influence coconut flavor and quality.

5. Discussion

The results illuminate the complex interplay of genes involved in lipid metabolism and their impact on the aroma compound profile of young '*Nam-Hom*' coconuts during storage. The superior clustering performance of K-means, as evidenced by the higher Silhouette Score, facilitated a clearer delineation of gene groups and their respective functions. The identified specific clusters associated with flavor-related genes provide a foundation for further exploration into genetic engineering and storage practices, aimed at enhancing the aromatic qualities of young '*Nam-Hom*' coconuts while mitigating off-flavors.

This study's findings contribute to the expanded understanding of post-harvest young '*Nam-Hom*' coconut flavor dynamics, emphasizing the importance of storage conditions on gene expression and flavor development. Future research could investigate the regulatory mechanisms governing these flavor-related genes, potentially uncovering new avenues for improving young '*Nam-Hom*' coconut fragrance through genetic manipulation or optimized storage strategies.

6. Conclusions

This research marks a significant step towards understanding the complexities of gene expression related to the flavor of young coconuts, with a focus on the application and comparative efficacy of K-means and Fuzzy C-means clustering techniques. By meticulously analyzing gene expression data under varying storage temperatures, the study unveils the intricacies of young *'Nam-Hom'* coconut aroma development but also showcases the strengths and nuances of two widely used clustering algorithms in bioinformatics.

Our analysis demonstrated that K-means clustering slightly outperformed Fuzzy C-means, as indicated by the higher Silhouette Score of 0.996. This comparative approach highlighted the subtle differences in clustering efficacy and underscored the suitable K-means for this specific application in genetic data analysis. The optimal clustering identified through K-means facilitated a nuanced understanding of the genetic factors at play, distinguishing between genes associated with various lipid metabolism processes and those directly impacting the aroma profile of young *'Nam-Hom'* coconuts.

The significance of this study extends beyond the specific insights into young *'Nam-Hom'*coconut gene expression. It underscores the pivotal role of advanced clustering techniques in deciphering complex biological data, offering a methodological framework for future genetic research. The findings reinforce the utility of clustering algorithms in bioinformatics, particularly in unraveling the genetic basis of phenotypic traits.

Furthermore, this research illuminates the broader applicability of clustering algorithms in agricultural genetics and post-harvest preservation strategies. By optimizing clustering approaches, researchers can gain deeper insights into the genetic mechanisms governing crop quality traits, enabling targeted interventions to enhance desirable characteristics and mitigate negative ones.

In conclusion, while the study provides valuable insights into the genetic determinants of coconut scent, its broader contribution lies in demonstrating the power and potential of clustering algorithms in genetic research. The comparative analysis between K-means and Fuzzy C-means offers a reference point for selecting appropriate clustering methods in similar genomic studies, paving the way for more refined and effective analyses of the quest to link genetic data with phenotypic traits.

Acknowledgements

This research was supported by Nakhon Pathom Rajabhat University (NPRU), Thailand.

References

Alagukumar, S., & Lawrance, R. (2015). A selective analysis of microarray data using association rule mining. *Procedia Computer Science*, *47*, 3-12. doi:10.1016/j.procs.2015.03.177

Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy *c*-means clustering algorithm. *Computers & Geosciences*, *10*(2-3), 191-203. doi:10.1016/0098-3004(84)90020-7

Hengpraprohm, S., Jungjit, S., Hengpraprohm, K, & Thammasiri, D. (2019). Molecular marker discovery for ovarian maturation level of the black tiger shrimp from microarray data using genetic algorithm. *International Journal of the Computer, the Internet and Management*, *27*(2), 43-51.

Meethaworn, K. (2021). Cracking characteristic on polished young coconut and its prevention. *Proceedings of the 13th NPRU National Academic Conference* (pp. 209-216). Nakhon Pathom, Thailand (in Thai).

Meethaworn, K., Imsabai, W., Zhang, B., Chen, K., & Siriphanich, J. (2022). Off-flavor and loss of aroma in young coconut fruit during cold storage are associated with the expression of genes derived from the LOX pathway and Badh2. *The Horticulture Journal*, *91*(2), 209-220. doi:10.2503/hortj.UTD-309

Meethaworn, K., Luckanatinwong, V., Zhang, B., Chen, K., & Siriphanich, J. (2019). Off-flavor caused by cold storage is related to induced activity of LOX and HPL in young coconut fruit. *LWT*, *114*, 108329. doi:10.1016/j.lwt.2019.108329

Mohpraman, K., & Siriphanich, J. (2012). Safe use of sodium metabisulfite in young coconuts. *Postharvest Biology and Technology*, *65*, 76-78.

Office of Agricultural Economics. (2024). *Amount and economic value of aromatic coconut*. Retrieved from http://www.infoservice@oae.go.th

Pakcharoen, A., Meethaworn, K., & Mohpraman, K. (2012). *The occurrence and deterrence of fruit cracking and off-flavor in aromatic coconut during storage at low temperature* (Report No. KU.R.1/2011). Postharvest Technology Innovation Center (in Thai).

Saensuk, C., Wanchana, S., Choowongkomon, K., Wongpornchai, S., Kraithong, T., Imsabai, W., … Arikit, S. (2016). De novo transcriptome assembly and identification of the gene conferring a "pandan-like" aroma in coconut (*Cocos nucifera* L.). *Plant Science*, *252*, 324-334. doi:10.1016/j.plantsci.2016.08.014

Shahapure, K. R., & Nicholas, C. (2020). Cluster quality analysis using Silhouette Score. *2020 IEEE 7th International Conference on Data Science and Advanced Analytics* (pp. 747-748). Sydney, Australia. doi:10.1109/DSAA49011.2020.00096

Sinaga, K. P., & Yang, M.-S. (2020). Unsupervised k-means clustering algorithm. *IEEE Access*, *8*, 80716-80727. doi:10.1109/ACCESS.2020.2988796

Siriphanich, J., Saradhuldhat, P., Romphophak, T., Krisanapook, K., Pathaveerat, S., & Tongchitpakdee, S. (2011). Coconut (*Cocos nucifera* L.). In E. M. Yahia (Ed.), *Postharvest biology and technology of tropical and subtropical fruits*. Woodhead Publishing.

Yong, J. W. H., Ge, L., Ng, Y. F., & Tan, S. N. (2009). The chemical composition and biological properties of coconut (*Cocos nucifera* L.) water. *Molecules*, *14*, 5144-5164. doi:10.3390/molecules14125144

Zhu, P., Zhu, W., Hu, Q., Zhang, C., & Zuo, W. (2017). Subspace clustering guided unsupervised feature selection. *Pattern Recognition*, *66*, 364-374. doi:10.1016/j.patcog.2017.01.016

Available online at https://li02.tci-thaijo.org/index.php/ssstj

Suan Sunandha Science and Technology Journal

A New Type of Extended Soft Set Operation: Complementary Extended Union Operation

Aslıhan Sezgin[1*], Ahmet Mücahit Demirci[2], Emin Aygün[3]

[1]Department of Mathematics and Science Education, Faculty of Education, Amasya University, Amasya, Türkiye
[2]Department of Mathematics, Graduate School of Natural and Applied Sciences, Amasya University, Amasya, Türkiye
[3]Department of Mathematics, Faculty of Science, Erciyes University, Kayseri, Türkiye
***Corresponding author e-mail: aslihan.sezgin@amasya.edu.tr**

Received: *19 April 2024* / **Revised:** *18 June 2024* / **Accepted:** *13 July 2024*

Abstract

Soft set theory was proposed by Molodtsov in 1999 to model some problems involving uncertainty. It has a wide range of theoretical and practical applications. Soft set operations constitute the basic building blocks of soft set theory. Many kinds of soft set operations have been described and applied in various ways since the inception of the theory. In this paper, to contribute to the theory, a new soft set operation, called complementary extended union operation, is defined, its properties are discussed in detail to obtain the relationship of each operation with other soft set operations, and the distributions of these operations over other soft set operations are examined. We obtain that the complementary extended union operation along with other certain types of soft set operations construct some well-known algebraic structure such as Boolean Algebra, De Morgan Algebra, semiring, and hemiring in the set of soft sets with a fixed parameter set. Since Boolean Algebra is fundamental in digital logic design, computer science, information retrieval, set theory and probability; De Morgan Algebra in logic and set theory, computer science, artificial intelligence, circuit design; semirings in theoretical computer science, optimization problems, economics, cryptography and coding theory, and hemirings in combinatorics, mathematical economics, theoretical computer science, these algebraic structures provide essential tools for various applications, facilitating the analysis, design, and optimization of systems across many disciplines, and thus this study is expected to contribute to decision-making methods and cryptography based on soft sets.

Keywords: Soft set, Complementary extended soft set operations, Semiring, Hemiring, Algebras

1. Introduction

Many theories have been put forward to explain uncertainties for years. One of the most important theories in this field is the theory of fuzzy sets, proposed by Zadeh in 1965. A fuzzy set is defined through its membership function. The degree of membership is the real number between [0,1]. 0 indicates no membership, and 1 indicates full membership. However, due to the structural problems of the membership function, Molodtsov (1999) proposed soft set theory. The soft set theory eliminated the structural problems of the membership function. Molodtsov successfully applied soft set theory to continuous differentiable functions, operation research, Riemann integration, and many other fields. Soft set operations constitute the basis of soft set theory, and studies on both soft algebraic structures and soft decision-making methods are based on soft set operations. In this regard, Maji, Biswas, and Roy (2003) started inspiring studies on soft set operations. A more widely accepted definition of soft subset than the one defined by Maji et al. (2003) was given by Pei and Miao (2005).

When the studies of soft set operations such as Maji et al. (2003), Ali et al. (2009), Feng et al. (2010), Jiang et al. (2010), Ali et al. (2011), Fu (2011), Ge and Yang (2011), Neog and Sut (2011), Sezgin and Atagün (2011), Park et al. (2012), Singh and Onyeozili (2012), Zhu and Wen (2013), Onyeozili and Gwary (2014), Sen (2014), Husain and Shivani (2018), Sezgin et al. (2019), and Stojanovic (2021) are examined, it is seen that soft set operations proceed under two separate headings: restricted and extended operations. Moreover, Eren and Çalışıcı (2019) defined a new form of soft set operation for the literature, called the soft binary piecewise difference operation, Sezgin and Çalışıcı (2024) improved the work of Eren and Çalışıcı (2019) and studied the properties of the soft binary piecewise difference operation by comparing it with the difference operation in classical sets.

Çağman (2021) and Sezgin et al. (2023c) introduced new binary set operations, and these operations were transferred to soft sets as new restricted soft set operations and extended soft set operations by Aybek (2024). Besides, some new form of soft set operations, different from the restricted and extended forms of operations, were introduced by various authors (Akbulut, 2024; Demirci, 2024; Sarıalioğlu, 2024; Sezgin & Atagün, 2023; Sezgin & Aybek, 2023; Sezgin et al., 2023a; Sezgin et al., 2023b; Sezgin & Çağman, 2024; Sezgin & Dagtoros, 2023; Sezgin & Demirci, 2023; Sezgin & Sarıalioğlu, 2024; Sezgin & Yavuz, 2023a), and soft set operations, one of the most fundamental elements of soft set theory, has been studied by researchers since the theory was introduced. For other applications of soft sets as regards algebraic structures, we refer to: Çağman et al. (2012), Sezer (2014), Sezer et al. (2015), Muştuoğlu et al. (2016), Sezgin (2016), Tunçay and Sezgin (2016), Sezgin et al. (2017), Atagün and Sezgin (2018), Mahmood et al. (2018), Sezgin (2018), Jana et al. (2019), Özlü and Sezgin (2020), and Sezgin et al. (2022).

Moreover, different types of soft eqaulities were defined and some important equivalance relations were obtained with these different types of soft equalites as Jun and Yang (2011), Liu et al. (2012), Feng and Li (2013), Abbas et al. (2014), Abbas et al. (2017), Al-shami (2019), Al-shami and El-Shafei (2020). Studying the soft algebraic structures of an algebraic structure and other types of soft sets has been of interest by the researchers as Ali et al. (2015), Iftikhar and Mahmood (2018), and Mahmood (2020). For possible applications of graphs and network research concerning soft sets, we refer to Pant et al. (2024).

In the scope of algebra, one of the most important mathematical issues is analyzing the properties of the operation defined on a set to classify algebraic structures. Two primary categories of soft set collections are investigated within the context of soft sets as algebraic structures: The first represents a class of soft sets with a fixed set of parameters, whereas the second represents a class of soft sets with changing parameter sets. Depending on the extra properties, these two types of collections have different algebraic properties.

Boolean Algebras, De Morgan Algebras, semirings and hemirings provide essential tools for various applications, facilitating the analysis, design, and optimization of systems across many disciplines. Boolean algebra is fundamental in designing and simplifying digital circuits, algorithm design, data structure optimization, and coding theory, information retrieval, set theory and probability. De Morgan Algebra is essential in logic and set theory, especially in formal logic, helping to simplify complex logical expressions and set operations, computer science, design and optimization of algorithms, artificial intelligence, simplification of logic circuits by transforming expressions to their simplest forms. Semirings are crucial in theoretical computer science, automata theory, formal languages, and the analysis of algorithms, optimization problems, dynamic programming and shortest path algorithms, such as those used in network routing, mathematical modeling of economic systems, including cost-benefit analysis and resource allocation, constructing cryptographic protocols and coding theory. Hemiring are of importance as they generalize rings and semirings, providing a framework for more complex algebraic structures. They are used in the study of combinatorial structures and their properties, economic systems where subtraction may not always be feasible or meaningful and theoretical computer science, especially in automata theory.

In this study, we propose a new type of soft set operation called the complementary extended union operation and thoroughly discuss its properties to contribute to the theory of soft sets. To determine the relationship between the operation and other soft set operations, the distribution of complementary extended union operations over other kinds of soft set operations is examined. It is demonstrated that in the set of soft sets with a fixed parameter, the complementary extended union operation forms many well-known and significant algebraic structures in classical algebra, including semiring, hemiring, Boolean Algebra, and De Morgan Algebra, along with other specific kinds of soft set operations.

2. Preliminaries

Definition 2.1. Let U be the universal set, E be the parameter set, $P(U)$ be the power set of U, and $V \subseteq E$. A pair (F, V) is called a soft set over U where F is a set-valued function such that $F: V \to P(U)$ (Molodtsov, 1999).

Throughout this paper, the set of all the soft sets over U (no matter what the parameter set is) is designated by $S_E(U)$. Let V be a fixed subset of E. Then, $S_V(U)$ represents the collection of all soft sets over U with the fixed parameter set V.

Definition 2.2. $(G,\aleph)$ is called a relative null soft set (with respect to the parameter set $\aleph$), denoted by $\emptyset_\aleph$, if $G(t)=\emptyset$ for all $t\in\aleph$ and $(G,\aleph)$ is called a relative whole soft set (with respect to the parameter set $\aleph$), denoted by $U_\aleph$ if $G(t)=U$ for all $t\in\aleph$. The relative whole soft set U_E with respect to the universe set of parameters E is called the absolute soft set over U (Ali et al. 2009).

Definition 2.3. For two soft sets $(F,\aleph)$ and (G,R), we say that $(F,\aleph)$ is a soft subset of (G,R) and it is denoted by $(F,\aleph)\tilde{\subseteq}(G,R)$, if $\aleph\subseteq R$ and $F(t)\subseteq G(t)$, for all $t\in\aleph$. Two soft sets $(F,\aleph)$ and (G,R) are said to be soft equal if $(F,\aleph)$ is a soft subset of (G,R) and (G,R) is a soft subset of $(F,\aleph)$ (Pei & Miao, 2005).

Definition 2.4. The relative complement of a soft set $(F,\aleph)$, denoted by $(F,\aleph)^r$, is defined by $(F,\aleph)^r=(F^r,\aleph)$, where $F^r:\aleph\to P(U)$ is a mapping given by $(F,\aleph)^r=U\backslash F(t)$ for all $t\in\aleph$ (Ali et al. 2009). From now on, $U\backslash F(t)=[F(t)]'$ will be designated by $F'(t)$ for the sake of designation.

Çağman (2021) defined two new complements as inclusive and exclusive complements. Let + and θ denote inclusive and exclusive complements, respectively, and let V and Y be two sets. Then, these binary operations are as follows: $V+Y=V'\cup Y$, $V\theta Y=V'\cap Y'$. Sezgin et al. (2023c) analyzed the relations between these two operations and also defined three new binary operations and examined their relations with each other. Let V and Y be two sets $V*Y=V'\cup Y'$, $V\gamma Y=V'\cap Y$, $V\lambda Y=V\cup Y'$ and $\circledcirc$ denote the operations on sets. Then, restricted operations on soft sets, extended operations, extended operations with complement, soft binary piecewise operations, and soft binary piecewise operations with complement can be given in general form with the following generalized definitions:

Definition 2.5. Let $(F,V),(G,Y)\in S_E(U)$. The restricted $\circledcirc$ operation of (F,V) and (G,Y) is the soft set $(H,Ƶ)$, denoted to be $(F,V)\circledcirc_{\Re}(G,Y)=(H,Ƶ)$, where $Ƶ=V\cap Y\neq\emptyset$ and for all $\aleph\in Ƶ$, $H(\aleph)=F(\aleph)\circledcirc G(\aleph)$. Here, if $Ƶ=V\cap Y=\emptyset$, then $(F,V)\circledcirc(G,Y)=\emptyset_\emptyset$ (Ali et al., 2009; Aybek, 2024; Sezgin & Atagün, 2011).

Definition 2.6. Let $(F,V),(G,Y)\in S_E(U)$. The extended $\circledcirc$ operation (F,V) and (G,Y) is the soft set $(H,Ƶ)$, denoted by $(F,K)\circledcirc_\varepsilon(G,Y)=(H,Ƶ)$, where $Ƶ=V\cup Y$ and for all $\aleph\in Ƶ$,

$$H(v)=\begin{cases}F(v), & v\in V-Y\\ G(v), & v\in Y-V\\ F(v)\circledcirc G(v), & v\in V\cap Y\end{cases}$$

(Ali et al., 2009; Aybek, 2024; Maji et al., 2003; Sezgin et al., 2019; Stojanovic, 2021).

Definition 2.7. Let $(F,V),(G,Y)\in S_E(U)$. The complementary extended $\circledcirc$ operation (F,V) and (G,Y) is the soft set $(H,Ƶ)$, denoted by $(F,K)\overset{*}{\circledcirc}_\varepsilon(G,Y)=(H,Ƶ)$, where $Ƶ=V\cup Y$ and for all $\aleph\in Ƶ$,

$$H(v)=\begin{cases}F'(v), & v\in V-Y\\ G'(v), & v\in Y-V\\ F(v)\circledcirc G(v), & v\in V\cap Y\end{cases}$$

(Akbulut, 2024; Demirci, 2024; Sarıalioğlu, 2024).

Definition 2.8. Let $(F,V),(G,Y)\in S_E(U)$. The soft binary piecewise $\circledcirc$ of (F,V) and (G,Y) is the soft set (H,V), denoted by $(F,V)\underset{\circledcirc}{\sim}(G,Y)=(H,V)$, where for all $\aleph\in V$,

$$H(v)=\begin{cases}F(v), & v\in V-Y\\ F(v)\circledcirc G(v), & v\in V\cap Y\end{cases}$$

(Eren & Çalışıcı, 2019; Sezgin & Çalışıcı, 2024; Sezgin & Yavuz, 2023b; Yavuz, 2024).

Definition 2.9. Let $(F,V),(G,Y) \in S_E(U)$. The complementary soft binary piecewise ⊙ of (F,V) and (G,Y) is the soft set (H,V), denoted by $(F,V) \underset{\odot}{\overset{*}{\sim}} (G,Y) = (H,V)$, where for all $\aleph \in V$,

$$H(v) = \begin{cases} F'(v), & v \in V - Y \\ F(v) \odot G(v), & v \in V \cap Y \end{cases}$$

(Sezgin & Atagün, 2023; Sezgin & Aybek, 2023; Sezgin et al., 2023a; Sezgin et al., 2023b; Sezgin & Çağman, 2024; Sezgin & Dagtoros, 2023; Sezgin & Demirci, 2023; Sezgin & Sarıalioğlu, 2024; Sezgin & Yavuz, 2023a).

Definition 2.10. Let (S, ⊙) be an algebraic structure. An element $s \in S$ is called idempotent if $s^2=s$. If $s^2=s$ for all $s \in S$, then the algebraic structure (S,⋆) is said to be idempotent. An idempotent semigroup is called a band, an idempotent and commutative semigroup is called a semilattice, and an idempotent and commutative monoid is called a bounded semilattice (Clifford, 1954).

Definition 2.11. Let S be a non-empty set, and let "+" and "⊙" be two binary operations defined on S. If the algebraic structure (S, +, ⋆) satisfies the following properties, then it is called a semiring:

i. (S, +) is a semigroup,
ii. (S, ⊙) is a semigroup,
iii. $x \odot (y + z) = x \odot y + x \odot z$ and $(x + y) \odot z = x \odot z + y \odot z$ for all $x, y, z \in S$.

If $x+y=y+z$ for all $x,y \in S$, then S is called an additive commutative semiring. If $x \odot y = yx$ for all $x,y \in S$, then S is called a multiplicative commutative semiring. If there exists an element $1 \in S$ such that $x \odot 1 = 1 \odot x = x$ for all $x \in S$ (multiplicative identity), then S is called semiring with unity. If there exists $0 \in S$ such that $0 \odot x = x \odot 0 = 0$ and $0+x=x+0=x$ for all $x \in S$, then 0 is called the zero of S. A semiring with commutative addition and a zero element is called a hemiring (Vandiver, 1934).

Definition 2.12. Let L be a non-empty set, and let "∨" and "∧" be two binary operations defined on L. If the algebraic structure. (L,∨,∧) satisfies the following properties, then it is called a lattice:

i. (L, ∨) a commutative, idempotent semigroup (semilattice)
ii. (L, ∧) a commutative, idempotent semigroup (semilattice)
iii. $\Upsilon \vee (\Upsilon \wedge y) = \Upsilon \wedge (\Upsilon \vee y)$, for all $\Upsilon, y \in L$ (absorption law)
iv.

A lattice with an identity element according to both operations is called a bounded lattice. In a bounded lattice, the identity element of L with respect to the ∧ operation is usually denoted by 1, while the identity element with respect to the ∨ operation is denoted by 0. If the bounded lattice L has an element Υ' such that $\Upsilon \wedge \Upsilon' = 0$ and $\Upsilon \vee \Upsilon' = 1$ for all $\Upsilon \in L$, then L is called a complemented lattice. A lattice holding distribution law is called a distributive lattice. A lattice that is bounded, distributive, and at the same time complemented is called a Boolean algebra. The lattice with De Morgan's laws, i.e., $(\Upsilon \vee y)' = \Upsilon' \wedge y'$ and $(\Upsilon \wedge y)' = \Upsilon' \vee y'$ for all $\Upsilon, y \in L$ is called De Morgan algebra (Birkhoff, 1967).

Definition 2.13. Let M be a non-empty set with the binary operation "⊙" and the unary operation "*" defined on M. If 0 is a constant that satisfies the following axioms for any Υ, $y \in M$, then the structure (M, ⊙,∗ ,0) is called an MV-algebra:

i. (M, ⊕, 0) is a commutative monoid
ii. $(\Upsilon^*)^* = \Upsilon$
iii. $0^* \odot \Upsilon = 0^*$
iv. $(\Upsilon^* \odot y)^* \odot y = (y^* \odot \Upsilon)^* \odot \Upsilon$

(Chang, 1959).

3. Complementary Extended Union Operation

In this section, a new soft set operation called the complementary extended union operation is introduced, full algebraic properties of the operation are analyzed by comparing its properties with the union operation in classical set theory.

Definition 3.1. Let $(F, Ƶ)$ and (G, B) be two soft sets over U. The complementary extended union operation of $(F, Ƶ)$ and (G, B) is the soft set $(H, Ş)$, denoted by $(F, Ƶ) \overset{*}{\cup}_{\varepsilon} (G, B) = (H, Ş)$, where for all $\aleph \in Ş = Ƶ \cup B$,

$$H(\aleph)=\begin{cases} F'(\aleph), & \aleph \in Ƶ - B \\ G'(\aleph), & \aleph \in B - Ƶ \\ F(\aleph) \cup G(\aleph), & \aleph \in Ƶ \cap B \end{cases}$$

Example 3.2. Let $E=\{e_1,e_2,e_3,e_4\}$ be the parameter set and $Ƶ=\{e_1, e_3\}$ and $B=\{e_2, e_3, e_4\}$ be two subsets of E, and $U=\{h_1,h_2,h_3,h_4,h_5\}$ be the universal set.
Assume that $(F,Ƶ)=\{(e_1,\{h_2,h_5\}),(e_3,\{h_1,h_2,h_5\})\}$, $(G,B)=\{(e_2,\{h_1,h_4,h_5\}),(e_3,\{h_2,h_3,h_4\}),(e_4,\{h_3,h_5\})\}$ be soft sets over U. Let $(F, Ƶ) \overset{*}{\cup}_{\varepsilon} (G,B)=(H,Ƶ\cup B)$, where for all $\aleph \in Ƶ\cup B$,

$$H(\aleph)=\begin{cases} F'(\aleph), & \aleph \in Ƶ - B \\ G'(\aleph), & \aleph \in B - Ƶ \\ F(\aleph) \cup G(\aleph), & \aleph \in Ƶ \cap B \end{cases}$$

Here, since $Ƶ\cup B=\{e_1,e_2,e_3,e_4\}$, $Ƶ-B=\{e_1\}$, $B-Ƶ=\{e_2,e_4\}$, $Ƶ\cap B=\{e_3\}$,
$H(e_1) =F'(e_1)=\{h_1,h_3,h_4\}$, $H(e_2) =G'(e_2)=\{h_2,h_3\}$, $H(e_4) =G'(e_4)=\{h_1,h_2,h_4\}$ and $H(e_3)=F(e_3)\cup G(e_3)=\{h_1,h_2,h_5\}\cup\{h_2,h_3,h_4\}=\{h_1,h_2,h_3,h_4,h_5\}$.
Thus, $(F, Ƶ) \overset{*}{\cup}_{\varepsilon}(G,B)=\{(e_1,\{h_1,h_3,h_4\}), (e_2,\{h_2,h_3\}),(e_3,\{h_1,h_2,h_3,h_4,h_5\}),(e_4,\{h_1,h_2,h_4\})\}$

Theorem 3.3. Algebraic Properties of Operation

1) $\overset{*}{\cup}_{\varepsilon}$ is closed in $S_E(U)$.

Proof: It is clear that $\overset{*}{\cup}_{\varepsilon}$ is a binary operation in $S_E(U)$. Indeed,

$$\overset{*}{\cup}_{\varepsilon} : S_E(U) \times S_E(U) \rightarrow S_E(U)$$

$$((F, Ƶ), (G,B)) \rightarrow (F,Ƶ) \overset{*}{\cup}_{\varepsilon} (G,B)=(H, Ƶ \cup B)$$

Similarly,

$$\overset{*}{\cup}_{\varepsilon} : S_Ƶ(U) \times S_Ƶ(U) \rightarrow S_Ƶ(U)$$

$$((F, Ƶ), (G, Ƶ)) \rightarrow (F, Ƶ) \overset{*}{\cup}_{\varepsilon}(G,Ƶ)=(T, Ƶ \cup Ƶ)=(T, Ƶ)$$

That is, when Ƶ is a fixed subset of the set E and (F, Ƶ) and (G, Ƶ) be elements of $S_Ƶ(U)$, then so is $(F,Ƶ) \overset{*}{\cap}_{\varepsilon} (G, Ƶ)$. Namely, $S_Ƶ$ is closed under $\overset{*}{\cup}_{\varepsilon}$ either.

2) $[(F, Ƶ) \overset{*}{\cup}_{\varepsilon} (G,B)] \overset{*}{\cup}_{\varepsilon} (H,Ş) \neq (F, Ƶ) \overset{*}{\cup}_{\varepsilon} [(G,B) \overset{*}{\cup}_{\varepsilon} (H,Ş)]$.

Proof: First, let's consider the left-hand side (LHS). Suppose $(F, Ƶ) \overset{*}{\cup}_{\varepsilon} (G,B)=(T,Ƶ\cup B)$, where for all $\aleph \in Ƶ\cup B$,

$$T(ℵ)=\begin{cases} F'(ℵ), & ℵ∈Z-B \\ G'(ℵ), & ℵ∈B-Z \\ F(ℵ)∪G(ℵ), & ℵ∈Z∩B \end{cases}$$

Let $(T,Z∪B) \overset{*}{∪}_{ε} (H,Ş) = (M,Z∪B∪Ş)$, where for all $ℵ∈Z∪B∪Ş$,

$$M(ℵ)=\begin{cases} T'(ℵ), & ℵ∈(Z∪B)-Ş \\ H'(ℵ), & ℵ∈Ş-(Z∪B) \\ T(ℵ)∪H(ℵ), & ℵ∈(Z∪B)∩Ş \end{cases}$$

Hence,

$$M(ℵ)=\begin{cases} F(ℵ), & ℵ∈(Z-B)-Ş=Z∩B'∩Ş' \\ G(ℵ), & ℵ∈(B-Z)-Ş=Z'∩B∩Ş' \\ F'(ℵ)∩G'(ℵ), & ℵ∈(Z∩B)-Ş=Z∩B∩Ş' \\ H'(ℵ) & ℵ∈Ş-(Z∪B)=Z'∩B'∩Ş \\ F'(ℵ)∪H(ℵ), & ℵ∈(Z-B)∩Ş=Z∩B'∩Ş \\ G'(ℵ)∪H(ℵ), & ℵ∈(B-Z)∩Ş=Z'∩B∩Ş \\ (F(ℵ)∪G(ℵ))∪H(ℵ), & ℵ∈(Z∩B)∩Ş=Z∩B∩Ş \end{cases}$$

Now consider the RHS, i.e., $(F,Z) \overset{*}{∪}_{ε} [(G,B) \overset{*}{∪}_{ε} (H,Ş)]$. Let $(G,B) \overset{*}{∪}_{ε} (H,Ş)=(K,B∪Ş)$. So, for all $ℵ∈B∪Ş$,

$$K(ℵ)=\begin{cases} G'(ℵ), & ℵ∈B-Ş \\ H'(ℵ), & ℵ∈Ş-B \\ G(ℵ)∪H(ℵ), & ℵ∈B∩Ş \end{cases}$$

Let $(F,Z) \overset{*}{∪}_{ε} (K,B∪Ş) = (S,Z∪B∪Ş)$. Thus, for all $ℵ∈Z∪B∪Ş$,

$$S(ℵ)=\begin{cases} F'(ℵ), & ℵ∈Z-(B∪Ş) \\ K'(ℵ), & ℵ∈(B∪Ş)-Z \\ F(ℵ)∪K(ℵ), & ℵ∈Z∩(B∪Ş) \end{cases}$$

Hence,

$$S(ℵ)=\begin{cases} F'(ℵ), & ℵ∈Z-(B∪Ş)=Z∩B'∩Ş' \\ G(ℵ), & ℵ∈(B-Ş)-Z=Z'∩B∩Ş' \\ H(ℵ), & ℵ∈(Ş-B)-Z=Z'∩B'∩Ş \\ G'(ℵ)∩H'(ℵ), & ℵ∈(B∩Ş)-Z=Z'∩B∩Ş \\ F(ℵ)∪G'(ℵ), & ℵ∈Z∩(B-Ş)=Z∩B∩Ş' \\ F(ℵ)∪H'(ℵ), & ℵ∈Z∩(Ş-B)=Z∩B'∩Ş \\ F(ℵ)∪(G(ℵ)∪H(ℵ)), & ℵ∈Z∩(B∩Ş)=Z∩B∩Ş \end{cases}$$

Thus, $M≠S$. That is, in the set $S_E(U)$, $\overset{*}{∪}_{ε}$ is not associative. However, we have the following:

3) $[(F,Z) \overset{*}{∪}_{ε} (G,Z)] \overset{*}{∪}_{ε} (H,Z) = (F,Z) \overset{*}{∪}_{ε} [(G,Z) \overset{*}{∪}_{ε} (H,Z)]$.

4) $(F,Z) \overset{*}{∪}_{ε} (G,B)=(G,B) \overset{*}{∪}_{ε} (F,Z)$.

Proof: Firstly, the parameter sets of the soft set on both sides of the equation is $Z∪B$, and thus the first condition of the soft equality is satisfied. Now let's consider the LHS. Let $(F,Z) \overset{*}{∪}_{ε} (G,B)=(H,Z∪B)$, where for all $ℵ∈Z∪B$,

$$H(ℵ)=\begin{cases} F'(ℵ), & ℵ∈Z-B \\ G'(ℵ), & ℵ∈B-Z \\ F(ℵ)∪G(ℵ), & ℵ∈Z∩B \end{cases}$$

Now consider the RHS, i.e., $(G,B) \overset{*}{∪}_{ε} (F,Z)$. Let $(G,B) \overset{*}{∪}_{ε} (F,Z)=(T,B∪Z)$, where for all $ℵ∈B∪Z$,

$$T(\aleph)=\begin{cases} G'(\aleph), & \aleph\in B-Z \\ F'(\aleph), & \aleph\in Z-B \\ G(\aleph)\cup F(\aleph), & \aleph\in B\cap Z \end{cases}$$

Thus, it is seen that H=T. Similarly, one can show that (F, Z) $\overset{*}{\cup}_{\varepsilon}$(G,Z)=(G, Z) $\overset{*}{\cup}_{\varepsilon}$(F,Z. That is, $\overset{*}{\cap}_{\varepsilon}$ operation is commutative in both $S_E(U)$ and $S_Z(U)$.

5) (F, Z) $\overset{*}{\cup}_{\varepsilon}$(F,Z)=(F,Z).

Proof: Let (F, Z) $\overset{*}{\cup}_{\varepsilon}$ (F,Z)=(H,Z∪Z), where for all ℵ∈Z,

$$H(\aleph)=\begin{cases} F'(\aleph), & \aleph\in Z-Z=\emptyset \\ F'(\aleph), & \aleph\in Z-Z=\emptyset \\ F(\aleph)\cup F(\aleph), & \aleph\in Z\cap Z=Z \end{cases}$$

Hence, H(ℵ)= F(ℵ)∪F(ℵ)=F(ℵ), for all ℵ∈Z. Thus, (H,Z)=(F,Z). That is, $\overset{*}{\cup}_{\varepsilon}$ is idempotent in $S_E(U)$.

6) (F, Z) $\overset{*}{\cup}_{\varepsilon}$ $\emptyset_Z$=$\emptyset_Z$ $\overset{*}{\cup}_{\varepsilon}$(F,Z)=(F,Z).

Proof: Let $\emptyset_Z$ =(S,Z). Then, for all ℵ∈Z, S(ℵ)= ∅. Let (F, Z) $\overset{*}{\cup}_{\varepsilon}$ (S,Z)=(H,Z∪Z), where for all ℵ∈Z,

$$H(\aleph)=\begin{cases} F'(\aleph), & \aleph\in Z-Z=\emptyset \\ S'(\aleph), & \aleph\in Z-Z=\emptyset \\ F(\aleph)\cup S(\aleph), & \aleph\in Z\cap Z=Z \end{cases}$$

Hence, H(ℵ) = F(ℵ)∪S(ℵ)= F(ℵ)∪∅=F(ℵ), for all ℵ∈Z. Thus, (H,Z)=(F,Z). That is, in $S_Z(U)$, the identity element of $\overset{*}{\cup}_{\varepsilon}$ is the soft set $\emptyset_Z$.

Corollary 3.3.1. By Theorem 3.3 (1), (3), (4), (5), and (6), ($S_Z(U)$, $\overset{*}{\cup}_{\varepsilon}$) is a commutative, idempotent monoid, that is, a bounded semilattice, whose identity element is $\emptyset_Z$, where Z ⊆ E is a fixed set of parameters. Moreover, from Theorem 3.3 (2), $\overset{*}{\cup}_{\varepsilon}$ cannot form a semigroup as it is not associative in S_E (U). Thus, ($S_E(U)$, $\overset{*}{\cup}_{\varepsilon}$) is a groupoid.

7) (F, Z) $\overset{*}{\cup}_{\varepsilon}$ U_Z=U_Z $\overset{*}{\cup}_{\varepsilon}$(F,Z)= U_Z.

Proof: Let U_Z =(T,Z). Thus, T(ℵ)=U for all ℵ∈Z. Let (F, Z) $\overset{*}{\cup}_{\varepsilon}$(T, A) =(H,Z∪Z), where for all ℵ∈Z,

$$H(\aleph)=\begin{cases} F'(\aleph), & \aleph\in Z-Z=\emptyset \\ T'(\aleph), & \aleph\in Z-Z=\emptyset \\ F(\aleph)\cup T(\aleph), & \aleph\in Z\cap Z=Z \end{cases}$$

Hence, H(ℵ) = F(ℵ)∪T(ℵ)= F(ℵ)∪U=U, for all ℵ∈Z, and (H,Z)= U_Z. That is, the absorbing element of $\overset{*}{\cup}_{\varepsilon}$ in $S_Z(U)$ is the soft set U_Z.

8) (F, Z) $\overset{*}{\cap}_{\varepsilon}$ $\emptyset_\emptyset$= $\emptyset_\emptyset$ $\overset{*}{\cap}_{\varepsilon}$(F,Z) = $(F,Z)^r$.

Proof: Let $\emptyset_\emptyset$= (K, ∅) and (F, Z) $\overset{*}{\cup}_{\varepsilon}$(K, ∅) = (Q,Z∪∅) = (Q, Z), where for all ℵ∈ Z,

$$Q(\aleph)=\begin{cases} F'(\aleph), & \aleph\in Z\text{-}\emptyset=Z \\ K'(\aleph), & \aleph\in\emptyset\text{-}Z=\emptyset \\ F(\aleph)\cup K(\aleph), & \aleph\in Z\cap\emptyset=\emptyset \end{cases}$$

Hence, $Q(\aleph) = F'(\aleph)$, for all $\aleph\in Z$, so $(Q,Z) = (F,Z)^r$.

9) $(F, Z) \overset{*}{\cup_\varepsilon} (F,Z)^r = (F,Z)^r \overset{*}{\cup_\varepsilon} (F,Z)=U_Z$.

Proof: Let $(F,Z)^r = (H,Z)$, hence $H(\aleph)=F'(\aleph)$ for all $\aleph\in Z$. Let $(F, Z) \overset{*}{\cup_\varepsilon} (H,Z)=(T,Z\cup Z)$, where for all $\aleph\in Z$,

$$T(\aleph)=\begin{cases} F'(\aleph), & \aleph\in Z\text{-}Z=\emptyset \\ H'(\aleph), & \aleph\in Z\text{-}Z=\emptyset \\ F(\aleph)\cup H(\aleph), & \aleph\in Z\cap Z=Z \end{cases}$$

Here, $T(\aleph)= F(\aleph)\cup H(\aleph)= F(\aleph) \cup F'(\aleph)=U$ for all $\aleph\in Z$. Thus, $(T,Z)= U_Z$.

10) $[(F, Z) \overset{*}{\cup_\varepsilon} (G,B)]^r =(F,Z)^r \overset{*}{\cap_\varepsilon} (G,B)^r$

Proof: Let $(F, Z) \overset{*}{\cup} (G,B)=(H,Z\cup B)$, where for all $\aleph\in Z\cup B$,

$$H(\aleph)=\begin{cases} F'(\aleph), & \aleph\in Z\text{-}B \\ G'(\aleph), & \aleph\in B\text{-}Z \\ F(\aleph)\cup G(\aleph), & \aleph\in Z\cap B \end{cases}$$

Let $(H,Z\cup B)^r =(T,Z\cup B)$, where for all $\aleph\in Z\cup B$,

$$T(\aleph)=\begin{cases} F(\aleph), & \aleph\in Z\text{-}B \\ G(\aleph), & \aleph\in B\text{-}Z \\ F'(\aleph)\cap G'(\aleph), & \aleph\in Z\cap B \end{cases}$$

Hence, $(T,Z\cup B) =(F,Z)^r \overset{*}{\cap_\varepsilon} (G,B)^r$

11) $(F, Z) \overset{*}{\cup_\varepsilon} (G, Z)= \emptyset_Z \Leftrightarrow (F, Z) = (G, Z) = \emptyset_Z$.

Proof: Let $(F, Z) \overset{*}{\cup_\varepsilon} (G, Z) = T,(Z\cup Z)$, where for all $\aleph\in Z$,

$$T(\aleph)=\begin{cases} F'(\aleph), & \aleph\in Z\text{-}Z=\emptyset \\ G'(\aleph), & \aleph\in Z\text{-}Z=\emptyset \\ F(\aleph)\cup G(\aleph), & \aleph\in Z\cap Z=Z \end{cases}$$

Since $(T,Z)= \emptyset_Z$, $T(\aleph)=F(\aleph)\cup G(\aleph)=\emptyset$, for all $\aleph \in Z$. Hence, $F(\aleph)=G(\aleph)=\emptyset$ for all $\aleph \in Z$. Thus, $(F, Z) = (G, Z) = \emptyset_Z$.

12) $\emptyset_Z \,\tilde{\subseteq}\, (F, Z) \overset{*}{\cup_\varepsilon} (G,B)$, $\emptyset_B \,\tilde{\subseteq}\, (F, Z) \overset{*}{\cup_\varepsilon} (G,B)$, $\emptyset_{Z\cup B} \,\tilde{\subseteq}\, (F, Z) \overset{*}{\cup_\varepsilon} (G,B)$, $(F, Z) \overset{*}{\cup_\varepsilon} (G,B) \,\tilde{\subseteq}\, U_{Z\cup B}$.

13) $(F,Z) \,\tilde{\subseteq}\, (F, Z) \overset{*}{\cup_\varepsilon} (G,Z)$ and $(G,Z) \,\tilde{\subseteq}\, (F, Z) \overset{*}{\cup_\varepsilon} (G,Z)$.

Proof: Le $(F, Z) \overset{*}{\cup_\varepsilon} (G,Z)=(H, Z \cup Z)$, where for all $\aleph\in Z$,

$$H(\aleph)=\begin{cases} F'(\aleph), & \aleph\in Z\text{-}Z=\emptyset \\ G'(\aleph), & \aleph\in Z\text{-}Z=\emptyset \\ F(\aleph) \cup G(\aleph), & \aleph\in Z\cap Z=Z \end{cases}$$

Since $H(\aleph)=F(\aleph) \subseteq F(\aleph) \cup G(\aleph)$, for all $\aleph\in Z$, $(F,Z)\,\tilde{\subseteq}\,(F, Z) \overset{*}{\cup_\varepsilon} (G,Z)$. Similarly, since $H(\aleph)=G(\aleph) \subseteq F(\aleph) \cup G(\aleph)$, for all $\aleph\in Z$, $(G,Z) \,\tilde{\subseteq}\, (F, Z) \overset{*}{\cup_\varepsilon} (G,Z)$.

14) $(F,Z) \,\tilde{\subseteq}\, (G, Z) \Leftrightarrow (F, Z) \overset{*}{\cup_\varepsilon} (G,Z) =(G,Z)$.

Proof: Let $(F,Ƶ) \cong (G,Ƶ)$. Then, $F(ℵ) \subseteq G(ℵ)$ for all $ℵ \in Ƶ$, and so $G'(ℵ) \subseteq F'(ℵ)$. Let $(F,Ƶ) \overset{*}{\cup_\varepsilon} (G,Ƶ)=(H,Ƶ\cupƵ)$, where for all $ℵ \in Ƶ$,

$$H(ℵ)=\begin{cases} F'(ℵ), & ℵ \in Ƶ-Ƶ=\emptyset \\ G'(ℵ), & ℵ \in Ƶ-Ƶ=\emptyset \\ F(ℵ) \cup G(ℵ), & ℵ \in Ƶ\capƵ=Ƶ \end{cases}$$

Thus, $H(ℵ)=F(ℵ) \cup G(ℵ)=G(ℵ,)$ for all $ℵ \in Ƶ$. Hence, $(F,Ƶ) \overset{*}{\cup_\varepsilon} (G,Ƶ)=(G,Ƶ)$.

Conversely, let $(F,Ƶ) \overset{*}{\cup_\varepsilon} (G,Ƶ)=(G,Ƶ)$. Hence, $F(ℵ) \cup G(ℵ)=G(ℵ)$, for all $ℵ \in Ƶ$ and thus, $F(ℵ) \subseteq G(ℵ)$. Therefore, $(F,Ƶ) \cong (G,Ƶ)$.

15) $(F,Ƶ) \overset{*}{\cap_\varepsilon} (G,B) \cong (F,Ƶ) \overset{*}{\cup_\varepsilon} (G,B)$.

Proof: Let $(F,Ƶ) \overset{*}{\cap_\varepsilon} (G,B)=(H,Ƶ\cup B)$, where $ℵ \in Ƶ\cup B$,

$$H(ℵ)=\begin{cases} F'(ℵ), & ℵ \in Ƶ-B \\ G'(ℵ), & ℵ \in B-Ƶ \\ F(ℵ) \cap G(ℵ), & ℵ \in Ƶ\cap B \end{cases}$$

Let $(F,Ƶ) \overset{*}{\cup_\varepsilon} (G,B)=(T,Ƶ\cup B)$, where for all $ℵ \in Ƶ\cup B$,

$$T(ℵ)=\begin{cases} F'(ℵ), & ℵ \in Ƶ-B \\ G'(ℵ), & ℵ \in B-Ƶ \\ F(ℵ) \cup G(ℵ), & ℵ \in Ƶ\cap B \end{cases}$$

Since $H(ℵ)=F'(ℵ) \subseteq F'(ℵ)=T(ℵ)$, for all $ℵ \in Ƶ-B$, $H(ℵ)=G'(ℵ) \subseteq G'(ℵ)=T(ℵ)$, for all $ℵ \in B-Ƶ$, and $F(ℵ) \cap G(ℵ) \subseteq F(ℵ) \cup G(ℵ)$, for all $ℵ \in Ƶ\cap B$, $(F,Ƶ) \overset{*}{\cap_\varepsilon} (G,B) \cong (F,Ƶ) \overset{*}{\cup_\varepsilon} (G,B)$.

16) $(F,Ƶ) \overset{*}{\cap_\varepsilon} (G,B) = (F,Ƶ) \overset{*}{\cup_\varepsilon} (G,B) \Leftrightarrow (F,Ƶ\cap B) = (G,Ƶ\cap B)$.

Proof: Let $(F,Ƶ) \overset{*}{\cap_\varepsilon} (G,B) \cong (F,Ƶ) \overset{*}{\cup_\varepsilon} (G,B)$ and $(F,Ƶ) \overset{*}{\cap_\varepsilon} (G,B) = (H,Ƶ\cup B)$, where for all $ℵ \in Ƶ\cup B$,

$$H(ℵ)=\begin{cases} F'(ℵ), & ℵ \in Ƶ-B \\ G'(ℵ), & ℵ \in B-Ƶ \\ F(ℵ) \cap G(ℵ), & ℵ \in Ƶ\cap B \end{cases}$$

Let $(F,Ƶ) \overset{*}{\cup_\varepsilon} (G,B) = (K,Ƶ\cup B)$, where for all $ℵ \in Ƶ\cup B$,

$$K(ℵ)=\begin{cases} F'(ℵ), & ℵ \in Ƶ-B \\ G'(ℵ), & ℵ \in B-Ƶ \\ F(ℵ) \cup G(ℵ), & ℵ \in B\capƵ \end{cases}$$

Since $(H,Ƶ\cup B)=(K,Ƶ\cup B)$, $F'(ℵ)=F'(ℵ)$, for all $ℵ \in Ƶ-B$, $G'(ℵ)=G'(ℵ)$, for all $ℵ \in B-Ƶ$, and $F(ℵ) \cap G(ℵ)= F(ℵ) \cup G(ℵ)$, for all $ℵ \in Ƶ\cap B$. So, $F(ℵ) = G(ℵ)$ for all $ℵ \in Ƶ \cap B$. Thus, $(F,Ƶ \cap B) = (G,Ƶ \cap B)$.

Conversely, let $(F,Ƶ\cap B)=(G,Ƶ\cap B)$. Hence, $F(ℵ) = G(ℵ)$, for all $ℵ \in Ƶ \cap B$. So, $F(ℵ) \cap G(ℵ)= F(ℵ) \cup G(ℵ)$, for all $ℵ \in Ƶ \cap B$. Moreover, since $F(ℵ)=F'(ℵ)$, for all $ℵ \in Ƶ-B$, and $G'(ℵ)=G'(ℵ)$, and for all $ℵ \in B-Ƶ$, $H(ℵ)=K(ℵ)$, for all $ℵ \in Ƶ\cup B$. Thus, $(H,Ƶ\cup B) = (K,Ƶ\cup B)$ and $(F,Ƶ) \overset{*}{\cap_\varepsilon} (G,B) = (F,Ƶ) \overset{*}{\cup_\varepsilon} (G,B)$.

17) If $(F,Ƶ) \cong (G,Ƶ)$, then $(F,Ƶ) \overset{*}{\cup_\varepsilon} (H,Ƶ) \cong (G,Ƶ) \overset{*}{\cup_\varepsilon} (H,Ƶ)$.

Proof: Let $(F,Ƶ) \cong (G,Ƶ)$. Hence, $F(ℵ) \subseteq G(ℵ)$, for all $ℵ \in Ƶ$. Let $(F,Ƶ) \overset{*}{\cup_\varepsilon} (H,Ƶ)=(W,Ƶ)$. Thus, for all $ℵ \in Ƶ$,

$$W(ℵ)=\begin{cases} F'(ℵ), & ℵ \in Ƶ-Ƶ=\emptyset \\ H'(ℵ), & ℵ \in Ƶ-Ƶ=\emptyset \\ F(ℵ) \cup H(ℵ), & ℵ \in Ƶ\capƵ=Ƶ \end{cases}$$

Let (G, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H,Ƶ) = (L, Ƶ). Thus, for all ℵ∈Ƶ,

$$L(ℵ)=\begin{cases} G'(ℵ), & ℵ∈Ƶ\text{-}Ƶ=\emptyset \\ H'(ℵ), & ℵ∈Ƶ\text{-}Ƶ=\emptyset \\ G(ℵ)\cup H(ℵ), & ℵ∈Ƶ\capƵ=Ƶ \end{cases}$$

Thus, W(ℵ)=F(ℵ)∪H(ℵ) ⊆ G(ℵ)∪H(ℵ)=L(ℵ), for all ℵ∈Ƶ. Hence, (F,Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H,Ƶ) $\widetilde{\subseteq}$ (G,Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H,Ƶ).

18) If (F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H, Ƶ) $\widetilde{\subseteq}$ (G, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H, Ƶ), then (F,Ƶ) $\widetilde{\subseteq}$(G,Ƶ) needs not be true. That is, the converse of Theorem 3.3 (17) is not true.

Proof: Let us give an example to show that the converse of Theorem 3.3 (17) is not true. Let E={e_1,e_2,e_3,e_4, e_5} be the parameter set, Ƶ={e_1,e_3} be a subset of E, and U={h_1,h_2, h_3,h_4, h_5} be universal set, (F,Ƶ), (G,Ƶ) and (H,Ƶ) be soft sets over U such that (F,Ƶ)={(e_1,{h_2, h_5}), (e_3, {h_1,h_2,h_5})}, (G,Ƶ)={(e_1,{h_2}),(e_3, {h_1,h_2})}, (H,Ƶ)={(e_1,U),(e_3,U)}

Let (F,Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H,Ƶ)=(L,Ƶ), thus (L,Ƶ)={(e_1,U),(e_3,U)} and let (G, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H, Ƶ)=(K,Ƶ), thus (K,Ƶ)={(e_1,U),(e_3,U)}. Hence, (F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H, Ƶ) $\widetilde{\subseteq}$ (G, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H, Ƶ) but it is obvious that (F,Ƶ) $\widetilde{\subseteq}$(G,Ƶ) is not satisfied.

19) If (F,Ƶ) $\widetilde{\subseteq}$ (G, B) and (K,Ƶ) $\widetilde{\subseteq}$ (L, B), then (F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (K,Ƶ) $\widetilde{\subseteq}$(G,B) $\overset{*}{\cup}_{\varepsilon}$ (L,B).

Proof: Let (F, Ƶ) $\widetilde{\subseteq}$ (G, B) and (K,Ƶ) $\widetilde{\subseteq}$ (L, B). Thus, Ƶ ⊆ B and for all ℵ∈Ƶ, F(ℵ)⊆ G(ℵ) and K(ℵ)⊆ L(ℵ). Let (F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (K,Ƶ) = (W,Ƶ). Thus, for all ℵ∈Ƶ,

$$W(ℵ)=\begin{cases} F'(ℵ), & ℵ∈Ƶ\text{-}Ƶ=\emptyset \\ K'(ℵ), & ℵ∈Ƶ\text{-}Ƶ=\emptyset \\ F(ℵ)\cup K(ℵ), & ℵ∈Ƶ\capƵ=Ƶ \end{cases}$$

Let (G,B) $\overset{*}{\cup}_{\varepsilon}$ (L, B) = (S,B). Thus, for all ℵ∈B,

$$S(ℵ)=\begin{cases} G'(ℵ), & ℵ∈B\text{-}B=\emptyset \\ L'(ℵ), & ℵ∈B\text{-}B=\emptyset \\ G(ℵ)\cup L(ℵ), & ℵ∈B\cap B=B \end{cases}$$

Hence, for all ℵ ∈Ƶ, W(ℵ)=F(ℵ) ∪K(ℵ) ⊆ G(ℵ) ∪L(ℵ)=S(ℵ) and so (F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (K,Ƶ) $\widetilde{\subseteq}$(G,B) $\overset{*}{\cup}_{\varepsilon}$ (L,B).

20) (F, Ƶ) $\overset{*}{\cap}_{\varepsilon}$ ([(F,Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (G,Ƶ)]= (F, Ƶ) and (F,Ƶ) $\overset{*}{\cup}_{\varepsilon}$ [(F,Ƶ) $\overset{*}{\cap}_{\varepsilon}$ (G,Ƶ)] = (F,Ƶ) (absorbtion laws)

Proof: Let's consider the LHS. Let (F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (G,Ƶ)=(T,Ƶ), where for all ℵ∈Ƶ,

$$T(ℵ)=\begin{cases} F'(ℵ), & ℵ∈Ƶ\text{-}Ƶ=\emptyset \\ G'(ℵ), & ℵ∈Ƶ\text{-}Ƶ=\emptyset \\ F(ℵ)\cap G(ℵ), & ℵ∈Ƶ\capƵ=Ƶ \end{cases}$$

Let (F, Ƶ) $\overset{*}{\cap}_{\varepsilon}$ (T,Ƶ)=(M,Ƶ), where for all ℵ∈Ƶ,

$$M(ℵ)=\begin{cases} F'(ℵ), & ℵ∈Ƶ\text{-}Ƶ=\emptyset \\ T'(ℵ), & ℵ∈Ƶ\text{-}Ƶ=\emptyset \\ F(ℵ)\cap T(ℵ), & ℵ∈Ƶ\capƵ=Ƶ \end{cases}$$

Thus,

$$M(ℵ)=\begin{cases} F'(ℵ), & ℵ∈Ƶ\text{-}Ƶ=\emptyset \\ F(ℵ)\cup[F(ℵ)\cap G(ℵ)] & ℵ∈Ƶ\capƵ=Ƶ \end{cases}$$

Hence,

$$M(\aleph)=\begin{cases}F'(\aleph), & \aleph\in Ƶ-Ƶ=\emptyset\\ F(\aleph), & \aleph\in Ƶ\cap Ƶ=Ƶ\end{cases}$$

Thus, (M,Ƶ)=(F,Ƶ) and so $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(F,Ƶ)\overset{*}{\cap}_{\varepsilon}(G,Ƶ)] = (F,Ƶ)$.

Similarly, $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(F,Ƶ)\overset{*}{\cap}_{\varepsilon}(G,Ƶ)] = (F,Ƶ)$ can be shown. Here, if $\overset{*}{\cap}_{\varepsilon}$ is replaced by the restricted intersection operation or the extended intersection operation, it is evident that Theorem 3.3. (20) holds again in $S_{Ƶ}$ (U), since these operations coincide with each other in the collection $S_{Ƶ}$(U).

Theorem 3.4. $(S_{Ƶ}(U),^{r}, \overset{*}{\cup}_{\varepsilon}, \emptyset_{Ƶ})$ is an MV-algebra.

Proof: Let us show that it satisfies the MV-algebra conditions.

- (MV1) $(S_{Ƶ}(U), \overset{*}{\cup}_{\varepsilon}, \emptyset_{Ƶ})$ is a commutative monoid (Corollary 3.3.1).
- (MV2) $((F,Ƶ)^{r})^{r}=(F,Ƶ)$ (Ali et al., 2011).
- (MV3) $(\emptyset_{Ƶ})^{r}\overset{*}{\cup}_{\varepsilon}(F,Ƶ) = U_{Ƶ}\overset{*}{\cup}_{\varepsilon}(F,Ƶ) = U_{Ƶ} = (\emptyset_{Ƶ})^{r}$.
- (MV4) $[(F,Ƶ)^{r}\overset{*}{\cup}_{\varepsilon}(G,Ƶ)]^{r}\overset{*}{\cup}_{\varepsilon}(G,Ƶ)=((G,Ƶ)^{r}\overset{*}{\cup}_{\varepsilon}(F,Ƶ)^{r})^{r}\overset{*}{\cup}_{\varepsilon}(F,Ƶ)$. Indeed, $[(F,Ƶ)^{r}\overset{*}{\cup}_{\varepsilon}(G,Ƶ)]^{r}\overset{*}{\cup}_{\varepsilon}(G,Ƶ)=[((F,Ƶ)^{r})^{r}\overset{*}{\cap}_{\varepsilon}(G,Ƶ)^{r}]\overset{*}{\cup}_{\varepsilon}(G,Ƶ)=[(F,Ƶ)\overset{*}{\cap}_{\varepsilon}(G,Ƶ)^{r}]\overset{*}{\cup}_{\varepsilon}(G,Ƶ)=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,Ƶ)]\overset{*}{\cap}_{\varepsilon}[(G,Ƶ)^{r}\overset{*}{\cup}_{\varepsilon}(G,Ƶ)]=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,Ƶ)]\overset{*}{\cap}_{\varepsilon}[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(F,Ƶ)^{r}]=(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(G,Ƶ)\overset{*}{\cap}_{\varepsilon}(F,Ƶ)^{r}]=(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(G,Ƶ)^{r}\overset{*}{\cup}_{\varepsilon}(F,Ƶ)]^{r}=[(G,Ƶ)^{r}\overset{*}{\cup}_{\varepsilon}(F,Ƶ)]^{r}\overset{*}{\cup}_{\varepsilon}(F,Ƶ)$

Thus, $(S_{Ƶ}(U),^{r}, \overset{*}{\cup}_{\varepsilon}, \emptyset_{Ƶ})$ is an MV-algebra.

4. Distribution Rules

In this section, the distribution rules of complementary extended union operation over other types of soft set operations are studied, and many algebraic structures are obtained in the collection of soft sets with a fixed parameter set with complementary extended soft set operations and other types of soft set operations.

Theorem 4.1. Let (F,Ƶ), (G,B), (H,Ş) be soft sets over U. The complementary extended union operation has the following distributions over restricted soft set operations:

i) LHS Distributions of the Complementary Extended Union Operation over Restricted Soft Set Operations:

1)If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, then $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(G,B)\cap_{R}(H,Ş)]=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,B)]\cup_{R}[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ş)]$.

Proof: Consider first the LHS. Let $(G,B)\cap_{R}(H,Ş)=(M,B\cap Ş)$, where for all $\aleph\in B\cap Ş$, $M(\aleph) = G(\aleph)\cap H(\aleph)$. Let $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(M, B\cap Ş)=(N, Ƶ\cup(B\cap Ş))$, where for all $\aleph\in Ƶ\cup(B\cap Ş)$,

$$N(\aleph)=\begin{cases}F'(\aleph), & \aleph\in Ƶ-(B\cap Ş)\\ M'(\aleph), & \aleph\in(B\cap Ş)-Ƶ\\ F(\aleph)\cup M(\aleph), & \aleph\in Ƶ\cap(B\cap Ş)\end{cases}$$

Thus,

$$N(\aleph)=\begin{cases}F'(\aleph), & \aleph\in Ƶ-(B\cap Ş)=Ƶ-(B\cap Ş)\\ G'(\aleph)\cup H'(\aleph), & \aleph\in(B\cap Ş)-Ƶ=Ƶ'\cap B\cap Ş\\ F(\aleph)\cup(G(\aleph)\cap H(\aleph)), & \aleph\in Ƶ\cap(B\cap Ş)=Ƶ\cap B\cap Ş\end{cases}$$

Now consider the RHS, that is, $[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,B)]\cap_{R}[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ş)]$. Let $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,B) = (V,Ƶ\cup B)$, where for all $\aleph\in Ƶ\cup B$,

$$V(ℵ)=\begin{cases} F'(ℵ), & ℵ∈Ƶ\text{-}B \\ G'(ℵ), & ℵ∈B\text{-}Ƶ \\ F(ℵ)∪G(ℵ), & ℵ∈Ƶ∩B \end{cases}$$

Let $(F,Ƶ) \overset{*}{∪_ε} (H,Ş)=(W,Ƶ∪Ş)$, where for all $ℵ∈Ƶ∪Ş$,

$$W(ℵ)=\begin{cases} F'(ℵ), & ℵ∈Ƶ\text{-}Ş \\ H'(ℵ), & ℵ∈Ş\text{-}Ƶ \\ F(ℵ)∪H(ℵ), & ℵ∈Ƶ∩Ş \end{cases}$$

Let $(V,Ƶ∪B) ∩_R(W,Ƶ∪Ş)=(T,(Ƶ∪B)∩ (Ƶ ∪ Ş)))$, where for all $ℵ∈Ƶ∪(B∩ Ş)$, $T(ℵ)=V(ℵ)∩W(ℵ)$. Hence,

$$T(ℵ)=\begin{cases} F'(ℵ)∩F'(ℵ), & ℵ∈(Ƶ\text{-}B)∩(Ƶ\text{-}Ş)=Ƶ∩B'∩Ş' \\ F'(ℵ)∩H'(ℵ), & ℵ∈(Ƶ\text{-}B)∩(Ş\text{-}Ƶ)=∅ \\ F'(ℵ)∩(F(ℵ)∪H(ℵ)) & ℵ∈(Ƶ\text{-}B)∩(Ƶ∩Ş)=Ƶ∩B'∩Ş \\ G'(ℵ)∩F'(ℵ), & ℵ∈(B\text{-}Ƶ)∩(Ƶ\text{-}Ş)=∅ \\ G'(ℵ)∩H'(ℵ), & ℵ∈(B\text{-}Ƶ)∩(Ş\text{-}Ƶ)=Ƶ'∩B∩Ş \\ G'(ℵ)∩(F(ℵ)∪H(ℵ)), & ℵ∈(B\text{-}Ƶ)∩(Ƶ∩Ş)=∅ \\ (F(ℵ)∪G(ℵ))∩F'(ℵ), & ℵ∈(Ƶ∩B)∩(Ƶ\text{-}Ş)=Ƶ∩B∩Ş' \\ (F(ℵ)∪G(ℵ))∩H'(ℵ), & ℵ∈(Ƶ∩B)∩(Ş\text{-}Ƶ)=∅ \\ (F(ℵ)∪G(ℵ))∩(F(ℵ)∪H(ℵ)), & ℵ∈(Ƶ∩B)∩(Ƶ∩Ş)=Ƶ∩B∩Ş \end{cases}$$

Thus,

$$T(ℵ)=\begin{cases} F(ℵ), & ℵ∈(Ƶ\text{-}B)∩(Ƶ\text{-}Ş)=Ƶ∩B'∩Ş' \\ F'(ℵ)∩H(ℵ) & ℵ∈(Ƶ\text{-}B)∩(Ƶ∩Ş)=Ƶ∩B'∩Ş \\ G'(ℵ)∩H'(ℵ), & ℵ∈(B\text{-}Ƶ)∩(Ş\text{-}Ƶ)=Ƶ'∩B∩Ş \\ G'(ℵ)∩F'(ℵ) & ℵ∈(Ƶ∩B)∩(Ƶ\text{-}Ş)=Ƶ∩B∩Ş' \\ F(ℵ)∪G(ℵ))∩(F(ℵ)∪H(ℵ)), & ℵ∈(Ƶ∩B)∩(Ƶ∩Ş)=Ƶ∩B∩Ş \end{cases}$$

Here, when considering the $Ƶ\text{-}(B∩Ş)$ in the function N, since $Ƶ\text{-}(B∩Ş)=Ƶ\text{-}(B∩Ş)'$, if an element is in the complement of $(B∩Ş)$, it is either in $B\text{-}Ş$, in $Ş\text{-}B$, or in $(B∪ Ş)'$. Thus, if $ℵ ∈ Ƶ -(B∩Ş)$, then $ℵ ∈ Ƶ ∩B∩Ş'$ or $ℵ ∈ Ƶ ∩B'∩Ş$ or $ℵ ∈ Ƶ ∩B'∩Ş'$. Thus, N=T under the conditions $Ƶ'∩B∩Ş=Ƶ∩B'∩Ş=Ƶ∩B∩Ş=Ƶ∩B∩Ş'=∅$. It is obvious that the condition $Ƶ'∩B∩Ş=Ƶ∩B'∩Ş=∅$ is equivalent to the $(Ƶ∆B)∩Ş=∅$.

2)If $(Ƶ∆B)∩Ş=Ƶ∩B∩Ş'=∅$, then $(F,Ƶ)\overset{*}{∪_ε}[(G,B)∪_R(H,Ş)]=[(F,Ƶ)\overset{*}{∪_ε}(G,B)]∪_R[(F,Ƶ)\overset{*}{∪_ε}(H,Ş)]$.

3)If $(Ƶ∆B)∩Ş=Ƶ∩B∩Ş'=∅$, then $(F,Ƶ)\overset{*}{∪_ε}[(G,B)*_R(H,Ş)]=[(F,Ƶ)\overset{*}{λ_ε}(G,B)∪_R[(F,Ƶ)\overset{*}{λ_ε}(H,Ş)]$.

4)If $(Ƶ∆B)∩Ş=Ƶ∩B∩Ş'=∅$, then $(F,Ƶ)\overset{*}{∪_ε}[(G,B)θ_R(H,Ş)]=[(F,Ƶ)\overset{*}{λ_ε}(G,B)]∩_R[(F,Ƶ)\overset{*}{λ_ε}(H,Ş)]$.

ii) RHS Distribution of Complementary Extended Union Operation over Restricted Soft Set Operations

1)If $(Ƶ∆B)∩Ş=Ƶ∩B∩Ş'=∅$, then $[(F,Ƶ)∪_R(G,B)]\overset{*}{∪_ε}(H,Ş)=[(F,Ƶ)\overset{*}{∪_ε}(H,Ş)]∪_R[(G,B)\overset{*}{∪_ε}(H,Ş)]$.

2)If $(Ƶ∆B)∩Ş=Ƶ∩B∩Ş'=∅$, then $[(F,Ƶ)∩_R(G,B)]\overset{*}{∪_ε}(H,Ş)=[(F,Ƶ)\overset{*}{∪_ε}(H,Ş)]∩_R[(G,B)\overset{*}{∪_ε}(H,Ş)]$.

3)If $(Ƶ∆B)∩ Ş=Ƶ∩B∩Ş'=∅$, then $(F,Ƶ)*_R(G,B)]\overset{*}{∪_ε}(H,Ş)=[(F,Ƶ)\overset{*}{+_ε}(H,Ş)]∪_R[(G,B)\overset{*}{+_ε}(H,Ş)]$.

4)If $(Ƶ∆B)∩Ş=Ƶ∩B∩Ş'=∅$, then $(F,Ƶ)θ_R(G,B)]\overset{*}{∪_ε}(H,Ş)=[(F,Ƶ)\overset{*}{+_ε}(H,Ş)]∩_R[(G,B)\overset{*}{+_ε}(H,Ş)]$.

Note 4.1.1 Considering the distributions in 4.1 and the conditions under which they are satisfied, it is obvious that the following distributions are satisfied without any conditions the set $S_Ƶ(U)$, where Ƶ is a fixed subset of the parameter set E.

- $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(G,Ƶ)\cap_R(H,Ƶ)]=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,Ƶ)]\cap_R[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]$.
- $[(F,Ƶ)\cap_R(G,Ƶ)]\overset{*}{\cup}_{\varepsilon}(H,Ƶ)=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]\cap_R[(G,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]$.
- $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(G,Ƶ)\cup_R(H,Ƶ)]=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,Ƶ)]\cup_R[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]$.
- $[(F,Ƶ)\cup_R(G,Ƶ)]\overset{*}{\cup}_{\varepsilon}(H,Ƶ)=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]\cup_R[(G,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]$.

Theorem 4.1.2. $(S_Ƶ(U),\cup_R,\overset{*}{\cup}_{\varepsilon})$ is a commutative, idempotent semiring without zero but with unity.

Proof: Ali et al. (2011) showed that $(S_Ƶ(U),\cup_R)$ is a commutative, idempotent monoid with identity $\emptyset_Ƶ$, thus forming a bounded semilattice (and therefore, a semigroup). By Corollary 3.3.1, $(S_Ƶ(U),\overset{*}{\cup}_{\varepsilon})$ is a commutative, idempotent monoid with identity $\emptyset_Ƶ$, that is a bounded semilattice (thus, a semigroup From Note 4.1.1, in $S_Ƶ(U)$, $\overset{*}{\cup}_{\varepsilon}$ distributes over $\cup_R$ from both sides. Thus, $(S_Ƶ(U),\cup_R,\overset{*}{\cup}_{\varepsilon})$ is a commutative, idempotent semiring without zero, but with unity.

Theorem 4.1.3 $(S_Ƶ(U),\cap_R,\overset{*}{\cup}_{\varepsilon})$ is a commutative, idempotent hemiring with unity.

Proof: Ali et al. (2011) showed that $(S_Ƶ(U),\cap_R)$ is a commutative idempotent monoid with identity $U_Ƶ$, that is a bounded semi-lattice (hence a semigroup). By Corollary 3.3.1, $(S_Ƶ(U),\overset{*}{\cup}_{\varepsilon})$ is a commutative, idempotent monoid with identity $\emptyset_Ƶ$, that is a bounded semilattice (thus, a semigroup). Also, by Note 4.1.1, in $S_Ƶ(U)$, $\overset{*}{\cup}_{\varepsilon}$ distributes over $\cap_R$ from both sides. Moreover, since $(F,Ƶ)\cap_R U_Ƶ=U_Ƶ\cap_R(F,Ƶ)=(F,Ƶ)$ and by Theorem 3.3 (7) $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}U_Ƶ=U_Ƶ\overset{*}{\cup}_{\varepsilon}(F,Ƶ)=U_Ƶ$, $(S_Ƶ(U),\cap_R,\overset{*}{\cup}_{\varepsilon})$ is a commutative, idempotent hemiring with unity.

Theorem 4.1.4. $(S_Ƶ(U),U_Ƶ,\emptyset_Ƶ,\cap_R,\overset{*}{\cup}_{\varepsilon})$ Boolean Algebra and De Morgan Algebra.

Proof: Ali et al. (2011) showed that $(S_Ƶ(U),\cap_R)$ is an idempotent, commutative monoid with identity $U_Ƶ$, hence a bounded semilattice (thus, a semigroup). By Theorem 3.3.1, $(S_Ƶ(U),\overset{*}{\cup}_{\varepsilon})$ is a commutative, idempotent monoid with identity $\emptyset_Ƶ$, thus a bounded semi-lattice (hence a semigroup). By 3.3. Theorem (21), $\overset{*}{\cap}_{\varepsilon}$ and $\cap_R$ hold the absorbing law. Hence, $(S_Ƶ(U),U_Ƶ,\emptyset_Ƶ,\cap_R,\overset{*}{\cup}_{\varepsilon})$ is a bounded lattice. Moreover, since $(F,Ƶ)\cap_R(F,Ƶ)^r=\emptyset_Ƶ$ and $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(F,Ƶ)^r=U_Ƶ$ for all $(F,Ƶ)\in S_Ƶ(U)$, $(S_Ƶ(U),U_Ƶ,\emptyset_Ƶ,\cap_R,\overset{*}{\cup}_{\varepsilon})$ is a complemented bounded lattice. Furthermore, by 3.4.1.1. Corollary, $\overset{*}{\cup}_{\varepsilon}$ distributes over $\cap_R$ from both sides. Thus, $(S_Ƶ(U),U_Ƶ,\emptyset_Ƶ,\cap_R,\overset{*}{\cup}_{\varepsilon})$ is a distributive, complemented bounded lattice, hence a Boolean Algebra.

Moreover, since $[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,Ƶ)]^r=(F,Ƶ)^r\cap_R(G,Ƶ)^r$ and $[(F,Ƶ)\cap_R(G,Ƶ)]^r=(F,Ƶ)^r\overset{*}{\cup}_{\varepsilon}(G,Ƶ)^r$, that is, De Morgan laws hold, $(S_Ƶ(U),U_Ƶ,\emptyset_Ƶ,^c,\cap_R,\overset{*}{\cap}_{\varepsilon})$ is a De Morgan Algebra.

Here note that since $(S_E(U),\overset{*}{\cup}_{\varepsilon})$ is a non-commutative idempotent semigroup in $S_E(U)$, $\overset{*}{\cup}_{\varepsilon}$ does not form a lattice together with $\cap_R$.

Theorem 4.2. Let (F, Z), (G, B), (H,Ş) be soft sets over U. Then, the following distributions of the complementary extended union operation over extended soft set operations hold:

i) LHS Distributions of the Complementary Extended Union Operation on Extended Soft Set Operations

1)If (Z∆B)∩Ş=Z∩B∩Ş'=∅, then $(F,Z) \overset{*}{\cup}_{\varepsilon}$ [(G,B)$\cap_{\varepsilon}$(H,Ş)]=[$(F,Z) \overset{*}{\cup}_{\varepsilon}$ (G,B)]$\cap_{\varepsilon}$[$(F,Z) \overset{*}{\cup}_{\varepsilon}$ (H,Ş)].

Proof: Consider first the LHS. Let (G,B)$\cap_{\varepsilon}$ (H,Ş)=(M,B∪Ş), where for all ℵ∈B∪Ş,

$$M(ℵ)=\begin{cases} G(ℵ), & ℵ∈B\text{-}Ş \\ H(ℵ), & ℵ∈Ş\text{-}B \\ G(ℵ)∩H(ℵ), & ℵ∈B∩Ş \end{cases}$$

Let $(F,Z) \overset{*}{\cup}_{\varepsilon}$ (M,B∪Ş)=(N,Z∪(B∪Ş)), where for all ℵ∈Z∪B∪Ş,

$$N(ℵ)=\begin{cases} F'(ℵ), & ℵ∈Z\text{-}(B∪Ş) \\ M'(ℵ), & ℵ∈(B∪Ş)\text{-}Z \\ F(ℵ)∪M(ℵ), & ℵ∈Z∩(B∪Ş) \end{cases}$$

Hence,

$$N(ℵ)=\begin{cases} F'(ℵ), & ℵ∈Z\text{-}(B∪Ş)=Z∩B'∩Ş' \\ G'(ℵ), & ℵ∈(B\text{-}Ş)\text{-}Z=Z'∩B∩Ş' \\ H'(ℵ), & ℵ∈(Ş\text{-}B)\text{-}Z=Z'∩B'∩Ş \\ G'(ℵ)∪H'(ℵ), & ℵ∈(B∩Ş)\text{-}Z=Z'∩B∩Ş \\ F(ℵ)∪G(ℵ), & ℵ∈Z∩(B\text{-}Ş)=Z∩B∩Ş' \\ F(ℵ)∪H(ℵ), & ℵ∈Z∩(Ş\text{-}B)=Z∩B'∩Ş \\ F(ℵ)∪(G(ℵ)∩H(ℵ)), & ℵ∈Z∩(B∩Ş)=Z∩B∩Ş \end{cases}$$

Now consider the RHS, i.e., [$(F,Z) \overset{*}{\cup}_{\varepsilon}$ (G,B)]$\cap_{\varepsilon}$[$(F,Z) \overset{*}{\cup}_{\varepsilon}$ (H,Ş)]. Let $(F,Z) \overset{*}{\cup}_{\varepsilon}$ (G,B)=(V,Z∪B), where for all ℵ∈Z∪B,

$$V(ℵ)=\begin{cases} F'(ℵ), & ℵ∈Z\text{-}B \\ G'(ℵ), & ℵ∈B\text{-}Z \\ F(ℵ)∪G(ℵ), & ℵ∈Z∩B \end{cases}$$

Let $(F,Z) \overset{*}{\cup}_{\varepsilon}$ (H,Ş)=(W,Z∪Ş), where for all ℵ∈Z∪Ş,

$$W(ℵ)=\begin{cases} F'(ℵ), & ℵ∈Z\text{-}Ş \\ H'(ℵ), & ℵ∈Ş\text{-}Z \\ F(ℵ)∪H(ℵ), & ℵ∈Z∩Ş \end{cases}$$

Let (V,Z∪B) $\cap_{\varepsilon}$(W,Z∪Ş)=(T,(Z∪B)∪Ş), where for all ℵ∈Z∪B∪Ş,

$$T(ℵ)=\begin{cases} V(ℵ), & ℵ∈(Z∪B)\text{-}(Z∪Ş) \\ W(ℵ), & ℵ∈(Z∪Ş)\text{-}(Z∪B) \\ V(ℵ)∩W(ℵ), & ℵ∈(Z∪B)∩(Z∪Ş) \end{cases}$$

Hence,

$$
T(\aleph)=\begin{cases}
F'(\aleph), & \aleph\in(Ƶ-B)-(Ƶ\cup Ş)=\emptyset \\
G'(\aleph), & \aleph\in(B-Ƶ)-(Ƶ\cup Ş)=Ƶ'\cap B\cap Ş' \\
F(\aleph)\cup G(\aleph), & \aleph\in(Ƶ\cap B)-(Ƶ\cup Ş)=\emptyset \\
F'(\aleph), & \aleph\in(Ƶ-Ş)-(Ƶ\cup B)=\emptyset \\
H'(\aleph), & \aleph\in(Ş-Ƶ)-(Ƶ\cup B)=Ƶ'\cap B'\cap Ş \\
F(\aleph)\cup H(\aleph), & \aleph\in(Ƶ\cap Ş)-(Ƶ\cup B)=\emptyset \\
F'(\aleph)\cap F'(\aleph), & \aleph\in(Ƶ-B)\cap(Ƶ-Ş)=Ƶ\cap B'\cap Ş' \\
F'(\aleph)\cap H'(\aleph), & \aleph\in(Ƶ-B)\cap(Ş-Ƶ)=\emptyset \\
F'(\aleph)\cap(F(\aleph)\cup H(\aleph)), & \aleph\in(Ƶ-B)\cap(Ƶ\cap Ş)=Ƶ\cap B'\cap Ş \\
G'(\aleph)\cap F'(\aleph), & \aleph\in(B-Ƶ)\cap(Ƶ-Ş)=\emptyset \\
G'(\aleph)\cap H'(\aleph), & \aleph\in(B-Ƶ)\cap(Ş-Ƶ)=Ƶ'\cap B\cap Ş \\
G'(\aleph)\cap(F(\aleph)\cup H(\aleph)), & \aleph\in(B-Ƶ)\cap(Ƶ\cap Ş)=\emptyset \\
(F(\aleph)\cup G(\aleph))\cap F'(\aleph), & \aleph\in(Ƶ\cap B)\cap(Ƶ-Ş)=Ƶ\cap B\cap C' \\
(F(\aleph)\cup G(\aleph))\cap H'(\aleph), & \aleph\in(Ƶ\cap B)\cap(Ş-Ƶ)=\emptyset \\
(F(\aleph)\cup G(\aleph))\cap F(\aleph)\cup H(\aleph)), & \aleph\in(Ƶ\cap B)\cap(Ƶ\cap Ş)=Ƶ\cap B\cap Ş
\end{cases}
$$

Thus,

$$
T(\aleph)=\begin{cases}
G'(\aleph), & \aleph\in(B-Ƶ)-(Ƶ\cup Ş)=Ƶ'\cap B\cap Ş' \\
H'(\aleph), & \aleph\in(Ş-Ƶ)-(Ƶ\cup B)=Ƶ'\cap B'\cap Ş \\
F'(\aleph), & \aleph\in(Ƶ-B)\cap(Ƶ-Ş)=Ƶ\cap B'\cap Ş' \\
F'(\aleph)\cap H(\aleph), & \aleph\in(Ƶ-B)\cap(Ƶ\cap Ş)=Ƶ\cap B'\cap Ş \\
G'(\aleph)\cap H'(\aleph), & \aleph\in(B-Ƶ)\cap(Ş-Ƶ)=Ƶ'\cap B\cap Ş \\
G(\aleph)\cap H'(\aleph), & \aleph\in(Ƶ\cap B)\cap(Ƶ-Ş)=Ƶ\cap B\cap Ş' \\
(F(\aleph)\cup G(\aleph))\cap F(\aleph)\cup H(\aleph), & \aleph\in(Ƶ\cap B)\cap(Ƶ\cap Ş)=Ƶ\cap B\cap Ş
\end{cases}
$$

N=T if $Ƶ'\cap B\cap Ş=Ƶ\cap B'\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$. It is obvious that the condition $Ƶ'\cap B\cap Ş=Ƶ\cap B'\cap Ş'=\emptyset$ is equivalent to the condition $(Ƶ\Delta B)\cap Ş=\emptyset$.

2)If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, then $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(G,B)\cup_{\varepsilon}(H,Ş)]=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,B)]\cup_{\varepsilon}[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ş)]$.

3)If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, then $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(G,B)\ast_{\varepsilon}(H,Ş)]=[(F,Ƶ)\overset{*}{\lambda}_{\varepsilon}(G,B)]\cup_{\varepsilon}[(F,Ƶ)\overset{*}{\lambda}_{\varepsilon}(H,Ş)]$.

4)$(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, then $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(G,B)\theta_{\varepsilon}(H,Ş)]=[(F,Ƶ)\overset{*}{\lambda}_{\varepsilon}(G,B)]\cap_{\varepsilon}[(F,Ƶ)\overset{*}{\lambda}_{\varepsilon}(H,Ş)]$.

ii) RHS Distributions of Complementary Extended Union Operation over Extended Soft Set Operations

1)If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, then $([(F,Ƶ)\cup_{\varepsilon}(G,B)]\overset{*}{\cup}_{\varepsilon}(H,Ş)=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ş)]\cup_{\varepsilon}[(G,B)\overset{*}{\cup}_{\varepsilon}(H,Ş)]$.

2)If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, then $[(F,Ƶ)\cap_{\varepsilon}(G,B)]\overset{*}{\cup}_{\varepsilon}(H,Ş)=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ş)]\cap_{\varepsilon}[(G,B)\overset{*}{\cup}_{\varepsilon}(H,Ş)]$.

3)If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, then $[(F,Ƶ)\theta_{\varepsilon}(G,B)]\overset{*}{\cup}_{\varepsilon}(H,Ş)=[(F,Ƶ)\overset{*}{+}_{\varepsilon}(H,Ş)]\cap_{\varepsilon}[(G,B)\overset{*}{+}_{\varepsilon}(H,Ş)]$.

4) If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, then $[(F,Ƶ)\ast_{\varepsilon}(G,B)]\overset{*}{\cup}_{\varepsilon}(H,Ş)=[(F,Ƶ)\overset{*}{+}_{\varepsilon}(H,Ş)]\cup_{\varepsilon}[(G,B)\overset{*}{+}_{\varepsilon}(H,Ş)]$.

Note 4.2.1 Considering the distributions in Theorem 4.2. and the conditions under which they are satisfied, it is obvious that the following distributions are satisfied in the set $S_Ƶ(U)$ without any conditions, where Ƶ is a fixed subset of the parameter set E:

- $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(G,Ƶ)\cap_{\varepsilon}(H,Ƶ)]=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,Ƶ)]\cap_{\varepsilon}[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]$.

- $[(F,Ƶ)\cap_{\varepsilon}(G,Ƶ)]\overset{*}{\cup}_{\varepsilon}(H,Ƶ)=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]\cap_{\varepsilon}[(G,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]$.
- $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(G,Ƶ)\cup_{\varepsilon}(H,Ƶ)]=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,Ƶ)]\cup_{\varepsilon}[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]$.
- $[(F,Ƶ)\cup_{\varepsilon}(G,Ƶ)]\overset{*}{\cup}_{\varepsilon}(H,Ƶ)=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]\cup_{\varepsilon}[(G,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]$.

Theorem 4.2.2. $(S_{Ƶ}(U),\cup_{\varepsilon},\overset{*}{\cup}_{\varepsilon})$ is a commutative, idempotent semiring without zero but with unity.

Theorem 4.2.3 $(S_{Ƶ}(U),\cap_{\varepsilon},\overset{*}{\cup}_{\varepsilon})$ is a commutative, idempotent hemiring with unity.

Theorem 4.2.4. $(S_{Ƶ}(U),U_{Ƶ},\emptyset_{Ƶ},\cap_{\varepsilon},\overset{*}{\cup}_{\varepsilon})$ Bool Algebra and De Morgan Algebra.

Theorem 4.3. Let $(F,Ƶ),(G,B),(H,Ş)$ be soft sets over U. The following distributions of the complementary extended union operation over complementary extended operations hold:

i) LHS Distributions of Complementary Extended Union Operations over Complementary Extended Soft Set Operations

1)If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, then $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(G,B)\overset{*}{\cap}_{\varepsilon}(H,Ş)]=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,B)]\overset{*}{\cap}_{\varepsilon}[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ş)]$.

Proof: Consider first the LHS. Let $(G,B)\overset{*}{\cap}_{\varepsilon}(H,Ş)=(M,B\cup Ş)$, where for all $\aleph\in B\cup Ş$,

$$M(\aleph)=\begin{cases} G'(\aleph), & \aleph\in B-Ş \\ H'(\aleph), & \aleph\in Ş-B \\ G(\aleph)\cap H(\aleph), & \aleph\in B\cap Ş \end{cases}$$

Let $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(M,B\cup Ş)=(N,Ƶ\cup(B\cup Ş))$, where for all $\aleph\in Ƶ\cup B\cup Ş$,

$$N(\aleph)=\begin{cases} F'(\aleph), & \aleph\in Ƶ-(B\cup Ş) \\ M'(\aleph), & \aleph\in (B\cup Ş)-Ƶ \\ F(\aleph)\cup M(\aleph), & \aleph\in Ƶ\cap(B\cup Ş) \end{cases}$$

Thus,

$$N(\aleph)=\begin{cases} F'(\aleph), & \aleph\in Ƶ-(B\cup Ş)=Ƶ\cap B'\cap Ş' \\ G(\aleph), & \aleph\in (B-Ş)-Ƶ=Ƶ'\cap B\cap Ş' \\ H(\aleph), & \aleph\in (Ş-B)-Ƶ=Ƶ'\cap B'\cap Ş \\ G'(\aleph)\cup H'(\aleph), & \aleph\in (B\cap Ş)-Ƶ=Ƶ'\cap B\cap Ş \\ F(\aleph)\cup G(\aleph), & \aleph\in Ƶ\cap(B-Ş)=Ƶ\cap B\cap Ş' \\ F(\aleph)\cup H(\aleph), & \aleph\in Ƶ\cap(Ş-B)=Ƶ\cap B'\cap Ş \\ F(\aleph)\cup(G(\aleph)\cap H(\aleph)), & \aleph\in Ƶ\cap(B\cap Ş)=Ƶ\cap B\cap Ş \end{cases}$$

Now consider the RHS, i.e., $[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,B)]\overset{*}{\cap}_{\varepsilon}[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ş)]$. Let $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,B)=(V,Ƶ\cup B)$, where for all $\aleph\in Ƶ\cup B$,

$$V(\aleph)=\begin{cases} F'(\aleph), & \aleph\in Ƶ-B \\ G'(\aleph), & \aleph\in B-Ƶ \\ F(\aleph)\cup G(\aleph), & \aleph\in Ƶ\cap B \end{cases}$$

Let $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ş)=(W,Ƶ\cup Ş)$, where for all $\aleph\in Ƶ\cup Ş$,

$$W(\aleph)=\begin{cases} F'(\aleph), & \aleph\inƵ-Ş \\ H'(\aleph), & \aleph\inŞ-Ƶ \\ F(\aleph)\cup H(\aleph), & \aleph\inƵ\capŞ \end{cases}$$

Let $(V,Ƶ\cup B)\overset{*}{\cap}_{\varepsilon}(W,Ƶ\cupŞ)=(T,(Ƶ\cup B)\cupŞ)$, where for all $\aleph\inƵ\cup B\cupŞ$,

$$T(\aleph)=\begin{cases} V'(\aleph), & \aleph\in(Ƶ\cup B)-(Ƶ\cupŞ) \\ W'(\aleph), & \aleph\in(Ƶ\cupŞ)-(Ƶ\cup B) \\ V(\aleph)\cap W(\aleph), & \aleph\in(Ƶ\cup B)\cap(Ƶ\cupŞ) \end{cases}$$

Hence,

$$T(\aleph)=\begin{cases} F(\aleph), & \aleph\in(Ƶ-B)-(Ƶ\cupŞ)=\emptyset \\ G(\aleph), & \aleph\in(B-Ƶ)-(Ƶ\cupŞ)=Ƶ'\cap B\capŞ' \\ F'(\aleph)\cap G'(\aleph), & \aleph\in(Ƶ\cap B)-(Ƶ\cupŞ)=\emptyset \\ F(\aleph), & \aleph\in(Ƶ-Ş)-(Ƶ\cup B)=\emptyset \\ H(\aleph), & \aleph\in(Ş-Ƶ)-(Ƶ\cup B)=Ƶ'\cap B'\capŞ \\ F'(\aleph)\cap H'(\aleph), & \aleph\in(Ƶ\capŞ)-(Ƶ\cup B)=\emptyset \\ F'(\aleph)\cap F'(\aleph), & \aleph\in(Ƶ-B)\cap(Ƶ-Ş)=Ƶ\cap B'\capŞ' \\ F'(\aleph)\cap H'(\aleph), & \aleph\in(Ƶ-B)\cap(Ş-Ƶ)=\emptyset \\ F'(\aleph)\cap(F(\aleph)\cup H(\aleph)), & \aleph\in(Ƶ-B)\cap(Ƶ\capŞ)=Ƶ\cap B'\capŞ \\ G'(\aleph)\cap F'(\aleph), & \aleph\in(B-Ƶ)\cap(Ƶ-Ş)=\emptyset \\ G'(\aleph)\cap H'(\aleph), & \aleph\in(B-Ƶ)\cap(Ş-Ƶ)=Ƶ'\cap B\capŞ \\ G'(\aleph)\cap(F(\aleph)\cup H(\aleph)), & \aleph\in(B-Ƶ)\cap(Ƶ\capŞ)=\emptyset \\ (F(\aleph)\cup G(\aleph))\cap F'(\aleph), & \aleph\in(Ƶ\cap B)\cap(Ƶ-Ş)=Ƶ\cap B\capŞ' \\ (F(\aleph)\cup G(\aleph))\cap H'(\aleph), & \aleph\in(Ƶ\cap B)\cap(Ş-Ƶ)=\emptyset \\ (F(\aleph)\cup G(\aleph))\cap(F(\aleph)\cup H(\aleph)), & \aleph\in(Ƶ\cap B)\cap(Ƶ\capŞ)=Ƶ\cap B\capŞ \end{cases}$$

Therefore,

$$T(\aleph)=\begin{cases} G(\aleph), & \aleph\in(B-Ƶ)-(Ƶ\cupŞ)=Ƶ'\cap B\capŞ' \\ H(\aleph), & \aleph\in(Ş-Ƶ)-(Ƶ\cup B)=Ƶ'\cap B'\capŞ \\ F'(\aleph) & \aleph\in(Ƶ-B)\cap(Ƶ-Ş)=Ƶ\cap B'\capŞ' \\ F'(\aleph)\cap H(\aleph) & \aleph\in(Ƶ-B)\cap(Ƶ\capŞ)=Ƶ\cap B'\capŞ \\ G'(\aleph)\cap H'(\aleph), & \aleph\in(B-Ƶ)\cap(Ş-Ƶ)=Ƶ'\cap B\capŞ \\ F'(\aleph)\cap F'(\aleph) & \aleph\in(Ƶ\cap B)\cap(Ƶ-Ş)=Ƶ\cap B\capŞ' \\ (F(\aleph)\cup G(\aleph))\cap(F(\aleph)\cup H(\aleph)), & \aleph\in(Ƶ\cap B)\cap(Ƶ\capŞ)=Ƶ\cap B\capŞ \end{cases}$$

N=T is satisfied under $Ƶ'\cap B\capŞ=Ƶ\cap B'\capŞ=Ƶ\cap B\capŞ'=\emptyset$. It is obvious that the condition $Ƶ'\cap B\capŞ=Ƶ\cap B'\capŞ=\emptyset$ is equivalent to the condition $(Ƶ\Delta B)\capŞ=\emptyset$.

2)If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\capŞ'=\emptyset$, then $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(G,B)\overset{*}{\cup}_{\varepsilon}(H,Ş)]=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,B)]\overset{*}{\cup}_{\varepsilon}[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ş)]$.

3)If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\capŞ'=\emptyset$, then $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(G,B)\overset{*}{*}_{\varepsilon}(H,Ş)]=[(F,Ƶ)\overset{*}{\lambda}_{\varepsilon}(G,B)]\overset{*}{\cup}_{\varepsilon}[(F,Ƶ)\overset{*}{\lambda}_{\varepsilon}(H,Ş)]$.

4) If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\capŞ'=\emptyset$, then $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(G,B)\overset{*}{\theta}_{\varepsilon}(H,Ş)]=[(F,Ƶ)\overset{*}{\lambda}_{\varepsilon}(G,B)]\overset{*}{\cap}_{\varepsilon}[(F,Ƶ)\overset{*}{\lambda}_{\varepsilon}(H,Ş)]$.

ii) RHS Distributions of Complementary Extended Union Operations over Complementary Extended Soft Set Operations

1)If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\capŞ'=\emptyset$, then $[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,B)]\overset{*}{\cup}_{\varepsilon}(H,Ş)=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ş)]\overset{*}{\cup}_{\varepsilon}[(G,B)\overset{*}{\cup}_{\varepsilon}(H,Ş)]$.

2)If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\capŞ'=\emptyset$, then $[(F,Ƶ)\overset{*}{\cap}_{\varepsilon}(G,B)]\overset{*}{\cup}_{\varepsilon}(H,Ş)=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ş)]\overset{*}{\cap}_{\varepsilon}[(G,B)\overset{*}{\cup}_{\varepsilon}(H,Ş)]$.

3)If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, then $[(F,Ƶ)\overset{*}{\theta}_{\varepsilon}(G,B)]\overset{*}{\cup}_{\varepsilon}(H,Ş)=[(F,Ƶ)\overset{*}{\lambda}_{\varepsilon}(H,Ş)]\overset{*}{\cap}_{\varepsilon}[(G,B)\overset{*}{\lambda}_{\varepsilon}(H,Ş)]$.

4)If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, then $[(F,Ƶ)\overset{*}{*}_{\varepsilon}(G,B)]\overset{*}{\cup}_{\varepsilon}(H,Ş)=[(F,Ƶ)\overset{*}{+}_{\varepsilon}(H,Ş)]\overset{*}{\cup}_{\varepsilon}[(G,B)\overset{*}{+}_{\varepsilon}(H,Ş)]$.

Note 4.3.1 If we consider the distributions in Theorem 4.3 and the conditions under which they are satisfied, it is obvious that the following distributions are satisfied without any conditions in the set $S_Ƶ(U)$, where Ƶ is a fixed subset of the parameter set E.

- $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(G,Ƶ)\overset{*}{\cap}_{\varepsilon}(H,Ƶ)]=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,Ƶ)]\overset{*}{\cap}_{\varepsilon}[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]$.
- $[(F,Ƶ)\overset{*}{\cap}_{\varepsilon}(G,Ƶ)]\overset{*}{\cup}_{\varepsilon}(H,Ƶ)=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]\overset{*}{\cap}_{\varepsilon}[(G,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]$.
- $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(G,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,Ƶ)]\overset{*}{\cup}_{\varepsilon}[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]$.
- $[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,Ƶ)]\overset{*}{\cup}_{\varepsilon}(H,Ƶ)=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]\overset{*}{\cup}_{\varepsilon}[(G,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ƶ)]$.

Theorem 4.3.2. $(S_Ƶ(U),\overset{*}{\cup}_{\varepsilon},\overset{*}{\cup}_{\varepsilon})$ is a commutative, idempotent semiring without zero but with unity.

Theorem 4.3.3 $(S_Ƶ(U),\overset{*}{\cap}_{\varepsilon},\overset{*}{\cup}_{\varepsilon})$ is a commutative, idempotent hemiring with unity.

Theorem 4.3.4. $(S_Ƶ(U),U_Ƶ,\emptyset_Ƶ,\overset{*}{\cap}_{\varepsilon},\overset{*}{\cup}_{\varepsilon})$ Bool Algebra and De Morgan Algebra.

Theorem 4.4. Let (F,Ƶ), (G,B), (H,Ş) be soft sets over U. The following distributions of the complementary extended union operation over soft binary piecewise operations hold:

i) LHS Distributions of Complementary Extended Union Operations over Soft Binary Piecewise Soft Set Operations

1)$(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$ if $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}[(G,B)\overset{\sim}{\cap}(H,Ş)]=[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,B)]\overset{\sim}{\cap}[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ş)]$.

Proof: Consider first the LHS. Let $(G,B)\overset{\sim}{\cap}(H,Ş)=(M,B)$, where for all $\aleph\in B$,

$$M(\aleph)=\begin{cases} G(\aleph), & \aleph\in B-Ş \\ G(\aleph)\cap H(\aleph), & \aleph\in B\cap Ş \end{cases}$$

Let $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(M,B)=(N,Ƶ\cup B)$, where for all $\aleph\in Ƶ\cup B$,

$$N(\aleph)=\begin{cases} F'(\aleph), & \aleph\in Ƶ-B \\ M'(\aleph), & \aleph\in B-Ƶ \\ F(\aleph)\cup M(\aleph), & \aleph\in Ƶ\cap B \end{cases}$$

Hence,

$$N(\aleph)=\begin{cases} F'(\aleph), & \aleph\in Ƶ-B \\ G'(\aleph), & \aleph\in(B-Ş)-Ƶ=Ƶ'\cap B\cap Ş' \\ G'(\aleph)\cup H'(\aleph), & \aleph\in(B\cap Ş)-Ƶ=Ƶ'\cap B\cap Ş \\ F(\aleph)\cup G(\aleph), & \aleph\in Ƶ\cap(B-Ş)=Ƶ\cap B\cap Ş' \\ F(\aleph)\cup(G(\aleph)\cap H(\aleph)), & \aleph\in Ƶ\cap(B\cap Ş)=Ƶ\cap B\cap Ş \end{cases}$$

Now consider the RHS, i.e., $[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,B)]\overset{\sim}{\cap}[(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(H,Ş)]$. Let $(F,Ƶ)\overset{*}{\cup}_{\varepsilon}(G,B)=(V,Ƶ\cup B)$, where for all $\aleph\in Ƶ\cup B$,

$$V(\aleph)=\begin{cases} F'(\aleph), & \aleph\in Ƶ-B \\ G'(\aleph), & \aleph\in B-Ƶ \\ F(\aleph)\cup G(\aleph), & \aleph\in Ƶ\cap B \end{cases}$$

Let $(F,Ƶ) \overset{*}{\cup}_{\varepsilon} (H,Ş)=(W,Ƶ\cup Ş)$, where for all $ℵ\in Ƶ\cup Ş$,

$$W(ℵ)=\begin{cases} F'(ℵ), & ℵ\in Ƶ\text{-}Ş \\ H'(ℵ), & ℵ\in Ş\text{-}Ƶ \\ F(ℵ)\cup H(ℵ), & ℵ\in Ƶ\cap Ş \end{cases}$$

Let $(V,Ƶ\cup B) \tilde{\cap} (W,Ƶ\cup Ş)=(T,(Ƶ\cup B))$, where for all $ℵ\in Ƶ\cup B$,

$$T(ℵ)=\begin{cases} V(ℵ), & ℵ\in(Ƶ\cup B)\text{-}(Ƶ\cup Ş) \\ V(ℵ)\cap W(ℵ), & ℵ\in(Ƶ\cup B)\cap(Ƶ\cup Ş) \end{cases}$$

Hence,

$$T(ℵ)=\begin{cases} F'(ℵ), & ℵ\in(Ƶ\text{-}B)\text{-}(Ƶ\cup Ş)=\emptyset \\ G'(ℵ), & ℵ\in(B\text{-}Ƶ)\text{-}(Ƶ\cup Ş)=Ƶ'\cap B\cap Ş' \\ F(ℵ)\cup G(ℵ), & ℵ\in(Ƶ\cap B)\text{-}(Ƶ\cup Ş)=\emptyset \\ F'(ℵ)\cap F'(ℵ), & ℵ\in(Ƶ\text{-}B)\cap(Ƶ\text{-}Ş)=Ƶ\cap B'\cap Ş' \\ F'(ℵ)\cap H'(ℵ), & ℵ\in(Ƶ\text{-}B)\cap(Ş\text{-}Ƶ)=\emptyset \\ F'(ℵ)\cap(F(ℵ)\cup H(ℵ)), & ℵ\in(Ƶ\text{-}B)\cap(Ƶ\cap Ş)=Ƶ\cap B'\cap Ş \\ G'(ℵ)\cap F'(ℵ), & ℵ\in(B\text{-}Ƶ)\cap(Ƶ\text{-}Ş)=\emptyset \\ G'(ℵ)\cap H'(ℵ), & ℵ\in(B\text{-}Ƶ)\cap(Ş\text{-}Ƶ)=Ƶ'\cap B\cap Ş \\ G'(ℵ)\cap(F(ℵ)\cup H(ℵ)), & ℵ\in(B\text{-}Ƶ)\cap(Ƶ\cap Ş)=\emptyset \\ (F(ℵ)\cup G(ℵ))\cap F'(ℵ), & ℵ\in(Ƶ\cap B)\cap(Ƶ\text{-}Ş)=Ƶ\cap B\cap Ş' \\ (F(ℵ)\cup G(ℵ))\cap H'(ℵ), & ℵ\in(Ƶ\cap B)\cap(Ş\text{-}Ƶ)=\emptyset \\ (F(ℵ)\cup G(ℵ))\cap(F(ℵ)\cup H(ℵ)), & ℵ\in(Ƶ\cap B)\cap(Ƶ\cap Ş)=Ƶ\cap B\cap Ş \end{cases}$$

Hence,

$$T(ℵ)=\begin{cases} G'(ℵ), & ℵ\in(B\text{-}Ƶ)\text{-}(Ƶ\cup Ş)=Ƶ'\cap B\cap Ş' \\ F'(ℵ), & ℵ\in(Ƶ\text{-}B)\cap(Ƶ\text{-}Ş)=Ƶ\cap B'\cap Ş' \\ F'(ℵ)\cap H(ℵ), & ℵ\in(Ƶ\text{-}B)\cap(Ƶ\cap Ş)=Ƶ\cap B'\cap Ş \\ G'(ℵ)\cap H'(ℵ), & ℵ\in(B\text{-}Ƶ)\cap(Ş\text{-}Ƶ)=Ƶ'\cap B\cap Ş \\ G(ℵ)\cap H'(ℵ) & ℵ\in(Ƶ\cap B)\cap(Ƶ\text{-}Ş)=Ƶ\cap B\cap Ş' \\ (F(ℵ)\cup G(ℵ))\cap(F(ℵ)\cup H(ℵ)), & ℵ\in(Ƶ\cap B)\cap(Ƶ\cap Ş)=Ƶ\cap B\cap Ş \end{cases}$$

Here, if we consider Ƶ-B in the function N, since $Ƶ\text{-}B=Ƶ\cap B'$, if an element is in the complement of B, then the element is either in Ş-B or in $(B\cup Ş)'$. If $ℵ\in Ƶ\text{-}B$, then $ℵ\in Ƶ\cap B'\cap Ş$ or $ℵ\in Ƶ\cap B'\cap Ş'$. N=T under the $Ƶ'\cap B\cap Ş=Ƶ\cap B'\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$. It is obvious that the condition $Ƶ'\cap B\cap Ş=Ƶ\cap B'\cap Ş=\emptyset$ is equivalent to the condition $(Ƶ\Delta B)\cap Ş$.

2)If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, then $(F,Ƶ) \overset{*}{\cup}_{\varepsilon} [(G,B)\tilde{\cup}(H,Ş)]=[(F,Ƶ) \overset{*}{\cup}_{\varepsilon} (G,B)]\tilde{\cup} [(F,Ƶ) \overset{*}{\cup}_{\varepsilon} (H,Ş)]$.

3)If $(Ƶ\Delta B)\cap C=Ƶ\cap B\cap Ş'=\emptyset$, then $(F,Ƶ) \overset{*}{\cup}_{\varepsilon} [(G,B \overset{\sim}{*} (H,Ş)]=[(F,Ƶ) \overset{*}{\lambda}_{\varepsilon} (G,B)] \tilde{\cup} [(F,Ƶ) \overset{*}{\lambda}_{\varepsilon} (H,Ş)]$.

4) If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, then $(F,Ƶ) \overset{*}{\cap}_{\varepsilon} [(G,B)\tilde{\theta}(H,Ş)]=[(F,Ƶ) \overset{*}{\lambda}_{\varepsilon} (G,B)]\tilde{\cap} [(F,Ƶ) \overset{*}{\lambda}_{\varepsilon} (H,Ş)]$.

ii) RHS Distributions of Complementary Extended Union Operations over Soft Binary Piecewise Soft Set Operations

1) If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, $(F,Ƶ)\tilde{\cup} (G,B)] \overset{*}{\cup}_{\varepsilon} (H,Ş)=[(F,Ƶ)\overset{*}{\cap}_{\varepsilon} (H,Ş)]\tilde{\cup}[(G,B) \overset{*}{\cup}_{\varepsilon} (H,Ş)]$.

2)If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, then $(F,Ƶ)\tilde{\cap} (G,B)] \overset{*}{\cup}_{\varepsilon} (H,Ş) =[(F,Ƶ) \overset{*}{\cup}_{\varepsilon} (H,Ş)]\tilde{\cap}[(G,B) \overset{*}{\cup}_{\varepsilon} (H,Ş)]$.

3)If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, then $[(F,Ƶ)\tilde{\theta} (G,B)] \overset{*}{\cup}_{\varepsilon}(H,Ş)=[(F,Ƶ)\overset{*}{+}_{\varepsilon} (H,Ş)]\tilde{\cap}[(G,B)\overset{*}{+}_{\varepsilon} (H,Ş)]$.

4)If (ƵΔB)∩Ş=Ƶ∩B∩Ş'=∅, then [(F, Ƶ) $\underset{*}{\sim}$ (G,B)] $\overset{*}{\cap}_{\varepsilon}$ (H,Ş)=[(F, Ƶ) $\overset{*}{+}_{\varepsilon}$ (H,Ş)] $\overset{\sim}{\cup}$ [(G,B) $\overset{*}{+}_{\varepsilon}$ (H,Ş)].

Note 4.4.1 If we consider the distributions in Theorem 4.4 and the conditions under which they are satisfied, it is obvious that the following distributions are satisfied in $S_{Ƶ}(U)$ without any conditions, where Ƶ is a fixed subset of the parameter set E.

- (F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ [(G,Ƶ) $\overset{\sim}{\cap}$ (H,Ƶ)]=[(F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (G,Ƶ)] $\overset{\sim}{\cap}$ [(F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H,Ƶ)].
- [(F,Ƶ) $\overset{\sim}{\cap}$ (G,Ƶ)] $\overset{*}{\cap}_{\varepsilon}$ (H,Ƶ)=[(F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H,Ƶ)] $\overset{\sim}{\cap}$ [(G, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H,Ƶ)].
- (F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ [(G,Ƶ) $\overset{\sim}{\cup}$ (H,Ƶ)]=[(F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (G,Ƶ)] $\overset{\sim}{\cup}$ [(F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H,Ƶ)].
- [(F,Ƶ) $\overset{\sim}{\cup}$ (G,Ƶ)] $\overset{*}{\cup}_{\varepsilon}$ (H,Ƶ)=[(F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H,Ƶ)] $\overset{\sim}{\cup}$ [(G, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H,Ƶ)].

Theorem 4.4.2. ($S_{Ƶ}$(U), $\overset{\sim}{\cup}$, $\overset{*}{\cup}_{\varepsilon}$) is a commutative, idempotent semiring without zero but with unity.

Theorem 4.4.3 ($S_{Ƶ}$(U), $\overset{\sim}{\cap}$, $\overset{*}{\cup}_{\varepsilon}$) is a commutative, idempotent hemiring with unity.

Theorem 4.4.4. ($S_{Ƶ}$(U), $U_{Ƶ}$, $\emptyset_{Ƶ}$, $\overset{\sim}{\cap}$, $\overset{*}{\cup}_{\varepsilon}$) Bool Algebra and De Morgan Algebra.

Theorem 4.5. Let (F, Ƶ), (G, B), (H,Ş) be soft sets over U. The following distributions of the complementary extended union operation over complementary soft binary piecewise operations hold:

i) LHS Distributions of Complementary Extended Union Operations over Complementary Soft Binary Piecewise Soft Set Operations

1)If (ƵΔB)∩Ş=Ƶ∩B∩Ş'=∅, then (F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ [(G,B) $\underset{\cap}{\overset{*}{\sim}}$ (H,Ş)]=[(F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (G,B)] $\underset{\cap}{\overset{*}{\sim}}$ [(F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H,Ş)].

Proof: Consider first the LHS. Let (G,B) $\underset{\cap}{\overset{*}{\sim}}$ (H,Ş)=(M,B). Hence, for all ℵ∈B,

$$M(ℵ)=\begin{cases} G'(ℵ), & ℵ\in B-Ş \\ G(ℵ)\cap H(ℵ), & ℵ\in B\cap Ş \end{cases}$$

Let (F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (M,B)=(N,Ƶ∪B), where for all ℵ∈Ƶ∪B,

$$N(ℵ)=\begin{cases} F'(ℵ), & ℵ\in Ƶ-B \\ M'(ℵ), & ℵ\in B-Ƶ \\ F(ℵ)\cup M(ℵ), & ℵ\in Ƶ\cap B \end{cases}$$

Hence,

$$N(ℵ)=\begin{cases} F'(ℵ), & ℵ\in Ƶ-B \\ G(ℵ), & ℵ\in (B-Ş)-Ƶ=Ƶ'\cap B\cap Ş' \\ G'(ℵ)\cup H'(ℵ), & ℵ\in (B\cap Ş)-Ƶ=Ƶ'\cap B\cap Ş \\ F(ℵ)\cup G'(ℵ), & ℵ\in Ƶ\cap(B-Ş)=Ƶ\cap B\cap Ş' \\ F(ℵ)\cup(G(ℵ)\cap H(ℵ)), & ℵ\in Ƶ\cap(B\cap Ş)=Ƶ\cap B\cap Ş \end{cases}$$

Now consider the RHS, i.e., [(F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (G,B)] $\underset{\cap}{\overset{*}{\sim}}$ [(F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H,Ş)]. Let (F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (G,B)=(V,Ƶ∪B), where for all ℵ∈Ƶ∪B,

$$V(ℵ)=\begin{cases} F'(ℵ), & ℵ\in Ƶ-B \\ G'(ℵ), & ℵ\in B-Ƶ \\ F(ℵ)\cup G(ℵ), & ℵ\in Ƶ\cap B \end{cases}$$

Let $(F,Ƶ) \overset{*}{\cup}_{\varepsilon} (H,Ş)=(W,Ƶ\cup Ş)$, where for all $\aleph\in Ƶ\cup Ş$,

$$W(\aleph)=\begin{cases} F'(\aleph), & \aleph\in Ƶ-Ş \\ H'(\aleph), & \aleph\in Ş-Ƶ \\ F(\aleph)\cup H(\aleph), & \aleph\in Ƶ\cap Ş \end{cases}$$

Let $(V,Ƶ\cup B) \overset{*}{\underset{\cap}{\sim}} (W,Ƶ\cup Ş)=(T,(Ƶ\cup B))$, where for all $\aleph\in Ƶ\cup B$,

$$T(\aleph)=\begin{cases} V'(\aleph), & \aleph\in (Ƶ\cup B)-(Ƶ\cup Ş) \\ V(\aleph)\cap W(\aleph), & \aleph\in (Ƶ\cup B)\cap(Ƶ\cup Ş) \end{cases}$$

Hence,

$$T(\aleph)=\begin{cases} F(\aleph), & \aleph\in(Ƶ\text{-}B)-(Ƶ\cup Ş)=\emptyset \\ G(\aleph), & \aleph\in(B\text{-}Ƶ)-(Ƶ\cup Ş)=Ƶ'\cap B\cap Ş' \\ F'(\aleph)\cap G'(\aleph), & \aleph\in(Ƶ\cap B)-(Ƶ\cup Ş)=\emptyset \\ F'(\aleph)\cap F'(\aleph), & \aleph\in(Ƶ\text{-}B)\cap(Ƶ\text{-}Ş)=Ƶ\cap B'\cap Ş' \\ F'(\aleph)\cap H'(\aleph), & \aleph\in(Ƶ\text{-}B)\cap(Ş\text{-}Ƶ)=\emptyset \\ F'(\aleph)\cap(F(\aleph)\cup H(\aleph)), & \aleph\in(Ƶ\text{-}B)\cap(Ƶ\cap Ş)=Ƶ\cap B'\cap Ş \\ G'(\aleph)\cap F'(\aleph), & \aleph\in(B\text{-}Ƶ)\cap(Ƶ\text{-}Ş)=\emptyset \\ G'(\aleph)\cap H'(\aleph), & \aleph\in(B\text{-}Ƶ)\cap(Ş\text{-}Ƶ)=Ƶ'\cap B\cap Ş \\ G'(\aleph)\cap(F(\aleph)\cup H(\aleph)), & \aleph\in(B\text{-}Ƶ)\cap(Ƶ\cap Ş)=\emptyset \\ (F(\aleph)\cup G(\aleph))\cap F'(\aleph), & \aleph\in(Ƶ\cap B)\cap(Ƶ\text{-}Ş)=Ƶ\cap B\cap Ş' \\ (F(\aleph)\cup G(\aleph))\cap H'(\aleph), & \aleph\in(Ƶ\cap B)\cap(Ş\text{-}Ƶ)=\emptyset \\ ((F(\aleph)\cup G(\aleph))\cap(F(\aleph)\cup H(\aleph)), & \aleph\in(Ƶ\cap B)\cap(Ƶ\cap Ş)=Ƶ\cap B\cap Ş \end{cases}$$

Hence,

$$T(\aleph)=\begin{cases} G(\aleph), & \aleph\in(B\text{-}Ƶ)-(Ƶ\cup Ş)=Ƶ'\cap B\cap Ş' \\ F'(\aleph), & \aleph\in(Ƶ\text{-}B)\cap(Ƶ\text{-}Ş)=Ƶ\cap B'\cap Ş' \\ F'(\aleph)\cap H'(\aleph), & \aleph\in(Ƶ\text{-}B)\cap(Ƶ\cap Ş)=Ƶ\cap B'\cap Ş \\ G'(\aleph)\cap H'(\aleph), & \aleph\in(B\text{-}Ƶ)\cap(Ş\text{-}Ƶ)=Ƶ'\cap B\cap Ş \\ G(\aleph)\cap F'(\aleph), & \aleph\in(Ƶ\cap B)\cap(Ƶ\text{-}Ş)=Ƶ\cap B\cap Ş' \\ ((F(\aleph)\cup G(\aleph))\cap(F(\aleph)\cup H(\aleph)), & \aleph\in(Ƶ\cap B)\cap(Ƶ\cap Ş)=Ƶ\cap B\cap Ş \end{cases}$$

Here, if we consider Ƶ-B in the function N, since $Ƶ\text{-}B=Ƶ\cap B'$, then if an element is in the complement of B, that element is either in Ş-B or in $(B\cup Ş)'$. From here, if $\aleph\in Ş\text{-}B$, then $\aleph\in Ƶ\cap B'\cap Ş$ or $\aleph\in Ƶ\cap B'\cap Ş'$, hence we see that N=T with the condition $Ƶ'\cap B\cap Ş=Ƶ\cap B'\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$. It is obvious that the condition $Ƶ'\cap B\cap Ş=Ƶ\cap B'\cap Ş=\emptyset$ is equivalent to the condition $(Ƶ\Delta B)\cap Ş$.

2)If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, then $(F,Ƶ) \overset{*}{\cup}_{\varepsilon} [(G,B) \overset{*}{\underset{\cup}{\sim}} (H,Ş)]=[(F,Ƶ) \overset{*}{\cup}_{\varepsilon} (G,B)] \overset{*}{\underset{\cup}{\sim}} [(F,Ƶ) \overset{*}{\cup}_{\varepsilon} (H,Ş)]$.

3) If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, then $(F,Ƶ) \overset{*}{\cup}_{\varepsilon} [(G,B) \overset{*}{\underset{*}{\sim}} (H,Ş)]=[(F,Ƶ) \overset{*}{\lambda}_{\varepsilon} (G,B)] \overset{*}{\underset{\cup}{\sim}} [(F,Ƶ) \overset{*}{\lambda}_{\varepsilon} (H,Ş)]$.

4) If $(Ƶ\Delta B)\cap Ş= Ƶ\cap B\cap Ş'=\emptyset$, then $(F,Ƶ) \overset{*}{\cup}_{\varepsilon} [(G,B) \overset{*}{\underset{\theta}{\sim}} (H,Ş)]=[(F,Ƶ) \overset{*}{\lambda}_{\varepsilon} (G,B)] \overset{*}{\underset{\cap}{\sim}} [(F,Ƶ) \overset{*}{\lambda}_{\varepsilon} (H,Ş)]$.

ii) RHS Distributions of Complementary Extended Union Operations over Complementary Soft Binary Piecewise Soft Set Operations

1)$(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$ if $(F,Ƶ) \overset{*}{\underset{\cup}{\sim}} (G,B)] \overset{*}{\cup}_{\varepsilon} (H,Ş)=[(F,Ƶ) \overset{*}{\cup}_{\varepsilon} (H,Ş)] \overset{*}{\underset{\cup}{\sim}} [(G,B) \overset{*}{\cup}_{\varepsilon} (H,Ş)]$.

2) If $(Ƶ\Delta B)\cap Ş=Ƶ\cap B\cap Ş'=\emptyset$, then $(F,Ƶ) \overset{*}{\underset{\cap}{\sim}} (G,B)] \overset{*}{\cup}_{\varepsilon} (H,Ş) =[(F,Ƶ) \overset{*}{\cup}_{\varepsilon} (H,Ş)] \overset{*}{\underset{\cap}{\sim}} [(G,B) \overset{*}{\cup}_{\varepsilon} (H,Ş)]$.

3) If (ƵΔB)∩Ş=Ƶ∩B∩Ş'=∅, then (F, Ƶ) $\overset{*}{\underset{\theta}{\sim}}$ (G,B)] $\overset{*}{\cup}_{\varepsilon}$ (H,Ş) =[(F, Ƶ) $\overset{*}{+}_{\varepsilon}$ (H,Ş)] $\overset{*}{\underset{\cap}{\sim}}$ [(G,B) $\overset{*}{+}_{\varepsilon}$ (H,Ş)].

4) If (ƵΔB)∩Ş=Ƶ∩B∩Ş'=∅, then [(F, Ƶ) $\overset{*}{\underset{*}{\sim}}$ (G,B)] $\overset{*}{\cup}_{\varepsilon}$ (H,Ş)=[(F, Ƶ) $\overset{*}{+}_{\varepsilon}$ (H,Ş)] $\overset{*}{\underset{\cup}{\sim}}$ [(G,B) $\overset{*}{+}_{\varepsilon}$ (H,Ş)].

Note 4.5.1. If we consider the distributions in Theorem 4.5 and the conditions under which they are satisfied, it is obvious that the following distributions are satisfied in the set $S_Ƶ(U)$ without any conditions, where Ƶ is a fixed subset of E.

- (F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ [(G,Ƶ) $\overset{*}{\underset{\cap}{\sim}}$ (H,Ƶ)]=[(F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (G,Ƶ)] $\overset{*}{\underset{\cap}{\sim}}$ [(F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H,Ƶ)].
- [(F,Ƶ) $\overset{*}{\underset{\cap}{\sim}}$ (G,Ƶ)] $\overset{*}{\cup}_{\varepsilon}$ (H,Ƶ)=[(F, Ƶ) $^{*}\underset{\cap_{\varepsilon}}{\overset{*}{\cup_{\varepsilon}}}$ (H,Ƶ)] $\overset{*}{\underset{\cap}{\sim}}$ [(G, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H,Ƶ)].
- (F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ [(G,Ƶ) $\overset{*}{\underset{\cup}{\sim}}$ (H,Ƶ)]=[(F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (G,Ƶ)] $\overset{*}{\underset{\cup}{\sim}}$ [(F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H,Ƶ)].
- [(F,Ƶ) $\overset{*}{\underset{\cup}{\sim}}$ (G,Ƶ)] $\overset{*}{\cup}_{\varepsilon}$ (H,Ƶ)=[(F, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H,Ƶ)] $\overset{*}{\underset{\cup}{\sim}}$ [(G, Ƶ) $\overset{*}{\cup}_{\varepsilon}$ (H,Ƶ)].

Theorem 4.5.2. ($S_Ƶ$(U), $\overset{*}{\underset{\cup}{\sim}}$, $\overset{*}{\cup}_{\varepsilon}$) is a commutative, idempotent semiring without zero but with unity.

Theorem 4.5.3 ($S_Ƶ$(U), $\overset{*}{\underset{\cap}{\sim}}$, $\overset{*}{\cup}_{\varepsilon}$) is a commutative, idempotent hemiring with unity.

Theorem 4.5.4. ($S_Ƶ$(U),$U_Ƶ$,$\emptyset_Ƶ$, $\overset{*}{\underset{\cap}{\sim}}$, $\overset{*}{\cup}_{\varepsilon}$) Bool Algebra and De Morgan Algebra.

5. Discussion

In this paper, we introduced the complementary extended union operation, and showed that the collection of all soft sets with a fixed parameter set together with the complementary extended union operation and also with other certain types of soft set operations form many important algebraic structures such as semiring, hemiring, Boolean Algebra, De Morgan Algebra. Let $S_Ƶ$(U) be the collection of all soft sets over U with the fixed parameter set Ƶ, where Ƶ ⊆ E. Then,

- ($S_Ƶ$(U), $\overset{*}{\cup}_{\varepsilon}$) is a commutative, idempotent monoid, that is, a bounded semilattice, whose identity element is $\emptyset_Ƶ$.
- (S_E(U), $\overset{*}{\cup}_{\varepsilon}$) is a groupoid.
- ($S_Ƶ$(U),r, $\overset{*}{\cup}_{\varepsilon}$, $\emptyset_Ƶ$) is an MV-algebra.
- ($S_Ƶ$(U),$\cup_R$, $\overset{*}{\cup}_{\varepsilon}$), ($S_Ƶ$(U),$\cup_\varepsilon$, $\overset{*}{\cup}_{\varepsilon}$), ($S_Ƶ$(U), $\overset{*}{\cup}_{\varepsilon}$, $\overset{*}{\cup}_{\varepsilon}$), ($S_Ƶ$(U), $\overset{\sim}{\cup}$, $\overset{*}{\cup}_{\varepsilon}$), ($S_Ƶ$(U), $\overset{*}{\underset{\cup}{\sim}}$, $\overset{*}{\cup}_{\varepsilon}$) are commutative, idempotent semirings without zero but with unity.

- $(S_Z(U), \cap_R, \overset{*}{\cup}_\varepsilon)$, $(S_Z(U), \cap_\varepsilon, \overset{*}{\cup}_\varepsilon)$, $(S_Z(U), \overset{*}{\cap}_\varepsilon, \overset{*}{\cup}_\varepsilon)$, $(S_Z(U), \tilde{\cap}, \overset{*}{\cup}_\varepsilon)$, $(S_Z(U), \overset{*}{\underset{\cap}{\sim}}, \overset{*}{\cup}_\varepsilon)$ are commutative, idempotent hemirings with unity.
- $(S_Z(U), U_Z, \emptyset_Z, \cap_R, \overset{*}{\cup}_\varepsilon)$, $(S_Z(U), U_Z, \emptyset_Z, \cap_\varepsilon\, \overset{*}{\cup}_\varepsilon)$, $(S_Z(U), U_Z, \emptyset_Z, \overset{*}{\cap}_\varepsilon, \overset{*}{\cup}_\varepsilon)$, $(S_Z(U), U_Z, \emptyset_Z, \tilde{\cap}, \overset{*}{\cup}_\varepsilon)$, $(S_Z(U), U_Z, \emptyset_Z, \overset{*}{\underset{\cap}{\sim}}, \overset{*}{\cup}_\varepsilon)$ are Boolean Algebras and De Morgan Algebras.

6. Conclusion

Soft set operations are the foundational elements of soft set theory, crucial for its advancement in both theoretical and practical realms. Since its inception, numerous restricted and extended operations have been introduced for soft sets. However, this study introduces and explores the algebraic properties of a new soft set operation, which we call “the complementary extended union operation”, specifically comparing it to the union operation in classical set theory. We examine the distribution of the complementary extended union operation over various other soft set operations. By considering the distribution rules and algebraic properties of these operations, we provide an in-depth analysis of the algebraic structures formed by soft sets using this new operation. We demonstrate that the set of all soft sets with a fixed parameter set, along with the complementary extended union operation and other specific soft set operations, form many significant algebraic structures, including semirings, hemirings, Boolean algebras, and De Morgan algebras. As the concepts related to soft set operations are as vital to soft sets as basic operations are to classical set theory, examining their algebraic structures in relation to new soft set operations enhances our understanding of their applications and introduces new examples of algebraic structures. We believe this work contributes to the literature on both classical algebra and soft set theory. Future studies may explore different types of complementary extended soft set operations, along with their distributions and properties, to further identify the algebraic structures formed within the collection of soft sets with a fixed parameter set. Additionally, we think that this study will inspire researchers to propose new encryption methods based on soft sets, and thus it will serve as a foundation for various applications, particularly in decision-making and cryptography as soft sets are a powerful mathematical tool for detecting uncertain objects. Furthermore, the algebraic properties of soft algebraic structures can be re-examined and further developed in the sense of the soft set operation defined in this paper.

Acknowledgements

This paper is derived from the second author’s Master Thesis supervised by the first author at Amasya University, Türkiye.

Conflict of Interest

The authors stated that there are no conflicts of interest regarding the publication of this article.

Publication Ethic

This manuscript has not been previously published by or has been under review by another print or online journal or source.

References

Abbas, M., Ali, B., & Romaguera, S. (2014). On generalized soft equality and soft lattice structure. *Filomat*, *28*(6), 1191-1203.

Abbas, M., Ali, M. I., & Romaguera, S. (2017). Generalized operations in soft set theory via relaxed conditions on parameters. *Filomat*, *31*(19), 5955-5964.

Akbulut, E. (2024). *New type of extended operations of soft sets: Complementary extended lambda and difference operation* (Master’s thesis). Amasya University, Graduate School of Natural and Applied Sciences, Amasya, Türkiye.

Ali, M. I., Feng, F., Liu, X., Min, W. K., & Shabir, M. (2009). On some new operations in soft set theory. *Computers and Mathematics with Applications*, *57*(9), 1547-1553.

Ali, M. I., Mahmood, T., Rehman, M. M., & Aslam, M. F. (2015). On lattice ordered soft sets. *Applied Soft Computing*, *36*, 499-505.

Ali, M. I., Shabir, M., & Naz, M. (2011). Algebraic structures of soft sets associated with new operations. *Computers and Mathematics with Applications*, *61*(9), 2647-2654.

Al-shami, T. M. (2019). Investigation and corrigendum to some results related to *g*-soft equality and *gf*-soft equality relations. *Filomat*, *33*(11), 3375-3383.

Al-shami, T. M., & El-Shafei, M. E. (2020). *T*-soft equality relation. *Turkish Journal of Mathematics*, *44*(4), 1427-1441.

Atagün, A. O., & Sezgin, A. (2018). Soft subnear-rings, soft ideals and soft N-subgroups of near-rings. *Mathematical Sciences Letters*, *7*(1), 37-42.

Aybek, F. (2024). *New restricted and extended soft set operations* (Master's thesis). Amasya University, Graduate School of Natural and Applied Sciences, Amasya, Türkiye.

Birkhoff, G. (1967). *Lattice theory*. Providence, Rhode: American Mathematical Society.

Çağman, N. (2021). Conditional complements of sets and their application to group theory. *Journal of New Results in Science*, *10*(3), 67-74.

Çağman, N., Çıtak, F., & Aktaş, H. (2012). Soft int-group and its applications to group theory. *Neural Computing and Applications*, *21*, 151-158.

Chang, C. C. (1959). A new proof of the completeness of the łukasiewicz axioms. *Transactions of the American Mathematical Society*, *93*(1), 460-490.

Clifford, A. H. (1954). Bands of semigroups. *Proceedings of the American Mathematical Society*, *5*(3), 499-504.

Demirci, A. M. (2024). *New type of extended operations of soft sets: Complementary extended union, plus and theta operations* (Master's thesis). Amasya University, Graduate School of Natural and Applied Sciences, Amasya, Türkiye.

Eren, Ö. F., & Çalışıcı, H. (2019). On some operations of soft sets. *The Fourth International Conference on Computational Mathematics and Engineering Sciences (CMES 2019)*, Antalya.

Feng, F., Li, C., Davvaz, B., & Ali, M. I. (2010). Soft sets combined with fuzzy sets and rough sets: A tentative approach. *Soft Computing*, *14*, 899-911.

Feng, F., & Li, Y. (2013). Soft subsets and soft product operations. *Information Sciences*, *232*, 44-57.

Fu, L. (2011). Notes on soft set operations. *ARPN Journal of Systems and Software*, *1*(6), 205-208.

Ge, X., & Yang, S. (2011). Investigations on some operations of soft sets. *World Academy of Science, Engineering and Technology, International Journal of Mathematical, Computational, Physical, Electrical and Computer Engineering*, *5*(3), 370-373.

Husain, S., & Shivani, K. (2018). A study of properties of soft set and its applications. *International Research Journal of Engineering and Technology*, *5*(1), 363-372.

Iftikhar, M., & Mahmood, T. (2018). Some results on lattice ordered double framed soft semirings. *International Journal of Algebra and Statistics*, *7*(1-2), 123-140.

Jana, C., Pal, M., Karaaslan, F., & Sezgin, A. (2019). (α, β)-Soft intersectional rings and ideals with their applications. *New Mathematics and Natural Computation*, *15*(2), 333-350.

Jiang, Y., Tang, Y., Chen, Q., Wang, J., & Tang, S. (2010). Extending soft sets with description logics. *Computers and Mathematics with Applications*, *59*(6), 2087-2096.

Jun, Y. B., & Yang, X. (2011). A note on the paper "Combination of interval-valued fuzzy set and soft set" [Comput. Math. Appl. 58 (2009) 521-527]. *Computers and Mathematics with Applications*, *61*(5), 1468-1470.

Liu, X., Feng, F., & Jun, Y. B. (2012). A note on generalized soft equal relations. *Computers and Mathematics with Applications*, *64*(4), 572-578.

Mahmood, T. (2020). A novel approach toward bipolar soft sets and their applications. *Journal of Mathematics*, *2020*, 1-11.

Mahmood, T., Rehman, Z. U., & Sezgin, A. (2018). Lattice ordered soft near rings. *Korean Journal of Mathemtics*, *26*(3), 503-517.

Maji, P. K, Biswas, R., & Roy, A. R. (2003). Soft set theory. *Computers and Mathematics with Applications*, *45*(4-5), 555-562.

Molodtsov, D. (1999). Soft set theory-First results. *Computers and Mathematics with Applications*, *37*(4-5), 19-31.

Muştuoğlu, E., Sezgin, A., & Türk, Z. K. (2016). Some characterizations on soft uni-groups and normal soft uni-groups. *International Journal of Computer Applications*, *155*(10), 1-8.

Neog, T. J., & Sut, D. K. (2011). A new approach to the theory of soft sets. *International Journal of Computer Applications*, *32*(2), 1-6.

Onyeozili, I. A., & Gwary, T. M. (2014). A study of the fundamentals of soft set theory. *International Journal of Scientific & Technology Research*, *3*(4), 132-143.

Özlü, Ş., & Sezgin, A. (2020). Soft covered ideals in semigroups. *Acta Universitatis Sapientiae, Mathematica*, *12*(2), 317-346.

Pant, S., Dagtoros, K., Kholil, M. I., & Vivas, A. (2024). Matrices: Peculiar determinant property. *Optimum Science Journal*, *1*, 1-7.

Park, J. H., Kim, O. H., & Kwun, Y. C. (2012). Some properties of equivalence soft set relations. *Computers and Mathematics with Applications*, *63*(6), 1079-1088.

Pei, D., & Miao, D. (2005). From soft sets to information systems. *IEEE International Conference on Granular Computing*, *2*, 617-621.

Sarıalioğlu, M. (2024). *New type of extended operations of soft sets: Complementary extended intersection, gamma and star operations* (Master's thesis). Amasya University, Graduate School of Natural and Applied Sciences, Amasya, Türkiye.

Sen, J. (2014). On algebraic structure of soft sets. *Annals of Fuzzy Mathematics and Informatics*, *7*(6), 1013-1020.

Sezer, A. S. (2014). Certain characterizations of LA-semigroups by soft sets. *Journal of Intelligent and Fuzzy Systems*, *27*(2), 1035-1046.

Sezer, A. S., Çağman, N., & Atagün, A. O. (2015). Uni-soft substructures of groups. *Annals of Fuzzy Mathematics and Informatics*, *9*(2), 235-246.

Sezgin, A. (2016). A new approach to semigroup theory I: Soft union semigroups, ideals and bi-ideals. *Algebra Letters*, *2016*, 1-46.

Sezgin, A. (2018). A new view on AG-groupoid theory via soft sets for uncertainty modeling. *Filomat*, *32*(8), 2995-3030.

Sezgin, A., Ahmad, S., & Mehmood, A. (2019). A new operation on soft sets: Extended difference of soft sets. *Journal of New Theory*, *27*, 33-42.

Sezgin, A., & Atagün, A. O. (2011). On operations of soft sets. *Computers and Mathematics with Applications*, *61*(5), 1457-1467.

Sezgin, A., & Atagün, A. O. (2023). A new soft set operation: Complementary soft binary piecewise plus (+) operation. *Matrix Science Mathematic*, *7*(2), 125-142.

Sezgin, A., Atagün, A. O., Çağman, N., & Demir, H. (2022). On near-rings with soft union ideals and applications. *New Mathematics and Natural Computation*, *18*(2), 495-511.

Sezgin, A., & Aybek, F. N. (2023). A new soft set operation: Complementary soft binary piecewise gamma operation. *Matrix Science Mathematic*, *7*(1), 27-45.

Sezgin, A., Aybek, F. N., & Atagün, A. O. (2023a). A new soft set operation: Complementary soft binary piecewise intersection operation. *Black Sea Journal of Engineering and Science*, *6*(4), 330-346.

Sezgin, A., Aybek, F. N., & Güngör, N. B. (2023b). A new soft set operation: Complementary soft binary piecewise union operation. *Acta Informatica Malaysia*, *7*(1), 38-53.

Sezgin, A., & Çağman, N. (2024). A new soft set operation: Complementary soft binary piecewise difference operation. *Osmaniye Korkut Ata University Journal of the Institute of Science and Technology*, *7*(1), 58-94.

Sezgin, A., Çağman, N., & Atagün, A. O. (2017). A completely new view to soft intersection rings via soft uni-int product. *Applied Soft Computing*, *54*, 366-392.

Sezgin, A., Çağman, N., Atagün, A. O., & Aybek, F. N. (2023c). Complemental binary operations of sets and their application to group theory. *Matrix Science Mathematic*, *7*(2), 114-121.

Sezgin, A., & Çalışıcı, H. (2024). A comprehensive study on soft binary piecewise difference operation. *Eskişehir Teknik Üniversitesi Bilim ve Teknoloji Dergisi B - Teorik Bilimler*, *12*(1), 32-54.

Sezgin, A., & Dagtoros, K. (2023). Complementary soft binary piecewise symmetric difference operation: A novel soft set operation. *Scientific Journal of Mehmet Akif Ersoy University*, *6*(2), 31-45.

Sezgin, A., & Demirci, A. M. (2023). A new soft set operation: Complementary soft binary piecewise star operation. *Ikonion Journal of Mathematics*, *5*(2), 24-52.

Sezgin, A., & Sarıalioğlu, M. (2024). A new soft set operation: Complementary soft binary piecewise theta operation. *Journal of Kadirli Faculty of Applied Sciences*, *4*(2), 325-357.

Sezgin, A., & Yavuz, E. (2023a). A new soft set operation: Complementary soft binary piecewise lambda operation. *Sinop University Journal of Natural Sciences*, *8*(2), 101-133.

Sezgin, A., & Yavuz, E. (2023b). A new soft set operation: Soft binary piecewise symmetric difference operation. *Necmettin Erbakan University Journal of Science and Engineering*, *5*(2), 189-208.

Singh, D., & Onyeozili, I. A. (2012). Some results on distributive and absorption properties on soft operations. *IOSR Journal of Mathematics*, *4*(2), 18-30.

Stojanovic, N. S. (2021). A new operation on soft sets: Extended symmetric difference of soft sets. *Military Technical Courier*, *69*(4), 779-791.

Tunçay, M., & Sezgin, A. (2016). Soft union ring and its applications to ring theory. *International Journal of Computer Applications*, *151*(9), 7-13.

Vandiver, H. S. (1934). Note on a simple type of algebra in which the cancellation law of addition does not hold. *Bulletin of the American Mathematical Society*, *40*(12), 914-920.

Yavuz, E. (2024). *Soft binary piecewise operations and their properties* (Master's thesis). Amasya University, Graduate School of Natural and Applied Sciences, Amasya, Türkiye.

Zhu, P., & Wen, Q. (2013). Operations on soft sets revisited. *Journal of Applied Mathematics*, *2013*, 1-7.

Available online at https://li02.tci-thaijo.org/index.php/ssstj

Suan Sunandha Science and Technology Journal

Analysis of Land Use Changes with the Google Earth Engine (GEE) Platform: A Case Study in Saraburi Province

Tanachart Inpuron[1], Rattanawat Chaiyarat[1], Poonperm Vardhanabindu[1], Pattana Suwansumrit[2], Monthira Yuttitham[1*]

[1]Faculty of Environment and Resource Studies, Mahidol University, Phutthamonthon Sai 4 Road, Nakhon Pathom 73170, Thailand
[2]Faculty of Architecture, Rajamangala University of Technology Thanyaburi, Pathum Thani 12110, Thailand
***Corresponding author e-mail: monthira.yut@mahidol.ac.th**

Received: *9 April 2024* / **Revised:** *24 June 2024* / **Accepted:** *30 July 2024*

Abstract

The increasing trend of industrial development has led to rapid changes in land use over the past few years. The changes in land use in Saraburi Province were studied during three periods: the years 2000, 2010, and 2021, using Landsat satellite imagery processed with the Google Earth Engine (GEE) platform. The Random Forest (RF) machine learning algorithm was employed to classify land use activities into five categories based on the criteria set by the Department of Land Development. The overall accuracy values for land use were 83.59%, 82.42%, and 80.78%, respectively. The changes in land use have clearly affected the urban and built-up area. In 2000, there were only 7,561.15 ha, but by 2021, the area had increased to 34,221.11 ha. Forest areas also saw an increase. In 2000, there were only 79,626.26 ha, which increased to 163,542.05 ha in 2021, reflecting that Saraburi Province places importance on green areas. In contrast, agricultural areas decreased, which was attributed to the increasing number of households, according to the statistical report from the Department of Provincial Administration. In 2000, there were 168,979 households, and by 2021, they had increased to 288,275 households, resulting in a change in land use from agricultural areas to urban and built-up areas, totaling 22,114.95 ha. This information can be used as a guideline for planning development and land use to promote Saraburi Province, aiming towards becoming an eco-industrial city that is livable and has sustainable environmental quality in the future.

Keywords: Land-use changes, Landsat satellite, Google Earth Engine, Sustainability, Saraburi

1. Introduction

At present, Thailand has continuously increased land use as a result of economic expansion in many regions. The main activities are in the industrial sector, which is a factor that causes changes in land use. In addition, the trend of industrial development also stimulates the economy. As a result, the population has increased, leading to the expansion of community areas and buildings. All these changes are considered important factors that greatly influence land use. Saraburi Province is one of the provinces with the most continuous economic development.

The province's location has geographic characteristics that are conducive to the expansion of the investment sector and convenient transportation, which can connect the transportation sector to many regions. This supports the expansion of the industrial sector and the size of the city, which have grown together, causing continuous changes in land use. In developing the province, the 20-year national strategy, 23 master plans under the national strategy, and government administration plans have been adopted. These include the National Economic and Social Development Plan No. 12 (Office of the National Economic and Social Development Board, 2017), the Upper Central Region Development Plan, and the Strategy of the Ministry of Interior 2018-2022, as well as an analysis of the potential of the provinces and the needs of the people to be used in connection with the determination of the provincial vision.

The current development goals of Saraburi Province are focused on developing it into an eco-industrial city. The objective is to develop Saraburi Province into a livable, eco-industrial city with balanced and continuous growth. Therefore, geo-informatics technology has been introduced for remote sensing that is suitable for analyzing large spatial data. Using Landsat satellite image data, which is considered a natural resource survey satellite (Geo-Informatics and Space Technology Development Agency (Public Organization) [GISTDA], 2009), it is appropriate to classify land uses. When analyzing data on the Google Earth Engine platform, which is a platform used to analyze geospatial data provided in an open-source manner, it can be used to analyze spatial data free of charge (Supatra & Puttanawarat, 2021). In this study, the objective was to examine land use changes in Saraburi Province. Geo-informatics technology was applied to analyze the context and factors affecting land use change. The aim is to provide guidelines for promoting Saraburi Province towards becoming a livable eco-industrial city with sustainable environmental quality in the future.

2. Materials and Methods

2.1 Data acquisition

2.1.1 Satellite datasets

This research uses satellite image data from Landsat-5 TM and Landsat-8 OLI, covering Saraburi Province for three periods: 2000, 2010, and 2021. These images were selected based on their availability on the Google Earth Engine platform, which covers the study period. Details are shown in Table 1.

Table 1. Satellite datasets used in the analysis.

Satellites	Data characteristics		
	Resolution (m)	**Band**	**Acquisition**
Landsat-5 TM	30	Visible (band 1 - 3) NIR (band 4) SWIR (band 5)	2000, 2010,
Landsat-8 OLI	30	Visible (band 2 - 4) NIR (band 5) SWIR (band 6)	2021

Source: GISTDA (2009)

2.1.2 Digital data files on the administrative boundaries of Saraburi Province are from the Geo-Informatics and Space Technology Development Agency (Public organization).

2.1.3 Geographic Information System data analysis program.

2.2 Methods for analyzing land use and land use change

Land use analysis involves selecting satellite image data from the study area, which is the boundary of Saraburi Province. Landsat 5 and Landsat 8 satellite images that are free of clouds or other obstructions affecting the analysis were selected. Using the Google Earth Engine (GEE) platform and interpreting satellite images with supervised classification, the efficiency of the Random Forest (RF) machine learning algorithm in land use classification is demonstrated by an overall accuracy of 92.00% (Intarat, 2022). It was used for three time periods: 2000, 2010, and 2021. Land use activity data were classified into five categories based on the criteria of the Land Development Department: urban and built-up area, agricultural area, forest area, water body, and miscellaneous area.

Land use changes were studied using the land use data interpreted from satellite images, classified into three periods: the first period between 2000 and 2010, the second period between 2010 and 2021, and the third period between 2000 and 2021.

2.3 Methods of verifying the correctness of land use

Validation of land use classification involves comparing the results of the classification of land use in each category obtained from interpreting satellite images with land use data from the closest year provided by the Land Development Department. This comparison is conducted using the Confusion Matrix method by selecting checkpoints to compare the interpretations with reference data from the Land Development Department. The desired overall accuracy percentage (P) is set to be equal to 80%, the acceptable error value is set to 5%, and the confidence level is set to 95% (Z = 1.96). For this study, a total of 256 points (according to binomial probability theory and stratified distribution) were examined using the equalized stratified random sampling format in the GIS program (Fitzpatrick-Lins, 1981). The error matrix is a statistical tool used to measure the accuracy of data classification. By comparing the data obtained from classification with reference data, it helps to explain the overall accuracy, which can be easily calculated. It also provides useful information for evaluating accuracy. To calculate overall accuracy (OA), the ratio of the sum of the number of correctly classified points (the diagonal elements of the matrix) to the total number of points examined is used. This ratio is then expressed as a percentage. The overall accuracy value, in percentage form, indicates how accurately the entire classification system classifies the data compared to the reference data. An OA value closer to 100% indicates a more accurate classification of the data.

The Kappa coefficient is calculated using statistical methods to test the consistency between two sets of data. It does not rely on assumptions about whether the data is normally distributed. This differs from the calculation of overall accuracy (OA), which measures the accuracy of all classifications. According to McHugh (2012), the Kappa coefficient is divided as follows:

A confidence value of 0.81 and above	means very high confidence.
Confidence values from 0.61 to 0.80	indicate high confidence.
Confidence values ranging from 0.41 to 0.60	indicate moderate confidence.
Confidence values ranging from 0.21 to 0.40	indicate low confidence.
A confidence value less than 0.20	means no confidence.

3. Results and Discussion

The study of land use for the years 2000, 2010, and 2021, using Landsat satellite images and analyzing land use on the Google Earth Engine (GEE) platform with the Random Forest (RF) machine learning algorithm, categorized land use activity data into five categories based on the criteria of the Land Development Department, as shown in Table 2. These categories are as follows:

As shown in Figure 1, land use in 2000 revealed that the agricultural area was the largest, covering 224,021.20 ha (64.34%), followed by forest area with 79,626.26 ha (22.87%), miscellaneous area with 31,668.67 ha (9.10%), urban and built-up area with 7,561.77 ha (2.17%), and water body with 5,263.19 ha (1.51%). In 2010, it was found that the agricultural area had the largest extent, covering 181,850.80 ha (52.23%), followed by the forest area with 119,630.29 ha (34.36%), the miscellaneous area with 24,917.79 ha (7.16%), the urban and built-up area with 15,139.94 ha (4.35%), and the water body with 6,621.64 ha (1.90%).In 2021, it was found that the forest area had the largest extent, covering 163,542.05 ha (46.97%), followed by the agricultural area with 115,576.70 ha (33.20%), the miscellaneous area with 23,705.06 ha (6.81%), the urban and built-up area with 34,221.11 ha (9.83%), and the water body with 11,115.54 ha (3.19%).

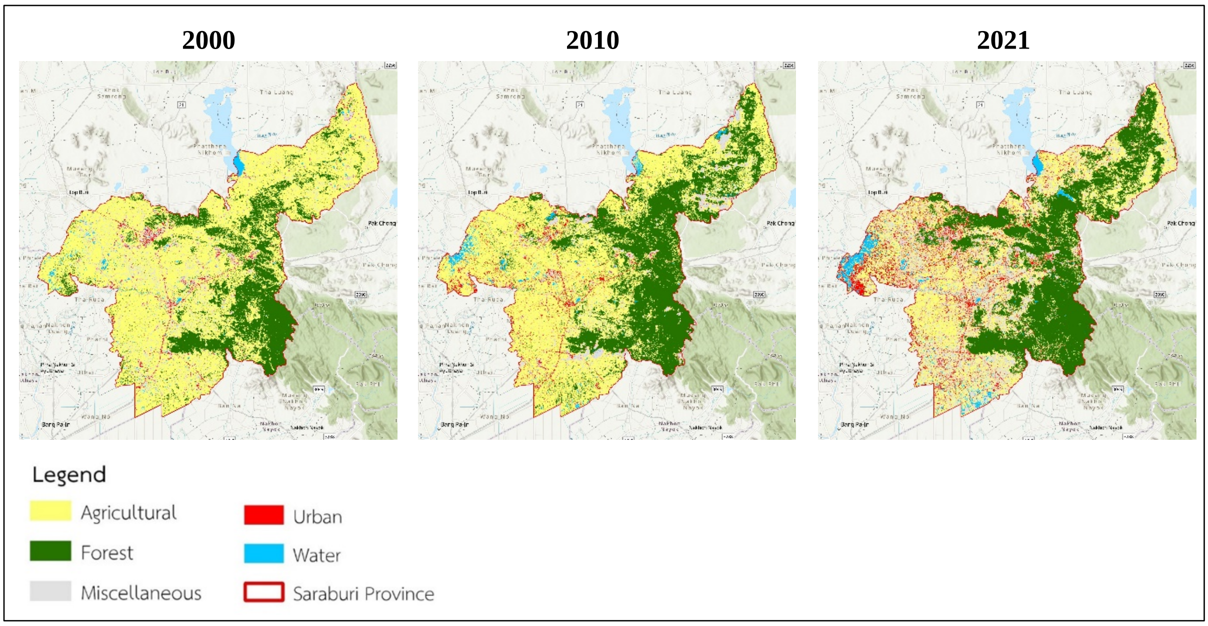


Figure 1. Land use map of Saraburi Province in 2000, 2010, and 2021.

Table 2. Land use in 2000, 2010, and 2021.

Land use	Land use area (ha)		
	2000	**2010**	**2021**
Agricultural area	224,021.20	181,850.80	115,576.70
Forest area	79,626.26	119,630.29	163,542.05
Miscellaneous area	31,688.67	24,917.79	23,705.06
Urban and built-up area	7,561.15	15,139.94	34,221.11
Water body	5,263.19	6,621.64	11,115.54
Total	**348,160.46**	**348,160.46**	**348,160.46**

When comparing the accuracy of land use in this study with that of the Land Development Department, the overall accuracy in this study is close to that of the Land Development Department. The overall accuracy was 83.59%, 82.42%, and 80.78% for the years 2000, 2010, and 2021, respectively (Table 3).

Table 3. Accuracy values of land use data classification.

Validation	Land use		
	2000	2010	2021
Kappa Coefficient	0.74	0.72	0.72
Overall Accuracy	83.59	82.42	80.78

The land use changes over the three time periods using data from satellite images are as follows:

In the first period, between 2000 and 2010, it was observed that the agricultural and miscellaneous areas decreased, while the forest area, water body, and urban and built-up areas increased. The proportion of land use areas was higher when interpreting satellite images, as shown in Figure 2. The urban and built-up area increased to 7,578.79 ha, which is consistent with the Department of Provincial Administration statistics indicating that the number of households in 2010 increased by 55,089 from 2000, leading to a change in the agricultural area. Concurrently, the forest area increased during this period, which aligns with the objectives of the 8th National Economic and Social Development Plan. This plan focused on natural resource and environmental management, aiming to conserve and restore natural resources to maintain ecosystem balance and enhance the quality of life for communities, thus serving as a crucial foundation for the country's long-term development process (Table 4).

Table 4. Land-use in 2000 and 2010.

Land use	Land use area (ha)		
	2000	2010	Land use change
Agricultural area	224,021.20	181,850.80	(-) 42,170.40
Forest area	79,626.26	119,630.29	(+) 40,004.03
Miscellaneous area	31,688.67	24,917.79	(-) 6,770.88
Urban and built-up area	7,561.15	15,139.94	(+) 7,578.79
Water body	5,263.19	6,621.64	(+) 1,358.45
Total	**348,160.46**	**348,160.46**	

Note: (+) means the area has increased. (-) means the area has decreased.

Table 5. Land-use change matrix as observed between in 2000 and 2010.

	Land use class	2010 (ha)					
		A	F	M	U	W	Total 2000
2000 (ha)	**A**	145,223.10	52,144.43	15,735.28	7,209.89	3,708.49	**224,021.20**
	F	10,927.40	63,614.79	3,653.97	890.85	539.25	**79,626.26**
	M	20,431.44	2,888.89	4,133.51	3,869.91	364.92	**31,688.67**
	U	2,810.79	456.91	1,211.11	3,009.83	72.50	**7,561.15**
	W	2,458.06	525.27	183.92	159.46	1,936.48	**5,263.19**
	Total 2010	**181,850.80**	**119,630.29**	**24,917.79**	**15,139.94**	**6,621.64**	**348,160.46**

Note: A = Agricultural area, F = Forest area, M = Miscellaneous area, U = Urban and built-up area, W = Water body

In the second period, between 2010 and 2021, it was observed that the agricultural area decreased. Conversely, the forest area, miscellaneous area, water body, and urban and built-up areas are likely to have increased. During this period, the population in Saraburi Province increased significantly by 30,818 people. The proportion of land use areas was determined by interpreting satellite images. The urban and built-up area increased to 19,081.17 ha, which is consistent with Department of Provincial Administration statistics showing that the number of households in 2021 increased by 64,207 from 2010, leading to changes in the agricultural area. Simultaneously, the forest area increased due to the objectives of the 11th National Economic and Social Development Plan, which aimed to manage natural resources and the environment adequately to maintain ecological balance and provide a solid foundation for national development (Table 6).

Table 6. Land-use change in 2010 and 2021.

Land use	Land use area (ha)		
	2010	2021	Land use change
Agricultural area	181,850.80	115,576.70	(-) 66,274.10
Forest area	119,630.29	163,542.05	(+) 43,911.75
Miscellaneous area	24,917.79	23,705.06	(-) 1,212.73
Urban and built-up area	15,139.94	34,221.11	(+) 19,081.17
Water body	6,621.64	11,115.54	(+) 4,493.90
Total	**348,160.46**	**348,160.46**	

Note: (+) means the area has increased. (-) means the area has decreased.

Table 7. Land-use change matrix as observed between in 2010 and 2021.

	Land use class	2021 (ha)					
		A	F	M	U	W	Total 2010
2010 (ha)	**A**	86,872.74	51,832.04	15,571.06	21,969.86	5,605.09	**181,850.80**
	F	13,811.98	100,039.87	2,669.65	2,097.35	1,011.45	**119,630.29**
	M	9,395.28	8,658.91	3,368.83	3,186.13	308.64	**24,917.79**
	U	4,150.94	2,196.74	1,933.66	6,502.61	356.00	**15,139.94**
	W	1,345.76	814.49	161.87	465.16	3,834.37	**6,621.64**
	Total 2021	**115,576.70**	**163,542.05**	**23,705.06**	**34,221.11**	**11,115.54**	**348,160.46**

Note: A = Agricultural area, F = Forest area, M = Miscellaneous area, U = Urban and built-up area, W = Water body

In the third period, between 2000 and 2021, it was observed that agricultural areas and miscellaneous areas had noticeably decreased. Meanwhile, forest areas, water bodies, and urban and built-up areas exhibited an increasing trend. This economic growth was a result of continuous industrial development initiatives within Saraburi Province. Since the Industrial Estate Authority of Thailand established the Nong Khae Industrial Estate in 1990 and the WHA Saraburi Industrial Estate in 1988, a significant amount of employment and land use has been created. Consequently, statistics from the Registration Office, Department of Provincial Administration, report that the number of households increased from 168,979 in 2000 to 288,275 in 2021, an increase of 41.3%. During this period, Saraburi Province focused on developing the area into a city of international trade and investment in green industry while creating a happy society. This aligns with the objectives of the provincial development plan, which emphasizes promoting industrial economic development while considering environmentally friendly industrial practices and increasing green areas within the province (Table 8).

Table 8. Land-use change from 2000 to 2021.

Land use	Land use area (ha)		
	2000	2021	Land use change
Agricultural area	224,021.20	115,576.70	(-) 108,444.50
Forest area	79,626.26	163,542.05	(+) 83,915.78
Miscellaneous area	31,688.67	23,705.06	(-) 7,983.60
Urban and built-up area	7,561.15	34,221.11	(+) 26,659.97
Water body	5,263.19	11,115.54	(+) 5,852.35
Total	**348,160.46**	**348,160.46**	

Note: (+) means the area has increased. (-) means the area has decreased.

Table 9. Land-use change matrix as observed between in 2000 and 2021.

Land use class		2021 (ha)					
		A	**F**	**M**	**U**	**W**	**Total 2000**
2000 (ha)	**A**	94,183.81	83,565.27	16,934.61	22,114.95	7,222.56	**224,021.20**
	F	6,035.76	70,296.73	1,236.52	1,381.87	675.38	**79,626.26**
	M	12,546.24	7,647.62	4,287.61	6,766.43	440.76	**31,688.67**
	U	1,668.79	1,292.76	1,095.03	3,427.96	76.60	**7,561.15**
	W	1,142.10	739.66	151.30	529.89	2,700.24	**5,263.19**
	Total 2021	**115,576.70**	**163,542.05**	**23,705.06**	**34,221.11**	**11,115.54**	**348,160.46**

Note: A = Agricultural area, F = Forest area, M = Miscellaneous area, U = Urban and built-up area, W = Water body

Table 10. Land-use change over 3 periods.

Land use	Year			Land use area (ha)		
	2000	**2010**	**2021**	**2000-2010**	**2010-2021**	**2000-2021**
Agricultural area	224,021.20	181,850.80	115,576.70	(-) 42,170.40	(-) 66,274.10	(-) 108,444.50
Forest area	79,626.26	119,630.29	163,542.05	(+) 40,004.03	(+) 43,911.75	(+) 83,915.78
Miscellaneous area	31,688.67	24,917.79	23,705.06	(-) 6,770.88	(-) 1,212.73	(-) 7,983.60
Urban and built-up area	7,561.15	15,139.94	34,221.11	(+) 7,578.79	(+) 19,081.17	(+) 26,659.97
Water body	5,263.19	6,621.64	11,115.54	(+) 1,358.45	(+) 4,493.90	(+) 5,852.35
Total	**348,160.46**	**348,160.46**	**348,160.46**			

Note: (+) means the area has increased. (-) means the area has decreased.

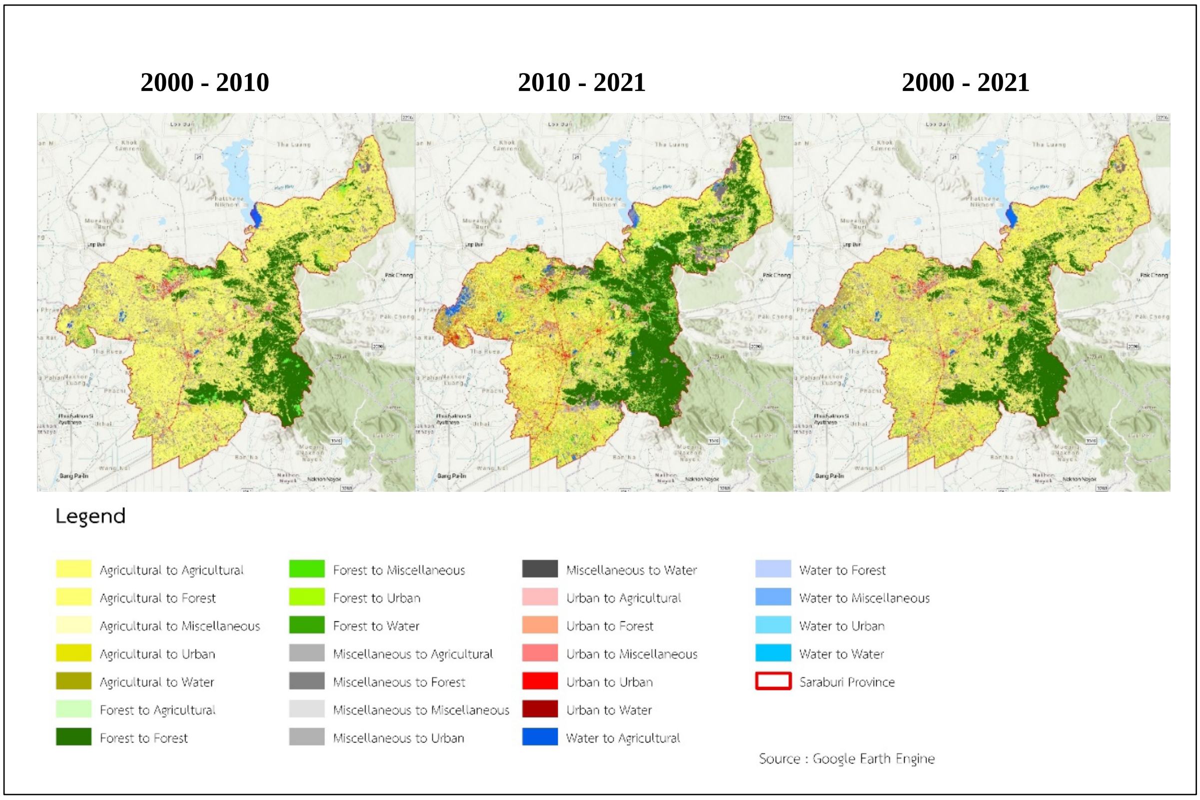


Figure 2. Map of Land-use change, Saraburi Province, 3 periods.

4. Conclusions

The results of the study of land use in Saraburi Province for the years 2000, 2010, and 2021 showed overall accuracies of 83.59%, 82.42%, and 80.78%, respectively. The analysis of land-use changes over these three periods clearly indicated changes in area, particularly in the urban and built-up areas, which have been continuously increasing. The area of urban and built-up land was only 7,561.15 ha in 2000 but increased to 34,221.11 ha by 2021, representing a growth of 26,659.96 ha (Table 10). This increase is attributed to the continuous economic growth in Saraburi Province, which has developed into an industrial city. Conversely, agricultural land use decreased significantly, from 224,021.20 ha in 2000 to 115,576.70 ha in 2021, a reduction of 108,444.50 ha over 20 years. This trend is consistent with Pannon's (2018) study on the Bang Pakong River Basin, which reported significant decreases in agricultural land due to urbanization driven by population growth. Similarly, in Saraburi Province, agricultural areas have been repurposed for other uses.

The forest area, however, showed a tendency to increase due to Saraburi Province's focus on developing itself into an international trade and investment hub for green industries. The provincial development plan emphasizes industrial and economic growth while increasing green spaces, leading to an expansion in forest areas. This indicates that Saraburi Province prioritizes its development plans and forest conservation.

The data obtained from this research can be used for spatial development and land use planning in the province. It can also serve as a policy guideline to promote Saraburi Province towards becoming a livable eco-industrial city with sustainable environmental quality in the future.

Acknowledgements

The author expresses gratitude to the Geo-Informatics and Space Technology Development Agency (Public Organization) and the Department of Land Development for providing data support for this study.

Conflict of Interest

There is no conflict of interest.

ORCID

Author: Tanachart Inpuron
https://orcid.org/0009-0009-6568-6975

References

Fitzpatrick-Lins, K. (1981). Comparison of sampling procedures and data analysis for a land-use and land-cover map. *Photogrammetric Engineering and Remote Sensing*, *47*(3), 343-351.

Geo-Informatics and Space Technology Development Agency (Public Organization). (2009). *Space technology and geo-informatics*. Bangkok, Thailand: Amarin Printing and Publishing Public Company.

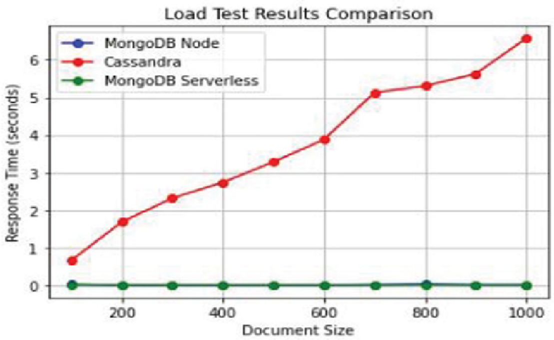
Intarat, K. (2022). Land use classification in Nakhon Nayok Province using machine learning algorithms and Sentinel-2 image. *Burapha Science Journal*, *27*(2), 1153-1171.

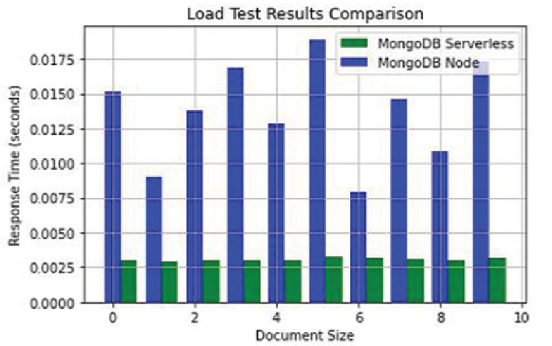
McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, *22*(3), 276-282.

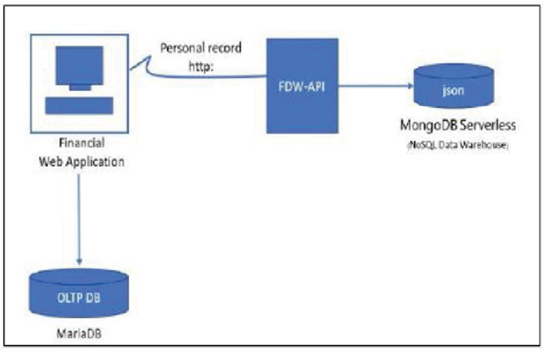
Office of the National Economic and Social Development Board. (2017). *The Twelfth National Economic and Social Development Plan (2017-2021).* Bangkok, Thailand: Office of the Prime Minister.

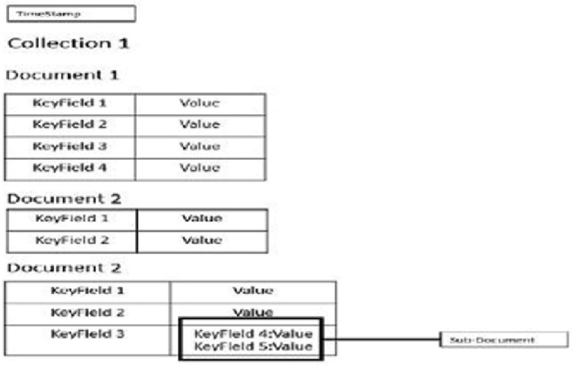
Pannon, S. (2018). *Land use change and its impacts on water pumping in Bang Pakong River basin, Thailand* (Master's thesis). Asian Institute of Technology, Thailand.

Supatra, P., & Puttanawarat, N. (2021). *Development of geographic information systems and spatial data analysis*. Faculty of Science and Technology, Prince of Songkhla University, Surat Thani Campus.









SSSTJ

Suan Sunandha Science and Technology Journal
https://li02.tci-thaijo.org/index.php/ssstj/index

Crossref
TCI
ศูนย์ดัชนีการอ้างอิงวารสารไทย
Thai-Journal Citation Index Centre
Google scholar
DOAJ DIRECTORY OF OPEN ACCESS JOURNALS