Progress in IS


Volker Wohlgemuth
Hamdy Kandil
Amna Ramzy *Editors*


# Advances and New Trends in Environmental Informatics

## Transboundary Environmental Challenges: Digital Inclusion for Sustainable Development—Environinfo 2024

Springer

# Progress in IS

Progress in IS encompasses the various areas of Information Systems in theory and practice, presenting cutting-edge advances in the field. It is aimed especially at researchers, doctoral students, and advanced practitioners. The series features both research monographs, edited volumes, and conference proceedings that make substantial contributions to our state of knowledge and handbooks and other edited volumes, in which a team of experts is organized by one or more leading authorities to write individual chapters on various aspects of the topic. Individual volumes in this series are supported by a minimum of two external reviews.

The Series is SCOPUS-indexed.

Volker Wohlgemuth · Hamdy Kandil ·
Amna Ramzy
Editors

# Advances and New Trends in Environmental Informatics

## Transboundary Environmental Challenges: Digital Inclusion for Sustainable Development—EnviroInfo 2024

Springer

*Editors*
Volker Wohlgemuth
HTW Berlin, University of Applied Sciences
Berlin, Germany

Hamdy Kandil
German University in Cairo
New Cairo City, Egypt

Amna Ramzy
German University in Cairo
New Cairo City, Egypt

ISSN 2196-8705 ISSN 2196-8713 (electronic)
Progress in IS
ISBN 978-3-031-85283-1 ISBN 978-3-031-85284-8 (eBook)
https://doi.org/10.1007/978-3-031-85284-8



This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# Preface

This book contains the proceedings of the 38th International and Interdisciplinary Conference on Environmental Informatics (EnviroInfo 2024).

EnviroInfo 2024 took place from November 12 to 14, 2024 in Cairo, Egypt. It was organized and hosted by the German University in Cairo in collaboration with the Technical Committee of Environmental Informatics of the German Informatics Society—GI (Gesellschaft für Informatik e.V.).

The EnviroInfo conference series combines national and international initiatives in applied informatics and sustainability to present and discuss the latest developments in ICT for sustainability.

This book compiles peer-reviewed research papers showcasing new scientific approaches and latest developments in environmental informatics. Papers in this volume cover a broad range of scientific perspectives, including environmental informatics technologies and methods such as Artificial Intelligence Applications, Sustainable Mobility, Green Coding, ICT, Circular Economy and more. The main theme of the conference is how Environmental Informatics contribute to the United Nations' Sustainable Development Goals (SDGs) with a focus on which goals are addressed by the Environmental Informatics community. This year's conference also focused on the theme "Transboundary Environmental Challenges: Digital Inclusion for Sustainable Development" to highlight the potential of environmental informatics to facilitate cross-border cooperation aimed at addressing shared environmental challenges and bridging the technological gap for sustainable development.

We thank all authors for their papers. Special thanks are due to the program and organizing committee members for their work in reviewing the papers. We would like to thank our cooperation partners at the German University in Cairo, Faculty of Engineering and Materials Science: Prof. Hamdy Kandil, Assoc. Prof. Dr.-Ing. Amna Ramzy and M.Sc. Simone Khalil, and the team of Prof. Volker Wohlgemuth from the German Informatics Society for their great support.

Finally, we express our heartfelt thanks to our sponsors for their generous support of the conference. Their contributions have been instrumental in making this event a success.

New Cairo City, Egypt
November 2024

Volker Wohlgemuth
Hamdy Kandil
Amna Ramzy

# Contents

**AI & ML in Environmental Monitoring and Sustainability I**

**A Method for Potential Analysis to Identify Application Scenarios for Machine Learning** . . . . . 3
Frank Fuchs-Kittowski, Paul Schulze, Andreas Abecker, Jonas Lachowitzer, Stefan Lossow, Heino Rudolf, and Erik Rodner

**A Geo-Parser for German Documents with Environmental Context** . . . . 21
Nicolas Doms, Thorsten Schlachter, and Lisa Hahn-Woernle

**Low-Cost Air Quality Sensor Nodes in a Network Setup: Using Shared Information to Impute Missing Values** . . . . . 35
Theodosios Kassandros, Evangelos Bagkis, and Kostas Karatzas

**Living with Future AI: Studying Experience, Attitudes and Expectations of Greek Smart Home Users** . . . . . 47
Lu Jin, Sotirios Athanasoulias, Dominik Pins, Alexander Boden, Britta Essing, and Nikolaos Ipiotis

**Business Intelligence for Green Decision-Making**

**Towards Enhanced Manufacturing Precision: The Role of AI-Driven Quality Grading Systems** . . . . . 71
Mohamed Ghoneim, Radwa Hussein, and Nada Sharaf

**Environmental Informatics for Industry**

**Use of Artificial Intelligence to Optimize Processes and Increase Resource Efficiency in Small and Medium-Sized Enterprises** . . . . . 91
Martina Willenbacher and Volker Wohlgemuth

**Low-Power and High-Perceptibility Underwater Drone Implementation for Inshore Aquaculture** . . . . . 105
Tian Song, Takafumi Katayama, Takashi Shimamoto, and Xiantao Jiang

**Climate and Ecosystems Modeling I**

**Integration of Process Modeling and LCA for the Assesment of the Environmental Impact of Pharmaceutical Industries: Case Study** . . . . . 125
Shaimaa Ragab and Amna Ramzy

**Modeling for the Design and Manufacturing of Superabsorbent Polymers (SAPs) for Agriculture in Arid Areas by Using the RSM Method** . . . . . 141
Maysa Muhammad, M. L. Tawfic, Mohamed Taha, and Ahmed Elsabbagh

**AI & ML in Environmental Monitoring and Sustainability II**

**A GIS-Based Decision Support System for Renewable Energy Investments in Egypt: A Machine Learning Approach** . . . . . 161
Omar S. Salem, Moustafa A. Baraka, and Ahmed M. Abdel Sattar

**Analytics for Environmental Sustainability I**

**Optimizing Foundry Operations: A Case Study on Improving Material and Energy Flow Efficiency in Line with the SDGs** . . . . . 183
Ahmed ElFeky, Hazem Elshiekh, and Amna Ramzy

**Achieving Sustainable Design Utilizing Software Programs and Tools in Architecture Field** . . . . . 201
Hebatulla H. Elsharawy

**Green Coding Techniques and Methodologies**

**Investigating the Use of GitHub Copilot for Green Software** . . . . . 219
Maria Stivala, Iffat Fatima, and Patricia Lago

**Enhancing Energy Efficiency in AI: A Multi-faceted Analysis Across Time Series, Semantic AI and Deep Learning Domains** . . . . . 237
Lejla Begic Fazlic, Berkay Cetkin, Achim Guldner, Matthias Dziubany, Julian Heinen, Stefan Naumann, and Guido Dartmann

**Towards an Energy Estimation Model for ABAP Statements on SAP S/4HANA Software** . . . . . 257
Ann-Christin Fleischle, Laurenz Baumgart, Holger Wittges, and Helmut Krcmar

**Sustainability in Transportation and Logistics**

**Green Delivery Analytics: Software Vision of a Planning Tool for Sustainable Last Mile Logistics** . . . . 275
Benjamin Wagner vom Berg, Richard Schulte, Mattes Leibenath, Senad Hasanspahic, Arne Kruse, Christoph Drieling, Kian Seelaff, Aina Andriamananony, Uta Kühne, and Franziska Giese

**Climate and Ecosystems Modelling II**

**Strategic Sensor Placement for the Identification of Disinfection Byproducts from Chlorinated Drinking Water: Case Study in the Water Distribution Network of Coimbra, Portugal** . . . . 295
Aristotelis Magklis and Andreas Kamilaris

**Development of 5G Smart Farming Dashboard to Detect Wild Animals on Pasture by Using Convolutional Neural Network** . . . . 313
Ali Akyol, Rami Chahin, Jorge Marx Gómez, Hendrik Schwabe, Henrika Schwanke, and Nora Uderstadt

**Supply Chain and Sustainability Planning**

**Overview, Evaluation and Comparison of Current CCS and CCU Technologies** . . . . 331
Josua Benz, Konstantin Blank, Silvio Hötzel, Jonas Just, Charlotte Lau, Franziska Schicks, Nicolas von der Ahe, and Jochen Wittmann

**Towards Sustainable Training Frameworks: Investigating Collaborative Learning in Virtual Reality** . . . . 349
Salsabil Lotfy Monib, Nada Nasser, and Nada Sharaf

**Analytics for Environmental Sustainability II**

**Mobile Application for Citizens for Information About Low Water and Drought** . . . . 367
Fabian Bressel, Frank Fuchs-Kittowski, Tim Seegert, Maximillian Deharde, Ruben Müller, Bernd Pfützner, Moritz Zemann, and Andreas Abecker

**Eco-Innovation and Circular Economy Performance in European Union Countries** . . . . 389
Pawel Bartoszczuk

**Empowering Sustainable Development: The Role of E-Commerce Adoption in Advancing SDGs Among Youth** . . . . 401
Noha Tarek Amer, Hany Abdelghaffar, and Laila Amr

# AI & ML in Environmental Monitoring and Sustainability I

# A Method for Potential Analysis to Identify Application Scenarios for Machine Learning

## Case Study in a State Environmental Agency

**Frank Fuchs-Kittowski, Paul Schulze, Andreas Abecker, Jonas Lachowitzer, Stefan Lossow, Heino Rudolf, and Erik Rodner**

**Abstract** This article presents a method for potential analysis for identifying application potentials for application of machine learning (ML) in organizations. This method describes a systematic approach that emphasizes both the requirements of employees and business processes. The structure and artefacts of the method are described in this paper. Furthermore, the application of this method at an environmental agency as pilot user is presented. The results show that this method helped the environmental agencies to quickly develop ML solutions and select beneficial ML solutions effectively.

**Keywords** Potential analysis · Machine learning · Artificial intelligence · Use cases · State environmental agency

F. Fuchs-Kittowski (✉) · P. Schulze · E. Rodner
HTW Berlin, University of Applied Sciences, Wilhelminenhofstraße, Berlin, Germany
e-mail: Frank.Fuchs-Kittowski@htw-berlin.de

P. Schulze
e-mail: Paul.Schulze@htw-berlin.de

E. Rodner
e-mail: Erik.Rodner@htw-berlin.de

A. Abecker · J. Lachowitzer · S. Lossow
Disy Informationssysteme GmbH, Ludwig-Erhard-Allee, Karlsruhe, Germany
e-mail: Andreas.Abecker@disy.net

J. Lachowitzer
e-mail: Jonas.Lachowitzer@disy.net

S. Lossow
e-mail: Stefan.Lossow@disy.net

H. Rudolf
Simplex4Data GmbH, Am Waldschlößchen, Dresden, Germany
e-mail: heino.rudolf@simplex4data.de



V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_1

# 1 Introduction

With the continuing development of powerful information and communication technologies (storage media, computer networks, processors, etc.) and the increasing availability of huge amounts of data, interest in machine learning (ML) technologies has increased over the years to support business processes as well as to increase their efficiency. ML technologies are nowadays used in many areas, such as autonomous driving [1], security technology [2, 3], medical technology [4, 5], and the environmental sector, including water management [6, 7], energy management [8, 9] forestry [10], and agriculture [11, 12].

There are already numerous publications on successfully implemented solutions in various organizations. This shows what range of activities can be supported (efficiently and effectively) in organizations by using ML technologies. However, it remains an open question how these organizations select their business processes or procedures for using ML and design their IT systems to support new ML processes.

At present, there is no reliable knowledge in research on how machine learning (ML) use cases can be identified in organizations. This makes the development of ML applications in organizations way more difficult than needed so that the potential of these new technologies cannot exploited or just in an insufficient way.

The decision-making process for the use and design of ML applications in an organization must be systematic and comprehensible. The requirements of potential users and their work processes must align with the technical capabilities of ML technology. Therefore, it's essential that the introduction of ML is not only orientated toward the technology, but rather toward the requirements—the challenges and problems—of a specific field of application or business process. At the same time, potential users often lack necessary knowledge and experience to realistically assess the specifics and potential of ML technologies for solving identified challenges.

This article presents a method for an analysis of the potential for identifying ML applications. The experiences and results from using this method in various organizations (notably a state environmental agency) are discussed. The method is not focusing on the available ML technologies, but instead on the requirements of potential users and their field of work in their organization and the business processes of their organization. According to the design thinking approach, all three perspectives have been considered during the process [13]: Economic viability, technical feasibility, and human desirability.

The article is structured as follows: Sect. 2 presents the technical principles and the state of research on identifying ML use cases in organizations. Section 3 describes the method for identifying ML use cases in organizations. This is followed in Sect. 4 by a presentation of the results and experience gained from applying this method in a state environmental agency as pilot organization. The article concludes with a summary and an outlook in Sect. 5.

## 2 Technical Principles and State of Research—Identifying ML Use Cases in Organizations

This chapter presents the technical principles and the current state of research on identifying ML use cases in organizations.

### 2.1 *Technical Principles*

The aim of this article is to present a systematic approach for identifying scenarios for ML applications (potential analysis). This method is one of the first steps or a stage in a more comprehensive process model that provides systematic support and guidance throughout the entire process of developing ML applications in organizations (see Fig. 1).

Regardless of the general software development process model in which the project is managed, such as the waterfall model [14], the "V-Modell" [15] or agile models [16] (with requirements analysis, design, implementation and testing, deployment and operation), or the use of a specific process model for the development of ML applications, such as CRISP-DM [17], CRISP-ML(Q) [18], PAISE [19], ML4P [20] (with business understanding, data understanding, data preparation, modeling, evaluation and deployment), the analysis of potentials (i.e., identifying use cases) is a preliminary or first stage in a ML project.

These upstream stages of comprehensive process models are presented below:

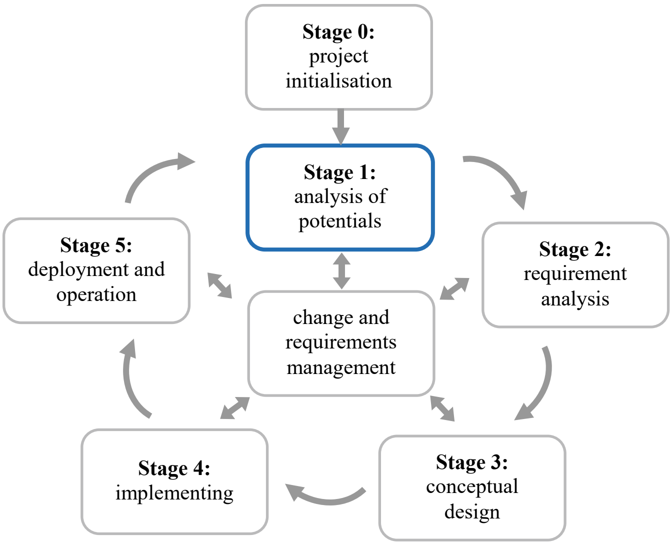



**Fig. 1** Integration of the potential analysis into the software development process

- **Project Initialization:** In this preliminary stage of a project the used framework will be defined (often only for this stage) and the project is set up. Therefore, this is classified as a project management task (see DIN 69901-2 [21] or PRINCE2 [22]). Roles and responsibilities must be clarified, the used project management processes must be selected, and the overall objectives must be defined. This preliminary stage generally ends with approval by management.
- **Potentials Analysis:** In this first project stage, fields of ML application are identified. This includes the determination and evaluation of specific scenarios ("case studies"), the identification of objectives as well as the planning of an overall process strategy. The result is a list of potentially possible ML applications for the considered organization. The individual examined steps are:
  - identification and selection of potential application scenarios,
  - assessment and selection of application scenarios, and
  - final agreement on applications scenarios and defining objectives.

## *2.2 State of Research*

Specific process models for ML do not provide guidance for the identification of potentially possible use cases for organizations as described above. Models such as the widely used industry standard CRISP-DM typically start with the stage "business understanding". In this stage, a comprehensive understanding of the objectives and requirements of an ML project is created. This means that the considered use case has already been identified. Even latest agile ML methods do not include a phase for identifying potentially possible use cases.

To date, there is very little scientific literature on identification of potentially possible ML use cases:

- In [23], six activities: isolating, preparing, discovering, understanding, designing, and implementing are integrated into an iterative process and have been supported by continuous project management. The focus lies on discovering (identifying problems), understanding (searching for causes), and designing (finding solutions). However, how these phases should be carried out exactly (including specific methods, artefacts, and results) remains largely vague (methods like stakeholder analysis, requirements definition, etc., should be considered here).
- In [24], the stages are differentiated into preparation, generating ideas, evaluation, prioritization, and operation. However, this method is aiming at large organizations or enterprises. For example, in the stage preparation, AI vision has been developed top down and an AI maturity level of the organization has been evaluated. The overall stage of identifying potentially possible ML use cases remains also vague and seems to be overstrained. The reason for this is based on specific problems in their respective strategic areas of the considered organization which must be considered to solve these problems with appropriate AI technologies (e.g., computer vision, computer audition, computer linguistics, etc.). For identifying

and evaluating, a structured canvas is used, and furthermore, an evaluation matrix is used as a method to support prioritizing (with benefits $\times$ implementation costs as scale).

In addition to established process models and existing scientific literature, there exist several guidelines and tutorials for introducing ML in organizations. Guidelines (such as [25]) usually provide a general introduction to the topic of ML, refer to the specific ML process models (usually CRISP-DM), and sometimes list a series of possible use cases. However, no instructions are given on how to identify and examine such use cases. In some cases, the focus is even placed on analyzing existing data to find use cases. This approach is generally regarded as problematic when ML is established for its own sake and there is no reference to existing business processes (see the stage "business understanding" in CRISP-DM).

In summary, a method for identifying potentially possible ML applications in organizations is still missing. Furthermore, specific conditions and requirements of smaller organizations should be considered, focusing on the stages of developing ideas, evaluation, and selection for new ML applications.

## 3 Method for Identifying ML Application Scenarios in Organizations (Potentials Analysis)

In the following section the first stage—the analysis of potentials—is described in detail. The aim of this stage is to identify, evaluate, and select potentially possible ML applications as well as to define objectives for the selected ML applications which need to be achieved for organizations under consideration. These stages are:

1. identification of case studies and potential ML application scenarios,
2. assessment and selection of ML application scenarios and
3. final agreement on ML applications scenarios as well as defining of objectives.

### *3.1 Identification of Potential Case Studies*

The aim of this stage is to identify case studies or use cases for the application of ML methods in the respective organization. To achieve these goals, potential users should develop ideas where a software application with the integration of ML technology could be meaningful used in their typical workflow or in the business processes of their organization. The described method consists of three stages to accomplish this goal:

1. **Building a knowledge base:** Future users of the desired ML application must be familiar with the topic of ML and need to have a basic technical knowledge to identify relevant fields and processes for new ML applications. In some cases,

this requires building up a knowledge base of ML technologies and their potential applications for the user in the first place.

a. **Survey of the state of knowledge:** Therefore, the individual knowledge of basic ML concepts and their applications should be assessed. This can be done with the help of a short questionnaire at the beginning of the workshop. The questionnaire contains of the following key questions: "Do you know what ML is?", "Have you already seen/read about ML applications?", "Do you already use ML applications (in a private situation)?", "Do you already use ML applications (at work)?", and "Do you already have ideas for new ML applications for your typical workflow or in business processes of your organization?"
b. **Impulse lecture:** This is followed by a talk on ML methods and their potentially possible applications in common workflows or business processes, taking into account the previous results of the common knowledge level. Prior to this, a research of existing ML applications in the organization's domain was conducted, and the applications found were classified into several dimensions (e.g., task type, see Fig. 2). The talk is structured by these dimensions of common ML tasks (e.g., task type "regression", see Fig. 3) and presents a spectrum of ML technologies and their applications.

2. **Identification of possible ML applications:** The creativity technique of brainwriting [26] is used on the participants of the workshop (employees) to identify and enhance ideas for the application of ML methods to improve typical workflows or business processes. Brainwriting involves participants from the different departments of the organization. Each participant can actively develop and improve the ideas of the others, generating many complementary ideas.

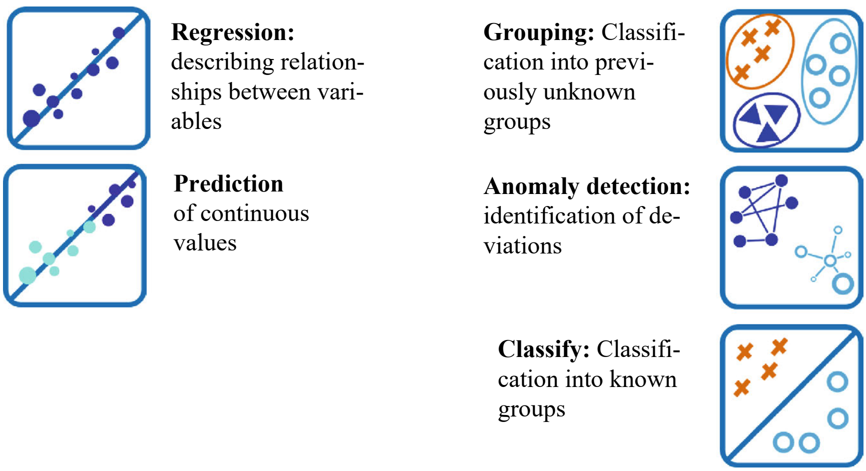



**Fig. 2** Morphological box "type of ML tasks" for classifying ML applications

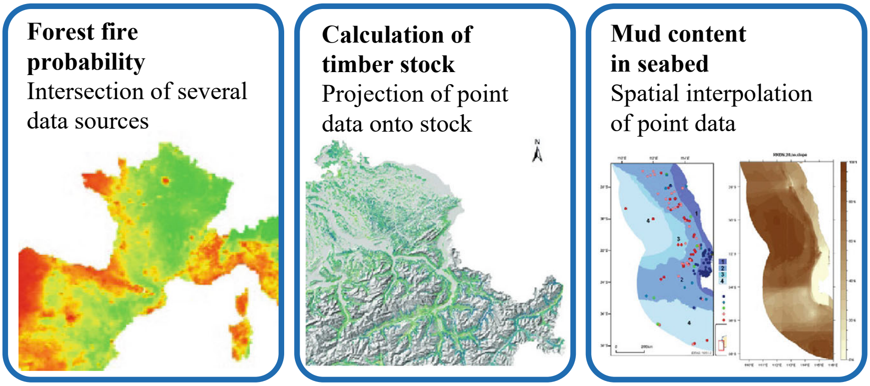



**Fig. 3** Examples of ML applications for the ML task type "regression"

3. **Structuring and discussing ideas:** To structure the identified ideas, they are collected, named one after the other, and put in relation to each other. These ideas were discussed and criticized in the group afterwards. However, the participants of the workshop are not allowed in the first place to evaluate or comment on any of the ideas while the process of "mining" is still running, to keep the focus on gaining new ideas. In the next step, the ideas should be critically discussed using a classification grid (morphological box, see Fig. 2) to filter out ideas that do not focus on ML or are unrelated to the participant's domain or day-to-day work. This requires ML expert knowledge and cannot generally be done by the participants.

## *3.2 Evaluation and Selection of Case Studies and Scenarios for ML Applications*

The aim of this stage is to select case studies and scenarios to be realized as ML applications. This involves two steps:

1. **Evaluation of the identified application scenarios:** The application scenarios obtained from the user's perspective need to be evaluated from both an organizational and a technical perspective (design thinking approach).

    a. Each scenario is evaluated based on "contribution to success" and "need for action" as two-dimensional categories from an organizational perspective (see Fig. 4). The category "contribution to success" relates to the potential benefit for the respective organization (i.e., is the scenario relevant and promising success?), whereas the category "need for action" relates to required effort in a functional and organizational sense (i.e., is there a high potential for benefit or improvement compared to the status quo?). Economic

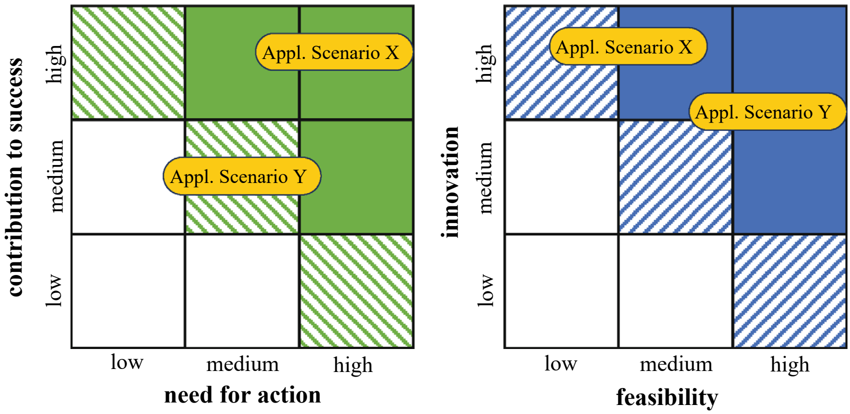



**Fig. 4** Evaluation matrices for the evaluation, prioritization, and selection of application scenarios

efficiency (see Design Thinking approach) is not addressed here, as ML technologies have the potential for disruptive transformation and no reliable cost information can be provided at this early stage of project identification.

b. On the other hand, the "feasibility" in terms of technical realization and implementation is assessed from an ML project management perspective to determine whether the required effort may exceed the expected benefit from an economic point of view.

2. **Prioritization and selection of application scenarios:** Based on the evaluation above, the identified application scenarios can be selected, and prioritized, or eliminated by the potential users.

   a. Application scenarios that are not feasible technically or economically can be excluded from further consideration.
   b. Application scenarios that have been assessed as highly beneficial (and technically feasible) and classified with a high need for action can be selected for further consideration. This way of prioritization of new application scenarios supports the acceptance of the system by the future users.

## *3.3 Assessment of Scenarios and Objectives*

The aim of this stage is to finalize the decision on an application scenario and define the objectives that must be achieved by implementing the scenario using ML technology. This stage consists of the following three steps.

1. **Identification of objectives and requirements for the application scenarios:** The organization must define the requirements and objectives that must be achieved during implementation of the previously selected scenarios or during

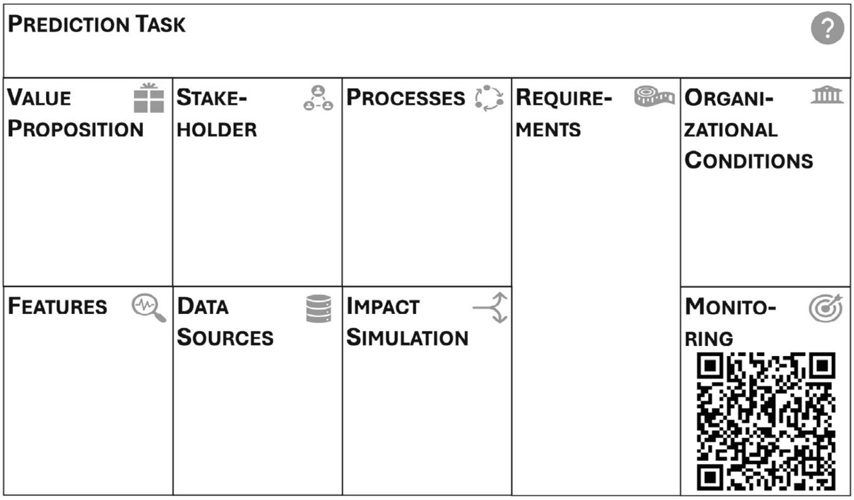



**Fig. 5** ML/Use canvas for analyzing the objectives and benefits of ML application scenarios (simplified for print, available for download at https://github.com/simplex4learningProject/ ML-UseCase-Canvas or by QR-Code)

development and operation of the respective ML application. The canvas is a frequently used method to describe the idea of an application scenario in more detail by developing and presenting all relevant points on one page. For this purpose, an ML/use case canvas based on [27] was developed, which takes up ML-specific points and includes the perspective of analyzing business processes use cases. The following points are queried: "prediction task", "value proposition", "stakeholders", "processes", "requirements", "organizational conditions", "features" (parameters of the learning task), "data sources", "impact simulation", and "evaluation of success". This canvas is discussed and developed with the workshop participants for each selected application scenario (Fig. 5).

2. **Finalization (final definition and selection) of the application scenarios:** In this step, the selected use cases should be finalized, i.e., checked for content, enhanced, completed (checked for completeness and supplemented if necessary), consolidated, and finally selected to be prepared for the following analysis.
3. **Involvement of management:** To obtain their support is the final step.

## 4 Applying the Method to Identify ML Applications

The described method was applied in an environmental agency of a German federal state, and the results are presented below.

The implementation of ML applications to assist and improve the work and business processes is highly relevant to this environmental agency. The two key objectives are as follows: first, employees should become more effective (e.g., routine tasks could be eliminated), and second, the agency aims to make better use of the available (environmental) data (e.g., because relevant environmental data is extensively available and can be used for several business processes).

The analysis to identify potential application scenarios was carried out in three onsite workshops with employees of this agency (see Fig. 6). Each workshop addressed a different part of the method (see Sect. 3). Appropriate instruments (questionnaires, evaluation matrices, canvas, etc.) were created and used at each stage of the process. Workshop participants included 4–5 employees from different departments of the agency.

## *4.1 First Analysis Workshop: Identification and Selection of Potential ML Scenarios*

During the first workshop, several ML application scenarios were developed together collaboratively with the participants. In accordance with the developed method, the steps taken included determining the participants' level of ML knowledge, delivering an impulse lecture on ML applications and their potentials, identifying possible application scenarios in the agency by using brainwriting, and discussing the resulting ideas.

1. **Participants' level of knowledge:** The survey revealed that the participants had a basic understanding of ML and were familiar with a few applications. Some participants had only used ML applications privately and for selected purposes. In a professional context, no ML application was currently used by the surveyed employees of the agency, nor was the introduction of an application imminent. Some ideas for ML applications in a business context were already proposed by the participants at the beginning of the workshop.
2. **Impulse lecture:** As part of the talk, typical ML learning tasks were presented to the participants along the dimension "ML learning tasks" (see Fig. 2), as well as a variety of potential applications in their subject domain (cf. Figure 3).
3. **Brainwriting:** During the workshop, participants from various specialist departments (including air quality, water management, nature conservation, and climate) developed a total of 21 ideas for ML applications in their fields of work in approximately 2 ½ hours.
4. **Discussion:** It turned out that a high number and outstanding quality of the possible ML application scenarios were developed by the workshop participants. The identified application scenarios were classified and structured afterwards. Using the dimension "ML learning tasks" as classifier (see impulse lecture) helped to group similar learning tasks together. A strong interest in application scenarios relating to the automatic generation of texts, which had not been covered in advance by the talk, was expressed by the participants during the discussion. As a result, these application scenarios were also considered in the

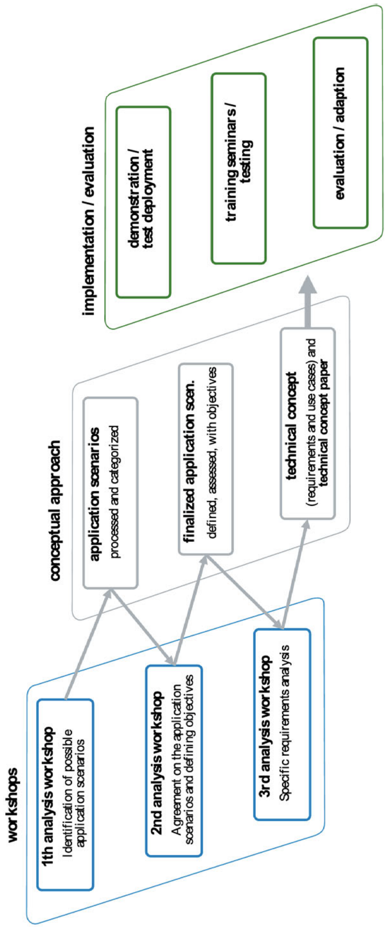



**Fig. 6** Workshop concept

**Table 1** ML application scenarios developed using the brainwriting method (extract)

| No | Brief description | Type of learning task |
|---|---|---|
| 1 | Creation of water level forecasts for various rivers; nationwide in real time | Forecast |
| 2 | Recording of living roofs. Or in general: classification of roofs using aerial photos according to predefined categories | Classify |
| 3 | Prediction of the distribution of known species in an area or prediction of known invasive species | Forecast |
| 4 | Outlier detection in time series with environmental measurement data (e.g., air and water quality) for downstream and further analyses | Anomaly detection |
| 5 | Interpolation of point data for spatial analyses of environmental data (e.g., air and water quality) | Regression |
| 6 | Determination of aquatic organisms in water samples for the identification and quantification of these organisms | Classification |
| 7 | Recognizing photos with data protection-relevant content to anonymize or not publish these photos in the following step when data protection is violated | Classification |
| 8 | Generation of scientific and generally understandable text snippets to support publishing of documents for the agency in the context of public relations work or public statements | Language model |
| 9 | Generation of textual descriptions for charts to support publishing of documents for the agency in the context of public relations work as well as to support accessibility | Language model |
| 10 | Generation of automatic or semi-automatic translations of texts into “easy language” to support the accessibility of documents | Language model |

subsequent analysis. Some of the proposed application scenarios could not be solved using ML methods; these scenarios were excluded from further considerations after the workshop. In Tab. 1 an excerpt of the ML application scenarios developed is shown.

5. **Evaluation of the scenarios:** The previously discussed ML application scenarios were evaluated by the participants using the evaluation matrix in a second step. The aim of the evaluation was to find application scenarios with a high benefit (cf. Figure 4: contribution to success) that also solve a relevant problem (cf. Figure 4: need for action).

### *4.2 Second Analysis Workshop: Agreement on the Application Scenarios and Defining Objectives*

In the second workshop, the application scenarios were finally evaluated and prioritized by the participants and two application scenarios were selected for further action. A comprehensive requirements analysis was carried out for these selected scenarios during the workshop. The second workshop involves the following steps:

1. **Preparation of the ML Application Scenarios:** In preparation of the 2nd workshop, the previously selected ML application scenarios from the 1st workshop were evaluated from the perspective of the (research) project regarding the dimensions of innovation and feasibility/project relevance (cf. Figure 4 right). From the application scenarios that achieved at least a medium rating from the users and researchers, four were selected, and a brief description was created beforehand. These descriptions included the following topics: "description of the application scenario verbatim", "learning task", "data sources", "approach strategy", "obstacles and restrictions", and finally "questions for the agency/organization". This preparation aimed to achieve a common understanding of the selected application scenarios with the users of the considered agency and was sent to them beforehand.
2. **Discussion and final selection of the ML application scenarios:** During the workshop, a common understanding of the ML application scenarios based on the brief description was achieved through discussion with the users. This discussion was crucial for the subsequent final selection of ML application scenarios. From the four remaining application scenarios, the participants selected two that would generate a high benefit (contribution to success) and also solve a relevant problem (need for action). The scenarios with the numbers 3 and 9 (see Tab. 1) have been selected.
3. **Identification of goals and requirements for the application scenarios:** An ML/use case canvas (cf. Figure 5) was used and developed with the workshop participants within 2 h. During the discussion, it became clear that there was a need to modify the ideas from the first workshop, specifically the content of the learning task, although the type of learning task was not altered.

## *4.3 Third Analysis Workshop: Specific Requirements Analysis (Technical Concept)*

The technical concept for the previously selected application scenarios, consisting of requirements and use cases as well as a technical concept, is developed in the 3rd workshop. For this purpose, a specific requirements analysis is conducted with the participants, which includes the following steps:

1. **Preparation of the ML/use case canvas:** The completed canvas from the 2nd workshop was transcribed and analyzed in preparation for the final workshop, i.e., it's checked for completeness, enhanced, consolidated, and prepared for the following analysis. A summary of the two application scenarios is shown in Fig. 7.
2. **Final agreement (final specification of the application scenarios):** A 3rd workshop is being prepared to conduct a specific requirements analysis with the users.

**Scenario: "Predicting the distribution of known species..." (No. 3)**

Benefit: Prediction of the species distribution of diatoms at monitoring sites to optimize the water sampling strategy

Learning task: Prediction of a distribution (diatoms in this case) for monitoring sites (waters sections in this case)

Stakeholder: relevant department can optimize the work tasks of employees; relevant employees may be less burdened

Processes: Affects the existing sampling strategy process

Requirements: Consideration of all monitoring sites covered by Water Framework Directive (WFD); user-friendly UI

organizational conditions: State agency has data sovereignty over data for the relevant monitoring sites ➔ data acquisition is simplified.

Features and data collection: will be finalized during the 3rd workshop.

Impact simulation: influence on the sampling strategy; "wrong" measuring points may be sampled, cost savings.

evaluation of success: A scientific study of the correctness of this new method is necessary. I.e. new sampling strategy vs. old sampling strategy

**Scenario: "Generation of textual descriptions for diagrams ..." (No. 9)**

Benefit: Time savings when creating "generic" texts to describe illustrations (e.g. charts) and increased accessibility. For example, when creating the monthly climate report

Learning task: Training of a specific large language model (LLM) with sample texts from the state agency.

Stakeholder: respective users in the departments who write such texts

Processes: existing dual control strategy remains in place; existing data pipeline (e.g. R scripts) remains in place

Requirements: The language model only verbalizes the pre-selected and pre-evaluated data. There is no image recognition to extract data from a diagram.

organizational conditions: Preservation of data sovereignty and integrity of the IT infrastructure is crucial (on-premises solution)

Features and data collection: Sample texts are available; existing "R scripts" can be used as input.

Impact simulation: Employees can be relieved of the routine creation of texts.

evaluation of success: Expert users are still responsible for the final editorial control of the generated texts. These people also carry out a review.

**Fig. 7** Summary of the transcribed ML/use case canvas

The previously selected and enhanced application scenarios with their elaborated specifications will be finally assessed. To achieve this, the existing business processes and domain procedures will be analyzed (this includes IT infrastructure, application view, and process view). The entire application scenario is broken down into sub-problems, and an ideal process is defined for each sub-problem. This is followed by proposing solutions for each sub-problem and specifying their requirements. The results of this analysis will be summarized in a concept paper (technical concept), where benefits for the test users are also outlined.

3. **Involvement of management:** Support from top management has yet to be obtained.

## 5 Summary and Outlook

The proposed approach outlines the essential steps for identifying ML application scenarios in organizations. These steps are particularly crucial for the successful design, implementation, and realization of an ML solution. This work is motivated by the lack of such a method in existing research. To close this gap, this paper proposes an approach for the systematic identification of ML application scenarios in organizations. This proposed method was successfully applied in a state environmental agency in Germany. The presented method contributes to methodological approaches for planning and designing ML solutions. The authors believe that this method is also applicable to other domains that require a high degree of creativity through user input, while also consideration the contribution to success and the need for action.

With the presented case study, this method has proven its effectiveness in practice, as well as in two other case studies. It has been shown that this method can help develop ML solutions for organizations by focusing, in the first stage of the analysis, on specific ML application scenarios that can support or improve typical workflows or business processes. Additionally, involving future users at an early stage of the process helps to demonstrate the benefits of using ML technology in their daily business. Furthermore, points of contact have been identified for the subsequent introduction of the ML solution in considered organization.

After the final specification of the ML application scenarios is completed, an initial prototype of an ML platform for environmental data will be developed in collaboration with the respective users. Thus, the identified and selected application scenarios will be generalized to enable the transfer of desired ML solution approaches to different organizations. Using the (environmental) data provided by considered organization, a suitable ML architecture will be developed, and ML models will be created for their respective use cases. After a successful evaluation, it's planned to test and evaluate the transferability of these generalized ML solutions with the other project partners form other federal states of Germany.

For further practical use, the method should be expanded to include two aspects: a profitability analysis and a risk analysis. Incorporating a profitability analysis into the potential analysis method helps to assess the financial impact and expected benefits of the identified ML applications, enabling informed decision-making and the efficient use of resources. Integrating a risk analysis into the potential analysis method would help to identify and assess potential problems at an early stage, enabling more informed decision-making and increasing the chances of ML projects being a success.

**Acknowledgements** The content presented is part of the collaborative project simplex4learning ("Intelligent environmental data analysis through automated machine learning for domain users") with Disy Informationssysteme GmbH, simplex4data GmbH, and HTW Berlin, University of Applied Science in cooperation with Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen (LANUV), LUBW Landesanstalt für Umwelt Baden-Württemberg, and Landesbetrieb Forst Brandenburg as associated pilot users. Simplex4learning project is sponsored by Federal Ministry of Education and Research (BMBF) as part of the "KMU-innovativ IKT" funding program under the funding code 01IS23041A-C.

## References

1. Kuutti, S., Bowden, R., Jin, Y., Barber, P., Fallah, S.: A Survey of Deep Learning Applications to Autonomous Vehicle Control. IEEE Trans. Intell. Transp. Syst. **22**(2), 712–733 (2021). https://doi.org/10.1109/TITS.2019.2962338
2. Ahmad, R., Alsmadi, I.: Machine learning approaches to IoT security: A systematic literature review. Internet of Things **14**(100365) (2021). https://doi.org/10.1016/j.iot.2021.100365.
3. Atadoga, A., Sodixa, E., Umoga, U., Amoo, O.: A comprehensive review of machine learning's role in enhancing network security and threat detection. World J. Adv. Res. Rev. **21**(2), 877–886 (2024). https://doi.org/10.30574/wjarr.2024.21.2.0501
4. Ali, O., Abdelbaki, W., Shrestha, A., Elbasi, E., Abdallah Alryalat, M., Dwivedi, Y.: A systematic literature review of artificial intelligence in the healthcare sector: Benefits, challenges, methodologies, and functionalities. J. Innov. Knowl. **8**(1), 100333 (2023). https://doi.org/10.1016/j.jik.2023.100333
5. Barragán-Montero, A., Javaid, U., Valdés, G., Nguyen, D., Desbordes, P., Macq, B., Willems, S., Vandewinckele, L., Holmström, M., Löfman, F., Michiels, S., Souris, K., Sterpin, E., Lee, J.A.: Artificial intelligence and machine learning for medical imaging: A technology review. Physica Med. **83**, 242–256 (2021). https://doi.org/10.1016/j.ejmp.2021.04.016
6. Ahmed, A., Sayed, S., Abdoulhalik, A., Moutari, S., Oyedele, L.: Applications of machine learning to water resources management—A review of present status and future opportunities. J. Cleaner Prod. **441**, Art. 140715(2024). https://doi.org/10.1016/j.jclepro.2024.140715
7. Ghobadi, F., Kang, D.: Application of Machine Learning in Water Resources Management – A Systematic Literature Review. Water **15**(4), Art. 620, (2023). https://doi.org/10.3390/w15040620
8. Mosavi, A., Salimi, M., Faizollahzadeh Ardabili, S., Rabczuk, T., Shamshirband, S., Varkonyi-Koczy, A.R.: State of the Art of Machine Learning Models in Energy Systems – a Systematic Review. Energies **12**(7), Art. 1301 (2019). https://doi.org/10.3390/en12071301
9. Ifaei, P., Nazari-Heris, M., Charmchi, A., Asadi, S., Yoo, C.: Sustainable energies and machine learning—An organized review of recent applications and challenges. Energy **266**, Art. 126432 (2023). https://doi.org/10.1016/j.energy.2022.126432
10. Alkhatib, R., Sahwan, W., Alkhatieb, A., Schütt, B.: A brief review of machine learning algorithms in forest fires science. Applied Sci. **13**(14), Art. 8275 (2023). https://doi.org/10.3390/app13148275
11. Attri, I., Awasthi, L., Sharma, T., Rathee, P.: A review of deep learning techniques used in agriculture. Ecological Info. **77**, Art. 102217 (2023). https://doi.org/10.1016/j.ecoinf.2023.102217
12. Benos, L., Tagarakis, A.C., Dolias, G., Berruto, R., Kateris, D., Bochtis, D.: Machine learning in agriculture—A comprehensive updated review. Sensors **21**(11), Art. 3758 (2021). https://doi.org/10.3390/s21113758
13. Vaishnavi, V.K., Kuechler, W.: Design science research methods and patterns—Innovating information and communication technology, 2nd edn. CRC Press, Boca Raton (2015).
14. Royce, W.: Managing the Development of Large Software Systems. In: Proceedings of IEEE WESCON, 26, pp. 1–9. (1970).
15. Friedrich, J., Hammerschall, U., Kuhrmann, M.: Das V-Modell XT – Informatik im Fokus. Springer, Heidelberg (2009).
16. Schwaber, K.: Agile Software Development with Scrum. Prentice Hall (2001). https://dl.acm.org/doi/10.5555/559553

17. Wirth, R., Hipp, J.: CRISPDM—Towards a standard process model for data mining. In: Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining (4), pp. 29–40. http://cs.unibo.it/~danilo.montesi/CBD/Beatriz/1.1.198.5133.pdf (2000).
18. Studer, S., Bui, B., Drescher, C., Hanuschkin, A., Winkler, L., Peters, S., Müller, K.-R.: Towards CRISP-ML(Q)—A Machine learning process model with quality assurance methodology. Mach. Learn. Knowledge Extraction **3**(2), 392–413 (2021). https://doi.org/10.3390/make3020020
19. Hasterok, C., Stompe, J., Pfrommer, J., Reiter, S., Ziehn, J., Usländer, T., Weber, M., Riedel, T.: PAISE—Das Vorgehensmodell für KI-Engineering. Fraunhofer IOSB, Karlsruhe, (2021). https://doi.org/10.24406/publica-fhg-301357
20. Eisenträger, M., Frey, C., Herzog, A., Moghiseh, A., Morand, L., Pfrommer, J., Stephani, H., Stoll, A., Wessels, L.: ML4P—Vorgehensmodell Machine Learning for Production, Fraunhofer IOSB, Karlsruhe (2022). https://doi.org/10.24406/publica-485
21. DIN: Projektmanagement – Projektmanagementsysteme – Teil 2: Prozesse, Prozessmodell (DIN 69901–2:2009–01). Deutsches Institut für Normung, Ausgabe 2009–01 (2009). https://www.din.de/de/mitwirken/normenausschuesse/nqsz/veroeffentlichungen/wdc-beuth:din21:113428357
22. Axelos Ltd.: Managing Successful Projects with PRINCE2. 6th edn. The Stationery Office, Norwich (2009).
23. Hofmann, P., Jöhnk, J., Protschky, D., Stähle, P., Urbach, N., Buck, C.: KI-Anwendungsfälle zielgerichtet identifizieren. Wirtschaftsinformatik & Management **12**(3), 184–193 (2020). https://doi.org/10.1365/s35764-020-00257-z
24. Brakemeier, H., Gerbert, P., Hartmann, P., Liebl, A., Schamberger, M., Waldmann, A.: Applying AI—How to find and prioritize AI use cases. appliedAI Initiative, München (2020). https://www.appliedai.de/assets/files/How-to-identify-and-prioritize-AI-use-cases_2024-03-01-120504_yrme.pdf
25. Röhler, M., Haghi, S.: Leitfaden Künstliche Intelligenz – Potenziale und Umsetzungen im Mittelstand. VDMA, München (2020). https://publica.fraunhofer.de/handle/publica/300499
26. Rohrbach, B.: Kreativ nach Regeln – Methode 635, eine neue Technik zum Lösen von Problemen. Absatzwirtschaft **19**(12), 73–76 (1969).
27. Dorard, L.: From Data to AI with the Machine Learning Canvas (Part I)—A framework to connect the dots between data collection, machine learning, and value creation. Medium (2016). https://medium.com/louis-dorard/from-data-to-ai-with-the-machine-learning-canvas-part-i-d171b867b047. Last accessed 17 March 2025

# A Geo-Parser for German Documents with Environmental Context

**Nicolas Doms, Thorsten Schlachter, and Lisa Hahn-Woernle**

**Abstract** Environmental information is often bound to some geographic entity, be it continent, country, city, or a smaller entity like a forest or a water body. Consequently, documents with environmental context also contain geographic entities. A geo-parser can help to understand what geographic information is present in the document. This information can then be used to display the geographic entities on a map, to group related data as a facet in an internet search, or to enable links between these documents to other documents that refer to the same geographic entity.

There have been numerous geo-parsers in the past, however, none of them dealt explicitly with German documents with environmental context. This scenario features a number of challenges that will be explained before a solution is proposed in this publication. As the geo-parser requires some sort of a reference dataset with geographic names and geographic areas, different datasets are analyzed before the most fitting one is picked for the implementation. Furthermore, an evaluation dataset containing pre-tagged geographical entities is created and presented briefly, before the proposed solution is evaluated against the very same dataset.

**Keywords** Geo-Parsing · Machine Learning · Natural Language Processing

N. Doms (✉) · T. Schlachter
Institute for Automation and Applied Informatics (IAI), Karlsruhe Institute of Technology (KIT), Hermann-Von-Helmholtz-Platz 1, 76344 Karlsruhe, Germany
e-mail: nicolas.doms@kit.edu

T. Schlachter
e-mail: thorsten.schlachter@kit.edu

L. Hahn-Woernle
Center of Excellence for Environmental Informatics, Baden-Württemberg State Institute for the Environment, Survey and Nature Conservation, Grießbachstraße 1, 76185 Karlsruhe, Germany
e-mail: lisa.hahn-woernle@lubw.bwl.de

 
V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_2

## 1 Introduction, Motivation, and Structure of This Publication

The German state Baden-Wuerttemberg has initialized a digitalization strategy called "digital.LÄND" with associated research projects, including the project "Umweltinformationen digital 4.0" [1]. The latter has numerous goals in the sector of environmental information, including the enhancement of metadata to increase search quality, a chatbot specialized in questions regarding the Rhine meadows of Karlsruhe, and a mobile application that enhances the learning trail of said meadows by augmented reality experiences.

Tackling the challenge of increasing search quality by enhancing metadata, it can be stated that for documents or websites with environmental context, semantic, geographic, and time information are present for most of the cases. The challenge of extracting time information has been introduced in an earlier paper [2], and the semantic information is being extracted by using an already existing service (sns.uba.de). Thus, the challenge of extracting geographic information remains.

The challenge of finding geographic information in text and resolving it into an explicit geometry, also known as geo-tagging, is a challenge that was repeatedly faced by numerous authors in the past [3, 4]. However, for documents with environmental context, this challenge adds a new layer of difficulty as geographic information does not only include cities, countries, and so on, but also mountains, water bodies, forests, field names, and more. Reading through the first paragraph of this chapter, one can find the expression "Rhine meadows of Karlsruhe" which most geo-parsers would probably resolve into the city of "Karlsruhe" with its corresponding geometry. However, the meadows that are being mentioned have a different geographic position that is not included in the geometry of Karlsruhe city.

In order to correctly identify geographical information in an environmental context, we will develop a geo-parser that can identify different types of geographic information and tag them with the appropriate geometry. This parser will be specialized for German documents, but the principles apply to documents of any language.

This geo-parser can find potential applications in different use cases. For example, the results page of an enterprise or internet search can display information as a map, if the geometries determined by the geo-parser are saved in the corresponding metadata fields of the search engine index.

When presenting documents in a search results page, spatially related documents can be presented as a common facet, allowing the user to refine the query.

Since to our knowledge, no pre-tagged German datasets that include environmental information exist, we will also build up our own dataset for evaluation purposes. This dataset does not only contain geographic tags, but also TIMEX3 annotations [5] and will be published in the future.

Concerning the structure of this publication, Sect. 2 will show important considerations regarding the technical architecture of the geo-parsers. These include the introduction of different reference datasets as well as the challenges when dealing with German text documents, resulting in a few special additions to the geo-parser later.

Section 3 covers the dataset used for the evaluation of this geo-parser. As this is subject to another publication, this introduction is kept rather short, but all the important steps from building up the dataset to the structure of the resulting dataset are explained.

The technical architecture and implementation of the geo-parser is covered in Sect. 4. This also includes an explanation of the comprehensive classification service (explained later), a short exploration of the chosen reference dataset, as well as an overview of the tagging logic.

The performance of the geo-parser is then evaluated in Sect. 5. Expressions that the parser cannot currently tag correctly are also analyzed before Sect. 6 finally summarizes the results and highlights implications for the future.

## 2 Considerations and Technical Implementation

This section is subdivided into three subsections that deal with different design aspects of the geo-parser. Subsection 2.1 compares different datasets that are available for usage. These datasets contain the geographic information needed to display the geographic entity on a map, as well as the names associated with it.

Subsection 2.2 shows some challenges that have been faced while building up the parser. Some of these are language specific, others might also be present in other languages.

### *2.1 Choosing a Gazetteer*

Geo-tagging requires some sort of a lookup table, known as a gazetteer, that contains the geographic names that need to be recognized as well as the corresponding geographic information. A single entry in a gazetteer is also known as a toponym, representing a geographic entity in the real world. For the German language, a number of datasets are available that feature different advantages and disadvantages. Some of them will now be presented.

#### 2.1.1 GeoNames

GeoNames is a geographical database that contains over 25 million geographical names from all over the world. The data originates from different sources, mostly websites of governmental nature. For Germany, this includes the German Federal Agency for Cartography and Geodesy which is also mentioned later in this paper. This dataset is licensed under the Creative Commons Attribution 4.0 License [6].

The data is split up into different countries, where each country has its own text file (.txt). There are almost 200,000 geographical names for Germany—including forests, bays, water bodies, etc. However, this dataset only contains the coordinates of the center of the corresponding entity, not the shape of the area itself. Therefore, the corresponding toponym can not be displayed with its whole area on a map.

### 2.1.2 OpenStreetMap

OpenStreetMap is a project established in 2004. The project tries to build up a map of the entire world, including houses, water bodies, railways, etc., by letting its users contribute and cross-validate its data. Currently, there are 8 million user accounts registered on OpenStreetMap. The data itself is usable under the Open Database License 1.0 [7].

While this dataset is vast and freely usable, it comes in its own data format (.osm). To use it in a different application, this data needs to be opened and queried in a tool like Osmium before being exported in a more standardized data format.

### 2.1.3 Gn250

The German Federal Agency for Cartography and Geodesy (Bundesamt für Kartographie und Geodäsie, BKG) offers various geographical data services. This includes the dataset "Geographische Namen 1:250 000" (engl. "Geographical Names 1:250 000"). It is licensed under the Open Data policy [8].

This dataset contains over 160,000 geographical names with the corresponding shapes of the area as boundary boxes. Also included is a categorization of the geographical name (water bodies, railways, forests, cities, …) and a hierarchical structure for cities, districts, states, etc. As this data is authoritative, a high precision can be expected.

The data is available for download in either shape or.csv format and can therefore be consumed directly after downloading.

### 2.1.4 Conclusion for Reference Dataset

All three presented datasets come with their own advantages and disadvantages. While GeoNames is very extensive and easily accessible, it does not include the necessary geographic area data to display the corresponding entities on a map, eliminating it as a candidate for this geo-parser.

OpenStreetMap has a high user count, which ensures both a high actuality of the data as well as a good precision because of the cross-validation.

The GN250 dataset contains exactly what is required for this geo-parser, already providing a good categorization and hierarchical structure in the data. Made available as either shape or.csv file ensures that this data can instantly be used and updated.

Both of these datasets (GN250 and OpenStreetMap) are strong candidates for this geo-parser with their very own smaller advantages and disadvantages. In the end, GN250 has been chosen over OpenStreetMap as its authoritative nature ensures a higher data consistency and therefore quality over the crowd-sourced OpenStreetMap data.

### *2.2 Challenges of Identification of Geographic Entities*

When trying to identify geographic entities in a text, a number of challenges arise.

One challenge which is not language specific are homonyms. For example, there exists a village called “Luft” in Germany. However, “Luft” also means “air” in German. When analyzing text that contains the expression “Luft”, the geo-parser must decide if the semantic meaning is the village or something different.

Another non-language-specific challenge are different geographic entities that share a name. An example would be “Karlsruhe”, which can be a city or a district (or even an administrative district, a larger structure than a district). “Neustadt” is a geographic name that, without further context, appears over 40 times in Germany. Resolving these names into the correct geographic entity will certainly pose another challenge.

As the chosen language for this geo-parser is German, one of the language-specific challenges includes the different word endings depending of the case used in the sentence. For example, the geographic entity “Schwarzwald” (engl. “Black forest”) might appear as “Schwarzwalds”, “Schwarzwaldes”, or even “Schwarzwälder” in the text, depending on the context. As all of these point to the same geographic entity, the corresponding expression needs to be reduced to its stem form before trying to assign the geographic tag.

## 3 Evaluation Dataset

As mentioned in Sect. 1, to our knowledge, a fitting dataset that can be used for evaluation currently does not exist for this use case. Therefore, we had to build it up ourselves.

Since this dataset does not only contain geographic tags, but also time tokens annotated by TimeML guidelines, and as this dataset features a unique subject as it contains only documents with environmental context, the dataset will be fully explained in a future publication. This dataset will also be extended in the future to provide an even larger basis for the evaluation of models that extract and/or normalize time and/or geographic information.

However, as this dataset is still being used for the evaluation of this geo-parser, we will briefly present what kind of documents it consists of in Subsection 3.1. In

Subsection 3.2, the method of annotation is explained before Subsection 3.3 shows the structure of the resulting evaluation dataset that has been used for this parser.

### *3.1 Publication Service (PUDI)*

The "Publikationsdienst" (engl. "Publication Service"), or in short PUDI, is the publication service of the Baden-Wuerttemberg State Institute for the Environment. It contains about 3000 publications dealing with different environment topics, including topics like water, air pollution, or radioactivity [9].

These documents come in different shapes and formats: some are posters that show different information on the corresponding topic mostly visually, some are official reports containing around 200 pages of facts and figures, and some might be smaller flyers that contain local information for a specific region.

Most of the 3000 publications are accessible as a PDF. The documents are supplemented by some metadata entered manually by the respective editors, e.g. the date of publication and the environmental topics to which it relates.

### *3.2 Annotation of the Dataset*

To help with the annotation of the dataset, we have implemented a small tool. It consists of a graphical user interface where a directory of documents can be displayed. A single document can then be selected and annotated. Figure 1 shows the user interface of this tool.

The tool allows the annotation of both geo-tags, requiring the selection of a geographic entity of the aforementioned reference dataset GN250, as well as the annotation of a TIMEX3 expression. Validation logic checks if the tags are well-defined.

To indicate a geographic entity in the text, the tool adds inline XML tags as per SpatialML definition [11]. These are similar to the TIMEX3 expression logic: geographic expressions are surrounded by a leading < PLACE > and a closing < / PLACE > tag. The leading < PLACE > tag contains attributes describing the exact geographic entity that the geographic expression is referencing.

In order to provide a dataset of sufficient quality, all documents are individually annotated by three different human annotators. The proposed annotations are then compared, discussed, and finally harmonized.

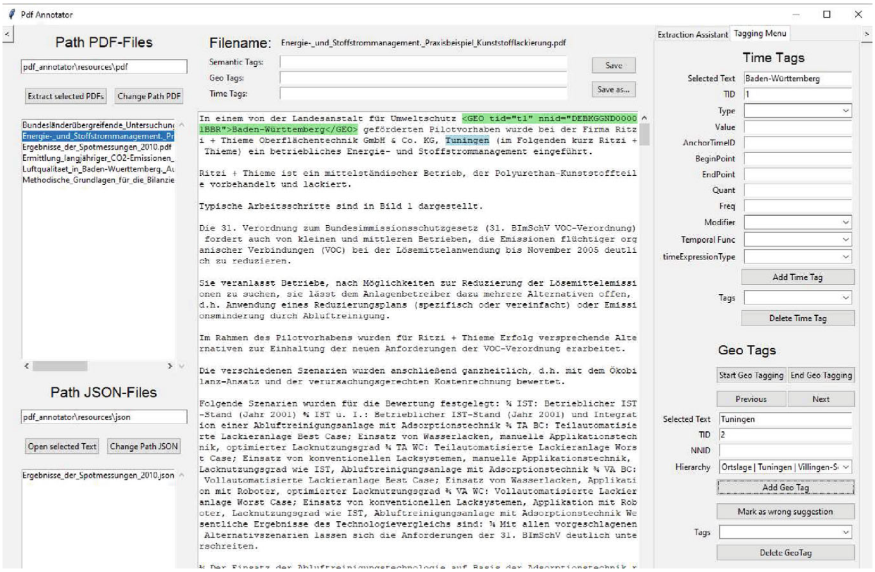



**Fig. 1** User interface annotation tool

## *3.3 Structure of the Evaluation Dataset*

As mentioned before, for the evaluation of the geo-parser, only an excerpt of the final dataset has been used so far. However, as this dataset already contains a sufficient amount of different geographic expressions, this is a good indication of how the geo-parser performs.

To put it into numbers, the geo-parser has been evaluated on a set of 10 documents. These documents contained a total of around 600 annotated geographic expressions.

Depending on the environmental topic that the document covered, the annotated tags have a different composition. For example, if the document deals with water or water quality, tags with rivers, seas, and other environmental categories have been tagged about as often as cities, districts, and other structures. For the topic of air or air pollution, a disproportionally high number of cities have been tagged, while there are only some tags with environmental entities like forests or parks.

When the 10 documents for the evaluation of this geo-parser have been assembled, documents with topics that promote the usage of tags of the environmental category have been preferably chosen over topics that presumably not feature many of these tags in order to accurately test the capabilities of this parser.

# 4 Technical Architecture and Implementation

As previously mentioned, this geo-parser is meant to be a part of a classification service that also recognizes and normalizes semantic (environment-specific) and temporal expressions. Therefore, the question of what architecture will be used to implement this geo-parser has already been answered. This is swiftly covered in the upcoming Subsection 4.1.

Afterward, Subsection 4.2 covers the structure of the reference dataset as well as data preparation steps in order to resolve some of the challenges mentioned in Subsection 2.2, before the final tagging logic is presented in Subsection 4.3.

## *4.1 Technical Architecture*

The geo-parser has been implemented as part of a Java-based microservice, which is accessible via a REST interface. A user can place a POST-request at the service and optionally split up the document into different parts (title = title of document or website, body = content of document or website, path = e.g. URL of website or breadcrumbs of webpage/document, abstract = short abstract of document) which all receive a different weight factor. These weights can be provided by the user and will receive a default value if the user does not provide them. The classification service will then run all three subservices and return each classification result separately as part of a JSON document. Figure 2 shows what this request would look like.

This architecture allows the usage of the classification service from anywhere. For example, a web crawler can use it to enrich crawled documents with additional metadata, the web front end can then use the very same classification service to classify search requests from users. Afterward, the results are compared.

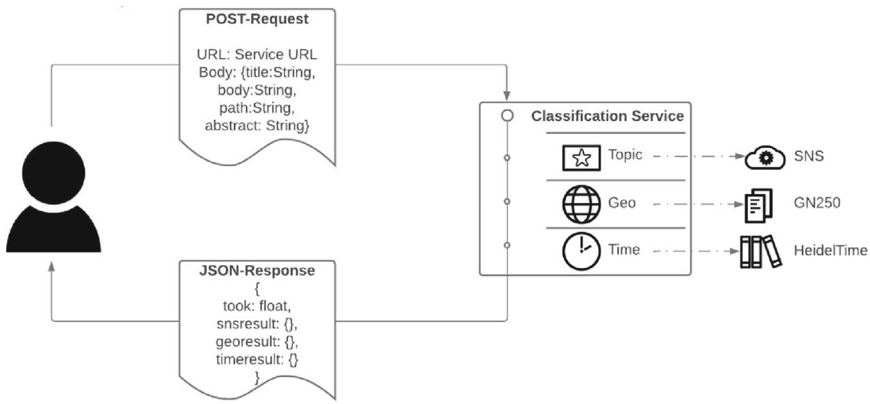



**Fig. 2** Interaction with the classification service

The classification service can also be used by the editors of the publication service (explained in Subsection 3.1) to provide some proposals for the metadata. The editors can then either accept the proposed metadata tags or exchange them for other standardized tags. This will leverage the comparability between different systems and allow cross-reference between them.

## *4.2 Structure of the Reference Dataset and Data Preparation*

As mentioned in Subsection 2.1, the GN250 dataset also comes with a categorization of the geographic entities. For cities, districts, and so on, these entities naturally come with a hierarchy. For example, districts are made up of several cities, a state consists of multiple districts, and so on. Figure 3 shows the hierarchical structure.

There are more categories in the GN250 dataset than displayed here—these are the categories for environmental entities like forests, water bodies, etc. These are summarized as "Sonstige" (engl. "Others") in Fig. 3.

In order to mitigate the language-specific challenge of different word endings, the Stanza library [10] with the German language model is used in this geo-parser. This library allows the usage of "Pipelines", which apply different transformations to the text that is being inserted. We will use the "Lemmatization" module to transform each token into its corresponding lemma, or base form.

**Fig. 3** Hierarchical structure of the categories

### 4.3 Tagging Logic

We use a rule-based approach to determine the correct geographical entity. To help understand the logic, we will assume our reference dataset has three entries: “Frankfurt am Main”, “Frankfurt an der Oder”, and “Heidelberg”. We will show what happens with these entries over the course of this subsection.

To speed up the lookup process, we will extract the first words of all geographical entities and put them into a dictionary as keys, adding all their possible values as a list of strings to their corresponding key. The value-list is then sorted descending by the number of tokens. We now have the dictionary *geoFirstWords* with the following key-value-pairs:

- “Frankfurt” → [“Frankfurt an der Oder”, “Frankfurt am Main”],
- “Heidelberg” → [“Heidelberg”].

The parser then looks up every normalized word of the text in *geoFirstWords*. If a key is found, the logic displayed in Fig. 4 applies. “Hierarchical Matching” means that every value from the corresponding list of strings is matched against the text, starting with the value with the most tokens. So if the parser finds the key “Frankfurt”, it first builds up a string with “Frankfurt” and the next three tokens and compares it to the value “Frankfurt an der Oder”. If this is not successful, it concatenates “Frankfurt” with the next two tokens and compares it to “Frankfurt am Main” and so on.

If no single matching value is found, the most recently tagged geographic entity is chosen and a distance function to all possible values is calculated. Then, the minimum of these values is chosen as tag. In the possible case that there are no prior tags (for

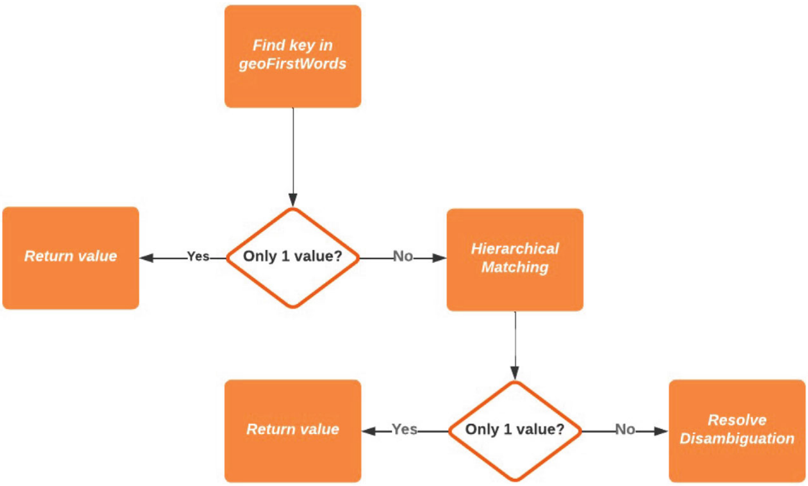



**Fig. 4** Tagging logic

example at the beginning of the document), this tag is marked and resolved with the next fully resolved geographic tag.

In case the disambiguation cannot be resolved because there are multiple entities with the same tokens, but in different hierarchical positions (for example: “Karlsruhe [City]” and “Karlsruhe [District]”), the surrounding area of the expression is searched for key words that help in resolving this conflict (for example: “Landkreis” [engl. “District”]). If no such words are found, the lowest hierarchical level is chosen. In this example case, “Karlsruhe [City]” would be chosen over “Karlsruhe [District]”.

## 5 Evaluation and Current Challenges

As mentioned in Sect. 3, a dataset of 10 documents with around 600 tags has been prepared for the purpose of evaluating this geo-parser. The evaluation is performed on the raw text of the documents, and the annotated tags from this geo-parser are then compared to the original annotation.

Overall, this geo-parser reaches an F1-Score of around 0.88. While this is a satisfactory result, it is worth looking at what is keeping this parser from performing even better.

From the challenges mentioned in Subsection 2.2, the geo-parser seems to have next to no issues with both the challenge where the same geographical name can refer to different geographic entities as well as the language-specific challenge with different word endings. It does however struggle with homonyms.

Currently, possible homonyms have been extracted by comparing the GN250 reference dataset to an OpenThesaurus dictionary of German words. Expressions that exist in both datasets have been extracted, annotated manually, and fed to a neural network that tries to classify whether the presented expression is a geographical entity or not. However, these expressions seem to be too difficult to accurately predict by the neural network, resulting in wrong predictions by the geo-parser.

Another challenge that has not been mentioned before are composite words that are typical for the German language. For example, the expression “branch of the Rhine”, where a geo-parser would most certainly identify “Rhine” as a geographic entity, translates to “Rheinarm” in German. Splitting up these composite wordss would help in identifying the geographic entity, but at the same time, it would create even more possible homonyms that would need to be predicted by the model mentioned in the previous chapter. As long as this model does not improve, it is probably not advisable to do so.

Lastly, this geo-parser is heavily dependent on the quality of both the OpenThesaurus dictionary (as this is needed to identify possible homonyms) as well as the GN250 dataset. Missing data in the prior may result in markups of non-geographic expressions as geographical entities by the parser, while missing data in the latter will lead to no markups at all for these specific expressions.

## 6 Conclusion and Future Work

This paper has shown what kind of geographical information is included in documents with environmental context and how a geo-parser specifically for this use case can be implemented.

As this geo-parser has been implemented for the German language specifically, in addition to challenges like homonyms and shared names across different geographic entities, we also had to tackle language-specific challenges like getting the base form of a word in order to maximize the recognition of the geo-parser. Different reference datasets with geographical names have been analyzed, and the presumably best one for this use case has been chosen.

After building up a dataset that can be used for the evaluation of this and other geo-parsers, the architecture of this parser has been aligned with existing infrastructure in order to assure a smooth implementation. The reference dataset has been explored, and finally, the tagging logic itself has been implemented and evaluated.

As shown in the evaluation section, this parser can recognize all sorts of geographic entities as featured in the GN250 dataset at a satisfying accuracy. These entities are recognized with their corresponding area, enabling the display in maps, the comparison of areas, or even a geographical drill-up into the next largest hierarchical structure.

Regarding the evaluation dataset, the number of documents and therefore the number of tokens will be increased. There will be a separate paper explaining the annotation process and final results in detail in the near future. As mentioned earlier, this dataset features both geographical as well as temporal expressions for German documents, specifically for those with environmental reference.

The environment domain has been chosen specifically as a high occurrence of toponyms which are not city names (like forests, rivers, other points of interest in nature) has been expected for this domain, posing a bigger challenge for the geo-parser. However, these results can be applied to other domains. This might require additional domain-specific information in the gazetteer (for example: power plants for the energy domain), but should work out-of-the-box for a lot of German toponyms.

As Large Language Models (LLM) become increasingly popular for natural language processing tasks, an LLM-based approach might also perform well on the recognition of geographic names in text. However, it should be noted that a gazetteer is still required to resolve the recognized geographic names into their corresponding toponyms.

The geo-parser itself will be used in the classification service shown in Fig. 2.

## Literature

1. Hahn-Woernle, L., Schillinger, W., Schlachter, T., Doms, N., Trefzger, M., Schlegel, T., Wolf, A., Preiß, A.: Umweltinformationen digital 4.0. In: Umweltinformationssysteme – Vielfalt. Offenheit, Komplexität, pp. 203–205. Springer, Wiesbaden (2022)

2. Doms, N., Schlachter, T., Hahn-Woernle, L.: Graph-based time token recognition using graph neural networks and stanza library. In: EnviroInfo 2023—Complete Volume, pp. 135–142. Gesellschaft für Informatik e.V., Garching (2023).
3. Brunsting, S., De Sterck, H., Van Sprundel, T.:GeoTextParser: High-Precision location tagging of textual documents using a natural language processing approach (2016). arXiv:1601.05893
4. Lieberman, M., Samet, H., Sankaranarayanan, J.: Geotagging with local lexicons to build indexes for textually-specified spatial data. In: 2010 IEEE 26th International Conference on Data Engineering (ICDE 2010), pp. 201–212. Long Beach, CA, USA (2010).
5. Pustejovsky, J., Lee, K., Bunt, H., Romary, L.: ISO-TimeML: An international standard for semantic annotation. In: Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10). European Language Resources Association (ELRA), Malta (2010).
6. GeoNames: http://geonames.org/. Last accessed 15 May 2024
7. Curran, K., Fisher, G., Crumlish, J.: OpenStreetMap. Int. J. Interact. Commun. Sys. Technol. **2**(1), 69–78. IGI Global, Hershey, PA, USA (2012).
8. Geographische Namen 1:250 000, https://gdz.bkg.bund.de/index.php/default/open-data/geographische-namen-1-250-000-gn250.html. Last accessed 15 May 2024
9. Publikationsdienst – Landesanstaltung für Umwelt Baden-Württemberg, https://pudi.lubw.de/. Last accessed 15 May 2024
10. Qi, P., Zhang, Y., Zhang, Y., Bolton, J., Manning, C.: Stanza: A python natural language processing toolkit for many human languages. In: CoRR abs/2003.07082. Computing Research Repository (2020).
11. Mani, I., Hitzeman, J., Richer, J., Harris, D., Quimby, R., Wellner, B.: SpatialML: Annotation scheme, corpora, and tools. In Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08). European Language Resources Association (ELRA), Marrakech, Morocco (2008).

# Low-Cost Air Quality Sensor Nodes in a Network Setup: Using Shared Information to Impute Missing Values

**Theodosios Kassandros, Evangelos Bagkis, and Kostas Karatzas**

**Abstract** Low-Cost Air Quality Sensor Nodes (LCAQSN) are being widely deployed across numerous cities worldwide as a new way for assessing air quality. Despite facing challenges related to accuracy and consistency, these nodes offer valuable insights, substantially reducing the costs associated with monitoring air pollutants. A significant hurdle to address is the occurrence of missing values. In this study, we hypothesize that a network of LCAQSN within the same urban environment can effectively retain and utilize shared information to accurately impute missing values, even in cases with substantial gaps in the time-series data of individual nodes. Employing various Machine Learning techniques, our analysis reveals that a network comprising 26 LCAQSN in the Greater Thessaloniki Area, Greece, with 40.93% missing values, can achieve an imputation accuracy of 0.7 $R^2$ on a simulated test set of 10% of missing values. These findings exhibit great promise and unveil numerous opportunities for leveraging LCAQSN networks further, including data fusion and downscaling applications.

**Keywords** Air Quality · Sensors · Imputation · Machine Learning

## 1 Introduction

Air pollution is a persistent environmental problem presenting a significant challenge in urban environments, impacting millions of people worldwide. In 2019, outdoor air pollution was estimated to be responsible for approximately 4.2 million premature

T. Kassandros (✉) · E. Bagkis · K. Karatzas
Environmental Informatics Research Group, School of Mechanical Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece
e-mail: tkassand@physics.auth.gr

E. Bagkis
e-mail: mpagkise@meng.auth.gr

K. Karatzas
e-mail: kkara@meng.auth.gr



V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_3

deaths [1]. Much of this mortality is specifically attributed to fine particulate matter, i.e. to those particles with a mean aerodynamic diameter of less than 2.5 μm ($PM_{2.5}$), penetrating deep into the lungs and posing severe health risks. Despite significant efforts that reduced mortality by 41% from 2005 to 2021, premature deaths in the EU attributable to $PM_{2.5}$ remained alarmingly high at 253,000 in 2021 [2]. The persistence of these high mortality rates underscores the urgent need for innovative solutions in air quality monitoring.

Data on air quality (AQ) are routinely collected from official, ground-based monitoring stations. These networks, commonly developed and managed by national environmental agencies, provide robust and precise information on AQ concentration levels. In Europe the European Environment Agency (EEA) is an organization of the European Union (EU) responsible, among others, for delivering data to support Europe's environment and climate goals. On this basis, the EEA oversees a network comprising over 4000 stations operating in European countries, dedicated to monitoring the exceedance of AQ limits [3], as stipulated by EU directives [4], to protect human health and equip decision-makers with the necessary data to implement effective mitigation strategies. Although these reference AQ stations deliver precise readings, their associated high costs and intensive maintenance demands restrict the number of measurement points. This limitation makes it difficult to achieve a comprehensive representation of AQ values across diverse areas within urban environments, highlighting the necessity for more accessible and versatile measurement methods [5].

In response to these limitations, the past few decades have seen the emergence of Low-Cost Air Quality Sensor Nodes (LCAQSN), designed to fill the gaps left by sparse reference station networks [6]. These sensors not only offer real-time measurements at a reduced cost but also enhance spatial coverage, accessing areas previously beyond the reach of conventional methods. However, despite their benefits, low-cost sensors frequently exhibit shortfalls in measurement quality, a concern thoroughly documented in literature [7]. To mitigate these accuracy issues, a range of studies have proposed various calibration techniques, employing in-situ, data-driven approaches [8] and capitalizing on advances in Machine Learning [9].

Although extensive research has been conducted on calibrating and enhancing the accuracy of LCAQSN, addressing the problem of missing values has received relatively little attention. LCAQSN networks frequently face problems such as communication failures, power shortages, and malfunctions, including sensor faults and network disruptions. These challenges often lead to significant gaps in the time-series data, complicating the maintenance of a sufficient number of nodes in the network. This is necessary for applying methods which enhance the spatial resolution of air quality mapping, like data fusion techniques [10].

Handling missing values is a critical step in time-series analysis and the construction of predictive models, and extensive research has been conducted in this area. Machine learning has proven to be a powerful tool for such tasks, with methods such as k-NN, MICE, Random Forest, and XGBoost being the most prevalent in the field [11]. In a comparative study, Shaadan and Rahim concluded that Kalman filters combined with ARIMA modeling are the best methods for imputing $PM_{10}$ values, achieving an $R^2$ of 0.75 for datasets with 20% missing values [12]. Agbo et al. proposed a novel technique called Best Fit Missing Value Imputation (BFMVI), based on the k-Means algorithm, for imputing missing values from an Internet of Things air quality sensor network. Their study demonstrated that the imputed dataset provided better calibration [13]. Additionally, several studies have utilized networks of sensors to address missing values: one used Gaussian Processes [14], another employed multiple techniques such as k-NN, MICE, and missForest, and a third applied Deep Matrix Factorization [15].

The present study aims to utilize the shared information within an LCAQSN network to impute missing values from sensors based on data from operational ones. We hypothesize that in a given area, the information from functioning nodes can be used to estimate the values at locations with missing data. To test this hypothesis, we develop a spatio-temporal model and compare various machine learning methods.

## 2 Materials

Data were collected from a network of 26 LCAQSN, as part of the KASTOM project in the Greater Thessaloniki Area (GTA) [Fig. 1; http://app.air4me.eu/city/thessaloniki/sensors]. KASTOM is an operational platform designed for the systematic monitoring and prediction of air quality. It utilizes, among other technologies, an Internet of Things (IoT) infrastructure to deliver relevant information to stakeholders and the general public. The GTA is the largest urban area in Northern Greece, home to over 1,000,000 residents and characterized by a Mediterranean climate with warm, dry summers and cold, wet winters, interspersed with cold air masses [16]. From 2009 to 2022, there was only one year in which the pollution levels did not exceed the EU limits for $PM_{10}$, while exceedances of other major pollutants have continued to be observed.

Each node within the network measures particulate matter concentrations across three size categories—coarse ($PM_{10}$), fine ($PM_{2.5}$), and ultra-fine ($PM_{0.1}$)—utilizing an optical particle counter sensor (Manufacturer: Plantower, Model: PMS5003). Meteorological parameters such as temperature (T), relative humidity (RH), and barometric pressure (P) are also monitored, employing sensors from Bosch Sensortec (Fig. 2).

The LCAQSN network is communicating the measurements through a LoRaWAN (Long Range Wide Area Network) protocol. The focus of this study is on imputing missing values for $PM_{2.5}$ during the period from April 1, 2020, to December 31, 2022. During this time, a significant portion of the data was missing, with 40.93% unrecorded as indicated by individual node analysis (Fig. 3).

It is evident that missing values in the dataset primarily arise from three distinct factors. Firstly, the complete failure of the information system results in the absence of data, rendering any imputation efforts unfeasible due to the total lack of available information. Secondly, delays in deployment or premature removal of sensors create

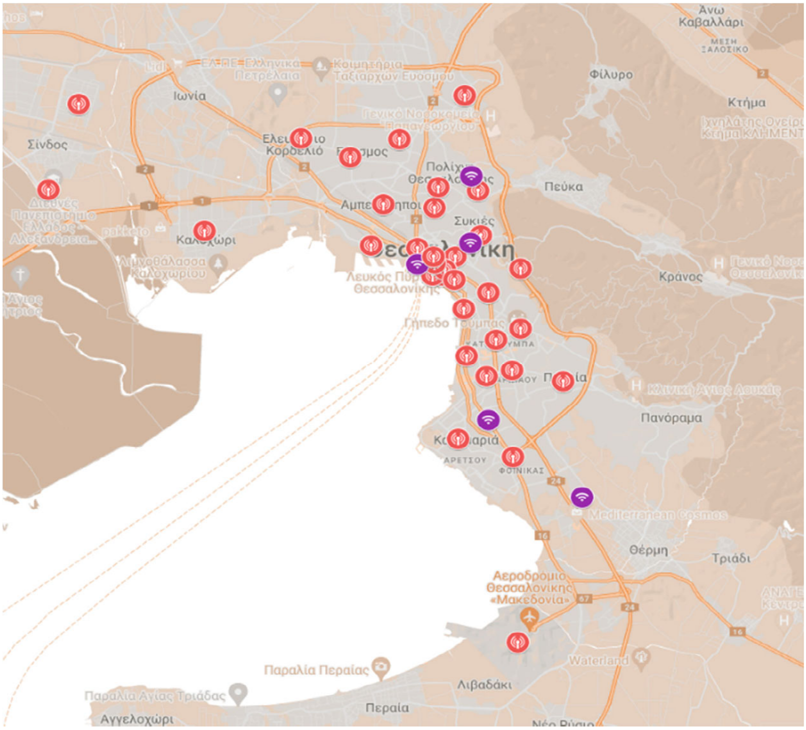



**Fig. 1** The LCAQSN network of KASTOM in the Greater Thessaloniki Area. Nodes are marked in red and LoRaWAN gateways in purple

**Fig. 2** Examples of installed LCAQSN nodes of the KASTOM project in the field

extended periods of missing data, posing significant challenges for accurate imputation. Lastly, spontaneous missing values, likely caused by communication or power issues, represent the most manageable scenario for data replication, as these gaps are typically shorter and more sporadic.

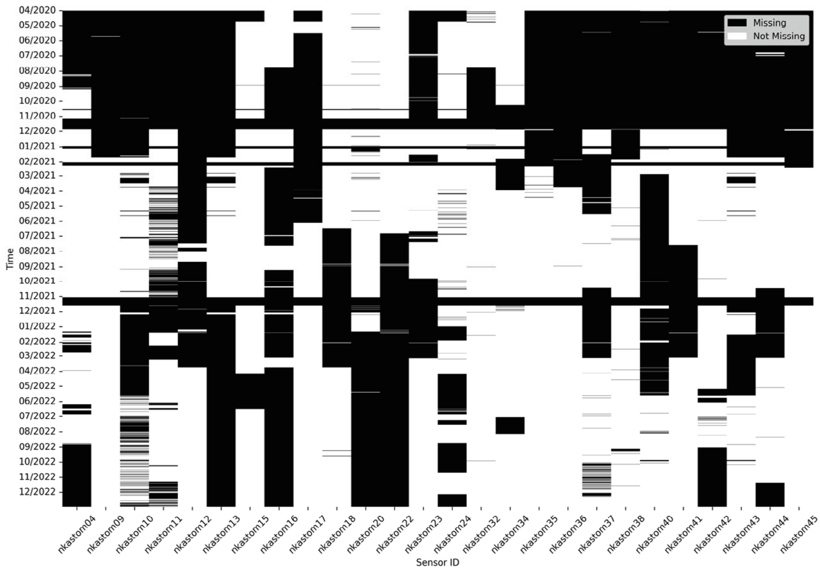



**Fig. 3** Missing Values (black) for every LCAQSN in the GTA. Rows represent days and columns represent different LCAQSNs

## 3 Methods

Three different imputation techniques were applied to address missing values. The first technique applied was simple mean imputation, which serves as a reference method. In this approach, missing values are replaced with the mean value of the observed data for that variable. While this method is straightforward to implement, it does not account for any underlying patterns or relationships within the data, which can lead to biased estimates and reduced variability.

The second technique utilized was k-Nearest Neighbors (k-NN) [17]. This method imputes missing values based on the values of the k-nearest data points, identified using a distance metric such as Euclidean distance. By considering the similarity between observations, k-NN can provide more accurate and contextually relevant imputations compared to simple mean imputations. However, the effectiveness of k-NN depends on the choice of k and the nature of the data. The k-NN imputation was performed using the *KNNImputer* from the sklearn library [18], with the number of neighbors (k) set to 5, after performing a grid search to optimize performance.

The third technique employed was Multiple Imputation by Chained Equations (MICE) [19] using the Bayesian Ridge Regression algorithm. MICE generates multiple imputed datasets by iteratively filling-in missing values using predictive models for each variable. Bayesian Ridge Regression, a type of regularized linear regression, is used to create these models, incorporating prior distributions to improve estimates. This method accounts for the uncertainty of the imputations and provides a more robust solution, especially for complex datasets with intricate relationships between variables. MICE employed using the IterativeImputer class from the sklearn library, with a maximum of 10 iterations and a fixed random seed (random_state = 0) was used to ensure reproducibility of results.

All methods were evaluated using the tenfold Random Cross-Validation technique. To ensure a robust assessment, an additional 10% of missing values were artificially introduced to the dataset, which already contained real missing values. Consequently, the dataset used for modeling had a total of 50.93% missing values. After imputing all missing values, the results were validated specifically on the 10% of artificially created missing values to assess the accuracy and robustness of each imputation technique.

We evaluate the performance of our imputation models using several metrics. The R-squared ($R^2$) value measures the proportion of variance explained by the model. The Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) assess the accuracy of predictions by calculating the average differences between observed and predicted values. The Pearson correlation coefficient indicates the strength and direction of the linear relationship between observed and predicted values. Finally, the Index of Agreement (IA) gauges the overall agreement between the observed and predicted values, providing a comprehensive assessment of model performance.

## 4 Results and Discussion

Table 1 shows the results achieved by the imputation methods based on the artificially generated data gaps and can be considered a robust way of estimating their overall performance. It is evident that the MICE and k-NN algorithms exhibit strong performance, with MICE showing the best metric values overall. In contrast, the simple mean imputation, used as a basic reference method, performed poorly and did not yield satisfactory results. Therefore, in further analysis, we will exclude the mean imputation method from consideration.

Figure 4 illustrates the performance of the two most efficient algorithms in relation to the percentage of missing values per station. Initially, it can be observed that there is no clear relationship between the percentage of missing values and the performance metrics. For example, even a node with 72% missing values demonstrates very high performance in terms of $R^2$ values. Conversely, there are instances where nodes

**Table 1** Evaluation of imputation techniques performance for $PM_{2.5}$

| Method | $R^2$ | RMSE | Pearson | MAE | IA |
|---|---|---|---|---|---|
| Mean | 0.03 | 38.46 | 0.18 | 26.17 | 0.21 |
| MICE | 0.73 | 21.35 | 0.84 | 11.21 | 0.91 |
| k-NN | 0.70 | 21.41 | 0.84 | 10.13 | 0.91 |

with a lower percentage of missing values exhibit lower $R^2$ values. This generally indicates that the imputation process can effectively recover the "lost" information even in extreme cases.

From an algorithmic performance perspective, k-NN shows a lower $R^2$ compared to MICE but consistently performs well across all nodes, never falling below 0.5. On the other hand, while MICE performs better in several instances, it completely fails for three nodes (nkastom11, nkastom12, nkastom15). This might be due to the iterative nature of MICE and the inherent assumptions of the Bayesian Ridge Regression used in each iteration, which might not have captured the more complex missing patterns in these nodes. On the other hand, the k-NN imputer consistently performed well across all nodes, likely due to its ability to base imputations on local information from neighboring data points, making it robust to varying missing data patterns. Since it does not rely on iterative processes or model assumptions, k-NN was more stable, even when large portions of data were missing. Consequently, the imputed values from the k-NN algorithm will be used for the subsequent analysis.

Since statistical performance metrics alone may not provide a complete picture of the behavior of imputation models, Fig. 5 presents the distributions of concentrations

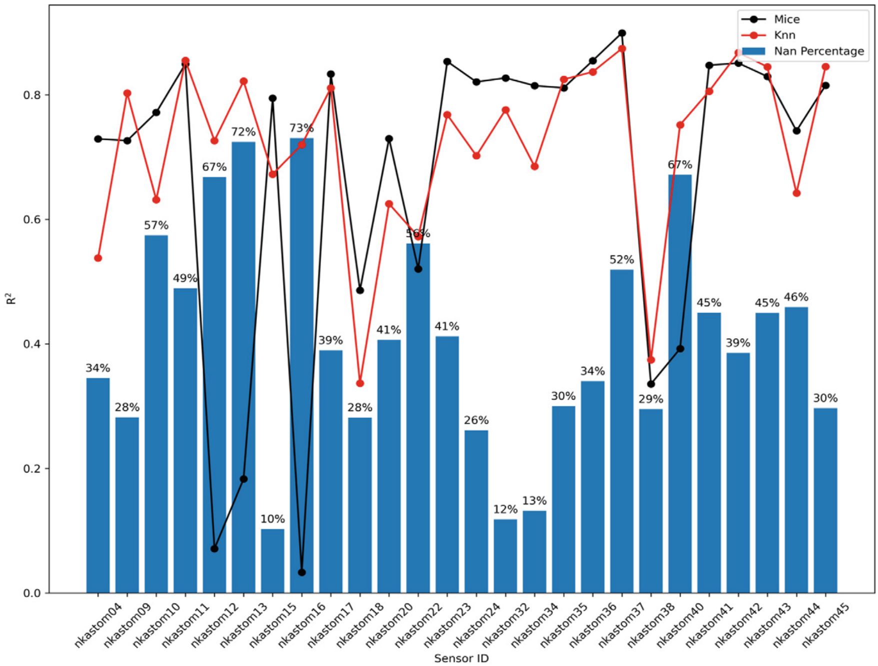



**Fig. 4** Performance of imputation algorithms ($R^2$) by percentage of missing values per LCAQSN in relation to $PM_{2.5}$

before and after imputation, while Fig. 6 shows the difference in the Pearson correlation coefficient between pairs of nodes before and after imputation for $PM_{2.5}$ (PearsonTableAfter—PearsonTableBefore). It is evident that there are minimal differences in the distributions of $PM_{2.5}$ concentration values, therefore verifying that the imputation process not only succeeded in regenerating missing values, but the latter obeyed the basic distribution characteristics of the node values they referred to. In addition, results suggest very small changes in the Pearson correlation coefficient, which is particularly important considering that 50.93% of the data were missing and subsequently imputed. The differences in the Pearson correlation coefficient between node pairs were chosen for display because, in many cases, LCAQSN data are used in data fusion processes to produce estimates at unknown locations, where the spatial correlation between nodes is of utmost importance. In our case results show that the Pearson correlation coefficient between estimated and real values is zero in most of the cases, while for the rest of the node couples it is higher after imputation in comparison to the situation before imputation.

In the analysis presented, Fig. 7 showcases the temporal dynamics of the sensor "nkastom10", comparing raw concentration measurements with those derived from imputation. Although no definitive conclusions can be drawn from this visual representation, it effectively demonstrates the consistency and trends present in the data, providing insight into the sensor's performance over the specified time frame.

## 5 Conclusions

The findings of this study underscore the significant potential of LCAQSN in preserving information and supporting the accurate imputation of missing values in urban air quality monitoring networks, with the aid of proper Machine Learning algorithms. Despite individual nodes experiencing frequent malfunctions and data gaps, the network as a whole can maintain data integrity and successfully fill in missing time-series data, even when more than 50% of the data is missing.

Among the imputation methods evaluated, the k-NN algorithm emerged as the most effective and consistent technique. While the MICE algorithm showed strong performance in several cases, it failed for three nodes. Conversely, k-NN demonstrated robust performance across all nodes, never falling below an $R^2$ value of 0.4. These results highlight the utility of LCAQSN networks in providing high-resolution air quality data through effective data imputation. The successful application of k-NN in this context suggests that it is a viable tool for enhancing the spatial resolution of air quality mapping. This capability is particularly important for data fusion processes aimed at producing estimates in unknown locations, where maintaining spatial correlation between nodes is critical.

Moreover, this study reveals that imputation methods like k-NN are resilient even in extreme cases of missing data, which is encouraging for future deployment in real-world urban environments where network malfunctions and data gaps are inevitable. The ability to accurately impute data even for sensors with significant losses can

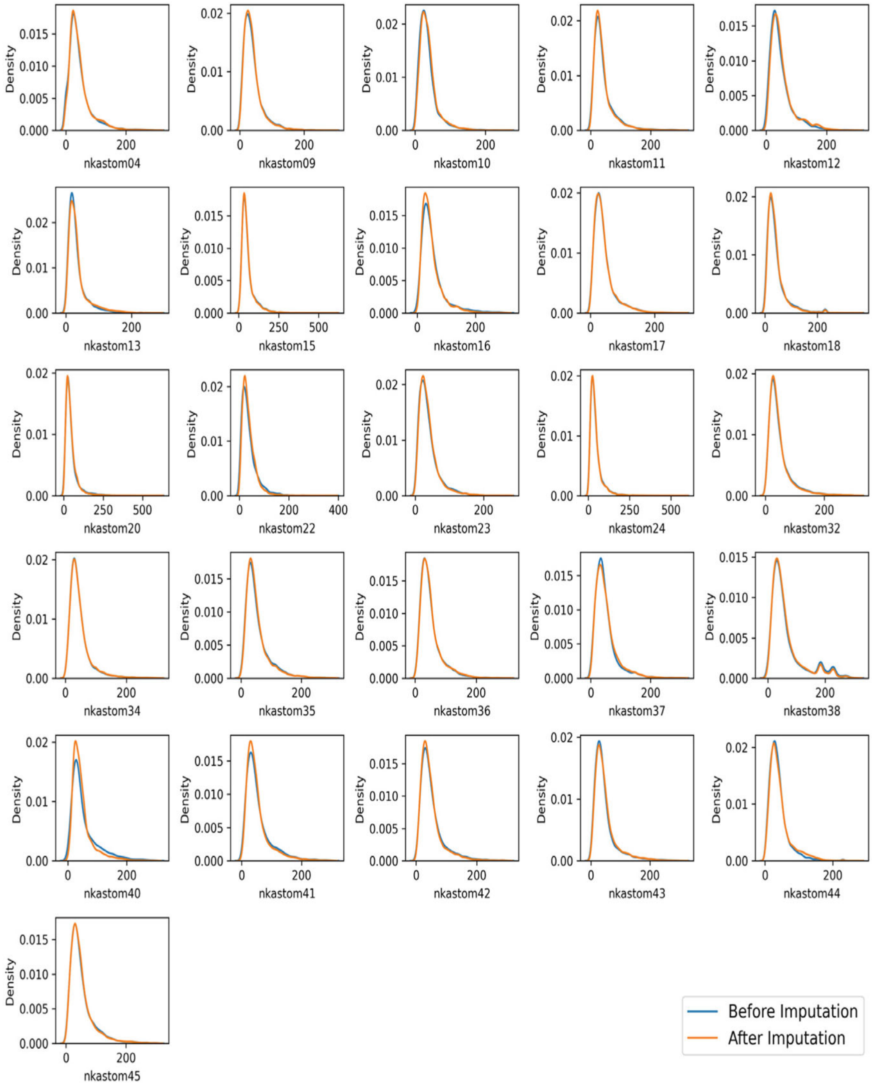



**Fig. 5** Distribution of $PM_{2.5}$ concentrations before and after imputation of missing values per LCAQSN. The horizontal axis reports the number of data points while the vertical axis their value density

help avoid the high costs associated with sensor maintenance and replacement, contributing to the scalability and sustainability of air quality monitoring systems.

Lastly, this study focused on data imputation for particulate matter (PM2.5). Expanding this analysis to other pollutants and meteorological parameters could further validate the imputation methods and reveal different dynamics across air

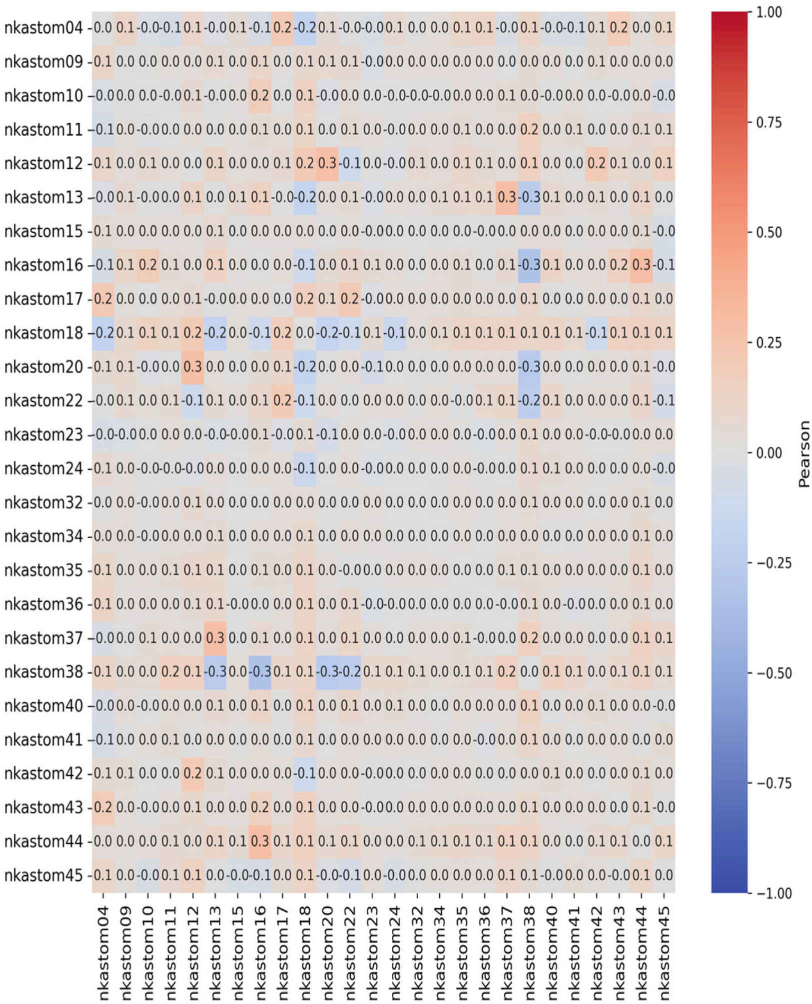


**Fig. 6** The difference in Pearson correlation coefficient between pairs of nodes for $PM_{2.5}$ (PearsonTableAfter—PearsonTableBefore). Blue indicates negative difference while red indicates positive difference

quality indicators. In terms of practical applications, integrating these imputation techniques with real-time monitoring systems could offer enhanced decision-making tools for stakeholders aiming to mitigate air pollution in urban environments.

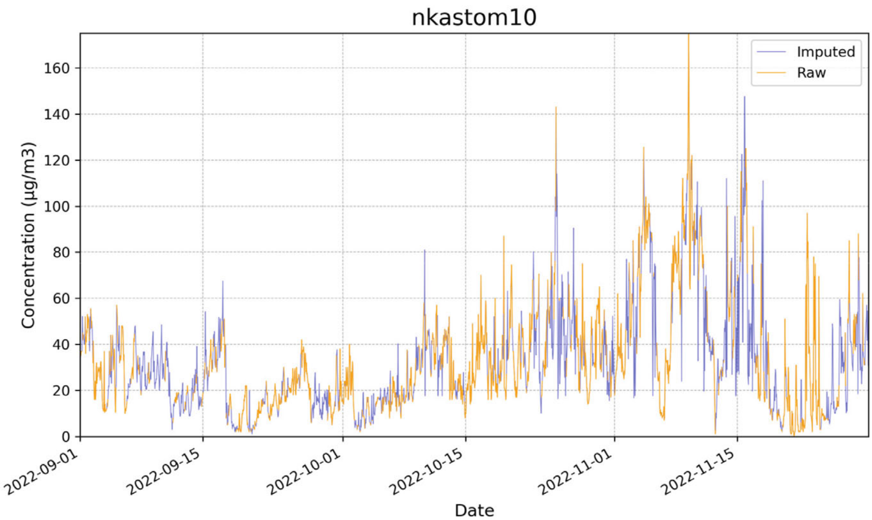



**Fig. 7** Time-series comparison of raw and imputed concentration values for the sensor "nkastom10". The blue line represents the imputed data, while the orange line corresponds to the raw measurements

**Acknowledgements** This work is supported by the EnviTwin project, which is carried out within the framework of the National Recovery and Resilience Plan Greece 2.0, funded by the European Union—NextGenerationEU (Implementation body: HFRI).

## References

1. WHO: https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health (2022). Last accessed 24 June 2024
2. EEA: https://www.eea.europa.eu/en/analysis/indicators/health-impacts-of-exposure-to. Last accessed 24 June 2024
3. EEA: https://www.eea.europa.eu/highlights/air-pollution-goes-down-as. Last accessed 24 June 2024
4. Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe. OJ L 152, 11.6.2008, pp. 1–44. http://data.europa.eu/eli/dir/2008/50/oj. Last accessed 24 June 2024
5. Sokhi, R.S., Moussiopoulos, N., Baklanov, A., Bartzis, J., Coll, I., Finardi, S., Friedrich, R., Geels, C., Grönholm, T., Halenka, T., Ketzel, M., Maragkidou, A., Matthias, V., Moldanova, J., Ntziachristos, L., Schäfer, K., Suppan, P., Tsegas, G., Carmichael, G., et al.: Advances in air quality research—Current and emerging challenges. Atmos. Chem. Phys. **22**(7), 4615–4703 (2022)
6. Kumar, P., Morawska, L., Martani, C., Biskos, G., Neophytou, M., Di Sabatino, S., Bell, M., Norford, L., Britter, R.: The rise of low-cost sensing for managing air pollution in cities. Environ. Int. **75**, 199–205 (2015)
7. Kang, Y., Aye, L., Ngo, T.D., Zhou, J.: Performance evaluation of low-cost air quality sensors: A review. Sci. Total. Environ. **818**, 151769 (2022)

8. Borrego, C., Costa, A.M., Ginja, J., Amorim, M., Coutinho, M., Karatzas, K., Sioumis, Th., Katsifarakis, N., Konstantinidis, K., De Vito, S., Esposito, E., Smith, P., André, N., Gérard, P., Francis, L.A., Castell, N., Schneider, P., Viana, M.C., et al.: Assessment of air quality microsensors versus reference methods: The EuNetAir joint exercise. Atmos. Environ. **147**, 246–263 (2016)
9. Bagkis, E., Kassandros, T., Karatzas, K.: Learning calibration functions on the fly: Hybrid batch online stacking ensembles for the calibration of low-cost air quality sensor networks in the presence of concept drift. Atmosphere **13**(3), 416 (2022)
10. Gressent, A., Malherbe, L., Colette, A., Rollin, H., Scimia, R.: Data fusion for air quality mapping using low-cost sensor observations: Feasibility and added-value. Environ. Int. **143**, 105965 (2020)
11. Suh, H., Song, J.: A comparison of imputation methods using machine learning models. Commun. Stat. Appl. Meth **30**(3), 331–341 (2023)
12. Shaadan, N., Rahim, N.A.M.: Imputation analysis for time series air quality (PM10) Data set: A comparison of several methods. J. Phys.: Conf. Series **1366**(1), 012107 (2019)
13. Agbo, B., Al-Aqrabi, H., Hill, R., Alsboui, T.: Missing data imputation in the internet of things sensor networks. Future Internet **14**(5), 143 (2022)
14. Van Zoest, V., Liu, X., Ngai, E.: Data quality evaluation, outlier detection and missing data imputation methods for IoT in smart cities. In: Studies in Computational Intelligence, pp. 1–18. Springer International Publishing (2021)
15. Rivera-Muñoz, L.M., Giraldo-Forero, A.F., Martinez-Vargas, J.D.: Deep matrix factorization models for estimation of missing data in a low-cost sensor network to measure air quality. Ecol. Inform **71**, 101775 (2022)
16. Hellenic National Meteorological Service, http://www.emy.gr/emy/en/climatology/climatology_city?perifereia=Central%20Macedonia&poli=Thessaloniki_Mikra. Last accessed 24 June 2024
17. Altman, N.S.: An introduction to kernel and nearest-neighbor nonparametric regression. The Am. Statistician **46**(3), 175–185 (1992)
18. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al.: Scikit-learn: Machine learning in Python. J. Mach. Learn. Res. **12**(Oct), 2825–2830 (2011)
19. Buuren, S.v., Groothuis-Oudshoorn, K.: MICE: Multivariate imputation by chained equations in R. J. Stat. Softw. **45**(3) (2011)

# Living with Future AI: Studying Experience, Attitudes and Expectations of Greek Smart Home Users

**Lu Jin, Sotirios Athanasoulias, Dominik Pins, Alexander Boden, Britta Essing, and Nikolaos Ipiotis**

**Abstract** As the European Union advances towards its strategic goal of achieving a cleaner environment, member states are intensifying their energy efficiency initiatives. Among these, the integration of artificial intelligence (AI) holds considerable promise for enhancing sustainable energy management practices. While much research has focused on refining the accuracy of AI in enhancing energy efficiency, less has been conducted on exploring users' perceptions of AI. Our study seeks to explore the perceptions of early adopters of smart home technologies in Greece regarding the role of AI in future smart homes. Employing a mixed-methods approach, we have analyzed the views of 9 individuals, who represent a diverse range of living arrangements from single occupants to multi-person households. The in-depth analysis highlights a generally positive outlook on the benefits of AI in energy management. Nonetheless, there is a discernible desire to balance automation and user autonomy within AI systems. Participants preferred AI interactions that do not mimic human behaviour, emphasizing the importance of retaining a home-like ambience rather than shifting towards full automation. Moreover, our findings indicate that the perceived benefits of AI correlate with household size, with larger households showing a greater appreciation for AI's capacity to optimize energy usage and reduce expenses. These insights provide valuable implications for policymakers and developers aiming to customize AI solutions better to meet the needs and preferences of Greek households, thereby facilitating more effective AI integration in homes.

**Keywords** Smart homes · AI · Energy

L. Jin (✉) · D. Pins · A. Boden · B. Essing
Fraunhofer-Institute of Applied Information Technology FIT, Sankt Augustin, Germany
e-mail: lu.jin@student.uni-siegen.de

L. Jin
University of Siegen, Siegen, Germany

S. Athanasoulias
National Technical University of Athens, School of Rural, Surveying and Geoinformatics Engineering, Athens, Greece

S. Athanasoulias · N. Ipiotis
Plegma Labs, Marousi, Greece



V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_4

## 1 Introduction

AI is perceived as a cutting-edge technology that can bring state-of-the-art management techniques to ordinary households. Smart homes [2, 3], as a focal point of AI applications, utilize systems that automate daily tasks and optimize energy consumption. AI is currently being integrated into various aspects of daily living in smart homes to increase home-automation [4, 5] . Devices like smart thermostats, lights, and security systems can be controlled remotely or set to operate autonomously based on user preferences and routines. Smart meters and home energy management systems monitor real-time energy usage, providing insights and recommendations to homeowners. AI can provide advanced security systems such as facial recognition, motion detection, and real-time alerts [3].

The European Union (EU) has actively promoted the adoption of smart meters to enhance energy efficiency and transition to a more sustainable and secure energy system. According to the [6] , the installation of smart meters is expected to have 74% penetration in Europe by 2027.

Greece, with its unique climate and abundant natural resources such as wind and solar energy, stands out among other European countries. These conditions make Greece particularly suitable for the implementation of smart homes that optimize energy use through integration with smart grids. Recognizing this potential, the Greek government has distributed smart plugs, paving the way for the widespread adoption of smart home technologies. Therefore, in this paper, we have on-site interview with Plegma Living Lab—early smart home adopter in Greece to discover the potential future benefits and challenges of AI in smart homes.

In addition, due to the ubiquitous presence of AI in smart homes and other areas, human-centered Artificial Intelligence (HCAI) has emerged as a different perspective on AI, focusing on human needs [7] and perceptions [8–10], human-centered AI should be built on three levels: user, community, and society. In this paper, we cooperated with an existing Living Lab in Greece which it is a part of the Athenian community, a newly established non-profit energy community located in the municipality of Attika, Greece. It provides us a unique chance to explore the HCAI concept from all of the three levels—user, community and society.

Therefore, our paper starts from the users' perspective. By understanding their views and experiences, we can develop AI technologies that are both innovative and socially responsible, ensuring a smooth and positive transition to future smart living environments. Based on the background of HCAI, we adopted a subject-context-motivation [12] approach. For the subject, we select Greece's early smart home adopters. We investigate their perception of AI-infused smart homes. For the context, we are interested in experiences of living in smart homes. For the motivation, we aim to investigate what the concerns, perceptions and expectations of the users are, in order to better understand how to design AI in a human-centered way.

Therefore, our main research questions are:

1. **RQ1:** What are the concerns, perceptions, and expectations of Greek residents regarding the integration of AI technologies in smart home environments?

2. **RQ2:** What are the differences in acceptance and trust towards AI technologies in smart homes among different household compositions in Greece?

By focusing on Greece—a Mediterranean country with unique energy challenges—On the one side, the trend towards smart homes in Greece is unstoppable. The government has promoted smart plugs, which have been adopted by many Greek families, paving the way for smart homes. Also, Greece's unique natural resources, such as solar and wind energy, offer significant potential benefits for adopting smart home technologies and AI to optimize energy use. On the other side, the Greek population shows the moderate level of trust of AI and smart home technologies [13].

Therefore, our study offers insights into local Greek attitudes towards innovative technologies. Through 1 on 1 interview (qualitative) and data analysis from their energy consumption patterns (quantitative), we seek to provide a comprehensive understanding of perception towards AI in smart homes.

The structure of the paper is as follows: Sect. 2 lists the related work; Sect. 3 describes the methods; Sect. 4 analyzes the results; Sect. 5 discusses the research questions.

## 2 Related Work

In this paper, we mainly aim to examine users' attitudes, expectations, and perceptions towards AI in smart homes. Based on our context, we formed our related work in three areas: User attitudes, expectations, and perceptions towards AI in smart homes, human-centered AI in smart homes, and social dynamics influence the perception of AI adoption.

### *2.1 Users' Attitude, Perception, Expectation Towards AI in Smart Homes*

Research has highlighted the importance of understanding users' perceptions and preferences in the adoption of these technologies [14]. Users' attitudes, perceptions, and expectations towards AI in smart homes are influenced by various factors [15]. These attitudes towards AI in smart homes are predominantly positive, with many appreciating the enhanced convenience and efficiency these systems bring to daily life [16]. The technical performance [17, 19, 20], the explainable artificial intelligence(XAI) [21], the potential benefit of energy efficiency [18, 22] can impact residents' attitudes towards adopting smart home technology. In [23] , he mentioned that the lifestyle benefits and perceived compatibility with modern living are also significant motivators for users adopting smart home technology. Besides, in [24], characteristics such as personal beliefs, energy expenses, and environmental protection values play a significant role in users' preferences for interactions with smart home assistants, affecting their attitude towards proactive SHAs ( smart home assistants). Studies also have shown that more engagement in AI systems enhances their feeling of control, perceived understanding, and usefulness of AI [16].

On the contrary, skepticism and concerns also persist among users [25, 26]. Privacy and data security are major issues, with many users wary of how their personal information is collected, stored, and utilized by AI systems [27, 28]. The potential for over-dependence on technology and the fear of AI systems malfunctioning or being hacked add to this apprehension. Additionally, there is a segment of users who are unsure about the reliability and accuracy of AI, preferring to maintain a level of control over their home systems rather than relying entirely on automated solutions [29].

Our work is built on these studies and aims to have a comprehensive understanding of early smart home adopters' perceptions towards AI in future smart homes in Greece.

## 2.2 Human-Centered AI in Smart Homes

Human-centered AI in smart homes focuses on prioritizing user empowerment, ethical considerations, and enhancing user experiences [30, 31, 34]. This approach emphasizes the importance of user-centric design principles, ethical guidelines, and interdisciplinary collaboration to create AI technologies that align with human values and promote user agency.

In [32], he introduced three innovative concepts: (1) a two-dimensional HCAI framework illustrating how both high levels of human control and high levels of automation can coexist; (2) a shift in focus from mimicking human behaviour to empowering individuals, advocating for a change in language, imagery, and metaphors from depicting AI as autonomous teammates to portraying them as powerful tools and teleoperated devices; (3) a three-tier governance structure that outlines how software engineering teams can create more reliable systems, how managers can foster a safety-oriented culture within organizations, and how industry-wide certification can enhance the trustworthiness of HCAI systems. Wei et al. [33] put the human-centered AI philosophy into real-world application and propose a comprehensive HCAI methodological framework integrating seven components, stressing the need for multidisciplinary professional involvement. In [35], it shows that by involving users in the learning phase of smart home AI systems, individuals can enhance their feeling of control, understanding, and usefulness of AI, leading to better personalization and increased user control over intelligent agents for home automation.

However, there lack of empirical studies to investigate human-centered AI in smart homes in the real world. Our work is trying to investigate it based on the perceptions of early smart home adopters.

### 2.3 The Influence of Social Dynamics towards its Perception on Technology Adoption

The adoption of technology is influenced by various social dynamics and perceptions. Studies have shown that the dynamics of household demographic attributes such as socio-economic statistics, culture, religion, social support, education, age, social influence, environmental consciousness, and technology savviness can affect technology adoption within households [36]. Social networks also play a significant role in technology adoption, with information exchange within social groups influencing individual adoption decisions [37, 38].Studies have shown that factors such as socio-economic statistics, culture, religion, social support, education, and age play a significant role in shaping individuals' decisions to adopt technology [39]. This paper only focuses on three perspectives: early smart home adopters, Greek, and households with various co-inhabitants.

Adami et al. [40] shows the early adopters often perceive AI as ubiquitous, whereas industrial engineers rarely feel the need to explicitly highlight its intelligence. This contrast could stem from the conservative stance of end-users, primarily older individuals, who may not fully grasp the benefits of AI and might even be deterred by it.

User perception of AI in smart homes in Greece is generally positive, with Greek consumers showing a growing interest in smart home technology due to factors like perceived usefulness, compatibility, and ease of use [13]. Health Smart Homes are seen as beneficial for elderly users in Greece, offering proactive monitoring and customization of the user environment to enhance autonomy and quality of life [41].

We can see limited information on AI perception in Greece's smart homes and user perspectives on AI in Greek smart homes are not specified. Our research uses a mixed method to investigate the perception of AI in smart homes based on the Living Lab data in Greece.

## 3 Method

### 3.1 Participants

Our participants come from users in the Plegma living lab. The 13 distinct houses involved in data collection belong to the Athenian community, a newly established non-profit energy community located in the municipality of Attika, Greece. A group of experienced technology and engineering professionals with extensive backgrounds in research and development founded this community. The households include a wide variety of demographics, from retired and working couples to families and single-person households. For the living lab, the researchers installed smart plugs for the inhabitants appliances and connected them to an energy monitoring system with a graphical user interface. This application allows participants to monitor and visualize their energy consumption in real-time an provides web-based or mobile based access to their energy data and can control the device remotely. The Plegma dataset and its corresponding research are in [42].

The living lab got ethical approval from the National Technical University of Athens' ethical committee(See Sect. 3.3). We also got consent for the interview and recorded the interview with oral permission. To create an environment conducive to open and interactive dialogue, seven interviews were planned to take place in person at Plegma Labs in Greece, the remaining two interviews were conducted online. The interviews were all conducted in English. In order to make the questionnaire fully comprehensive, we conducted two Pilot Testers (R1+R2). Two Pilot Test is varied in age and living situation. One is living with family, one is living alone.

## 3.2 Procedure

*Quantitative Data.* The Quantitative data were collected in 2023 with the Plegma based on an energy monitoring platform. As we were interested in comparing different living situations, we selected three use cases based on different family sizes—4-size family (Fig. 1), 2-size family (Fig. 2) and 1-size family (Fig. 3) from the Plegma living lab data. These types show very different energy consumption patterns as can be seen in the figures. 4-size families show most energy activity over the day, while 1-size families show most activities at night or on weekends. This is due to different life situations: if the caretakers and children live at home mostly, you can see their energy active during most of the day. A single person who is always working at the office, on the other hand, will be only home at night or on weekends. For families with adult children, finally, there is another situation where different patterns intersect as parents lose their influence over the child's behaviours.

**Table 1** Demographic and living situation of interviewees

| User | Gender | Living situation | Co-inhabitants | Property type |
|---|---|---|---|---|
| R1 | Male | Lives with family | Family members | Apartment |
| R2 | Male | Lives alone | No co-inhabitants | House |
| R3 | Male | Lives alone | No co-inhabitants | Apartment |
| R4 | Male | Lives alone | Animal | Apartment |
| R5 | Male | Lives with family | Family members | House |
| R6 | Female | Lives alone | No co-inhabitants | Apartment |
| R7 | Male | Lives with family | Family members | House |
| R8 | Male | Lives with partner | Partner | Apartment |
| R9 | Male | Lives alone | No co-inhabitants | Apartment |

*Interview Data.* Qualitative interview data were collected in a month in Feb 2023. The first author visited Plegma Lab to conduct nine in-depth semi-structured interviews. The nine participants, are detailed in Table 1. With the interviews, we aim to deeply understand user's attitudes, perceptions and expectations towards AI in smart homes. The interviews also commenced with questions about demographic details and living conditions to place each participant's experiences within their specific environmental context. The overall discussion included:

- Part 0: demographic question towards background information and house information.
- Part 1: energy behaviour practice questions
- Part 2: Smart home questions related the smart home experience and investigation towards perception and attitude towards AI-enable smart home
- Part 3: The peer influence of energy-saving practice

## *3.3 Ethical Considerations*

To ensure ethical integrity and secure necessary approvals for our data collection, we implemented strict protocols that emphasize the respect and privacy of participants. The Ethical Committee of the National Technical University of Athens (https://www.elke.ntua.gr/en/ethics-committee/) approved our study under file number 15623. Participation was entirely voluntary, and each household provided detailed informed consent before we commenced data collection. This consent outlined the study's goals, the nature of the data collected, and its use in the European Commission-funded Marie Skłodowska-Curie Action GECKO project for research and innovation purposes. We guaranteed participants the right to withdraw at any point without any negative consequences. To maintain confidentiality, we anonymized all data by removing personal identifiers and securely stored it in Plegma Labs' databases, demonstrating our strong commitment to ethical research and the protection of participant privacy.

## *3.4 Data Analysis*

*Energy usage pattern analysis.* In this study, we utilized the data from Plegma Living Lab platform to analyze energy consumption patterns across three distinct family sizes. Data was extracted and processed using the Colab data analysis tools, facilitating the generation of data visualizations. The Plegma dataset and its corresponding research [42] are available in a public repository at [https://pureportal.strath.ac.uk/en/datasets/plegma-dataset]. We then correlated the quantitative data with interviews conducted with families, employing qualitative methodologies to craft narrative stories reflective of their respective energy consumption behaviours. These narratives were integrated with the energy usage data to develop three user case studies, each tailored to different family size, thus providing comprehensive insights into the dynamics of household energy consumption.

*Interview data analysis.* In our data analysis, we employed a qualitative methodology, executing a thematic analysis to methodically uncover patterns in users' perception of AI in smart homes. The process began with the first author and third author converting all dialogues into user stories. Subsequently, the first author delved into these narratives to identify and extract viewpoints. This collaborative analysis led to the emergence of distinct patterns. Our focus is centred on: the functions of AI in smart homes, and the users' attitudes, perceptions, and expectations towards AI in smart homes.

# 4 Results

## *4.1 Energy Consumption Pattern Based on Different Co-inhabitant Composition*

*Single-Person Households:* Single-person households have the simplest energy consumption patterns due to the lack of additional family members influencing decision-making. They typically follow a routine work schedule, resulting in predictable energy usage. They work 9am to 5 pm, and the 6pm to 8am at home, their custom usage of energy is between 6 pm and 10 pm which is the peak energy usage among all of the countries. Their lifestyle habits are often straightforward, and they tend to show little interest in energy savings because they believe their consumption is not significant. However, they occasionally compare their energy consumption with other individuals in similar situations to seek improvements.

*Traditional Four-Person Family:* In traditional four-person families, energy consumption patterns are characterized by diverse usage behaviours due to varying schedules, needs, and preferences of each family member. The shared use of common spaces like the kitchen and living room leads to higher daily energy consumption. Additionally, there is a tendency to forget to turn off devices because of reliance on other family members, contributing to unnecessary energy usage.

*Multi-generational Families:* Multi-generational families often consist of multiple generations living under one roof, leading to unique energy consumption behaviours. Overlapping schedules and diverse habits among family members result in significant variations in daily energy usage. Elders may have different comfort and usage preferences compared to younger members, further contributing to the complexity of managing energy consumption patterns in such households.

### 4.2 *Use cases: Nuclear Family Versus Nuclear Family Special Versus Single Occupancy*

For the total energy consumption data, we identified the three types of families based on their co-inhabitants—families with children, single and families with adult children and parents (Table 2). We used narrative to describe their daily routine.

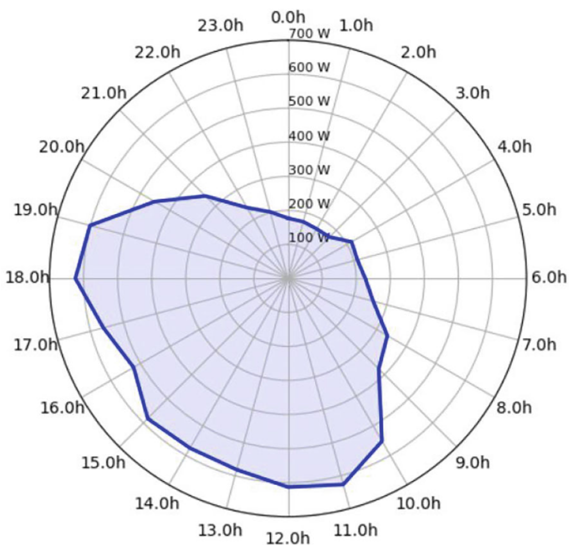
*Use Case 1: Nuclear Family—Two parents and one or two children.* U1 (Fig. 1) is a technologically savvy individual living in a two-story house with his family, which includes two children and his wife. U1's household has basic smart home components, including smart meters and plugs. These devices are connected to a central smart home system capable of monitoring and controlling energy usage across various devices and appliances. The output of energy consumption pattern data is shown in Fig. 1.

Every morning, U1 reviews energy consumption data from the previous day using a mobile app linked to the smart home system. The system provides notifications and reminders to turn off unused appliances. For example, if the boiler is left on unnecessarily, U1 receives an alert on his phone. *"I check, I have automated to turn*

**Table 2** Categorization of household compositions

| Category | Number of coinhabitants | Common characteristics |
|---|---|---|
| Nuclear family | 3–4 | Two parents and one or two children |
| Nuclear family (special) | 2–3 | Adult child with parents |
| Single occupancy | 1 | Individual living alone |

**Fig. 1** Family with children



*on the boiler and turn off, yeah. From the app that we have. So I check that we don't leave it on."*

U1 attempts to engage his family in energy-saving practices by explaining the benefits and showing them how to use the smart system for monitoring their personal energy consumption. Despite his efforts, family members occasionally forget to turn off lights and appliances, leading to unnecessary energy waste. *"They leave the light open, they leave the boilers open, yeah to laundry, they just have a few clothes, they don't wait until it's a good load." "I try to organize, we take a shower like 2 people or three people so that the water doesn't go and get cold again. I have automated to turn on the boiler and turn off."*

The smart home system provides weekly reports on energy consumption, highlighting trends and areas for improvement. He adjusts the settings on the smart devices based on seasonal changes and feedback from the system to further reduce energy consumption.

Since he is a technologically savvy individual. His expectation towards future smart home is that he can program the smart home system to automatically turn off lights in unoccupied rooms using motion sensors. And also he expects that he can set the heating system to adjust automatically based on the weather forecast and typical usage patterns to optimize energy use. Also for other special situations, for example, if a device fails or behaves unexpectedly, U1 expect to receive immediate notifications and can troubleshoot or contact support. Or, if energy consumption increases unexpectedly, the system alerts him to potential issues, like a malfunctioning appliance or deviation from typical usage patterns, allowing for quick interventions. Besides from his vision, *" I think the solution to the energy problem is a smart grid, not just a smart home because if you see the energy that people consume it's pretty low."*

Also, he takes limited value towards intelligent personal assistants like Google Assistant connecting to smart devices. *" Google Home, Google Assistant, yeah, little round one... No, it's not that I don't need it."*

In his family, he is the primary user of the smart devices, who is responsible for setting up and monitoring smart home devices. His wife is the secondary users, who only check the data from the smart devices. There is not too much communication between he and his wife towards the usage of smart devices.

*Use Case 2: Nuclear Family Special—Adult child with parents.* U2 (Fig. 2) is a software engineer living with his parent. His family's home is equipped with basic smart home devices like smart meters, smart thermostats and smart plugs. He is the primary user who is responsible for setting up and maintaining the smart home system. He has a high level of technical expertise, whereas his parents do not. His parent is a secondary user who interacts with the system primarily through simplified interfaces.

It exposes the potential of two customized interfaces in the future smart home for two kinds of users—an advanced interface for him, offering detailed analytics, manual controls, and customization options and a simplified interface for his parents, focusing on basic controls and notifications.

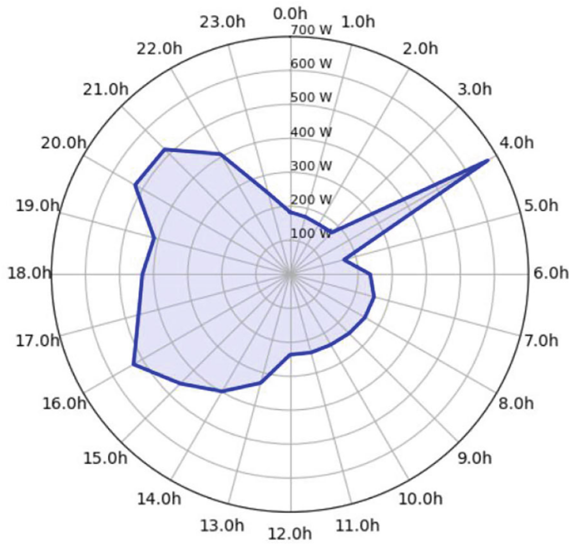



**Fig. 2** Adult children and parent

Since the data is in the platform, he can do the practical actions towards energy saving. For example, *"When I turn on the boiler or air conditioner, I know exactly how much it consumes and I am aware that I need to turn it off in a while."*

He is frustrated with energy policies in Greece because they can not have dynamic prices for energy. *"One thing is the legislation in Greece. So as I told you, we don't have dynamic prices per hour."*

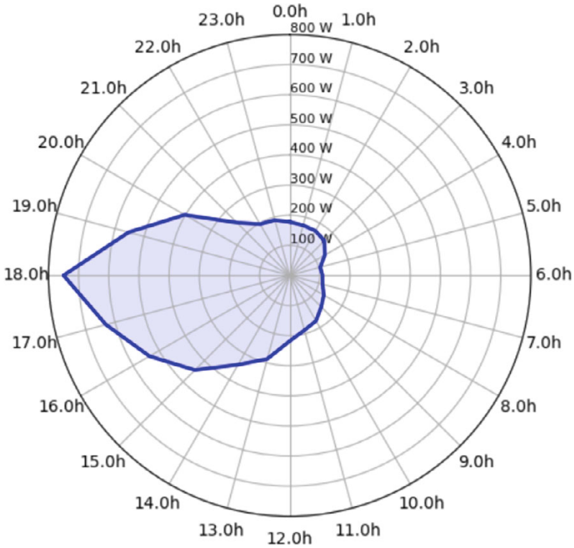
As he is an adult, he can influence his family members' energy habits. *" try to motivate my parents to save energy by telling them not to leave things open for too long."*

For future smart homes, he expects automation in some devices, for example, water heaters and air conditioners. However, he is biased towards AI's role in energy efficiency from AI professional perspective because he thinks it is too early to consider it. At the moment, smart home is mostly aiming at controlling the house. Most people is just living with all the devices on and not fully aware of it. *" In an ideal smart home... I would like to have actuation for some of the devices, like the water heater and air conditioners."*

*Use Case 3: Single Occupancy.* U3 (Fig. 3), a homeowner living alone with a dog, seeks to optimize her household's energy consumption while maintaining convenience and comfort. U3 is familiar with the technology but not with AI. She does not consider herself an energy-sensitive person because she thinks she does not consume too much energy.

U3 also installed a comprehensive smart home energy management system. Smart meters provide real-time data on her total energy consumption, while smart plugs enable her to monitor and control individual appliances remotely. Using the data collected from her smart meters, She gains insights into her household's energy

**Fig. 3** Single




usage patterns. She identifies high-consumption appliances and areas where energy waste occurs.

She takes action to save energy based on the data from the system and expects direct feedback from energy providers. *"And you what are my most consuming appliances? There's still use the use of them. OK. In that way if I use the more energy I would."*

As a Greek Resident, she explores options to integrate renewable energy sources, such as solar panels, into her smart home ecosystem. She investigates incentives for renewable energy adoption and evaluates the feasibility of incorporating green energy solutions into her energy management strategy.

Her expectation towards future smart home will be a voice-activated control that can interact with her smart home system effortlessly. It includes adjusting thermostat settings, turning off lights, or scheduling appliance usage simply by speaking to her smart assistant. *"And with voice commands to schedule tasks, for example, start the washing machine at specific time in order to be ready to be dried when I get home. It can save a lot of time."*

### *4.3 ViewPoint 1: AI as a Personal Assistant for Energy Efficiency*

*A Facilitator of New Energy-saving Habits.* Integrating AI as a personal assistant for energy efficiency is anticipated to transform how energy is managed in homes. Users expect these systems to be intelligent, adaptive, and sensitive to their preferences and routines. AI is seen as a facilitator of new energy-saving habits. Users believe AI should serve as a discreet yet effective manager of home energy systems, capable of making sophisticated decisions based on a comprehensive understanding of user habits and preferences, as suggested by the desire for AI to *"understand my patterns and optimize the house accordingly without my constant input (R6)."*

*A Crucial Role for Energy Optimization.* AI's ability to learn and adapt to household patterns plays a crucial role in energy optimization, suggesting the most efficient times to use appliances and adjusting settings for optimal energy use. *"AI could optimize energy use by learning our household patterns and suggesting the best times to use energy-intensive appliances (R4)." "AI should ideally manage energy usage by scheduling the operation of devices based on optimal energy consumption times (R3)."*

## 4.4 Viewpoint 2: Desiring Control Over AI Functions

*Override AI Decisions.* The need for an override function is a recurring theme, illustrating users' insistence on being able to intervene when AI does not align with their expectations.It emphasises the importance of being able to intervene when AI operations do not align with their specific requirements or situations. *"I need to be able to override AI decisions if they don't align with my current needs(R3)."* As one user pointedly mentioned, *"I like the idea of AI, but I must have the ability to override it anytime I feel like it's not doing what I want (R6)."*

*Customization of AI Settings.* Customized AI functions according to personal preferences is crucial for users to feel comfortable and satisfied with the smart technology in their homes. *"I would like to have detailed control over the AI's functions, ensuring it acts within the parameters I set(R4)."* Due to the background of these interviewees, most of them have a technical AI background, therefore, they prefer to control their AI systems extends to very personal preferences and situational needs. *"It's important that I can manually adjust what the AI proposes, especially when it comes to heating(R5)." "I want to be able to program the AI myself to ensure it meets my standards for energy savings (R9)."*

## 4.5 Viewpoint 3: Skepticism Towards Over-Automation

*Concerns over inefficiencies in Automation.* Users express concerns about the ability of AI to handle older systems and infrastructure effectively. These concerns are not just about the adaptability of AI but also about its efficiency in managing non-modern equipment: *"We also have very old radiators here with you know...that don't work very well (R1)."* It exposed the gap between the AI capabilities and the realities of existing home infrastructure. Users find that AI-driven systems are not only hard to integrate but often fail to improve energy efficiency due to their inability to effectively interact with old, inefficient heating components.

*Concerns on excessive AI control.* Many users are wary of AI systems that may exert too much control over home environments, potentially overriding personal preferences. *"I'm cautious about letting AI handle everything; there needs to be a balance so it doesn't override our personal preferences(R5)." "I'm wary of AI that takes too much control over home systems, like adjusting the thermostat without my consent (R3)."* These concerns reflect a fear that users are particularly sensitive to changes made without their consent, indicating a need for systems that offer greater transparency and user control. *"There's a risk that AI could become too controlling, making changes that might not be suitable for every situation(R3)."* The unease emotion among the users exposes that these systems to exert excessive control over home environments, which might infringe upon personal autonomy and preferences. *"I'm skeptical of any system that tries to automate every aspect of home life without sufficient user interaction(R6)."*

*Concerns over loss of personal feeling.* A significant aspect of skepticism relates to the potential for automation to make homes feel less personal and more machine-driven. It exposes the potential dehumanization of home environments through over-automation. Users fear that too much reliance on AI might strip away the personal, human touches that make a house a home. *"Too much automation can make a home feel less personal; I worry that AI might strip away the homely feeling (R5)."* It reveals a concern that AI, while improving efficiency, may reduce the home to a series of automated processes devoid of personal intervention or warmth. This is the special concern that AI will face when being immersed into the home environment.

## 4.6 Viewpoint 4: Desire for Non-Human AI Interactions

*Practicality Over Personality in AI.* Users clearly favour AI systems that enhance device functionality and provide useful assistance rather than attempting to replicate human characteristics. *"I want AI to be a tool that provides data and recommendations, not something that tries to emulate human interaction (R3)."* .This highlights a widespread preference for AI that serves specific, practical purposes, focusing on the delivery of information and recommendations that assist in decision-making processes. Users value AI for its ability to improve the efficiency and functionality of tools and systems they use daily, advocating for a practical, utilitarian approach to AI design and interaction. *"AI should focus solely on operational efficiency without attempting to engage on a personal level (R9)."*

*Functional AI without anthropomorphic characteristics.* According to the in-depth interview, users prefer AI systems that prioritize task-oriented functionality without attempting to forge emotional or social connections. It is grounded in a practical approach to technology where the primary value of AI lies in its ability to streamline daily tasks and improve efficiency within the home. For instance, one user pointedly remarks, *"AI should assist with tasks rather than trying to engage* *socially..(R3)"* highlighting a widespread desire for AI to remain a utilitarian tool rather than a social companion. Another strong theme among users is the desire for AI to remain non-anthropomorphic, focusing solely on task efficiency and effectiveness without adopting any human-like interactions. For example, one user explicitly states their discomfort with AI systems that exhibit human characteristics: *"I would prefer it to be mains a computer and not put it human characteristics. Why? Because that's creepy (R6)."* The user expressed a clear preference for AI systems that operates in a strictly utilitarian capacity, performing necessary tasks without any attempts to emulate human behaviors.

## 5 Discussion

Our key findings indicate that early smart home adopters in Greece acknowledge that AI can be an effective way for energy efficiency to help build new energy-saving behaviour and optimize energy usage, yet they continue to consider control over AI decisions. This is due to their lack of trust towards AI and the problem of finding a balance between automation and autonomy. While AI is a critical changer in future smart homes, such as optimization, automation, and non-touch interaction, users also express some concern about a loss of the traditional home feeling in smart homes. Moreover, they advocate for AI to evolve to be customized for different users, especially for elder users. This evolution means that AI systems do not only need to optimize energy management by acquiring user habits efficiently, but also balance the AI control vs Human control, and AI decision vs human decision, emphasizing putting the human in the centre of AI system. We delve into these aspects below, discussing balancing control and independence in AI systems and the three levels of human-centered AI in smart homes.

### *5.1 Automation Versus Autonomy: Balancing Control and Independence in AI Systems in Smart Homes*

Automation and Autonomy are two recurring keywords in our interviews. On the one side, the users prefer AI to achieve automation, on the other side, users also want autonomy, especially when automation breaks down. According to the [43], automation cannot simply be categorized into "manual" or "automatic". Instead, it should find the right balance between the points that we might believe are meaningful to automate and tasks that are valuable for users to be involved. It is easy to embrace the two extremes—full automation and non-automation. The best ideal solution will exist in between—"as a duality between automation and human interaction, between autonomous technology and the tools we wield [43]." We argue that the future of smart homes is located in between full automation and full autonomy, even for different smart home users with various ages, cultures, and technical backgrounds, the duality point is different. Therefore, the balance is dynamic. Recent work has aligned with this point. In [44], it suggests that semi-automated systems are more prevalent than fully automated smart homes, indicating a preference for a certain level of user control and input in the operation of these systems.

We have found in our study how automation and autonomy show themselves as two characteristics of AI-enabled smart homes. From our results, humans aim to retain control and have the right to override AI's decisions. The ideal solution is to create a balanced integration where AI systems can operate autonomously to improve efficiency and convenience, while still allowing users to maintain ultimate control. There are three ways to balance automation and autonomy:

*User preferences and Customization.* According to our study, different smart home users hold various ages, cultures, and technical backgrounds, so the duality point is different. Some tech-savvy person, they even want to customize AI functions themselves. Therefore, users should have the ability to define and adjust the level of automation and autonomy according to their comfort levels and specific needs. Given the diverse demographic of smart home users, including variations in age, culture, and technical background, customization is crucial. Tech-savvy individuals may desire extensive control and customization capabilities, while others may prefer simple, automated solutions.

*Context-Aware Functionality.* AI systems should be capable of understanding and responding appropriately to various scenarios to maintain a balance between automation and user control. For example, AI can take full control when users are away from home, ensuring security and energy efficiency while providing more manual control when tech-savvy inhabitants are present. This balance ensures that smart home technologies enhance daily living without compromising the user's autonomy or ability to intervene when necessary.

*Overall user experience.* Our results challenge the full automation, interviewees are worried about a loss of the traditional home feeling in smart homes. This stance also raises new ethical concerns in the AI research field. How to keep the human element alive? As we all know, home is an environment that evokes more personal feelings, warmth, and a sense of inner sanctity. Whether smart home empower these feelings or diminish them? Therefore, the balance between full automation and full autonomy can keep the human element alive, for example, AI systems should prioritize personalization to reflect the unique preferences and lifestyles of their users. For example, allowing users to customize lighting, temperature, and sound settings to create a warm and comfortable atmosphere that feels uniquely theirs. Based on different households, the solution should based on their traditional home appliances and lifestyle.

### 5.2 Human-centered AI in smart homes: User-level, Community-level, and Society-level

According to James Landa [11], human-centered AI should be built on three levels: user, community, and society. Our findings support this perspective. For the users, our findings indicate that the recurring themes from the interviewees' requirement align with the three categories of the Human-Centered AI (HCAI) framework established by Xu et al. [45]: ethically aligned design, human factors design, and technology enhancement. For the community, our results reveal that perceptions of AI effectiveness vary with the size of the household. For society, our observations identify that the outdated buildings and infrastructure in Greece are the main significant barrier to AI implementation, most of which were constructed in the 1970s. We will describe these findings in detail below.

*User Level: human-centered AI.* Xu [45] mentioned HCAI with three primary elements: 1) Ethically aligned design 2) Advanced technology that mirrors human intelligence and 3) Human factors design. We can see how the Xu's HCAI framework can cover our results. For the ethically aligned design, it ensures AI in smart homes maintains fairness, accountability, and user control. Oversight of AI decisions is crucial, allowing for human intervention to prevent unchecked actions. Addressing concerns over excessive AI control involves implementing safeguards to ensure users retain ultimate authority over their environment. Additionally, refining AI processes to eliminate inefficiencies in automation ensures that AI enhances rather than complicates daily life. These measures foster trust and confidence in AI technologies, aligning with ethical principles. For the technology enhancement, it plays a vital role in energy optimization and promoting energy-saving habits. AI systems can monitor and manage energy use, adjusting systems for greater efficiency, reducing costs, and contributing to sustainability. For the human factors design, it prioritizes user-centric approaches, emphasizing customization, practicality, and clear interactions. Customization allows users to tailor AI settings to their preferences, enhancing satisfaction. Focusing on practicality over personality ensures AI functions effectively without unnecessary human-like traits. Non-anthropomorphic interactions ensure AI remains a useful tool that enhances daily life without emotional complications.

*Community Level: Different Attitudes towards AI adoption based on different Family Size.* For the Traditional Four-Person Family, the technologically savvy individual often assumes the role of the primary decision-maker for smart home technologies. Their vision for smart homes includes a remote control and overview system that provides a comprehensive status of all devices, ensuring nothing is left on unnecessarily. This individual holds high expectations for automation, recognizing AI's potential to streamline smart systems, conserve energy, and alleviate the operational burden on family members. In many families, one person typically manages all smart devices as the "master user", while others function merely as users. For Multi-generational Families, attitudes toward AI adoption are shaped by differing preferences across age groups. The demand for AI is not solely focused on functionality but also on the user interface, which must be intuitive and user-friendly to accommodate all generations. Elders may resist adopting AI technology due to unfamiliarity and concerns about complexity, whereas younger generations are more inclined to embrace AI solutions for managing energy because of their technological comfort. This creates challenges in finding a balance between simplicity and advanced features to accommodate all family members. It's crucial to ensure that the smart system caters to diverse preferences without compromising efficiency. Our study shows the potential of AI in the large family with more generations. Participants, who live in the households with more co-inhabitants, expressed the need to involve AI in their energy management system. This can be indicated they host more devices, consume more energy and involve more necessity to improve energy efficiency. The gap between different relationships with co-habitants fosters more design space for the designers and engineers to provide a seamless, user-friendly AI system. This can be a diverse interface based on different users, shared control and identification.

*Society Level: Special Characteristic of Greece's Society on AI integration in smart homes.* Recent data indicates that family composition and household dynamics in Greece have been undergoing significant changes, reflecting broader social and economic trends. The traditional extended family model, once prevalent, is gradually giving way to nuclear families and alternative living arrangements. Economic factors have led to a rise in multi-generational households as a strategy to manage resources more effectively. Due to the rise of multi-generational households, our findings show the efficiency of AI in smart homes in large families which is a good hint for adopting AI in Greece's smart homes.

Besides, households in Greece face issues related to the energy performance of their dwellings [1]. The prevalence of old dwellings, many of which were built in the 1970s, presents a significant barrier to adopting modern smart home technologies. It has been reflected in our interview. These outdated buildings often lack the necessary infrastructure for efficient AI integration and energy management systems, complicating efforts to implement smart home solutions effectively. The government's efforts to refurbish old buildings present another challenge and barrier to the adoption of smart home and AI technologies in Greece.

## 6 Limitation and future work

There are several limitations in our study. First, the unbalanced gender of living lab users, most of whom are male and might lack a perspective from women; this may lead to a biased vision towards results. Second, the interview was conducted in the winter season of Greece, which is the least energy-consuming season under the warm winter weather conditions in Greece, this also may lead to a biased answer. Third, most of the interviews are conducted in their workplace; the tension and constraint of work might reduce their openness and braveness in their expression. Fourth, the number of interviewees might be a limitation of our study, preventing a comprehensive understanding of the perceptions across all Greek families. Finally, because this work is under a broad project of Plegma Living Lab, we do not need to pay for the interviewees; this might reduce their engagement in the interview process and dialogue. Future work will involve more viewpoints and investigate another perspective on AI in smart homes, with the ambition of identifying the barriers to the implementation of AI in smart homes and predicting their future.

## 7 Conclusion

Our study aimed to address two research questions: the concerns, perceptions, and expectations of Greek residents regarding the integration of AI technologies in smart homes (RQ1) and the differences in acceptance and trust towards AI technologies in smart homes among different household compositions(RQ2). Through first-hand interviews with Living Lab users in Athens,Greece, we answer the research question both quantitative and qualitatively.

For RQ1, our findings reveal that users want to balance control and dependence in AI-integrated smart homes. They want AI systems can operate autonomously to improve efficiency and convenience, while still allowing users to maintain ultimate control. In the smart home context, they also want to keep the smart home traditional home feeling. For RQ2, our findings reveal that AI holds significant potential in accommodating large family sizes; there are many design spaces for this. However, barriers to adoption exist, particularly in the outdated infrastructure of Greek homes. Despite this, the benefits are notable, particularly due to Greece's abundant solar and wind resources. By harnessing green energy generation plans, redistributing energy among households, and using AI in the smart grid system, significant gains in energy efficiency are achievable.

Greece, with its natural advantages and government promotion of smart meters, is progressing toward smart homes. Moving forward, addressing infrastructure challenges and broadening research to more Greek households are essential for maximizing AI's potential for energy efficiency in smart homes.

**Acknowledgements** This work has been funded by the EU Horizon 2020 Marie Skłodowska-Curie International Training Network GECKO, Grant number 955422 (https://gecko-project.eu/).

## References

1. Chatzikonstantinou, E.: Housing, energy performance and social challenges in Greece. A quantitative approach. In: E3S Web of Conferences, vol. 436, p. 01015 (2023). https://doi.org/10.1051/e3sconf/202343601015
2. Hsu, Y., Chou, P., Chang, H., Lin, S., Yang, S., Su, H., Chang, C., Cheng, Y., Kuo, Y.: Design and implementation of a smart home system using multisensor data fusion technology. Sensors **17**, 1631 (2017). https://doi.org/10.3390/s17071631

3. Pala, Z., Ozkan, O.: Artificial intelligence helps protect smart homes against thieves. Dümf Mühendislik Dergisi (2020). https://doi.org/10.24012/dumf.700311
4. Olufisoye, A., O, O.: Design of a voice-based intelligent prototype model for automatic control of multiple home appliances. Trans. Mach. Learn. Artif. Intell. 4 (2016). https://doi.org/10.14738/tmlai.42.1899
5. Jin, L., Boden, A., Shajalal, M.: Automated decision-making systems in smart homes: a study on user engagement and design (2022)
6. Cannata, M.: The digitalization of the electricity system: Impact assessment of digital technologies on the electricity system and its main stakeholders (2023)
7. Jin, L., Boden, A.: The design space of building user-centered AI user interfaces for smart heating systems (2023)
8. Bach, T., Khan, A., Hallock, H., Beltrão, G., Sousa, S.: A systematic literature review of user trust in AI-enabled systems: An HCI perspective. Int. J. Human-Comput. Interact. **40**, 1251–1266 (2022). https://doi.org/10.1080/10447318.2022.2138826
9. Cirqueira, D., Nedbal, D., Helfert, M., Bezbradica, M.: Scenario-based requirements elicitation for user-centric explainable AI, 321–341 (2020). https://doi.org/10.1007/978-3-030-57321-8_18
10. Sheth, J.: Designing an empathetic user-centric customer support organization: Practitioners' perspectives. Eur. J. Mark. **58**, 845–868 (2024). https://doi.org/10.1108/ejm-05-2022-0350
11. Landay, J.: For more human-centered AI, ethics should be a part of the development process (2023). https://english.hani.co.kr/arti/english_edition/e_business/1097275
12. Lehtiö, A., Hartikainen, M., Ala-Luopa, S., Olsson, T., Väänänen, K.: Understanding citizen perceptions of AI in the smart city. AI & Soc. **38**(3), 1123–1134 (2023)
13. Pliatsikas, P., Economides, A.A.: Factors influencing intention of Greek consumers to use smart home technology. Appl. Syst. Innovat. **5**(1), 26 (2022)
14. Yang, H., Lee, H., Zo, H.: User acceptance of smart home services: an extension of the theory of planned behavior. Indust. Manag. Data Syst. **117**, 68–89 (2017). https://doi.org/10.1108/imds-01-2016-0017
15. Chin, C.: Exploring the usage intention of AI-powered devices in smart homes among millennials and zillennials: the moderating role of trust. Young Consum. Insight Ideas Respons. Market. **25**, 1–27 (2023). https://doi.org/10.1108/yc-05-2023-1752
16. Sieger, L.N., Hermann, J., Schomäcker, A., Heindorf, S., Meske, C., Hey, C.-C., Doğangün, A.: User involvement in training smart home agents: increasing perceived control and understanding. In: Proceedings of the 10th International Conference on Human-Agent Interaction, pp. 76–85 (2022)
17. Lau, J., Lau, J., Harun, Z.: Factors influencing the intention to adopt smart home technology among households in Johor Bahru, Malaysia. Int. J. Acad. Res. Business Soc. Sci. 12 (2022). https://doi.org/10.6007/ijarbss/v12-i12/16016
18. Athanasoulias, S., Katsari, A., Savvakis, M., Kalogridis, S., Ipiotis, N.: An interoperable and cost-effective IoT-based framework for household energy monitoring and analysis. In: Proceedings of the 16th International Conference on PErvasive Technologies Related to Assistive Environments, 589–595 (2023)
19. Athanasoulias, S., Sykiotis, S., Kaselimi, M., Protopapadakis, E., Ipiotis, N.: A first approach using graph neural networks on non-intrusive-load-monitoring. In: Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environments, pp. 601–607 (2022)
20. Athanasoulias, S., Philippides, A.: Autonomous visual navigation: a biologically inspired approach (2022)
21. Dobrovolskis, A., Kazanavičius, E., Kizauskiene, L.: Building XAI-based agents for IoT systems. Appl. Sci. **13**, 4040 (2023). https://doi.org/10.3390/app13064040
22. Kumar, N., Sundaram, K., R, N., Madhumathi, S.: Optimizing energy consumption in smart homes using machine learning techniques. In: E3S Web of Conferences, vol. 387, p. 02002 (2023). https://doi.org/10.1051/e3sconf/202338702002

23. Wright, D., Shank, D.: Smart home technology diffusion in a living laboratory. J. Tech. Writ. Commun. **50**, 56–90 (2019). https://doi.org/10.1177/0047281619847205
24. He, T., Jazizadeh, F.: Proactive smart home assistants for automation-user characteristic-based preference prediction with machine learning techniques. Comput. Civil Engin. **2021**, 271–278 (2021)
25. Zhou, T.: The impact of privacy concern on user adoption of location-based services. Indust. Manag. Data Syst. (2011). https://doi.org/10.1108/02635571111115146
26. Kang, Y., Su, Y.-S.: A literature analysis of consumer privacy protection in augmented reality applications in creative and cultural industries: A text mining study. Front. Psychol. (2022). https://doi.org/10.3389/fpsyg.2022.869865
27. Wang, L., Sun, Z., Dai, X., Zhang, Y., Hu, H.: Retaining users after privacy invasions. Inf. Technol. People (2019). https://doi.org/10.1108/itp-01-2018-0020
28. Maseeh, H.I.: Exploring the privacy concerns of smartphone app users: a qualitative approach. Market. Intell. Plann. (2023). https://doi.org/10.1108/mip-11-2022-0515
29. He, J.: The impact of users' trust on intention to use the mobile medical platform: evidence from China. Front. Public Health (2023). https://doi.org/10.3389/fpubh.2023.1076367
30. Yao, Y., Huang, L., He, Y., Ma, Z., Xu, X., Mi, H.: Reviewing and reflecting on smart home research from the human-centered perspective. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, pp. 1–21 (2023)
31. Houzé, É.: Explainable artificial intelligence for the smart home: enabling relevant dialogue between users and autonomous systems (2019)
32. Shneiderman, B.: Human-centered artificial intelligence: three fresh ideas. AIS Trans. Human-Comput. Interact. **12**(3), 109–124 (2020)
33. Wen, F., Wang, C., Lee, C.: Progress in self-powered sensors-moving toward artificial intelligent and neuromorphic system. Nano Res. **16**(9), 11801–11821 (2023). https://doi.org/10.1007/s12274-023-5879-4
34. Jin, L., Boden, A.: Review on the application areas of decision-making algorithms in smart homes. In: HHAI 2023: Augmenting Human Intellect, pp. 74–92. IOS Press (2023)
35. Usmani, U.A., Happonen, A., Watada, J.: Human-centered artificial intelligence: Designing for user empowerment and ethical considerations. In: 2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), pp. 1–5. IEEE (2023)
36. Nazari, F.: Electric vehicle adoption behavior and vehicle transaction decision: estimating an integrated choice model with latent variables on a retrospective vehicle survey. Transp. Res. Record J. Transp. Res. Board (2023). https://doi.org/10.1177/03611981231184875
37. Negi, D.S., Birthal, P.S., Kumar, A., Tripathi, G.: Farmers' social networks and the diffusion of modern crop varieties in India. Int. J. Emerg. Mark. (2020). https://doi.org/10.1108/ijoem-04-2020-0407
38. Ogada, M.J., Mwabu, G., Muchai, D.: Farm technology adoption in Kenya: a simultaneous estimation of inorganic fertilizer and improved maize variety adoption decisions. Agricult. Food Econ. (2014). https://doi.org/10.1186/s40100-014-0012-3
39. Munyigi, L.: Factors influencing management of human excrement through biogas technology in Mukuru Kwa Njenga slums, Nairobi City County, Kenya. East African J. Environ. Nat. Resource (2024). https://doi.org/10.37284/eajenr.7.1.1918
40. Bonett, M.C., Castillo, N.G., Sanz, J.J.G.: An analysis of the perception of intelligence by different stakeholders in the ambient assisted living domain. XVIII Conferencia de la Asociación Española Para la Inteligencia Artificial (CAEPIA 2018): Avances en Inteligencia Artificial, 87–92 (2018)
41. Adami, I., Foukarakis, M., Ntoa, S., Partarakis, N., Stefanakis, N., Koutras, G., Kutsuras, T., Ioannidi, D., Zabulis, X., Stephanidis, C.: Monitoring health parameters of elders to support independent living and improve their quality of life. Sensors **21**(2), 517 (2021)
42. Athanasoulias, S., Guasselli, F., Doulamis, N., Doulamis, A., Ipiotis, N., Katsari, A., Stankovic, L., Stankovic, V.: The Plegma dataset: domestic appliance-level and aggregate electricity demand with metadata from Greece. Sci. Data **11**(1), 376 (2024)

43. Wang, G.: Humans in the loop: the design of interactive AI systems. https://hai.stanford.edu/news/humans-loop-design-interactive-ai-systems. Accessed 20.10.2019
44. Liao, J., Xiao, C., Kim, H.-Y.: Mapping a decade of smart homes for the elderly in web of science: a scientometric review in CiteSpace. Buildings (2023). https://doi.org/10.3390/buildings13071581
45. Schmager, S., Pappas, I., Vassilakopoulou, P.: Defining human-centered AI: a comprehensive review of HCAI literature. In: Proceedings of the Mediterranean Conference on Information Systems (2023)

# Business Intelligence for Green Decision-Making

# Towards Enhanced Manufacturing Precision: The Role of AI-Driven Quality Grading Systems

**Mohamed Ghoneim, Radwa Hussein, and Nada Sharaf**

**Abstract** This paper introduces a comprehensive grading system to enhance product quality control in industrial and manufacturing processes. The system includes live color detection, damage identification, quality classification, and robotic arm manipulation. Live color detection allows real-time analysis of product colors, while damage detection uses computer vision algorithms to identify defects. Quality classification uses machine learning techniques to categorize products based on predefined parameters. The SVM model, with RBF kernel, was found to be the most effective for quality classification, achieving an accuracy of 94.78% for oranges and 84.77% for pallets. YOLOv8 was used for defect detection, with the most successful run resulting in 88.55% precision, recall, F-Score, and mAP at 50.70%. The system also integrates a robotic arm for real-time manipulation and removal of detected defects. FlexiGrade, the general-purpose automatic product grading system, is a versatile framework that can be adapted and enhanced for various products, making it a robust and scalable solution for different needs.

**Keywords** Product Grading · Artificial intelligence (AI) · Machine learning (ML) · Computer vision

M. Ghoneim · R. Hussein · N. Sharaf (✉)
The German International University, New Administrative Capital, Egypt
e-mail: nada.hamed@giu-uni.de

M. Ghoneim
e-mail: mohamed.ghoneim@student.giu-uni.de

R. Hussein
e-mail: radwa.hassan@giu-uni.de



V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_5

## 1 Introduction

The increasing reliance on automated systems for assessing and classifying manufactured goods plays a critical role in making production processes more efficient while maintaining consistent quality standards in an increasingly competitive market. These systems leverage advanced technologies such as artificial intelligence (AI), machine learning (ML), and computer vision to analyze and classify products, offering accurate, reliable assessments with minimal human intervention. Despite these advancements, significant challenges remain, especially in the context of diverse industries and product types. These challenges include creating robust datasets, adapting models to a wide variety of product categories, addressing variations in appearance and quality, and ensuring scalability across different manufacturing environments.

This research seeks to employ a range of machine learning models and computer vision techniques for automatic product grading systems, with a focus on evaluating different architectures to optimize both accuracy and efficiency. The ultimate objective is to contribute valuable insights and methodologies that enhance the development and practical deployment of these systems, while also ensuring that the interfaces are user-friendly for seamless integration into industrial operations.

While most research on automatic product grading has primarily focused on perishable items such as fruits and vegetables, a significant gap remains in applying this technology to industrial products like machinery parts, car components, and manufactured goods like plastic pallets. This lack of research makes it difficult to fully understand the potential improvements automated systems could bring to manufacturing efficiency, quality control, and supply chain optimization. Bridging this gap offers the potential to significantly enhance industrial processes by making quality assessments more efficient and optimizing production workflows in the context of Industry 4.0. Additionally, there is a need to ensure the adaptability of automated grading technologies across diverse industries, as many specialized systems may be limited in their ability to scale and adjust to the varying demands of different sectors.

Furthermore, adopting these automated grading systems can have a positive impact on sustainability. By improving accuracy in quality control, manufacturers can reduce waste from defective products, minimize energy consumption by optimizing production processes, and ensure that resources are used more efficiently. Sustainable practices, such as predictive maintenance and enhanced lifecycle management of industrial products, can be integrated into the automated workflows, ultimately contributing to reduced environmental footprints. The AI-driven grading system can also reduce the waste by ensuring that only products, with a certain quality level are accepted into the next stages of production. This in turn xan reduce, material waste as well as along with energy and carbon footprint.

## 2 Product Grading Systems

Product grading systems have evolved over time, with early studies focusing on automated assessment in sectors like agriculture and food production. Computer vision has been used in various industries, including agriculture, food production, and fruit grading [2]. Researchers have developed machine vision techniques for assessing and sorting fruits based on quality attributes, such as size, color, stem location, and blemishes. Image processing techniques have also been explored, showcasing the potential of image processing for food quality evaluation [1].

In the early 2010s, product grading systems saw significant strides, with the integration of sophisticated technologies such as computer vision, machine learning, and sensor technologies. Systems like the one developed for strawberries achieved high accuracy rates in size detection, color grading, and shape classification within a short time [9]. The system also addressed complex classification tasks with multiple classes and properties, reducing errors by up to 15 percentage points.

In recent years, product grading systems have continued to evolve rapidly, driven by advancements in AI, data analytics, and IoT. These systems are capable of handling complex grading tasks across diverse industries, showcasing the potential for transformative impacts within Industry 4.0 and smart manufacturing. Data processing methods, such as machine vision and spectroscopy, are crucial for assessing fruit quality throughout different processing stages [11]. Automated product grading systems have been developed to automate the grading process of areca nuts, minimizing manual inconsistencies [10].

Fruit grading systems are crucial in determining the quality of fruits using various methods and technologies. Different grading methods are applied based on their distinctive appearance and structure, such as eggplants, which use standard measurements and glossiness to categorize fruits into five sizes and four grades [3]. On the other hand, strawberries undergo a different grading process using machine-vision technology, which relies on automated assessments of shape, size, and color [9]. The research focuses on the development and evaluation of an automated strawberry grading system using image processing technology, aiming to test and validate the accuracy and efficiency of the grading algorithm and mechanical components through experimental trials involving a variety of strawberries. Numerous methodologies have been explored for the assessment of even a single fruit, such as MLP-Neural Networks, geometrical models, each contributing unique perspectives and techniques to the field [12].

Research on grading systems in manufacturing is advancing, focusing on automated solutions to ensure product quality and streamline production processes. Advancements in Computer-Aided Design/Computer-Aided Manufacturing (CAD/CAM) have led to innovative grading techniques that enhance efficiency and reduce errors [6]. The integration of robotic systems into the shoe manufacturing industry is explored as a means to automate intricate manual tasks, showcasing potential applications in polishing, cleaning, packaging, and visual inspection. The influence of Artificial Intelligence (AI) within the Industry 4.0 paradigm transforms manufacturing by enhancing efficiency through predictive insights and intelligent decision support systems. The integration of robotic systems in the shoe manufacturing industry emphasizes the need for cooperative robotic systems capable of collaboration with human workers [5]. In the ever-evolving landscape of Industry 4.0, 3D printing technology has gained prominence, offering cost and time efficiency advantages. This study [8] proposes a comprehensive procedure for anomaly detection at each stage of 3D product development using machine learning algorithms and pre-trained models.

## 3 FlexiGrade

This section introduces FlexiGrade, a general-purpose grading system, detailing its creation, dataset creation, and user interface design, providing detailed technical processes.

Figure 1 serves as a visual roadmap showing all the system components. The decision of the robotic arm is based on all the processing done throughout the different stages.

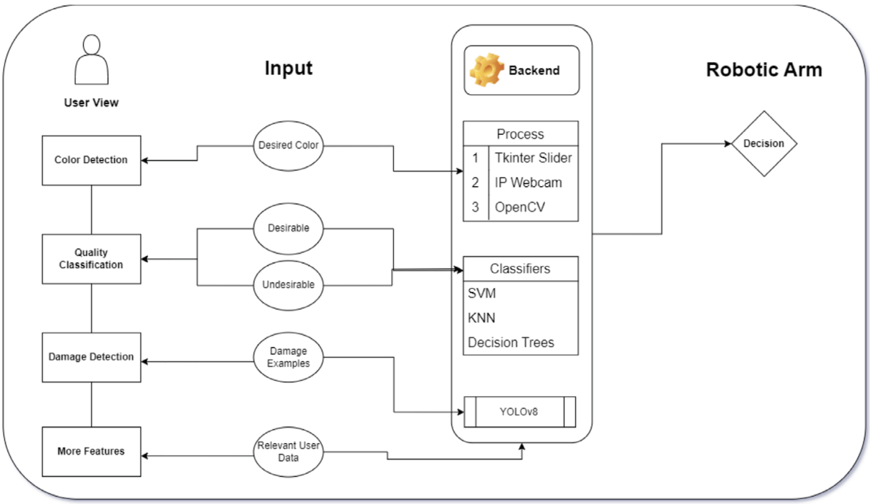



**Fig. 1** System overview: interface, backend processing, and robotic arm decision diagram

## *3.1 User Interface Design*

The user interface section of the provided code is responsible for creating a graphical interface for a product grading system. This system incorporates color detection, damage detection, and custom quality assessment functionalities. The UI is built using the tkinter library in Python, and it features buttons and labels for user interaction (Fig. 2).

The user interface of the product grading system facilitates a straightforward process for assessing different grading criteria. Upon launching the application, the main window titled "Product Grading System" presents a set of buttons representing various grading aspects, including "Color Detection," "Damage Detection," and "Custom Assessment."

To initiate the evaluation for a specific criterion, the user clicks the corresponding button. For example, selecting "Color Detection" opens a new window titled "Color Detection and Capture." Within this window, the user can fine-tune the color range by adjusting sliders for lower and upper hue values. Clicking the "Detect Color!" button initiates the image capture process, and a pop-up message promptly displays the result, such as "Good" or "Bad." Following this, the user can close the color detection window.

The user can then proceed to assess other criteria by clicking the respective buttons on the main screen, such as "Damage Detection" or "Custom Assessment." Placeholder actions are executed for these criteria, allowing flexibility for future implementation.

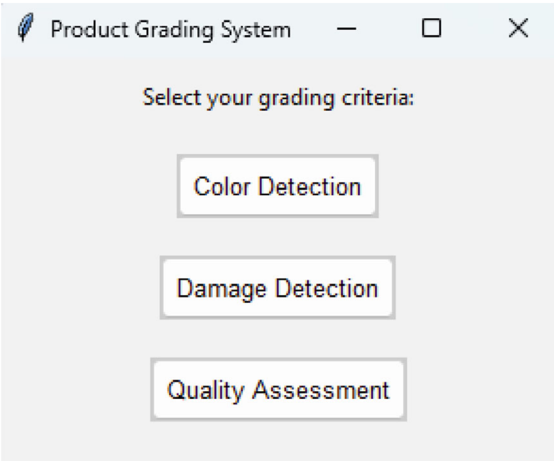



**Fig. 2** Simple user interface for the general purpose grading system

## 3.2 Feature Implementation

This section outlines the integration of live color detection, damage detection, custom quality classification, and additional attributes into a versatile grading system for diverse applications.

### 3.2.1 Live Color Detection

The live color detection feature integrated into the grading system is an application of computer vision principles. This functionality utilizes Python along with OpenCV and Tkinter libraries to access live video streams transmitted from an IP Webcam app. Once the system retrieves the video frames, it initiates a series of image processing steps. Initially, these frames undergo a conversion process from the conventional RGB color space to the HSV (Hue, Saturation, Value) color model. This conversion proves instrumental in isolating and analyzing specific colors within the image due to HSV's inherent ability to separate color information effectively.

The feature boasts a user-friendly GUI which allows users to interactively modify and fine-tune the color detection parameters. Through sliders, users can adjust the lower and upper hue thresholds, delineating the specific range of colors they aim to detect and evaluate. This ensures compatibility with various products and situations where significant color variations or standards may exist.

Upon setting the desired color range, the system proceeds to identify and highlight the detected colors in real-time within the video stream. In the following figure, the HSV ranges for the color Green have been chosen. The user is able to a changing spectrum as they move the slider, which follows this reference reffig:colormaphsv.

**lower** = `np.array([40, 150, 20])`

**upper** = `np.array([77, 255, 255])`

The Hue component represents the type of color. It typically ranges from 0 to 360 degrees (or 0 to 180 in OpenCV or other software that uses the 0-180 scale). It's essentially the base color, such as red, green, blue, etc. In the code, `lower[0] = 40` and `upper[0] = 77` represent the lower and upper limits of the Hue component, respectively. The values `0` to `30` in Hue correspond to a range of colors between yellowish-green and blueish-green hues in the HSV color space, for example (Fig. 3).

Saturation represents the intensity or purity of the color and ranges from 0 to 100 %. Higher saturation means more vivid colors, while lower saturation tends towards shades of gray. In the code, `lower[1] = 150` and `upper[1] = 255` represent the lower and upper limits of the Saturation component, respectively. The values `150` to `255` in Saturation represent a relatively high saturation level, indicating more vivid and intense colors within the specified Hue range.

Value or Brightness represents the brightness of the color, ranging from 0 (black) to 100%(full brightness or white). In the code, `lower[2] = 20` and `upper[2] = 255` represent the lower and upper limits of the Value component, respectively.

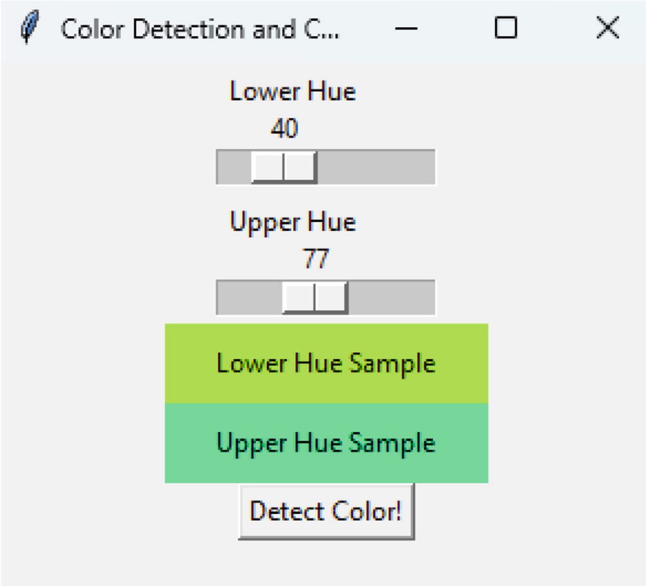



**Fig. 3** Color range picker

The values `20` to `255` in Value indicate a range from relatively dark colors (brightness level `20`) to maximum brightness (`255`), within the specified Hue and Saturation range.

To link between the phone's camera and the computer judgement system, it is mandatory to open a server on the Ip Webcam app. It is also important to make sure that the IPV4 address that is visible on the bottom corner of the interface is the same as the URL listed in the code for both systems to connect.

Employing robust color detection algorithms, the system places bounding boxes around the identified color regions, allowing users to visually confirm the system's accuracy in color detection. A vibrant green color is discernible within a red bounding box on the judgement platform (laptop), denoting successful color detection within the color range selector.

Additionally, it uses a temporal constraint mechanism to distinguish between desired and undesired color presence over a specified duration. If the system detects the presence of the color consistently for a predefined time frame, it promptly provides a notification denoting the quality of the detected color as either "Good" or "Bad."

If no color is consistently detected for a user-defined duration of x seconds consecutively, a popup indicating the issue appears. This is followed by the movement of the robotic arm to remove the object.

Provided there is a consecutive feed where an object has the same color as a hue inside the spectrum range chosen by the user for x (determined by the user) seconds straight, then the user receives a popup with the message "Good" appearing, instead.

### 3.2.2 Quality Classification

Including a quality assessment feature in the system, the implementation allows for the distinction between good and defective oranges. The orange dataset encompasses 1,466 instances of fresh oranges and 1,595 instances of rotten or defective ones, sourced from Kaggle. Utilizing SVM, KNN, and decision tree classifiers, the system processes these images. Integrated with an IP webcam, the live video feed involves initial object detection, followed by classification. If the model identifies an object as being poor quality, the robotic arm intervenes, removing it to maintain an uninterrupted feed.

The purpose of the quality classification feature in a general-purpose grading system is to accurately distinguish between different levels of quality or condition among items being evaluated, thereby facilitating efficient sorting or categorization based on predetermined criteria.

In this context, the quality classifications were limited to either "good" or "bad," as the specific thresholds for what constitutes acceptable or unacceptable quality in pallets vary among manufacturers. Generally, some manufacturers refrain from releasing a pallet for commercial use if it exhibits even minor imperfections. Conversely, for more general purposes such as storage, some manufacturers might tolerate multiple fractures as long as they do not compromise the structural integrity or load-bearing capacity. This variability allows for the provision of multiple grades within the system. It highlights the system's flexibility, as users can customize the number of quality delineations based on their specific requirements.

### 3.2.3 Damage Detection

The Damage Detection Feature implemented within the grading system is a pivotal module employing advanced object detection techniques, specifically employing YOLOv8 architecture. This feature aims to identify and delineate various types of damage present on plastic pallets through image analysis. The system processes images containing plastic pallets, scanning for distinct damage categories: Voids, encompassing missing pieces or substantial damage; blemishes, representing scratches or minor defects; discolorations, indicating deviations from standard color or hue; and scorches, indicative of burn-related damage. Upon receiving these images, YOLOv8 detects and localizes damaged areas.

By integrating YOLOv8, the system efficiently creates bounding boxes around the identified damaged segments within the plastic pallets. This feature enables precise mapping of the damaged regions, aiding in swift and accurate assessments of the type and severity of the detected damage. The distinct classification of damage categories facilitates an organized and systematic inspection process, allowing for the efficient categorization of damage types. This functionality is instrumental in quality assessment procedures across various industrial sectors, contributing significantly to quality control measures and ensuring product integrity and reliability.

In implementing the damage detection segment of the project, the system incorporates two distinct use cases-one tailored for organic products and the other representing manufacturing or non-bio products. As for organic items, oranges were the produce of choice. the system demonstrates a nuanced ability to identify various imperfections on the skin and label them as “defects”. Those defects encompass any imperfection visible on the skins of the oranges (Fig. 4).

Transitioning to the manufacturing or non-bio context, where plastic pallets are the focal point, the system showcases a robust capability to detect a range of potential issues. These include scorch damage, blemishes, manglements, voids (missing parts), and discolorations. Notably, the system extends its proficiency to identify the blending of colors resulting from transitions between pallets (Fig. 5).

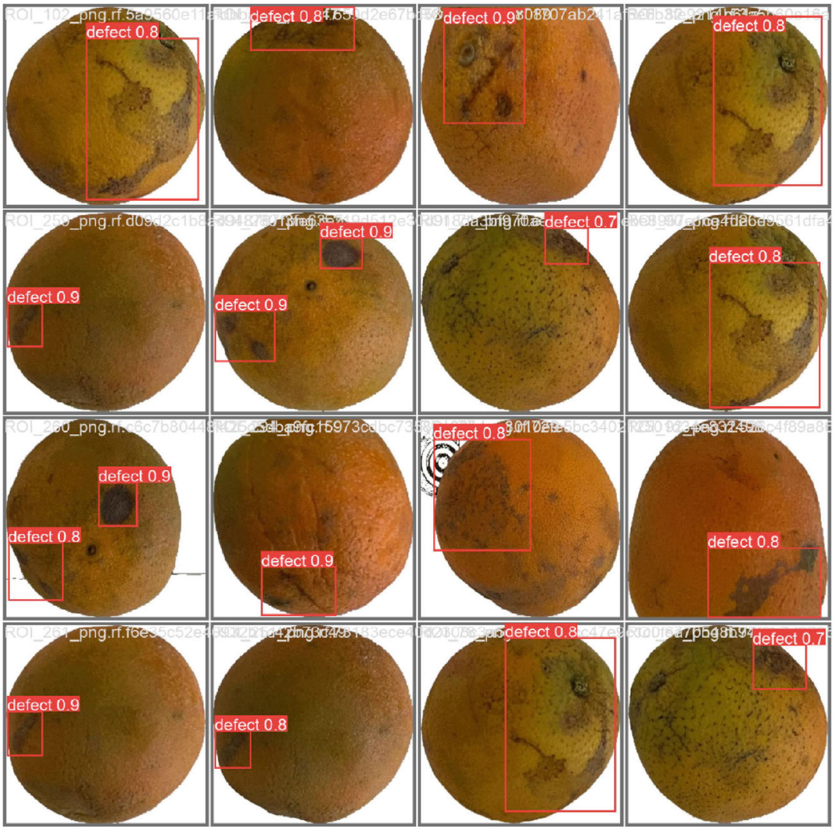



**Fig. 4** Orange defect detection

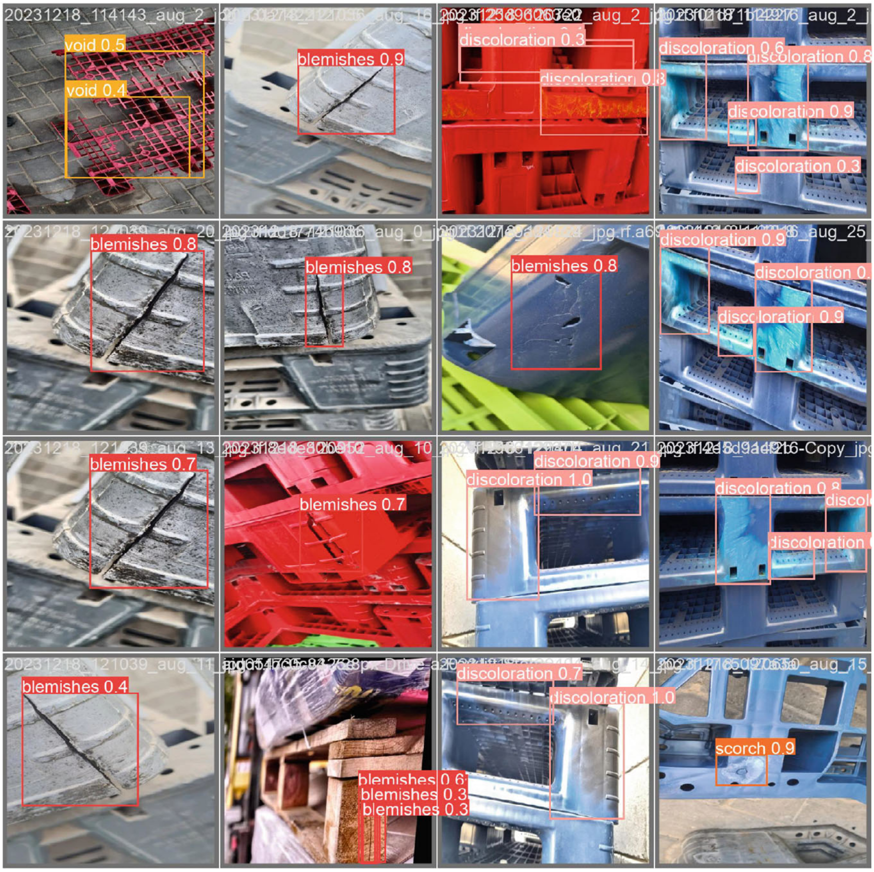



**Fig. 5** Pallet damage detection

## *3.3 Dataset Creation and Preparation*

A comprehensive dataset was created for analyzing plastic pallets, encompassing various attributes and conditions. The dataset includes images of damaged and pristine pallets, including voids, blemishes, discolorations, and scorches. The dataset was curated to capture diverse perspectives, angles, lighting conditions, and contextual variations.

The dataset was gathered from over 130 images of plastic pallets from the “PAFT” factory, a regional leader in supply chain solutions. The images were carefully captured, documenting various types and conditions of plastic pallets. The dataset serves as a foundational resource for training and validating systems for assessing plastic pallet quality, offering a diverse spectrum of visual data.

The Dataset Labeling process involved a manual annotation approach to accurately delineate and catalog the various attributes and damages present in the

**Fig. 6** Damage types detected

images of plastic pallets. This involved carefully examining each image to identify and demarcate specific regions associated with damages: blemishes, discolorations, voids, and scorches (Fig. 6).

Blemishes are imperfections that affect an object's appearance or surface quality, while discolorations are alterations in color or tone that deviate from the expected appearance. Voids are empty or unfilled spaces within a material or object that should ideally be solid or filled, often considered defects. Scorches are areas of localized burning or charring on a material's surface caused by excessive heat exposure.

To accurately represent damages, manual bounding was used to highlight each type of damage in the dataset. Roboflow's web-based annotation platform was used for labeling, simplifying the annotation task. Data augmentation is a technique in machine learning and computer vision that artificially expands a dataset by creating modified versions of existing data samples. Preprocessing of images involved an Auto-Orient process for proper orientation and resizing to a standardized dimension. Augmentation techniques during training included various transformations, such as horizontal and vertical flipping, rotations, cropping, rotation adjustments, saturation, brightness, exposure modifications, blur effects, and noise introduction, to enhance model robustness, generalize learning, and improve adaptability to diverse real-world scenarios.

It is noteworthy that datasets for quality classification and defect detection differ due to their distinct purposes and data types. Quality classification distinguishes quality levels, while defect detection identifies flaws or anomalies, necessitating tailored datasets for each task.

Some criteria of removal included confidentiality, privacy, and uniformity. Images were removed due to top-down viewpoints, quality, resolution, or content deviations to create a cohesive, focused dataset, free from inconsistencies and irrelevant variations.

Through augmentation, the dataset has been expanded to 166 images, divided into three subsets: training (70%), validation (11%), and test (19%). The training set consists of 117 images, the validation set includes 18 images, and the test set held 31.

The dataset comprises annotated images for pallets and oranges. For pallets, there are 166 annotated images for damage detection and 753 images for quality assessment, including 331 good pallets and 422 bad pallets. For oranges, there are 719 annotated images for damage detection and 3,061 images for quality assessment, with 1,466 rotten oranges and 1,595 fresh ones. Annotation was performed using Roboflow for consistency and accuracy.

# 4 Results

## 4.1 Classification Results

This table compares the classification performance of different models for identifying oranges. Each row corresponds to a specific model, with metrics such as accuracy, precision, recall, and F-score provided. The models considered in this analysis include SVM, KNN, and Decision Trees.

We have two datasets for experimentation. In the case of oranges, the classification is between fresh oranges, comprising 1,466 images, and rotten oranges, which consists of 1,595 images. Regarding pallets, there are 331 images representing good pallets and 422 images depicting bad pallets.

The models were trained with default settings. Specifically, for the SVM, the kernel used was linear. In the case of the KNN algorithm, the number of neighbors was set to 2 for both pallets and oranges. **Starting with Orange Classification:**

The KNN model outperformed SVM and Decision Trees in orange classification, achieving the highest accuracy of 92.5%. Its simplicity and proximity-based approach make it well-suited for the task. Decision Trees showed slightly lower performance than both models across multiple metrics as shown in Table 1.

**Table 1** Baseline classification results for oranges

| Model | Accuracy (%) | Precision (%) | Recall (%) | F-score (%) |
|---|---|---|---|---|
| SVM | 92.3 | 92.0 | 92.0 | 92.0 |
| KNN | 92.5 | 93.0 | 93.0 | 93.0 |
| D.T | 92.3 | 87.0 | 87.0 | 87.0 |

**Table 2** Classification results for pallets

| Model | Accuracy (%) | Precision (%) | Recall (%) | F-score (%) |
|---|---|---|---|---|
| SVM | 61.6 | 62.0 | 61.0 | 61.0 |
| KNN | 67.5 | 77.0 | 67.0 | 72.0 |
| D.T | 61.6 | 78.0 | 77.0 | 77.0 |

As shown in Table 2, the KNN model is the most effective for pallet classification, achieving the highest accuracy at 67.5% and minimizing false positives. Compared to Decision Trees, KNN's proximity-based approach captures intricate patterns and variations, making it the preferred choice for accurate and reliable classification. SVM slightly lagged behind in accuracy and F-score. As per our best knowledge, there is a notable absence of classification models in the existing literature specifically tailored to plastic pallets. Upon analyzing the classification results for pallets in this study, it becomes evident that the KNN model distinguishes itself as the most effective among the three considered models from baseline. These results could serve as a minimum threshold for future research in this specific domain.

The next analysis will focus on the performance of three models, namely YOLOv8n, YOLOv8s, and YOLOv8m, in object detection within images.

The baseline settings for object detection for both oranges and pallets were as follows: a default batch size of 16, an automatic optimizer (AdamW), and a learning rate of 0.00125. The chosen model was YOLOv8n. The training process involved running 15 epochs for the detection of oranges, while for the detection of damaged pallets, we ran a total of 50 epochs.

As shown in Table 3, the object detection metrics reveal that the model performs well in identifying defects in oranges despite low epochs, with high precision (74.5%) and recall (82.3%). It also shows good accuracy in localizing oranges (mAP50 at 84.12%). However, there's room for improvement in detecting oranges across different IoU thresholds (mAP50-95 at 55.47%). For pallets, the model achieves moderate precision (78.81%) but struggles with recall (51.29%), resulting in a balanced F-Score of 62.09%. Accuracy in pallet localization at a 50% IoU threshold is moderate (mAP50 at 59.36%), but there's considerable room for improvement across IoU thresholds (mAP50-95 at 29.88%).

**Table 3** Object/damage detection metrics

| | Precision | Recall (%) | F-score (%) | mAP50 (%) |
|---|---|---|---|---|
| Orange | 74.5 | 82.3 | 78.3 | 84.1 |
| Pallets | 78.8 | 51.3 | 62.1 | 59.4 |

# 5 Hyperparameter Tuning

Hyperparameter tuning is the process of adjusting parameters in a model, such as architecture, number of neighbors, kernel, epoch number, and tree depth, to influence the learning process.

## 5.1 *Classification Tuning*

### 5.1.1 SVM Parameters: Kernel

The Random Basis Function (RBF) outperformed Linear in classifying orange quality, with an accuracy rate of 94.78%, precision of 95.9%, and recall rate of 95.9%. This indicates that RBF's enhanced capability in identifying defects outperforms Linear's 92.3% accuracy rate and 92.3% recall rate, indicating superior performance in orange quality classification.

The Random Basis Function (RBF) kernel outperformed Linear in classifying pallets, demonstrating superior accuracy (84.77%), higher precision (86.8%), better recall (85%), and an 85% F-Score (85%). This indicates enhanced predictive capabilities, avoiding false positives, and capturing more relevant instances. The RBF kernel outperformed Linear in all measured metrics, indicating superior performance in pallet classification for this use-case.

### 5.1.2 KNN Parameter

The evaluation of KNN models for an automatic product grading system using an orange dataset reveals varying performance based on the number of neighbors. With n_neighbors set to 2, the model achieves 92.50% accuracy, precision, recall, and F-Score at 93.00%. Even with higher n_neighbors, it maintains strong performance at 91.35%.

The performance of KNN models in an automatic product grading system varies with n_neighbors. Initially, models show moderate accuracy at 67.55%, low false positive rate, and recall at 67%. Increased n_neighbors leads to lower accuracy.

### 5.1.3 Decision Trees Parameters

The orange quality classification model's maximum depth parameter is analyzed, with a maximum depth of 5 achieving an overall accuracy of 86.30%. Increasing the depth to 5 slightly improves accuracy to 86.62% and precision, recall, and F-Score to 87.00%. A further increase to 10 results in a marginal increase to 87.28%, suggesting a pragmatic choice.

The orange classification model's performance is influenced slightly by the 'max_features' parameter. Setting it to 'None' results in an accuracy of 86.13%, but may lead to sub-optimal generalization. However, setting it to 'sqrt' improves the model's performance by achieving an accuracy of 88.09% and precision, recall, and F-Score of 88.00%. Setting it to 'log2' balances complexity and effectiveness.

The data presented illustrates a pattern in the accuracy of the pallet classification model concerning different depth configurations. Initially, with a shallow tree depth set at 0, the model attains an accuracy of 76.16%. However, as the depth increases to 3, the accuracy slightly decreases to 74.83%, suggesting potential difficulty in discerning intricate features. Expanding the depth to 5, enabling more nuanced analysis, results in an accuracy boost to 75.50%. Notably, with a depth of 10, the model achieves its peak accuracy of 78.15%, indicating that a deeper tree structure enhances performance.

In analyzing the impact of tuning the max_features parameter on the pallet classification model, several key findings emerge. Initially, setting max_features to 'None' yields a notable accuracy of 75.50%, although it may introduce heightened model complexity. Conversely, opting for 'sqrt' or 'log2' significantly reduces accuracy to 62.25% and 66.23%, respectively, despite the imposition of feature constraints.

## 5.2 *Object Detection Parameter Tuning*

### 5.2.1 Model architecture

YOLOv8 models are object detection tasks with varying sizes, computational complexity, and performance. They range from YOLOv8n (Nano) to YOLOv8x (Extra Large), offering a trade-off between model size, inference speed, and accuracy. The training dataset consisted of 501 images, with a batch size of 16 images per iteration and an optimal learning rate of 0.00125.

YOLOv8n has the shortest training time at 3.1 min, followed by YOLOv8s at 3.4 min, and YOLOv8m with the longest training time of 5.4 min.

Table 4 compares orange defect detection scores across three YOLOv8 models: YOLOv8n, YOLOv8s, and YOLOv8m. YOLOv8n has the highest precision (0.93794), low false-positive rate (0.82292), and recall (0.84292), indicating a larger proportion of true positives. It also has the highest average precision across different overlap scenarios, suggesting better performance.

**Table 4** Orange defect detection scores versus chosen model

| Model | Precision (%) | Recall (%) | F-score (%) | mAP50 (%) |
|---|---|---|---|---|
| YOLOv8n | 74.5 | 82.3 | 78.3 | 84.1 |
| YOLOv8s | 81.6 | 67.7 | 74.1 | 82.6 |
| YOLOv8m | 93.8 | 63.0 | 75.4 | 82.4 |

**Table 5** Pallet damage detection scores versus chosen model

| Model | Precision (%) | Recall (%) | F-score (%) | mAP50 (%) |
|---|---|---|---|---|
| YOLOv8n | 78.8 | 51.3 | 62.1 | 59.4 |
| YOLOv8s | 82.2 | 62.1 | 70.7 | 64.0 |
| YOLOv8m | 59.5 | 62.1 | 60.8 | 61.9 |

The YOLOv8 models have distinct performance metrics, with YOLOv8m being a balanced choice for preprocess, inference, and postprocess stages. YOLOv8n, with its inference speed being ideal for real-time applications. Both models are efficient for resource constraints, offering a trade-off between speed and performance. YOLOv8n is the most suitable model for defect detection on oranges, demonstrating competitive F-scores and favorable speed metrics.

**Pallet Damage Detection:** The dataset is divided into three sets for training, validation, and testing. The training set has 117 images, the validation set has 18 images, and the test set has 31 images. The training process uses default settings and the AdamW optimizer.

YOLOv8n has, as expected, the shortest training time at 2.5 min, followed by YOLOv8s at 3.1 min, and YOLOv8m with the longest training time of 5.8 min.

Table 5 compares YOLOv8 models for pallet damage identification tasks, revealing YOLOv8s outperforms other models in precision, recall, F-score, mean Average Precision, and training time. YOLOv8s is more effective in identifying damaged areas on pallets, with a reasonable training time. YOLOv8n and YOLOv8m have lower precision and F-score, making them less suitable for this task.

### 5.2.2 Epochs

The Pallet Scores Versus Epochs evaluation used a dataset split of 70% for training, 11% for validation, and 19% for testing, with the YOLOv8n model focusing on Nano.

The oranges detection model showed its best performance at epoch 75, with precision increasing from 81.5% to 88.55%. Recall fluctuated, peaking at 73.96% in both the initial and final epochs. The F-Score and mAP50 varied over time, reflecting the model's improving accuracy.

For Pallet Damage Detection, using the YOLOv8s model over 50 epochs, precision started at 70.77%, while recall was low at 20.36%. By epoch 50, recall improved to 48.77%, and the F-Score reached 55.08%. At epoch 75, precision dropped to 60.02%, with recall at 62.14%, yielding an F-Score of 61.02%. By epoch 100, precision declined to 42.72%, but recall remained high at 62.99%, resulting in an F-Score of 50.99% and mAP50 of 54.86%.

## 6 Discussion

The RBF Kernel was identified as the most effective model for classifying oranges and pallets due to its ability to handle complex, nonlinear data. Its flexibility and robustness against overfitting made it preferable to Linear models.

The best object detection results were achieved using YOLOv8n for oranges and YOLOv8s for pallets at 75 epochs. The smaller models often outperformed larger ones, highlighting the importance of network architecture and parameter tuning over size. Dataset characteristics, like class imbalance, also influenced performance.

YOLOv8 was selected for its real-time processing capabilities, accuracy, and ongoing updates, making it ideal for practical object detection tasks. For detecting plastic pallet defects, SVM, KNN, and Decision Trees were chosen due to their ability to handle complex patterns and generalize across diverse datasets.

The study highlighted the need for dedicated damage detection models for plastic pallets, with YOLOv8 showing strong initial performance in terms of precision, recall, F-scores, and mAP50-95. These insights set the stage for future research in defect detection.

## 7 Conclusions

A general-purpose automatic grading system that integrates live color detection, quality classification, and damage detection, combined with the precision of a robotic arm, is a significant step forward in streamlining manufacturing processes. This system is particularly beneficial for those looking to implement automated quality control systems across diverse product types. Additionally, it is important to note that the results obtained in this study are subject to further improvement with the utilization of a substantially larger dataset. The encouraging aspect is that these baseline results for the quality classification of oranges align closely with the findings presented in the literature [4]. The comparison indicates a strong correspondence between this model's performance and established benchmarks.

FlexiGrade is a comprehensive system that offers versatile product grading across various types and metrics. Its modular design allows for easy customization, allowing users to suit different product types and grading requirements. This adaptability is crucial in dynamic manufacturing environments. FlexiGrade's seamless integration with robotic arms enhances its utility in automated processes, reducing manual intervention and enhancing efficiency. This comprehensive system stands out in the absence of existing general-purpose solutions for versatile product grading.

Future work should be directed at enhancing the functionality of the system. This includes enhancing the algorithms for better performance as well as enabling easy integration of the system into different manufacturing environments. Furthermore, larger datasets of plastic pallet defects used on this work should be availed. This is expected to produce better performing models. Automated systems for data generation should be implemented similar to the one provided in [7]. In addition, in the future, effects from shadows and illumination should be also considered in the live color detection phase.

## References

1. Blasco, J., Aleixos, N., Moltó, E.: Machine vision system for automatic quality grading of fruit. Biosys. Eng. **85**(4), 415–423 (2003)
2. Brosnan, T., Sun, D.-W.: Inspection and grading of agricultural and food products by computer vision systems-a review. Comput. Electron. Agric. **36**(2–3), 193–213 (2002)
3. Chong, V.K., Kondo, N., Ninomiya, K., Nishi, T., Monta, M., Namba, K., Zhang, Q.: Features extraction for eggplant fruit grading system using machine vision. Appl. Engin. Agricult **24**(5), 675–684 (2008)
4. Gill, J., Sandhu, P., Singh, T.: A review of automatic fruit classification using soft computing techniques (2014)
5. Gracia, L., Perez-Vidal, C., Mronga, D., et al.: Robotic manipulation for the shoe-packaging process. Int. J. Adv. Manuf. Technol. **92**(1–4), 1053–1067 (2017)
6. Hinojo-Pérez, J.J., Davia-Aracil, M., Jimeno-Morenilla, A., et al.: Automation of the shoe last grading process according to international sizing systems. Int. J. Adv. Manuf. Technol. **85**(1–4), 455–467 (2016)
7. Hussein, R., Kahar, K., Mashaly, M., Sharaf, N.: Datacrop: a generic tool for crop data set generation. In: Banissi, E., Siirtola, H., Ursyn, A., Pires, J.M., Datia, N., Nazemi, K., Kovalerchuk, B., Andonie, R., Nakayama, M., Temperini, M., Sciarrone, F., Nguyen, Q.V., Mabakane, M.S., Rusu, A., Cvek, U., Trutschl, M., Müller, H., Francese, R., Bouali, F., Venturini, G., (eds.), 27th International Conference Information Visualisation, IV 2023, Tampere, Finland, July 25–28, 2023, pp. 43–47. IEEE (2023)
8. Kumar, K.K., Srikanth, V., Prasad, G.N.R., Hazela, B., Tamrakar, A.K.: Fault detection on the 3-d printed objective surface by using the svm algorithm. In: Materials Today: Proceedings (2023)
9. Liming, X., Yanchao, Z.: Automated strawberry grading system based on image processing. Comput. Electron. Agric. **71**, S32–S39 (2010)
10. Salunke, A., Honnungar, S.: Quality grading of areca nuts harvested and processed in goa using image processing and lab view. In: Advances in Mechanical Design, Materials and Manufacture: Proceeding of the Second International Conference on Design, Materials and Manufacture (ICDEM 2019) (2020)
11. Srivastava, S., Sadistap, S.: Data processing approaches and strategies for non-destructive fruits quality inspection and authentication: a review. J. Food Measur. Characterizat. (2018)
12. Unay, D., Gosselin, B.: Apple defect detection and quality classification with mlp-neural networks 02 (2004)

# Environmental Informatics for Industry

# Use of Artificial Intelligence to Optimize Processes and Increase Resource Efficiency in Small and Medium-Sized Enterprises

**Martina Willenbacher and Volker Wohlgemuth**

**Abstract** An important driver for increasing productivity in manufacturing small and medium-sized enterprises (SMEs) is the digitization and digitalization of production processes. The associated increase in data volume offers enormous potential for analyzing and optimizing processes. Data from a variety of devices and systems increases the need for intelligent, dynamic analysis models. However, SMEs have a low to very low degree of digitalization. This is the result of a combination of various factors, such as scarce financial and human resources for research and development activities, lack of IT expertise, and a reluctance to introduce new digital technologies and artificial intelligence. Furthermore, the production processes of processing SMEs are very individual and sometimes highly specialized, so existing AI modules cannot be adapted to the existing production structure without increased adaptation effort. As part of this doctoral project, two machine learning methods were developed for practical use in a processing SME. The aim was to identify connections between energy consumption and plastic scrap and the machine settings as well as to find optimal parameter settings to increase energy efficiency and reduce the waste rate. The focus was on the simplicity of the solution and the easy adaptability to changing production processes. It could be shown that significant increases in productivity can also be achieved with less complex AI processes, the selection of which is based on a clear definition of goals.

**Keywords** Small and medium-sized enterprises (SME) · Artificial Intelligence (AI) · Machine Learning (ML) · Random Forest (RF) · Artificial Neural Network (ANN) · Resource Efficiency

---

M. Willenbacher (✉) · V. Wohlgemuth
Industrial Environmental Informatics Unit, Department Engineering-Technology and Life, Hochschule Für Technik Und Wirtschaft Berlin (HTW), University of Applied Sciences Berlin, Berlin, Germany
e-mail: martina.willenbacher@htw-berlin.de



V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_6

## 1 Introduction

Small and medium-sized enterprises (SMEs) are the most important economic drivers in Germany and also in the countries of the EU. In Germany, there are about 2.5 million SMEs, which account for 42% of the country's total value added and employ 55% of socially insured employees [1]. In 2020, SMEs generated 33.7% of all sales in Germany [2]. However, the degree of digitization of SMEs in Germany and also in the EU is very low. The Institute for SME Research Bonn determined that in 2022, 82% of German SMEs have a low to very low digitization intensity [3] (Figs. 1, 2).

Compared to large companies, SMEs must accept some hurdles in introducing and implementing new technologies and concepts, which leads to SMEs investing significantly less in research and development as well as in practical implementation. Rammer et al. identified the following main barriers to R&D activities in SMEs [5]:

- **Minimum project sizes and minimum costs:** SMEs have to invest a higher proportion of their total resources in R&D projects compared to larger companies, as the technological effort does not allow projects to be scaled down at will. However, these resources are needed in other areas of the company (e.g., innovation activities in process automation or marketing). Furthermore, many SMEs refrain from these activities due to the high minimum costs of R&D projects.
- **High initial and fixed costs in research and development:** To implement R&D projects, there must be a minimum of technical and personnel equipment, which can usually only be used for their project-specific application area. After the end of

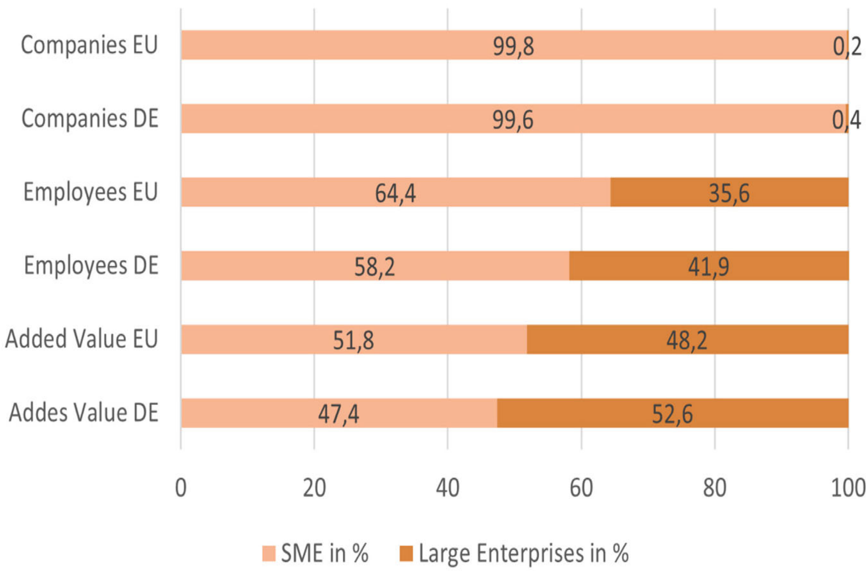



**Fig. 1** Percentage of SMEs, employees in SMEs, and value growth of SMEs in Germany and the EU 2021/2022 [4]

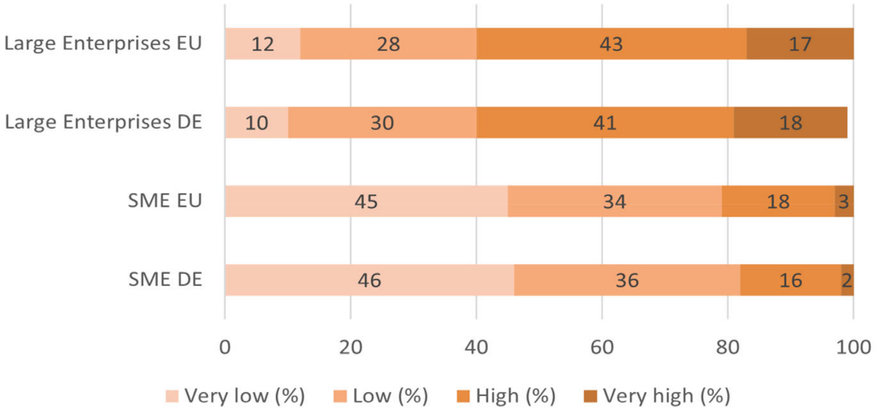



**Fig. 2** Degree of digitization of SMEs and large companies in 2022 [3]

the research activities, these investments are often no longer usable for other areas of the company ("sunk costs"). Furthermore, R&D costs have a high proportion of fixed costs, which can only be economically distributed in the case of several R&D projects.

- **Limited debt financing:** SME-related research is usually very specifically tailored to the respective company and its processes and strategies. On the one hand, this makes it difficult for external parties to assess the chances of success and, on the other hand, transferability to a broad mass for long-term economic exploitation is often not possible. These factors reduce the willingness of external donors to finance SME R&D projects.
- **High-risk exposure:** Due to the high fixed costs and minimum project sizes, SMEs can only pursue a few R&D projects at a time. Therefore, unlike large companies, they can't spread the innovation risk internally. The failure of one of these projects can quickly lead to a threat to the existence of the company for SMEs, as a considerable proportion of resources has been invested in the R & D project and any deficits that may have arisen are difficult to compensate for.
- **Research cooperation:** The use of external knowledge to expand one's know-how is particularly important for SMEs. However, the formation of cooperations requires a high level of investment and at the same time also carries the risk of knowledge leakage. In contrast to large companies (LSEs), R&D projects of SMEs are largely activities that concern the central core processes and are therefore subject to a high level of secrecy to protect the company.

Therefore, it is of great economic importance to support digitalization and the use of intelligent data processing mechanisms in SMEs. In addition to the lack of expertise in the IT sector, many decision-makers in SMEs find it difficult to recognize the optimization opportunities offered by artificial intelligence (AI) for their business sector and to specify the changes and expectations of successfully used AI. Thus, the introduction of these procedures to provide the necessary resources appears to be a high time and cost factor. These factors include first IT personnel with AI skills, acquisition costs for end-to-end digitization of the respective area of application, possible machine failures due to tests to be carried out, training of existing personnel in the use of new technologies, and finally compliance with legal requirements for data protection. It is also not clear to companies at the beginning whether and to what extent optimization potential can be recognized at all and how it can then be implemented. To successfully carry out the digitization and automation process, it is a good idea for SMEs to make gradual changes to the IT and production infrastructure. To digitize the individual production areas and use artificial intelligence methods, decision-makers and employees must be actively involved in these processes and convinced of them. Furthermore, for the technical implementation, there must be knowledge of the work and production processes and their interrelationships. This results in the need to be in close contact with the respective production employees and to understand the individual work processes. On-site inspections and communication across all levels serve to reduce critical key factors:

- **Non-acceptance of the "new" technologies:** Many employees fear the loss of jobs. Automation eliminates the need for certain manual operations. This primarily concerns activities that are easy to perform and could be replaced by machines. Furthermore, the introduction of digitization is associated with new competencies and skills of employees. In addition, there are also reservations regarding data protection and IT security.
- **Incomplete database:** Especially in SMEs, older machines are often still used. In addition, not entire machine parks are replaced during the conversion, so communication between the individual devices is not always given. As a result, you must deal with different communication channels and their combination between the machines and databases. In many cases, there are also media breaks in the higher-level departments (Fig. 3).

- **Insufficient definition of the objective:** The exact objective is essential for the selection of the data, the meaningful calculation by the AI algorithms, and thus the generation of meaningful models. Only with a well-prepared database can an optimal result be achieved, which brings significant added value for the SME.
- **Insufficient knowledge of processes and workflows:** For AI that is optimally tailored to the company, it is essential to have a good knowledge of the workflows that affect the achievement of the goal. Often, certain steps do not seem sensible to the developers, but they have proven themselves over the years for the employees in production.

However, AI applications that relate specifically to the field of corporate environmental protection have so far rarely been found in manufacturing SMEs [7]. Due to the above-mentioned reasons for the difficulties of the introduction of AI and the associated digitalization, manufacturing SMEs are focusing, if at all, on the economic benefits of using these technologies. It therefore makes sense to couple economic

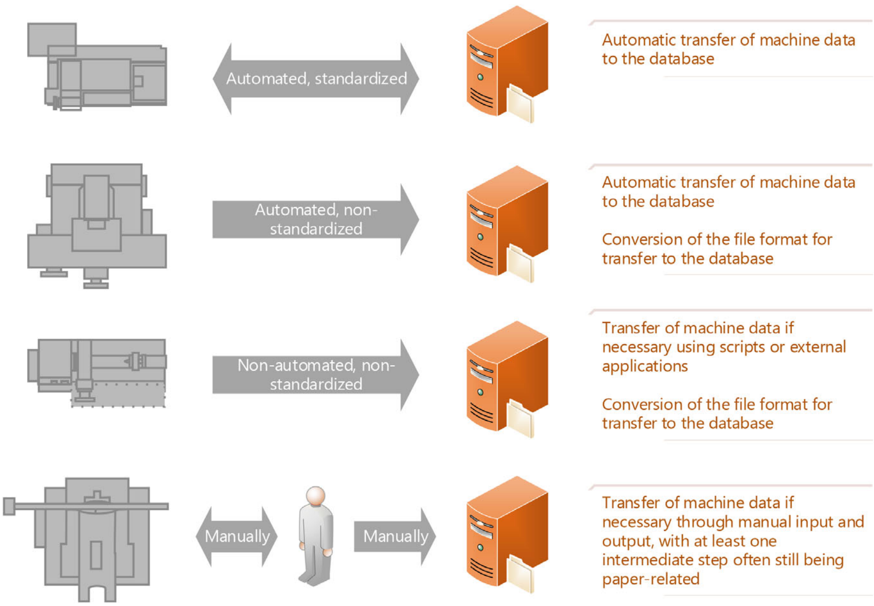



**Fig. 3** Communication channels used between machines and databases in SMEs [6]

with ecological advantages, as was carried out in this doctoral project. For the companies, a cost advantage is achieved through material and energy efficiency, which at the same time serve to reduce environmental pollution by lowering greenhouse gas emissions and conserving resources.

## 2 Materials and Methods

This project aimed to develop AI processes for use in manufacturing SMEs to save energy and waste for the entire manufacturing process. In addition to a reduction in environmental pollution and damage, the increase in material and energy efficiency leads to a reduction in production costs in companies, which in turn leads to an increase in competitiveness. Using an example of an application in a plastics manufacturing company, an adaptable solution for the resource optimization of materials and energy is being developed using the machine learning method Random Forest (RF) [8] and the development of an artificial neural network (ANN) [9]. The following research questions were addressed:

- How can machine learning methods be used to establish relationships between the parameter settings of the machines, the energy consumption and the probability of incorrectly produced plastic parts, and which parameter changes reduce energy consumption and reduce the probability of incorrect production?

Ultimately, the RF model should be able to make statements that machine x can reduce the energy consumption per produced part by z kWh by changing the parameters y1, y2, …, yn.

Based on the RF model, an ANN is then developed to increase the accuracy of these statements. The development of the ANN addresses the following research question:

- How can neural networks be used to find ideal machine settings to reduce energy consumption and improve quality?

For practical use, the ANN is to make statements on the following questions:

- How much energy could be saved if the optimal machine parameter settings proposed by the ANN were used?
- To what extent can the quality be improved considering the reduction of plastic waste?

The statements of these two AI methods should also be interpretable for humans, as many AI models, but especially ANN, exhibit black-box behavior.

In the run-up to the development work, a structured investigation was carried out to determine whether the use of AI for the objective is at all in an appropriate cost–benefit ratio. Therefore, the following process of the introduction of AI was identified and kept as general as possible so that it can be adapted to other problem definitions of SMEs (Fig. 4).

First, the relevant data was identified with their sources. It turned out that the data were distributed heterogeneously or, in some cases, not collected at all. Therefore, the machines were equipped with energy meters, and the OPC UA architecture was implemented for the data exchange of the machines. OPC UA (Open Platform Communications Unified Architecture) is a communication protocol designed for seamless data exchange between industrial devices and systems, regardless of platform or vendor. It enhances interoperability, security, and scalability, enabling real-time data sharing and advanced information modeling in automation environments. This data was then stored centrally in a database. The starting point was 4 machines with 34 features and a total of 777,458 data over a recording period of 4 months. In the first step, the Random Forest model was developed [8], which served as a basis for comparison for the subsequently developed ANN [9]. The development processes are shown in the following figures (Figs. 5, 6).

The evaluation of the RF model results in an energy saving of approx. 58,100 kWh/year and $CO_2$ savings of 27,540 kg/year at a factor of 474 g/kWh [8]. Due to the small amount of data of the M68 machine, which is due to the low production, it is not possible to make accurate statements through the RF model (Table 1).

In contrast to the RF method, the ANN considered both energy and quality simultaneously when determining the recommended machine settings. However, the network's hyperparameters were optimized for energy saving. In the end, only those settings that save energy without reducing quality and have a certain significance are considered. These recommended parameters are shown in **Table 2**.

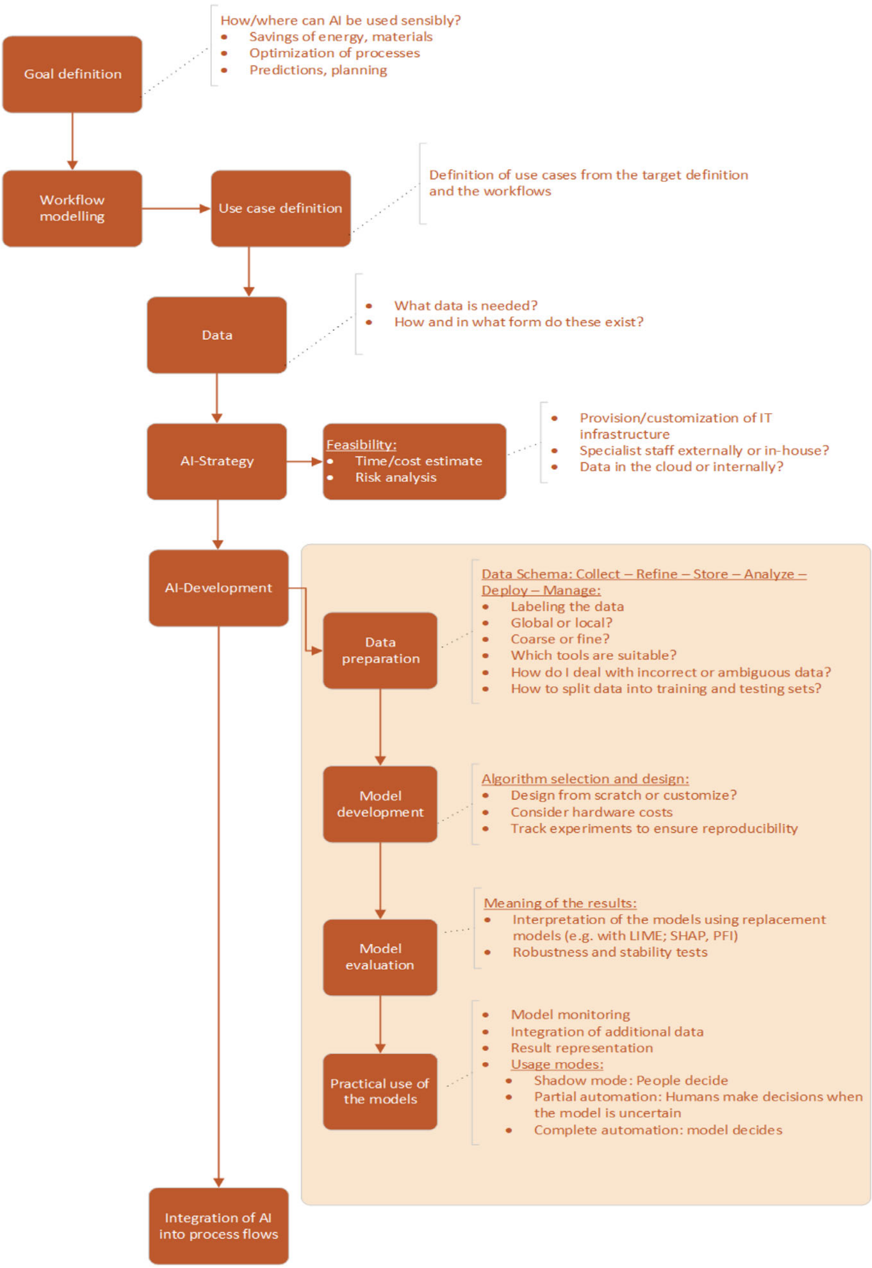



**Fig. 4** Conceptual process of the introduction of AI in SMEs [6]

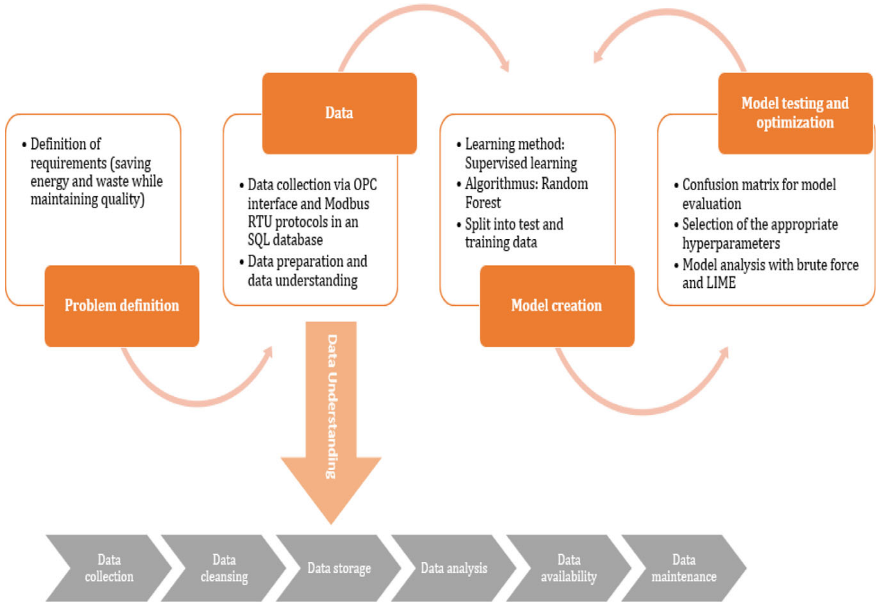



**Fig. 5** Process of the RF [6]

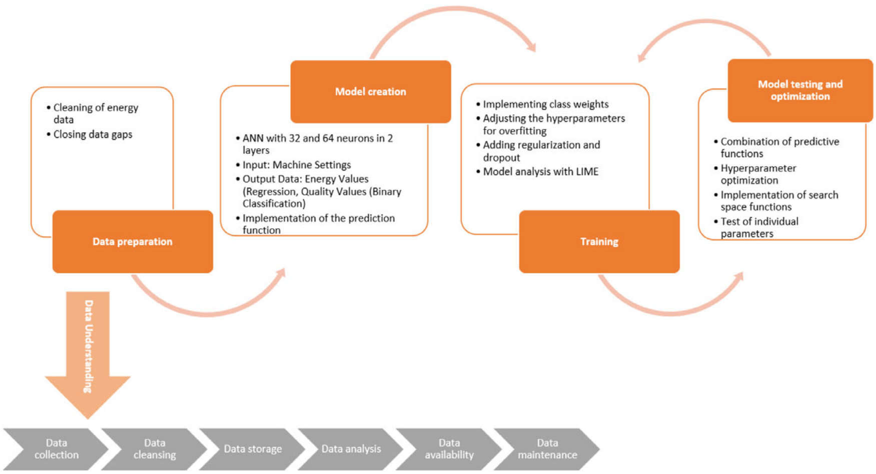



**Fig. 6** Process of the ANN [6]

This results in energy savings of approximately 5,327 kWh/year and $CO_2$ savings of 2,525 kg/year at a factor of 474 g/kWh [10].

**Table 1** Projected energy savings per year (RF)

| | Predicted Energy Savings Per Part [Wh] | Parts produced during the measurement period | Parts produced per year | Energy savings per year [kWh] |
|---|---|---|---|---|
| M64 | 60 | 168,917 | 550,488.438 | 33,029.31 |
| M67 | 20 | 247,052 | 805,124.821 | 16,102.5 |
| M68 | 0 | 86,203 | 280,929.42 | 0 |
| M69 | 10 | 275,286 | 897,137.411 | 8,971.37 |
| **Total** | | | | **58,103.18** |

**Table 2** Projected Energy Savings Per Year (ANN) [9]

| | Recommended parameter | Energy savings per part [Wh] | Parts produced per year (interpolated) | Energy savings per year [kWh] |
|---|---|---|---|---|
| M64 | Cylinder_heating_zone_2_K1 | 6.29 | 550,488.438 | 3,462.57 |
| M67 | Maximum_injection_Pressure_K1, Tool_heating_circuit_1 | 1.67 | 805,124.821 | 1,344.56 |
| M69 | Cylinder_heating_zone_1_K1, Cylinder_heating_zone_4_K1 | 0.58 | 897,137.411 | 520.34 |
| **Total** | | | | **5,327.47** |

## 3 Results

Each method—the Random Forest and the ANN—approaches the problem differently, leading to notable variations in results, which highlight the strengths and weaknesses of both algorithms [8] [9].

The RF algorithm treats energy and quality separately, offering distinct parameter adjustments for each objective. For example, when applied to the M64 machine, RF identified that adjusting the switching injection pressure could result in a 6% improvement in quality and save 15.5 Wh of energy. However, it recommended different pressure values: 685 mbar for quality optimization and 742 mbar for energy efficiency. This means that achieving both improvements simultaneously would require manual adjustments and compromises between the two goals.

In contrast, the ANN algorithm simultaneously optimized both energy and quality, providing unified recommendations for settings. For the M64 machine, it suggested changing the cylinder heating zone temperature to 205.44°C, leading to a 4.87% improvement in quality and a 6.29 Wh energy saving per part. This ability to offer a balanced improvement in both metrics makes ANN a more efficient approach in many cases. The following graphics show the corresponding improvements of both methods using the example of the M64 machine with the respective parameters (Figs. 7, 8).

The same trend is observed for the M67 machine, where RF found that adjusting the switching volume could yield a 12% quality improvement and a 10 Wh energy

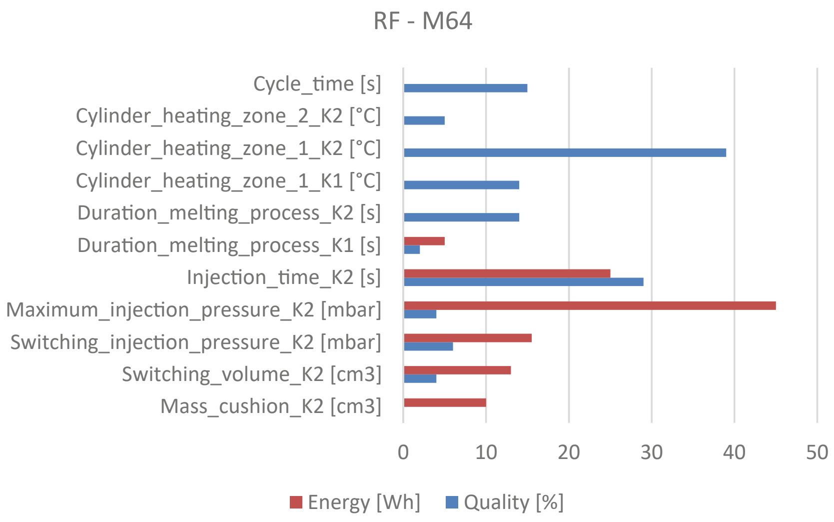



**Fig. 7** Results of the RF algorithm for M64 [6]

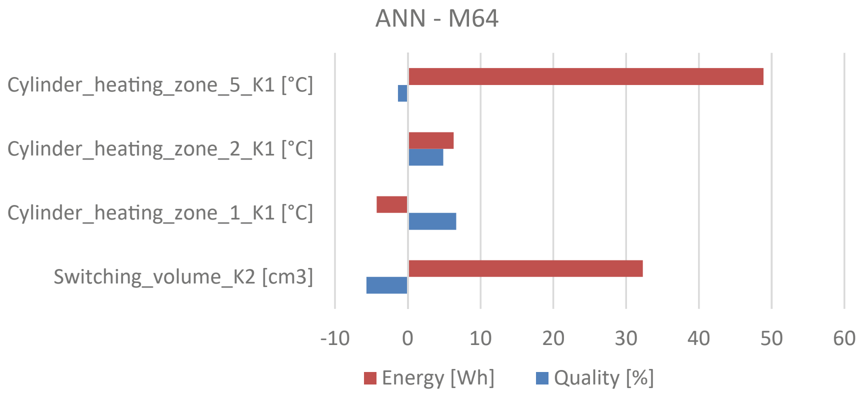



**Fig. 8** Results of the ANN for M64

saving. However, the RF recommendations for quality and energy optimization were not aligned. For example, it proposed that adjusting the mass cushion to 8.4 $cm^3$ would result in a 4% quality improvement, while adjusting it to over 11 $cm^3$ would save 5 Wh of energy. The ANN, however, suggested that the same parameter could yield a 0.13% quality improvement and 7.14 Wh energy savings when set to 17.052 $cm^3$, showing a more balanced optimization approach (Figs. 9, 10).

For the M68 machine, RF identified some energy-saving potential, such as a 2.5 Wh reduction, but also found significant quality improvements—up to 20%—by adjusting certain parameters like the tool heating circuit and cylinder heating zones.

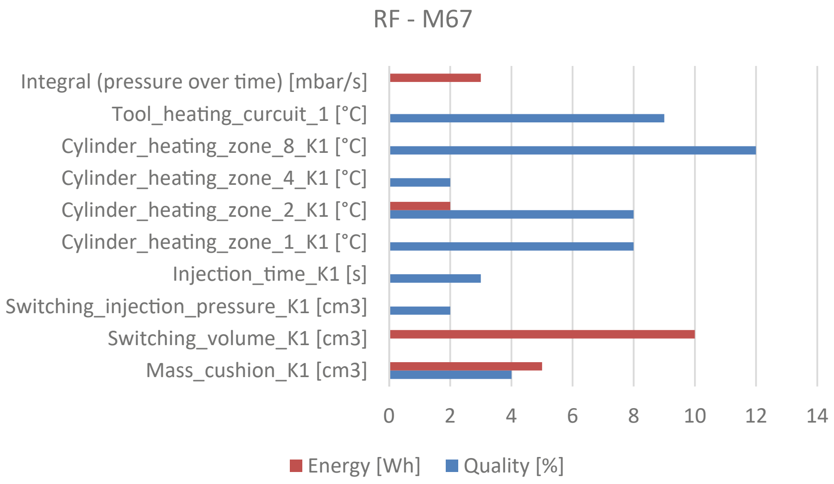



**Fig. 9** Results of the RF algorithm for M67

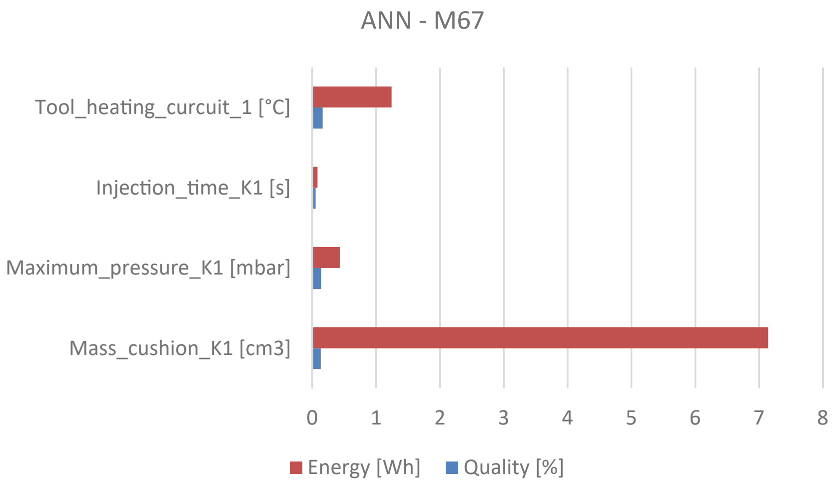



**Fig. 10** Results of the ANN for M67

However, due to limited data (86,203 data sets), the ANN could not be trained for this machine, and thus no conclusions could be drawn.

In the case of M69, RF found no simultaneous improvements for energy and quality, but it suggested significant quality enhancements, such as an 18% increase by adjusting the cylinder heating zone to 243.3°C. ANN, on the other hand, recommended minor changes, such as adjusting the same heating zone to 245.338°C, resulting in a 0.06% quality improvement and 0.11 Wh energy saving (Figs. 11, 12).

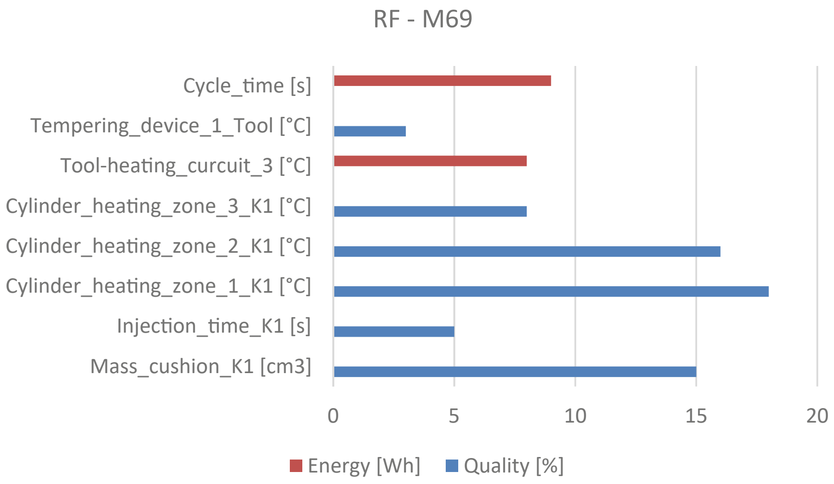



**Fig. 11** Results of the RF algorithm for M69

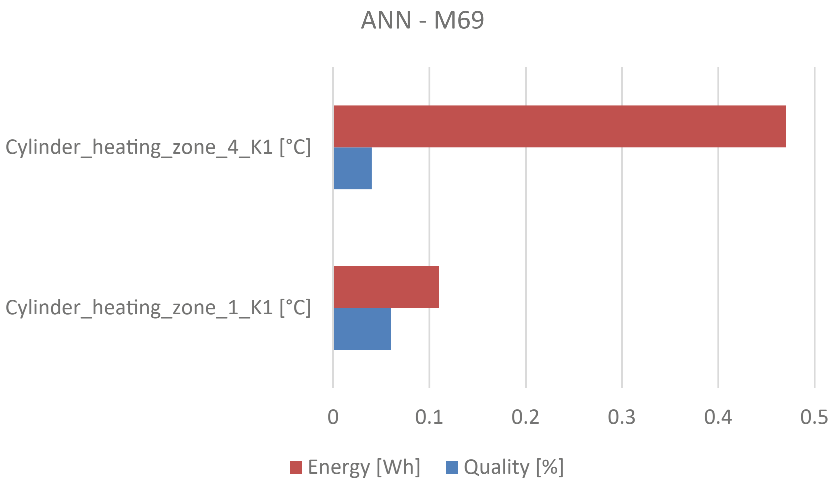



**Fig. 12** Results of the ANN for M69

That leads to the conclusion that while RF can identify valuable parameter adjustments, its limitations lie in treating quality and energy independently. ANN, on the other hand, is more flexible and accurate because it learns from data and can simultaneously optimize for multiple objectives. However, it requires a larger data set and more development time. Despite the challenges, ANN offers a more comprehensive and accurate method for optimizing industrial processes, making it particularly useful for improving energy efficiency and quality in small and medium-sized enterprises (SMEs).

In summary, both methods have their advantages: RF is quicker to implement and useful when data is limited, but ANN provides superior results when sufficient data is available, making it the preferred approach for more complex optimization tasks in industrial settings.

## 4 Conclusion

As part of this research project and through the constant intensive exchange with other medium-sized companies, it became clear again and again that they are generally open to the introduction of AI. However, AI technologies are perceived as very confusing. Existing lighthouse projects cannot be adapted to the individual needs of the respective SMEs without expert knowledge. The subject of further research would be to provide standardized and usable AI modules based on the most common use cases of manufacturing companies. For example, customized company-related software solutions can be offered on the “IT2match” platform [11]. There, companies and software developers are networked according to their needs and offers for the digital workflow. An AI platform for various use cases with corresponding networking options for companies and developers would be a future possibility to introduce intelligent solutions in SMEs.

First and foremost, less complex technical systems should be used when introducing AI. The algorithms presented here can be integrated into the IT of most companies without considerable personnel and financial effort. Furthermore, this makes it easier to adapt to similar problems. In particular, the coupling of an intelligent circular economy with intelligent process optimization can be promising for ecological production from an economic point of view. By linking it to the processing technology (combination of sensors and actuators) on the machines and vehicles, an automated process flow is formed that can uncover further ecological potential using AI processes. There is no question that digitization has a decisive influence on efficient production and serves to optimize and thus save costs for companies because even with the inclusion of a few machines, possible reductions of about 2.500 kg of $CO_2$ per year could be determined. If you keep in mind that the company has 20 times the number of plants, it becomes clear what enormous economic and ecological advantages this technology holds. If these research results and their practical application are extended to other SMEs, a significant contribution can be made to reducing $CO_2$ emissions.

## References

1. Statistisches Bundesamt: Statistik für kleine und mittlere Unternehmen: Code: 48121. https://www-genesis.destatis.de/genesis/online (2023). Accessed 12 Jan 2023

2. Braun, S., Kay, R. Unternehmensgrößenstatistik 2019: Auswirkungen der Berücksichtigung geringfügig Beschäftigter im Unternehmensregister auf die KMU-Kennzahlen. Institut für Mittelstandsforschung (IfM), Bonn (2021)
3. Institut für Mittelstandsforschung Bonn: Mittelstand im Einzelnen. https://www.ifm-bonn.org/statistiken/mittelstand-im-einzelnen (2022). Accessed 12 Jan 2023
4. European Commission, Directorate-General for Internal Market, Industry, Entrepreneurship, SMEs et al.: Annual report on European SMEs 2021/2022 : SMEs and environmental sustainability. Publications Office of the European Union (2022)
5. Rammer, C., Gottschalk, S., Peters, B. et al.: Die Rolle von KMU für Forschung und Innovation in Deutschland: Studie im Auftrag der Expertenkommission Forschung und Innovation. Studien zum deutschen Innovationssystem, Nr. 10–2016. Expertenkommission Forschung und Innovation (EFI), Berlin (2016)
6. Willenbacher, M.: Einsatz von künstlicher Intelligenz zur Prozessoptimierung und Steigerung der Ressourceneffizienz in kleinen und mittleren Unternehmen. Leuphana Universität Lüneburg, Medien- und Informationszentrum (2024)
7. Friedrich, R.: Potenziale der schwachen künstlichen Intelligenz für die betriebliche Ressourceneffizienz. VDI ZRE Publikationen: Studien. VDI Zentrum Ressourceneffizienz GmbH (VDI ZRE), Berlin (2021)
8. Willenbacher, M., Scholten, J., Wohlgemuth, V.: Machine learning for optimization of energy and plastic consumption in the production of thermoplastic parts in SME. Sustainability **13**, 6800 (2021). https://doi.org/10.3390/su13126800
9. Willenbacher, M., Wohlgemuth, V., Risch, L.: Neural networks for energy optimization of production processes in small and medium sized enterprises. In: Wohlgemuth, V., Naumann, S., Behrens, G., et al. (eds.) Advances and New Trends in Environmental Informatics, pp. 129–145. Springer International Publishing, Cham (2023)
10. Icha, P., Kuhs, G.: Entwicklung der spezifischen Kohlendioxid-Emissionen des deutschen Strommix in den Jahren 1990–2018 (2019)
11. Mittelstand 4.0-Kompetenzzentrum—IT Wirtschaft. IT2Match. Version 2.1-LIVE. https://www.it2match.de/. Accessed 14 Mar 2023

# Low-Power and High-Perceptibility Underwater Drone Implementation for Inshore Aquaculture

**Tian Song, Takafumi Katayama, Takashi Shimamoto, and Xiantao Jiang**

**Abstract** In this work, a total solution is proposed for a typical underwater drone with high quality and low power consumption dedicated to inshore aquaculture. The proposed method introduced a high-perceptibility underwater image correction algorithm using a structural similarity evaluation to improve the video quality and object detection accuracy. The proposed algorithm considered real-time performance and low-power consumption by redesigning a compact model. It also employed automatic mixed precision (AMP) to effectively reduce the computational redundancy. An efficient implementation is also introduced using a low-cost GPU board, namely Jetson Xavier NX. The proposed implementation demonstrates improved object detection performance with a processing speed up to 60 fps.

**Keywords** Underwater drone · Inshore aquaculture · Object detection · Low power implementation

## 1 Introduction

Inshore aquaculture plays a critical role in addressing the global food security challenge and mitigating the environmental impacts of land-based agriculture. However, inshore aquaculture which employs open-net cages carries a high potential for

T. Song (✉) · T. Katayama · T. Shimamoto
Graduate School of Technology, Industrial and Social Sciences, Tokushima University, Tokushima City, Japan
e-mail: tiansong@ee.tokushima-u.ac.jp

T. Katayama
e-mail: t.katayama@ee.tokushima-u.ac.jp

T. Shimamoto
e-mail: simamoto@ee.tokushima-u.ac.jp

X. Jiang
Department of Information Engineering, Shanghai Maritime University, Shanghai, China
e-mail: xtjiang@shmtu.edu.cn



V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_7

environmental degradation. To avoid these potential risks, land-based aquaculture is promoted with very high productive costs in Japan. On the other hand, open-water aquaculture is an environmentally friendly alternative to traditional aquaculture methods, but it comes with a significant cost. Furthermore, the high cost of open-water aquaculture highly depends on state-of-the-art technology, particularly in the development of low-cost, high-performance drones for environmental monitoring and management.

In recent years, small underwater remotely operated vehicles (ROVs) have emerged as powerful tools in a wide range of applications in marine industries, scientific research, and entertainment [1–3]. They allow for cost-effective and efficient inspection, monitoring, and exploration of underwater environments without the need for human divers. To address the challenges of traditional aquaculture, developments of inshore aquaculture utilizing underwater drones become increasingly important. In a camera-equipped underwater drone, object detection plays a crucial role in autonomous driving in various environments, such as avoiding obstacles, locating, tracking targets, and navigating in complex underwater environments. Some previous works show good performance using updated learning-based models to enhance object detection performance. Moreover, real-time processing is important for chasing moving objects in underwater situations. Considering the battery-driven underwater operating a low power consumption implementation is also required to extend the drone's operating time. All these factors are especially important for low-cost underwater drones, which are often utilized for tasks that require long-term monitoring.

Many efficient models have been proposed for object detection, such as deformable part models (DPM) and R-CNN [4, 5]. Emerging in recent years, the YOLO (You Only Look Once) object detection model has rapidly become an acceptable choice due to its advantages, leading to widespread use [6]. The YOLO algorithm employs a grid-based partitioning of the entire image, where each grid cell is utilized to estimate the object class and bounding box. YOLO simplifies the CNN architecture, resulting in faster processing. Unlike approaches that rely on sliding windows and region proposals, YOLO employs the entire image for training, allowing it to learn contextual information simultaneously. This helps to reduce background false positives and improves object discrimination. Considering the high requirement for highly accurate real-time object detection in underwater drones, YOLO is used as the basic model in our object detection tasks.

However, underwater videos can suffer from various factors such as light attenuation, water turbidity, backscatter, and camera noise, which can significantly degrade the image quality and affect the accuracy of object detection algorithms. These factors can also cause color distortion and reduce the contrast and sharpness of the images, making it more difficult for the algorithms to distinguish between objects and the background. Therefore, the color correction algorithm is highly required for an underwater image. Additionally, computational efficiency is essential for object detection, as both image correction and object detection algorithms must operate within the constraints of computing power and limited memory on an embedded GPU. Therefore, achieving real-time processing capabilities becomes crucial.

Our key contributions in this work can be concluded as follows:

1. A total solution is proposed for a low-cost underwater drone dedicated to inshore aquaculture.
2. An efficient color correction algorithm is proposed to enhance the video quality. Consequently, the object detection accuracy is improved. The proposed algorithm accurately performs color correction, even for severely deteriorated images.
3. The reduction in the number of channels and automatic mixed precision (AMP) have notably elevated the computational speed through the mitigation of redundancy.
4. Finally, we have introduced an updated model capable of providing accurate real-time correction for severely degraded images.

## 2 Related Work

Several related works are introduced in this chapter including a GAN-based color correction algorithm named FUnIEGAN, an object detection model named YOLO, and a SLAM system named ORB-SLAM.

### *2.1 FUnIEGAN*

FUnIEGAN is a conditional GAN (CGAN) that learns the correspondence between an image and a condition vector as input to correct the color drifts [7]. FUnIEGAN learns the correspondence between an input image and a condition image. Moreover, to facilitate real-time processing, FUnIEGAN is redesigned as a simple model with significantly fewer parameters compared to typical similar models. This simplicity is critical for applications requiring efficient and fast image generation without sacrificing the quality of the generated image. Figure 1 shows the overall structure of FUnIEGAN.

The generator of FUnIEGAN follows a UNet architecture and does not include fully connected layers [8]. UNet is an image segmentation network that achieves both feature extraction and spatial information retention through convolution. By

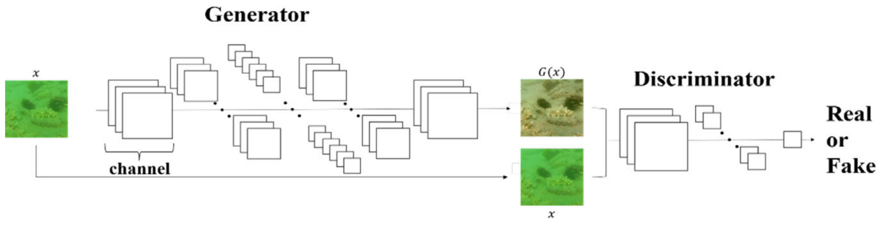



**Fig. 1** FUnIEGAN's overall structure

replacing the fully connected layers of a CNN with convolutional layers, the output shifts from identifying what an object is to determining where the object is located. FUnIEGAN is introduced in this work for its fast processing and high correction performance.

### 2.2 Object Detection

For image detection in this work, we employ YOLO, a CNN capable of fast and accurate object detection [6]. YOLO can detect over 9,000 different objects and operates much faster than competitive methods like SSD and Faster R-CNN while achieving equal or better accuracy. Object detection algorithms typically consist of three main stages: Region Proposal, which detects potential object regions; Classification, which labels these detected regions; and Bounding Box Regression, which estimates the precise coordinates of the objects. In most previous works, regions are proposed using various methods before performing classification. However, YOLO takes a different approach by dividing the entire image into a grid and predicting classes and bounding boxes for each region. This simplifies the CNN architecture, allowing for increased speed. Unlike methods that use sliding windows or region proposals, YOLO uses the entire image during training, enabling the network to learn contextual information surrounding objects. This reduces false detections of background elements. Despite its advantages, YOLO has some limitations. The grid size is fixed, and each grid cell can detect only one class and a maximum of two objects. This restriction makes it challenging to detect numerous objects within a single image.

In this work, we use YOLOX, an improved model of YOLO, as the basic model [9]. The baseline network is YOLOv3-SPP, which uses DarkNet53 as its backbone, connected to an SPP (Spatial Pyramid Pooling) layer [10]. YOLOX incorporates additional features such as EMA (Exponential Moving Average) weight updating, a cosine learning rate schedule, IoU loss, and an IoU-aware branch to form its baseline. One of the significant features of YOLOX is its anchor-free approach. Unlike traditional anchor-based methods, anchor-free means that each feature map cell predicts a single bounding box, similar to the mechanism in YOLOv3 but with far superior accuracy and speed. As a result, YOLOX achieves higher precision than the baseline method, demonstrating enhanced performance in object detection.

### 2.3 ORB-SLAM

Simultaneous Localization and Mapping (SLAM) involves accurately determining one's position and orientation while simultaneously gathering various information to understand the surrounding environment. Generally, "Localization" is often expressed as self-position estimation, and "Mapping" as environment map creation."

Hence, SLAM is a collective term for the technology that performs self-position estimation and environment map creation simultaneously.

Visual SLAM refers to SLAM technology that utilizes image information. Cameras used in Visual SLAM include monocular cameras with a single lens and stereo cameras with multiple lenses. The cost advantage of using relatively inexpensive cameras, combined with advancements in technology that have improved accuracy, has led to its rapid adoption. Moreover, Visual SLAM can be combined with advanced technologies such as video analysis and machine learning to achieve more accurate self-position estimation and environment mapping.

A representative Visual-SLAM system is ORB-SLAM [11]. ORB-SLAM is a real-time, ORB feature-based monocular Visual-SLAM that alternates between self-position estimation and environment mapping. The algorithm begins by initializing a 3D point map from two video frames. This is achieved using triangulation based on correspondences of 2D ORB features, creating the initial camera pose and 3D points, a process known as Map Initialization.

After the map is initialized, for each new frame, the current frame's features are matched with the features of the last keyframe to estimate the camera pose. The estimated camera pose is then refined through local map tracking, a process called Tracking. When the current frame is identified as a keyframe, it is used to create a new 3D map. At this stage, bundle adjustment is employed to minimize reprojection error by optimizing the camera poses and 3D points. This step is known as Mapping. Finally, ORB-SLAM uses a bag-of-features approach to detect loops by comparing each keyframe with all previous keyframes. When a loop closure is detected, the pose graph is optimized, adjusting the camera poses of all keyframes. This final step is referred to as Loop Closing.

## 3 The Proposed Method

In this work, we propose new approaches for FUnIEGAN, which is capable of real-time inference, to solve challenges in aquaculture industry applications.. These approaches include creating new datasets, improving the network, and modifying computation methods. An overview of the proposed method is shown in Fig. 2.

In this work, we implement five key improvements to FUnIEGAN, as illustrated in Fig. 2. Firstly, to accurately correct images with severe underwater turbidity, we create a new dataset specifically tailored for such conditions. This dataset help train the model to handle the unique challenges posed by underwater environments, including varying levels of turbidity and lighting conditions. Secondly, to enhance correction accuracy, we introduce the Structural Similarity Index (SSIM) loss function. The SSIM loss function is a perceptual metric that evaluates image quality by comparing structural information, making it more effective in preserving important features and textures in the corrected images compared to traditional loss functions like Mean Squared Error (MSE). Thirdly, to improve processing speed, we reduce the

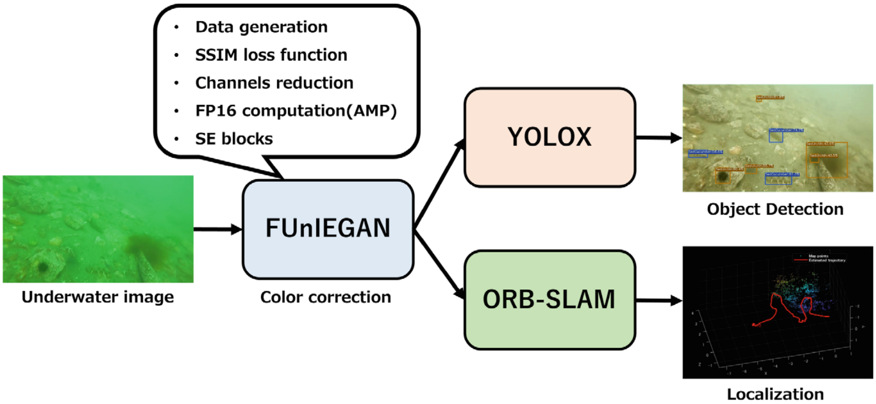



**Fig. 2** Overview of the proposed method

number of channels in the generator and perform computations using 16-bit floating-point arithmetic. This optimization balances computational efficiency and precision, which is crucial for real-time applications where speed is essential. Fourthly, redundant channels in the convolutional layers are reduced to achieve fast processing and further low-power consumption. Finally, for even higher precision in image correction, we incorporate Squeeze-and-Excitation (SE) Blocks into the network. SE Blocks enhance the network's representational power by adaptively recalibrating channel-wise feature responses, leading to better-quality image generation.

The new FUnIEGAN model is applied to underwater image correction, and the output images are evaluated. After assessing the quality of the corrected images, the effectiveness of underwater image correction in object detection and SLAM is verified.

## *3.1 Data Generation*

Preparing a training dataset consisting of paired underwater and ground truth images is extremely challenging. Therefore, in this work, Data Generation (DG) is conducted by creating new ideal training pairs, where images generated by Water-Net are treated as ground truth images [13].

The original underwater images are sourced from the Underwater Robot Picking Contest dataset, namely URPC2018, provided by the National Natural Science Foundation of China and the Dalian Municipal Government. This dataset is designed for object detection of nearshore aquaculture organisms using underwater drones [14]. URPC2018 consists of 7,600 images, which include actual footage from the underwater robot picking contest and various underwater scenes captured by underwater drones.

Figure 3 shows some examples of images included in the URPC2018 dataset.

**Fig. 3** Example of URPC2018 dataset

Since URPC2018 images are captured by underwater drones, many images contain motion blur due to drone movement or are too blurry to recognize the objects. Therefore, images with significant blur or haze that obscure object detection are removed from the URPC2018 dataset. Additionally, some images in the URPC2018 dataset were already corrected. Therefore, it was necessary to distinguish between underwater images and corrected images. Specifically, the RGB values of each pixel in each image were obtained, and then the average RGB values for the entire image were calculated. By focusing on the average G and R values, any image with a subtraction result of "average G value" and "average R value" that smaller than 33 is considered a corrected image and removed from the dataset. As a result, a refined URPC2018 dataset consisting of 2,359 images is generated.

Next, the 2,359 images from the URPC2018 dataset are corrected using Water-Net, creating an ideal training dataset composed of paired underwater and ground truth images. FUnIEGAN was then trained on this dataset and the inference results are shown in Fig. 4.

As shown in Fig. 4, images that could not be correctly corrected by the existing FUnIEGAN were successfully corrected using Data Generation. Hereafter, the FUnIEGAN model enhanced by Data Generation will be referred to as the DG model.

### 3.2 SSIM Loss Function

The Structural Similarity (SSIM) index is a measure used to calculate the structural similarity between two images and is widely used in image correction and

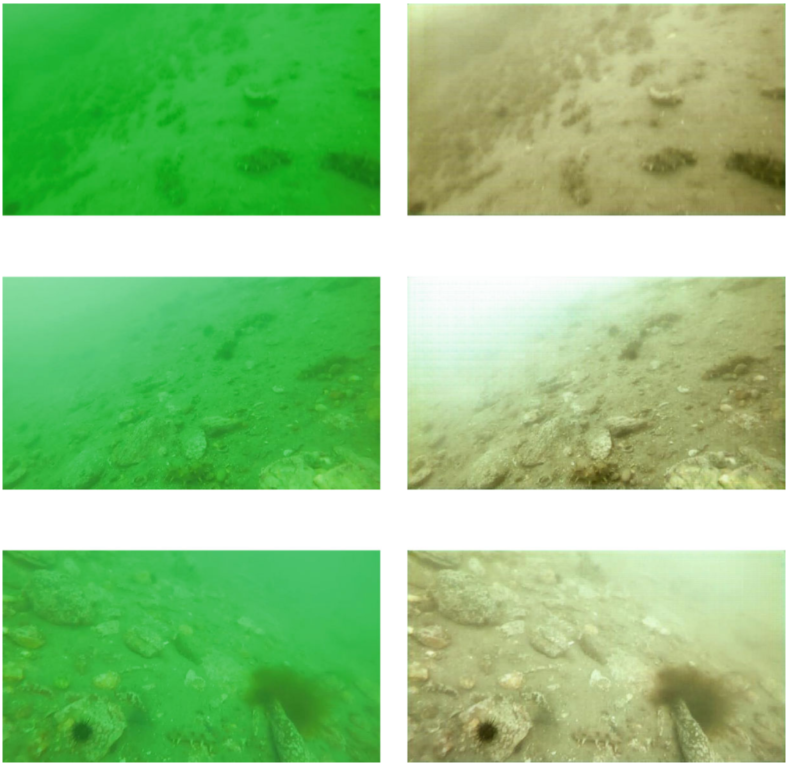


**Fig. 4** Correction results by the DG model; Left : Raw data Right:FUnIEGAN with DG model

restoration tasks to evaluate image quality. However, in the existing FUnIEGAN model, SSIM is not included as a loss function during the training process, potentially neglecting structural similarity during image generation. To address this issue, this work proposes a new training method that incorporates the SSIM loss function into the learning process of the FUnIEGAN model.

The SSIM calculation formula is shown in Eq. 1. Since SSIM calculations are performed in local regions, it has a high correlation with subjective image quality.

$$SSIM\,(p) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{1}$$

Here, $p$ denotes a pixel, $\mu_x$ and $\mu_y$ are the means, $\sigma_x$ and $\sigma_y$ are the standard deviations, $\sigma_{xy}$ is the covariance, and $C_1$ and $C_2$ are constants set to avoid instability when the denominator is very small. In this work, $C_1$ is set to 0.01 and $C_2$ is set to 0.03.

By introducing the SSIM loss function, the objective evaluation value of the generated images can be improved, emphasizing structural similarity. As a result, the generator can produce images that are more structurally similar to the original images. Hereafter, the FUnIEGAN model trained with the SSIM loss function will be referred to as the SSIM model.

### 3.3 Channels Reduction

In FUnIEGAN, image generation is performed by the Generator during inference. Reducing the number of parameters in the Generator can speed up computation, making real-time inference more feasible.

FUnIEGAN is a fully convolutional model, and the number of parameters in the convolutional layers is determined by the filter size and the number of input and output channels. The number of parameters in the Generator heavily depends on the number of channels. Therefore, this work proposes to reduce the number of channels in the convolutional layers of the Generator by half to decrease the number of parameters.

Figure 5 shows the channel representation of the existing FUnIEGAN Generator, while Fig. 6 shows the channel representation of the Channel Reduction (CR) model. Here, the channel reduction is referred to as CR (Channel Reduction), and the model with reduced channels is referred to as the CR model. The conventional FUnIEGAN has 7,019,587 parameters, while the CR model has 1,756,707 parameters. The number of channels in the Generator of the CR model is reduced by approximately 75% compared to the existing model.

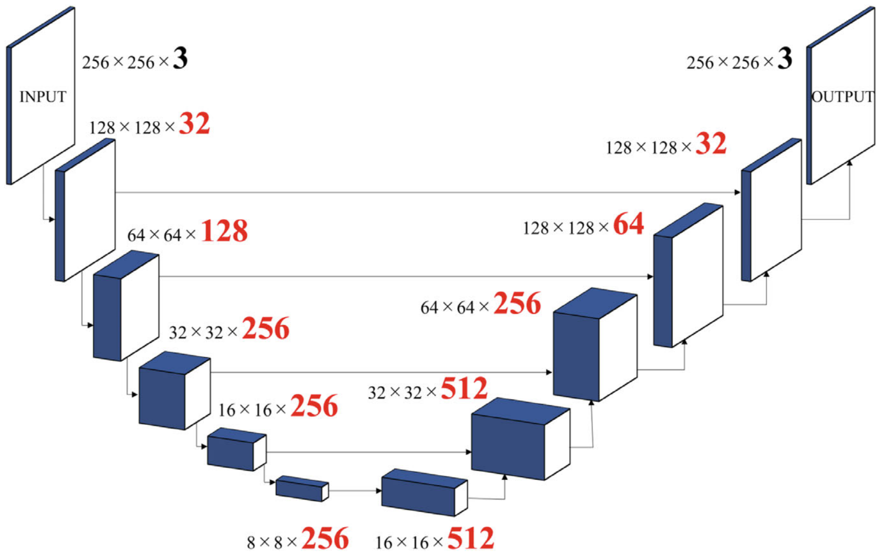



**Fig. 5** Number of channels of FUnIEGAN Generator

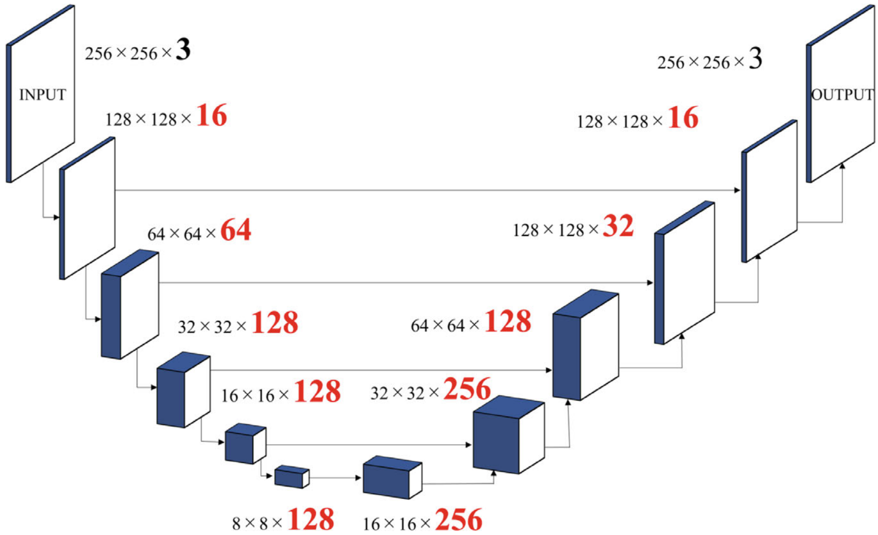



**Fig. 6** Number of channels of FUnIEGAN Generator with CR model

### *3.4 Automatic Mixed Precision (AMP)*

To further enhance real-time inference, we propose using Automatic Mixed Precision (AMP) to perform calculations in 16-bit floating-point (FP16) format, similar to the channel reduction approach.

Traditionally, neural network models have performed calculations using 32-bit floating-point (FP32) format. By using FP16, which has half the bit width of FP32, the required memory size is halved, and bandwidth savings during weight transfer lead to faster computations. However, performing all calculations in FP16 can lead to underflow during training, which can decrease accuracy. Underflow occurs when the exponent part of the floating-point operation becomes too small to be represented, resulting in the value being rounded to zero.

AMP addresses this issue by dynamically determining which parts of the computation should be performed in FP16 and which in FP32. It classifies operations into three categories: those that must be computed in FP32 to avoid underflow, those that can safely be computed in FP16, and those that should be computed in FP16 for performance benefits without loss of precision. By selectively applying FP16 and FP32, AMP ensures that training proceeds efficiently without sacrificing accuracy.

Similarly, performing inference in FP16 speeds up computations, enabling real-time inference with lower power consumption. Hereafter, the FUnIEGAN model that uses FP16 for inference will be referred to as the FP16 model.

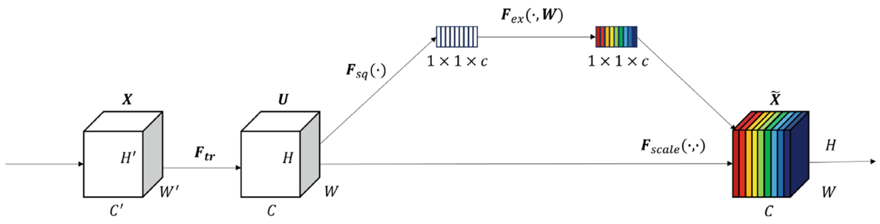



**Fig. 7** The structure of SE block

## *3.5 SE Blocks*

The Squeeze-and-Excitation (SE) Block functions by automatically recalibrating the channel-wise feature responses of convolutional layers. Unlike traditional convolutional layers, SE Block emphasizes important features, enabling the model to focus on relevant information and improve image correction.

In this work, SE Block is incorporated into the architecture of the Generator to mitigate the decrease in objective evaluation metrics caused by the reduction in the number of channels. SE Blocks efficiently distinguish and prioritize important channels during the image generation process. This functionality enhances the accuracy of the generated images while maintaining real-time processing speed.

Hereafter, the FUnIEGAN model trained with the SE Block will be referred to as the SE Block model (Fig. 7).

# 4 Simulation and Implementation Results

## *4.1 Simulation Environment*

FUnIEGAN was trained on the URPC2018 dataset with a batch size of 8, using the Adam optimization method and a learning rate of 0.0003. The training process lasted for 200 epochs. Additionally, the SSIM model was developed by adding the SSIM loss function to FUnIEGAN, with a hyperparameter set to 50.

The CR model modifies FUnIEGAN by halving the number of generator channels in all layers. Due to the lack of convergence after 200 epochs, the training was extended to 300 epochs. This adjustment reduced the number of generator parameters from 7,019,587 to 1,756,707.

The FP16 model utilizes FP16 operations for computations. Similar to the CR model, the training for the FP16 model was extended to 300 epochs due to non-convergence after 200 epochs.

The SE Blocks model incorporates SE Blocks processing layers into FUnIEGAN. As with the other models, the training was extended to 300 epochs because it did

**Table 1** The detail of configuration in YOLOX

| | | | | |
|---|---|---|---|---|
| Input image size | 720 × 405 | | Batch size | 32 |
| Input format | RGB | | Optimizer | Momentum |
| Epoch | 300 | | Trainning data | URPC2018 |
| Learning rate | $\frac{0.1}{64}$ | | Test data | URPC2018 |
| Back borne arch. | YOLOx_s | | | |

not converge after 200 epochs. When SE Blocks were added to the CR model, the number of generator parameters increased from 1,756,707 to 1,781,827.

#### 4.1.1 The Hyperparameter Configuration in YOLOX

To perform detection, validation was conducted using YOLOX. The various parameters used during validation are shown in Table 1.

The dataset used for YOLOX consisted of the optimized URPC2018 dataset, which contains 2,359 images for training and test.

### *4.2 Comparison of Color Correction Methods*

The results of the correction performed on the optimized URPC2018 dataset, consisting of 2,359 images, are shown in Fig. 8 and Table 2. The processing speed is measured on a Jetson Orin Nano, a small GPU module.

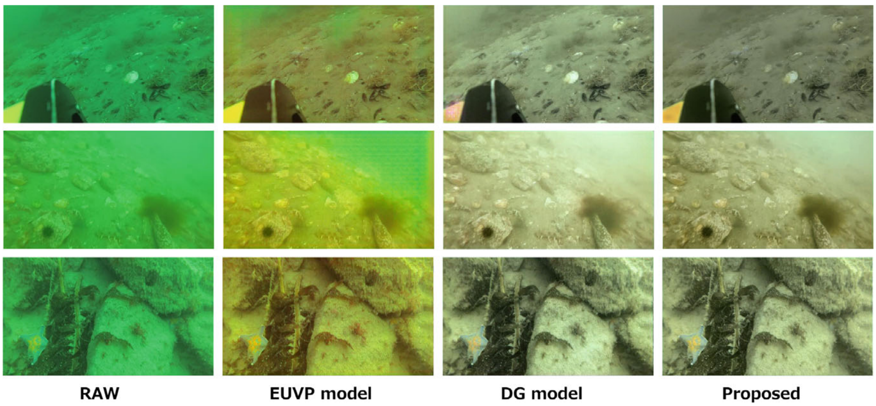



**Fig. 8** Comparision of color correction performance

**Table 2** Quantitative evaluation in each model

| DG | SSIM | CR | FP16 | SE block | MSE | PSNR[dB] | SSIM | FPS | epoch |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | 1350.545 | 17.663 | 0.811 | 66.788 | 200 |
| ✓ | | | | | 379.300 | 23.970 | 0.927 | 66.551 | 200 |
| ✓ | ✓ | | | | 225.480 | 26.172 | 0.950 | 66.944 | 200 |
| ✓ | | ✓ | | | 312.821 | 24.482 | 0.923 | 78.909 | 300 |
| ✓ | | | ✓ | | 288.885 | 25.522 | 0.938 | 82.953 | 300 |
| ✓ | | | | ✓ | 159.605 | 27.714 | 0.955 | 45.491 | 300 |
| ✓ | ✓ | ✓ | | | 238.553 | 25.152 | 0.942 | 82.260 | 300 |
| ✓ | ✓ | | ✓ | | 239.267 | 25.722 | 0.945 | 81.616 | 300 |
| ✓ | ✓ | | | ✓ | 114.372 | 29.307 | 0.963 | 46.634 | 300 |
| ✓ | ✓ | ✓ | ✓ | | 328.637 | 24.721 | 0.933 | 98.861 | 300 |
| ✓ | ✓ | ✓ | ✓ | ✓ | 145.292 | 28.026 | 0.954 | 62.153 | 300 |

The DG model was able to correctly perform color correction even on images with severe underwater turbidity. Similarly, the proposed method also achieved accurate color correction on images with significant turbidity.

As shown in Table 2, the EUVP model performed the worst in terms of MSE, PSNR, and SSIM, which indicate the extent of visual degradation. The DG model showed significant improvement in these evaluation metrics compared to the EUVP model, indicating better visual quality with less degradation. The SSIM model further improved all evaluation metrics, achieving a substantial increase in SSIM values as intended.

Although the CR model and FP16 model showed decreased evaluation metrics compared to the SSIM model, they each demonstrated processing speed improvements of approximately 12.4 FPS and 16.00 FPS, respectively, over the DG model. The SE Block model achieved the best results among the individual models for all evaluation metrics. However, it showed a decrease in processing speed by 21.06 FPS compared to the DG model, despite only a 1.4% increase in the number of parameters. This is likely due to the computational cost and increased memory usage resulting from the introduction of SE Blocks in all layers of the Generator.

The combined DG & SSIM & CR & FP16 model achieved a processing speed of 98.861 FPS, representing an improvement of 32.31 FPS over the DG model. The final proposed method (DG & SSIM & CR & FP16 & SE Block) maintained a processing speed of over 60 FPS, achieved the second-best results in MSE and PSNR, and the third-best result in SSIM. These results demonstrate that the proposed method is a high-accuracy correction model capable of real-time processing.

### 4.3 Object Detection Results

The results of training and inference using images corrected by Water-Net from the 2,124-image URPC2018 dataset are shown in Fig. 9 and Table 3.

As shown in Fig. 9, the original images allow for some detection, but only for objects in the foreground or those that are large. Objects in the background are not detected. In the EUVP model, the failure in correction adversely affects the detection results. The DG model improves detection, allowing for the identification of objects that are not detected in the original images. The proposed method further enhances the detection performance, identifying objects in the background that even Water-Net failed to detect.

From Table 3, the mean Average Precision (mAP) for the original images is 80.55%. However, the Average Precision (AP) for sea cucumbers and scallops is relatively low, at 68.20% and 79.93%, respectively. The EUVP model, due to failed correction, shows the worst detection performance with an mAP of 73.89%. The DG model, having correctly performed correction, improves the mAP to 82.76%, with an 8.55% increase in the AP for scallops.

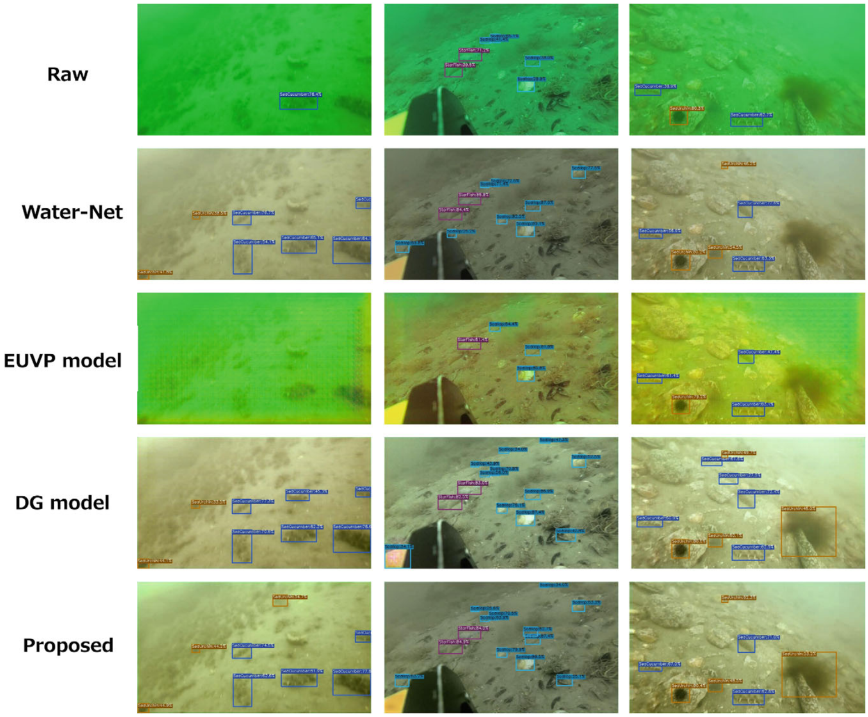



**Fig. 9** Qualitative evaluation of detection accuracy for each model

**Table 3** Quantitative evaluation of object detection for each model

| DG | SSIM | CR | FP16 | SEBlock | Average precision for each class (%) | | | | mAP(%) |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | Sea urchin | Sea cucumber | Scallop | Starfish | |
| | | | | | 85.24 | 61.10 | 80.33 | 68.91 | 73.89 |
| ✓ | | | | | 89.97 | 69.57 | 88.48 | 83.02 | 82.76 |
| ✓ | ✓ | | | | 89.99 | 73.23 | 90.47 | 86.03 | 84.93 |
| ✓ | | ✓ | | | 89.35 | 72.19 | 90.58 | 83.13 | 83.81 |
| ✓ | | | ✓ | | 89.64 | 73.32 | 89.86 | 85.70 | 84.63 |
| ✓ | | | | ✓ | 90.21 | 71.64 | 90.52 | 84.96 | 84.33 |
| ✓ | ✓ | ✓ | | | 89.70 | 70.89 | 89.87 | 84.62 | 83.77 |
| ✓ | ✓ | | ✓ | | 90.23 | 71.85 | 88.46 | 85.74 | 84.07 |
| ✓ | ✓ | | | ✓ | 90.10 | 73.32 | 91.94 | 86.21 | 85.39 |
| ✓ | ✓ | ✓ | ✓ | | 89.97 | 71.10 | 88.27 | 84.15 | 83.37 |
| ✓ | ✓ | ✓ | ✓ | ✓ | 91.26 | 73.51 | 91.47 | 83.24 | 84.87 |
| Water-net | | | | | 90.84 | 74.59 | 92.67 | 86.15 | 86.06 |
| Raw | | | | | 90.03 | 68.20 | 79.93 | 84.05 | 80.55 |

The DG & SSIM & SE Block model achieves the highest mAP of 85.39% among all models, also attaining the highest AP for scallops and starfish. The final proposed method achieves an mAP of 84.87%, the third highest among all models. It also records the highest AP for sea urchins and sea cucumbers, making it highly suitable for applications in the aquaculture industry.

These results demonstrate the effectiveness of underwater image correction in object detection and the efficacy of the proposed method.

Additionally, the processing time for YOLOX, when run on the Jetson Orin Nano, achieves 15.246 FPS. When FUnIEGAN and YOLOX are combined in a pipeline, a processing speed of 9.469 FPS is achieved.

### *4.4 ORB-SLAM Results*

To demonstrate the effectiveness of underwater image correction, evaluations were conducted using not only object detection but also SLAM (Simultaneous Localization and Mapping). Figure 10 shows the feature point extraction process in SLAM.

Figure 10 highlights a portion of the feature point extraction. As shown in the figure, the original images suffer from severe underwater turbidity, resulting in very few feature points being extracted. The extracted few points are all located in the foreground without feature points detected in the background.

In contrast, the images corrected using the proposed method show significant improvement. The turbidity is effectively removed, allowing for the extraction of a

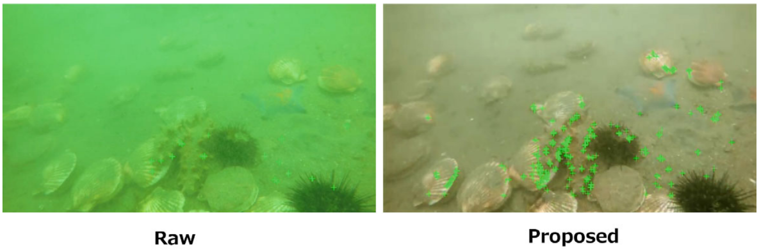



**Fig. 10** Feature point extraction process in SLAM

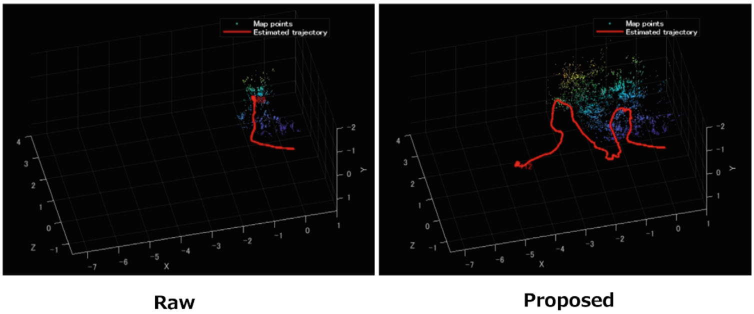



**Fig. 11** Mapping results of ORB-SLAM

large number of feature points. Unlike the original images, feature points are extracted not only in the foreground but also in the background.

Next, Fig. 11 shows the mapping process using SLAM.

Figure 11 illustrates the mapping results for each method. From the figure, it is evident that mapping stops due to the shortage of feature points with the original images. As shown in Fig. 10, the failure to extract sufficient feature points led to the inability to continue mapping.

In contrast, the images corrected with the proposed method allowed for the extraction of a substantial number of feature points, enabling the mapping process to continue without interruption until completion. These results clearly demonstrate the effectiveness of underwater image correction for SLAM applications.

## 5 Conclusion

In this paper, we proposed an underwater image correction method optimized for real-time inference on small underwater drones. While the existing Water-Net method offers high-accuracy correction, it has the disadvantage of a substantial computational load. FUnIEGAN allows for real-time processing but fails to perform accurate color correction in aquaculture environments. To address these issues, we created a dataset suitable for the aquaculture industry and made improvements to the network. As a result, the proposed method achieved high-performance correction in aquaculture environments while maintaining real-time capabilities. Additionally, the proposed method demonstrated high-accuracy object detection, particularly effective for detecting sea urchins and sea cucumbers, highlighting the effectiveness of underwater image correction. The effectiveness of the proposed method was also validated in SLAM applications. Future work will focus on applying the results of object detection and SLAM to autonomous control.

## References

1. Yang, H.-H., Huang, K.-C., Chen, W.-T.: LAFFNet: a lightweight adaptive feature fusion network for underwater image enhancement. In: ICRA2021 (2021). arXiv:2105.01299v2
2. Bingham, B., Foley, B., Singh, H., Camilli, R., Delaporta, K., Eustice, R., Mallios, A., Mindell, D., Roman, C., Sakellariou, D.: Robotic tools for deep water archaeology: surveying an ancient shipwreck with an autonomous underwater vehicle. J. Field Robot. (2010)
3. Qin, X., Zhang, Z., Huang, C., Gao, C., Dehghan, M. and Jagersand, M.: BASNet: boundary-aware salient object detection. In: CVPR2019 (2019)
4. Crandall, D., Felzenszwalb, P., Huttenlocher, D.: Spatial priors for part-based detection using statistical models. In: IEEE (2005)
5. Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation Tech report (v5)" CVPR2014 (2014). arXiv:1311.2524v5
6. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: IEEE (2016). arXiv:1506.02640v5
7. Ialam, J., Xia, Y., Sattar, J.: Fast underwater image enhancement for improved visual perception. IEEE Robot. Autom. Lett. 3227–3234 (2020)
8. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. Comput. Vis. Pattern Recognit. (cs.CV) (2015)
9. Ge, Z., Liu, S., Wang, F., Li, Z., Sun, J.: Megvii technology "YOLOX: exceeding YOLO series in 2021,". In: 2021 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2021)
10. Redmon, J., Farhadi, A.: University of Washington "YOLOv3: an incremental improvement. In: 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2018)
11. Mur-Artal, R., Montiel, J.M.M. and Tardos, J.D.: Member. "ORB-SLAM: a versatile and accurate monocular SLAM system, Robotics (cs.RO). In: Computer Vision and Pattern Recognition (cs.CV) (2015)
12. Imada, A., Katayama, T., Song, T., Shimamoto, T.: YOLOX based underwater object detection for inshore aquaculture. In: OCEANS 2022, Hampton Roads, Hampton Roads, pp. 1–5 (2022)
13. Li, C., Guo, C., Ren, W., Cong, R., Hou, J., Kwong, S., Tao, D.: An underwater image enhancement benchmark dataset and beyond. IEEE Trans. Image Process. 4376–4389 (2019)
14. Chen, L., Liu, Z., Tong, L., Jiang, Z., Wang, S., Dong, J., Zhou, H.: Underwater object detection using invert multi-class adaboost with deep learning. In: Computer Vision and Pattern Recognition (cs.CV); Machine Learning (cs.LG) (2020)
15. Hu, J., Shen, L., Albanie, S., Sun, G., Wu, E.: Squeeze - and - excitation networks. In: Computer Vision and Pattern Recognition (cs.CV) (2017)

# Climate and Ecosystems Modelling I

# Integration of Process Modeling and LCA for the Assesment of the Environmental Impact of Pharmaceutical Industries: Case Study

**Shaimaa Ragab and Amna Ramzy**

**Abstract** Life Cycle Assessment (LCA) is utilized to evaluate the environmental, economic, and social impacts of a product or process across all stages of its life cycle, from raw material extraction to disposal. The carbon footprint, a measure of the total carbon dioxide emissions directly and indirectly caused by an activity or accumulated over a product's life stages, is a key component of this assessment. This study aims to demonstrate the application of LCA in an Egyptian pharmaceutical factory producing oral liquid drugs. Initially, Material and Energy Flow Analysis (MEFA) was conducted using Umberto Efficiency + before establishing the inventory for LCA using Umberto LCA + . The assessment followed ISO 14040/14044 standards, with a cradle-to-gate system boundary that considered raw material transportation, energy use, waste handling, and carbon footprint estimation. Results show that 78.25 kg $CO_2$-Eq is generated to produce 3,597.21 kg of oral liquid drug. Syrup production contributes over 85% to terrestrial ecotoxicity and climate change, with more than 25% impact on fossil depletion. This study highlights a significant advantage by being based on actual values from a real case study within a pharmaceutical factory. Gaining access to such data is particularly valuable, given the stringent regulations that make it difficult to obtain information from pharmaceutical factories. As such, this research fills a critical gap by applying LCA to non-sterilized liquid drugs, an area that has been largely unexplored. Moreover, LCA modeling links technology with sustainability, leveraging computer science to assess and enhance the environmental sustainability of complex systems.

**Keywords** Life Cycle Assessment · Liquid drug · Carbon footprint · Umberto software

S. Ragab
German International University, New Administrative Capital, Cairo, Egypt
e-mail: Shaimaa.ragab@student.giu-uni.de

A. Ramzy (✉)
German University in Cairo, New Cairo, Cairo, Egypt
e-mail: Amna.ramzy@guc.edu.eg

 
V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_8

## 1 Introduction

Life Cycle Assessment (LCA) has become a crucial tool for evaluating the environmental impacts of pharmaceutical products across their life cycles. Previous studies have explored different aspects of LCA in the pharmaceutical industry, focusing on various stages from cradle-to-gate.

LCA modeling serves as a method to link technology with sustainability, utilizing computer science to evaluate the environmental impact of our systems. By leveraging comprehensive life cycle data and integrating advanced modeling techniques, LCA allows us to quantify and improve the sustainability of pharmaceutical processes, bridging the gap between technological advancement and ecological responsibility.

For example, Gonzalez et al. characterized the environmental impact of Active Pharmaceutical Ingredient (API) manufacturing, revealing that solvent use significantly contributes to environmental impacts, particularly when incineration is used instead of recovery [1]. Kim et al. applied LCA to assess the life cycle inventories of enzymes, emphasizing the importance of modular inventory estimation techniques for accuracy in biocatalyzed processes [2].

Additionally, Parvatker et al. provided insights into the greenhouse gas emissions associated with anesthetic drugs, highlighting the correlation between synthesis complexity and environmental impact [3]. Other studies, such as those by Sherman et al., compared the environmental impacts of pharmaceutical packaging materials and anesthetic drugs, respectively, demonstrating the potential for significant environmental gains through material choice and process optimization [4].

However, despite these advances, there remain methodological challenges and gaps, particularly in the application of LCA to non-sterilized liquid drugs. Siegert et al. proposed harmonized rules to address these challenges, yet highlighted the need for further research, particularly in modeling the use and end-of-life phases of pharmaceutical products [5].

Finally In the pharmaceutical field, Sharma et al. compared the environmental effects of two paracetamol dosage forms using GaBi v8.0. The study revealed that syrup production contributes significantly to climate change (90%) and other impact categories such as fossil fuel depletion, freshwater consumption, and ecotoxicity. In contrast, tablet production mainly affects human toxicity and metal depletion. The findings indicate that syrup production has a more substantial impact across multiple categories compared to tablet production. This study provides valuable insights for environmental managers to better manage the sustainability of pharmaceutical products [6].

Building on this foundation, our study aims to address the underexplored area of LCA application in non-sterilized liquid drugs. By leveraging comprehensive life cycle data and addressing existing methodological challenges, this research contributes to a more sustainable and environmentally responsible pharmaceutical industry.

## 2 Experimental Work

This section describes the materials, methods, and tools that are used in this work. Experimental steps are described in detail.

### *2.1 Materials*

The manufacturing of the oral liquid drug, Halorange, a food supplement produced by an Egyptian factory, involves the use of carefully selected materials to ensure both safety and efficacy. The Active Pharmaceutical Ingredients (APIs) include high-quality vitamins such as vitamin C (ascorbic acid), vitamin E, and Cod liver oil (Halibut) sourced from reputable suppliers in China not mentioned according to factory privacy. Vegetable oils are the main dietary sources of vitamin E components. In the manufacturing of liquid drug, various materials play crucial roles in ensuring product stability, palatability, and overall efficacy. Ingredients such as Orange flavor and Citric acid contribute to enhancing the taste and aroma of the liquid drug, making it more palatable for patients. Orange oil is obtained from orange citrus oils. Sweeteners like Sucrose and Dextrose are often used to improve the overall flavor profile, especially in pediatric formulations. Preservatives such as Methylparaben sodium and Propylparaben sodium aid in preventing microbial contamination, extending the shelf life of the liquid drug. Methyl paraben sodium and Propyl paraben sodium are sodium salt of methyl paraben and sodium salt of propyl paraben respectively. Methyl paraben and propyl paraben are an ester of hydroxybenzoic acid. Thickening agents like Xanthan gum help achieve the desired viscosity, ensuring proper suspension of active pharmaceutical ingredients. Xanthan gum is composed of D-glucose, D-mannose, and D-glucuronic acid by a ratio of 2:2:1. Coloring agents such as Sun Set Yellow enhance visual appeal. The color of sunset yellow is produced by diazotizing 4-amino benzene sulfonic acid using hydrochloric acid and sodium nitrite and sodium nitrite then coupled with 6-hydroxy-2-napthalene sulfonic acid. Chelating agents like EDTA and emulsifiers like Tween 80 contribute to the stability of the formulation. Tween 80 is composed of 20 ethylene oxide, 1 sorbitol, and 1 oleic acid. Additionally, Sodium sulfite may be employed to maintain chemical stability. Collectively, these materials are carefully selected and integrated into the manufacturing process to ensure the safety, efficacy, and overall quality of the liquid drug. Finally packaging process, nitrogen-filled amber glass bottles serve as containers for liquid drugs, ensuring product stability and preventing degradation. Amber glasses are high in sand silica about 70%. The amber glass bottles are filled with syrup, and plastic caps are used to seal them securely. Plastic caps made of Polyethylene and Polypropylene. Each bottle is adorned with paper labels containing crucial information such as manufacturing date, expiry date, and batch number, applied using printing ink through an injection machine. Accompanying the drug bottles are leaflets and informative papers that provide instructions for use. To facilitate organized distribution, the bottles, along with their respective leaflets, are packed into sturdy boxes. Boxes are made of cellulose fiber. These boxes, in turn, are carefully arranged and secured within cartons, forming a comprehensive and well-protected packaging system for the liquid drugs produced in the factory. This meticulous packaging process ensures the integrity of the pharmaceutical products from production through distribution, safeguarding their quality until they reach the end user.

## 2.2 Methods and Tools

UMBERTO Efficiency + is software specifically made for Material and Energy flow simulation of a factory's production lines developed by the German company ifu Hamburg (member of iPoint Group). Umberto is used worldwide in industry, consulting as well as in Research Development, and at universities for various purposes. Because of the ease of use, UMBERTO was chosen as a program to represent the simulation tool in this thesis. UMBERTO LCA is used to analyze and reduce the environmental impact of products in less time. Take advantage of the opportunity to automate the creation of LCAs to a greater extent to save time and resources. Aggregating and assessing relevant data throughout the entire product life cycle for an LCA is an enormous task. Even more so, since most assessments start with the finished product and work their way down to components and substances. This often entails intensive efforts of manual data collection. With its existing granular database of product-related compliance and sustainability information, the iPoint LCA software solution enables companies to achieve a much more detailed assessment and reach a much higher level of automated data integration from different sources. Our Life Cycle Assessment software collects and combines the relevant data from the supplier network and internal production as well as data from the use phase and recycling and reuse objectives. That way companies get an overview of potential environmental risks and improvements. The LCA has been conducted according to ISO 14040/14044 standards. The LCA boundary of the two cases is cradle-to-gate, considering the transportation of raw materials and the manufacturing stage, cradle-to-grave is considered the transportation of raw materials, manufacturing stage, usage of product, and end of life. Ecoinvent Database enables users to gain a deeper understanding of the environmental impacts of their products and services. It is a repository covering a diverse range of sectors on a global and regional level. It currently contains more than 18,000 activities, otherwise referred to as 'datasets', modeling human activities or processes. Ecoinvent datasets contain information on the industrial or agricultural process they model, measuring the natural resources withdrawn from the environment, the emissions released to the water, soil, and air, the products demanded from other processes (electricity), and of course, the products, co-products, and wastes produced.

### 2.3 Data Acquisition and Collection

The purpose of the pharmaceutical factory visit is to observe the production lines, select a suitable production line for the manufacturing of liquid pharmaceutical drugs, map it, and collect relevant data. This information will enable the comprehension of the sequence and subsequently translate it into a simulation model using UMBERTO. The factory houses two production lines for liquid drugs: non-sterile (oral) drugs and sterile (injectable) drugs. For the purpose of this research, we have opted to focus on the non-sterile (oral) drug production line. The functional unit is linked to the goal and scope of the study. Since the focus of the present study was the production of liquid drug, the functional unit was taken as one batch (one batch of syrup = 12,500 bottles, each containing 120 ml of syrup) of Halorange liquid drug.

This particular production line comprises seven main machines: Dispensing machine, mixing and preparation tanks, storage tank, automatic liquid Filler and Capper, blow Air Cleaner, wrap-around Labeler, and ink injection.

Each machine has specific inputs and outputs (Fig. 1).

For syrup production, the process begins with raw materials being weighed using a dispensing machine. The raw materials are then added in a specific order to the mixer according to the batch manufacturing record. The syrup is stored in a storage tank with continuous mixing to prevent precipitation. Subsequently, the automatic filler and capper fill the syrup into cleaned bottles from the blow air cleaner and seal them with caps. The filled bottles are labeled using a wrap-around labeler. Following this, small boxes are labeled by ink injection before the final step, which is the packaging.

The amperes (A) of some machines were measured using a Clamp Meter, knowing the voltage (V) of all machines, and by considering the power factor (PF) of the factory in three phases, which is equal to 92%. We calculated power (kW) by applying Eq. 1 and then applying Eq. 2. For other machines such as the water station, compressor,

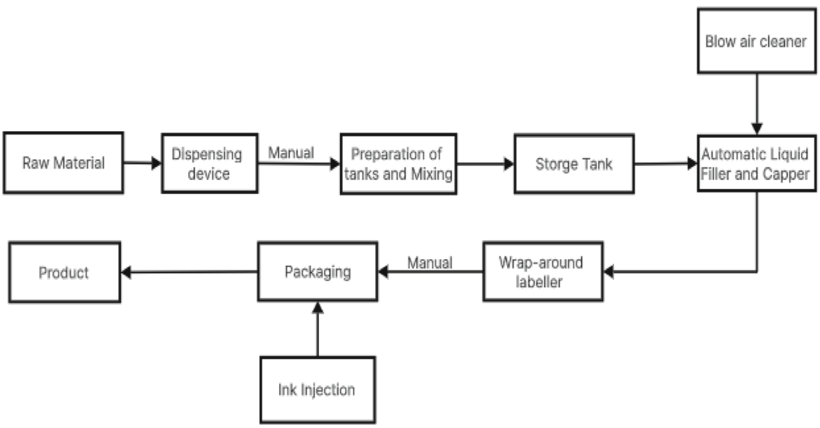



**Fig. 1** Mapping of the production line

chiller, and boiler, we couldn't measure the current (amperes). Therefore, we obtained their energy consumption from operational experience per hour or per day.

$$\text{Power(kW)} = \text{Reciprocal } 1000 * \sqrt{3} * \text{A} * \text{V} * \text{PF}$$

$$\text{Energy Consumption(kWh)} = \text{Power(kW)} * \text{hours(h)}$$

## *2.4 Modeling of the Process Using UMBERTO*

Umberto Efficiency + software was initially used to understand the flow of materials and energy throughout the systems. Material and Energy Flow Analysis (MEFA) using Umberto Efficiency + was carried out initially before LCA using Umberto LCA + to establish the inventor data. These data were taken into account while creating a relation function that defined each process in respect to its inputs, outputs, and wastes. For all the models, the inputs, outputs, and processes are represented in their according symbols.

The production line was divided into 5 phases (Dispensing, preparation of tanks and mixing, filling and capping, packaging, and product) in modeling the production line on UMBERTO Efficiency + and insert data collected on inputs, outputs, and connections (Fig. 2).

An LCA approach is used to calculate and assess the environmental impact associated with a product across its life cycle. By quantifying the environmental consequences, stakeholders can have a better understanding of the changes that should be made at various stages of the manufacturing process of a product or service to enhance overall environmental outcomes.

For this section, an LCA was performed on the production line of a liquid (oral) drug, taking into consideration the burdens imposed on the environment from each phase under the cradle-to-gate scope. Cradle-to-gate indicates measuring the environmental impact of raw materials, going through the production line until the final product. The key to undertaking a successful LCA is defining the phases and system boundaries and having a complete dataset of inputs from the Ecoinvent Database.

For this LCA, a two-phase model was chosen:

1. **Raw materials:** represents how the raw materials were delivered to the factory and from where.
2. **Manufacture:** represents all the processes these raw materials undergo until they reach their final form as a product.

Raw Materials: Active pharmaceutical ingredients (APIs) are imported from abroad, which will have a significant effect on LCA, such as Cod liver oil, Vitamin C (ascorbic acid), and Vitamin E imported from China. While printing ink is imported from England (Table 1, Fig. 3).

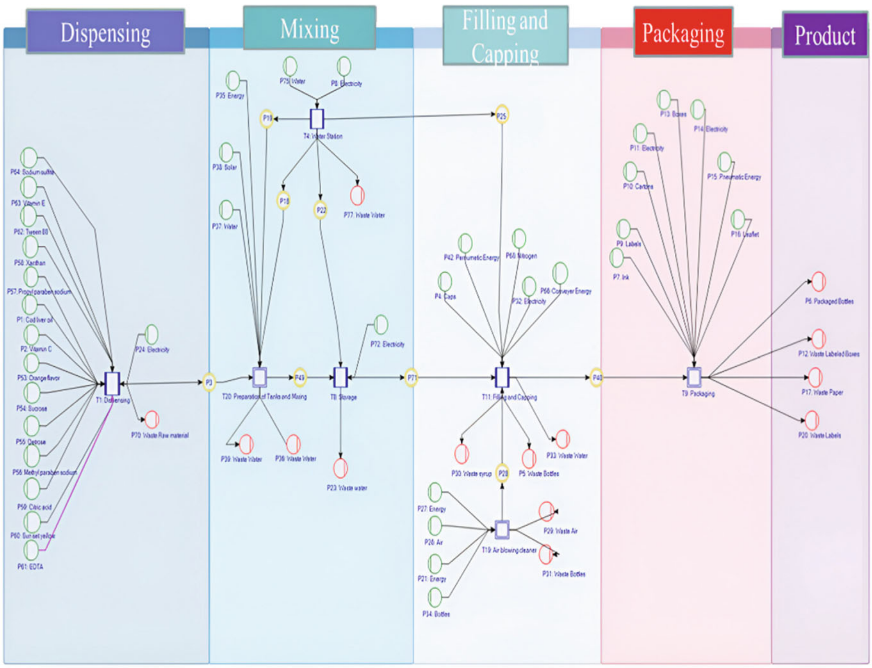



**Fig. 2** Liquid (oral) production line model

**Table 1** Transportation Data

| Name of Transportation | Mass of raw materials (kg) | Distance (km) | Value of transportation (metric) |
|---|---|---|---|
| APIs by lorry | 72.79575 | 429.1 | 31.23665633 |
| APIs by ship | 72.79575 | 15,604.95 | 1135.974039 |
| Ink by lorry | 0.104 | 287.5 | 0.0299 |
| Ink by ship | 0.104 | 3364.038462 | 0.34986 |
| Nitrogen | 1.0485 | 24.1 | 0.02526885 |
| Pharmaceutical Excipients | 813.354515 | 24.1 | 19.60184381 |
| Packaging materials | 1862.68218 | 24.1 | 44.89064054 |

Transportation value is calculated by this equation:

$$\text{Transportation(metric)} = \text{Mass of raw materials (kg)}*\text{Distance(km)}/1000$$

**Manufacture:** Environmental impact data for all raw materials, auxiliary materials, and operating materials must be extracted from the Ecoinvent database using

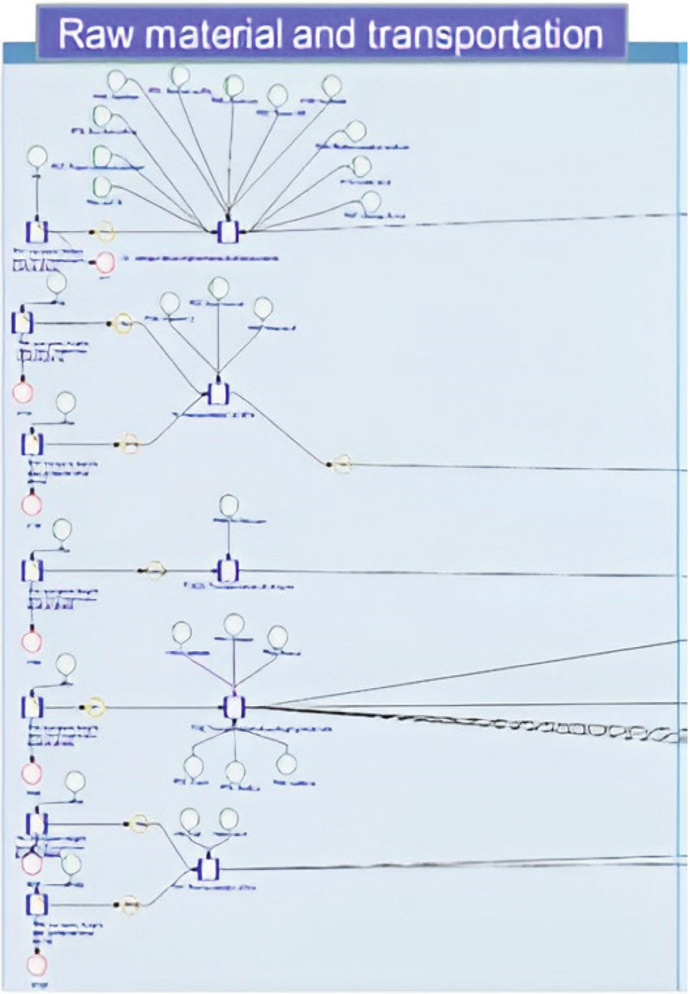



**Fig. 3** Raw materials and transportation phase

ReCiPe 2016 v1.03, midpoint (H) cutoff by classification, and uploaded to Umberto LCA software. In cases where certain materials are not available in the Ecoinvent database, material compositions or the nearest comparable compound/material are inserted instead. For example, liquid nitrogen is used in place of nitrogen gas when it is not found in the Ecoinvent database. This process involves transferring the environmental impact data from the Ecoinvent database to the Umberto LCA software.

# 3 Results and Discussion

This section explains the findings obtained from testing the model with a Sankey diagram after modeling the production systems on Umberto software, entering all required data, and defining all the processes. We then insert material flow.

## 3.1 Sankey Diagram on Umberto Efficiency +

After a network has been calculated, all material flows may be shown using color-coded Sankey arrows, with the understanding that larger quantities are represented visually by thicker lines. Yellow arrows indicate electricity, blue arrows indicate inputs and intermediate products, green indicates the final output outcome, and red indicates waste.

As shown in **Fig. 4**, the thickest arrow represents the preparation of tanks and the mixing process, indicating the highest energy consumption for this process. In contrast, the thinnest arrow represents the dispensing process, indicating the lowest energy consumption for this process.

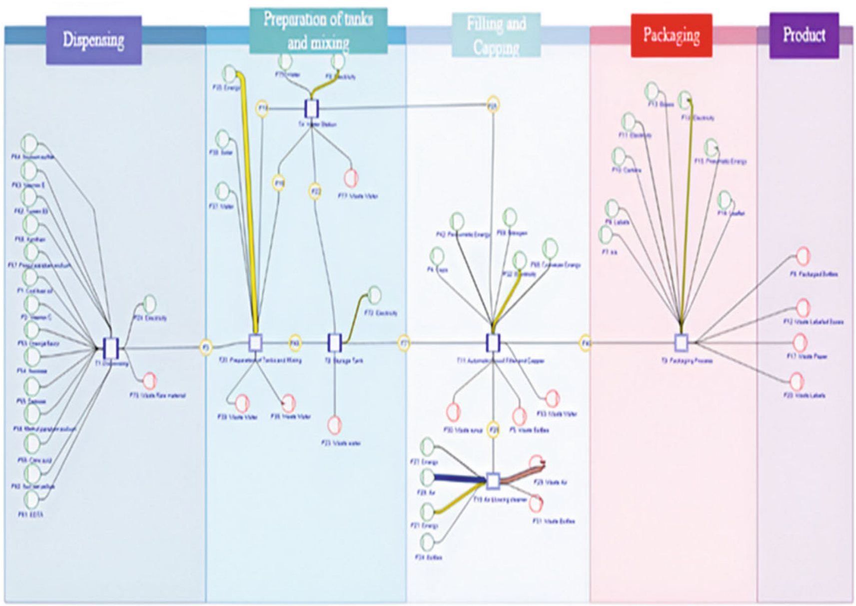



**Fig. 4** Sankey diagram on Umberto efficiency +

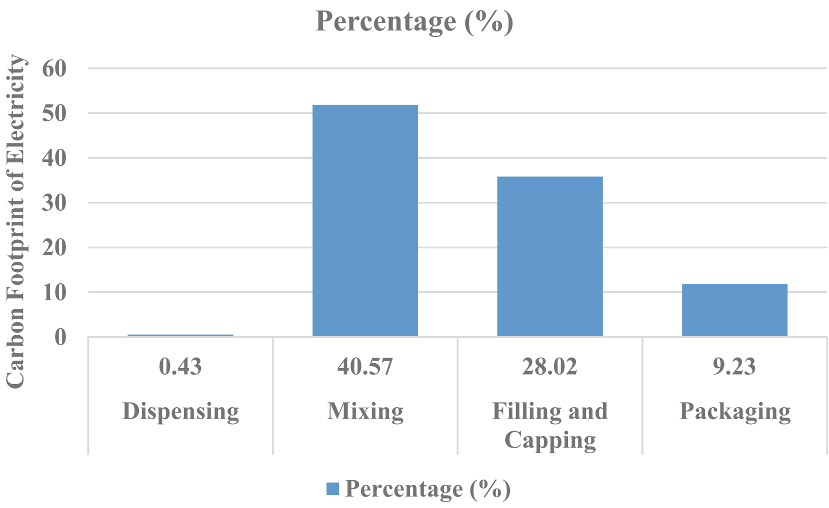



**Fig. 5** Carbon footprint of electricity by phases

## *3.2 Carbon Footprint Analysis of Electricity*

Inserting the value of the global warming potential (GWP100) of electricity in Egypt from the Ecoinvent database into the carbon footprint cell, calculate the total carbon footprint of the liquid production line using Umberto Efficiency. A total of 78.25 kg $CO_2$-Eq is produced to manufacture 3597.21 kg of oral liquid drug (Fig. 5).

## *3.3 LCIA Details*

After inserting the environmental impacts of each raw material, the LCA results are organized according to 18 environmental impacts (Table 2, Fig. 6).

### 3.3.1 LCIA Details by Phases

LCA results are categorized into two phases: Raw Material and Manufacture. So, LCA results show each environmental impact in each phase.

**Figure 7.** shows the environmental impact of a product across its life cycle, from raw material extraction to the final product. The impact is measured in twelve impact categories, such as climate change, human toxicity, and water depletion. For each category, the figure shows the impact of two phases: the raw material phase and the manufacture phase. The raw material phase includes the extraction, processing of the materials used in the product, and transportation of the product. The manufacture

**Table 2** LCIA results

| Impact Category | VALUE |
|---|---|
| Agricultural land occupation, ALOP ($m^2$*a crop-Eq) | 320.17 |
| Freshwater ecotoxicity, FETPinf (kg 1,4-DCB-Eq) | 64.34 |
| Metal depletion, MDP (kg Cu-Eq) | 24.3 |
| Urban land occupation, ULOP ($m^2$a) | 0.89 |
| Terrestrial acidification, TAP100 (kg $SO_2$-Eq) | 6.38 |
| Natural land transformation, NLTP ($m^2$) | 9.66E-03 |
| Terrestrial ecotoxicity, TETPinf (kg 1,4-DCB-Eq) | 5,718.49 |
| Photochemical oxidant formation, POFP (kg NOx-Eq) | 4.48 |
| Human toxicity, HTPinf (kg 1,4-DCB-Eq) | 97.96 |
| Ionizing radiation, IRP_HE (kBq Co-60-Eq) | 53.15 |
| Water depletion, WDP ($m^3$) | 51.7 |
| Ozone depletion, ODPinf (kg CFC-11-Eq) | 2.62E-03 |
| Marine ecotoxicity, METPinf (kg 1,4-DCB-Eq) | 81.92 |
| Particulate matter formation, PMFP (kg PM2.5-Eq) | 2.55 |
| Fossil depletion, FDP (kg oil-Eq) | 496.98 |
| Marine eutrophication, MEP (kg N-Eq) | 0.65 |
| Climate change, GWP100 (kg $CO_2$-Eq) | 1,710.62 |
| Freshwater eutrophication, FEP (kg P-Eq) | 1.26 |

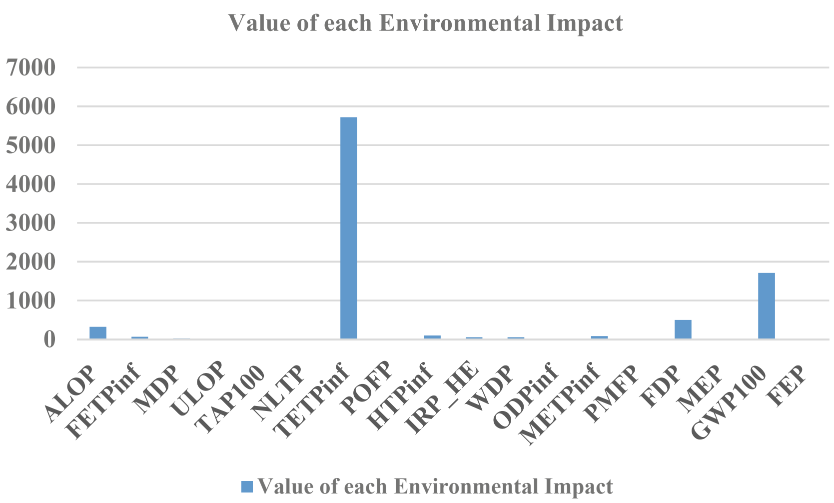



**Fig. 6** LCIA Results of environmental impacts

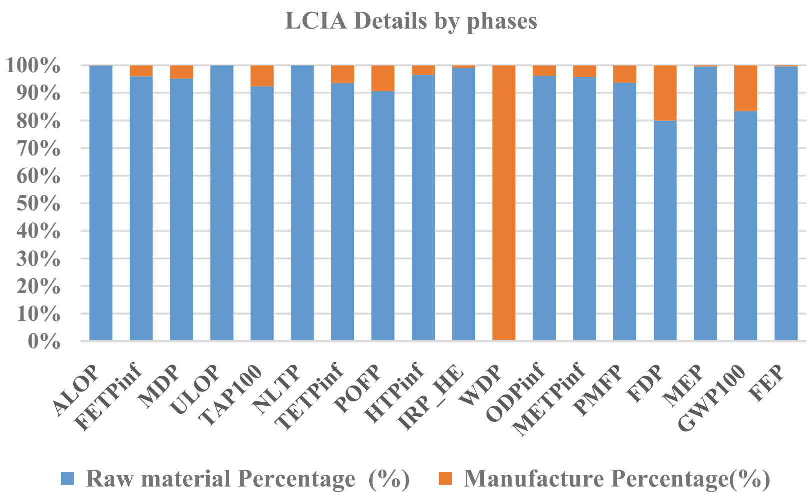



**Fig. 7** LCIA Results by Phases

phase includes only the manufacturing of a product. Notably, some impact categories, like Urban Land Occupation and Natural Land Transformation, have undefined values in the Manufacture phase. For instance, fossil depletion sees a higher percentage in the Raw Material phase, while Climate Change and Water Depletion have a larger impact during the Manufacture phase. These findings highlight the distinct environmental footprints associated with the Raw Materials and Manufacture phases for the specified impact categories.

### *3.4 Discussion*

A comparison with the research done by Sharma et al. [6] offers important context and helps assess the importance of the current study. While Sharma et al.'s research covers the environmental impacts of paracetamol dosage forms using life cycle assessment, the focus in our study is on the oral aspect; however, their techniques and main conclusions differ.

Sharma et al.'s [6] research focused on evaluating and comparing the environmental impacts of two different dosage forms of paracetamol—tablet and syrup. The field data were collected in real-time, utilizing technical documents such as batch manufacturing records (BMRs) employed by the pharmaceutical manufacturing organization located in the northern part of India, engaged in manufacturing pharmaceutical formulations.

This study utilized GaBi™ v8.0 life cycle assessment software for life cycle impact assessment and life cycle inventory compilation. The data were obtained from the GaBi v8.0 life cycle assessment software database (ReCiPe midpoint method). The syrup production process comprised two major constituent processes: syrup preparation and packaging, with an intermediate storage process between them. Additionally, there were two utility processes—steam production and deionized water production.

The functional unit was defined as one batch (one batch of syrup = ten thousand bottles, each containing 100 ml of syrup) of paracetamol dosage forms. Raw materials, including APIs and excipients, totaled 721.37 kg.

In contrast, the current study delves into the Life Cycle Assessment of liquid drug material (oral drug) utilizing field data collected in real-time and technical documents such as batch manufacturing records (BMRs) from a pharmaceutical manufacturing organization located in Egypt, involved in manufacturing pharmaceutical formulations.

The current study uses Umberto Efficiency + and Umberto LCA, with data sourced from the Ecoinvent database v3.9.1 cutoff (ReCiPe midpoint H method). The liquid drug production involves seven processes: dispensing, preparation of tanks and mixing, filling and capping, air blower cleaner, labeling, ink injection, and storage tank. Additionally, there are four utility processes: boiler, chiller, compressor, and water station.

The functional unit was defined as one batch (one batch of syrup = 12,500 bottles, each containing 120 ml of syrup) of halorange dosage forms. Raw materials, including APIs and excipients, totaled 886.150265 kg.

A noteworthy distinction lies in the energy consumption, equal to 124.5703924 kWh, and water consumption, equal to 4050 kg. In contrast, Sharma et al. (2021) relied on an energy consumption of 68.21 kWh and water consumption of 3467 kg.

Despite these differences, a shared observation between the two studies is the involvement of two processes (mixing and packaging until labeling), certain utilities (water station and steam), and specific raw materials (citric acid, methyl paraben sodium, propyl paraben sodium, and xanthan) but with varying quantities.

After eliminating a transportation section from the current research to compare its results with Sharma et al.'s findings, we created a functional unit that is the same (per one bottle = 100 ml) in both studies. This allows us to compare the environmental impacts between the two studies.

## 4 Conclusion

In conclusion, the Life Cycle Assessment (LCA) conducted on the unsterile (oral) liquid syrup production in an Egyptian pharmaceutical factory, utilizing Umberto Efficiency + and Umberto LCA software tools with the Ecoinvent database v3.9.1, has provided comprehensive insights into the environmental impacts associated with the entire life cycle, from cradle-to-gate, including the transportation of raw materials.

The results of the LCA analysis revealed a total of 18 environmental impact categories, with notable indicators such as agricultural land occupation, freshwater ecotoxicity, metal depletion, urban land occupation, terrestrial acidification, natural land transformation, terrestrial ecotoxicity, photochemical oxidant formation, human

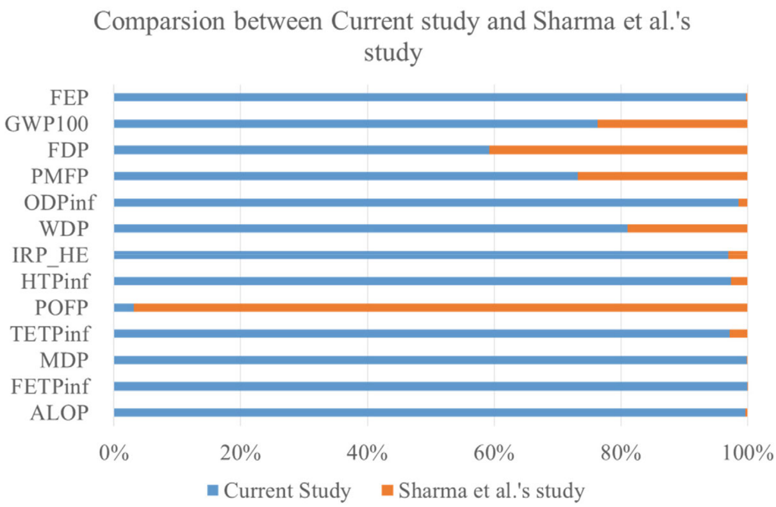



**Fig. 8** Comparison between current study and Sharma et al.'s study

toxicity, ionizing radiation, water depletion, ozone depletion, marine ecotoxicity, particulate matter formation, fossil depletion, marine eutrophication, climate change, and freshwater eutrophication. The results show that syrup production has more than a 50% contribution to the terrestrial ecotoxicity impact category. Furthermore, the results indicate that syrup production has more than 85% contribution to terrestrial ecotoxicity and climate change. The results indicate more than 25% contribution to climate change and fossil depletion. In contrast, Sharma et al.'s study LCA analysis revealed a total of 13 environmental impact categories, with notable indicators such as agricultural land occupation, freshwater ecotoxicity, metal depletion, terrestrial ecotoxicity, photochemical oxidant formation, human toxicity, ionizing radiation, water depletion, ozone depletion, particulate matter formation, fossil depletion, climate change, and freshwater eutrophication. The results indicate more than 70% contribution to climate change and fossil depletion. Comparing the current study with Sharma et al.'s study involves an objective 'cradle-to-gate' assessment of the production of paracetamol dosage forms. The results of the two papers differ from each other due to differences in certain conditions such as batch capacity, software used, and the database from which the data is imported, the number of processes, the number of utilities, the final product, geographical location, raw materials, APIs, and energy consumption (Fig. 8).

## List of Abbreviations

ALOP: Agricultural land occupation.
FETPinf: Freshwater ecotoxicity.

| | |
|---|---|
| MDP: | Metal depletion. |
| ULOP: | Urban land occupation. |
| TAP100: | Terrestrial acidification. |
| NLTP: | Natural land transformation. |
| TETPinf: | Terrestrial ecotoxicity. |
| POFP: | Photochemical oxidant formation. |
| HTPinf: | Human toxicity. |
| IRP_HE: | Ionizing radiation. |
| WDP: | Water depletion. |
| ODPinf: | Ozone depletion. |
| METPinf: | Marine ecotoxicity. |
| PMFP: | Particulate matter formation. |
| FDP: | Fossil depletion. |
| MEP: | Marine eutrophication. |
| GWP100: | Climate change. |
| FEP: | Freshwater eutrophication |

## References

1. Jiménez-González, C., Curzons, A. D., Cunningham, V. L.: Expanding GSK's Solvent Selection Guide: Application of Life Cycle Assessment to Enhance Solvent Selections. Clean Technol. Environ. Policy. Springer Science+Business Media (2004)
2. Kim, S., Dale, B. E.: Enzymes for Pharmaceutical Applications—A Cradle-to-Gate Life Cycle Assessment. Int. J. Life Cycle Assess., Springer Science+Business Media (2009)
3. Parvatker, A.G., Tunceroglu, H., Sherman, J.D., Coish, P., Anastas, P., Zimmerman, J.B., Eckelman, M.J.: Cradle-to-gate greenhouse gas emissions for twenty anesthetic active pharmaceutical ingredients based on process scale-up and process design calculations. ACS Sustainable Chem. Eng. **7**(7), 6580–6591 (2019)
4. Sherman, J., Le, C., Lamers, V., Eckelman, M.: Life cycle greenhouse gas emissions of anesthetic drugs. Anesth. Analg. **114**(5), 1086–1090 (2012)
5. Siegert, M.-W., Lehmann, A., Emara, Y., Finkbeiner, M.: Harmonized rules for future lcas on pharmaceutical products and processes. Int. J. Life Cycle Assess. **24**, 1040–1057 (2019)
6. Sharma, R., Gupta, P., Singh, A.: Environmental Impacts of Paracetamol Dosage Forms Using Life Cycle Assessment, 2nd edn. Springer, New York (2021)

# Modeling for the Design and Manufacturing of Superabsorbent Polymers (SAPs) for Agriculture in Arid Areas by Using the RSM Method

**Maysa Muhammad, M. L. Tawfic, Mohamed Taha, and Ahmed Elsabbagh**

**Abstract** The objective of this study is to use JMP software for modeling by using the RSM method to design and analyze an engineering product that has the potential to assist and change agriculture in semi-arid and arid areas. Designing and analyzing an engineering product in the form of absorboost tiles (ATs), these ATs assist in the replacement of the functions of soil, for easy interchangeability these AT dimensions are according to the effect root zone of the plant type.

**Keywords** JMP · Modeling · RSM method · Superabsorbent polymers · Design · Manufacturing · Agriculture · Arid Areas · Absorboost Tile · Hydrogel · Sodium Poly Acrylate (NaPA)

## 1 Introduction

Mankind civilization began with agriculture, agriculture necessity extends from nutrition and health aspects to other aspects like reducing environmental pollution, economic development, and protecting animals from extinction [1]. Water consumption in irrigated agriculture is 87% percent of global freshwater [2]. By 2050, feeding a planet of 9 billion people will require about a 50% percent increase in the agriculture sector and an increase in water consumption by 15% percent [3]. The aridity problem is in about 50% of the world countries [4]. One promising way to solve the problem is associated with the use of modern AT technologies to water conservation,

M. Muhammad · M. Taha · A. Elsabbagh (✉)
Design and Production Engineering Department, Ain Shams University, Abbasiya, Cairo, Egypt
e-mail: elsabbagh.ahmed@eng.asu.edu.eg

M. L. Tawfic
Chemistry Department, Ain Shams University, Abbasiya, Cairo, Egypt

Polymers Chemistry and Pigments Department, National Research Center (NRC), El Buhouth, Giza, Egypt

A. Elsabbagh
Design and Production Engineering Department, Galala University, New Galala, Egypt

 
V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_9

improved soil health, improved crop yields, versatility, cost-effective, eco-friendly, biodegradable, and environmental benefits [5]. ATs have proven beneficial for growth and sandy soil health [6]. Design of experiments (DOE) involves numerical methods and model techniques to analyze and optimize experimental data. The numerical data, such as regression analysis, analysis of variance (ANOVA), optimization algorithms, Monte Carlo simulation, and finite element analysis (FEA). The modeling techniques, such as response surface methodology (RSM), Taguchi methods, factorial design, central composite design (CCD), and box-Behnken design. The important benefits of modeling by using JMP (a statistical software suite), such as comprehensive data analysis, predictive modeling, comparison and selection, ease of use, cross-validation, and real-time data visualization [7]. The potential engineering solution is AT manufactured from superabsorbent polymers (SAPs), ground rice straw (RS) utilization, and natural adhesive (NA). SAPs are hydrophilic polymers that absorb and retain large amounts of water [8]. They have functions in the agriculture sector for water conservation, swelling capacity, soil improvement, seed coating, and drought resistance [9]. RS is a byproduct of rice milling and can benefit agriculture. The benefits of ground rice straw for agriculture are soil amendment, water retention, sustainable mulch, silica content, and pest control [10]. The characteristics of NA for agriculture are a natural binder that helps bind RS fibers length together and enhances the mechanical strength of RS [11]. The study objective was achieved by passing some steps. The first step is to manufacture SAP powders by bulk polymerization process [12] such as (NaPA) with five varying cross-linker contents, (wt.) and evaluate the performance for SAPs powder after their manufacture according to ASTM standards. The second step is to prepare the specimens with three manufacturing processes such as preparing raw material, RH crushing, RS shredding, and RS screening with three different thicknesses classified as small RS fibers length (74–100 $\mu$m) [13]. The third step is to NA solution. The fourth step is to design and manufacture a manual press (test rig) to produce AT. The fifth step is to design and manufacture absorboost tile according to the design of experimental (DOE) parameters. Finally, the experimental analysis for absorboost tile is very important for evaluating the absorboost tile performance used in agriculture in arid areas. Experimental analysis of absorboost tile involves mechanical [14], physical [15], and chemical properties [16], SEM, FTIR, TGA, pF curve [17]. Experimental analysis of SAPs involves various methods to characterize their mechanical properties and performance. This study focuses on modeling and analyzing the compressive strength of absorboost tile with five varying of cross-linker content, (wt.) and with four varying of weights (g) of NaPA. Modeling and analyzing for absorption capacity for absorboost tile with five varying cross-linker content, (wt.) and with three varying interval times (min). Modeling and analyzing for durability with four varying SAP weights (g) and with four varying interval times (day).

## 2 Materials and Methods

### *2.1 Materials*

The materials investigated in this study were classified into three categories: First manufacturing by free radical bulk polymerization SAPs powder from chemical materials with five varying cross-linker contents, (wt.). Next, the preparation of the following agricultural residues RS specimens with small fibers length dimensions, and the final preparation of natural adhesive (NA). Hydrogels were manufactured based on a hygroscopic type (cross-linker).

### *2.2 Design and Manufacturing of Absorboost Tile*

First, the manufacturing of five samples of sodium polyacrylate (NaPA) with varying cross-linker content, (wt.). RS are agricultural residues from rice grain processing, with an optimum moisture content of 8% to 12%, with the limitation of low lignin content. RS was obtained from Mashtohor Toukh Village, Qalyubia Governorate. RH are collected after milling the rice, dried to reduce moisture content, crushed on a chipper, shredded, and screened for small fibers length [18]. The important characteristic of the NA binder like high water absorption, where it absorbs and retains water due to its branched amylopectin structure [19], the thickening agent creates a clear and glassy texture stabilizer helps stabilize emulsions and prevents the separation of liquids [20], and low acid levels [21]. Binders enhance the physical and mechanical properties of tiles, making them more durable, NA based binders improve physical and mechanical properties [22, 23]. Design and manufacturing manual press depend on the required effect root zone dimensions, and required compressing pressure. The manual press consists of AT mold, hydraulic jack, pressure gauge, rods, three plates, spring, and bearing. Mixing and formulation for RS, NA solution, and NaPA. Compacting and shaping for AT according to the effect root zone for plant type.

### *2.3 SAP Materials Used for Agriculture*

SAPs have various applications in agriculture due to their water absorption and retention properties. The most commonly used types of SAPs in agriculture are biodegradable (natural) SAP is primarily used as a water retention agent in agriculture, it helps to improve soil moisture retention in arid regions, important natural materials used in agriculture such as cellulose, starch, alginate, chitosan, pectin, and gelatin [24]. Semi synthetic SAP is widely used in agriculture, and contributes to water management by enhancing soil water retention, important semi synthetic SAP materials used in agriculture such as starch and acrylonitrile graft copolymer and polyvinyl alcohols

(PVA) [25, 26], or synthetic SAP help in water Polyvinyl alcohols (PVA) [25, 26], or synthetic SAP help in water conservation, support crop growth, the important synthetic SAP used in agriculture such as polyacrylates and polyacrylamide [27].

# 3 Experimental Method

## *3.1 Compressive Strength (Mpa)*

Measuring compressive strength involves determining the maximum load a material can withstand before failing. Prepare the Specimen:

The experimental method steps:

1. Prepare the specimen of absorboost tile to be tested is typically according to ASTM D695 standard shape.
2. Set up the Testing machine (AMETEK LLOYD Instruments, Universal Testing Machine -5 KN, Faculty of Engineering, Ain Shams University).
3. A compression testing machine is used for this test. The machine applies a uniaxial compressive load to the specimen.
4. Place the Specimen: Position the specimen between the platens of the testing machine. Ensure it is aligned properly to avoid uneven loading.
5. Gradually apply the compressive load to the specimen. The load should be applied continuously and without shock, until the specimen fails.

## *3.2 Absorbency Capacity (g/g)*

The water absorbency of SAP hydrogel defined the ability of absorb and retain of water or other liquid. Water Absorption ASTM D570.

The experimental method steps:

1. Take absorboost tile sample.
2. Immerse the absorboost in 300 ml water and allow it fully absorb water.
3. Use filter paper to remove any excess water and weight the final swollen SAP sample.
4. Calculate the water absorbency using the weight gain method Eq. (1).

$$AC = \frac{W_w - W_d}{W_d} \tag{1}$$

where:

$W_w$: The wet weight of absorboost tile sample.

$W_d$: The dry weight of absorboost tile sample.

### 3.3 Durability

Measuring the weight of absorboost tiles with respect to time can be important for several reasons, especially in the context of manufacturing, quality control, and application agriculture. Over time, tiles can absorb moisture from the environment, which can affect their weight. Monitoring this change helps in understanding the absorboost tile's durability and suitability for different environments. Consistent weight measurements over time ensure that the tiles are manufactured to the specified standards. Any significant deviation might indicate issues in the production process [28]. Absorboost tiles can degrade over time due to environmental factors such as humidity, temperature changes, and chemical exposure. Measuring weight changes can provide insights into the rate of degradation and the longevity of the tiles. For agricultural absorboost tiles, such as those used in irrigation systems or soil stabilization, weight changes can indicate the absorption of water or other substances. This information is vital for optimizing their performance and ensuring they meet the required specifications. By regularly measuring the weight of tiles over time, manufacturers and users can ensure quality, durability, and safety in their applications.

## 4 Design of Experiment (DOE) Process

DOE allows for the investigation of multiple factors simultaneously, reducing the number of experiments needed compared to one-factor-at-a-time approaches and saves time and resources. It helps in identifying which factors have the most significant impact on the response variable. This is crucial for optimizing processes and improving product quality. DOE can reveal interactions between factors that might not be apparent when studying factors individually. Understanding these interactions is essential for complex systems. By using response surface methods and other advanced techniques, DOE helps in finding the optimal settings for factors to achieve the desired outcome. This is particularly useful in manufacturing, agriculture, and product development. DOE helps in developing robust processes and products that perform well under a variety of conditions [29]. By optimizing processes and reducing variability, DOE can lead to significant cost savings in production and development.

DOE encompasses various methods to systematically investigate the effects of multiple factors on a response variable [30].

- Full Factorial Design
- Fractional Factorial Design
- Response Surface Methodology (RSM)
- Taguchi Methods
- Central Composite Design (CCD)
- Box-Behnken Design
- Plackett–Burman Design
- Mixture Design

In this study use response surface methodology (RSM) method, RSM is used to explore the relationships between several explanatory variables and one or more response variables. It helps in optimizing processes by fitting a polynomial model to the experimental data. The RSM method passes through three stages: construction of the experimental design, response modeling, and graphic representations. The modeling used in an RSM study is quadratic design such as the central composite design units (Box-Wilson) or Box-Behnken. RSM is used to estimate the effects of individual parameters, the interaction of variables, efficiency, identification of key factors, and the optimum conditions for responses.

## *4.1 Importance of Using the RSM Method for Absorboost Tile*

RSM helps find the optimal conditions for absorboost tile and reduces the experimental chemical waste. By modeling the relationship between the response such as compressive strength (Mpa), absorption capacity (g/g), and durability, and the input variables such as cross-linker content wt., weight, and interval time (min, days), it allows for the identification of the best settings to achieve desired outcomes. RSM can significantly improve the quality and performance of absorboost tile. This is particularly valuable in the manufacturing and product development of absorboost tile. RSM provides a mathematical model that can predict the response for any combination of input variables within the studied range. This predictive capability is useful for making informed decisions without additional experiments. RSM helps in making processes more robust by identifying conditions that minimize variability and enhance reliability for absorboost tile. RSM techniques often include graphical representations like contour plots and surface plots. These visual tools make it easier to understand the relationships between variables and the response [30]. This study focuses on three factors, such as cross-linker content (wt.), interval time (min, day), and weight (g) of SAP. These factors are those that have a direct initial effect on compressive strength (Mpa), absorption capacity (g/g), and durability. In JMP software, optimization involves finding the best values for certain variables to maximize or minimize an objective function, often subject to constraints.

## *4.2 Equations and Methods Used in JMP for Optimization*

- Objective function is the Eq. (2) to optimize and to maximize profit.

$$\text{Profit} = \text{Revenue - Cost} \tag{2}$$

- Constraints are conditions that the solution by Eq. (3).

$$A\,x = b \tag{3}$$

where:

A is a matrix of coefficients,
x is the vector of variables,
b is a vector of constants,

- Constrained maximize or constrained minimize functions to find the optimal values of variables subject to constraints. The general form is:

  optimize f (x)subject to g(x) $\leq 0$

  Where:
  f(x) is the objective function,
  g(x) represents the constraints,

- Desirability functions are used to optimize multiple responses simultaneously. Each response is transformed into a desirability value ranging from 0 (completely undesirable) to 1 (fully desirable). The overall desirability is then maximized in the following Eq. (4).

$$D = \left(\prod_{k=1}^{n} d_i\right) \tag{4}$$

where:

$d_i$ is the desirability of the (i)the response,

- RSM is used to model and optimize processes. The second-order polynomial model is commonly used by Eq. (5):

$$y = \beta_0 + \sum_{i=n}^{n} \beta_i x_i + \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} x_i x_j + \sum_{i=n}^{n} \beta_{ii} x^{n-1_i} + \in \tag{5}$$

where:

y is the predicted response (lead removal capacity), $x_i$ and $x_j$ are the independent variables, $\beta_0$, $\beta_i$, $\beta_{ii}$, and $\beta_{ij}$ are the model constant, the linear coefficients, the quadratic Coeff., and the cross-product coefficients, respectively ε represents [31]. Table 1 illustrates the experimental design factors with two variables, five levels for cross-linker content, (wt.) (gram), and four levels for weight of NaPA (gram) and the one response is minimized value obtained from the experimental work for compressive strength for absorboost tile. Table 2 illustrates the experimental design factors with two variables, five levels for cross-linker content, (wt.) (gram), and three levels for time (min) and the one response is maximized value obtained from the experimental work for absorption capacity for absorboost tile. Table 3 illustrates the experimental design factors with two variables, four levels for weight of NaPA(gram), and swelling time (days) and the one response is maximized value obtained from the experimental work for durability for absorboost tile.

**Table 1** The levels and ranges of each variable (factors) for compressive strength

| Variable (Factors) | Code | Units | Levels | | | | |
|---|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 | 5 |
| Cross-linker content, (wt.) | X1 | gram | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 |
| Weight of NaPA | X2 | gram | 2 | 3 | 4 | 5 | |

**Table 2** The levels and ranges of each variable (factors) for absorption capacity

| Variable (Factors) | Code | Units | Levels | | | | |
|---|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 | 5 |
| Cross-linker content, (wt.) | X1 | gram | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 |
| Time | X2 | min | 20 | 40 | 60 | | |

**Table 3** The levels and ranges of each variable (factors) for durability

| Variable (Factors) | Code | Units | Levels | | | |
|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 |
| **Wight** | X1 | gram | 2 | 3 | 4 | 5 |
| **Time** | X2 | day | 1 | 2 | 3 | |

# 5 Results and Discussion

## *5.1 Experimental Design*

The importance of the experimental design lies in the experimental values and results after modeling and analysis via JMP software. Measure efficiency, as DOE allows multiple factors to be investigated simultaneously, reducing the number of experiments. This is crucial for optimizing processes and improving product quality. Table 4 illustrates the experimental matrix of compressive strength of the absorboost tile, the number of experiments is twenty DOE runs of two operators and one resonance. Table 5 illustrates the experiment matrix for the absorption capacity of the absorboost tile, the number of experiments is fifteen runs using DOE of two factors and one response. Table 6 illustrates the experiment matrix for the compressive strength of the absorboost tile, the number of experiments is twelve runs using DOE of two factors and one response [32].

**Table 4** Experiment matrix of the compressive strength for absorboost tile

| Run | Experimental value | | Absorboost Tile |
|---|---|---|---|
| | Cross-linker content, (wt.) (g) | Weight of NaPA (g) | Compressive strength |
| 1 | 0.4 | 2 | 60 |
| 2 | 0.4 | 3 | 54 |
| 3 | 0.5 | 2 | 77 |
| 4 | 0.5 | 3 | 73 |
| 5 | 0.1 | 5 | 51 |
| 6 | 0.1 | 5 | 51 |
| 7 | 0.1 | 4 | 63 |
| 8 | 0.5 | 4 | 60 |
| 9 | 0.1 | 3 | 78 |
| 10 | 0.2 | 4 | 39 |
| 11 | 0.1 | 2 | 99 |
| 12 | 0.5 | 5 | 44 |
| 13 | 0.1 | 2 | 99 |
| 14 | 0.4 | 4 | 46 |
| 15 | 0.4 | 5 | 30 |
| 16 | 0.2 | 3 | 48 |
| 17 | 0.5 | 2 | 77 |
| 18 | 0.2 | 5 | 32 |
| 19 | 0.2 | 2 | 48 |
| 20 | 0.5 | 5 | 44 |

### 5.2 Coefficient Determination $R^2$, Studentized Residuals, and Box-Cox Transformations

R2 explanation of variance of the coefficient of determination measures the proportion of the variance in the dependent variable that is predictable from the independent variables. It provides insight into how well the model explains the variability of the response data. By identifying observations with large studentized residuals, you can determine which data points have a significant influence on the model. The Box-Cox transformation is used to stabilize variance and make the data more closely conform to a normal distribution. This is important for meeting the assumptions of many statistical tests and models [33]. Figure 1 illustrates the experimental and predicted response, studentized residuals, and box-cox transformation for compressive strength for absorboost tile, where the significant value ($P$-value) > 0.0001 and $R^2$ value is 91%. Figure 2 illustrates the experimental and predicted response, studentized residuals, and box-cox transformation for absorption capacity for absorboost tile, where the significant value ($P$-value) > 0.0001 and $R^2$ value is 88%. Figure 3 illustrates

**Table 5** Experiment matrix of the absorption capacity (g/g) for absorboost tile

| Run | Experimental value | | Absorboost Tile |
|---|---|---|---|
| | Cross-linker content, (wt.) (g) | Time (min) | Absorption capacity (g/g) |
| 1 | 0.4 | 40 | 141 |
| 2 | 0.2 | 60 | 156 |
| 3 | 0.5 | 60 | 124 |
| 4 | 0.5 | 60 | 124 |
| 5 | 0.1 | 40 | 175 |
| 6 | 0.5 | 20 | 100 |
| 7 | 0.5 | 40 | 115 |
| 8 | 0.1 | 60 | 188 |
| 9 | 0.1 | 20 | 128 |
| 10 | 0.4 | 60 | 141 |
| 11 | 0.2 | 20 | 124 |
| 12 | 0.4 | 20 | 118 |
| 13 | 0.1 | 60 | 188 |
| 14 | 0.2 | 40 | 145 |
| 15 | 0.1 | 20 | 128 |

**Table 6** Experiment matrix of the durability for absorboost tile

| Run | Experimental value | | Absorboost Tile |
|---|---|---|---|
| | Weight of NaPA (g) | Time (day) | Durability |
| 1 | 2 | 3 | 98 |
| 2 | 2 | 2 | 110 |
| 3 | 4 | 3 | 103 |
| 4 | 5 | 1 | 129 |
| 5 | 4 | 1 | 127 |
| 6 | 3 | 1 | 126 |
| 7 | 3 | 2 | 112 |
| 8 | 3 | 3 | 101 |
| 9 | 5 | 2 | 117 |
| 10 | 4 | 2 | 115 |
| 11 | 2 | 1 | 125 |
| 12 | 5 | 3 | 107 |

the experimental and predicted response, studentized residuals, and box-cox transformation for durability for absorboost tile, where the significant value ($P$-value) > 0.0001 and $R^2$ value is 99%.

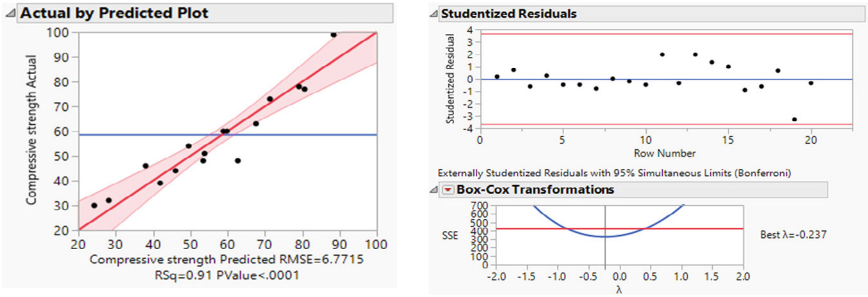



**Fig. 1** Experimental and predicted response, studentized residuals, and box-cox transformation for compressive strength for absorboost tile

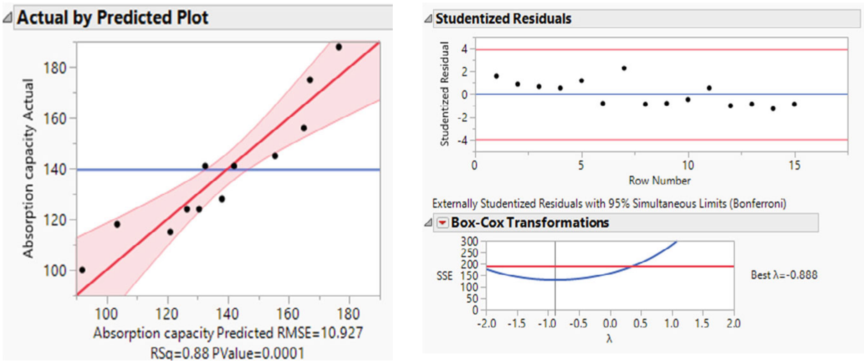



**Fig. 2** Experimental and predicted response, studentized residuals, and box-cox transformation for absorption capacity for absorboost tile

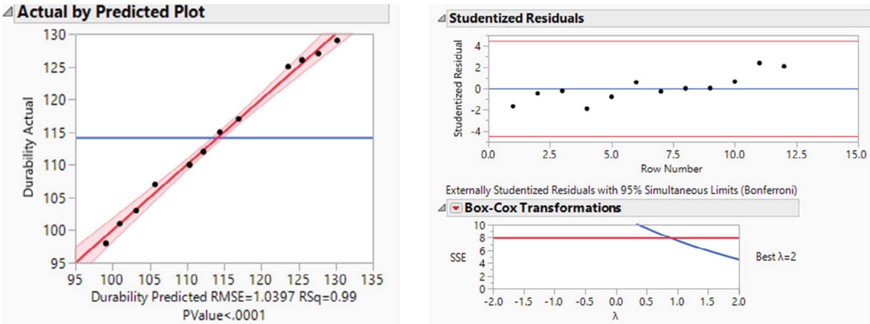



**Fig. 3** Experimental and predicted response, studentized residuals, and box-cox transformation for durability for absorboost tile

### 5.3 Analysis of Variance (ANOVA)

ANOVA is a statistical tool with several important applications and benefits [34]. Table 7 illustrates the quadratic polynomial model of compressive strength for absorboost tile,, such as comparing multiple groups, understanding variability, identifying significant differences, handling multiple variables, and improving experimental design where the results of AT of the polynomial model (central composite design) is significant at the confidence level of 0.05, given that $p$-values are less than 0.05 for absorboost tile. Weight of NaPA (2,5), cross-linker content, (wt.) * cross-linker content, (wt.), cross-linker content, (wt.) (0.1,0.5), and weight of NaPA*weight of NaPA with a $p$-value of 0.00000, 0.00000, 0.04818, and 0.48392, respectively. Table 8 illustrates the quadratic polynomial model of compressive strength for absorboost tile, where the absorption capacity results of absorboost tile of the polynomial model (central composite design) is significant at the confidence level of 0.05, given that $p$-values are less than 0.05 for AT. The cross-linker content, (wt.) (0.1–0.5) gram, time (20, 60) min, the time (min)* time (min), and the cross-linker content, (wt.) * cross-linker content, (wt.) with a $p$-value of 0.00005, 0.00018, 0.16101, and 1.00000, respectively. Table 9 illustrates the quadratic polynomial model of durability for AT, where the results of absorboost tile of the polynomial model (central composite design) is significant at the confidence level of 0.05, given that $p$-values are less than 0.05 for absorboost tile. Weight of NaPA (2,5), weight of NaPA* weight of NaPA, time (1,3), and time* time (day) with a $p$-value of 0.00000, 0.00008, 0.16026, and 0.59597, respectively.

**Table 7** Quadratic polynomial model of compressive strength for absorboost tile

| Term | DF | Estimate | Sum of Squares | t ratio | *P*-value |
|---|---|---|---|---|---|
| Weight of NaPA (2,5) | 1 | –17.30172 | 3858.2845 | –9.17 | 0.00000 |
| Cross-linker content, (wt.) * Cross-linker content (wt.) | 1 | –2.53125 | 23.6250 | –0.72 | 0.00000 |
| Cross-linker content, (wt.) (0.1,0.5) | 1 | –3.892857 | 212.1607 | –2.15 | 0.04818 |
| Weight of NaPA* weight of NaPA | 1 | 31.666667 | 2632.2917 | 7.58 | 0.48392 |

**Table 8** Quadratic polynomial model of absorption capacity (g/g) for absorboost tile

| Term | DF | Estimate | Sum of Squares | t ratio | *P*-value |
|---|---|---|---|---|---|
| Cross-linker content, (0.1–0.5) | 1 | –23 | 5424.7453 | –6.74 | 0.00005 |
| Time (min) (20, 60) | 1 | 1.99e-15 | 7.90e-30 | 0 | 0.00018 |
| Time (min)* Time (min) | 1 | 19.25 | 3983.2033 | 5.78 | 0.16101 |
| Cross-linker content, (wt.) * Cross-linker content, (wt.) | 1 | –9.75 | 273.639 | –1.51 | 1 |

**Table 9** Quadratic polynomial model of durability for absorboost tile

| Term | DF | Estimate | Sum of Squares | t ratio | *P*-value |
|---|---|---|---|---|---|
| Weight of NaPA (2,5) | 1 | 33 | 72.60000 | 8.20 | 0.00000 |
| Weight of NaPA* weight of NaPA | 1 | 0.375 | 0.3333 | 0.56 | 0.00008 |
| Time (1,3) | 1 | –12.25 | 1200.5000 | –33.33 | 0.16026 |
| Time * Time (day) | 1 | 1 | 2.667 | 1.57 | 0.59597 |

## *5.4 Optimization*

### 5.4.1 Prediction Profiler

It is important for interactive visualization, exploring factor effects, optimization, understanding model predictions, scenario analysis, communicating results, and model validation [35]. The objective of parameter optimization is to find the optimal setting of the two parameters that lead to maximizing the removal efficiency of 95% for absorboost tile, where the optimal operating conditions that indicated to this performance, minimizing the compressive strength, maximizing of the absorption capacity, and maximizing the durability of absorboost tile, for determine the optimal condition, will apply the function (desirability); it giving optimal adjustment, which varies between 0 and 1. Figure 4 illustrates prediction profiler for compressive strength for AT, where the results of analyzed is the suitable weight of NaPA = 2 g, cross-linker content, (wt.) = 0.1 g and desirability 0.19. Figure 5 illustrates prediction profiler for absorption capacity for AT, where the results of analyzed is the suitable cross-linker content, (wt.) = 0.1 g, swelling time = 20 min, and desirability 0.48. Figure 6 illustrates prediction profiler for durability for AT, where the results of analyzed is the suitable weight of NaPA = 2 g, swelling time = 1 day and desirability 0.84.

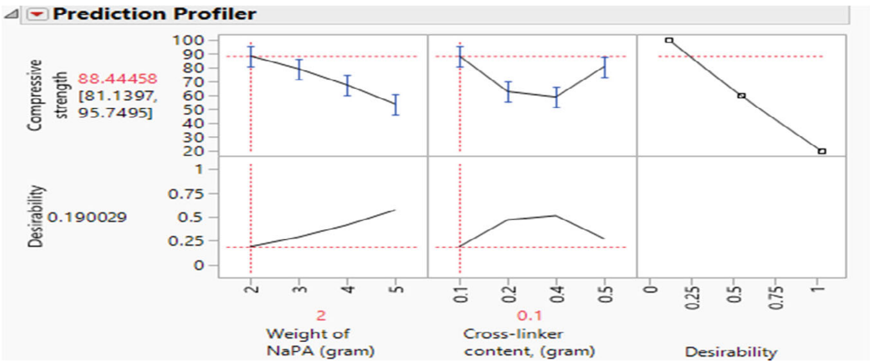



**Fig. 4** Prediction profiler for compressive strength (MPa) for absorboost tile

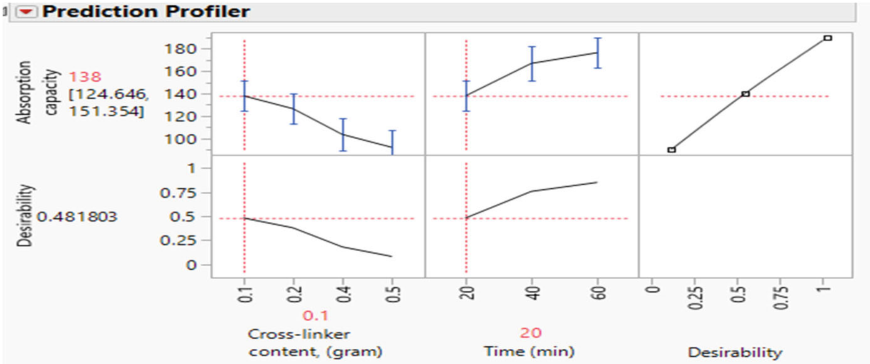



**Fig. 5** Prediction profiler for absorption capacity (g/g) for absorboost tile

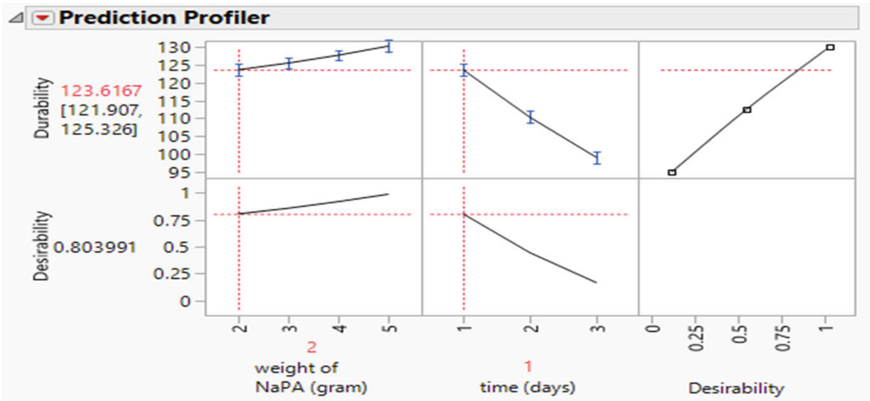



**Fig. 6** Prediction profiler for durability for absorboost tile

## 5.5 *Prediction Variance Surface (3D)*

The objective of this optimization is to find the set of values of the cross-linker content, (wt.), weight of NaPA (gram), and swelling time (min, days) leading to the desired responses, which is the optimum efficiency of compressive strength, absorption capacity, and durability for absorboost tile. The use of prediction variance surface (3D) is considered to be an easy technique for the analysis and optimization of the treatment ratio and even for identification between the studied parameters [35]. The results illustrate the uses of the response surface methodology (RSM) it possible to analyze and optimize correctly the influence of the two parameters such as weight of NaPA (gram), and cross-linker content, (wt.) for the efficiency of the compressive strength, for two parameters such as cross-linker content, (wt.), and swelling time (min) for the efficiency of the (days). The optimum values of these parameters give a minimum and maximum efficiency of 19%, and 52.38%. The reliability of the second-order model is to be considered based on multiple regressions and design experiments by the method of analysis (ANOVA). The results illustrate the models are highly significant and adequate with the experimental results. Figure 7a illustrates the prediction variance surface (3D) for compressive strength (Mpa), the results of the independent variables of the weight of NaPA and cross-linker content wt. are 3.5 g and 0.3 g, respectively. The efficiency and performance of compressive strength (Mpa) have a positive relationship with the concentration and weight of NaPA and cross-linker content wt. Figure 7b illustrates the prediction variance surface (3D) for absorption capacity (g/g), the results of the independent variables of the cross-linker content wt. and time are 0.3 g and 40 min, respectively. The efficiency and performance of absorption capacity (g/g) have a positive relationship with the concentration of cross-linker content wt. and time. Figure 8 illustrates the prediction variance surface (3D) for durability, the results of the independent variables of the weight of NaPA and time are 3.5 g and 2 days, respectively. The efficiency and performance of durability have a positive relationship with the concentration and weight of NaPA and time.

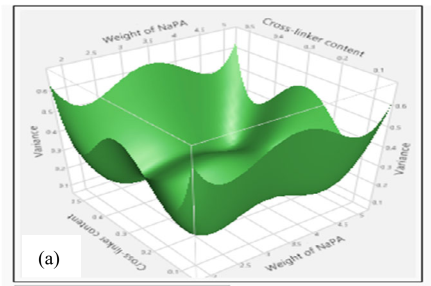
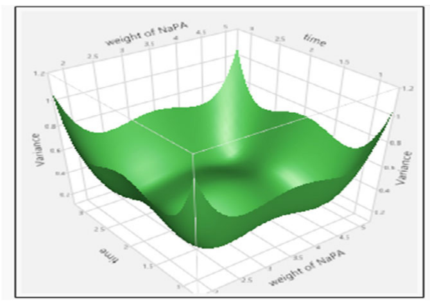



**Fig. 7** Prediction variance surface (3D) for (a) compressive strength (Mpa), and (b) absorption capacity (g/g) for absorboost tile

**Fig. 8** Prediction variance surface (3D) for durability for absorboost tile



## 6 Conclusion

Absorboost tiles have great potential for use in agriculture in semi-arid and arid areas. Absorboost tiles for agriculture in semi-arid and arid areas have high compressive stress resistance, high absorption capacity in hard water conditions, high durability and stability in swelling environments and during storage, low price, and gradual biodegradability without forming toxic species. This study, on the one hand, evaluates the absorboost tiles for agriculture in semi-arid and arid areas through designed and manufactured processes, and mechanical, agronomic, and chemical experimental analysis. On the other hand, by modeling, analyzing, and optimizing the compressive strength, absorption capacity, and durability of absorboost tiles, by studying the influence of the five-variable cross-content (by weight by the gram) of NaPA, the four variable weights of NaPA (by the gram), and different periods in minutes and days. Modeling and optimizing the absorboost tile of the compressive strength, absorption capacity, and durability using an experimental design by the response surface methodology (RSM). The second order of the model was tested by the analytical method (ANOVA). This analysis illustrates the model is significantly higher and more suitable for the experimental results. The prediction profiler indicates that the absorboost tile is optimal compressive strength the optimized parameter values are the weight of NaPA = 2 g, and cross-linker content = 0.1 g, the optimal absorption capacity for the optimized parameter values is cross-linker content = 0.1 g and swelling time = 20 min, and the optimal of durability for the optimized parameter values are weight of NaPA = 2 g and swelling time one day. These optimal values of the parameters lead to a minimum and maximum of 19%, 48.18%, and 80.39% respectively, for compressive strength, absorption capacity, and durability. The results illustrate the absorboost tile for agriculture in semi-arid and arid areas is more effective.

## References

1. Byerlee, D., Garcia, A.F., Giertz, A.: Growing Africa: Unlocking the Potential of Agribusiness. The World Bank, Washington. (2013).
2. Gerald Francis, J.: Towards Improved Application of Superabsorbent Polymers in Agriculture and Hydrology: A Cross-Disciplinary Approach. Auburn University (2011). Retrieved from http://hdl.handle.net/10415/2851
3. Yazdani, F., Allahadadi, I., Abas Akbari, G.: Impact of superabsorbent polymer on yield and growth analysis of soybean (Gly Gine max L.) under drought stress condition. Pakistan J. Bio. Sci., 4190–4196 (2008). https://doi.org/10.3923/pjbs.2007.4190.4196
4. Mortimore, M.: Dryland Opportunities. IUCN, Gland, Switzerland, IIED, London, UK and UNDP, New York, USA. (2009).
5. Kaith, B.S., Singh, A., Sharma, A.K., Sud, D.: Hydrogels: Synthesis, classification, properties and potential applications—A brief review. J Polym Environ **29**, 3827–3841 (2021). https://doi.org/10.1007/s10924-021-02184-5

6. Seo, Y.H., Lee, J.M., Park, S.Y., Kim, M.H., Kim, S.B., Oh, T.H.: PVA hydrogels supplemented with PLA mesh for tissue regeneration scaffold. Gels MDPI **10**(6), 364. https://doi.org/10.3390/gels10060364. (2024).
7. Data Mining and Predictive Modeling Prediction.: Retrieved from JMP statistical discovery: https://www.jmp.com/. (2024).
8. Damiri, F., Salave, S., Vitore, J., et al.: Properties and valuable applications of superabsorbent polymers: a comprehensive review. Polym. Bull. **81**, 6671–6701 (2024). https://doi.org/10.1007/s00289-023-05068-x
9. Malik, S., Chaudhary, K., Malik, A., Punia, H., Sewhag, M., Berkesia, N., Nagora, M., Kalia, S., Malik, K., Kumar, D., et al.: Superabsorbent polymers as a soil amendment for increasing agriculture production with reducing water losses under water stress condition. Polymers **15**(161) (2023). https://doi.org/10.3390/polym15010161
10. Masoumeh Kordi, Naser Farrokhi, Martin I. Pech-Canul, Asadollah Ahmadikhah.: Rice Husk at a Glance: From Agro-Industrial to Modern Applications[J]. Rice Science, **31**(1), 14–32 (2024). http://www.ricesci.org/EN/10.1016/j.rsci.2023.08.005
11. Supare, K., Mahanwar, P.A.: Starch-derived superabsorbent polymers in agriculture applications: an overview. Polym. Bull. **79**, 5795–5824 (2022). https://doi.org/10.1007/s00289-021-03842-3
12. Hoffman, J.D.: Theoretical aspects of polymer crystallization with chain folds: bulk polymers. Polym. Eng. Sci. **4**(4), 315–362 (1964)
13. Shahabazuddin, M., Banuvalli, B.K., Mulik, N., et al.: Comparative studies of the influence of particle size on various pretreatments of rice husk by assessment of chemical and structural components and wastewater characteristics of liquid fraction. Biomass Conv. Bioref. **13**, 5243–5252 (2023). https://doi.org/10.1007/s13399-021-01565-z
14. Viana, Q.S., Eugênio, T.M.C., Sabino, T.P.F., et al.: Physical, mechanical, and thermal properties of concrete roof tiles produced with vermiculite. Environ. Sci. Pollut. Res. **29**, 48964–48974 (2022). https://doi.org/10.1007/s11356-022-19337-1
15. Osman, K.T.: Physical Properties of Soil. In: Soils. Springer, Dordrecht (2013). https://doi.org/10.1007/978-94-007-5663-2_5
16. Osman, K.T.: Chemical Properties of Soil. In: Soils. Springer, Dordrecht (2013). https://doi.org/10.1007/978-94-007-5663-2_8
17. Chaudhari, S K., Somawanshi, R.B.: Modeling pF curve of clay, clay loam and silt loam soils under different quality waters. Agropedology. **10**, 183–194 (2000)
18. Anggraeni, S., Putri, S.R., Nandiyanto, A.B.D., Bilad, M.R.: Effect of particle size and tapioca starch content on performance of the rice husk and red bean skin briquettes. J. Eng. Sci. Technol. **16**(2), 1733–1745 (2021). https://www.scopus.com/inward/record.uri?eid=2-s2
19. Ngugen, N.: Native Tapioca Starch: Structure, Properties, and Applications. Retrieved from NGUYEN STRACH: https://nguyenstarch.com/native-tapioca-starch. (2024).
20. Seetapong, N., Mankaket, S., Rahem, S., Chanlert, P., Chulok, S.: Exploring binder efficacy in the fabrication of charcoal briquettes from Palmyra palm and oil palm shells: A comprehensive analysis. Bio Res. **19**(3), 5047–5057 (2024). https://doi.org/10.15376/biores.19.3.5047-5057
21. Adam, S.N.F.S., Zainuddin, F., Morgan, N.Z.S., Saroni, H.H.: Comparison of corn and tapioca starch binders on the characteristic of rice straw charcoal briquettes. In Emerging Technologies for Future Sustainability, Green Energy, 59–69 (2023). Retrieved from https://doi.org/10.1007/978-981-99-1695-5_5
22. Obi, O.F., Pecenka, R., Clifford, M.J: A review of biomass briquette binders and quality parameters. MDPI **15**(7), 2426 (2022). Retrieved from https://doi.org/10.3390/en15072426
23. Lubwama, M., Yiga, V.A., Lubwama, H.N.: Effects and interactions of the agricultural waste residues and binder type on physical properties and calorific values of carbonized briquettes. Biomass Conversion and Biorefinery **12**, 4979–4999 (2022). https://doi.org/10.1007/s13399-020-01001-8
24. Verma, K., Sarkar, C., Saha, S.: Biodegradable polymers for agriculture. In Sarkar, C., Saha, S. (eds.) Biodegradable Polymers and Their Emerging Applications (pp. 191–212) (2023). https://link.springer.com/chapter/10.1007/978-981-99-3307-5_9

25. Venkatachalam, D., Kaliappa. S.: Superabsorbent polymers: a state-of-art review on their classification, synthesis, physicochemical properties, and applications. Rev. Chem. Eng. **39**(1) (2021). https://doi.org/10.1515/revce-2020-0102
26. Ram Prasad, Manoj Kumar, and Vivek Kumar.: Nanoengineering Superabsorbent Materials: Agricultural Applications. In Peyravi, M., Selakjani, P.P. Khalili, S. (eds) (2017). https://link.springer.com/chapter/https://doi.org/10.1007/978-981-10-4573-8_6
27. Ostrand, M.S., DeSutter, T.M., Daigh, A.L.M., Limb, R.F., Steele, D.D.: Superabsorbent polymer characteristics, properties, and applications. Agrosystems, Geosci. Environ. **3**(1) (2020). https://doi.org/10.1002/agg2.20074
28. Understanding Material Strength, Ductility and Toughness.: Retrieved from The Efficient Engineer https://efficientengineer.com/material-strength-ductility-toughness/. (2024).
29. The Sequential Nature of Classical Design of Experiments, Prism, prismtc.co.uk. (2023).
30. Khoury, R., Wilhelm Harder, D.: Numerical Methods and Modeling for Engineering. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-21176-3
31. Keith M. Bower.: What is design of experiments (DOE). Retrieved from Learn about Quality https://asq.org/(2024)
32. Optimization Functions: Retrieved from JMP statistical discovery: https://www.jmp.com/ (2024).
33. Multiple Regression Residual Analysis and Outliers: Retrieved from JMP statistical discovery: https://www.jmp.com/ (2024)
34. What is Analysis of Variance (ANOVA): Retrieved from Investopedia: https://www.investopedia.com/ (2024)
35. Prediction Profiler and Prediction Variance Surface (3D): Retrieved from JMP statistical discovery: https://www.jmp.com/ (2024).

# AI & ML in Environmental Monitoring and Sustainability II

# A GIS-Based Decision Support System for Renewable Energy Investments in Egypt: A Machine Learning Approach

**Omar S. Salem, Moustafa A. Baraka, and Ahmed M. Abdel Sattar**

**Abstract** Predicting renewable power potentials is vital for generating clean sources of power and reducing the dependency on fossil fuels resulting in fewer greenhouse gas emissions. This paper illustrates a comparative analysis between two machine learning (ML) algorithms, the artificial neural networks (ANN) and the gene expression programming (GEP) models, in estimating the geospatial renewable wind and solar power potentials. The datasets, from which the ML models should learn, represent different locations obtained from Egypt's map. Error Analysis including the mean square error (MSE) and the coefficient of determination was utilized to specify the most significant geospatial generated model regarding accuracy. The results indicate that the ANN has shown better results against the GEP by showing fewer errors. Furthermore, sensitivity and uncertainty analysis were implemented in the ANN model providing further accuracy measurements. The predictive key results were used as a main power component of an underground water pumping system used for irrigation to demonstrate the aspect and the replicability of the study. A geographic information system (GIS) management algorithm drives the pumping system by optimizing the performance of the underground water system to function sufficiently using the estimated renewable power potentials.

**Keywords** Machine Learning · Geospatial · Renewable Power

O. S. Salem (✉) · M. A. Baraka
German University in Cairo (GUC), New Cairo, Egypt
e-mail: omar.salem@guc.edu.eg

M. A. Baraka
e-mail: moustafa.baraka@guc.edu.eg

A. M. A. Sattar
Cairo University, Giza, Egypt
e-mail: ahmoudy77@yahoo.com



V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_10

## 1 Introduction

As one of the priorities in countries' development, the construction of renewable power infrastructure is one of the solutions to face power shortage [1]. Egypt requires large power demand with a growth rate that exceeds 6% considering the reduction of ordinary power sources such as natural gas [2]. Egypt's geographical location is noteworthy as it is located in the Middle East and North Africa (MENA) region [3]. Egypt's large area is subjected to vast amounts of wind and solar potentials, placing Egypt with one of the highest solar potentials not only in the Middle East, that can be converted to renewable sustainable power sources covering the need for power [4]. Research forecasting renewable power generation suggests that solar power generation has a considerably greater potential compared to wind considering the political motivations that tend to encourage large-scale placement in many countries [5]. On the other hand, wind power has become extremely prevalent because it is highly effective, economical, and valuable for the environment. Current surveys showed that the yearly growth of installed wind power capacity is around 30% [6]. Egypt shows high primary energy use by about 11% of Africa [7].

Considering the irregular characteristics of renewable power sources, machine learning (ML) can predict wind and solar power potentials nowadays. Hence, it is vital to use ML methods that can learn from the historical provided data and generate models predicting both wind and solar power potentials. With the mounting of artificial intelligence (AI) in recent years, machine learning (ML) approaches that can solve nonlinear problems have been applied in the field of renewable power. Renewable power prediction models rely on the availability of the input data that influence the presence of the output power. The ML learning data can be extracted from empirical formulas or collected from real data [8]. The prediction models of the renewable power potentials are divided into two major groups the wind and solar power potentials. For each power potential, the examined ML methods are the artificial neural networks (ANN) and the gene expression programming (GEP). ANN and GEP work by finding relativity between the provided inputs and the outputs representing the relation in a predictive model as a result.

Different researchers suggest that several variables influence wind power generation [9, 10], using the ANN back-propagation method, proposed that the wind speed distribution factor, scale factor, and wind speed are input factors that highly affect the wind power potentials [11] proposes that ambient temperature, average panel temperature, average transformer temperature, average solar irradiance, and wind speed are the main inputs that influence the photovoltaic (PV) solar power potential generation using the ANN algorithm. However, [12] used daily air temperature, relative humidity, atmospheric pressure, wind speed, and earth temperature as parameters for the GEP model to learn from and generate PV power. The aforementioned researches, considering ANN and GEP, show accurate results when comparing the real obtained data to the ML-generated models' results. According to mentioned studies, the ML model's accuracy is evaluated using root mean square error (RMSE), coefficient of determination (R2), mean absolute error (MAE), and mean absolute deviation (MAD).

Choosing a suitable location to install either wind or solar farms requires many strategies for decision-makers however using GIS and ML facilitates the process [13]. Using ML to plan renewable power plants allows users to examine big data, predict the results, and simplify the process [14]. Underground water pumping systems are applications for AI models to be supplied by a renewable power source. [15] introduces the renewable water pumping system components which include PV solar panels, a submersible water pump, a battery bank, and a water tank. [15] suggests that to control the working flow of the system, a management algorithm should be introduced to guarantee that the water pump is sufficiently covered by the renewable power source, the battery in case of insufficient solar power, or by a combination of solar power and the battery. Wind and solar power are economically friendly and are sources of clean power that create a pollution-free environment. Renewable forms of energy can be used in a water pumping system as a source of power. [15] introduces an algorithm to manage power applied to the renewable water pumping system. The conversion of solar irradiation to electric power by photovoltaic panels is the most efficient option compared to traditional power produced from carbon emission methods to be used in a water pumping system [16]. According to [17], generated power from wind resources is a green and unrestricted form of power that became the initial choice of many countries all over the world and can be used to power up water pumping systems.

Egypt has high wind speed potential and is subjected to high solar radiation, making it one of the highest countries subjected to solar radiation. The two aforementioned sources can be converted to wind power and solar power which can cover Egypt's increasing need for power. Several studies have addressed the prediction of wind and solar power without introducing Egypt to its vastly natural available potential. Various studies that introduce predictive models of renewable power examine error analysis methods; however, few studies include uncertainty and sensitivity analysis in addition to performing error analysis. Moreover, a diverse of studies that considered the underground water pumping system examine adding a renewable power source to the system; therefore, few studies introduced using machine learning renewable power predictive models to their water pumping systems.

Although a diversity of ML models have been developed lately for wind or solar power predictions using the ANN or GEP algorithms, the mainstream ML models introduce forecasting power potentials without introducing further analyses. Moreover, it is essential to know that the ML model's input learning data could be uncertain or could be sensitive to the outputs introduced to the models. Therefore, it is crucial to introduce uncertainty and sensitivity analyses to check additional accuracy performance. This paper discusses the generation of ML models to predict the wind and solar power potentials using ANN and GEP methods and compares them regarding accuracy in addition to performing further analyses of the ML models including error analysis, uncertainty analysis, and sensitivity analysis. Additionally, the most performing ML predictive model is implemented as a power source for an underground water pumping system to be used for irrigation. The remainder of this paper was ordered as follows; Sect. 2 provides the materials and methods vital to generate ML models that function as power for an underground water pumping system. Section 3 presents the results of the AI models generation and the AI models' analysis, the water pumping system algorithm workflow, and case study results. Finally, Sect. 4 draws the conclusions.

## 2 Materials and Methods

### *2.1 Machine Learning Variables*

Machine learning algorithms learn from the inputs and outputs of data. The inputs and outputs can be either collected actual data or extracted data from numerical or theoretical expressions regarding both wind and solar power. The following sections discuss the numerical expression for calculating the renewable power potentials for wind and solar.

**Wind Power Variables.** It is difficult to determine the theoretical potential at every grid cell, so a measure of wind energy intensity is calculated. The mass of the air and its speed create the wind's kinetic energy. The kinetic energy of the wind is known as the wind power density [18].

The geographic potential is to remove all the restricted areas from the base map of Egypt, which are not applicable for adding wind farms. This stage requires defining the geographic constraint areas as individual layers on the GIS maps. Excluding those areas from the base map forms a single new exclusion map that is divided into grids of 1 km × 1 km. Each grid only includes the remaining areas that are suitable for wind turbine installation. Regarding onshore wind resources, the restricted excluded areas include the following [18]: urban built-up areas, water bodies, protected areas, highly elevated areas with elevations exceeding 2000 m, sloped areas with slopes more than 25°, roadways network, and areas with wind speeds lower than the lower limit of the wind turbine.

Wind speed distribution is another factor, which greatly affects the wind energy output. The Weibull function is applied to calculate this distribution. The Weibull function is a probability function applied to analyze the wind energy yield taking into consideration the wind probability of occurrence. The Weibull distribution is expressed with the following probability density function [18].

$$f(U) = \frac{k}{A}\left(\frac{U}{A}\right)^{k-1} e - \left(\frac{U}{A}\right)^{k} \tag{1}$$

$$k = \left(\frac{\sigma_U}{U_{mean}}\right)^{-1.086} \tag{2}$$

$$A = \frac{U_{mean}\, x\, k^{2.6674}}{0.184 + 0.816\, x\, k^{2.73855}} \tag{3}$$

where: U is the wind speed interval, known also as a class, with incrementing the value of 0.5 m/s from 0 to 25 m/s, k is the shape factor, A is the scale factor, $\sigma_U$ is the standard deviation and $U_{mean}$ is the mean wind speed in m/s.

The wind power curve is one factor of the most two important factors that calculate the wind turbine output. The power curve is characterized by three unique wind speeds: wind speeds titled cut-in, rated, and cut-out speeds. The wind turbine output is calculated by the following mathematical expression [18]:

$$E = \mu * T \sum_{wind\ class} P(U) f(U) \tag{4}$$

where: μ is the availability factor that accounts for turbine outages, $T$ is the hours of a year, P(U) is the power curve, f(U) is the Weibull function, and U is the wind class from 0 to 25 m/s.

The turbine's output is estimated by applying expression (4) over each point of the fishnet by linking the power curve and the Weibull probability function. This function is carried out by the submission of multiplying P(U) and f(U) where the U increases with 0.5 m/s interval until reaching the maximum wind speed of the wind class range. Certain classes may contain zero values due to the unavailability of wind speed data in specific locations. This function is applied at each grid cell on the GIS software extracting the wind turbine output of each grid.

**Solar Power Variables.** Solar irradiance is the most influencing factor that affects the amount of solar power potential that PV cells produce. Egypt is subjected to large amounts of solar irradiation that can be translated to power potential as a pure source of renewable energy. The map of Egypt is divided into grids of 1 km × 1 km using the ArcMap software using fishnet property [19]. The solar power output is expressed through the following equation [20].

$$P_{pv} = A.\, I_{pv}.\, V_{pv} \tag{5}$$

$$I_{pv} = I_r + \alpha(T_c - T_{cr})\frac{SR}{SR_r} + \left(\frac{SR}{SR_r} - 1\right) I_{sc} \tag{6}$$

$$V_{pv} = -\beta(T_c - T_{cr}) - R_s \Delta I + V_r \tag{7}$$

$$\Delta I = \alpha(T_c - T_{cr})\frac{SR}{SR_r} + \left(\frac{SR}{SR_r} - 1\right) I_{sc} \tag{8}$$

$$T_c = T + (NOCT - T_{cr})\frac{SR}{SR_r} \tag{9}$$

where: ($P_{pv}$) is the technical potential of the cell, ($I_{pv}$) is the photovoltaic panel current, (A) is the inclusion area within the same cell, and ($V_{pv}$) is the photovoltaic panel voltage, $V_{pv}$ and $I_{pv}$ are the photovoltaic panel voltage and current. $V_r$ is the reference voltage. $I_r$ is the reference current. SR is the solar irradiation. Tc is the cell's ambient temperature. $I_{sc}$ is the short circuit current. DI is the current variation. Ns is the number of panels in a series. NOCT is the nominal operating cell temperature that is associated with certain operating conditions. $P_{pv}$ is the photovoltaic power. T defines the ambient temperature. $\alpha$, $\beta$ are the short circuit current and the open circuit voltage ambient temperature coefficient. The $SR_r$ is1000W/m2 and $T_{cr}$ is 25 C at the Standard Test Conditions [20].

According to [20], the values of the ambient temperature of the map layer from the Global Solar Atlas are extracted over the grid cells fishnet that already have the values of global horizontal irradiation (GHI) to get the cell temperature ($Tc$). Finally, expression (5) is carried over all the remaining grid points for the available cells, and the solar power output is extracted from the ArcMap software to an Excel sheet.

## *2.2 Machine Learning Methods*

ML methods work to provide models that represent the inserted input data, which is derived, either from theoretical equations or from real-measured data. The ML methods learn from the provided data by trying to find a relation between the provided inputs and outputs to best represent the data. The generated models can be implemented for various needs among them is the prediction models. The two used ML methods are ANN and GEP. The objective is to reach the most accurate model by comparing and analyzing these models for the prediction of renewable power potentials in Egypt.

**Artificial Neural Networks.** ANN is a machine-learning algorithm that generates models that are influenced by the inputs, the output, and the neurons. The human brain approximately consists of 15 billion neurons; however, the number of neurons in the ANN model depends on the real need of the problem [21]. A feedforward neural network was used to forecast both wind and solar power. The structure of the feedforward neural network consists of an input layer, an output layer, and a hidden layer. The number of neurons in the hidden layer is chosen according to the performance of the model [22]. Therefore, neurons will be defined by trial and error method until reaching a proper number of RMSE [23]. The ANN is trained using the Levenberg–Marquardt algorithm which is a standard training algorithm [24].

As shown in .

Figure 1, the input layer gets data and gives it to neurons in the hidden layer and then passes the data to the output layers by multiplying weights [21]. From both the wind and solar power theoretical expressions mentioned earlier, the independent and dependent variables of the functions are used as the inputs and output parameters of the ANN algorithm.

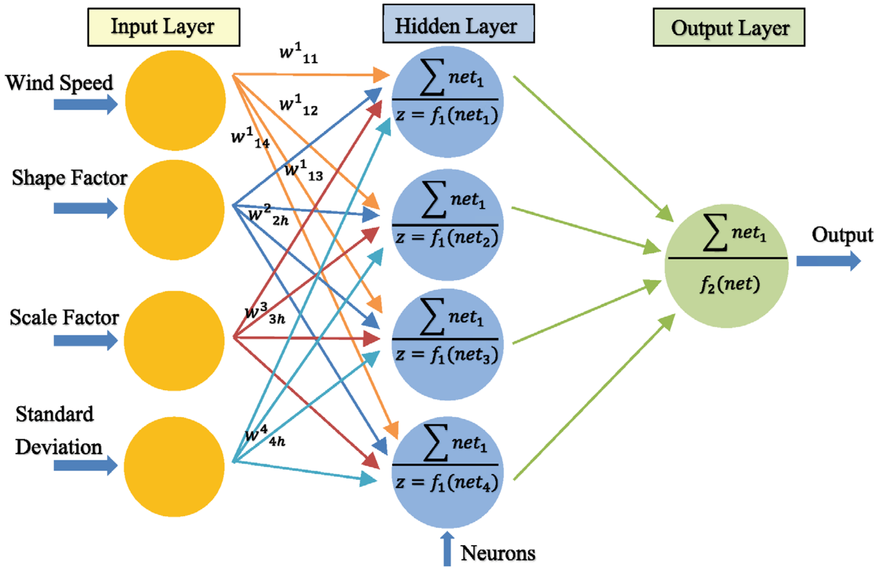



**Fig. 1** Artificial Neural Networks Algorithm

**Gene Expression Programming.** GEP consists of a structure called a chromosome which is a simple linear structure of fixed length. A chromosome can consist of one gene or more. To show the expression embedded in the genes, the GEP uses Expression Tree (ET) representations [25]. Hence, any mathematical function is converted to an ET expression. Consider the following algebraic expression also known as a program.

$$\frac{a\,x\,b}{\frac{c}{d}} \tag{10}$$

Expression (10) can be expressed as an ET representation as follows in Eq. (11).

$$\begin{matrix} / & x & / & a & b & c & d \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \tag{11}$$

Equation (11) shows the head of the gene which contains both the terminals as well as the function of the algebraic expression. Regarding the GEP, any gene consists of a head and a tail. The tail is part of the gene through which the terminal can be modified in the next generation using any kind of genetic operation without having any constraints [26]. However, GEP usually contains chromosomes of more than one gene that forms a complex multigenic chromosome [25]. The multigenic chromosome can be used to express more complex relations [26]. GEP keeps randomizing the terminals and the functions of the program, regarding the initial chromosome gene. The goal of the GEP is to reach an offspring, which can represent or best fit the input dataset. To reach this goal, the GEP algorithm uses the RMSE function to check accuracy. The RMSE ranges from zero to infinity. An acceptable range of the fitness test lies between 600 to 800 [25]. The GEP initiates by defining the number of variables to be tested. Then, the GEP architecture is defined by defining the number of chromosomes, the number of genes, the head size, and the linking function. Moreover, the GEP starts to develop the parent program that is checked by the fitness function. Furthermore, the GEP started to implement the second generation by randomizing the previously created genes to new ones. The evolution is repeated with more offspring evolutions until reaching an adequate GEP model.

## 2.3 Machine Learning Analysis

**Error Analysis.** The error analysis is the variation of the model's output about a mean. Error analysis indicates the error in model output says how far the forecasted value from the true value is to be [25]. The error analysis is a measure of how the model responds. The error analysis can be evaluated by more than one method. The number of neurons started by two and incremented by one neuron each time. The process of increasing the number of neurons stops when reaching the least MSE for that model. The least MSE is achieved when the MSE increases again after incrementing the number of neurons by one. The MSE is calculated using the following expression:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(X_p - X_o\right)^2 \tag{12}$$

where: $X_P$ is the predicted value of the ANN model, $X_O$ is the observed value of the ANN model, and n is the total number of samples. In addition to the MSE, the R-value after each trial is computed through the following expression [27]:

$$R = \frac{\sum_{i=1}^{n}\left(X_p - X_{p.mean}\right)(X_o - X_{o.mean})}{\sqrt{\sum_{i=1}^{n}\left(X_p - X_{p.mean}\right)^2}\sqrt{\sum_{i=1}^{n}(X_o - X_{o.mean})^2}} \tag{13}$$

where: $X_P$ is the predicted value of the ANN model, $X_{P.mean}$ is the mean of the predicted values of the ANN model, $X_O$ is the observed value of the ANN model, $X_{O.mean}$ is the mean of the observed values of the ANN model. After identifying the best trial for each combination of inputs, other error analyses are carried out to provide more validation for the models. The error analyses performed are as follows, calculating MAE, RMSE, and $R^2$ [25, 28].

$$MAE = \frac{\sum_{i=1}^{n}(X_P - X_o)}{N} \tag{14}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (X_P - X_O)^2}{N}} \tag{15}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(X_o - X_p\right)^2}{N \sum_{i=1}^{n} (X_o - X_{o.mean})^2} \tag{16}$$

where: $X_P$ is the predicted value of the ANN model, $X_O$ is the observed value of the ANN model, $X_{O.mean}$ is the mean of the observed values of the ANN model, and N is the total number of samples.

**Uncertainty Analysis.** The uncertainties of the input parameters have to be studied as they could influence the outcome of the AI models. Although using in this study a large data set, the uncertainty of prediction outcomes can occur, which can be decreased by gathering more data about the input parameters [25]. Upon knowing the architecture of the AI models, the number of neurons, and training the models using either ANN or GEP algorithms, the Monte-Carlo simulation (MCS) is used to examine the uncertainties in the final chosen AI model for forecasting wind and solar powers. MCS consists of the generation of the repeated random parameters of the inputs from their created probability density functions (PDFs) such as (e.g. Gaussian, log-normal, Weibull, etc.) and then computing the statistics of the resulting output. PDFs are generated for each input parameter for the chosen AI models. Each PDF is chosen based on the most fitted probability graph for the data of the inputs using the Kolmogorov–Smirnov test, which relates the probability distribution of input variables with a known probability distribution with a 5% significance level [29].

First, the frequency graph distribution of each input set of data is calculated. Then, the best-fitting PDF is chosen based on the generated frequency graph using the RISK software. From each PDF for each input parameter of the AI models, a certain number of random samples is extracted that is 250,000 samples. The 250,000 random samples are divided into 1000 samples resulting in 250 vectors. The first run of the AI model used the first vector of 1000 samples as the input. Then, for the upcoming run, the next vector was combined with the previous vectors cumulatively and the cumulative new vector was used as the next input for the new run of the AI model. Regarding each run, the variance was computed for the outputs using the following expression:

$$S^2 = \frac{\sum (x_i - \overline{x})^2}{n - 1} \tag{17}$$

where: $x_i$ is the output value, $\overline{x}$ is the average of the output values, and $n$ is the total number of samples. The number of variance is plotted versus the number of output samples and the convergence of the variance is extracted from the plotted graph [25]. The number of samples at which the convergence of variance occurred is the number to carry out the uncertainty analysis using the MCS. The reached number of samples of the input data sets is applied in the AI models again. The output of the AI models is used to conduct the uncertainty analysis using the following expressions [25]:

$$uncertainty = 100x \frac{MAD}{Median(X)} \tag{18}$$

$$MAD = \frac{1}{N} x \sum_{i=1}^{N} |X_i - Median(X)| \tag{19}$$

where: $X_i$ is the output of the AI models and $MAD$ is the mean absolute deviation.

**Sensitivity Analysis.** Sensitivity analysis is a measure of the influence of each input in the model on the output. The result of the sensitivity analysis is to give a rank from the most effective input to the least effective one. The sensitivity analysis is a measure of the relative influence of each input variable of the chosen ANN model. The sensitivity analysis is measured using Garson's equation which is derived from the neurons' weight matrices (IW1 and IW2). Garson's equation is shown as follows [30]:

$$I_{ik} = \frac{\sum_{j=1}^{L}\left(\left(\frac{|IW1_{ij}|}{\sum_{i=1}^{N}|IW1_{ij}|}\right)x|IW2_{jk}|\right)}{\sum_{i=1}^{N}\left(\sum_{j=1}^{L}\left(\left(\frac{|IW1_{ij}|}{\sum_{i=1}^{N}|IW1_{ij}|}\right)x|IW2_{jk}|\right)\right)} \tag{20}$$

where: $I_{ik}$ is the relative importance factor of the ith input neuron on the kth output neuron, N and L are the numbers of input and hidden neurons, respectively, $IW1_{ij}$ is the connection weight matrix between the ith input neuron and the jth hidden neuron, and $IW2_{jk}$ is the connection weight matrix between the jth hidden neuron and the kth output neuron.

**Water Pumping System.** The system works by finding a balance between the produced energy and the required energy. The water pumping system consists of either photovoltaic panels or wind turbines, a battery, a pump, a storage tank, and a multi-string inverter [20]. The system works by pumping the underground water to be used for irrigation purposes. The Siemens-Gamesa SG 2.1–114 wind turbine is examined for Egypt with heights of 80 and 100 m. The PV panels examined are the SOLARWORD-SW-250 Poly model. The management algorithm aims to use either the daily solar power or the daily wind power to run the water pump. To do so, first, the pump load required has to be calculated. The pump power is to be calculated by the following expression:

$$Pump\,Power = \frac{Q\,x\,H\,x\,g\,x\,\rho}{\eta} \tag{21}$$

where Q is the water discharge $m^3/day$, H water pump head, $g$ is the gravitational acceleration $9.81 m^2/s$, $\rho$ is the water density 1000 $kg/m^3$, and $\eta$ is the pump efficiency. The area of interest is Egypt's new delta project. The project is to be irrigated by underground water using water pumps that are to work by using wind or solar power as their source of power. A day in both the summer and the winter seasons is chosen to test the algorithm. After calculating the pump power using Eq. (21), the equivalent daily solar or wind power is calculated to check if the power source with be compatible or not. The inputs for the management algorithm are the power source ($P_{source}$) and the pump load ($P_L$).

## 3 Results and Discussion

### *3.1 Machine Learning Models*

**Artificial Neural Networks.** ANN algorithm uses several input combinations along with their outputs. The used inputs are attained from the theoretical methods equations as per [18]. The ANN chosen inputs are the wind speed, Weibull distribution shape factor, scale factor, and the annual wind speeds standard deviation resulting in four inputs with four different combinations for wind power potential. The previously mentioned studies encountered solar radiation and temperature as parts of their AI inputs; however, they did not encounter the effect of the PV panel's area and efficiency as mentioned by [20]. Therefore, the five inputs considered in this study for solar power prediction are the ambient temperature, PV cell temperature, area of installation, GHI, and PV panel efficiency. The efficiency is a function that is calculated from both the ambient temperature and the PV cell temperature, so they are not included in a single combination as that will form redundancy. The neurons transmit the data between the input layer the hidden layer, and the output layer. The number of neurons does not follow certain criteria for choosing them as using too many neurons can lead to overfitting and using too few neurons can lead to underfitting. They are defined by running several trials with different numbers of neurons as per [31]. The number of neurons achieving the lowest MSE value is the value showing the most accurate run as compared to [32]. This process is carried out for all the possible input combinations, mentioned earlier. The changing of neurons' number process resulted in 143 trials. According to [33], the smallest MSE values of all runs and R-values closer to 1 state higher the accuracy of the models. The least MSE values for the wind ANN models have the following input combinations: wind speed, shape factor, scale factor, and standard deviation. The results followed the previously mentioned studies including the wind speed as their main input parameters. The chosen ANN models for the wind and solar models are the models with 10 and 12 neurons, respectively. The corresponding MSE for the wind and solar ANN models are 0.0349 and 0.0427, respectively, which are similar to the MSE values of [34].

**Gene Expression Programming.** GeneXpro tools 5.0 software trial version was used to run the GEP model representing both wind power and solar power. The GEP architecture consists of the number of chromosomes, the number of genes, head size, the linking function, and the genetic operations [25, 35].

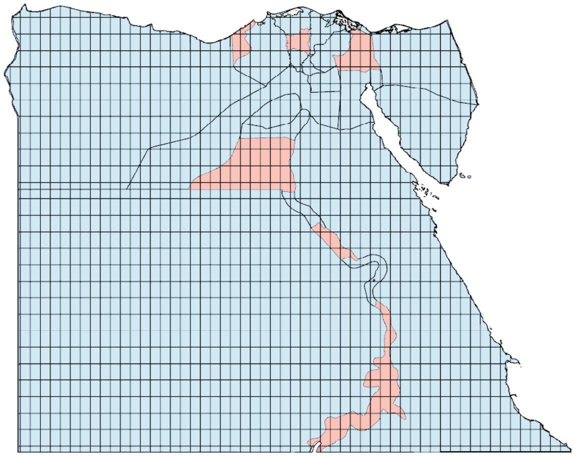


**Fig. 2.** 70% Testing in Blue and 30% Validation Sample in Red

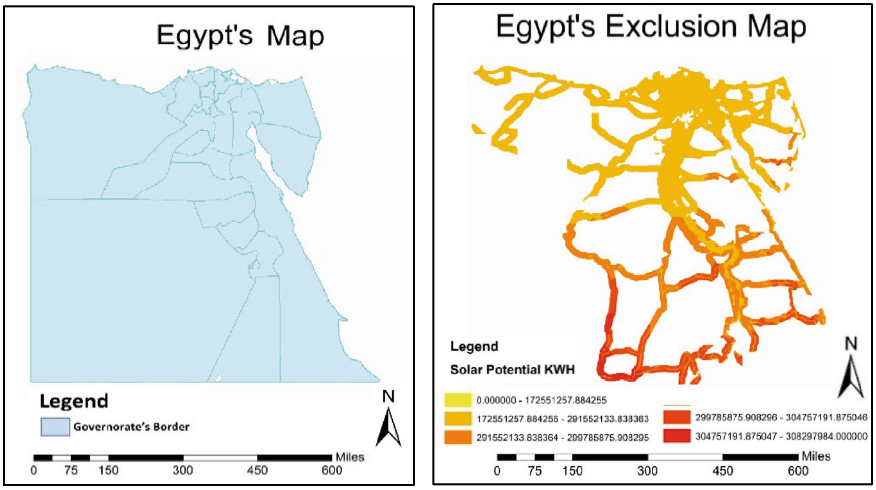
The GEP input dataset was divided into 70% training and 30% testing as shown in Fig. 2. GEP keeps randomizing the terminals and the operators along the same gene and between different genes. Every generation, the model is tested using the RMSE fitness test until reaching 20,000 generations, which is the stopping criteria [25]. The final GEP model for solar power had a fitness error of 640.693 and the wind GEP model had a fitness error of 919.538. The result of the solar power model is three expression trees. While the wind power model resulted in four expression trees. The expression trees are translated to the following equations:

$$Solar\ Power = \frac{3.608\, x\, Area\, x\, GHI^{\frac{3}{5}}}{T_c^{1/12}} \tag{22}$$

$$Wind\ Power = 2861.5257x\frac{WS^{1/6}x\sqrt{A}\, x\, K^{29/120}}{SD^{1/16}} \tag{23}$$

## 3.2 Machine Learning Models Output

The ML model output is a newly generated map presented in the ArcGIS software that shows places where the installation of both wind turbines and solar panels is applicable as well as practical. The newly generated map shown in Fig. 3 is the final exclusion map. The predictive model is generated based on the previously collected data that was provided for the ML algorithm to learn from and produce models representing the introduced data. The map can help decision-makers choose places that are ideal for the setting up of wind turbine or solar panel plants.




(a) (b)

**Fig. 3** (a) Egypt's map before exclusions (b) Egypt's exclusion renewable power potential map

## *3.3 Machine Learning Models Analysis*

To choose the best-performing ML models for both the wind and solar power potentials, several analyses are carried out to compare the models regarding accuracy. The performed analyses mainly compare the actual introduced data to the output results of the generated ML models. The error analysis is carried out between the outputs of the AI models, ANN and GEP. The $R^2$ and the RMSE are carried out for the final three chosen ANN models for the wind and solar as well as their GEP models. RMSE tells you how concentrated the data is around the best-fit line. RMSE is carried out for the ANN and GEP models for wind at 80 m, wind at 100 m, and solar. The models of the ANN and GEP are compared for each evaluated parameter using the error measurement tools, shown in Table 1. The $R^2$ values of the AI models clarify that the ANN has better results compared to the GEP model for all evaluated parameters, so the rest of the analysis will be carried out for the ANN final models. The $R^2$ values obtained in this study coincide with all the studies of [33, 36, 37], which are 0.97, 0.9944, and 0.999159, respectively.

The wide range of real data of the wind and solar used to construct the prediction ANN models inflicts uncertainties on model predictions. The uncertainty is

**Table 1** Error analysis results

| Parameter | Model | $R^2$ | RMSE | Model | $R^2$ | RMSE |
|---|---|---|---|---|---|---|
| Wind | ANN | 0.99841 | 0.21 | GEP | 0.99305 | 312.34 |
| Solar | ANN | 0.99997 | 0.22 | GEP | 0.99510 | 4.48 |

**Table 2** Uncertainty analysis results

| Parameter | Model | Median | MAD | Uncertainty % | Mean | SD | PDF |
|---|---|---|---|---|---|---|---|
| Wind 80 m | ANN | 11,956.9 | 944.71 | 7.9 | 11,825.5 | 1,448.59 | Weibull |
| Wind 100m | ANN | 12,254.8 | 1027.55 | 8.4 | 11,395.6 | 3,002.63 | Cauchy |
| Solar | ANN | 269.7 | 31.23 | 11.6 | 21,134.1 | 12,546.49 | Logistic |

performed based on the MCS analysis. Using the RISK software, for each input for the ANN models, the frequency bar charts are calculated. Using MCS, a probability distribution function is fitted for each input data. 250,000 Random samples from each PDF are extracted. For each random dataset of every input, the number of samples is divided into 250 vectors. The cumulative variances are extracted and plotted in a graph versing the number of samples. The plots formed showed a convergence of the variance after almost 100,000 samples of the inserted data. The result is the number of samples of the MCS for which the uncertainty analysis can be carried out. The results of performing Eqs. (18) and (19) for all the ANN models are summarized in Table 2. The uncertainty analysis results are acceptable compared to the values obtained by [25].

The sensitivity analysis measures how each of the input parameters influences the final output of the ANN model. Equation (20) is carried out for all the ANN models to determine the sensitivity analysis influence parameters. The influence of the input parameters are as follows ascendingly: wind speed, scale factor A, standard deviation, and shape factor K. On the other hand, the solar model parameters influence are as follows ascendingly: global horizontal irradiation, temperature, and area. Finally, the results show that the ANN models for both the wind and solar power potentials showed better results compared to the GEP models.

### *3.4 Water Pumping System Case Study*

The underground water pumping system that is used for irrigation is applied to a specific place in Giza, Egypt. The algorithm's criteria is to store excess renewable power in the battery, and if the battery is fully charged, the excess power is considered $P_{LOST}$. The inputs of the management algorithm are the $P_{SOURCE}$ and $P_L$. The outputs are P1, which is the power exchanged between the pump and the renewable energy source, P2 the exchanged power between the water pump and the battery, and $P_{LOST}$ which is the excess renewable power in case the battery is fully charged.

The management algorithm aims to quantify the amount of renewable power required to run water pumping systems to irrigate the New Delta project. To calculate the depth of the underground water table, the map of the underground water table was inserted into the ArcMap software and georeferenced to be placed in its original coordinates. The depth of the underground water, regarding the New Delta project, ranges from 150 to 200 m. According to [38], the static head is the difference between the depth of the water and the head of the receiving reservoir. The dynamic head can be estimated as 15% of the static head. The area of study consists of several irrigation circles, so only one circle of the irrigation area is evaluated. The examined circle has an underground water depth of 180 m. Two water pumps are going to work for 10 h a day which results in 158 $m^3$/hr. The suggested water pump is GRUNDFOS SP 10 inches. Knowing the pump head of 207 m (680 ft.), the corresponding Q results in Q of 66 gallons per minute (160 $m^3$/hr.). Substituting in Eq. (21) results in a required renewable power of 2148 kW/day. Two days are examined throughout the year, a day in summer and another in winter. Moreover, monthly data were also examined for both wind and solar models. Daily data for the wind speeds, global horizontal irradiation, and temperature were all provided online using the Photovoltaic Geographical Information System (PV GIS) through the European Commission website. Daily and monthly inputs, regarding the summer and winter days, were first calculated using a theoretical equation for both wind and solar power. On the other hand, the same data were used as inputs for the ANN of the solar model because the inputs for the daily solar power are the same compared to the inputs required for annual power calculations. However, the wind daily data is not inserted into the ANN model as the model represents the annual wind power output. Therefore, only monthly data can be inserted into the ANN model as the monthly wind power calculations follow the same steps as compared to the annual wind power[39].

Monthly wind speed distribution follows the Weibull probability density function, which is a reason to be inserted in the ANN model [39, 40]. The wind turbine used in that case study after following the methodological steps followed the same results as [41] regarding the wind turbine type, SG 2.1–114 turbine, and generated power. A comparison was made between theoretical equations and ANN models' outputs to validate that the ANN models can predict daily solar and wind powers [20]. As shown in Fig. 4, the output of both the theoretical equations and the ANN models demonstrates that the outputs of the theoretical equations and the ANN are nearly the same as the results [20]. A sample outcome of the algorithm for both summer seasons taking into consideration the daily solar power is shown in Fig. 5.

In Fig. 5, $P_L$ is always a constant straight line as it shows the water pump which is constant through the working 10 h. P1 increases as the power is exchanged between the pump and the source and is constant when the source power exceeds the required water pump load. On the other hand, P2 decreases when the power source is less than the pump load and the pump has to cover the remaining required load from the battery. However, P2 shows negative values when the source power exceeds the pump power and charges the battery until being fully charged. Finally, $P_{LOST}$ shows positive results when there is excess power after fully charging the battery. $P_{LOST}$ only shows positive results in the PV solar panels because decreasing the panels to an area of less than 2500 $m^2$ does not cover the required pump load. However, the wind turbine can serve 13 irrigation circles that when divided show small portions of $P_{LOST}$.

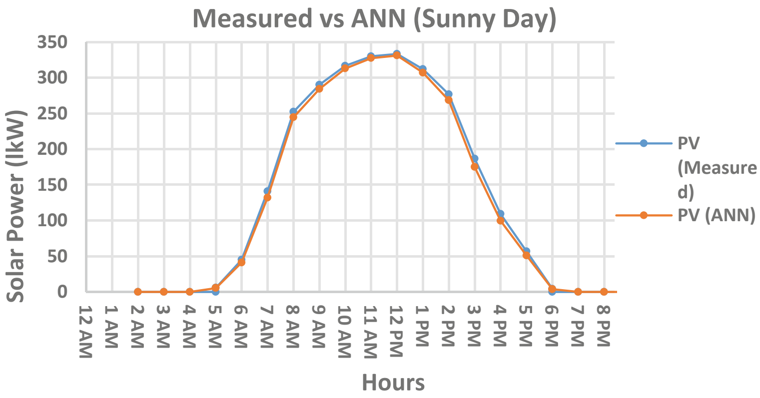



**Fig. 4** Solar power measured vs ann sunny day

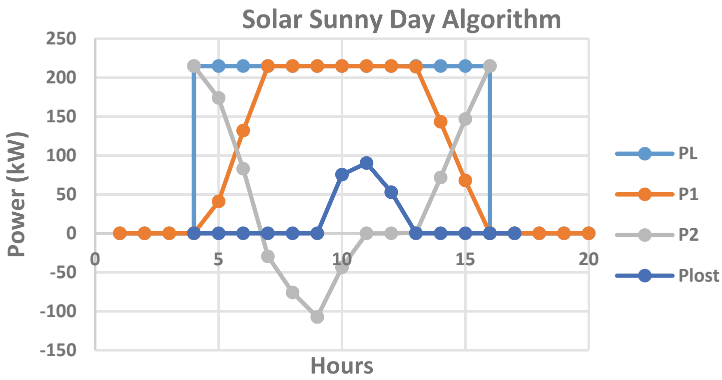



**Fig. 5** Solar power sunny day algorithm output

## 4 Conclusions

This study discusses the prediction of wind power at heights of 80 m and 100 m as well as solar power using ML methods including ANN and GEP. The ANN and GEP models for both wind and solar were generated and compared regarding accuracy. Moreover, uncertainty and sensitivity analyses were applied to the optimum ML model. The ML predictive renewable power model was used to power an underground water pumping system used to irrigate off-grid areas. From this study, it is concluded that the optimum ANN models regarding accuracy resulted from the 139 trials for wind and solar power generation the models with inputs of WS, A-factor, K-factor, and SD. The chosen ML models for the wind power at 80 m and 100 m above the ground have 10 neurons with corresponding MSE of 0.0349 and 0.0512, respectively.

On the other hand, the solar power predictive model is those having 12 neurons, with a corresponding MSE of 0.0427.

Regarding the GEP models, the used variables for wind power are WS, A-factor, K-factor, and SD. The used variables for solar power are GHI, cell temperature, and the area. The optimum GEP structure for both wind power and solar power was using 30 chromosomes, 3 genes, and 9 head sizes. After 20,000 generations of the GEP models, the optimum models for wind power and solar power showed fitness error values of 919.538 and 640.693, respectively.

ANN models of both wind power and solar power showed better accuracy when compared to the GEP models in terms of error analysis. The uncertainty analysis percentage using MAD for wind power at heights of 80 m and 100 m is 8% while the percentage for solar power is 11%. While the sensitivity analysis shows the influence of the input parameters on the output. ANN models for both wind power and solar power are used as the main power source for an underground water pumping system used for irrigation. The results of a management algorithm for the pumping systems stated that a pivot irrigation circle located in the case study area can be supplied with either PV panels with an area of 2500 m2 or one-tenth of the power of a wind turbine.

## References


1. Lei, Y., Lu, X., Shi, M., Wang, L., Lv, H., Chen, S., Hu, C., Yu, Q., da Silveira, S.D.H.: SWOT analysis for the development of photovoltaic solar power in Africa in comparison with China. Environ. Impact Assess. Rev. **77**, 122–127 (2019)
2. Shouman, E.R., Khattab, N.M.: Future economic of concentrating solar power (CSP) for electricity generation in Egypt. Renew. Sustain. Energy Rev. **41**, 1119–1127 (2015)
3. Elkadeem, M.R., Younes, A., Mazzeo, D., Jurasz, J., Elia Campana, P., Sharshir, S.W., Alaam, M.A.: Geospatial-assisted multi-criterion analysis of solar and wind power geographical-technical-economic potential assessment. Appl. Energy **322**, 119532 (2022)
4. Ahmed, A.S.: Investigation of wind characteristics and wind energy potential at Ras Ghareb Egypt. Renew. Sustain. Energy Rev. **15**, 2750–2755 (2011)
5. Pazikadin, A.R., Rifai, D., Ali, K., Malik, M.Z., Abdalla, A.N., Faraj, M.A.: Solar irradiance measurement instrumentation and power solar generation forecasting based on Artificial Neural Networks (ANN): A review of five years research trend. Sci. Total. Environ. **715**, 136848 (2020)
6. Hanifi, S., Liu, X., Lin, Z., Lotfian, S.: A critical review of wind power forecasting. Energies **13**, 1–24 (2020)
7. Aliyu, A.K., Modu, B., Tan, C.W.: A review of renewable energy development in Africa: A focus in South Africa, Egypt and Nigeria. Renew. Sustain. Energy Rev. **81**, 2502–2518 (2018)
8. Yadav, A.K., Chandel, S.S.: Solar radiation prediction using Artificial Neural Network techniques: A review. Renew. Sustain. Energy Rev. **33**, 772–781 (2014)
9. Shamshirband, S., Keivani, A., Mohammadi, K., Lee, M., Hamid, S.H.A., Petkovic, D.: Assessing the proficiency of adaptive neuro-fuzzy system to estimate wind power density: Case study of Aligoodarz Iran. Renew. Sustain. Energy Rev. **59**, 429–435 (2016)
10. Celik, A.N., Kolhe, M.: Generalized feed-forward based method for wind energy prediction. Appl. Energy **101**, 582–588 (2013)
11. Jency P., Praynlin.: Solar photovoltaic output power forecasting using back propagation neural network. ICTACT J. Soft Comput. **06**, 1144–52 (2016).

12. Mousavi, S.M., Mostafavi, E.S., Jaafari, A., Jaafari, A., Hosseinpour, F.: Using measured daily meteorological parameters to predict daily solar radiation. Meas. J. Int. Meas. Confed. **76**, 148–155 (2015)
13. Sachit, M.S., Shafri, H.Z.M., Abdullah, A.F., Rafie, A.S.M., Gibril, M.B.A.: Global spatial suitability mapping of wind and solar systems using an explainable AI-based approach. ISPRS Int. J. Geo-Info. **11**, (2022).
14. Rane, N., Choudhary, S., Rane, J. Artificial Intelligence and machine learning in renewable and sustainable energy strategies: a critical review and future perspectives. SSRN Electron. J., 7–9 (2024).
15. Ouachani, I., Rabhi, A., Yahyaoui, I., Tidhaf, B., Tadeo, T.F.: Renewable energy management algorithm for a water pumping system. Energy Procedia **111**, 1030–1039 (2017)
16. Poompavai, T., Kowsalya, M.: Control and energy management strategies applied for solar photovoltaic and wind energy fed water pumping system: A review. Renew. Sustain. Energy Rev. **107**, 108–122 (2019)
17. Rehman, S., Sahin, A.Z.: Wind power utilization for water pumping using small wind turbines in Saudi Arabia: A techno-economical review. Renew. Sustain. Energy Rev. **16**, 4470–4478 (2012)
18. Mentis, D., Hermann, S., Howells, M., Welsch, M., Siyal, S.H.: Assessing the technical wind energy potential in africa a GIS-based approach. Renew. Energy **83**, 110–125 (2015)
19. Global wind atlas. https://globalwindatlas.info/.
20. Ben Ammar, R., Ben Ammar, M., Oualha, A.: Photovoltaic power forecast using empirical models and artificial intelligence approaches for water pumping systems. Renew. Energy **153**, 1016–1028 (2020)
21. Viet, D.T., Van Phuong, V., Duong, M.Q., Tran, Q.T.: Models for short-term wind power forecasting based on improved artificial neural network using particle swarm optimization and genetic algorithms. Energies **13** (2020).
22. Sarić, R., Jokić, D., Beganović, N., Pokvić, L.G., Badnjević, A.: FPGA-based real-time epileptic seizure classification using Artificial Neural Network. Biomed. Signal Process. Control **62**, 102106 (2020)
23. Chen, Q., Folly, K.A.: Wind power forecasting. IFAC-PapersOnLine **51**, 414–419 (2018)
24. Abhishek, K., Singh, M.P., Ghosh, S., Anand, A.: Weather forecasting model using artificial neural Network. Procedia Technol. **4**, 311–318 (2012)
25. Sattar, A.M.A.: Gene expression models for prediction of dam breach parameters. J. Hydroinformatics **16**, 550–571 (2014)
26. Ferreira, C.: Gene expression programming: A new adaptive algorithm for solving problems. Gene Expr. Prog. **13**, 1–22 (2001).
27. Shboul, B., AL-Arfi, I., Michailos, S., Ingham, D., Ma, L., Hughes, K.J., Pourkashanian, M.: A new ANN model for hourly solar radiation and wind speed prediction: A case study over the north & south of the Arabian Peninsula. Sustain. Energy Technol. Assessments **46**, 101248 (2021).
28. Van Dao, D., Adeli, H., Ly, H.B., Le, L.M., Le, V.M., Le, T.T., Pham, B.T.: A sensitivity and robustness analysis of GPR and ANN for high-performance concrete compressive strength prediction using a monte carlo simulation. *Sustain.* **12**, 830 (2020).
29. Gao, M., Yin, L., Ning, J.: Artificial neural network model for ozone concentration estimation and Monte Carlo analysis. Atmos. Environ. **184**, 129–139 (2018)
30. Jiang, S., Sharafisafa, M., Shen, L.: Using artificial neural networks to predict influences of heterogeneity on rock strength at different strain rates. *Materials (Basel).* **14**, 3042 (2021)
31. Carolin Mabel, M., Fernandez, E.: Analysis of wind power generation and prediction using ANN: A case study. Renew. Energy **33**, 986–992 (2008)
32. Amarasinghe, G.: An artificial neural network for solar power generation forecasting using weather parameters. (2018).
33. Peiris, A.T., Jayasinghe, J., Rathnayake, U.: Forecasting wind power generation using artificial neural network: A case study from Sri Lanka. J. Electr. Comput. Eng. **2021** (2021)

34. Ekanayake, P., Peiris, A.T., Jayasinghe, J.M.J.W., Rathnayake, U.: Development of wind power prediction models for Pawan Danavi wind farm in Sri Lanka. Math. Probl. Eng. **2021**, 1–13 (2021).
35. Hashmi, M.Z., Shamseldin, A.Y., Melville, B.W.: Statistical downscaling of watershed precipitation using Gene Expression Programming (GEP). Environ Model Softw. **26**, 1639–1646 (2011)
36. Puri, V., Kumar, N.: Wind energy forecasting using artificial neural network in himalayan region. Model. Earth Syst. Environ. **8**, 59–68 (2022)
37. Yaniktepe, B., Tasdemir, S., Guher, A.b., Akcan, S. Wind power forecasting for the province of Osmaniye using artificial neural network method. Int. J. Intell. Syst. Appl. Eng. **4**, 114–7 (2016)
38. AECOM. The mathematics of pumping water AECOM design build civil, mechanical engineering. http://www.raeng.org.uk/publications/other/17-pumping-water. Last accessed 25 April 2024
39. Saleh, H., Abou El-Azm Aly, A., Abdel-Hady, S.: Assessment of different methods used to estimate Weibull distribution parameters for wind speed in Zafarana wind farm, Suez Gulf, Egypt. *Energy* **44**, 710–719 (2012)
40. Waewsak, J., Chancham, C., Landry, M., Gagnon, Y.: An analysis of wind speed distribution at Thasala, Nakhon Si Thammarat Thailand. J. Sustain. Energy Environ. **2**, 51–55 (2011)
41. Alham, M. H., Fathy Gad, M., Khalil Ibrahim, D.: Potential of wind energy and economic assessment in Egypt considering optimal hub height by equilibrium optimizer. Ain Shams Engin. J. **14**, 101816 (2022)

# Analytics for Environmental Sustainability I

# Optimizing Foundry Operations: A Case Study on Improving Material and Energy Flow Efficiency in Line with the SDGs

**Ahmed ElFeky, Hazem Elshiekh, and Amna Ramzy**

**Abstract** This study focuses on optimizing material and energy flows within a foundry to enhance sustainability and align with Sustainable Development Goals (SDGs) 9 (Industry, Innovation, and Infrastructure) and 12 (Responsible Consumption and Production). Using UMBERTO® Efficiency + software, key inefficiencies were identified, particularly in the induction furnaces, which consumed 85% of the facility's total energy. By implementing targeted optimization strategies, the foundry achieved significant improvements, including expected annual energy savings of 429,312 kWh and a reduction of 236 metric tons of $CO_2$ emissions. Additionally, reclassifying iron shavings by alloy type led to an increased revenue of 74,690 EGP, totaling 329,890 EGP. These results demonstrate that systematic resource management can greatly enhance both economic and environmental outcomes. This approach not only improves operational efficiency but also contributes to global efforts in achieving more sustainable industrial practices, in line with the SDGs.

**Keywords** Material and Energy Flow Optimization · Sustainable Development Goals (SDGs) · Industrial Efficiency · Greenhouse Gas Emissions Reduction · Foundry Operations · Resource Management · Modelling and simulation

## 1 Introduction

The high consumption of fossil fuels is a major driver of greenhouse gas (GHG) emissions, contributing significantly to global warming due to the substantial release of $CO_2$ and its equivalents. Since the mid-eighteenth century, the Industrial Revolution has led to a dramatic rise in carbon dioxide levels. By 2015, $CO_2$ emissions had surged by 43%, [1]. The increase in GHGs, including $CO_2$, nitrous oxide, methane,

A. ElFeky (✉) · A. Ramzy
German University in Cairo, Cairo, Egypt
e-mail: ahmed.el-feky@guc.edu.eg

H. Elshiekh
Sequoia for Industrial Investments, Cairo, Egypt



V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_11

and fluorinated gases, not only traps heat and raises global temperatures but also accelerates climate change, leading to rising sea levels [2]. These emissions originate from various human activities across all sectors of the global economy, including power production, transportation, manufacturing, and agriculture. Therefore, it is imperative for all these sectors to collaborate in reducing their GHG emissions.

The global challenge of reducing GHG emissions, driven largely by the high consumption of fossil fuels, is mirrored in the energy consumption patterns of individual countries. For instance, according to the United Nations Industrial Development Organization (UNIDO), Egypt is the third most populous country in Africa and the continent's largest consumer of oil and gas. As indicated in Table 1, Egypt's total energy consumption reached 1054 terawatt-hours (TWh) and has been increasing at a rate of nearly 7% annually since the early 2000s. However, this energy consumption is predominantly reliant on fossil fuels, with gas and oil together constituting over 90% of the total energy mix. This substantial dependence on carbon-intensive sources presents considerable challenges for the country in its efforts to enhance energy efficiency and transition towards a more sustainable energy system. Although there has been some progress in adopting renewable energy sources, such as hydro, wind, and solar, these currently account for only 6.2% of total energy consumption. The dominance of fossil fuels not only significantly impacts Egypt's greenhouse gas (GHG) emissions but also underscores the potential for substantial improvements in energy efficiency, particularly within the industrial sector, which is a major contributor to GHG emissions [3].

While Egypt's overall energy consumption is heavily reliant on fossil fuels, contributing significantly to its greenhouse gas (GHG) emissions, the industrial sector plays a particularly pivotal role. As illustrated in Fig. 1, the industrial sector is responsible for 31% of Egypt's total energy consumption, marking it as a significant contributor to the country's greenhouse gas (GHG) emissions. According to the Regional Center for Renewable Energy and Energy Efficiency (RCREEE), the sector's heavy reliance on fossil fuels, especially natural gas and oil, drives its substantial GHG output. The energy-intensive nature of industrial activities highlights the critical need for enhanced energy efficiency and the adoption of cleaner technologies. Addressing energy consumption in this sector offers Egypt a considerable opportunity to reduce its GHG emissions and move closer to more sustainable energy practices. The figure

**Table 1** The breakdown of Egypt's energy consumption by source in 2021

| Source | Consumption (TWh) | Percentage of Total (%) |
|---|---|---|
| Gas | 619 | 58.70 |
| Oil | 355 | 33.70 |
| Hydro | 38 | 3.60 |
| Wind | 20 | 1.90 |
| Coal | 14 | 1.30 |
| Solar | 7 | 0.70 |
| Total | 1054 | 100.00 |

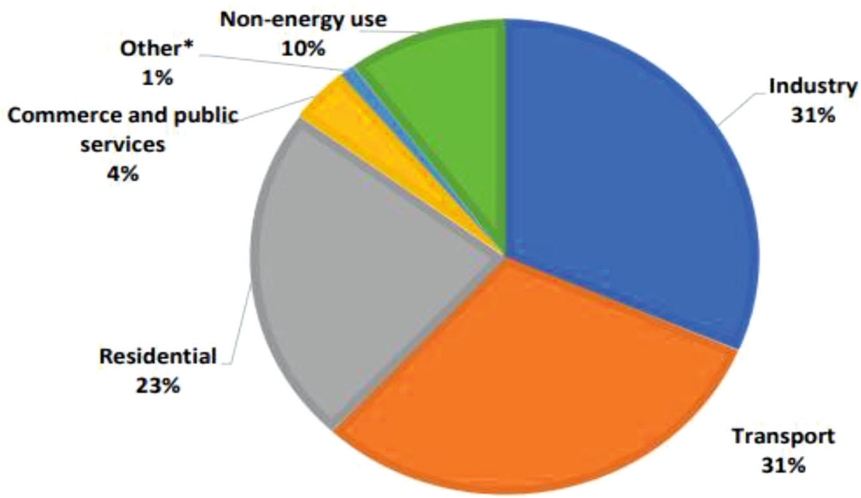



**Fig. 1** Share of electricity usage in Egypt by sector [4]

underscores the urgency of implementing targeted strategies to mitigate environmental impacts and improve the overall efficiency of energy use within the industrial domain.

Optimizing industrial energy systems is vital for reducing energy consumption, lowering operating costs, and minimizing greenhouse gas emissions. Given the complexity of industrial processes and the vast amounts of data involved, manual calculations for energy integration have become increasingly impractical. As a result, specialized software tools have been developed to model and optimize energy recovery systems, however, the diverse needs of energy integration mean that no single software can comprehensively address all aspects, making the selection process a critical task for researchers. Therefore, it is essential to select the appropriate software based on the specific objectives and requirements of each case study, as well as whether the focus is on a large company or a small and medium-sized enterprises (SMEs).

While larger industries have successfully implemented these tools, SMEs often face challenges in adopting them. This disparity in sustainability practices underscores the necessity for developing energy integration solutions tailored to the specific needs and constraints of SMEs. Small and medium-sized enterprises (SMEs) are globally recognized as vital contributors to economic growth, accounting for a substantial share of GDP and employment across various regions. According to the Organization for Economic Co-operation and Development (OECD), SMEs contribute up to 60% of the total GDP within the European Union (EU) [5]. Despite their economic significance, SMEs often trail behind larger corporations in adopting sustainability practices, leading to considerable environmental impacts. Globally, SMEs are responsible for approximately 60 to 70% of industrial environmental impact, with certain regions, such as the Asia–Pacific, facing even more severe consequences. For instance, De et al. (2020) found that SMEs in this region contribute over 60% of industrial pollution and consume more than 13% of the world's energy. In Europe, SMEs account for around 70% of industrial emissions, making them significant contributors to pollution [6].

In Egypt, SMEs similarly play a crucial role in the economy, significantly contributing to both employment and GDP. Research indicates that SMEs account for approximately 80% of total employment in the private sector, underscoring their importance in job creation and economic stability [7]. However, these enterprises face substantial challenges, including limited access to finance, bureaucratic hurdles, and inadequate infrastructure, which impede their growth and sustainability [8, 9]. According to the EU definition, an SME is an enterprise with fewer than 250 employees and a turnover of less than €50 million [8, 9]. To mitigate environmental impacts and reduce energy consumption, it is imperative that Egyptian SMEs increasingly adopt improve material and energy efficiency. This shift would not only enhance their sustainability but also alleviate the burden on the government to develop new power stations [4].

Achieving sustainability in SMEs requires a balanced approach that takes into account economic, environmental, and social factors [6]. This research focuses specifically on enhancing the environmental sustainability of SMEs. To achieve this, several key areas need to be addressed:

1. **Material Efficiency**: Implementing strategies to minimize waste generation and enhance recycling practices can significantly reduce the environmental impact of SMEs.
2. **Energy Efficiency**: Detecting energy consumption hotspots and conducting research and development (R&D) to address the identified issues. By targeting these specific areas, SMEs can implement more effective energy-saving strategies and enhance their overall sustainability.

Achieving sustainability in small and medium-sized enterprises (SMEs) calls for a thoughtful balance between economic, environmental, and social factors. This research highlights the importance of environmental sustainability, focusing on the efficient use of materials and energy to cut down waste and enhance resource utilization. By employing Material and Energy Flow Analysis (MEFA):

### *1.1 Material and Energy Flow Analysis (MEFA)*

Material and Energy Flow Analysis (MEFA) is a critical methodology in industrial engineering, designed to measure and optimize the flow of materials and energy within production processes. By focusing on both material mass and energy usage, MEFA provides a comprehensive understanding of how resources are consumed and transformed during manufacturing. The primary objective of MEFA is to maintain a material balance across all stages of production, from the initial sourcing of raw materials, through intermediate processes, to the final product. This balance is guided by the law of conservation of mass and energy, which states that the total mass and energy in a closed system must remain constant.

MEFA begins by mapping the entire production process, identifying sources of materials and energy, and tracking their pathways through various processing stages. Each operation within the flow process is analyzed to determine its inputs and outputs, enabling the identification of inefficiencies or losses. By comparing the input materials and energy with the outputs, including emissions and waste, MEFA ensures that all resources are accounted for, thereby minimizing waste and optimizing resource utilization. This analysis is essential not only for enhancing the efficiency of industrial processes but also for reducing environmental impacts, particularly in terms of emissions and waste.

Beyond balancing material and energy flows, MEFA also aids in identifying opportunities for recycling and reusing materials within the production process, contributing to more sustainable manufacturing practices. For instance, in the context of ductile iron production, as discussed earlier, MEFA can be applied to optimize the use of scrap iron and energy in melting processes, ensuring efficient resource utilization while minimizing waste and emissions. Overall, MEFA is a powerful tool that supports sustainable development goals by promoting responsible consumption and production within industrial systems.

This research aligns with the objectives of the Sustainable Development Goals (SDGs), 9 (Industry, Innovation, and Infrastructure), and 12 (Responsible Consumption and Production). The study aims to evaluate the efficiency of material and energy flow software in optimizing resource use within Egypt's industrial sector. The primary objective is to develop sustainable models centered on two core pillars: material efficiency and energy efficiency. By focusing on these areas, the research seeks to reduce greenhouse gas (GHG) emissions and contribute to global efforts in combating climate change. Advancing sustainable practices and improving resource efficiency in industrial operations are critical steps towards achieving a greener and more sustainable future for Egypt.

Building on the identification of key inefficiencies and the implementation of targeted optimization strategies within the foundry, the next chapter will dive into the detailed methods used to achieve these improvements. It will specifically cover how Material and Energy Flow Analysis (MEFA) and UMBERTO® Efficiency + software were applied to map and analyze the foundry's processes. Additionally, the chapter will discuss the systematic approach to optimizing material and energy flows, reducing waste, and improving overall sustainability. Through a comprehensive case study of a small to medium-sized enterprise (SME) in Egypt's industrial sector.

## 2 Methodology

As shown in Fig. 2, the methodology began with an on-site visit to Sequoia Company, followed by data gathering on production, energy, and material flows. This data was analyzed to model current operations, pinpointing inefficiencies and hotspots. From these insights, improvement opportunities were outlined to enhance the facility's economic and environmental performance.

### *2.1 SME Case Study*

This research paper examines the intricate production process within a casting facility dedicated to manufacturing ductile iron fittings and accessories. The process begins with two induction furnaces, each capable of melting 200 tons of scrap iron per month, setting the foundation for downstream operations. The molten iron is then directed into three distinct casting lines: fully automated, semi-automated, and manual.

In the fully automated line, advanced robotics and control systems oversee the entire casting process, ensuring precision and consistency from pouring to solidification. The semi-automated line, which combines automation with human oversight, offers greater flexibility, with critical tasks such as mold preparation and casting inspection handled by skilled operators. This blend of automation and human intervention ensures both efficiency and adaptability in production. The manual line is reserved for specialized or lower-volume orders, where experienced foundry workers manually execute the casting process, delivering high craftsmanship for custom products.

Following the casting process, the gating system is removed, and the castings undergo a thorough quality inspection to detect defects such as porosity or dimensional inaccuracies. Successful castings proceed to sandblasting, where abrasive cleaning removes surface impurities, preparing them for the subsequent machining

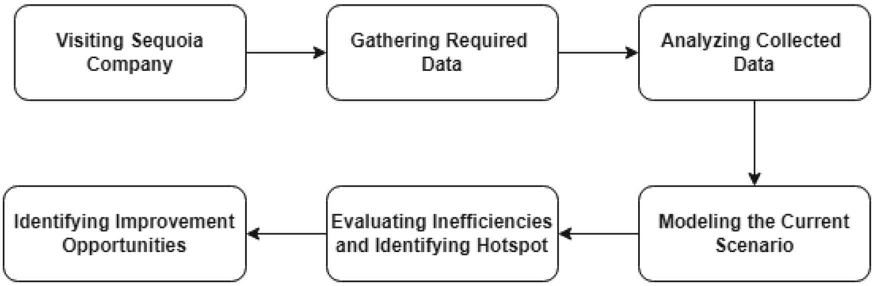



**Fig. 2** Flow diagram of experimental procedures

phase. During machining, the castings are meticulously shaped to meet precise specifications, which is crucial for ensuring the proper fit and functionality of the ductile iron fittings.

The final steps involve painting the products to prevent corrosion and enhance durability, after which the components are assembled into finished fittings and accessories. This study provides a comprehensive analysis of these processes, emphasizing efficiency, quality control, and the strategic integration of automation to optimize production within the casting facility. Figure 3 shows the whole process, starting from melting scrap iron in induction furnaces all the way to the final assembly.

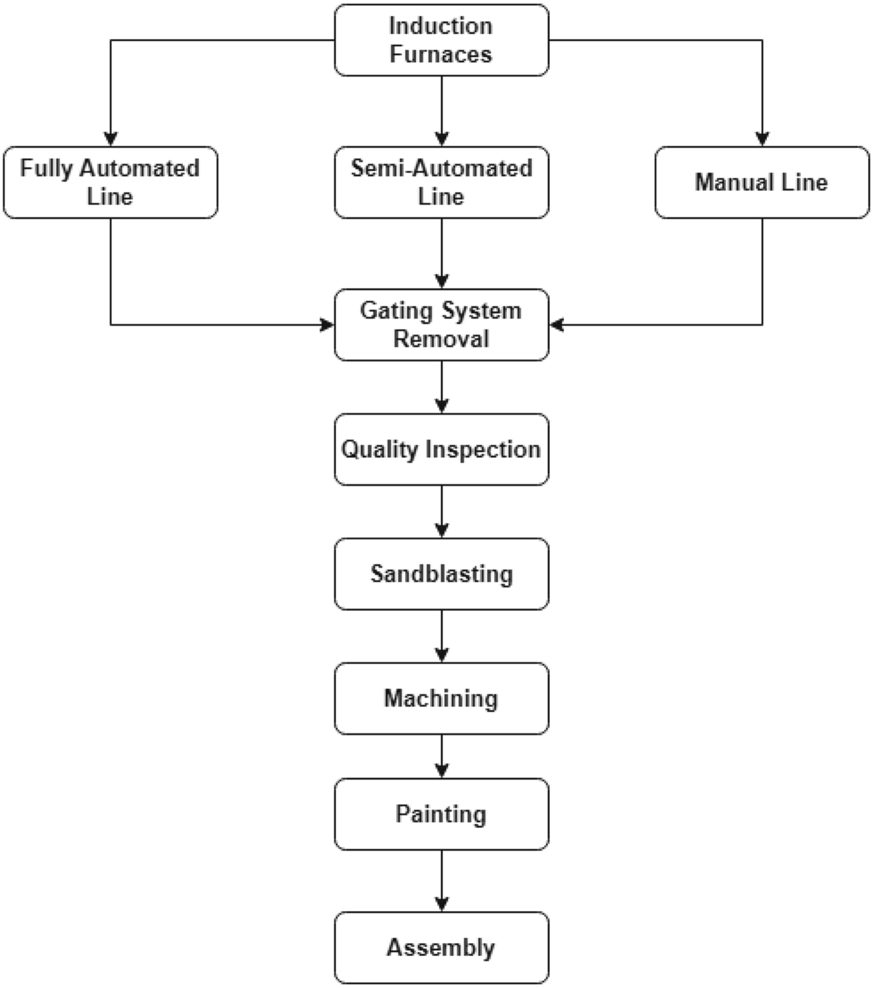



**Fig. 3** Process diagram: from melting scrap iron to final product assembly

### 2.2 Data Collection and Analysis

Material and energy data were collected from the company, using 2023 as the reference year.

- **Materials Consumption Analysis (2023):**

The data provided for the year 2023 in Table 2 offers a detailed overview of the material flow within the foundry, highlighting both the efficiency of the iron-casting process and areas where improvements could be made. Out of the total 1,461,000 kg of molten iron processed, 1,011,000 kg were successfully converted into cast products, demonstrating a high conversion rate that underscores the foundry's ability to efficiently produce final products. This aligns well with the goals of material and energy flow optimization.

However, the data also reveals that 345,000 kg of iron were utilized in the gating systems, risers, and sprues. While these are essential components of the casting process, they represent a significant portion of material that does not directly contribute to the final product. The fact that this material can likely be recycled within the facility is promising, as it suggests that a substantial amount of iron could be reintegrated into the production cycle, thereby reducing waste and conserving resources.

Additionally, 56,000 kg of iron were categorized as rejected products. Although these products did not meet the required standards, their potential for recycling is a positive aspect, as it ensures that the material is not entirely lost but can be reprocessed, further enhancing resource efficiency. Iron shavings, totaling 23,000 kg, were generated during various stages of the process, particularly during machining. The management of these shavings within the facility reflects a commitment to minimizing waste and maximizing the use of raw materials.

Despite these efforts, there were still 26,000 kg of waste losses, attributed to factors such as droplets, splashes, broken parts during casting, iron solidification in the furnace and ladles, and poor agglomeration of iron shavings. These losses highlight areas where further optimization could be pursued, potentially through improved process controls, enhanced recycling techniques, or better handling practices.

- **Energy Consumption Analysis (2023)**

In 2023, the foundry's total energy consumption reached 1,902,833 kWh, as detailed in Table 3. A significant portion of this energy (1,616,408 kWh) was

**Table 2** Material data (2023)

| | Total Molten Iron Processed (Kg) | Casted Products (kg) | Gating Systems, Risers, and Sprues (kg) | Rejected Products (kg) | Iron Shavings (kg) | Waste losses (Kg) |
|---|---|---|---|---|---|---|
| $\Sigma$Total | 1,461,000 | 1,011,000 | 345,000 | 56,000 | 23,000 | 26,000 |

**Table 3** Energy data (2023)

| Year 2023 | Total energy consumption (KWh) | Induction Furnaces energy consumption (KWh) | Remaining Processes energy consumption (KWh) |
|---|---|---|---|
| $\Sigma$Total | 1,902,833 | 1,616,408 | 286,424 |

**Table 4** Electricity cost data in Egypt [10]

| | Amount | Emissions | Price |
|---|---|---|---|
| Electricity specifications in Egypt | 1 (kWh) | 0.55 kg $CO_2$ | 1.55 EGP/ KWh |

consumed by the induction furnaces, which are essential for melting the iron used in the casting process. This indicates that approximately 85% of the total energy consumption was dedicated to the core activity of molten iron production, underscoring the energy-intensive nature of this stage. The remaining 286,424 kWh of energy was utilized across other aspects of the production process, including the operation of casting lines, machining, finishing processes, and facility maintenance. Although this represents a smaller portion of the total energy use, optimizing these auxiliary processes could contribute to overall energy efficiency improvements.

- **Environmental and Economic Impact**

The environmental impact of the foundry's energy consumption can be assessed by examining the associated carbon emissions. According to the electricity specifications for Egypt provided in Table 4, each kilowatt-hour (kWh) of electricity consumed results in 0.55 kg of $CO_2$ emissions. Given the total energy consumption of 1,902,833 kWh, the foundry's operations generated approximately 1,046,558 kg (1,046.6 metric tons) of $CO_2$ in 2023. This substantial carbon footprint highlights the importance of exploring energy efficiency measures and renewable energy options to reduce emissions.

From an economic perspective, the cost of electricity in Egypt is 1.55 EGP per kWh. Therefore, the total electricity cost for the foundry in 2023 was approximately 2,949,392 EGP. This significant expenditure underscores the potential financial benefits of improving energy efficiency within the facility. By reducing energy consumption, particularly in the induction furnaces, the foundry could lower operational costs and decrease its environmental impact, aligning with broader sustainability goals.

### *2.3 Modeling Tool*

This research aims to pursue the research objective which to offer practical insights customized to suit the specific requirements of SMEs, promoting their sustainability goals and fostering a more sustainable future through the following methodology:

1. Assessing the current state of sustainability management in a real case SME.
2. Determining the challenges, opportunities, and issues related to the adoption of sustainability practices from a managerial perspective.
3. Proposing relevant strategies geared towards facilitating the successful implementation of sustainability practices.

The modeling tool utilized in this research is UMBERTO® Efficiency + (version 10.0), a sophisticated software developed by the Ifu Hamburg and commercialized by iPoint-systems GmbH in Germany. UMBERTO® Efficiency + is designed to optimize and visualize production processes through advanced process mapping, providing a powerful platform for analyzing material and energy flows within industrial systems. This tool is particularly effective in assessing the environmental and economic impacts of production by incorporating detailed analysis methods such as Sankey diagrams, carbon footprint calculations, and cost analysis.

One of the key strengths of UMBERTO® Efficiency + is its ability to create highly accurate models of complex production systems. This is achieved by customizing the equations that govern the relationships between inputs and outputs at each stage of the production process. By defining specific parameters and variables, the software can simulate real-world operations with a high degree of precision, allowing for the identification of inefficiencies, bottlenecks, and areas where resource consumption can be reduced. This customization is crucial for tailoring the model to the specific characteristics of the production line being studied, ensuring that the results are relevant and actionable.

The process mapping capabilities of UMBERTO® Efficiency + are central to its functionality. The software allows users to build detailed process networks that visually represent the flow of materials and energy through the system. These networks are then analyzed using Sankey diagrams, which provide a clear and intuitive illustration of the relative magnitudes of these flows. Sankey diagrams are particularly useful for identifying where the most significant losses or inefficiencies occur, as they visually depict the distribution of energy, materials, and costs across the entire production line.

In addition to material and energy flow analysis, UMBERTO® Efficiency + supports carbon footprint and cost analyzes, enabling a comprehensive evaluation of the environmental and economic performance of the production process. By integrating these analyzes, the software helps researchers and industry professionals make informed decisions about process improvements and sustainability initiatives.

## *2.4 Mapping and Optimization of Material and Energy Flows*

In this research, a detailed mapping of the current production processes was conducted to accurately assess material and energy consumption within the foundry. Figure 4 shows utilizing UMBERTO® Efficiency + software, the actual processes were modeled, ensuring precise quantification of all material inputs and outputs, as well as the associated energy consumption. This comprehensive approach enables a thorough evaluation of the efficiency and sustainability of the production line, the first step in this process involves mapping the entire material flow from the melting of scrap iron in the induction furnaces to the final production of ductile iron fittings and valves. By closely monitoring each stage, including the handling of gating systems, risers, sprues, rejected products, and the generation of iron shavings and other waste, the model provided a clear picture of resource utilization within the foundry. Once the material and energy flows were mapped, a Sankey diagram was employed to visually represent these flows. The Sankey diagram proved invaluable in highlighting the distribution of materials and energy throughout the production process, making it easier to detect hotspots areas where inefficiencies or excessive resource consumption occur. These hotspots typically represent stages in the process with significant material losses, energy wastage, or high emissions.

After identifying these hotspots, targeted research was conducted to develop strategies for improvement. The focus was on optimizing the processes associated with the most significant inefficiencies. For instance, the research explored methods to reduce energy consumption in the induction furnaces, enhance recycling rates of gating systems and rejected products, and minimize waste generation during machining and casting. This approach aimed not only at improving material and energy efficiency but also at reducing the overall environmental impact of the foundry's operations.

## 3 Results and Discussion

This section analyzes material and energy flows to pinpoint waste and energy-intensive hotspots. Based on these insights, targeted optimizations were introduced to improve efficiency and minimize waste.

### *3.1 Material Flow Analysis*

The material flow analysis, represented by the green line in the Sankey diagram, shows that for every 1 kg of scrap, 0.966 kg of ductile iron is produced, resulting in only 3.35% of waste. Notably, 53% of this waste, amounting to 26 tons per year, originates from the three production lines. A process investigation revealed that this material loss is due to droplets, splashes, and broken parts during the casting process in these lines, which are subsequently lost during sand reclamation. Additionally, the remaining 47% of waste, totaling 23 tons per year, is generated during the machining process. Although this waste is unavoidable due to the nature of machining, it is managed effectively within the factory.

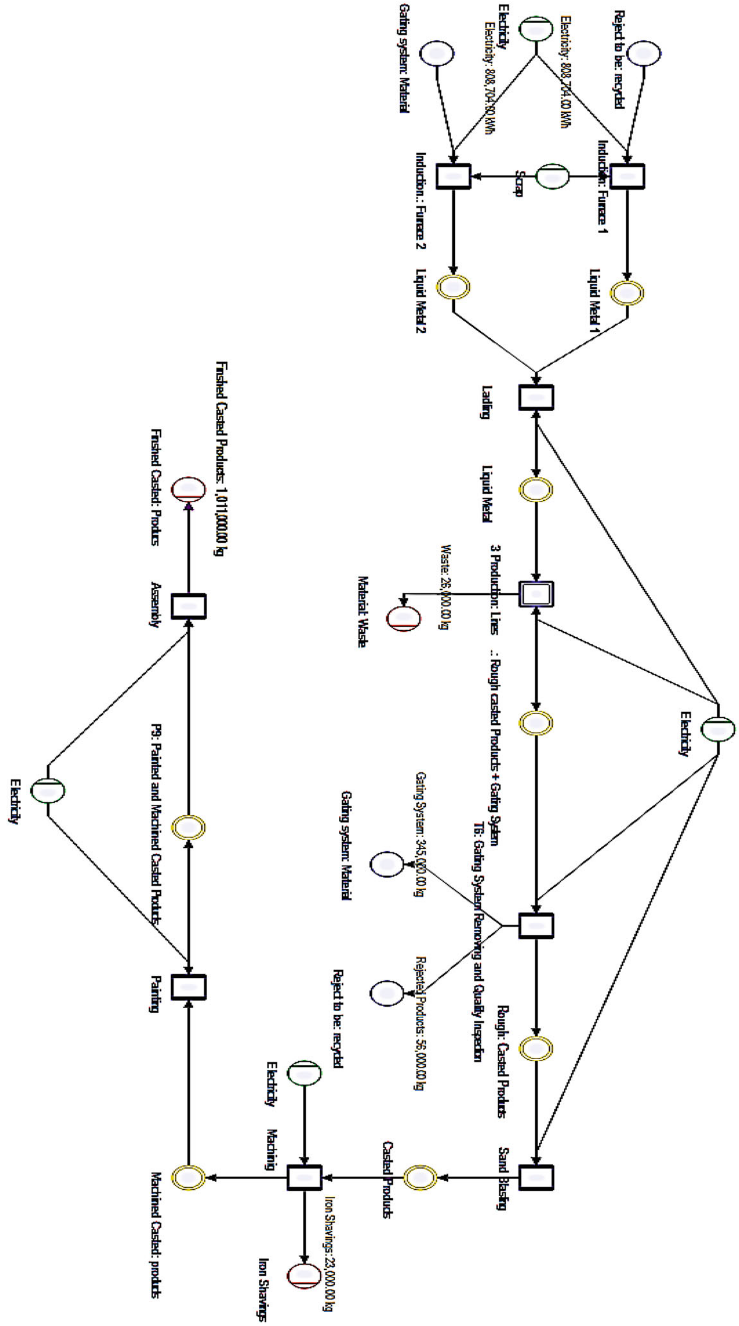



**Fig. 4** Actual Line Mapping

### 3.2 Energy Flow Analysis

After analyzing the energy flow (illustrated by the yellow line in the Sankey diagram), two induction furnaces are the primary energy hotspots, consuming nearly 85% of the facility's total energy, as shown in Fig. 5. The facility's overall energy usage is 1,902,833 kWh per year, with the induction furnaces alone consuming 1,616,408 kWh annually. Further investigation revealed that although the foundry is designed for a capacity of 400 tons per month, it typically operates at a reduced capacity, producing around 120 tons of molten metal monthly to meet market demands.

The furnaces generally operated 5 days a week or less, depending on demand, with each operational day involving 2.4 charges, resulting in 48 charges per month (2.4 charges $\times$ 5 days $\times$ 4 weeks = 48 charges per month). The average monthly energy consumption under this regime was 134,784 kWh (48 charges $\times$ 2806 kWh, where 2806 kWh is the weighted average energy per charge). This amounted to an annual consumption of 1,616,408 kWh under the usual operating conditions, highlighting the arc furnaces as significant energy hotspots that warrant close attention.

## 4 Optimization Approach

Based on the material and energy flow analyzes, enhanced by the software's high transparency, key hotspots were identified, allowing for the implementation of targeted optimization strategies:

- **Hotspot 1 (Production Lines Waste):**

To improve material flow within the cast iron foundry, the installation of a magnetic separator before the sand reclamation line was recommended. The primary objective of this installation was to recover 26,000 kg of metallic iron waste, resulting from droplets, splashes, and broken parts during the casting process in the three production lines that could be reintegrated into the production process, thereby reducing raw material consumption and minimizing waste. The magnetic separator was operational for four weeks, and the results are presented in Table 5. As detailed in the table, the separator consistently recovered significant quantities of metallic iron each week, with variations observed across the four-week period. The highest recovery occurred in the fourth week, with 267 kg of metallic iron reclaimed, while the lowest was in the second week, with 195 kg. These results underscore the effectiveness of the separator in capturing valuable material that would have otherwise been lost. On average, this approach can save approximately 925 kg per month, which equates to 11,100 kg annually, representing nearly 43% of the material loss from the three production lines.

- **Hotspot 2 (Iron Shavings):**

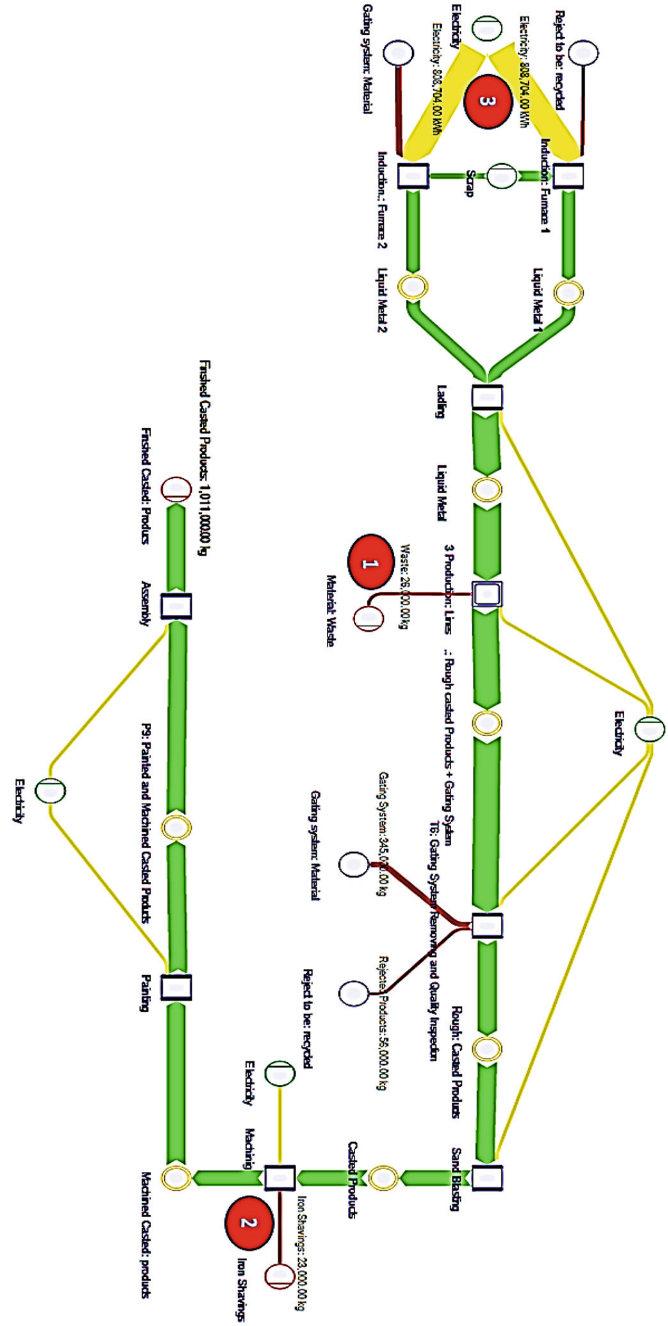



**Fig. 5** Actual Line Sankey Diagram

**Table 5** Material savings due to magnetic separator

| Week | Reclaimed Metallic Iron (kg) |
|---|---|
| 1st Week | 243 |
| 2nd Week | 195 |
| 3rd Week | 220 |
| 4th Week | 267 |

To enhance material value and maximize revenue, we proposed the classification of iron shavings by alloy type. Previously, the foundry sold all shavings at a uniform price of approximately 11,000 EGP per ton, regardless of their alloy composition. By classifying the shavings into specific grades namely Grey Cast Iron, Ductile Cast Iron, and Steel, the company was able to sell each type at its respective market value, significantly improving revenue. The breakdown of the shavings by type, along with the corresponding revenues, is presented in Table 6. This classification process allowed the foundry to achieve a total revenue of 329,890 EGP from the sale of 23.2 tons of iron shavings.

Table 7 compares the revenue generated from the classified shavings against the revenue that would have been earned if the shavings were sold at a uniform price of 11,000 EGP per ton. The classification strategy led to an increase in revenue of 74,690 EGP.

This classification strategy resulted in a total revenue of 329,890 EGP in 2024, compared to the 255,200 EGP that would have been generated under the previous uniform pricing strategy. This represents an increase of 74,690 EGP, which is a considerable gain for the foundry. The findings underscore the importance of material classification in maximizing revenue and optimizing resource utilization. The additional revenue generated from this approach could be reinvested into further process

**Table 6** Shavings composition and revenue analysis for 2023

| Shavings Type | Percentage | Weight (tons) | Revenue per Ton (EGP) | Total Revenue (EGP) |
|---|---|---|---|---|
| Grey Cast Iron | 15% | 3.48 | 11,000 | 38,280 |
| Ductile Cast Iron | 78% | 18.10 | 14,500 | 262,450 |
| Steel | 7% | 1.62 | 18,000 | 29,160 |
| Total | 100% | 23.2 | | 329,890 |

**Table 7** Comparison of revenue (classified vs. uniform price)

| Revenue Strategy | Total Weight (tons) | Total Revenue (EGP) |
|---|---|---|
| Classified Shavings | 23.2 | 329,890 |
| Uniform Price Shavings | 23.2 | 255,200 |
| Increased Revenue | | 74,690 |

**Table 8** New approach (proposed) total energy per operational day

| Parameter | Previous Operating System | Proposed Operating System |
|---|---|---|
| Operational Days per Month | 20 days (5 days/week) | 8 days (2 days/week) |
| Charges per Operational Day | 2.4 charges/day | 5 charges/day |
| Energy Consumption per Day | 6,739.2 kwh | 12,376 kWh |
| Monthly Energy Consumption | 6,739.2 * 20 = 134,784 kWh | 12,376 * 8 = 99,008 kWh |
| Annual Energy Consumption | 1,616,408 kWh | 1,188,096 kWh |

**Table 9** The final optimized model estimated savings,

| | Material (kg/year) | Electricity (kwh/year) | $CO_2$ Emissions (kg/year) | Cost (EGP/year) |
|---|---|---|---|---|
| Total | 11,000 | 429,312 | 236,121.6 | 665,433.6 |

improvements, energy efficiency measures, or other initiatives aimed at enhancing the overall sustainability and profitability of the foundry.

- **Hotspot 3 (Induction Furnaces):**

To optimize energy use, it was proposed to increase the number of daily charges to 5, as the first charge consumes 66% more energy to initiate melting. Specifically, the first charge requires 3640 kWh, while any subsequent charge consumes only 2180 kWh. By increasing the number of charges and reducing operational days to approximately 2 days per week, this strategy aims to maximize the foundry's unused capacity and concentrate energy usage into fewer, more intensive operating days.

From Table 8 above applying the proposed approach, the annual Energy Savings: 1,617,408 kWh - 1,188,096 kWh = 429,312 kWh. The final optimized model estimated savings shown in Table 9,

## 5 Conclusion

This study underscores the importance of optimizing material and energy flows in the foundry understudy, revealing both environmental and economic advantages. The material flow analysis showed that out of 1,461,000 kg of molten iron processed in 2023, 1,356,000 kg were effectively turned into cast products, highlighting efficient use of materials. However, some inefficiencies were noted, including 56,000 kg of rejected products. While these materials didn't directly contribute to the final product, they were identified as recyclable, offering opportunities for resource conservation. The magnetic separator has proven highly effective in minimizing material loss, recovering approximately 11,100 kg of metallic iron annually from a total of 23,000 kg processed.

On the energy front, the analysis pinpointed the induction furnaces as the main energy consumers, using 1,616,408 kWh out of a total of 1,902,833 kWh. This accounted for 85% of the foundry's total energy use, highlighting the need for focused energy-saving strategies. By increasing the number of charges per day and reducing the number of operating days, the foundry could potentially save 429,312 kWh annually, cut $CO_2$ emissions by 236,122 kg, and reduce energy costs by about 665,434 EGP.

The study also demonstrated the financial gains from classifying iron shavings by alloy type rather than selling them at a uniform price. This approach boosted the foundry's revenue by 74,690 EGP in 2023, bringing the total to 329,890 EGP.

## 6 Outlook

To enhance sustainability in the industry, future efforts could focus on integrating renewable energy sources, particularly solar power. With high solar irradiance in regions like Egypt, solar energy offers a strong opportunity to reduce reliance on fossil fuels, leading to lower energy costs and carbon emissions. Installing solar panels to support energy-intensive operations, such as those involving heavy machinery or high-temperature processes, could result in substantial energy savings and reduce the overall environmental impact. This approach aligns with the broader Sustainable Development Goals (SDGs) and positions the industry as a leader in sustainable practices.

## References


1. National Academy of Sciences: Environmental impact. Retrieved from http://needtoknow.nas.edu/energy/energy-costs/environmental/ (2020, June 5)
2. Carbon Footprint: Germany. Retrieved from https://www.carbonfootprint.com/warming.html (n.d.)
3. International Renewable Energy Agency (IRENA): Renewable Energy Outlook: Egypt. Abu Dhabi: IRENA. Retrieved from https://www.irena.org/publications/2018/Oct/Renewable-Energy-Outlook-Egypt (2018)
4. International Energy Agency (IEA): Clean Energy Transitions in North Africa. Paris: IEA. Retrieved from https://www.iea.org/reports/clean-energy-transitions-in-north-africa (2020)
5. European Commission: Annual Report on European SMEs 2017/2018. Retrieved from https://ec.europa.eu/growth/smes/business-friendly-environment/performance-review_el#annual-report (2018)
6. Brammer, S., Hoejmose, S., Marchant, K.: Environmental management in SMEs in the UK: Practices, pressures, and perceived benefits. Bus. Strategy Envi. **21**(7), 423–434 (2012). https://doi.org/10.1002/bse.717
7. Miniesy, R., FakhrEldin, H.: The impact of entrepreneurs' perceptions and social media usage on their intention to formalize their MSMEs in Egypt. J. Entrep. Public Policy. (2023). https://doi.org/10.1108/JEPP-04-2023-0037

8. Ahmed, N., Mlaker, S., Elgazzar, S.: Assessing the situation of micro, small, and medium-sized enterprises in the emerging markets of Egypt under COVID-19 implications. Stud. Bus. Econ. **18**(3), 270–289 (2023). https://doi.org/10.2478/sbe-2023-0037
9. Dey, P. K., Malesios, C., Abdelaziz, F. B., Chowdhury, S., & De, D.: Could lean and innovation enhance supply chain sustainability of small and medium-sized enterprises? Bus. Strat. Envi. **28**(4), 582–598 (2019). https://doi.org/10.1002/bse.2260
10. Abdallah, L., & El-Shennawy, T.: Evaluation of CO2 emission from Egypt's future power plants. Euro-Mediterranean J. Envi. Integr. **49**(5) (2020)

# Achieving Sustainable Design Utilizing Software Programs and Tools in Architecture Field

**Hebatulla H. Elsharawy**

**Abstract** As it become crucial to save the planet, mitigate climate change, keep natural resources, reduce pollution, improve quality of life, and overall live in a sustainable environment, a great concern and big efforts are executed worldwide to achieve this target in all fields of life. On the other hand, there is an exaggerated development in the computer science field and the production of software. Accordingly, in the architecture field and especially in the design stage vast number of computer programs came to duty with divergent functions.

This paper focuses on software programs and to what extent they can generate a sustainable building in its design stage. This primary stage is the most important phase of the building life, as it can control the building behavior, assures its performance, and mostly secure its end of life. With a proposed design project, the paper states an empirical study using chosen software to extract the real role of these computer programs in achieving sustainable design.

**Keywords** Architectural Software · Sustainable Design · Software sustainable potentials · Building Performance & Evaluation Tools · Mono-function Tools · Comprehensive Optimizing tools · Practical Study

## 1 Introduction

Formally, recorded sustainable buildings are extremely limited. The building sector emits 42% of the annual global GHG[1] and 40% of total energy consumption [1]. In addition, this field consumes enormous number of materials and resources and

[1] GHG: Green House Gasses. These gases consist of carbon dioxide, methane, ozone, nitrous oxide, chlorofluorocarbons, and water vapor. They are mainly produced from the excessive burning of fossil fuels, trapping the heat near the Earth's surface and creating "Greenhouse effect" [22]

H. H. Elsharawy (✉)
The German University in Cairo, Cairo, Egypt
e-mail: hebatulla.elsharawy@guc.edu.eg



V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_12

delivers great construction waste. Before time, to deal with this problem the architect bears long way of estimations and various sheets of calculations. Right now, a remarkable number of software and their complementary tools emerge which replace these complexities but unfortunately, a limited number of architects use and deal with them.

Accordingly, this research will review the possibility of achieving sustainable building design by using software programs and their complementary tools. Therefore, the research proposes providing a bundle of software and their tools in the architectural design phase which will facilitate and promote achieving sustainable building. Through this study, it is urged that architects adopt functional software and their complementary tools in the design phase; as an initial step; to produce sustainable buildings.

## 2 Software in Architecture Design

In the beginning, 2D drawings were the only available computational technology used in engineering drawings. Over time, software in architecture field extends to represent the third dimension for simple geometric shapes and fast produce organic forms [2]. Accompanying dynamic development, software not only targets visualizations, representation, and sometimes animation, such as in: 3ds Max, SketchUp, and Rhino 3D, but is also equipped with the true information and object specifications, such as in Building Information Modeling (BIM) software [3]. Due to the wide range of available software programs, it's hard for Architects to specify the exact needed application for the required tasks Fig. 1. Software in Architecture field [4, 5]. One of the revolutionary software in BIM is 'Revit', especially when extra dimensions were added to it to manage time, cost, sustainability, management, energy simulation, and further expected dimensions [6, 7]. Remarkably, Built-in, plug-in tools, and even compatible stand-alone software are successively released. These auxiliaries expand the main program capabilities. It becomes easier to specify, collaborate, evaluate, detect clashes, check performance, manage, maintain, and do more tasks along the whole project lifecycle [8].

Although many architects and firms prefer using the main BIM software, other software proves noticeable capability in specific functions. Non-BIM software and other tools can now do definite and reasonable functions and become able to add on or plugged in, or even integrate with other tools to gather software professionality and BIM audacity. Software competitors in architecture field, such as Graphisoft 'ArchiCAD', Autodesk 'Revit', and Bentley 'AECOsim Building Design' (ABD) extend their vision to track the project life cycle from the initial design stage till its end of life [4, 9]. Meanwhile, tracking has become important for software as a practicing tool, accordingly, computational technology has become an awesome requirement for the project design. Particularly, design stage is the stage that can identify the building lifecycle in advance. The initial design is built in 3D, then plug-ins& add-ons tools are run for specific duties Fig. 2. Furthermore, smart converters


**Fig. 1** Software in Architecture field [4]

(such as IFC[2] and COBie[3]) are utilized to continue in another stand-alone software [10]. The entire software systems are now ready to accomplish the worldwide target of sustainability.

## 3 Software Statistics and Market Share

Regarding architecture field, "Revit, ArchiCAD and AECOsim" are the most used software among architects and professional firms of architecture worldwide. In the Middle East, Revit became the first software used by the participants with 24.77%, then Bentley 10.09%, and finally ArchiCAD was 7.11% [12]. In its "10th Annual Report" the National BIM Library (NBS) stated that "Autodesk Revit" has nearly gained half the market share, and it is spreading fast. It achieved approximately 10% increase in market share in one year [13, 14].

[2] IFC: International Foundation Class. It is a neutral platform, that can be read and edited by any BIM software to enhance their interoperability [21, 20].

[3] COBie: Construction- Operations Building information exchange. It is a standard which defines information about assets used in a project, to help in documenting data for BIM process. These data can be found in.CVS or.XLS files [10].

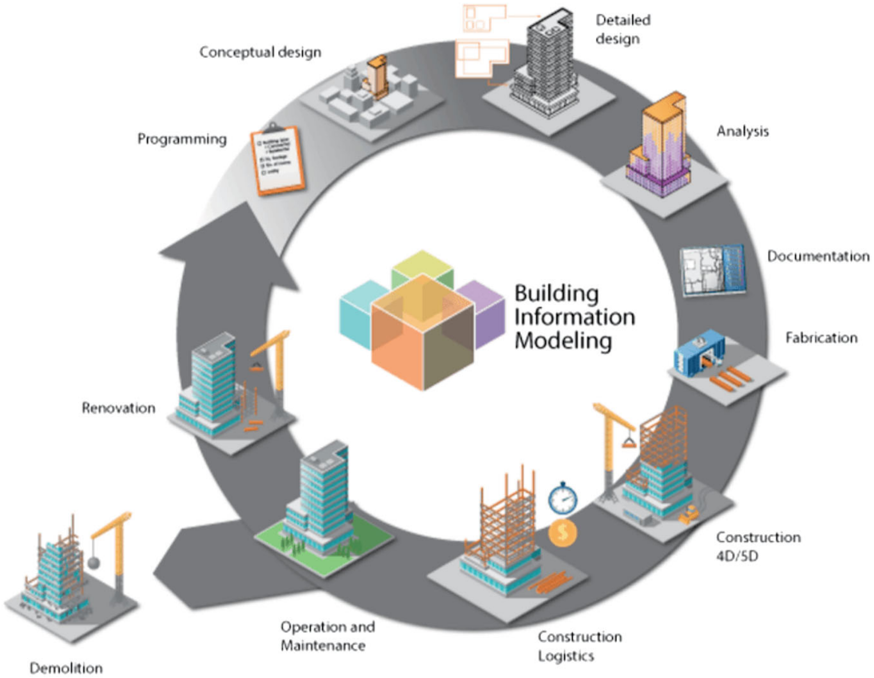



**Fig. 2** Project life cycle through BIM [11]

## 4 Software Sustainable Potentials

The paper studies the three software "Revit, ArchiCAD and AECOsim" software: general data, evolution, and generic potentials, Interface facilities, collaboration and technical interoperability, input options, and output forms and formats. The study reveals that the three main design software mostly share in achieving sustainability targets. Each software goes in a specific scenario to fulfill sustainability requirements in the whole project life cycle. Particularly, their compatible "plugin, add-on and compatible standalone programs" are playing a significant role in getting sustainability starting from the design stage. These compatible tools -whether they are specific or comprehensive work for definite functions such as: enhancing building performance, optimizing energy consumption, reducing environmental impact, and more.

The study also stated that there are common sustainable issues for software used in the architecture field. In addition, few of them go further in more specific issues related to sustainability.

### 4.1 Common Sustainable Issues

Software in architecture field promotes sustainability's main pillars (Environmental, Economic, and Social), and shares in the following [15]:

- As an initial stage, they generate conceptual models, analyze the climatic effect, test form efficiency, improve the design, and get proficient design.
- They work in a real climatic data source such as: "ASHRAE, StruSoft Climate Server, etc." and consequently get actual improvement feedback to the design.
- The final produced 3D model with its final level of development in the design stage gives precise details, save time, money, and effort. Furthermore, it avoids redesigning ambiguous parts of the project from the beginning.
- The availability of working in one platform and upon one model from all the design contributors (Architects, structure engineers 'RC, steel, and prefab', MEP, etc.) assures design integrity and avoids participants' conflict.
- Submitting a visualized 3D model with realized rendered material of the designed project can satisfy client desire and escape confusion.
- The precise description of the used materials, their quantity, manufacture specification, and suppliers help in identifying the whole project cost, and respectively the project viability to the owner.
- The varieties of outputs from different software—between: 2D "plans, elevations, etc.", 3D "whole project model, technical connection and specific details", bill of quantities, time schedules, and even an animated model—ensures compatibility and accuracy of all project documents in the design and post-design phases.
- In general, software allows output compatibility, which makes it flexible to produce free, fluid, and complex forms in one platform afterward they fulfill the design requirement in another one.
- Architecture software has the availability to export their files to "smart converters" (such as IFC, gbxml, etc.), which in return makes benefits of other software and tools. This maximizes applying sustainability functions when transferring the design among defined software.

### 4.2 Specific Sustainable Issues

Some software programs directly deal with these issues within its platform. While other software programs need integrated tools to perform the required tasks. Accordingly, a study of the available software shows immense potential and a wide range of specific functions that can serve definite sustainable points [15] such as:

- Analyzing heating and cooling loads.
- Optimizing heating and cooling systems.
- Reduce energy consumption to the minimum.
- Suggest solutions to increase design performance.

- Assure using certified green materials according to the criteria of “Environmental Product Declaration” (EPD)
- Control ozone depletion and greenhouse gas emissions.
- Assess the life cycle of the material and prevent environmental impact.
- Check conflict between MEP, structure, and architecture objects.
- Analyze environmental and climate effects on the design objects (walls, floors, roofs, and curtain walls.
- Consider topography, landscape, parking areas, etc. to get the minor climate precisely.
- Analyze solar radiation, adjust orientation, and calculate PV panels.
- Enhance passive design scenarios.
- Improve and deliver “Complement” natural light solutions.
- Distribute lighting units and calculate their efficiency.
- Integrate Green materials and products library.
- Identify the design progress due to LOD for different simulation purposes.
- Unify a wide range of tools, to do specific functions.
- Adapt some of the sustainability aspects for the green building rating system (such as: LEED).
- Practice “Hyper Reality” tools to give further possibility for client and society satisfaction.

## 5 Software Complementary Tools

Evaluation tools are used as “Plugins, Add-ons or Standalone” tools for main software. These tools are diverse and available in vast numbers, for example, Autodesk Inc. issued hundreds of plugins, add-ons, and standalone tools that extend main software potentials and cover wide range of requirements in several fields. Particularly, Microsoft produced and supported more than seventy tools to improve the building design performance and sustainability. Subsequently, building performance and evaluation tools are divided into direct and indirect tools relevant to their contribution and their definite task [16].

These complementary tools are also divided into two types which are “Mono-Function Tools” and “Comprehensive Optimizing Tools”. Some tools are responsible for one specific task, whereas others work for several collective tasks Fig. 3.

### 5.1 *Mono-Function Tools*

“Mono-function tools” are classified into three function fields, which regard the whole building and its components, external factors, and indoor spaces. The first field of functions that concern the whole building and its components dealing with Life Cycle Assessment, Passive design, Wind Simulation, Energy Consumption, Fire

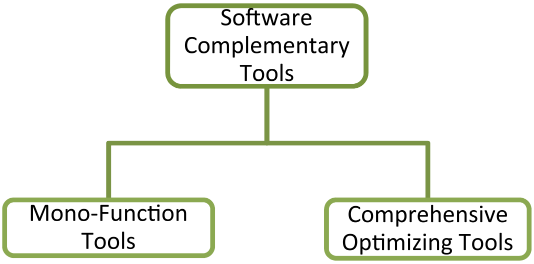



**Fig. 3** Division of simulation complementary tools (author)

Resistance, and Level of Development (LOD). The second field of functions targets external factors including Noise, Thermal Performance, and Green Communities& Buildings. The third field of functions focuses on indoor spaces introducing Interior Lighting, Heating and cooling systems, and Electrical loads [17].

## 5.2 *Comprehensive Optimizing Tools*

On the other hand, Comprehensive Optimizing tools are considered milestones of achieving sustainability in the building field. They play a key role in architectural design phase to avoid inefficiencies of the building along all its phases. The most common comprehensive tools in the building performance field are: Green Building Studio, Sefaira, IES VE, Insight 360, Design Builder, eQUEST, and FormIt (Ecotect & Vasari). Generally, Comprehensive optimizing tools are categorized according to their interoperability with their compatible main software programs or their commercial producers [18, 19].

Six of the optimizing tools: Insight, Green Building Studio, Ecotect, IES VE, eQuest, and Sefaira are compared. The six tools are compatible with REVIT software. The study indicates that the considered tools optimize the design performance with some factors such as energy, passive& active solutions, resources, emissions, and others. All factors analyze the building's performance in the design stage and intensify its sustainability in its upcoming lifecycle stages [15].

Although comprehensive optimizing tools are different in dealing with specific factors, they analyze the following aspects such as: contribute to managing, energy, passive and active solutions, resources, emissions, and others by one or more factors. In the energy section and by applying hybrid tools, a complete analysis is performed, energy consumption, cost, and total amount of energy are calculated, and the energy star scoring is determined. For passive design, some factors can be analyzed such as Natural Ventilation, Thermal Performance, Shadows and Reflections, Daylighting Tracker, Shading Design options, Infiltration Treatments, Orientation Suggestions, Weather and Solar Radiation studies, and the potential of daylighting credit. Additionally, active solutions are considered by these tools to calculate wind energy, photovoltaic installations, and suppose renewable energy plans. The last section of the comprehensive optimizing tools accommodates the analysis of: Carbon-Emissions Rates, Water Use Amount, Visual Impact Representation, Acoustic Analysis, HVAC Alternatives, User Operating Schedules, Lighting Efficiency, and Plug Loads Counting. As a result, the project design can benefit from all factors if hybrid tools are applied [15].

Overall, Complementary tools delve into specific functions to optimize the performance and enhance the sustainable architectural design.

## 6 Empirical Study

In this part, the research evaluates the role of software programs in achieving sustainable design. The chosen software and tools in this study are “Climatic Consultant, Revit, and Insight” Fig. 4. The practical study concentrates on designing a twin villa in New Capital City, as this typology consumes huge amount of energy and releases emissions. The design process goes through successive steps as follows:

### *6.1 Suggest Passive Design Solution*

The design process begins by applying Climatic Consultant program to suggest general passive solutions for the project prototype. In this step, the most convenient solutions are selected to be applied in the proposed design such as cross ventilation, adding outside planting area, using shades and light color materials in facades, and more Fig. 5.

### *6.2 Planning and Build the 3D-Model for Design Prototype*

In this step, the design sketches are prepared, then the proposed design prototype is built in “REVIT” software. All building elements are built with their real dimensions

| | | |
|---|---|---|
| C6 | AUTODESK REVIT | AUTODESK INSIGHT 360 |
| Climate Consultant tool | Revit software | Insight tool |

**Fig. 4** Software and tools used in practical study

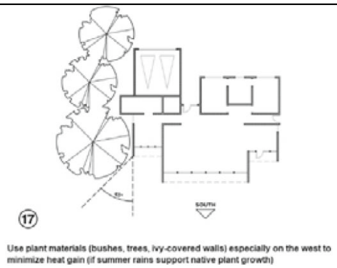
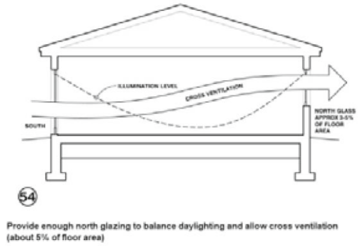
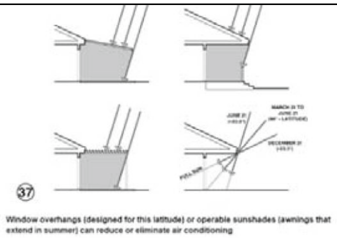
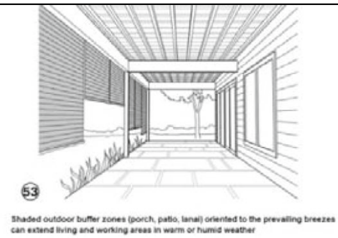
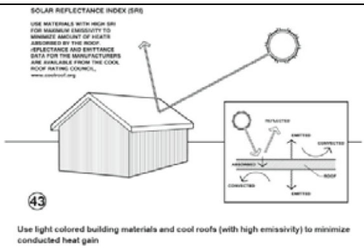
| | |
|---|---|
|  |  |
| Planting solution to reduce the heat gain | Cross Ventilation |
|  |  |
| Window shades can be just an extrusion over the window | Shaded outdoor Zones can be used as buffer zones |
|  |  |
| The usage of light-colored materials on the "Roof" to reduce the heat gain | Summary of the suggested Design Solutions, according to Ashrae 2005 |

**Fig. 5** Climatic consultant passive solutions (author)

and materials in their exact position. Ground, Typical, and Roof floor layers are specified. Exterior and Interior wall layers are chosen. Structural columns are determined in accordance with material, dimensions, and distribution. Structural stairs are also calculated and built in the model. Doors, windows, plumbing fixtures, and electrical lighting & Chandelier are added to the 3D model Fig. 6.

After fulfilling the stage of building the 3D model and completing all needed information, the model becomes a source of data and set for multiple outputs. In addition to the ability to walk through the model by using the suitable complementary software tools, working drawings with all specific details are available for printing. "Bill of Quantities (BOQ)" as well as specification tables of all the design elements

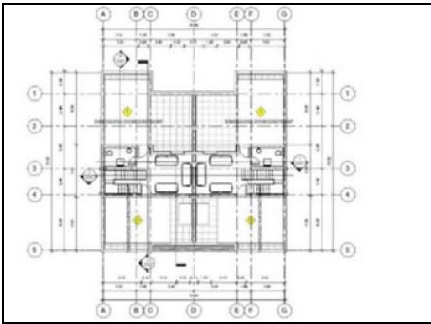
|  |  |
|---|---|
| Roof floor plan | 3D Model |

**Fig. 6** Proposed design (author)

and components are applicable Fig. 8. The overall cost of the project is produced and revised to achieve great extent of accuracy. The used software tracks shade and shadows along the 24h of the day and reports the amount of daylight that penetrates through the openings of each façade, regarding short, medium, and long sun paths in different seasons Fig. 7.

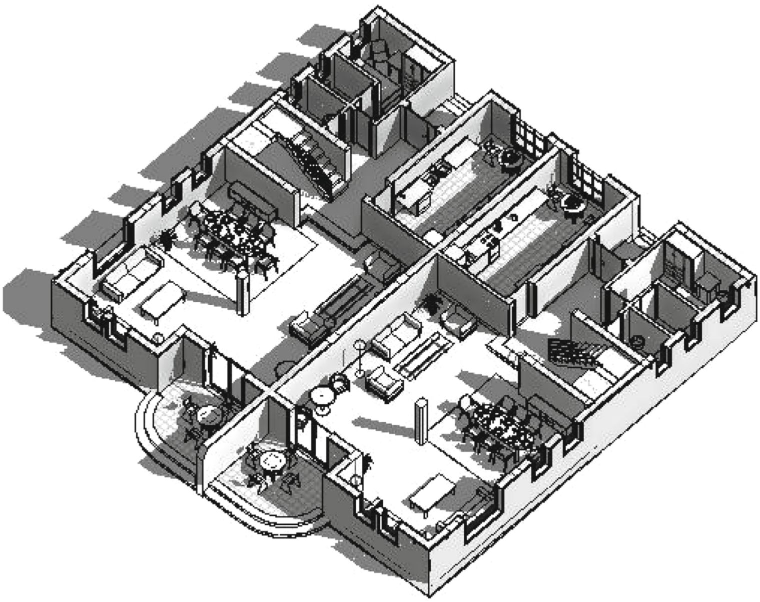
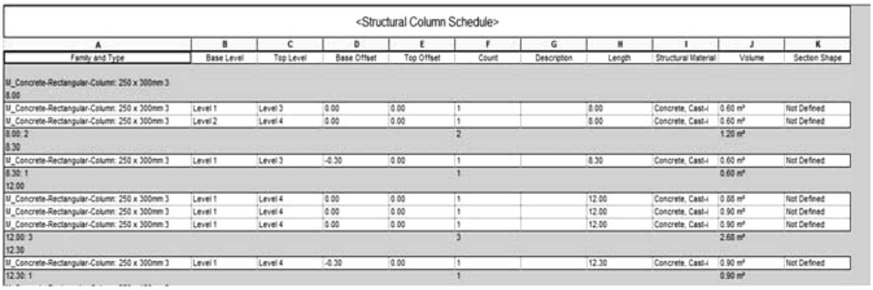


**Fig. 7** Daylighting and shadows analysis (Author)



<Structural Column Schedule>

| A | B | C | D | E | F | G | H | I | J | K |
|---|---|---|---|---|---|---|---|---|---|---|
| Family and Type | Base Level | Top Level | Base Offset | Top Offset | Count | Description | Length | Structural Material | Volume | Section Shape |
| M_Concrete-Rectangular-Column: 250 x 300mm 3 | | | | | | | | | | |
| 8.00 | | | | | | | | | | |
| M_Concrete-Rectangular-Column: 250 x 300mm 3 | Level 1 | Level 3 | 0.00 | 0.00 | 1 | | 8.00 | Concrete, Cast-i | 0.60 $m^3$ | Not Defined |
| M_Concrete-Rectangular-Column: 250 x 300mm 3 | Level 2 | Level 4 | 0.00 | 0.00 | 1 | | 8.00 | Concrete, Cast-i | 0.60 $m^3$ | Not Defined |
| 8.00: 2 | | | | | 2 | | | | 1.20 $m^3$ | |
| 8.30 | | | | | | | | | | |
| M_Concrete-Rectangular-Column: 250 x 300mm 3 | Level 1 | Level 3 | -0.30 | 0.00 | 1 | | 8.30 | Concrete, Cast-i | 0.60 $m^3$ | Not Defined |
| 8.30: 1 | | | | | 1 | | | | 0.60 $m^3$ | |
| 12.00 | | | | | | | | | | |
| M_Concrete-Rectangular-Column: 250 x 300mm 3 | Level 1 | Level 4 | 0.00 | 0.00 | 1 | | 12.00 | Concrete, Cast-i | 0.88 $m^3$ | Not Defined |
| M_Concrete-Rectangular-Column: 250 x 300mm 3 | Level 1 | Level 4 | 0.00 | 0.00 | 1 | | 12.00 | Concrete, Cast-i | 0.90 $m^3$ | Not Defined |
| M_Concrete-Rectangular-Column: 250 x 300mm 3 | Level 1 | Level 4 | 0.00 | 0.00 | 1 | | 12.00 | Concrete, Cast-i | 0.90 $m^3$ | Not Defined |
| 12.00: 3 | | | | | 3 | | | | 2.68 $m^3$ | |
| 12.30 | | | | | | | | | | |
| M_Concrete-Rectangular-Column: 250 x 300mm 3 | Level 1 | Level 4 | -0.30 | 0.00 | 1 | | 12.30 | Concrete, Cast-i | 0.90 $m^3$ | Not Defined |
| 12.30: 1 | | | | | 1 | | | | 0.90 $m^3$ | |

**Fig. 8** Structural column detailed schedule (Author)

## *6.3 Improving Design Performance*

Regarding simulating solar paths and tracking building shadows within the platform of Revit, the initial design of the prototype shows essential need for using shading elements for the west façade. While east and south facades demand convenient shading elements. Solar loads are also tested, and shading pergolas are proposed to modify heavy loads. The amount and ratio of natural lighting inside the space is checked and as a result areas of windows are adjusted. Heating and cooling loads are calculated for each surface and space Fig. 9. Peak of cooling and peak of heating are identified to be managed upon further solutions.

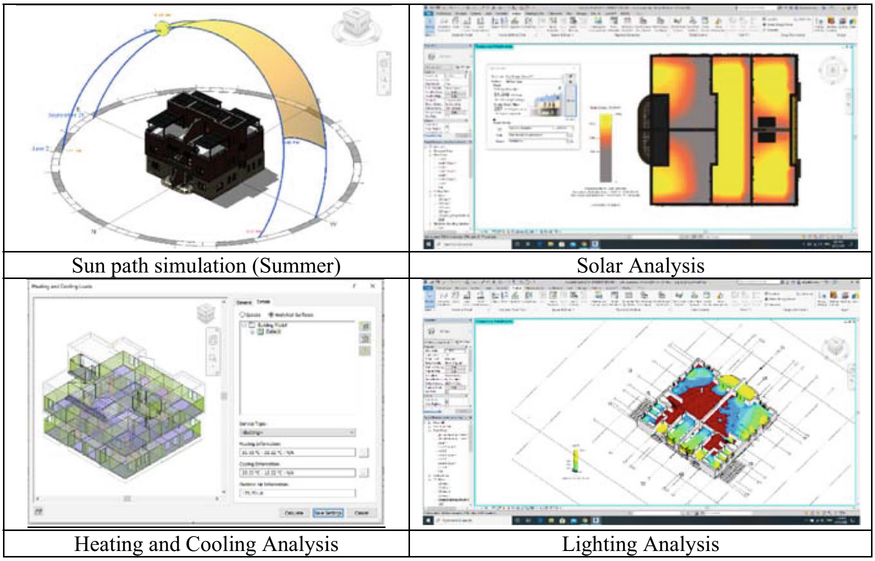



**Fig. 9** Main program analysis output (Author)

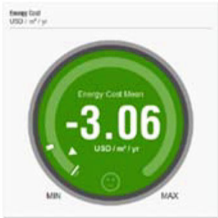
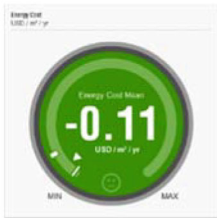
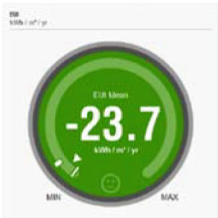
| Aspects | 3D Model scenario | Net Zero Energy scenario | Final adopted scenar-io |
|---|---|---|---|
| Cost |   |   |   |
| Energy consump-tion |   |   |   |

**Fig. 10** Final adopted scenario_ cost and energy consumption (author)

## *6.4 Control Energy Consumption and Cost*

After fulfilling the analysis and applying the suggested solution in the previous step, energy consumption and cost take place within the complementary tool “Insight”. This software applies “Ashrae 90.1 and Architecture 2030 codes and standards” to produce different scenarios reducing energy cost& consumption. The tool works automatically on the improved 3D-Model, where calculations for energy consumption and cost depend primarily on the “mean value” concept Fig. 10.

The practical study in this step goes through several scenarios by adjusting the software widgets to different alternatives: actual 3D-model scenario, Net Zero Energy (NZE) scenario, and Author adjusted scenario. The actual 3D Model scenario represents the real data extracted from the 3D model. While the NZE scenario shows the possible targets that can be reached within the 3D model. From both scenarios, the Author’s final adjusted scenario targets fulfilling the best achievable outcomes. In these three scenarios Insight tool gives detailed cost and energy consumption calculations for the following widgets: Plug load efficiency, Operating schedule time, Lighting efficiency, HVAC types, Roof construction kinds, windows to wall ratio for the four facades “north, east, west, and south”, the construction layers of roof and walls, building orientation, types of glass for windows, shading elements, Daylighting & occupancy controls, infiltration, and finally PV. “Panel efficiency, Payback limit, Surface coverage”. After running the program and getting results, the “author adjusted solutions” presents the best scenario with (-0.11 USD/m2/ yr.- for energy cost) and (29.8 kwh/m2/yr.- for energy consumption). The adjusted scenario saves 5.91 USD/m2/ yr. in cost, and 108.2 kwh/m2/yr. in energy consumption Fig. 11.

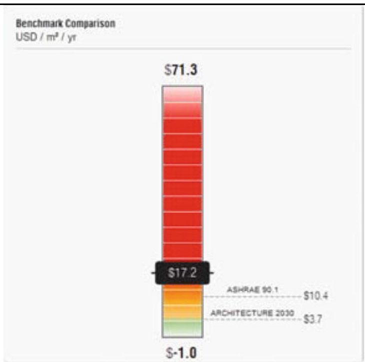
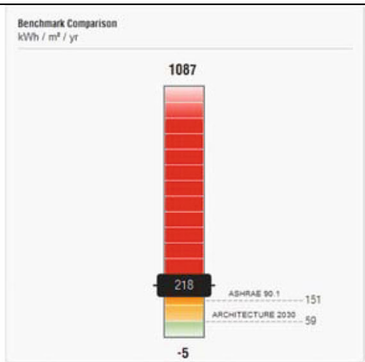
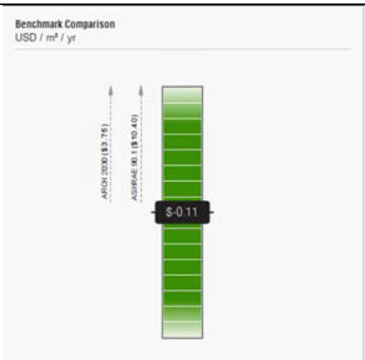
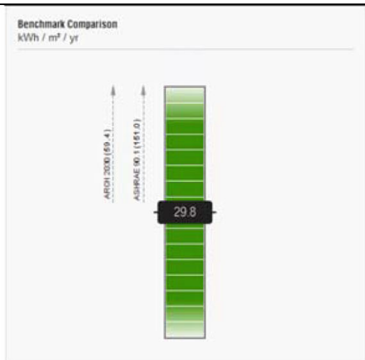
<table>
<tr>
<td>
</td>
<td>
</td>
</tr>
<tr>
<td><u>Default scenario:</u> the estimated cost (17.2 USD/m2/yr.) is higher than both of ASHRE90.1 (10.4 USD/m2/yr.) & ARCHITECTURE2030 (3.7 USD/m2/yr.) recommended benchmarks</td>
<td><u>Default scenario:</u> energy consumption equals (218 kWh/m2/yr.) which is higher than both of ASHRE90.1 (151 kWh/m2/yr.) & ARCHITECTURE2030 (59 kWh/m2/yr.) recommended benchmarks</td>
</tr>
<tr>
<td>
</td>
<td>
</td>
</tr>
<tr>
<td><u>Final adopted scenario:</u> the estimated cost in this case is (0.11 USD/m$^2$/yr) which is accepted as it is in-between ASHRE90.1 & ARCHITECTURE 2030 recommended benchmarks</td>
<td><u>Final adopted scenario:</u> the evaluated energy consumption (29.8 kWh /m$^2$/yr) is approved as it provides reasonable value on ASHRE90.1 & ARCHITECTURE 2030 recommended benchmarks</td>
</tr>
</table>

**Fig. 11** Default and final adopted scenarios_ cost & energy consumption results

## 7 Discussion and Findings

The paper studies reveal that software programs have noticeable potential in producing sustainable designs and promoting sustainability pillars (Environmental and Economic), while social pillar depends mainly on the design and user satisfaction. They offer design projects in real climatic data, avoid wasting time, money, and effort, or even go into a loop of redesigning. Using software in the design process offers working in one platform from all contributors, avoids conflict, confirms client satisfaction, and assures accuracy in cost, quantities, specifications, and furthermore. Software through their complementary tools, are also dealing with analyzing and optimizing heating and cooling loads, reducing energy consumption and cost, calculating PV panels needed, distributing lighting units, and calculating their efficiency, even software adapts sustainability aspects for the green building rating system certification.

On the other hand, the practical part of the study proves that software contributes to achieving sustainable design. This contribution varies between direct contribution and indirect support for sustainable criteria, where the latest refers to the design solutions itself. Some sustainable design criteria such as social criteria are difficult to be evaluated or improved depending on software. Transferring from one software to another in this practical study shows incomplete interoperability between the three used programs in addition to the different building standards used in each of them. Overall, the practical results verify the previous results of study regarding the strong contribution of software in achieving sustainable design.

The prototype design went through several stages of modification and improvement to get sustainable design. The design of the prototype begins with managing orientation and applying passive solutions offered by the first climatic software. The second software in addition to the build 3D-model and all benefits associated with it, improves the design by adding shading elements to the roof and facade openings depending on simulating solar path, tracking model shadow, and projecting solar load on roof. It also modifies opening size after checking the amount of penetrating natural light inside the proposed design, then it sets HVAC calculations. The third complementary software deals directly with specific sustainability issues. It gives several solutions to adjust the design automatically, then it allows user to modify the examined issues. The result of using this software is enhancing the performance of the design and minimize energy and cost consumption. (–0.11 USD/m2/ yr.- for energy cost) and (29.8 kwh/m2/yr.- for energy consumption).

Finally, the paper concludes that there is a great contribution of the used software in achieving sustainability and improving the design performance during its early stages. They also can enhance the design to match the requirements of green and sustainable rating systems.

## References

1. Architecture2030: Why the built environment? [Online]. Available: https://www.architecture2030.org/why-the-built-environment/ (2023). Accessed 06 October 2024
2. Tailoor: How we went from 2D to 3D visualization Tailoor [Online]. Available: https://tailoor.com/how-we-went-from-2d-to-3d-visualization/ (2024, April 01). Accessed 15 September 2024
3. Bentley: Bentley: Contact Us. Bentley Systems [Online]. Available: https://www.bentley.com/en/about-us/contact-us (2020). Accessed 13 August 2024
4. Augmintech: What is BIM Software? [Online]. Available: https://augmintech.com/what-is-bim-software/ (June, 2023 08)
5. Boiko, A.: BIM History Map. Big Data [Online]. Available: https://bigdataconstruction.com/history-of-bim/ (2024). Accessed 15 September 2024

6. Jordan, M.: A Brief History of 3D Software in Architecture [Online]. Available: https://www.linkedin.com/pulse/brief-history-3d-software-architecture-malik-jordan/ (2023, August 16). Accessed 15 September 2024
7. Autodesk: Multidisciplinary BIM software for higher-quality, coordinated designs. [Online]. Available: https://www.autodesk.com/products/revit/overview (2020). Accessed 24 July 2024
8. Cherkaoui, H.: A history of BIM. [Online]. Available: https://www.letsbuild.com/blog/a-history-of-bim (2017, March 22). Accessed 28 July 2024
9. Bentley: AECOsim Building Designer—Version History [Online]. Available: https://communities.bentley.com/products/building/building_analysis_design/w/building_analysis_and_design_wiki/25239/aecosim-building-designer-version-history (2020). Accessed 27 July 2024
10. IBM: Construction-Operations Building information exchange (COBie) standard data [Online]. Available: https://www.ibm.com/docs/en/mam-saas/7.6.0.9?topic=bim-cobie-standard-data (2021, March 05). Accessed 22 September 2024
11. BIMMDA: What is BIM [Online]. Available: https://bimmda.com/en/what-is-bim (2024)
12. Gerges, M., Austin, S., Mayouf, M., Ahiakwo, O., Jaeger, M., Saad, A., El Gohary, T.: An investigation into the implementation of building information modeling in the middle East. J. Inf. Technol. Constr **22**, 1874–4753 (2017)
13. NBS: National BIM Report 2019/ The definitive industry update. NBS, Newcastle upon Tyne (2019).
14. Autodesk: REVIT Software for Construction, Fabrication & Contractors. Autodesk [Online]. Available: https://www.autodesk.com/products/revit/construction?plc=RVT&term=1-YEAR&support=ADVANCED&quantity=1 (2020). Accessed 25 July 2024.
15. Elsharawy, H.H.: The Role and Integration of Simulation Models in achieving Sustainable Architectural Design. The German University in Cairo, Cairo (2021)
16. A. Inc: ADSK Sales vs. its Competitors Q2 2020 [Online]. Available: https://csimarket.com/stocks/compet_glance.php?code=ADSK (2020). Accessed 18 August 2024
17. Fahmy, D.A.E.: A guiding approach to make sustainable design decisions by Building Information Modelling (BIM). Int. J. Archit. Eng. Urban Res., 167–180 (2023)
18. Autodesk: Building performance analysis: plugin, Add-on, Extension for Revit [Online]. Available: https://apps.autodesk.com/RVT/en/List/Search?facet=_category%3a%3aBuilding+Performance+Analysis&page=1 (2011–2020). Accessed 30 August 2024
19. Autodesk: Autodesk: Insight "Building Performane Analysis Software" [Online]. Available: https://www.autodesk.com/products/insight/overview (2020). Accessed 18 August 2024
20. Autodesk: About the IFC File Format. Autodesk [Online]. Available: https://help.autodesk.com/view/RVT/2025/ENU/?guid=GUID-0D546BEA-6F88-4D4E-BDC1-26274C4E98AC (2024). Accessed 06 October 2024.
21. buildingSMART: Industry Foundation Classes (IFC). Dassault Systemes [Online]. Available: https://technical.buildingsmart.org/standards/ifc/ (2024). Accessed 06 October 2024.
22. Nasa: What is the greenhouse effect? Nasa Science [Online]. Available: https://science.nasa.gov/climate-change/faq/what-is-the-greenhouse-effect/ (2024, April). Accessed 06 October 2024.

# Green Coding Techniques and Methodologies

# Investigating the Use of GitHub Copilot for Green Software

**Maria Stivala, Iffat Fatima, and Patricia Lago**

**Abstract** This study empirically investigates the use of GitHub Copilot in developing energy-efficient software and the associated trade-offs. We compare the energy consumption and performance of human-written code to (i) the human-written code optimized by using GitHub Copilot, and (ii) code generated by GitHub Copilot. A set of 15 programming scenarios is used to test our hypothesis. GitHub Copilot is prompted for energy-efficient solutions. Our results demonstrate that GitHub Copilot-generated code can achieve significantly lower energy consumption when optimized, without compromising code quality. The study did not find statistically significant differences in energy consumption compared to human-written code. However, it highlights the potential of AI-assisted coding tools such as GitHub Copilot for developing energy-efficient software and the importance of explicitly prompting for energy consumption.

**Keywords** Software development · Sustainability · Green coding · Energy consumption · GitHub Copilot · AI-generated · AI-assisted

## 1 Introduction

Software development is continually evolving with technological advances and industry changes. The rise of AI-powered developer tools, such as GitHub Copilot, is revolutionizing how developers write code.[1]

[1] https://clickup.com/blog/software-engineering-trends/.

M. Stivala · I. Fatima · P. Lago (✉)
Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
e-mail: p.lago@vu.nl

M. Stivala
e-mail: m.stivala@student.vu.nl

I. Fatima
e-mail: i.fatima@vu.nl



V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_13

Traditionally, developers used techniques like *pair programming* to enhance code quality. Although pair programming improves code quality, it requires synchronous collaboration, making it difficult to scale [2]. GitHub Copilot builds on this by offering *AI pair programming* [4], which directly assists developers by generating code snippets, providing context-aware suggestions, and accelerating development workflows. This tool aims to boost productivity and democratize coding expertise [20]. According to Stack Overflow's 2023 survey of 90,000 developers and technologists, 44% have integrated AI tools into their workflow, with 83% using them mainly for code generation [19].

The rise in society's digitalization has increased the demand for computing resources, leading to high energy consumption and carbon emissions [23]. Given the growing environmental concerns, optimizing software energy use has become crucial. A major challenge in developing energy-efficient software is to minimize energy consumption while maintaining functionality [17].

To build energy-efficient software, techniques such as parallel and approximate programming, static source code analysis, efficient data structures, energy-efficient coding practices, and specific programming languages are being used [3]. Energy-efficient software not only contributes to reducing the carbon footprint, but also has the potential to enhance the performance of computing systems. Hence, the pursuit of energy consumption in software design and implementation has become a critical objective for developers, organizations, and society.

While the impact of the GitHub Copilot on productivity is a subject of ongoing investigation [15], its influence on other aspects, particularly energy consumption, remains largely unaddressed. Existing studies on the use of GitHub Copilot and other AI in software development have a primary focus on code quality [24] and functionality [10], overlooking the energy implications of the developed piece of software.

Within this context, this research seeks to explore the intersection of GitHub Copilot-generated code and energy-efficient software, aiming to uncover novel pathways for AI-driven tools in supporting sustainability efforts in the Information Technology industry.

To achieve this goal, we hypothesize that GitHub Copilot can significantly contribute to the development of energy-efficient software by suggesting code optimizations that reduce energy consumption. By exploring this potential, this study can provide insights and practical guidance to software developers seeking to create more sustainable applications. To achieve this, we define the following research questions (RQs).

**RQ1:** How does the energy consumption of code generated by GitHub Copilot compare to human-written code?

**RQ2:** To what extent (in terms of percentage change) can GitHub Copilot reduce the energy consumption of human-written code, when prompted?

**RQ3:** What trade-offs exist between code quality, performance, and energy consumption, when employing GitHub Copilot?

We answer these RQs by designing and executing an experiment to compare the trade-offs between different metrics for performance, code quality, and green software.

## 2 Related Work

GitHub Copilot represents a significant advancement in code completion tools, offering developers intelligent suggestions and context-aware recommendations. This overview of related literature explores the impact and effectiveness of GitHub Copilot in various scenarios, with a particular focus on its implications for energy-efficient software.

### *2.1 GitHub Copilot and Energy Efficiency*

The work by Vartziotis et al. [22] directly addresses the use of Large Language Models (LLMs), such as GitHub Copilot, in creating code that optimizes energy consumption alongside functionality. Their framework for measuring 'green code capacity' is particularly relevant, providing a foundational approach that aligns with the focus of this research. Thus, their study serves as a baseline for understanding how AI-driven code suggestions could be geared toward more sustainable software practices.

While not directly focused on AI-assisted development, Lago et al. [8] offer a comprehensive review of metrics for green software, including energy consumption, among other factors.

However, the existing literature lacks an evaluation of the trade-offs between energy consumption and other software qualities when using GitHub Copilot. This research aims to bridge this gap by systematically analyzing the impact of Github Copilot on energy consumption while also considering technical factors.

### *2.2 Code Quality and Energy Efficiency*

Imai [6] compares the effectiveness of GitHub Copilot in code development to traditional pair programming. The study reveals that while GitHub Copilot facilitated faster development, it often resulted in lower code quality. A tool that increases development speed may still require additional debugging and refactoring later, which can indirectly increase energy consumption, as outlined by Ournani et al. [12]. This indicates that AI-based suggestions may not consistently optimize code quality, leading to code smells, which could indirectly influence energy efficiency [13].

Yetistiren et al. [24] and Nguyen and Nadi [11] discuss the correctness and consistency of GitHub Copilot's code suggestions across various programming languages. Although not directly addressing energy efficiency, Yetistiren et al. provide insights on the importance of considering code quality. This underscores the importance of considering code quality when assessing the impact of GitHub Copilot on energy efficiency. The findings of Nguyen and Nadi highlight potential inconsistencies that could lead to less optimized, or more energy-consuming, code, stressing the importance of considering code quality in discussions about energy efficiency.

### *2.3 Technical and Environmental Considerations*

Pareira et al. [16] investigate the relationship between energy consumption, execution time, and memory usage in different programming languages. Understanding this relationship is crucial for analyzing the impact of GitHub Copilot on energy consumption, since the language chosen for the code generation process may inherently influence these factors.

Siavvas et al. [18] explore the potential connection between software security and energy consumption. They suggest that security vulnerabilities can sometimes lead to increased energy consumption. Furthermore, Pearce et al. [14] highlight potential security vulnerabilities that can be introduced through GitHub Copilot code suggestions. Thus, this may lead to increased energy use if additional computational resources are required to manage these risks.

Existing literature emphasizes the need for a holistic approach that considers trade-offs between environmental (energy consumption) and technical (code correctness, security, etc.) metrics when evaluating the impact of GitHub Copilot.

While Vartziotis et al. focus on the green capacity of code directly generated by LLMs, our work takes a different approach. Instead of evaluating the sustainability of AI-generated code, we concentrate on optimizing existing human-written code for sustainability through AI. This shift in focus allows us to leverage the strengths of human problem-solving and AI optimization. By introducing a sustainability metric optimization condition, we aim to bridge the gap between human code and AI-driven optimization, exploring the potential for significant sustainability improvements in existing software systems. Our approach differs from previous studies [22] that mainly focused on the green capacity of AI-generated code, offering a new perspective on the intersection of AI and sustainable software development. With this study, our aim is to explore how GitHub Copilot influences the energy consumption of the software it helps to develop, considering potential trade-offs.

# 3 Study Design and Execution

## 3.1 Research Approach

We employ an empirical approach to assess the impact of GitHub Copilot-generated code on energy consumption compared to human-written code. This enables the gathering of robust and quantifiable data to objectively assess the energy consumption of software developed with GitHub Copilot, thus directly addressing the research questions mentioned previously. We provide the data and code as a replication package online [21].

## 3.2 Data Acquisition

We leverage the publicly available HumanEval dataset [7] as a source of programming problem statements and their human-written solutions. The solutions provided are in Python. Related work shows that Vartziotis et al. [22] uses the top 0.05% LeetCode submissions as a baseline, the use of the publicly available dataset HumanEval was more optimal due to its diverse range of problems, mimicking real-world coding scenarios. To ensure a representative sample for analysis, the data set was categorized into three different types of problems: string manipulation, mathematical operations, and data structure and algorithms. We selected five problems within each category, resulting in 15 problems for evaluation. This approach ensures a representative sample while maintaining consistency with prior research methodologies [22]. The stratified random sampling approach aims to maintain a balanced representation of problem types, minimizing potential bias in the evaluation of GitHub Copilot's performance. The selection criteria prioritize problems with varying levels of complexity, measured using cyclomatic complexity, to offer a comprehensive assessment of GitHub Copilot's capabilities across the spectrum of coding tasks encountered in real-world development. In this study, we use a small data set that may not capture real-world programming task variability. Future studies could expand the scope by incorporating a larger and more diverse set of problems, potentially from other datasets or domains.

## 3.3 Experiment Design

We designed our experiment to compare human-written code, GitHub Copilot-generated code, and human-written code optimized by GitHub Copilot (see Fig. 1). We use the HumanEval dataset [7] as a source of programming problem statements and their human-written solutions as code. We used three versions of the code for the experiment. (i) Human-written code, (ii) GitHub Copilot-generated code based on

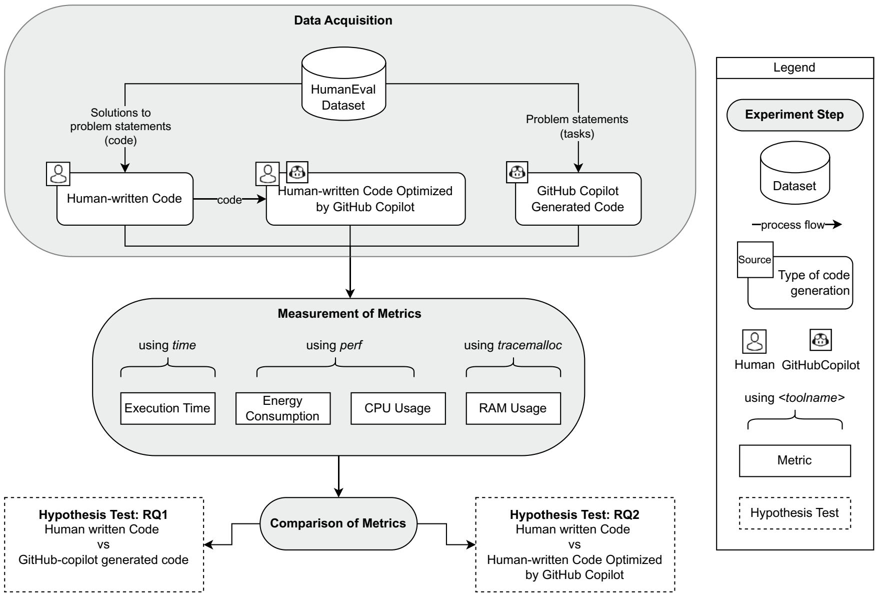



**Fig. 1** Experiment design

the problem statements from the HumanEval database, and (iii) Human-written code optimized by GitHub co-pilot by prompting it to generate an energy-efficient solution. These three versions of code are evaluated using metrics such as time, energy consumption, CPU usage, and RAM usage. Finally, hypothesis tests (for RQ1 and RQ2) compare the performance of human-written code, GitHub Copilot-generated code, and optimized human-written code across these metrics. Figure 1 shows our study design.

**Hypotheses** We present the hypotheses for both RQs to guide our experiment.

**Hypothesis for RQ1** To answer RQ1, we compare the average energy consumption of human-written code with GitHub Copilot-generated code. We aim to determine whether there is a statistically significant difference in energy consumption based on the code's source. To test this, a null hypothesis is established that asserts that there is *no statistically significant difference in the average energy consumption between the two code origins (Human-written and GitHub Copilot-generated)*

$$H_0 : \mu_{\text{human}} = \mu_{\text{Copilot}}$$

The alternative hypothesis is a two-tailed hypothesis asserting that *there is a statistically significant difference in the average energy consumption between the two code sources*

$$H_1 : \mu_{\text{human}} \neq \mu_{\text{Copilot}}$$

**Hypothesis for RQ2** To answer RQ2, we assess whether explicitly requesting energy-efficient solutions from GitHub Copilot improves the energy consumption of the generated code while maintaining overall quality and performance. It compares the energy consumption of human-written code (initial) to that of human-written code optimized by GitHub Copilot for energy consumption (optimized).

The null hypothesis tests whether *there is no statistically significant difference in the average energy consumption between the two code sources (Human-written and optimized Human-written)*

$$H_0 : \mu_{\text{initial}} = \mu_{\text{optimized}}$$

The alternative hypothesis is a one-tailed hypothesis that suggests *the mean energy consumption is statistically significantly lower for the code generated by the optimized human-written solution than for the initial Human-written solution*

$$H_1 : \mu_{\text{initial}} > \mu_{\text{optimized}}$$

## 3.4 Code Generation

We activate the GitHub Copilot code suggestions by adding each problem's task description from the HumanEval dataset's `prompt` property as a comment within the code. This prompts Copilot to recommend code snippets to solve the tasks. Human-written solutions from the `canonical_solution` property are placed in Python files containing the function definition from the `prompt`. The files are named according to the task number, with `_human` added for human-written solutions and `_human_opt` for optimized versions.

We used a single prompt "Make the code as energy-efficient as possible" to ensure consistency in the optimization approach throughout the experiment. We run unit tests using the `test` property from the HumanEval data set to ensure the correctness of the generated solution. We collect the initial human-written code and the optimized final code generated by Copilot for further analysis and provide the data in the online replication package [21].

The initial human-written code along with the final versions of the GitHub Copilot-generated code and the optimized human-written code are collected for further analysis. We provide the data in the online replication package [21].

## 3.5 Experimental Setup and Measurement

We generate code within a controlled environment to isolate the runtime environment and ensure that the suggestions of GitHub Copilot are contextually relevant.

The development environment is configured within a controlled server environment on a GL2, SuperMicro 813M-4 server, Intel(R) Xeon(R) CPU E3-1231 v3 @ 3.40GHz 8vCPUs, 32GB RAM, and Ubuntu 22.04 operating system. To ensure an appropriate standardized environment, we followed phases 2 and 3 of Mancebo et al.'s work [9], 'Measurement Environment Setting' and 'Measurement Environment Preparation'. We meticulously measure the energy consumption for initial human-written code, initial GitHub Copilot-generated code, and GitHub Copilot-optimized human-written code. This comparison among different versions addresses RQ1 and quantifies the effectiveness of energy-efficient prompting for RQ2.

## *3.6 Measurement of Metrics*

We perform metric measurements to evaluate the trade-offs between code quality, performance, and energy consumption.

**Code Quality**. We evaluate the quality of the code using the cyclomatic complexity of the code and the maintainability index using Radon.[2] We chose Radon because it runs seamlessly within the Python environment without requiring special adaptations and does not introduce additional runtime overhead that could affect energy consumption measurements. The command `$ radon cc filepath -s` shows the complexity score along with its rank, ranging from A (low risk) to F (very high risk). The command `$ radon mi filepath` displays the maintainability index score, where a score greater than 20 indicates very easy maintenance and a score below 9 indicates very hard maintenance.

**Code Performance**. Performance metrics include execution time, memory, or resource usage. The deterministic profiling tool *cProfile*, a built-in Python module, precisely measures the execution time of each function call, providing accurate data on the performance bottlenecks of the code. In comparison to other profiling tools, it has a relatively low overhead, minimizing any interference with the code's energy measurements.

**Green Metrics**. We assess the environmental impact of the code using a selection of green metrics described by Lago et al. [8] and elements from the Green Software Measurement Model (GSMM) [5]. The chosen metrics are Execution Time, Energy Consumption, CPU Usage, and RAM Usage. Ten sample runs are executed for each green metric to account for potential variations, ensuring the reliability of the measurements.

We measure *Execution time* using `perf`, a performance analysis tool for Linux environments, and the `time` module, a Python utility that measures execution time. We perform multiple runs and compute the average execution time to ensure accuracy and mitigate variability in execution time measurements.

We measure *Energy Consumption* by profiling the execution of code using `perf`, in particular, the `power/energy-pkg/` event. This reflects the system-wide

[2] https://radon.readthedocs.io/en/latest/.

energy use. To mitigate the sensitivity of the tool in measurement, the minimization of background tasks and repeated code executions is performed.

We measure *CPU Usage* using `perf` allowing for monitoring CPU Usage for the running process. The average CPU usage percentage over multiple sample runs is captured to ensure accurate measurement.

We measure *RAM Usage* during code execution using the memory analysis tool `tracemalloc`. The peak memory usage, or the average memory usage across multiple runs, is obtained through the `get_traced_memory()` function and is documented for consistency. The tracing begins before the code is run, outputting the peak memory usage in bytes after it is completed.

## 4 Results and Discussion

In this section, we analyze the evaluation results in different code samples. We present the results for each hypothesis test for each RQ. First, we compare the green metrics for the GitHub Copilot-generated code and the human-written code. Second, we explore potential improvements within the human-written code by using GitHub Copilot's assistance. Lastly, we discuss the implications of using GitHub Copilot, considering multiple factors beyond energy consumption.

**RQ1: How does the energy consumption of the code generated by GitHub Copilot compare to human-written code?**

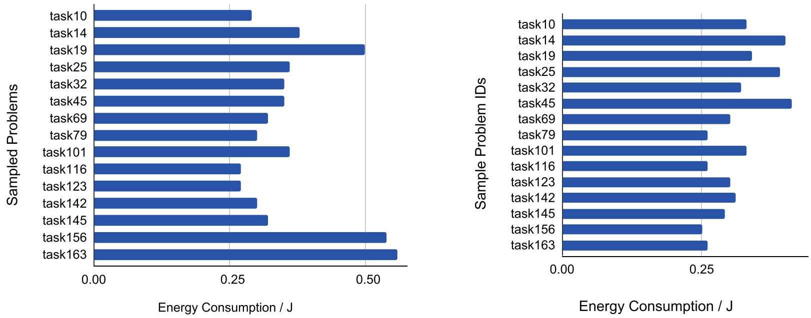
In addressing RQ1, we conduct a comparative analysis of the energy consumption between GitHub Copilot-generated code and Human-written code. Table 1 shows the average values of the metrics obtained from multiple sample runs.

The Shapiro-Wilk normality test indicates that the energy consumption for human-written code was not normally distributed, necessitating a non-parametric approach. Thus, statistical analysis is performed using the Mann-Whitney U test. This test revealed a U-statistic of 146.5 and a p-value of 0.1636. These results do not indicate statistically significant differences in energy consumption, at the significance level 10%, between the two types of code.

For this analysis, a significance level 10% was selected due to several considerations. Given the exploratory nature of this study, a less stringent criterion is appropriate. This allows for the identification of potential trends and preliminary insights. Additionally, the small sample size, consisting of only 15 programming problems,

**Table 1** Green metrics: human-written code versus GitHub copilot-generated code

| Green metric | Human-written code | GitHub copilot-generated code |
|---|---|---|
| Execution time | 87.06 s | 59.39 s |
| Energy consumption | 0.41 J | 0.32 J |
| CPU usage | 0.31% | 0.53% |
| RAM usage | 1.74E-06% | 1.76E-06% |




(a) Human-written Code (b) GitHub Copilot-generated Code

**Fig. 2** Energy consumption

requires a more flexible threshold to account for the increasing risk of false negatives. This approach provides a more flexible threshold for detecting possible differences in energy consumption.

However, the descriptive statistics showed that in 7 out of 15 problems, the GitHub Copilot generated code was more energy efficient than the Human code written, as shown in Fig. 2a and b.

Although GitHub Copilot-generated code generally shows improvements in Execution time and energy consumption, the differences are not statistically significant with our sample size. This finding suggests that while not statistically significant, there is a trend towards reduced energy consumption in code generated by AI tools like GitHub Copilot.

We did not perform statistical significance tests at the task level but observed significant variations in energy consumption across tasks. In the future, we aim to explore task-level significance testing to determine if these variations are statistically meaningful.

**RQ2: To what extent (in terms of percentage) can GitHub Copilot reduce the energy consumption of human-written code when prompted?**

To answer RQ2, we compare the energy consumption between the initial human-written code and the GitHub Copilot-optimized human-written code. Table 2 shows the results. The Mann-Whitney U test reveals a U-statistic of 138.5 and a p-value of 0.288. With a p-value of 0.288, greater than the significance level of 10%, we cannot reject the null hypothesis. This indicates that there are no statistically significant differences in energy consumption between the initial human-written code and the GitHub Copilot optimized code at the significance level 10%.

Despite this, the data showed improvements in energy consumption in several cases. In particular, the optimized Human-written code consumed an average of 18% less energy, as seen in Fig. 3, than its initial counterpart. In 9 out of 15 problems, the energy consumption of human-written code was significantly reduced after optimization with GitHub Copilot, as shown in Figs. 2a and 4. Although GitHub

Copilot can optimize code for a reduction in energy consumption, the degree of improvement may vary depending on the specific problem.

**RQ3: What trade-offs exist between code quality, performance, and energy consumption, when employing GitHub Copilot?**

To answer RQ3, we analyze the trade-offs between code quality, performance, and energy consumption when GitHub Copilot is used. To assess these trade-offs, the established metrics mentioned above are utilized.

**Code Quality** One of the key metrics analyzed regarding code quality is cyclomatic complexity. The results indicated that every function generated by GitHub Copilot received an 'A' rank, which indicates low complexity. This highly indicates the code produced by GitHub Copilot is straightforward to understand, indicating high code quality.

The maintainability index was also assessed. The results showed that again, all programs achieved an 'A' rank, signifying that the code generated by GitHub Copilot is highly maintainable.

This suggests that the use of GitHub Copilot does not introduce additional complexity or maintenance burden. Thus, this can alleviate concerns that optimizing for energy consumption or performance might lead to more complex and harder-to-maintain code.

The code generated by GitHub Copilot is not only simple but also easy to maintain. The low complexity and high maintainability ensure that the code is less prone to errors and easier to modify, supporting long-term sustainability and ease of use.

**Performance** Due to *cProfile* performing at optimum when the code's execution time is fairly long, the tool was not able to be used effectively due to smaller code

**Table 2** Green metrics: human-written code versus human-written optimized code

| Green metric | Human-written code | Optimized human-written |
|---|---|---|
| Execution time | 8.71E+10 s | 5.94E+10 s |
| Energy consumption | 0.41 J | 0.3373 J |
| CPU usage | 0.31% | 0.25% |
| RAM usage | 1.74E-06% | 1.76E-06% |

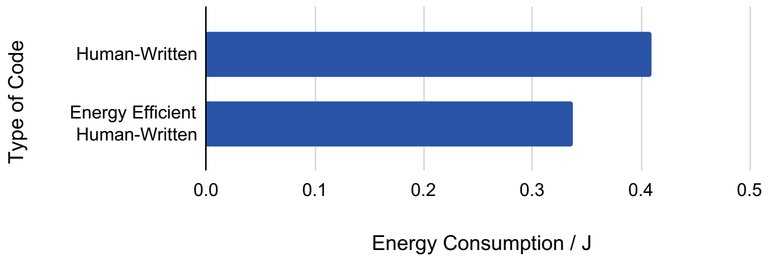



**Fig. 3** Comparison of average energy consumption of initial human-written and optimized human-written code

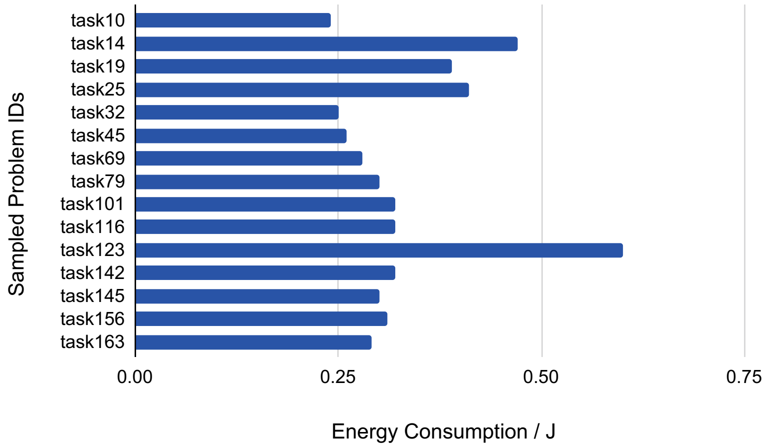
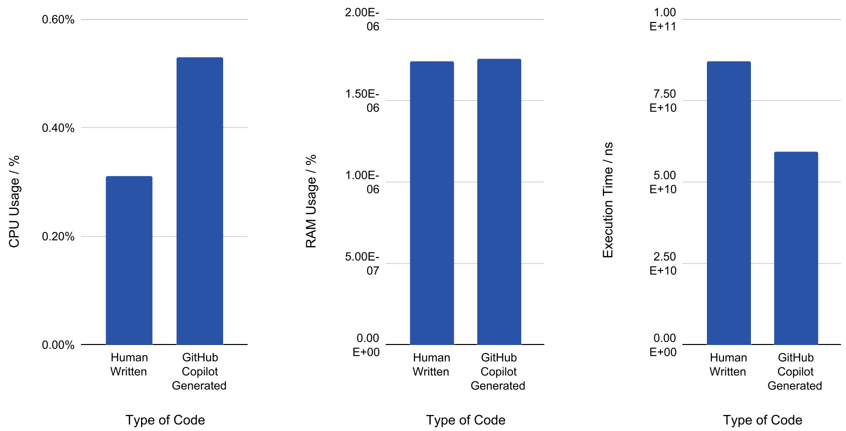



**Fig. 4** Individual energy consumption of optimized human-written code

snippets. Instead, the focus shifted to comparing Execution time, CPU Usage, and RAM Usage to evaluate performance.

Performance analysis of previously mentioned metrics reveals significant findings reflected in Fig. 5a, b, and c. Although GitHub Copilot-generated code generally utilized a high CPU and slightly higher RAM, it exhibited markedly faster execution times compared to human-written code. The trade-off here indicates that while GitHub Copilot can produce faster code, it does so by employing more computational resources. However, further analysis is needed to determine whether specific types of inefficiencies contribute to Copilot's performance optimizations.




(a) Average CPU Usage (b) Average RAM Usage (c) Average Execution time

**Fig. 5** Comparison of human-written code and GitHub copilot-generated code

### 4.1 Summary

We summarize the results of all RQs and present insights for practitioners who wish to include GitHub Copilot in their development workflow.

> GitHub Copilot-generated code achieves faster execution times, an advantage for applications where performance and responsiveness are critical.

The performance gain is primarily due to the optimized use of CPU resources, enabling more efficient processing and quicker execution of tasks. Although the increase in CPU usage is notable, this trade-off is often justified by the substantial reduction in execution time.

> GitHub Copilot-generated code consistently demonstrates lower energy usage compared to human-written code.

This reduction in energy consumption contributes to the development of more sustainable software, aligning with the growing emphasis on energy consumption in the technology industry. As energy costs and environmental concerns become increasingly significant, the ability to produce energy-efficient code is a valuable asset [1]. However, upon closer examination of individual tasks, we observed significant variations in performance metrics, where differences in energy and execution time appear more pronounced. While statistical significance was calculated for overall measurements, we did not perform individual significance tests at the task level. These task-specific variations may warrant further investigation in future research to understand whether they are statistically significant.

> GitHub Copilot-generated code may increase memory usage during code execution to provide a performance edge.

These performance and energy consumption improvements come with a slight increase in RAM Usage. While the difference in memory usage is relatively minor, it indicates a trade-off in memory management efficiency. In environments where memory resources are constrained, this could be a potential limitation.

> GitHub Copilot-generated code maintains low cyclomatic complexity and a high maintainability index in comparison to human-written code.

GitHub Copilot-generated code upholds a high standard of quality. Its low cyclomatic complexity shows the code is easy to navigate and understand. The high maintainability index indicates ease of modification for future changes. These qualities are particularly crucial for long-term software projects, where code maintainability significantly impacts development time and cost.

> The use of GitHub Copilot demonstrates a balanced approach to optimizing code for performance, energy consumption, and quality. While there are trade-offs in CPU and RAM Usage, the benefits in execution speed, energy savings, and maintaining high code quality, it also presents a compelling case for integrating AI-assisted coding tools in software development.

## 5 Challenges and Future Work

The study utilized a sample of 15 programming problems. While the sample size is sufficient for an initial exploratory analysis and offers preliminary insights into the potential of GitHub Copilot, it represents a limitation in terms of generalisability. As the small sample size may not be sufficient to draw definitive conclusions about the energy efficiency of AI-generated code, the results should be interpreted cautiously. Future research would benefit from a larger and more diverse set of problems to capture the full spectrum of programming tasks encountered in real-world software development.

The methodology included repeated executions of the code to mitigate the impact of anomalies and enhance the robustness of the results. However, using the controlled server environment, while beneficial for minimizing external influences, could also be considered a limitation. The findings may not fully reflect the performance variations that could occur in more diverse and dynamic real-world runtime environments.

While the study focuses on the energy efficiency of the generated code, the energy cost of running and training large AI models is a significant factor. Future research should investigate the total energy usage, including the overhead of AI model inference and training, to provide a more complete picture of the environmental impact of using AI-assisted coding tools. This could include identifying a break-even point where the energy savings from using more efficient, AI-generated code outweigh the energy consumed by the AI models themselves.

In addition to exploring energy consumption, future studies should expand the scope to examine other factors such as code quality, maintainability, and development time, and how they correlate with energy efficiency. Understanding the trade-offs between these factors will provide deeper insights into the broader impact of AI-assisted coding tools on the software development process. This would allow for a more comprehensive evaluation of the utility of tools like GitHub Copilot, particularly in balancing energy consumption with overall software development goals.

Despite these limitations, the study provides a basis for future research on AI-generated code's energy consumption and performance characteristics. The identified limitations can guide future studies to enhance the methodology, expand the sample size, and explore the impact of different run-time environments. By addressing these limitations, future research can contribute to a more comprehensive understanding of the potential of AI-assisted coding tools like GitHub Copilot and their implications for energy-efficient software development.

Future work should focus on exploring different types of programming problems, using alternative hardware and software tools, and using larger, more diverse datasets to fully understand the full capabilities and limitations of AI-assisted coding tools like GitHub Copilot. This will help to expand the applicability of these tools across various domains and enhance their effectiveness in real-world scenarios.

## 6 Conclusion

In this study, we conducted (i) a comparative analysis of energy consumption between human-written code and code generated by GitHub Copilot, (ii) examined the effect (on energy consumption and overall quality of the generated code) by prompting GitHub Copilot to optimize the human-written code, and (iii) explored how the trade-offs between performance and code quality metrics influence the overall utility of utilizing GitHub Copilot for software development.

Our findings indicate that code generated by GitHub Copilot, while using slightly more CPU and RAM, generally consumes less energy than human-written code while maintaining comparable execution times. This suggests that AI-generated code can be a viable solution for improving energy efficiency in software development without sacrificing performance. By explicitly prompting GitHub Copilot for energy-efficient solutions, developers can further reduce energy consumption. In our experiments, energy-optimized code demonstrated superior performance in terms of lower energy usage, while still maintaining functionality and adhering to quality standards. In terms of code quality, the generated code had lower cyclomatic complexity and high maintainability scores, indicating that the solutions produced by GitHub Copilot are easy to maintain. This is an important consideration for real-world applications, where the long-term sustainability of software is just as important as its initial performance. Using the strengths of AI-assisted coding tools, such as GitHub Copilot, developers can strike a balance between performance, energy consumption, and code quality.

**Acknowledgements** This work received funding from project SustainableCloud (OCENW.M20.243) of the research programme Open Competition which is (partly) financed by the Dutch Research Council (NWO). This work is partially supported by the LETSGO Project, promoted by the Netherlands Enterprise Agency (Rijksdienst voor Ondernemend Nederland). We thank our colleagues from Green Lab (s2group.cs.vu.nl/pages/greenlab/) for providing their support and sharing the lab resources for conducting experiments.

## References


1. Balanza-Martinez, J., Lago, P., Verdecchia, R.: Tactics for software energy efficiency: a review. In: Advances and New Trends in Environmental Informatics 2023, pp. 115–150. Springer Nature, Switzerland (2024)
2. Bird, C., Ford, D., Zimmermann, T., Forsgren, N., Kalliamvakou, E., Lowdermilk, T., Idan, G.: Taking flight with copilot: early insights and opportunities of AI-powered pair-programming tools. Queue **20**(6), 25–57 (2023)
3. Georgiou, S., Rizou, S., Spinellis, D.: Software development lifecycle for energy efficiency: techniques and tools. ACM Comput. Surv. **52**(4) (2019)
4. GitHub: About GitHub Copilot. https://docs.github.com/en/copilot/about-github-copilot. Accessed: 2024-05-29
5. Guldner, et al.: Development and evaluation of a reference measurement model for assessing the resource and energy efficiency of software products and components-Green Software Measurement Model (GSMM). Futur. Gener. Comput. Syst. **155**, 402–418 (2024)
6. Imai, S.: Is GitHub copilot a substitute for human pair-programming? An empirical study. In: 2022 IEEE/ACM 44th International Conference on Software Engineering: Companion Proceedings (ICSE-Companion), pp. 319–321 (2022)
7. Jun, H., Yuan, Q., Xu, Y.: HumanEval: hand-written evaluation set (2021). https://github.com/openai/human-eval. Accessed: 2024-05-29
8. Lago, P., Gu, Q., Bozzelli, P.: A systematic literature review of green software metrics. Technical report, Vrije Universiteit, Amsterdam (2014)
9. Mancebo, J., García, F., Calero, C.: A process for analysing the energy efficiency of software. Inf. Softw. Technol. **134** (2021)
10. Nascimento, N., Alencar, P., Cowan, D.: Artificial intelligence versus software engineers: an empirical study on performance and efficiency using ChatGPT. In: Proceedings of the 33rd Annual International Conference on Computer Science and Software Engineering, pp. 24–33. IBM Corp. (2023)
11. Nguyen, N., Nadi, S.: An empirical evaluation of GitHub Copilot's code suggestions. In: 19th International Conference on Mining Software Repositories (MSR '22) (2022)
12. Ournani, Z., Rouvoy, R., Rust, P., Penhoat, J.: Tales from the Code #1: the effective impact of code refactorings on software energy consumption. In: ICSOFT 2021 - 16th International Conference on Software Technologies. Virtual (2021)
13. Palomba, F., Nucci, D.D., Panichella, A., Zaidman, A., Lucia, A.D.: On the impact of code smells on the energy consumption of mobile applications. Inf. Softw. Technol. **105**, 43–55 (2019). https://doi.org/10.1016/j.infsof.2018.08.004
14. Pearce, H., Ahmad, B., Tan, B., Dolan-Gavitt, B., Karri, R.: Asleep at the keyboard? In: Assessing the Security of GitHub Copilot's Code Contributions (2021)
15. Peng, S., Kalliamvakou, E., Cihon, P., Demirer, M.: The Impact of AI on Developer Productivity: Evidence from GitHub Copilot (2023)
16. Pereira, R., Couto, M., Ribeiro, F., Rua, R., Cunha, J., Fernandes, J.P., Saraiva, J.: Ranking programming languages by energy efficiency. Sci. Comput. Program. **205**, 102,609 (2021). https://doi.org/10.1016/j.scico.2021.102609
17. Sabharwal, M.R.: Software power optimisation: analysis and optimisation for energy-efficient software. In: IEEE/ACM International Symposium on Low Power Electronics and Design, pp. 63–64 (2011)
18. Siavvas, M., Marantos, C., Papadopoulos, L., Kehagias, D., Tzovaras, D.: On the Relationship between Software Security and Energy Consumption (2019)
19. Stack Overflow: Hype or not? AI's benefits for developers explored in the 2023 Developer Survey. https://stackoverflow.blog/2023/06/14/hype-or-not-developers-have-something-to-say-about-ai/ (2023)
20. Stark, A.: Leverage AI for faster, more productive coding with GitHub Copilot (2023). https://techcommunity.microsoft.com/t5/microsoft-developer-community/leverage-ai-for-faster-more-productive-coding-with-github/ba-p/3993723. Accessed: 2024-05-29

21. Stivala, M., Fatima, I., Lago, P.: Replication package: investigating GitHub Copilot's impact in generating energy efficient software. https://github.com/S2-group/Investigating-GitHub-Copilot-s-Impact-in-Generating-Energy-Efficient-Software
22. Vartziotis, T., Dellatolas, I., Dasoulas, G., Schmidt, M., Schneider, F., Hoffman, T., Kotsopoulos, S., Keckeisen, M.: Learn to code sustainably: an empirical study on LLM-based green code generation (2024)
23. Woollacott, E.: Digital transformation set to drive huge rise in demand for critical resources (2024). https://www.itpro.com/business/digital-transformation/digital-transformation-set-to-drive-huge-rise-in-demand-for-critical-resources
24. Yetistiren, B., Ozsoy, I., Tuzun, E.: Assessing the quality of GitHub copilot's code generation. In: Proceedings of the 18th International Conference on Predictive Models and Data Analytics in Software Engineering, pp. 62–71. ACM (2022)

# Enhancing Energy Efficiency in AI: A Multi-faceted Analysis Across Time Series, Semantic AI and Deep Learning Domains

**Lejla Begic Fazlic, Berkay Cetkin, Achim Guldner, Matthias Dziubany, Julian Heinen, Stefan Naumann, and Guido Dartmann**

**Abstract** This research investigates strategies to enhance the energy efficiency of artificial intelligence (AI) algorithms, focusing on three pivotal domains: time series analysis, semantic AI, and deep learning (DL). Through a comprehensive examination of variables such as data size and the impact of hyper-parameter adjustments, the study aims to uncover nuanced insights into the relationship between algorithmic performance and energy consumption. By exploring the unique challenges and opportunities within each use case, this research provides valuable guidance for practitioners seeking to optimize energy efficiency in AI applications. The findings contribute to the ongoing discourse on sustainable AI development, offering practical overview to balance computational power with environmental considerations.

**Keywords** Artificial intelligence · Energy efficiency · Machine learning · Sustainability

## 1 Introduction

The rapid advancement of artificial intelligence (AI) has brought changes across various sectors. However, the increasing computational demands of AI algorithms pose significant energy efficiency challenges. Addressing these challenges is crucial to ensure sustainable AI development. This research focuses on analysing the resource and energy efficiency of AI algorithms, a key aspect of sustainable AI. It targets three pivotal domains: time series analysis, semantic AI, and deep learning (DL). Time series analysis is crucial in fields like finance and weather forecasting,

L. B. Fazlic (✉) · B. Cetkin · A. Guldner · S. Naumann · G. Dartmann
Institute for Software Systems (ISS), Trier University of Applied Sciences, Birkenfeld, Germany
e-mail: l.begic@umwelt-campus.de

M. Dziubany · J. Heinen
BITO CAMPUS GmbH, Meisenheim, Germany

 
V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_14

where AI can offer valuable insights. Each field offers distinct challenges and opportunities regarding energy efficiency. The goal of this research is to offer a detailed understanding of how different factors interact and affect the performance and energy efficiency of AI algorithms. We first conducted research on the energy aspects of natural language processing (NLP) in semantic AI. Next, we examined energy consumption related to complex neural network operations in DL. Finally, we focused on the computational demands of time series analysis. A key part of the study is to explore how the size of datasets affects energy efficiency, given that larger datasets generally demand more computational resources.

The research also examines how changes in hyper-parameters, like learning rate and batch size, can influence both algorithm performance and energy consumption, aiming to make AI models more energy-efficient. We conducted a study of the available literature to understand the relationship between energy use and algorithm efficiency. This research contributes to the discussion on sustainable AI by providing a practical view on balancing computing needs and environmental concerns. The structure of this paper is outlined as follows: Sect. 2 offers a review of relevant literature. The methodology and use case design is detailed in Sect. 3. Section 4 focuses on the evaluation and discussion, and the conclusions and the future work are provided in Sect. 5.

## 2 Related Work

In recent years, advancements in energy efficiency in the field of AI have focused on reducing the significant energy consumption of AI models. Advances in Green AI initiatives have been pivotal, emphasizing sustainable AI development by integrating energy-efficient practices in model training and deployment, thus balancing computational power with environmental considerations. Recent advancements not only focus on reducing the energy footprint of DL models [1–7] but also extend to optimizing semantic AI algorithms for better language understanding and enhancing time series models for more energy-efficient processing in different applications. In this study, we delve into the often-overlooked influence of software on the energy usage and overall environmental footprint of hardware systems. The rapidly growing field of AI, with a focus on machine learning (ML) and DL, has sparked a keen interest in evaluating their energy demands and ecological impacts during the training phase [5, 8, 9]. The assessment of the carbon footprint of AI [9], measurement of the energy requirements of AI systems [10], and evaluation of the efficiency of AI platforms such as PyTorch and TensorFlow [11] present one of the important research studies. Authors in [5] employ a life-cycle approach to estimate the carbon emissions generated from training NLP models. Recent advancements in Green AI emphasize the importance of energy-efficient ML where the contributions include the study on the impact of data preprocessing and feature selection [12], exploration of model optimization through weight pruning [13], and the discussion on Green AI's role in sustainable computing by Wang et al. [14]. In the field of energy- and resource-efficient software, various methodologies have been developed to assess the environmental sustainability of software products, as exemplified by the research of Naumann et al. [15] and Mancebo et al. [16]. Our analysis places special emphasis on energy consumption during the training and testing phases of various usage scenarios, comparing them using the energy and resource consumption metrics established by the authors in [17]. Recent advancements in time series analysis, especially regarding energy consumption, are key for efficient resource management and environmental sustainability. Sentiment analysis is a ML technique that interprets and classifies emotions expressed in text data, often used to understand opinions in customer feedback, social media, and other written sources. Recent research in sentiment analysis has explored the capabilities of large language models (LLMs). Studies like Zhong et al. [18] compared the zero-shot performance of LLMs with fine-tuned BERT models, while researchers in [19] investigated ChatGPT's proficiency in handling various sentiment analysis tasks, including polarity shifts and sentiment inference. Deng et al. [20] delved into fine-tuning a smaller model using a LLM to generate weak labels, achieving performances comparable to supervised models. These studies indicate a growing interest in understanding LLMs' effectiveness in sentiment analysis, but they also highlight the need for more comprehensive evaluations across diverse tasks and datasets. Recent study [21] introduces a new way to classify emotions in text using spiking neural networks (SNNs) to enhance energy efficiency. To the best of our knowledge, the investigation of energy efficiency in sentiment analysis tasks still remains a relatively untapped area in the field. Recent progress in image classification, led by convolutional neural networks (CNNs) models like ResNet [22] and VGG-16 [23], has significantly improved accuracy in various domains.

In our study, we examined three different use-case scenarios, exploring into aspects such as the volume of data and the consequences of modifying hyperparameters. This approach enabled us to discover complex details about the connection between the performance of algorithms and their energy consumption. In our first sentiment analysis use case scenario, we specifically investigate the energy efficiency of BERT models, examining how different data volumes and batch sizes influence their energy consumption. This focused approach enables us to understand the nuances of energy usage in these NLP models, ensuring that they not only maintain high accuracy but also adhere to sustainable computational practices. As a second use case scenario in our study, we extended our analysis of resource and energy efficiency to encompass various epoch and batch sizes while working with neural network models like ResNet [22], DenseNet [24], MobileNet [25], Inception [26], VGG-16 [23] and VGG-19 [23] for image classification. This approach allowed us to comprehensively assess how different training configurations impact the energy and computational demands of these models. Our findings offer a detailed perspective on the balance between training efficiency, model accuracy, and energy use, providing important guidance for enhancing neural network training in scenarios where energy efficiency is crucial. Transitioning to our third scenario, we shift our focus to the extensive landscape of sensor data. This case study involved analyzing time series data using tools from sktime, a Python library for time series analysis [27], specifically the Random Interval Spectral Ensemble (RISE), the KNeighbors

Time Series Classifier (kNNTime) and the Time Series Forest Classifier (TSFC). We aimed to evaluate the accuracy of detecting fill levels, while also incorporating energy assessments to determine the system's energy efficiency.

# 3 Methodology and Design

In this section, we outline the research questions we aim to address, describe the practical experiment conducted as part of this research, which includes a description of the case study's design, the experimental procedures employed, and the methods used to analyze the collected results. We assess algorithm efficiency by monitoring hardware usage and power consumption in the aforementioned scenarios. For reference, we have prepared a comprehensive replication package, which includes detailed system specifications, the scenarios explored, the data recorded, and the results of our analyses. All these materials are accessible in our Git repository https://gitlab.rlp.net/rgdsai/mfa.

## *3.1 Research Questions: AI Model Efficiency*

In light of our research objectives, we structured our inquiry into the energy consumption and efficiency of AI models, particularly in NLP and DL, through the following research questions:

**RQ1**: How does batch size influence energy consumption and accuracy in NLP model training? This research question aims to understand the relationship between batch size during the fine-tuning phase of NLP models and its impact on energy efficiency and model accuracy. By answering this, we intend to identify optimal batch sizes that balance energy consumption with performance efficacy.

**RQ2**: What is the effect of training data size on the energy consumption and accuracy of NLP models? We investigated the energy consumption and accuracy of models by varying the proportion of the training data, exploring how these metrics change with increasing data volumes. This seeks to determine the optimal data size that ensures efficient energy use without compromising the model's accuracy.

**RQ3**: How do different DL architectures compare regarding energy consumption and model performance? Focusing on architectures like DenseNet, ResNet, VGG-16, VGG-19, and Inception, this question explores how the structural variances in these models affect their energy efficiency and overall performance. The goal is to provide a comparative analysis that guides the selection of the most energy-efficient model without sacrificing performance.

**RQ4**: In the context of time series approaches using sensor data, what is the best configuration that exhibit the most energy-efficient processing? This question extends the study to time series analysis, evaluating various AI model scenarios on their energy efficiency when processing sensor data. The objective is to identify models that offer an optimal balance between energy efficiency and effective time series analysis.

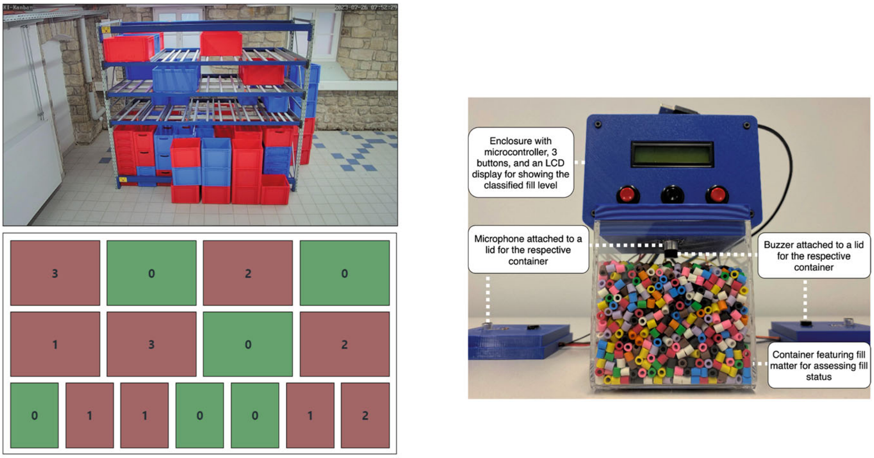
By exploring these research questions, our study aims to provide meaningful contributions into the energy-efficient implementation of AI systems, especially in a time where the environmental impact of technology is of paramount concern. The findings are expected to guide both practitioners and researchers in making informed decisions about AI model selection and optimization for sustainable usage.

### *3.2 Data Acquisition*

Our research embarks on a multifaceted exploration, traversing three distinct scenarios. For our first use case scenario we used Stanford's Large Movie Review Dataset IMDB [28, 29] that contains 50,000 movie reviews. Reviews are labeled as 1 or 0 corresponding to positive or negative sentiment, respectively. A minimal data preprocessing are done prior to tokenization as BERT was trained on complete sentences. To effectively utilize pre-trained BERT, we must utilize the library's tokenizer due to BERT's specific, fixed vocabulary and the tokenizer's unique handling of out-of-vocabulary words. Additionally, it's essential to add special tokens at the start and end of each sentence, standardize sentence length through padding or truncation, and explicitly identify padding tokens using the "attention mask". The developed code supports sentiment analysis on a variety of CSV datasets, making it adaptable to any text classification task, with the ability to analyze textual data and assign sentiment labels, regardless of the specific dataset's structure or content.

In our vision use-case a simple webcam took pictures of a flowrack with various number of bins in each lane. As the bins move forward after putting them into a lane, it is possible to determine the number of bins from the backside of the rack. The pictures of the whole rack were devided into smaller ones showing only one separated lane, which can store a maximal number of tree bins. In order to manage different lane sizes and camera angels the pictures of the lanes were mathematically transformed to equal size. Figure 1a shows a picture of the flowrack and the determined number of bins in each lane.

In our third application scenario, we developed a demonstrator for generating time series data. It's designed to determine the fill level of a small container through acoustic sound waves, with the capability to differentiate between five distinct levels of fill: 0%, 25%, 50%, 75%, and 100%. The construction of this demonstrator involved the use of 3D printing to create a structure that houses a central unit. This central unit incorporates an ESP32 [30] microcontroller, equipped with three buttons and a display that shows the classified fill level. Attached to this central unit are three lids, each fitted with a buzzer to generate acoustic signals and a microphone for recording the corresponding time-series data. The operational principle involves placing a lid over a container filled with material. Upon pressing the corresponding button associated with that lid, a sinusoidal sweep is emitted as an acoustic signal. The recording of this signal commences before the emission of the sound to capture




(a) (b)

**Fig. 1** Demonstrators utilized in the experimental use case

ambient noise for noise reduction purposes and continues after the signal has ended to include sound reflections. This process facilitates the approximate calculation of the room's impulse response, which is then transmitted to a Raspberry Pi via MQTT [31]. This Raspberry Pi is responsible for managing and storing the data. The impulse response of the room, captured in this scenario, serves as time-series data and forms the foundation for training the ML algorithms: RISE, kNNTime, and TSFC. Figure 1b illustrates the setup of the demonstrator, providing a visual representation of its configuration and components.

### *3.3 Case Study Overview*

In our first use case scenario, we conducted series of experiments specifically focusing on text classification using a BERT model. We iterate through a predetermined number of experiments, adjusting the dataset and feature set sizes based on specified percentages, to evaluate the impact on model performance. The process includes data preprocessing, model training, and validation stages, logging each step's start and end times for energy performance tracking and reproducibility. The different adjustments of the batch size (16 or 32) are used during training to maintain a balance between gradient noise and memory efficiency. The Adam optimizer with default hyperparameters was used in all scenarios. Additionally, we emphasize memory management through explicit garbage collection and system calls to clear RAM, ensuring efficient resource utilization during the experiments. Our methodological framework involves a detailed comparison of energy consumption metrics such as mean power (W) and energy usage (Wh), for the preprocessing and training phases, as well as for the GPU utilization. By incrementally increasing dataset sizes from 10% to 100% of the total volume, we could simulate different training intensities to observe their impacts on energy efficiency and model performance. This approach allowed us to capture a range of performance metrics, including processing times, CPU and GPU usage percentages, RAM and GRAM usage, and GPU temperature.

In our study's second use case scenario, we extended our analysis of resource and energy efficiency by examining various epoch and batch sizes when working with neural network models like ResNet [23], DenseNet [25], MobileNet [26], Inception [27], VGG-16 [24], and VGG-19 [24] for image classification, enabling a comprehensive assessment of the impact of different training configurations on energy and computational demands.

In our third scenario, we explore the application of Edge AI to determine container fill levels using acoustic analysis. Our analysis spans across distinct settings, aiming to identify the most efficient combination for precise fill level classification while optimizing energy consumption. We utilized three ML algorithms: RISE, kNNTime, and TSFC. These were employed to adjust two key parameters: sample length (the length of the room impulse response considered by the models) and the number of estimators/neighbors. Analogous to the second use case, we measured energy consumption using metrics such as mean power, energy usage during training and testing phases, as well as training and testing duration, CPU usage, and accuracy using the F1 score metric.

### *3.4 Tailored Hardware for Different Use Cases and Measurement Methodology*

For the semantic analysis and DL use-cases, we utilized a high-powered server configuration, powered by an Intel Xeon W-2295 processor with 18 cores running at 3.0 GHz, paired with 131.56 GB of DDR4-2933 RAM. It features an NVIDIA GeForce RTX 4090 24 GB GPU and is built on an ASUS WS C422 Pro/SE mainboard. This setup was specifically selected to optimize the processing power and memory requirements needed for these complex tasks, ensuring efficient and effective analysis. Conversely, for the time series analysis use-case, we choose a standard PC configuration. The experiments were carried out on a computer setup that included 4 GB of RAM, arranged in two modules of 2 GB each, and was driven by an Intel Core i5-650 processor. This system also boasted a dual-storage configuration, combining a 500 GB hard disk drive (HDD) for extensive storage capacity with a 250 GB solid-state drive (SSD) for rapid data retrieval and system responsiveness. This choice reflects the comparatively lower computational demands of time series analysis, which, while still requiring precision and accuracy, can be effectively conducted on less powerful hardware. This distinction in hardware choice between the use-cases allowed us to not only tailor our computational resources to the specific needs of each task but also to investigate the potential impacts of hardware capabilities on the efficiency and outcomes of different types of data analysis.

The measurement methodology for AI-based method was described in previous work [32] and in context of AI methods in [33]. The measurements are based on the methodology and guidelines outlined in the Green Software Measurement Model [34]. Data aggregation is automated, with users logging process start times, end times, and labels in a CSV file called the action log. A Standard Usage Scenario (SUS) outlines the basic workflow. Resource data, including CPU and RAM usage, is recorded using the Linux performance reporting tool collectl on a Linux Ubuntu system. Energy consumption data is obtained from a power meter, requiring synchronization with the executing computer's time. The accurate calculation of energy consumption for the process requires recording the baseline consumption of the executing system through a corresponding measurement, with the measured value adjusted accordingly, ideally conducted for a duration comparable to that of the actual process.

## 4 Results and Discussion

To address our first research question (**RQ1**), we aim to explore the impact of batch size on energy consumption and accuracy during the fine-tuning phase of NLP model training. Our goal is to try to identify the optimal batch sizes that balance energy efficiency with performance efficacy. We conducted a comprehensive analysis by systematically varying batch sizes and measuring their effects on energy consumption and accuracy for the BERT model across different dataset sizes. Detailed performance metrics, including energy efficiency (measured in Wh) and model accuracy (indicated by the F1 score) for two batch sizes (16 and 32), are presented in Table 1.

The relationships between batch size, mean power consumption and F1 score are presented in Fig. 2a.

Our final insights from this examination of batch size impacts on BERT model training underscore the intricate balance between computational resource utilization and model effectiveness. Our key findings are outlined below:

- We found a direct correlation between batch size and energy consumption. Larger batch sizes generally resulted in higher mean power consumption but also offered more efficient processing in terms of energy per data point processed, particularly with larger datasets. The analysis revealed that batch sizes of 32 often resulted in shorter training and preprocessing times compared to a batch size of 16, especially as the dataset size increased.
- Optimal model performance was observed at varying dataset percentages, highlighting the importance of selecting batch size based on the specific context of the model's application. The F1 score analysis indicated that while larger batch sizes can enhance training efficiency, they do not always correlate with improved model accuracy.

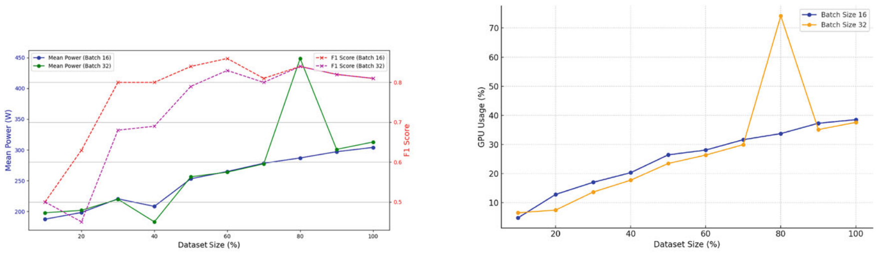
**Table 1** Combined performance metrics for Bert model

| Performance metrics for bert model—batch size 16 | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Metric | 10% | 20% | 30% | 40% | 50% | 60% | 70% | 80% | 90% | 100% |
| Mean power [W] | 187.516 | 198.40 | 220.544 | 208.318 | 253.173 | 264.890 | 278.309 | 286.963 | 297.075 | 304.070 |
| Time preprocessing [s] | 3.941 | 3.973 | 3.954 | 4.002 | 3.990 | 3.993 | 4.040 | 4.041 | 4.035 | 4.067 |
| Time train [s] | 1.410 | 1.911 | 2.433 | 3.044 | 3.421 | 3.971 | 4.426 | 4.949 | 5.379 | 5.891 |
| Time test [s] | 0.020 | 0.028 | 0.041 | 0.057 | 0.070 | 0.081 | 0.091 | 0.103 | 0.118 | 0.130 |
| Energy preprocessing [Wh] | 0.080 | 0.075 | 0.073 | 0.103 | 0.076 | 0.074 | 0.073 | 0.074 | 0.074 | 0.079 |
| Energy train [Wh] | 0.041 | 0.073 | 0.123 | 0.088 | 0.222 | 0.266 | 0.315 | 0.365 | 0.407 | 0.451 |
| CPU usage [%] | 3.250 | 3.448 | 3.437 | 3.307 | 3.464 | 3.410 | 3.409 | 3.385 | 3.350 | 3.266 |
| GPU usage [%] | 4.827 | 12.848 | 17.005 | 20.319 | 26.442 | 28.080 | 31.641 | 33.742 | 37.290 | 38.523 |
| RAM usage [%] | 1.072 | 1.122 | 1.141 | 1.242 | 1.180 | 1.189 | 1.242 | 1.228 | 1.294 | 1.306 |
| GRAM usage [%] | 12.338 | 12.329 | 12.630 | 12.399 | 12.490 | 12.488 | 12.492 | 12.548 | 12.595 | 12.582 |
| GPU temp [°C] | 53.177 | 42.558 | 46.417 | 49.798 | 47.827 | 45.068 | 45.756 | 46.346 | 47.625 | 40.608 |
| GPU energy [Wh] | 0.055 | 0.084 | 0.129 | 0.149 | 0.224 | 0.263 | 0.311 | 0.358 | 0.412 | 0.453 |
| F1 score | 0.5 | 0.63 | 0.80 | 0.80 | 0.84 | 0.86 | 0.81 | 0.84 | 0.82 | 0.81 |
| Performance metrics for bert model—batch size 32 | | | | | | | | | | |
| Metric | 10% | 20% | 30% | 40% | 50% | 60% | 70% | 80% | 90% | 100% |
| Mean power [W] | 197.94 | 201.840 | 219.91 | 183.181 | 256.350 | 263.929 | 277.332 | 448.520 | 300.924 | 312.802 |
| Time preprocessing [s] | 3.923 | 3.940 | 3.964 | 3.988 | 4.020 | 4.015 | 3.994 | 4.016 | 4.042 | 4.062 |
| Time train [s] | 1.253 | 1.560 | 1.921 | 2.463 | 2.581 | 2.873 | 3.246 | 3.282 | 3.906 | 4.208 |
| Time test [s] | 0.016 | 0.027 | 0.039 | 0.050 | 0.064 | 0.074 | 0.085 | 0.204 | 0.108 | 0.122 |
| Energy preprocessing [Wh] | 0.076 | 0.071 | 0.078 | 0.117 | 0.077 | 0.076 | 0.074 | 0.083 | 0.078 | 0.077 |
| Energy train [Wh] | 0.047 | 0.075 | 0.104 | 0.011 | 0.227 | 0.225 | 0.264 | 0.318 | 0.351 | 0.362 |

(continued)

**Table 1** (continued)

| Performance metrics for bert model—batch size 16 | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Metric | 10% | 20% | 30% | 40% | 50% | 60% | 70% | 80% | 90% | 100% |
| CPU usage [%] | 3.197 | 3.331 | 3.417 | 3.374 | 3.393 | 3.518 | 3.449 | 3.678 | 3.417 | 3.408 |
| GPU usage [%] | 6.553 | 7.482 | 13.658 | 17.744 | 23.481 | 26.364 | 29.937 | 74.155 | 35.123 | 37.605 |
| RAM usage [%] | 1.135 | 1.230 | 1.178 | 1.374 | 1.290 | 1.281 | 1.329 | 1.563 | 1.347 | 1.333 |
| GRAM usage [%] | 17.160 | 17.404 | 17.315 | 17.198 | 17.442 | 17.499 | 17.442 | 18.708 | 17.630 | 17.641 |
| GPU temp [°C] | 54.491 | 42.318 | 46.500 | 47.243 | 47.956 | 48.306 | 47.994 | 64.481 | 50.425 | 49.755 |
| GPU energy [Wh] | 0.055 | 0.059 | 0.088 | 0.113 | 0.169 | 0.198 | 0.237 | 9.247 | 0.316 | 0.346 |
| F1 score | 0.5 | 0.45 | 0.68 | 0.69 | 0.79 | 0.83 | 0.80 | 0.84 | 0.82 | 0.81 |




(a) (b)

**Fig. 2** Performance metrics and resource utilization for BERT models

- Batch size 32 (orange line in Fig. 2b) typically exhibits lower or comparable GPU usage compared to batch size 16 for most dataset sizes, except for a significant spike at the 80% data size. This suggests that using a larger batch size can be more GPU efficient for certain dataset sizes but may encounter inefficiencies or bottlenecks at specific points (like 80%). We found that as the dataset size increases, so does the energy required for training, particularly evident in smaller batch sizes.
- Identifying the optimal batch size for NLP model training involves balancing energy efficiency, training speed, and model accuracy. Our findings suggest that while larger batch sizes may enhance computational efficiency, they require careful consideration of the trade-offs involved, particularly regarding model performance and hardware limitations.

Our investigation into the effects of batch size on the energy consumption and accuracy of NLP model training with a specific focus on the BERT model provides critical insights for optimizing training processes. These insights emphasize the necessity of a nuanced approach to batch size selection, tailored to the specific goals of energy efficiency, computational resource management, and model accuracy.

In addressing our second research question (**RQ2**), we investigate the effects of scaling training dataset size on the energy consumption and accuracy of a BERT model. This study aims to elucidate the dual impact of dataset size on model efficiency-quantified through energy requirements-and on model performance, as gauged by the F1 score. To systematically explore these relationships, we employ Pearson correlation analysis. This method allows us to quantitatively assess the linear correlation between training data size and two key outcomes: the energy consumed throughout the training phase and the accuracy of the model. Our analysis revealed significant relationships for a batch size of 16:

- We found a strong positive correlation of 0.98 between the training data size and the energy consumed during training, highlighting a substantial increase in energy requirements as dataset size expands.
- We found a moderate positive link with a correlation of about 0.70, between the size of the training data and the F1 score. This means that making the dataset

larger tends to improve how accurate the model is, but not as much as it increases the energy needed for training.

From our analysis, we also found the following correlations for a batch size of 32:

– While a strong positive correlation of 0.93 was observed between the training data size and the energy consumed during training for batch size 32, this is slightly lower than the correlation noted for batch size 16, suggesting that while energy demands still increase with larger datasets, the rate of increase may be less steep for larger batch sizes.
– A strong positive correlation (0.85) between the training data size and the F1 score for batch size 32 was found, indicating a more distinct improvement in model accuracy with larger training datasets, more so than observed with batch size 16.

Our third research question (**RQ3**) focused on unraveling the intricate balance between energy consumption and model performance across prevalent DL architectures. This inquiry aimed to dissect how structural variations in models such as DenseNet, ResNet, VGG-16, VGG-19, and Inception influence their operational efficiency and effectiveness. We thoroughly compiled data reflecting each model's energy usage during training and testing phases, accompanied by their performance metrics, primarily measured through F1 scores. Additionally, we considered the computational time and resource utilization, to paint a comprehensive picture of each model's energy profile as it is shown in Table 2. The energy consumption across different models and the trade-off between global energy consumption and the best F1 score are depicted in Figs. 3 and 4, respectively.

Through this comparative analysis, our goal was to offer practical guidance for choosing the most energy-efficient model without sacrificing performance. In our analysis of energy efficiency across various DL architectures, we discovered that each model exhibits its own optimal configuration that harmonizes energy consumption with performance. For DenseNet and ResNet, the optimal configuration was determined to be a batch size of 32 with 5 training epochs, which significantly reduced energy usage during both the training and testing phases without compromising the models' accuracy, as evidenced by their F1 scores. Similarly, MobileNet demonstrated its energy efficiency under the same parameters, indicating a consistent pattern among these architectures for achieving operational efficiency. The Inception model, however, diverged slightly, finding its most energy-efficient configuration with a batch size of 16 and 5 epochs. This adjustment offered substantial energy savings across the board, while still securing a high F1 score, illustrating that a slight reduction in batch size could yield notable efficiency gains without sacrificing performance. VGG-16bn followed a similar pattern to Inception, opting for a batch size of 16 and 5 epochs, which was effective in reducing energy consumption while maintaining a commendable level of accuracy. Conversely, VGG-19bn aligned with DenseNet and ResNet, favoring a batch size of 32 and 5 epochs for its most energy-efficient performance. This setup allowed for minimized energy usage during

**Table 2** Comparison of algorithm performance and energy consumption—scenario image classification

| Model | Batch | Epoch | Mean | Time [s] | | Energy [Wh] | | CPU [%] | GPU[%] | RAM [%] | GRAM [%] | GPU [C°] | GPU [Wh] | F1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Name | Size | Numb. | Power[W] | Train | Test | Train | Test | Usage | Usage | Usage | Usage | Temp. | Energy | Score (%) |
| DenseNet | 32 | 9 | 271.06 | 104.7 | 74.35 | 6.39 | 1.74 | 11.85 | 40.26 | 1.40 | 21.17 | 44.52 | 5.92 | 99.65 |
| | 16 | 9 | 224.61 | 165.10 | 76.44 | 6.09 | 1.73 | 12.84 | 31.17 | 1.37 | 12.25 | 39.81 | 5.12 | 99.54 |
| | 32 | 5 | 245.50 | 53.66 | 73.80 | 3.20 | 1.67 | 9.76 | 33.71 | 1.40 | 21.24 | 40.90 | 3.30 | 99.34 |
| | 16 | 5 | 159.2 | 83.10 | 76.15 | 3.07 | 1.73 | 10.75 | 28.16 | 1.40 | 12.24 | 40.30 | 3.00 | 98.80 |
| ResNet | 32 | 9 | 282.32 | 46.65 | 14.76 | 2.58 | 0.40 | 26.77 | 34.64 | 0.72 | 6.69 | 44.86 | 1.77 | 99.09 |
| | 16 | 9 | 279.14 | 52.27 | 14.97 | 2.82 | 0.39 | 28.11 | 34.03 | 0.77 | 4.05 | 44.59 | 1.83 | 99.74 |
| | 32 | 5 | 265.14 | 24.70 | 15.04 | 1.33 | 0.40 | 22.77 | 31.55 | 0.80 | 6.64 | 49.00 | 1.02 | 99.80 |
| | 16 | 5 | 263.88 | 27.32 | 14.75 | 1.43 | 0.40 | 24.16 | 31.59 | 0.79 | 4.04 | 48.69 | 1.04 | 99.76 |
| MobileNet | 32 | 9 | 268.86 | 54.83 | 30.00 | 3.08 | 0.70 | 20.82 | 32.84 | 2.00 | 12.91 | 45.47 | 2.39 | 99.77 |
| | 16 | 9 | 232.04 | 72.46 | 30.65 | 2.85 | 0.71 | 22.25 | 26.82 | 2.06 | 7.74 | 44.61 | 2.11 | 99.56 |
| | 32 | 5 | 244.09 | 28.63 | 30.52 | 1.54 | 0.70 | 16.67 | 27.20 | 2.09 | 12.66 | 43.11 | 1.32 | 99.35 |
| | 16 | 5 | 226.24 | 36.00 | 29.65 | 1.44 | 0.70 | 18.43 | 24.30 | 2.11 | 7.74 | 47.77 | 1.24 | 99.82 |
| Inception | 32 | 9 | 286.27 | 138.85 | 53.84 | 7.72 | 1.82 | 19.64 | 43.25 | 0.85 | 18.81 | 47.20 | 6.95 | 99.82 |
| | 16 | 9 | 270.72 | 165.77 | 54.70 | 8.27 | 1.70 | 18.68 | 39.57 | 0.79 | 10.52 | 45.73 | 7.13 | 99.81 |
| | 32 | 5 | 253.88 | 71.30 | 54.16 | 3.25 | 1.83 | 16.48 | 37.67 | 0.87 | 18.79 | 44.77 | 3.72 | 99.31 |
| | 16 | 5 | 247.05 | 83.80 | 54.50 | 3.65 | 1.68 | 16.07 | 36.00 | 0.80 | 9.96 | 43.65 | 3.83 | 99.84 |
| VGG-16 | 32 | 9 | 355.50 | 140.54 | 14.26 | 9.68 | 0.95 | 13.61 | 74.27 | 1.02 | 30.88 | 58.16 | 9.23 | 64.05 |
| | 16 | 9 | 346.86 | 159.69 | 14.26 | 10.62 | 0.91 | 16.07 | 77.12 | 1.00 | 19.25 | 57.29 | 10.11 | 35.16 |
| | 32 | 5 | 333.10 | 71.18 | 14.11 | 4.36 | 0.98 | 13.34 | 71.65 | 1.05 | 31.05 | 56.46 | 4.97 | 71.68 |

(continued)

**Table 2** (continued)

| Model | Batch | Epoch | Mean | Time [s] | | Energy [Wh] | | CPU [%] | GPU[%] | RAM [%] | GRAM [%] | GPU [C°] | GPU [Wh] | F1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Name | Size | Numb. | Power[W] | Train | Test | Train | Test | Usage | Usage | Usage | Usage | Temp. | Energy | Score (%) |
| | 16 | 5 | 327.79 | 81.37 | 14.22 | 4.93 | 0.91 | 15.37 | 73.69 | 1.03 | 19.27 | 55.59 | 5.40 | 39.19 |
| VGG-16bn | 32 | 9 | 372.18 | 159.50 | 16.36 | 11.66 | 1.24 | 12.51 | 75.92 | 1.10 | 33.90 | 61.07 | 12.03 | 99.40 |
| | 16 | 9 | 370.19 | 176.72 | 16.14 | 12.88 | 1.19 | 14.59 | 78.68 | 1.07 | 23.06 | 60.98 | 13.14 | 98.43 |
| | 32 | 5 | 338.18 | 81.96 | 16.27 | 5.03 | 1.25 | 12.24 | 72.32 | 1.14 | 33.88 | 59.13 | 6.29 | 99.56 |
| | 16 | 5 | 340.11 | 89.79 | 16.37 | 5.64 | 1.20 | 14.03 | 75.09 | 1.13 | 23.04 | 59.35 | 6.83 | 93.75 |
| VGG-19 | 32 | 9 | 349.90 | 156.09 | 15.92 | 10.39 | 1.16 | 13.66 | 76.42 | 1.66 | 27.30 | 58.68 | 10.84 | 49.34 |
| | 16 | 9 | 271.38 | 44.26 | 15.88 | 2.01 | 0.72 | 12.67 | 71.91 | 1.22 | 19.57 | 53.16 | 3.29 | 33.15 |
| | 32 | 5 | 321.77 | 80.00 | 15.95 | 4.97 | 0.73 | 12.30 | 73.35 | 1.19 | 26.73 | 57.14 | 5.80 | 63.23 |
| | 16 | 5 | 316.62 | 90.95 | 15.90 | 5.80 | 0.43 | 14.00 | 76.47 | 1.16 | 19.54 | 56.08 | 6.20 | 28.71 |
| VGG-19bn | 32 | 9 | 376.54 | 178.10 | 17.99 | 13.28 | 1.35 | 11.64 | 77.52 | 1.16 | 35.35 | 61.00 | 13.58 | 92.14 |
| | 16 | 9 | 373.22 | 198.73 | 17.80 | 14.63 | 1.31 | 13.45 | 80.40 | 1.12 | 23.76 | 61.20 | 14.86 | 92.18 |
| | 32 | 5 | 348.06 | 90.72 | 18.00 | 5.89 | 1.36 | 11.26 | 74.39 | 1.11 | 35.33 | 60.15 | 7.18 | 98.00 |
| | 16 | 5 | 348.36 | 100.26 | 17.92 | 6.56 | 1.34 | 12.90 | 77.51 | 1.09 | 23.75 | 60.15 | 7.80 | 93.20 |

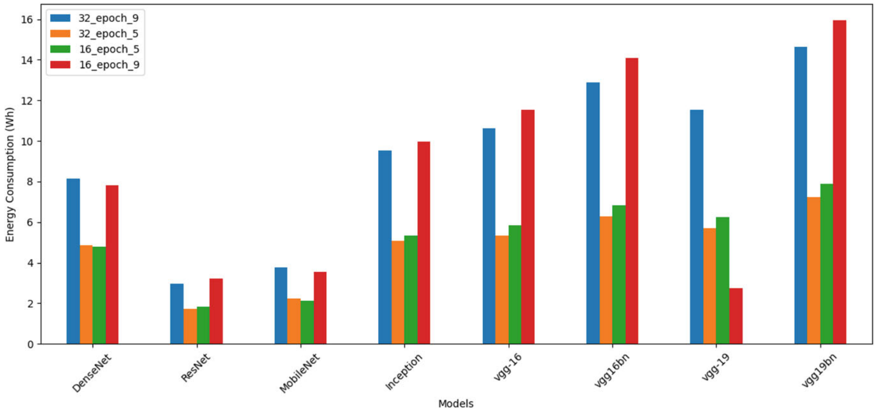



**Fig. 3** Total energy consumption versus models

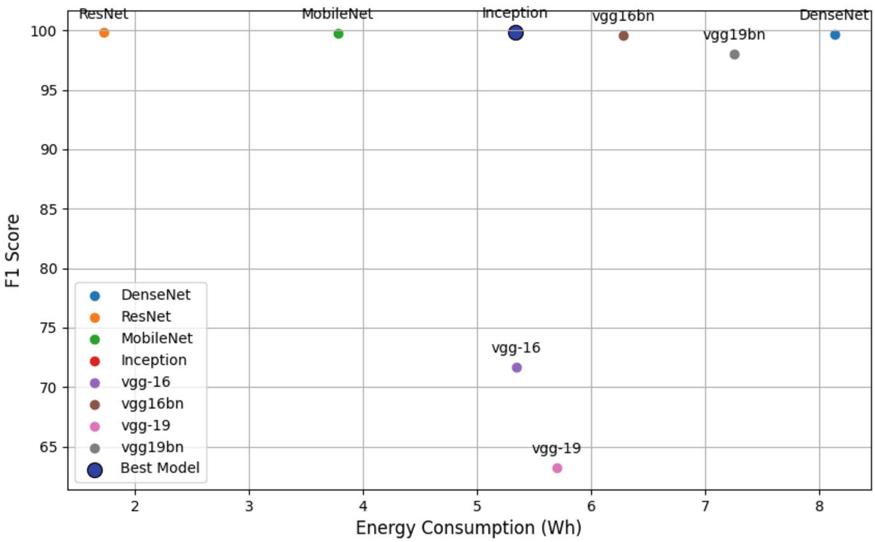



**Fig. 4** Trade-off between global energy consumption and best F1 score

operations while achieving a robust F1 score. This affirms that even among models with varying complexities, there is an opportunity to achieve equilibrium between energy efficiency and model accuracy. Our key insights are summarized as follows:

– The findings indicate a complex relationship between model depth and energy consumption, challenging the conventional belief that more complex models are always more energy-intensive.The investigation into F1 scores highlighted that architectural sophistication does not always translate to enhanced model accuracy.

For instance, despite VGG-19's depth, it did not consistently outperform the less complex VGG-16 in terms of accuracy.

– Through the lens of our analysis, it became evident that there isn't a universal optimal batch size or epoch count that maximizes energy efficiency across all models. Instead, each architecture demands a tailored approach to find its equilibrium point that harmonizes energy consumption with model performance. Across all models, a trend towards selecting a moderate batch size and a lower number of epochs (5) appears to be the sweet spot for optimizing energy efficiency without significantly impacting the model's performance. This analysis underscores the critical insight that, despite the diversity in architecture and design, DL models can achieve a delicate balance between energy efficiency and performance through strategic adjustments in batch size and epoch count.
– The research also highlighted a connection between the efficient use of hardware resources and energy consumption. Models that were able to make more effective use of hardware resources frequently demonstrated improved energy efficiency without necessarily sacrificing performance.

In addressing our last research question (**RF4**) regarding the energy efficiency of AI solution for processing time series sensor data, Table 3 outlines five distinct configurations of the kNNTime, TSFC and RISE algorithms. Each provides insights into how variations in sample lengths and the number of estimators/neighbors impact the algorithm's energy consumption and classification performance. Figure 5 illustrates the energy consumption and performance analysis of ML algorithms, using the first three entries for each algorithm from Table 3. The kNNTime algorithm demonstrates robust performance, particularly evident through its consistently high F1 scores, which peak at 0.92 for both 400 and 100 sample lengths when utilizing just one neighbor.

When examining energy and time efficiency, we observe that energy consumption during testing presents considerable variation, escalating as sample length increases. The energy required for training remains consistently low across various configurations, emphasizing the algorithm's efficiency during the learning phase.

This is primarily because kNN, unlike many other ML algorithms, does not require a separate training step to create a model; it simply stores the data and makes inferences directly from the entire dataset during the prediction phase. The most efficient configuration for kNNTime emerges with a sample length of 100 paired with a single neighbor. This setup also minimizes energy consumption and reduces testing time, representing an optimal balance for those seeking both precision and efficiency. For algorithm RISE, reducing the sample length from 700 to 100 leads to a decrease in energy consumption throughout the training and testing phase, aligning with expectations that less data requires less computational power. The F1 score remains consistently high across different sample lengths, suggesting that RISE effectively maintains predictive performance even with reduced data. Reducing the number of estimators from 100 to 50 decreases energy consumption in both training and testing phases without significantly compromising the F1 score. This indicates an efficient use of computational resources by RISE, as it maintains high accuracy with

**Table 3** Comparison of algorithm performance and energy consumption across different parameter configurations

| Algorithm Name | Sample Length | N-Estim./ Neighbors | Mean Power [W] | Energy [Wh] | | Time [s] | | CPU Usage [%] | F1 Score |
|---|---|---|---|---|---|---|---|---|---|
| | | | | Training | Testing | Training | Testing | | |
| kNNTime | 700 | 1 | 83.77 | 0.015 | 7518.32 | 0.00 | 174.56 | 24.67 | 0.87 |
| | 400 | 1 | 83.40 | 0.011 | 2466.45 | 0.00 | 57.02 | 24.67 | 0.92 |
| | 100 | 1 | 80.17 | 0.013 | 153.32 | 0.00 | 3.410 | 24.59 | 0.92 |
| | 700 | 3 | 83.56 | 0.017 | 7605.13 | 0.00 | 176.09 | 24.69 | 0.75 |
| | 400 | 3 | 83.23 | 0.012 | 2449.15 | 0.00 | 56.50 | 24.67 | 0.73 |
| RISE | 700 | 100 | 77.93 | 47.87 | 31.16 | 1.00 | 0.69 | 24.71 | 0.92 |
| | 400 | 100 | 76.89 | 42.97 | 28.06 | 0.89 | 0.62 | 24.52 | 0.92 |
| | 100 | 100 | 71.58 | 18.99 | 12.33 | 0.41 | 0.21 | 24.37 | 0.92 |
| | 700 | 50 | 73.75 | 22.92 | 14.83 | 0.45 | 0.32 | 24.43 | 0.92 |
| | 400 | 50 | 76.51 | 21.19 | 13.81 | 0.45 | 0.29 | 24.27 | 0.95 |
| TSFC | 700 | 100 | 57.32 | 1.90 | 0.89 | 0.03 | 0.00 | 20.61 | 0.98 |
| | 400 | 100 | 60.20 | 1.40 | 0.62 | 0.02 | 0.00 | 20.58 | 0.98 |
| | 100 | 100 | 62.40 | 0.77 | 0.28 | 0.00 | 0.00 | 15.93 | 0.95 |
| | 700 | 50 | 60.72 | 0.99 | 0.45 | 0.017 | 0.00 | 18.07 | 0.97 |
| | 400 | 50 | 56.57 | 0.74 | 0.31 | 0.00 | 0.00 | 15.38 | 0.97 |

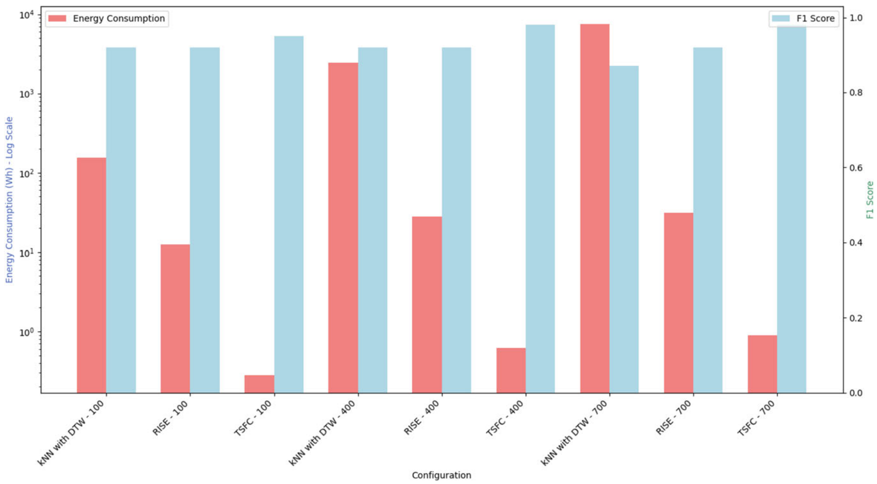



**Fig. 5** Energy consumption and performance analysis of ML algorithms

fewer estimators. RISE shows a consistent pattern of CPU usage across different configurations. A sample length of 400 with 50 estimators provides the best balance of high accuracy (F1 score of 0.95) with reduced energy and time consumption. TSFC shows remarkable energy efficiency across all sample lengths, with a significant decrease in energy consumption as sample length decreases. This suggests that TSFC is particularly suited for energy-efficient processing of time series data. The F1 score is very high across different configurations, indicating that TSFC does not compromise on predictive performance even when optimizing for energy efficiency. Similar to RISE, reducing the number of estimators for TSFC results in lower energy consumption without a notable drop in F1 score. This efficiency is especially remarkable, given the already low energy consumption of TSFC, underscoring its suitability for energy-constrained scenarios. TSFC presents an optimal scenario where energy efficiency and high predictive performance coexist. It demonstrates that careful algorithm design and parameter tuning can achieve high accuracy in ML tasks without incurring high computational costs.

## 5 Conclusion and Future Work

Our findings revealed a nuanced relationship between batch size and energy efficiency, where larger batch sizes led to increased mean power consumption but also enhanced energy efficiency per data point, especially with larger datasets. The study highlighted the importance of context-specific batch size selection, as the optimal balance between energy efficiency and performance varies across different scenarios. Investigating the impact of training dataset size, we observed a direct correlation between increased dataset sizes and higher energy requirements, alongside an improvement in model accuracy. This underscores a crucial trade-off between energy consumption and model performance, indicating that optimizing training processes necessitates a careful consideration of dataset size. Our analysis across various DL architectures demonstrated that each model exhibits its own optimal configuration that harmonizes energy consumption with performance. This finding challenges the conventional wisdom that more complex models are inherently more energy-intensive, advocating for a tailored approach to training parameter selection. In examining AI solutions for time series sensor data, strategic adjustments in sample lengths and the number of estimators/neighbors showed significant impacts on energy efficiency and model accuracy. This suggests that for energy-sensitive applications, choosing the right parameters can lead to substantial energy savings without major accuracy sacrifices.The collective insights from our research emphasize the critical role of strategic parameter selection in achieving energy-efficient ML practices. In our future work, we will investigate the cause of the observed GPU usage peak at the 80% dataset size, examining factors such as hardware configuration and potential system bottlenecks. Additionally, since our current energy efficiency assessment is based on specific hardware configurations, we recognize the importance of broadening the scope to improve generalizability. To achieve this, we plan to expand the datasets and include additional models in our analysis.

**Acknowledgment** This project was funded by German Federal Ministry for the Environment, Nature Conservation, Nuclear Safety, and Consumer Protection (BMUV) project "KIRA" under Grant 67KI32013 and Climate Action (BMWK) project "EASY" under Grant 01MD22002D. Parts of the text have been enhanced and linguistically revised using AI tools. We gratefully acknowledge students Marvin Schacht for providing necessary resources to accomplish this research work.

## References

1. Yarally, T., Cruz, L., Feitosa, D., Sallou, J., van Deursen, A.: Uncovering energy-efficient practices in deep learning training: Preliminary steps towards green ai. In: 2023 IEEE/ACM 2nd International Conference on AI Engineering - Software Engineering for AI (CAIN), pp. 25–36, IEEE Computer Society (2023)
2. Xu, T.: These simple changes can make ai research much more energy efficient. https://www.technologyreview.com/2022/07/06/1055458/ai-research-emissions-energy-efficient/. Accessed: 2023-12-01
3. Shrinking deep learning's carbon footprint. https://news.mit.edu/2020/shrinking-deep-learning-carbon-footprint-0807/. Accessed: 2023-11-01
4. Power-hungry ai: Researchers evaluate energy consumption across models. https://cse.engin.umich.edu/stories/power-hungry-ai-researchers-evaluate-energy-consumption-across-models. Accessed: 2023-12-11
5. Shiqiang, W.: Efficient deep learning. Nat. Comput. Sci. **1**, 181–182 (2021)
6. Luo, T., Wong, W.-F., Goh, R.S.M., Do, A.T., Chen, Z., Li, H., Jiang, W., Yau, W.: Achieving green ai with energy-efficient deep learning using neuromorphic computing. Commun. ACM **66**, 52–57 (2023)
7. Tu, X., Mallik, A., Chen, D., Han, K., Altintas, O., Wang, H., Xie, J.: Unveiling energy efficiency in deep learning: Measurement, prediction, and scoring across edge devices. In: 2023 IEEE Symposium on Edge Computing, pp. 80–93 (2023)
8. Strubell, E., Ganesh, A., McCallum, A.: Energy and policy considerations for modern deep learning research. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, pp. 13693–13696 (2020)
9. Lazzaro, D., Cinà, A., Pintor, M., Demontis, A., Biggio, B., Roli, F., Pelillo, M.: Minimizing energy consumption of deep learning models by energy-aware training, pp. 515–526. arXiv e-prints, 09 2023
10. García-Martín, E., Rodrigues, C.F., Riley, G., Grahn, H.: Estimation of energy consumption in machine learning. J. Parallel Distribut. Comput. **134**, 75–88 (2019)
11. Georgiou, S., Kechagia, M., Sharma, T., Sarro, F., Zou, Y.: Green ai: do deep learning frameworks have different costs?. In: Proceedings of the 44th International Conference on Software Engineering, ICSE '22, (New York, NY, USA), pp. 1082–1094, Association for Computing Machinery (2022)
12. Verdecchia, R., Cruz, L., Sallou, J., Lin, M., Wickenden, J., Hotellier, E.: Data-centric green ai an exploratory empirical study. In: 2022 International Conference on ICT for Sustainability (ICT4S), (Los Alamitos, CA, USA), pp. 35–45. IEEE Computer Society (2022)
13. Gholami, S., Omar, M.: Can pruning make large language models more efficient? (2023)
14. Wang, X., Wang, H., Bhandari, B., Cheng, L.: Ai-empowered methods for smart energy consumption: a review of load forecasting, anomaly detection and demand response. Int. J. Precis. Engin. Manufact.-Green Technol. 1–31 (2023)

15. Naumann, S., Dick, M., Kern, E., Johann, T.: The greensoft model: a reference model for green and sustainable software and its engineering. Sustain. Comput.: Inf. Syst. **1**(4), 294–304 (2011)
16. Mancebo, J., Calero, C., García, F.: GSMP: Green Software Measurement Process, pp. 43–67. Springer International Publishing, Cham (2021)
17. Johann, T., Dick, M., Naumann, S., Kern, E.: How to measure energy-efficiency of software: metrics and measurement results. In: 2012 1st International Workshop on Green and Sustainable Software, GREENS 2012 - Proceedings (2012)
18. Zhong, Y.S., Zhang, T., Ronzoni, G.: Sentimental analysis of facebook reviews: does hospitality matter in senior living?. Int. Hospit. Manag. **112**, 103384 (2023). Hospitality in Healthcare
19. Ji, Y., Wu, W., Zheng, H., Hu, Y., Chen, X., He, L.: Is chatgpt a good personality recognizer? a preliminary study (2023)
20. Deng, Y., Zhang, W., Pan, S.J., Bing, L.: SOUL: towards sentiment and opinion understanding of language. In: The 2023 Conference on Empirical Methods in Natural Language Processing (2023)
21. Chen, J., Ye, X., Sun, J., Li, C.: Towards energy-efficient sentiment classification with spiking neural networks. In: Iliadis, L., Papaleonidas, A., Angelov, P., Jayne, C. (eds.) Artificial Neural Networks and Machine Learning - ICANN 2023, pp. 518–529. Springer Nature Switzerland, Cham (2023)
22. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778 (2016)
23. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556 (2014)
24. Huang, G., Liu, Z., Maaten, L.V.D., Weinberger, K.Q.: Densely connected convolutional networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2261–2269, IEEE Computer Society (2017)
25. Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H.: Mobilenets: efficient convolutional neural networks for mobile vision applications. CoRR, arXiv:abs/1704.04861 (2017)
26. Tang, J.: Intelligent Mobile Projects with TensorFlow, Chap. 2. Pact Publishing (2018)
27. sktime. https://www.sktime.net/en/stable/. Accessed: 2024-01-08
28. imdb-dataset-of-50k-movie-reviews. https://www.kaggle.com/datasets/lakshmi25npathi/imdb-dataset-of-50k-movie-reviews. Accessed: 2023-11-01
29. Bert for sentiment-analysis. https://github.com/chriskhanhtran/bert-for-sentiment-analysis. Accessed: 2023-10-01
30. Esp32. https://www.espressif.com/en/products/socs/esp32
31. Mqtt. https://mqtt.org/. Accessed: 2024-01-08
32. Kern, E., Hilty, L.M., Guldner, A., Maksimov, Y.V., Filler, A., Gröger, J., Naumann, S.: Sustainable software products-towards assessment criteria for resource and energy efficiency. Futur. Gener. Comput. Syst. **86**, 199–210 (2018)
33. Guldner, A., Murach, J.: Measuring and assessing the resource and energy efficiency of artificial intelligence of things devices and algorithms. In: Wohlgemuth, V., Naumann, S., Behrens, G., Arndt, H.-K., Höb, M. (eds.) Advances and New Trends in Environmental Informatics, pp. 185–199. Springer International Publishing, Cham (2023)
34. Guldner, A., Bender, R., Calero, C., Fernando, G.S., Funke, M., Gröger, J., Hilty, L.M., Hörnschemeyer, J., Hoffmann, G.-D., Junger, D., Kennes, T., Kreten, S., Lago, P., Mai, F., Malavolta, I., Murach, J., Obergöker, K., Schmidt, B., Tarara, A., De Veaugh-Geiss, J.P., Weber, S., Westing, M., Wohlgemuth, V., Naumann, S.: Development and evaluation of a reference measurement model for assessing the resource and energy efficiency of software products and components-green software measurement model (gsmm). Futur. Gener. Comput. Syst. **155**, 402–418 (2024)

# Towards an Energy Estimation Model for ABAP Statements on SAP S/4HANA Software

**Ann-Christin Fleischle, Laurenz Baumgart, Holger Wittges, and Helmut Krcmar**

**Abstract** Business transactions are mainly conducted using enterprise software. The use of this software is responsible for a large percentage of the carbon emissions of an enterprise software provider such as SAP. To understand and reduce the energy consumption induced by software, quantifying it and relating it to the source code is crucial. This paper addresses this imperative by presenting a modular approach to measure the energy consumption of code statements written in SAP's proprietary programming language ABAP. Various measurement approaches were evaluated, leading to the development of a prototype based on an estimation model. The prototype, implemented in an SAP S/4HANA environment running on an IBM Power architecture, demonstrates a mean absolute percentage error of less than 1%. Initial measurements show that it is possible to measure differences in energy consumption between ABAP implementations, such as iterative versus recursive approaches for solving the Fibonacci problem. This research contributes an evaluated prototype tailored to ABAP statements and serves as a valuable resource for future empirical investigations in the domain.

**Keywords** Software energy consumption · Energy estimation model · Energy measurement approach

A.-C. Fleischle (✉) · L. Baumgart · H. Wittges · H. Krcmar
Technical University of Munich, Munich, Germany
e-mail: ann-christin.fleischle@tum.de

L. Baumgart
e-mail: laurenz.baumgart@tum.de

H. Wittges
e-mail: holger.wittges@tum.de

H. Krcmar
e-mail: helmut.krcmar@tum.de



V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_15

## 1 Introduction

The increasing reliance on information and communication technology (ICT) for digital innovation supporting economic growth [1] has a significant environmental impact, particularly in terms of energy consumption. According to Freitag et al. [2], peer-reviewed studies estimate ICT's share of global carbon emissions between 1.8% and 2.8%, with 0.6 to 0.9 gigatons in 2021. Podder et al. [3] estimate it to reach 14% by 2040. Among the various components of ICT, software plays a crucial role in determining the energy efficiency of hardware systems [4]. Energy smells in the source code can lead to significant variations in energy consumption [5]. Capra et al. [6] show that energy consumption for the same functionality in different enterprise software systems can vary by up to 145%.

Although such inefficiency is particularly relevant in large-scale systems like enterprise software such as SAP S/4HANA, research in this area with industrial relevance is low [7]. SAP SE [8] reports that 81% of the company's carbon emissions are attributed to the use of their sold software, highlighting the importance of optimizing software for energy efficiency. Since 77% of global transactions are processed using SAP systems [9], even minor improvements in software energy efficiency can significantly reduce global ICT-related carbon emissions.

To improve software energy efficiency, it is essential to identify energy smells [10, 11]. This requires measuring software energy consumption and relating it to the source code. Research investigates software energy consumption at different levels of granularity and in the context of other software systems. The researchers use various approaches to measure software energy consumption and to relate the energy consumption to the source code. Because of SAP S/4HANA's unique architecture, a specialized approach to determining the energy consumption of ABAP source code is required. Thus, we raise the following research question: How can a measurement approach be implemented to obtain the energy consumption of ABAP statements?

This research paper evaluates measurement approaches used in related work. Based on that, it proposes a measurement approach that applies an estimation model to measure the software energy consumption of an SAP S/4HANA system and relates the energy consumption to ABAP statements. To evaluate the developed measurement approach's applicability and underline its relevance, we carried out initial microbenchmarks demonstrating the varying influence of ABAP statements fulfilling the same functionality on the system's energy consumption. Our results contribute to green coding research by providing a measurement approach for a relevant use case. Future research can use the resulting energy estimation model approach to empirically investigate energy smells in ABAP-based SAP systems.

## 2 Background and Related Work

ICT consumes energy in two ways: while running software processes and while idling. Energy consumption while idling can be improved by optimizing hardware [12] and using virtualization [13]. However, executing software influences a system's hardware, which consumes energy [14]. Hardware components such as CPU, memory, disk storage, network adapters, and I/O ports are thus responsible for energy consumption [15, 16]. Each component completes work at a specific power consumption rate measured in watts (W). Energy is defined as power consumption over a certain time period measured in joules (J) [17].

Research uses different approaches to measure a software's energy consumption at various levels of granularity. Reviewing related work, we identified three approaches: hardware meters, software meters, and estimation models. While the estimation model approach is most used in the papers analyzed, especially for more granular measurements, each approach has advantages and disadvantages. Hardware meters are physically switched between the server plug and the power outlet, measuring the power running through. Thus, hardware meters precisely determine a complete server's energy consumption [18]. Only a few authors manage to measure the energy consumption of individual hardware components using hardware meters. Rieger and Bockisch [19] tried to map hardware meter energy measurements to code using machine learning techniques but delivered inconclusive results. Software meters provide software libraries to obtain a server's energy consumption. Intel RAPL is the software meter most represented in the literature reviewed [20–23]. However, it only works for Intel architectures developed after Intel Sandy Bridge. To overcome manufacturer-dependent software tools, the Intelligent Platform Management Interface (IPMI) was developed [24]. The IPMI accesses a server's Baseboard Management Controller, which monitors and manages system resources, including energy consumption. An advantage of software meters is that they are immediately available and come with low overhead during use [25]. However, Fieni et al. [26] state that software meters are often less accurate than physical measurements, especially when reading frequency increases. Kavanagh et al. [27] discuss the inaccuracies of IPMI-based power sensors, including errors due to averaging values and sensor inaccuracies.

While hardware meters and software meters provide energy values for entire computing systems, some purposes, such as optimizing software energy consumption, require a more granular understanding of energy consumption [15]. Noureddine et al. [28] argue that the energy consumption of individual software services can only be determined by applying estimation models. Estimation models use Hardware Performance Counters (HwPCs) that can be obtained at individual software service levels. Such HwPCs are then correlated with energy measurements through regression and machine learning techniques. The resulting model predicts energy based on the obtained HwPCs. HwPCs are widely supported on various operating systems (OS) [29]. Bellosa [30] shows that such models can be achieved with only a few HwPCs. Bircher and John [31] show that error rates of less than 10% can be achieved with simple linear regression models. Most authors use regression techniques to correlate energy and HwPCs. This result aligns with another review paper from García-Martín et al. [29]. Besides the possibility of predicting energy consumption with low error rates, research from Do et al. [32], Noureddine et al. [28], and Müller et al. [24] showed that estimation models can separate the energy behavior of individual software processes even when executed concurrently. This is a significant advantage over hardware and software meters, as they usually provide aggregated measurements on a processor or server level [25]. Fieni et al. [26] present a self-learning energy model that does not require any training. The prediction error is continuously assessed, and the model is improved. This approach achieves one of the lowest prediction errors with little overhead but requires extensive programming effort.

Besides the three energy measurement approaches, we identified two approaches to mapping the energy values to the source code: execution traces can be mapped to the energy values, or microbenchmarks can be conducted. Execution traces can be obtained through bytecode or source code instrumentation, or kernel traces can be used. Approaches from Noureddine et al. [18] and La Fosse et al. [21] instrument the bytecode to flag the start and exit points of executed methods and relate them to energy consumption. Instrumenting the bytecode of an application has the advantage of not having to modify the source code. This makes post-execution analysis easier and the code better to maintain. However, it seems less accurate than source code instrumentation [18]. An approach from Pathak et al. [33] uses source code instrumentation to mark methods' start and exit points during execution. Like bytecode profiling, this method creates an overhead but allows a straightforward annotation, which makes it easily understandable for humans. However, compilers often optimize written code, and the code injections would be part of the optimization [34]. Another technique, kernel-level tracing, is frequently naturally supported by modern computing infrastructures. The execution environment provides the possibility to trace code execution without further annotating it. In a recent paper from Alvi et al. [35], the kernel trace inherently supported by Android is used to map method-level traces to energy consumption. This approach has an approximate error rate of 5% and allows to identify methods with the need for energy improvement. Noureddine [36] leverages the Java Virtual Machine to read each thread's method stack trace and relates it to measured energy values. This method allows real-time energy consumption monitoring per method with a prediction error of around 0.3 to 3.8%. Kernel-level tracing provides a promising approach for obtaining highly accurate energy consumption on a code level. Due to its natural support in many programming languages, they come with low overhead, and trace dumps can be created on a microsecond level. If the energy consumption of more granular source code snippets is to be measured, researchers conduct microbenchmarks executing these granular source code snippets repeatedly and averaging the corresponding energy values. An example is the Computer Language Benchmark Game, which compares various implementations of the same problem in different programming languages [37]. Another publication from Kansal and Zhao [38] measures the energy consumption of different implementations of compression algorithms in Java to guide developers on which library to choose at design time. Such approaches provide insights on how to code more energy efficiently based on reference implementations.

## 3 Research Design and Setup

For this research, we used an SAP S/4HANA 2020 system. The system runs on an IBM Power8 Tyan server[1] with a SUSE Linux OS installed. The server has a 1200W redundant power supply unit and an integrated management port for IPMI. SAP Systems are implemented according to the client–server architecture. At the core of most SAP Systems is the NetWeaver application server (AS) [39]. SAP work processes (WPs) that execute every step of an ABAP program are essential to every SAP NetWeaver AS [39]. ABAP is SAP's proprietary programming language for developing and customizing business applications, including SAP S/4HANA deployed on the SAP NetWeaver AS ABAP is a high-level programming language that supports interoperable object-oriented and procedural programming models [40]. As described in Schneider [40], ABAP programs that the WP will execute are stored inside the extended memory of the AS. If a program was not previously executed, it will be read to the program buffer before it is available to be executed by the WP. During execution, the system creates statistical records stored in the AS database (DB). Statistical records include KPIs such as runtime and transmitted data amount that inform about the runtime behavior of a program. The information is provided by the C Kernel of the AS [40].

We followed the common modeling steps to implement an energy estimation model for ABAP statements on an SAP S/4HANA system. First, we defined the model inputs, in our case HwPCs. Second, training data was required to help the estimation model understand the relation between HwPCs and energy consumption. Once the estimation model was trained, another data set was needed to validate the model. For both data sets, HwPCs and corresponding actual energy values were required. For this, load on the system can either be created through sampling application data or stressing hardware components. A challenge with sampling application data is to create a dataset that fully utilizes existing hardware components. For example, the estimation model in Müller et al. [24] only predicts energy for up to 15% of DB usage due to training data limitations. Noureddine et al. [28] and Colmant et al. [41], however, use tools like Linux stress that create workloads for the individual hardware components of a server. The stress command varies the utility rate of the hardware component. Training with up to 100% resource utilization makes the energy model robust for varying workloads but requires a more extensive training period. Depending on the load generation and hardware setup, different software tools are available to extract the HwPCs. At the same time, software or hardware meters can be used to measure actual energy consumption [27]. Although such meters cannot

[1] mitaccomputing.com/Tyan/Barebones_TN71-BP012_BSP012T71V14HR-4 T-5_EN ~ Spec.

measure energy with the frequency or granularity of estimation models, their accurate measurements can assess prediction performance [28]. To validate estimation model performance, we used the measured and predicted values to calculate statistical metrics like the coefficient of determination ($R^2$), the Mean Absolute Error (MAE), the Mean Absolute Percentage Error (MAPE), the Mean Square Error (MSE) and the Root Mean Square Error (RMSE). Each of these metrics serves a distinct purpose in evaluating an estimation model. $R^2$ measures the overall goodness of fit, MAE and MAPE quantify average absolute and relative prediction errors, and MSE and RMSE take outliers into account [25, 35, 42]. Together, these metrics provide a comprehensive assessment of the model's predictive performance and error characteristics when estimating energy consumption.

To evaluate the applicability of the developed measurement approach and to underline its relevance, we conducted initial microbenchmarks that demonstrated the effect of programming concepts on energy consumption. We implemented ABAP statements fulfilling the same functionality but using different programming concepts. Running the implemented approach then provided us with an average energy consumption per ABAP statement.

## 4 Implementation of Measurement Approach

Based on the analysis of related work, we chose an estimation model approach to obtain energy consumption based on HwPCs per OS process. OS processes can be directly mapped to SAP WPs to relate the energy values to the source code. This is possible because one WP is responsible for executing one ABAP program at a time. Thus, as depicted in Fig. 1, four components are required to obtain the energy consumption of an executed ABAP program: HwPCs per OS process, an energy estimation model, the WP ID of the executed program, and a function to match the OS process ID to the WP ID. This way, the model estimates energy consumption based on the HwPCs, and the mapping of OS process ID to WP ID results in the energy consumption per ABAP program.

First, we implement the central component of this approach, the energy estimation model. The Linux stress command, also used by Noureddine et al. [28], synthetically creates CPU, memory, I/O, and disk workload on the available hardware resources. Varying the number of workers that execute the command achieves different levels of workload up to 100% CPU utilization. The Linux top command shows the HwPCs describing the system's resource consumption. The result is depicted in Fig. 2.

Linux divides the system load into user load (us), system load (sy), niced user processes load (ni), idle time (id), I/O wait time for completion (wa), time spent for servicing hardware interrupts (hi), time spent for servicing software interrupts (si) and time stolen from this virtual machine by the hypervisor (st). Figure 2 shows that different stress periods create different CPU loads in the system. The I/O operations require system load, the CPU phase creates user load, the disk phase uses wait load, and the memory allocation increases user and system load. While stressing the

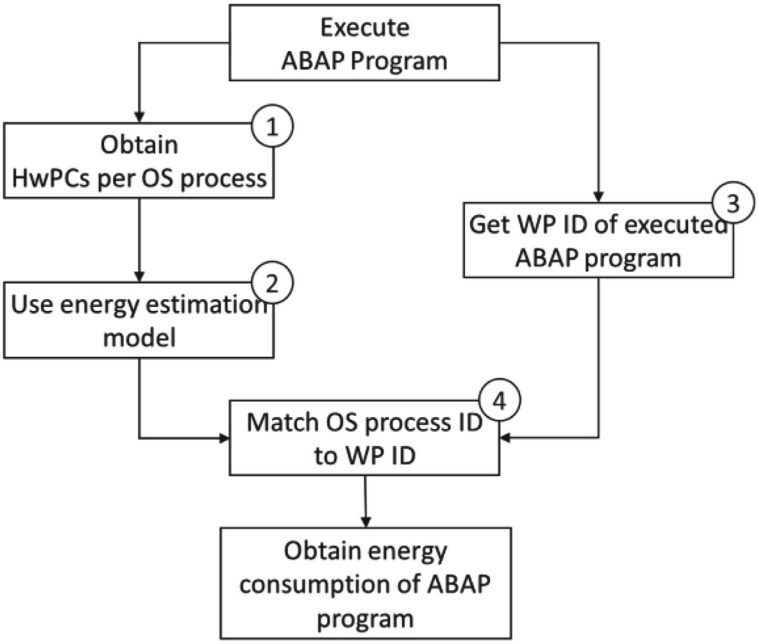



**Fig. 1** Components of an energy estimation approach for ABAP programs

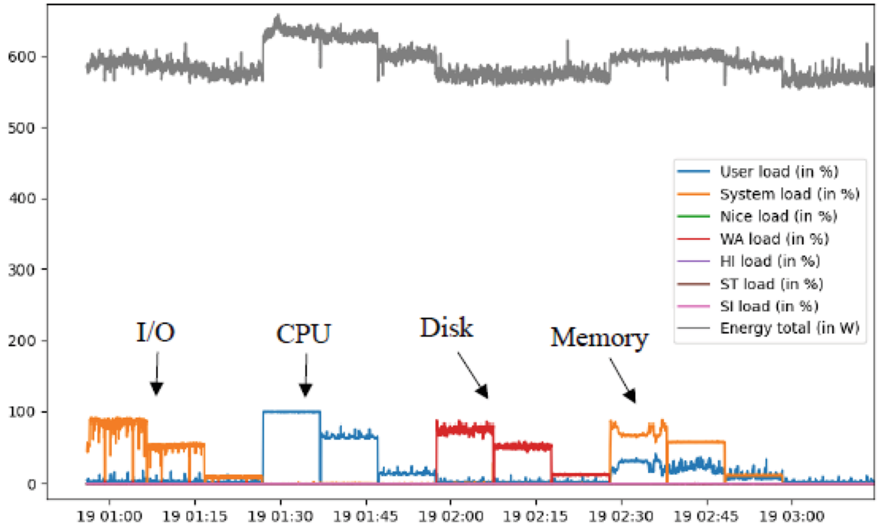



**Fig. 2** Synthetic load generated on the SAP S/4HANA system

**Table 1** Validation metrics obtained after training the model

| Metric | Value |
|---|---|
| $R^2$ | 0.720 |
| MAPE (in %) | 0.940 |
| MAE (in W) | 5.440 |
| MSE | 57.939 |
| RMSE | 7.612 |

CPU increases the system's energy consumption the most, memory allocation also increases energy consumption. I/O and disk operations, on the other hand, barely increase energy consumption. The graphs indicate two relevant HwPCs responsible for the system's energy increase: user and system loads. Although the wait time is high during disk operations, it does not increase energy consumption.

At the same time, two GUDE hardware meters[2] intercept the redundant power supply unit of the server and measure the actual energy consumption. The hardware meters provide APIs to extract the energy values. The HwPCs, system and user load, together with the actual energy values, serve as training data for the creation of the model.

With the training data at hand, we chose a simple linear regression to correlate the HwPCs user load and system load with the energy values to create the energy estimation model using an 80/20 train/test split. The Python library scikit-learn offers functions to correlate independent with dependent variables. After correlating the training data based on the residual sum of squares, the resulting energy formula is

$$E_t = 568.74 + 0.797 * US_t + 0.252 * SY_t$$

The energy formula consists of two parts: static energy consumption and variable energy consumption. The calculated idle energy consumption of the server is around 568 W. Per one unit increase in user load ($US_t$); the energy increases by 0.797 W and per one unit increase in system load ($SY_t$); the energy increases by 0.252 W. The user load is responsible for more than three times the energy consumption than the system load. This is in line with previous observations from Fig. 2.

After creating the model, we use it with a new data set to assess its accuracy and validate it. Accuracy describes how well the model estimates energy consumption based on the HwPCs. Good model fit and low error rates create high accuracy of the respective model and indicate a well rebuilt of actual energy measurements. Again, the scikit-learn library offers functions to calculate $R^2$, MAPE, MAE, MSE, and RMSE based on the estimated and measured energy values of the dataset. The results can be found in Table 1. The MAPE is less than 1% between estimated and measured values. The MAE between estimated and measured values is around 5W. An $R^2$ of 0.72 indicates a good correlation between the HwPCs and energy consumption (Table 1).

[2] https://gude-systems.com/en/products/expert-power-control-1105/

**Table 2** Energy consumption of insert and read operations on sorted vs. hashed tables gehört ins nächste kapitel 5 evaluation of measurement approach

| | Insert | | Read | |
|---|---|---|---|---|
| | Sorted Table | Hashed table | Sorted table | Hashed table |
| Number of runs | 2000 | 3500 | 2000 | 3000 |
| Duration (in seconds) | 32.382 | 36.833 | 40.131 | 31.081 |
| Total energy consumption (in J) | 30.601 | 34.918 | 38.967 | 30.118 |
| Energy consumption per run (in J) | 0.0153 | 0.0100 | 0.0195 | 0.0100 |
| Delta in energy consumption (hashed table as compared to sorted table in %) | –34.64 | | –48.72 | |

Having created and validated the energy estimation model, we integrated it with the other components to obtain the energy consumption of an ABAP statement. Since ABAP statements often execute faster than the measurement intervals, we use microbenchmarks. A microbenchmark consists of an ABAP program only containing the statement to be evaluated. The ABAP statement is executed repeatedly to gather at least 30 measurement points, as suggested by Koedijk and Oprescu [43]. Collecting around 30 measurement points results in a varying number of runs for different ABAP statements. To avoid the overhead of loading the ABAP program from the AS DB to the buffer, the program must be executed once before conducting the measurements. Once the microbenchmark is prepared and loaded into the buffer, collecting the HwPCs during its execution reflects the load generated by the microbenchmark. We used source code instrumentation to collect the information required to relate the energy values to the ABAP statement. This includes the ID of the WP that executed the ABAP program, the timestamps of the start and end of the execution, and the number of iterations of the ABAP statement. The energy estimation model requires the collected HwPCs as input. Since the HwPCs are collected per OS process per measuring point, the energy estimation model results in the corresponding energy values. OS processes can be directly mapped to SAP WPs. We mapped the OS process ID to the WP ID and summed up the energy values of all measurement points between the start and end time of the ABAP program execution. This value reflects the variable amount of energy consumed throughout all iterations of the ABAP statement, neglecting the fixed energy consumption of the server. Finally, we divided the energy consumption by the number of iterations to obtain the average energy consumption for one iteration of the ABAP statement.

## 5 Evaluation of Measurement Approach

To evaluate the applicability of the developed energy estimation approach, we estimated the energy consumption of some exemplary ABAP statements. Since the developed energy estimation model does not include the energy consumed by the DB server, we focused on statements that create load on the AS. We followed the process flow described above. In the following, we report some interesting findings to demonstrate the relevance and effectiveness of the measurement approach developed.

ABAP developers can choose between data structures such as sorted and hashed tables to store temporary data. Various authors [43–45] already found that different data structures and their operations consume varying amounts of energy. We compare the energy consumption of inserting data into and reading data from hashed and sorted tables. The results can be found in Table 2. Both inserting and reading data are more energy efficient when performed on hashed tables, respectively consuming 34.64% and 48.72% less energy than for sorted tables.

Unlike hashed and sorted tables, standard tables do not create an order or index. To create an order, a sorted table is used, or a standard table is sorted using a sort algorithm. Table 3 shows that the latter consumes 32.28% less energy. Constantly keeping items in order seems to be energy-intensive.

Another interesting example is the comparison of an iterative and a recursive implementation of the Fibonacci problem. The iterative implementation consumes 99.78% less energy than the recursive implementation. The numbers can be found in Table 4.

We can show that although fulfilling the same functionality, different programming concepts can consume varying amounts of energy.

**Table 3** Energy consumption of insert operation on standard table being sorted vs. sorted table

| | Insert into standard table, then sort | Insert into sorted table |
|---|---|---|
| Number of runs | 3000 | 2000 |
| Duration (in seconds) | 34.362 | 32.863 |
| Total energy consumption (in J) | 31.991 | 31.581 |
| Energy consumption per run (in J) | 0.0107 | 0.0158 |
| Delta in energy consumption (standard table as compared to sorted table in %) | –32.28 | |

**Table 4** Energy consumption of iterative vs. recursive implementation of the Fibonacci problem

| | Iteration | Recursion |
|---|---|---|
| Number of runs | 20,000 | 45 |
| Duration (in seconds) | 34.685 | 32.295 |
| Total energy consumption (in J) | 32.881 | 32.443 |
| Energy consumption per run (in J) | 0.0016 | 0.7210 |
| Delta in energy consumption (iteration as compared to recursion in %) | -99.78 | |

## 6 Discussion

In this work, we implemented an energy estimation approach for an SAP S/4HANA environment based on findings from previous work in the context of other software and hardware environments. We answer the research question: *How can a measurement approach be implemented to obtain the energy consumption of ABAP statements?* Reviewing related work led us to the conclusion that estimation models are most suitable for granular measurements. In addition to the choice of measurement approach we were able to derive further recommendations for our implementation. Applying these findings, we developed a measurement approach for an SAP S/4HANA environment that accurately estimates the energy consumption of ABAP statements.

Considering only CPU as input to estimate energy consumption is in line with research from Ferreira et al. [10] and Pereira et al. [45]. Both show that the CPU causes around 90% of energy consumption. The validation of our model also shows that it is a reliable input parameter. However, we built the model in a way that it can be extended with other HwPCs in the future. For example, Noureddine et al. [28], Fieni et al. [26], and Do et al. [32] also consider the effects of hard disk, memory, and network. Further, Fieni et al. [46] show that considering operating frequency could also reduce the estimation error to around 2 W.

Koedijk and Oprescu [43] find that there exist examples where less runtime does not necessarily imply less energy consumed. According to Noureddine et al. [28], this can be due to dynamic voltage and frequency scaling, saying that "A process at 100% CPU utilization will not necessarily consume more power than a process running at 50% CPU utilization but with a higher voltage." However, we could not observe such behavior in our environment. This may be due to the specific architecture since each SAP WP consumes 100% of one CPU core, meaning that more runtime also leads to more energy consumed.

When implementing the estimation model approach, we chose between using OS-level or application-level HwPCs. We decided to use OS-level HwPCs since the load was more accessible to generate on the OS level. The SD Benchmark used by Müller et al. [24] or an alternative for application data generation is no longer provided by SAP since SAP S/4HANA. Another reason for our decision was the limited capabilities of this benchmark to achieve various load levels up to 100%. Since the OS-level DB processes are multithreaded, we could not separate the load caused by the evaluated program from other processes running on the server. A single OS process ID represents the DB but comprises multiple workers handling DB requests. Thus, our current measurement approach only reflects the AS load, not the DB load. Müller et al. [24] correlated performance counters available in the SAP system, providing information about the DB load at the transaction level with energy consumption, and achieved promising results.

With their approach, Müller et al. [23] can profile the energy consumption of a whole SAP process at the transaction level. With the current approach, the most granular level we could achieve is the program level. Thus, estimating energy consumption for code statements is only possible by creating microbenchmarks. This seems to be limited first by the granularity of the HwPCs, which are only available on the WP level and can only be traced back to the program level. Another limiting factor seems to be the hardware meter, which can measure energy consumption only once per second. Therefore, training an estimation model for more granular units, such as code statements that only take a fraction of that to execute, is not straightforward.

To evaluate the application of the developed measurement approach, we chose code statements that only create AS load for the microbenchmarks. For example, we show that the iterative implementation of the Fibonacci Problem consumes 99.78% less energy than the recursive implementation. This indicates a contradiction to the finding of Noureddine et al. [28], who show that for the Towers of Hanoi, the iterative implementation consumes more energy than the recursive implementation. As other authors [28, 43, 45] have stated, our results indicate that the relative energy consumption of different programming concepts can differ depending on the programming language. While we achieved interesting results, the experiments could be extended to other programming concepts and to include more variations of the same programming concept. For example, different datasets for the same operation on the same table type can be used to examine whether there are differences in energy consumption. In an example from Noureddine et al. [47], the authors created a unit test for Java code examples that varies the input datasets. The findings show that larger datasets consume more energy; however, not always in a linear dependency.

## 7 Conclusion

This work addresses the need to understand the energy behavior of software to mitigate carbon emissions by reducing software energy consumption. Measuring and relating software energy consumption to the source code is essential to understand how software causes the overall energy consumption of a server system. This research applied an energy estimation model approach to an ABAP-based SAP system to determine the energy consumption caused by ABAP statements based on the CPU load of the AS. For this, we created a model by correlating HwPCs with actual energy values using a simple linear regression technique. After validating, we evaluated the model's applicability by applying it to microbenchmarks and comparing the energy consumption of different ABAP statements. The resulting measurement approach estimates energy consumption based on HwPCs per OS process collected while running an ABAP program. The OS process ID is mapped to the WP ID to obtain the energy consumption caused by the ABAP program. This approach has a prediction error of less than 1%. Applying the approach to microbenchmarks, we could show that ABAP statements fulfilling the same functionality but using different programming concepts consume different amounts of energy. With this research, we contribute a use case for energy estimation models with industrial relevance. With our approach, various ABAP programming concepts and their influence on software energy consumption can be empirically investigated to provide insights into the energy behavior of ABAP-based SAP systems.

Nevertheless, this research is not without limitations. The developed approach is only based on the CPU load of the AS and, therefore, does not include the energy consumption of the DB. So far, a program is the most granular level to which we could relate the energy consumption. Thus, microbenchmarks must be used to evaluate the energy consumption of ABAP statements, and profiling of whole software processes is not possible. Based on our findings and limitations, we have several suggestions for further research. Future research should explore using other performance counters, such as SAP's performance statistics, to include the DB energy consumption in the model. Further, HwPCs other than CPU could be included in the model to make it more robust. Finally, future research could conduct further microbenchmarks to investigate the energy consumption of programming concepts implemented in ABAP and be able to derive developer recommendations. It would also be interesting to explore how robust the developed approach is when applied to ABAP-based SAP systems running on other hardware architectures and how the energy consumption of ABAP statements behaves across these different hardware setups.

## References


1. EU4Digital: ICT Innovation is clearly vital to economic development in the digital age: How far behind are the Eastern partner countries compared to the European Union? https://eufordigital.eu/e-card/ict-innovation-is-clearly-vital-to-economic-development-in-the-digital-age-how-far-behind-are-the-eastern-partner-countries-compared-to-the-european-union/ (2020). Accessed 17 August 2024
2. Freitag, C., Berners-Lee, M., Widdicks, K., Knowles, B., Blair, G.S., Friday, A.: The real climate and transformative impact of ICT: A critique of estimates, trends, and regulations. Patterns **2** (2021). https://doi.org/10.1016/j.patter.2021.100340
3. Podder, S., Burden, A., Kumar Sing, S., Maruca, R.: How green is your software? https://hbr.org/2020/09/how-green-is-your-software (2020). Accessed 17 August 2024
4. Kern, E., Guldner, A., Naumann, S.: Including software aspects in green IT: How to create awareness for green software issues. In: Kharchenko, V., Kondratenko, Y., Kacprzyk, J. (eds), Green IT Engineering: Social, Business and Industrial Applications. Springer, Cham (2019)
5. Palomba, F., Di Nucci, D., Panichella, A., Zaidman, A., de Lucia, A.: On the impact of code smells on the energy consumption of mobile applications. Info. Software Techno. **105** (2019). https://doi.org/10.1016/j.infsof.2018.08.004
6. Capra, E., Francalanci, C., Slaughter, S.A.: Measuring application software energy efficiency. IT Professional **14** (2012). https://doi.org/10.1109/MITP.2012.39

7. Balanza-Martinez, J., Lago, P., Verdecchia, R.: Tactics for software energy efficiency: A review. In: Wohlgemuth, V., Kranzlmüller, D., Höb, M. (eds) Advances and New Trends in Environmental Informatics 2023. Springer (2024)
8. SAP SE: SAP Integrated Report 2023 (2024)
9. SAP SE: Sustainability. https://community.sap.com/topics/sustainability. Accessed 25 Mar 2023
10. Ferreira, M.A, Hoekstra, E., Merkus, B., Visser, B., Visser, J. Seflab: A lab for measuring software energy footprints. In: Lago, P., Meyer, N., Morisio, M. et al. (eds) Proceedings of the 2nd International Workshop on Green and Sustainable Software (GREENS). IEEE Press (2013)
11. Noureddine, A., Rouvoy, R., Seinturier, L.: A review of energy measurement approaches. Oper. Syst. Rev. **47** (2013). https://doi.org/10.1145/2553070.2553077
12. Shuja, J., Madani, S., Bilal, K., Hayat, K., Khan, S., Sarwar, S.: Energy-efficient data centers. Computing **94** (2012). https://doi.org/10.1007/s00607-012-0211-2
13. Chait, K., Juiz, C.: Research line on improving energy efficiency in web servers. In: 2013 World Congress on Computer and Information Technology (WCCIT). IEEE (2013)
14. Hilty, L.M., Lohmann, W., Behrendt, S., Evers-Wölk, M., Fichter, K., Hintemann, R.: Green software. Final report of the project: Establishing and exploiting potentials for environmental protection in information and communication technology (Green IT). Report commissioned by the Federal Environment Agency, Berlin, Förderkennzeichen 3710 95 302/3 (2015)
15. Singh, V.K., Dutta, K., VanderMeer, D.: Estimating the energy consumption of executing software processes. In: 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing. IEEE (2013)
16. Raisian, K., Yahaya, J., Ahmad Ibrahim, S.R., Deraman, A., Yunos, T.: The green software measurement structure based on sustainability perspective. In: 2021 International Conference on Electrical Engineering and Informatics (ICEEI). IEEE (2021)
17. Ling, S.J., Sanny, J., Moebs, W.: University Physics. OpenStax, Houston (2016)
18. Noureddine, A., Bourdon, A., Rouvoy, R., Seinturier, L.: Runtime monitoring of software energy hotspots. In: Goedicke M, Menzies T, Saeki M (eds) Proceedings of the 27th IEEE/ACM International Conference on Automated Software Engineering. ACM (2012a)
19. Rieger, F., Bockisch, C.: Evaluating techniques for method-exact energy measurements: Towards a framework for platform-independent code-level energy measurements. In: Hung, C-C., Cerny, T. (eds) Proceedings of the 35th Annual ACM Symposium on Applied Computing. ACM (2020)
20. Hähnel, M., Döbel, B., Völp, M., Härtig, H.: Measuring energy consumption for short code paths using RAPL. Perform. Eval. Rev. **40** (2012). https://doi.org/10.1145/2425248.2425252
21. La Fosse, T.B., Mottu, J-M., Tisi, M., Sunyé, G.: Characterizing a source code model with energy measurements. In: Condori-Fernandez, N., Bagnato, A., Kern, E. (eds) Proceedings of the 4th International Workshop on Measurement and Metrics for Green and Sustainable Software Systems (MeGSuS) (2018)
22. Weaver, V.M., Johnson, M., Kasichayanula, K., Ralph, J., Luszczek, P., Terpstra, D., Moore, S.: Measuring energy and power with PAPI. In: Proceedings of the 2012 41st International Conference on Parallel Processing Workshops (ICPPW). IEEE (2012)
23. Marincic, I., Vishwanath, V., Hoffmann, H.: PoLiMEr: An energy monitoring and power limiting interface for hpc applications. In: Proceedings of the 5th International Workshop on Energy Efficient Supercomputing (E2SC). ACM (2017)
24. Müller, H., Goerling, C., Hintsch, J., Splieth, M., Starke, S., Turowski, K.: Monitoring energy consumption on the service level: A procedure model for multitenant ERP systems. In: Cardoso, J., Ferguson, D., Méndez Munoz, V. et al. (eds) Proceedings of the 6th International Conference on Cloud Computing and Services Science (CLOSER). SCITEPRESS (2016)
25. LeBeane, M., Ryoo, J.H., Panda, R., John, L.K.: Watt Watcher: Fine-grained power estimation for emerging workloads. In: Proceedings of the 2015 27th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD). IEEE (2015)

26. Fieni, G., Rouvoy, R., Seinturier, L.: SmartWatts: Self-calibrating software-defined power meter for containers. In: 2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID). IEEE (2020)
27. Kavanagh, R., Armstrong, D., Djemame, K.: Accuracy of energy model calibration with IPMI. In: 2016 IEEE 9th International Conference on Cloud Computing (CLOUD). IEEE (2016)
28. Noureddine, A., Bourdon, A., Rouvoy, R., Seinturier, L.: A preliminary study of the impact of software engineering on GreenIT. In: Kazman, R., Lago, P., Meyer, N., Morisio, M. et al. (eds) Proceedings of the 2012 First International Workshop on Green and Sustainable Software (GREENS). IEEE (2012b)
29. García-Martín, E., Rodrigues, C.F., Riley, G., Grahn, H.: Estimation of energy consumption in machine learning. J. Parallel and Distrib. Comp. **134** (2019). https://doi.org/10.1016/j.jpdc.2019.07.007
30. Bellosa, F.: The benefits of event–driven energy accounting in power-sensitive systems. In: Shapiro, M. (ed) Proceedings of the 9th Workshop on ACM SIGOPS European Workshop (EW). ACM (2000)
31. Bircher, W.L., John, L.K.: Complete system power estimation: A trickle-down approach based on performance events. In: 2007 IEEE International Symposium on Performance Analysis of Systems & Software (ISPASS). IEEE (2007)
32. Do, T., Rawshdeh, S., Shi, W.: pTop: A process-level power profiling tool. In: Proceedings of the 2nd Workshop on Power Aware Computing and Systems (HotPower) (2009)
33. Pathak, A., Hu, C., Zhang, M.: Fine grained energy accounting on smartphones with eprof: Where is the energy spent inside my app? In: Felber, P. (ed) Proceedings of the 7th ACM European Conference on Computer Systems (EuroSys). ACM (2012)
34. Tallent, N.R., Mellor-Crummey, J.M., Fagan, M.W.: Binary analysis for measurement and attribution of program performance. In: Hind, M. (ed) Proceedings of the 30th ACM SIGPLAN Conference on Programming Language Design and Implementation. ACM (2009)
35. Alvi, H.M., Majeed, H., Mujtaba, H., Beg, M.O.: MLEE: Method level energy estimation—A machine learning approach. Sustain. Comput.: Inform. Syst. **32** (2021). https://doi.org/10.1016/j.suscom.2021.100594
36. Noureddine A.: PowerJoular and JoularJX: Multi-platform software power monitoring tools. In: Dustdar, S., Portet, F. (eds) 2022 18th International Conference on Intelligent Environments (IE). IEEE (2022)
37. The Computer Language Benchmarks Game. https://benchmarksgame-team.pages.debian.net/benchmarksgame/index.html. Accessed 17 August 2024
38. Kansal, A., Zhao, F.: Fine-grained energy profiling for power-aware application design. Perform. Eval. Rev. **36** (2008). https://doi.org/10.1145/1453175.1453180
39. SAP SE: Application server - ABAP keyword documentation. https://help.sap.com/doc/abapdocu_752_index_htm/7.52/en-US/abenapplication_server_glosry.htm. Accessed 12 March 2023
40. Schneider, T.: SAP-Performanceoptimierung: Analyse und Tuning von SAP-Systemen, 8th edn. SAP PRESS. Rheinwerk Verlag GmbH, Bonn (2017)
41. Colmant, M., Kurpicz, M., Felber, P., Huertas, L., Rouvoy, R. Sobe, A.: Process-Level power estimation in VM-Based systems. In: Réveillère L (ed) Proceedings of the Tenth European Conference on Computer Systems (EuroSys). ACM (2015)
42. Dolz, M.F., Kunkel, J., Chasapis, K., Catalán, S.: An analytical methodology to derive power models based on hardware and software metrics. Comput. Sci.—Res. Dev. **31** (2016). https://doi.org/10.1007/s00450-015-0298-8
43. Koedijk, L., Oprescu, A.: Finding significant differences in the energy consumption when comparing programming languages and programs. In: Somova, E., Fernandes, J.P. (eds) Proceedings of the 8th International Conference on ICT for Sustainability (ICT4S). IEEE (2022)
44. Munoz, D-J., Pinto, M., Fuentes, L.: Finding correlations of features affecting energy consumption and performance of web servers using the hadas eco-assistant. Computing **100** (2018). https://doi.org/10.1007/s00607-018-0632-7

45. Pereira, R., Couto, M., Ribeiro, F., Rua, R., Cunha, J., Fernandes, J.P., Saraiva, J.: Energy efficiency across programming languages: how do energy, time, and memory relate? In: Combemale, B., Mernik, M., Rumpe, B. (eds) Conference on Systems, Programming, Languages, and Applications (SPLASH). ACM (2017)
46. Fieni, G., Rouvoy, R., Seiturier, L.: SelfWatts: On-the-fly selection of performance events to optimize software-defined power meters. In: 2021 IEEE/ACM 21st International Symposium on Cluster, Cloud and Internet Computing (CCGrid). IEEE (2021)
47. Noureddine, A., Rouvoy, R., Seinturier, L.: Unit testing of energy consumption of software libraries. In: Cho, Y., Shin, S.Y., Kim, S. et al. (eds) Proceedings of the 29th Annual ACM Symposium on Applied Computing (SAC). ACM (2014)

# Sustainability in Transportation and Logistics

# Green Delivery Analytics: Software Vision of a Planning Tool for Sustainable Last Mile Logistics

**Benjamin Wagner vom Berg, Richard Schulte, Mattes Leibenath, Senad Hasanspahic, Arne Kruse, Christoph Drieling, Kian Seelaff, Aina Andriamananony, Uta Kühne, and Franziska Giese**

**Abstract** The field of logistics plays a pivotal role in modern society and the economy. However, it also has a significant environmental impact. In order to address this issue, it is imperative for logistics to become more sustainable. One effective solution is the use of cargo bikes and light electric vehicles for last mile delivery, as they are more eco-friendly compared to traditional delivery vehicles. To optimize the use of cargo bikes and create sustainable last mile logistics, delivery service providers must implement novel optimization strategies that account for the unique characteristics of cargo bikes and city hub architecture. The research project "Green Delivery Analytics" aims to develop a logistics planning tool with the goal of addressing the challenges of sustainable last mile logistics by incorporating advanced data analytics and optimization techniques regarding transport planning and the localization of micro-hubs. This paper presents a software vision for the aforementioned planning tool, developed using the Rational Unified Process (RUP). This process was employed to determine the necessary design of the software to fulfill the necessary functions and business processes. To this end, an analysis of relevant stakeholders and associated data was conducted. The software vision is presented and discussed for further development.

**Keywords** Data Analytics · Optimization Strategies · Software Vision · Logistics Planning · Sustainable Last Mile Logistics · Cargo Bikes · Micro-Hubs

B. Wagner vom Berg · R. Schulte (✉) · M. Leibenath · S. Hasanspahic · U. Kühne · F. Giese
University of Applied Sciences Bremerhaven, Smart Mobility Institute, Bremerhaven, Germany
e-mail: rschulte@hs-bremerhaven.de

A. Kruse · C. Drieling · K. Seelaff
RytleX Group SE, Am Weser-Terminal, Bremen, Germany

A. Andriamananony
COSMO UG Customer-Oriented Sustainable Mobility, Butteldorf, Elsfleth, Germany



V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*,
Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_16

# 1 Introduction

## *1.1 Motivation*

The growing importance of sustainability in logistics represents a pressing challenge, especially in major German cities, where the delivery of parcels on the "last mile" must increasingly be made emission-free. The increasing urbanization and the growing volume and heterogeneity of parcel deliveries in cities [1] lead to congestion, air pollution, and logistical complexity [2]. In this context, the development of innovative solutions for sustainable last mile logistics is of crucial importance [3]. For example, the use of electric cargo bikes in combination with lifting structures leads to less impact on traffic and the environment in cities [4].

The research project "Green Delivery Analytics" aims to contribute to solving this problem by designing and providing a tool for planning and implementing climate-neutral delivery systems in German cities. This paper presents the development of the respective software vision.

## *1.2 Research Project "Green Delivery Analytics"*

The research project, entitled "Green Delivery Analytics" (GDA), is a collaborative endeavor between the Smart Mobility Institute of the University of Applied Sciences in Bremerhaven, Germany, and the RytleX Group SE[1] (Urban Logistics Expert). Funded by the mFUND campaign for initiatives and innovation, a program of the Federal Ministry for Digital and Transport [5].

The GDA project is a data-driven initiative with the objective of enabling the design of carbon–neutral last mile logistics in the 80 largest German cities. The project's primary focus is the development of a partially automated system for planning and implementing climate-neutral delivery solutions. To achieve this, various data sets, including spatial structure data, delivery data, socio-demographic data, and traffic flow data, are integrated to design individual combinations of micro-hubs, e-transporters, and cargo bikes for each city [5].

The project employs an open data approach, whereby data from various sources, such as the Mobilithek,[2] can be utilized and made available to the general public. This enables cities and logistics companies to plan and implement sustainable logistics solutions that are both economically viable and ecologically beneficial [6].

By developing a software prototype, it will be possible to determine specific combinations of delivery systems and necessary infrastructure for each city [5].

1 https://rytlex.com/, last accessed 2024/10/02.

2 https://mobilithek.info/, last accessed 2024/10/02.

Furthermore, findings from previous mFUND projects, such as ZMo[3] and SIP,[4] will be employed to enhance the efficiency and effectiveness of the new system [7].

### 1.3 State of Art

**Sustainable Last Mile Logistics.** The last mile in terms of distribution logistics refers to the final part of the supply chain, which is highly complex due to the handover of goods to recipients of various types, among other things [8, 9]. This is where not only a large proportion of the costs are incurred, but also the emissions, which are described as "expensive and dirty" [9]. This is due to isolation and the number of stops, both of which have increased as a result of consumer trends and further increased the economic pressure on the courier, express, and parcel services (CEP) industry [9]. Optimizations in the CEP industry therefore often relate to the economic aspects [10], but the environmental impact is also receiving increasing attention [9].

The majority of concepts relate to the three dimensions of sustainability: the economic, ecological, and social dimensions [10]. However, the ecological dimension is given a more prominent role in the priority model of sustainability [11].

A variety of systems can be employed to enhance the sustainability of the last mile, with a reduced impact on the three dimensions, particularly the ecological dimension. Currently, there is a significant focus on the use of electrically powered vehicles and micro-hub structures, as this has the potential to reduce emissions, costs, and congestion [12].

The GDA research project is particularly focused on the use of electrically powered cargo bikes, which were "specially developed for the transportation of goods and/or people" [13] and can "deliver shipments of low weight and volume" [14] with the assistance of electric power at speeds of up to 25 km/h. These are primarily used in conjunction with micro-hubs, which are defined as "small, densely distributed areas" [15] in urban regions utilized for the handling of goods, temporary storage, and loading of smaller delivery vehicles. The end customer can be reached from a micro-hub within one kilometer. A fundamental distinction can be made between mobile and stationary micro-hubs [15].

The rationale behind the integration of cargo bikes and micro-depots lies in the potential to offset the comparatively limited transport capacity of the cargo bikes through the close proximity of the depots [14].

**Network planning in Last Mile Logistics.** Micro-hub structures are typically embedded within larger distribution networks, which are crucial components of logistics networks [16]. These logistics networks consist of nodes, edges, sources, and sinks. In the context of CEP, there are numerous sources and sinks, reflecting

[3] Zielgruppenorientierte Mobilitätsketten im Gesundheitswesen—ZMo-Gesund.

[4] Smart Industrial Parks (SIP).

the complex nature of these networks. Distribution networks, in particular, form the backbone of package delivery services. The planning of distribution networks is therefore relevant for the design of micro-hub structures and focuses on location planning, transport planning, and route planning [17].

Location planning has a significant impact on the functionality of logistics networks. The locations of warehouses, transshipment points, and means of transportation must be planned [18].

According to Mattfeld and Vahrenkamp, the models for optimal location decisions can be divided into discrete models, continuous models, and semi-discrete models. Furthermore, a distinction can be made between static or dynamic models or deterministic or stochastic models [18].

Transport planning, which encompasses the design of transport networks and control of transport processes, typically considers the pre-carriage, main leg, and onward carriage of the transport. The pre-carriage and onward carriage are local transport routes that are responsible for consolidating shipments and transferring them to transshipment points. In distribution networks, only the main leg and the onward leg are considered [19].

In essence, there is a distinction to be made between different terms. According to Fleischmann and Kopfer [19], the design of the transport network is a long-term task, the planning of transport routes and means is a medium-term task, and the planning of vehicle deployment is a medium-term and short-term task. The short-term tasks of vehicle deployment planning also include route planning. This process entails consolidating smaller transport orders and combining them into tours. The optimization of tours in this context pertains to the nodes and edge structure of the network. In practice, node-oriented planning, such as the Traveling Salesman Problem (TSP), is more prevalent than edge-oriented planning, such as the letter carrier problem. The standard objective is to minimize the distance to be traveled while maintaining vehicle capacities and fully meeting demand [19].

As Klose [17] notes, route planning and location planning are subject to interdependencies. Consequently, route planning approaches are approximately taken into account in location planning. Integrating approaches such as the Location Routing Problem (LRP) combines the solution of both problems and solves them using heuristic methods [20].

## *1.4 Vision Building*

Every software development project must engage in rigorous requirements engineering, as this is where the foundation for successful development is laid. This includes not only the collection of all requirements for the system but also the development of a (product) vision. The vision describes the goals of the product at a high level of abstraction. As outlined by Rupp [21], these are developed in a process that includes analyzing the environment for relevant information, documents, and systems that may be affected by the system development. In addition, an analysis of the current situation as well as a field of action analysis and identification of optimization potential should take place. The vision is to be derived from this. The analysis of customer benefits is particularly relevant for the development of new products [21].

The vision artifact of the RUP (Rational Unified Process) approach is employed to develop the software vision of the GDA project's analysis tool. RUP is a software development process framework developed by Rational Software Corporation, which was acquired by IBM in 2003 [22]. The RUP framework permits the development process to be adapted and the requisite artifacts to be selected, which are represented by diagrams in the Unified Modeling Language (UML) notation language [23].

In this paper, the "Artifact: Vision" is employed for the purpose of vision documentation, with particular emphasis on the "Vision Template (informal)," which delineates the structure of a vision development [24]. The RUP framework stipulates that the vision development process should include an introduction, a positioning (problem/product) section, a description of the relevant stakeholders, and a product overview [25].

For the implementation of the template in the GDA research project, a document analysis of the project description validated by all stakeholders was conducted. For this purpose, the document analysis software MAXQDA[5] was used to examine the relevant contents of the project description. Text passages were systematically coded and subcoded, which were then used in a quantitative analysis to generate insights for the requirements specification and vision. The results of this analysis represent the most significant source of information for the software vision and will not be reiterated for the sake of brevity. As is customary, other sources are duly referenced. The resulting software vision was validated through a review process with the stakeholders and adapted accordingly. For the documentation of the software vision, a system context diagram, architecture overview diagram, and a use case diagram were used. In future phases, the evaluation will expand to include external stakeholders such as local municipalities and independent logistics providers. These external evaluations will provide unbiased feedback on the system's efficacy, ensuring that it meets the needs of a broader audience and is adaptable to varying urban contexts.

[5] https://maxqda.com/, last accessed 2024/08/14.

# 2 Positioning

## 2.1 Problem Statement

The issue of managing and optimizing sustainable last mile logistics operations in urban areas involves a variety of stakeholders, including logistics companies, local authorities, economic stakeholders such as shippers and consumers, and the wider community. This issue is characterized by significant challenges due to increasing volumes and heterogeneity of shipments as well as increased traffic congestion in cities and the objective of reducing local emissions (such as nitrogen oxides, particulate matter, and noise) and greenhouse gas emissions while ensuring the cost-effectiveness and operational efficiency of cargo bikes and other light electric (battery or fuel cell) logistics vehicles. Although these vehicles are crucial for carbon–neutral logistics, they are currently not applicable in the same manner as conventional vehicles [26–28]. Their use requires a precise combination strategy based on comprehensive performance data to compensate for various limitations in terms of capacity and efficiency [26–28].

The decentralization of logistics processes due to the limited range and capacity of electric delivery systems further complicates the circumstances of the urban logistics chain over the last mile. This transformation presents new challenges for companies and municipalities who must navigate a complex web of logistical, infrastructural, and environmental considerations.

A successful solution [3, 19, 29–31] would involve the development and implementation of a robust, data-driven logistics system, that integrates modern information and communication technologies. Such a system should consist of new hardware such as cargo bikes and micro-hubs with adapted software and processes. It would use logistics, traffic, emissions, spatial structure, socio-demographic, and infrastructure data to uncover patterns and optimize network, transportation, and infrastructure planning. This approach would improve sustainability by reducing emissions, ensure profitability through improved operational efficiency, and support comprehensive urban planning. The interplay between operational logistics, IT, and policy is crucial here [32, 33].

### 2.2 Product Position Statement

For logistics companies, municipalities, and urban planners striving for sustainable and efficient last mile logistics, the software planned within the Green Delivery Analytics (GDA) project is a comprehensive and innovative solution that optimizes last mile logistics with data-driven insights on logistics, traffic, emissions, and infrastructure and specifically addresses the challenges of carbon–neutral logistics. Unlike traditional logistics management systems, which often lack a specific focus on sustainability and detailed integration of performance data [29, 34, 35], the planned product is designed to optimize the use of electric and hydrogen-powered delivery vehicles, taking into account their range and capacity constraints and ensuring profitability through advanced data analytics. It integrates extensive performance data with city, traffic, and emissions parameters to enable precise planning and execution of sustainable logistics operations. This enables companies and municipalities to achieve their environmental goals while maintaining operational efficiency and cost-effectiveness.

## 3 Stakeholders and User Environment

The stakeholders involved in the Green Delivery Analytics (GDA) project and the logistics planning tools to be developed include logistics companies that make last mile deliveries, municipal authorities responsible for urban planning, environmental authorities that monitor emissions, and society as a whole, which is concerned about sustainable urban development. Table 1 provides a more detailed breakdown of each stakeholder with descriptions. These stakeholders work in dynamic urban environments characterized by fluctuating delivery demands [3], evolving traffic patterns [36], and increasing environmental regulations [29]. The aim of the GDA project is to achieve efficiency gains and positive sustainability effects for these stakeholders through comprehensive software for the analysis, optimization, and implementation of sustainable last mile delivery solutions.

Table 2 illustrates the target users of the product, who are directly interacting with the software. These users are primarily logistics service providers, such as the project partner RytleX. This is due to the necessity of economically sustainable business models, which can be assured through an established and economically viable partner that implements the software. The target users of the planning tool are logistics planners who want to find the perfect combination of logistics system elements and network design to achieve efficiency and sustainability within last mile logistics operations. These logistics planners should be able to find these combinations taking into account their own logistics systems and the systems of external logistics companies. Logistics planners should also be able to advise urban planners on infrastructure planning, advising on which areas should be made possible for infrastructure in the land use plans and the setting of targets to achieve climate goals. The tool must

**Table 1** Stakeholders and their roles within the research project "Green Delivery Analytics"

| Name | Role |
|---|---|
| Project Sponsor mFUND | Funding for the GDA project within the framework of the mFUND innovation initiative by the German Federal Ministry for Digital and Transport |
| Project Manager RytleX | Project managers overseeing the project to ensure that practical and operational requirements are met from RytleX's perspective |
| Project Manager SMI Logistics | Project managers overseeing the project to ensure that logistical requirements are met |
| Project Manager SMI Software | Project managers overseeing the project to ensure that software requirements are met |
| Project Leaders | Senior project managers from both SMI and RytleX possessing top management skills |
| Research Associates SMI Software | Research associates specializing in software |
| Research Associates SMI Logistics | Research associates specializing in logistics |
| Software Architects and Developers | Development of the planned software by RytleX Group SE and COSMO UG[6] (Subcontract) |
| Technical Support Team | Team responsible for system maintenance |
| Transport Service Providers (CEP) | Organizations providing transport systems and delivery data. Recipients of last mile logistics (LML) planning results |
| Urban Traffic Managers | Organizations providing urban traffic data |
| Urban Planners | Organizations providing data on urban structure and requirements |
| Environmental Monitoring Organizations | Organizations providing emissions data |
| Consumers | Natural or legal persons utilizing delivery services |
| Shippers | Natural or legal persons aiming to meet the needs of consumers |

be able to be partially integrated into users' existing software systems and ensure data exchange with logistics and city management systems. The IT department and administrators only use the system for further developments and not for the logistics planning use case.

[6] https://cosmo-mobility.org/, last accessed 2024/08/14.

**Table 2** Responsibilities and roles of interacting users within the planning tool

| Name | Description | Role |
|---|---|---|
| Logistics Planner RytleX | Utilizes the tool for planning the LML network. Provides network planning results to traffic service providers for a fee and recommendations to urban planners. Exchanges emissions data with environmental monitoring organizations | Develops requirements, provides transport systems and delivery data, coordinates communication with cities and transport service providers, validates and tests requirements and prototype, develops business model, utilizes project results |
| IT Department RytleX | Manages IT infrastructure | Assists with software installation and deployment, ensures security, and maintains IT infrastructure |
| Administrator RytleX | Manages system settings | Manages access, configures settings, and monitors system status |

# 4 Product Overview

## *4.1 Product Perspective*

The planned logistics planning software is positioned as a central tool within the broader urban logistics and sustainability management ecosystem. This subsection explains how the software integrates with other related products and fits into the user's operating environment, detailing dependencies and interfaces.

The software can function as an independent, self-contained software solution, providing comprehensive data analytics and logistics optimization capabilities for last mile delivery operations.

The software interacts with multiple systems to exchange data and provide its functionalities effectively. These systems and data sources include:

- Urban traffic management systems to receive and analyze historical traffic flow information, congestion data, and incident reports
- Environmental monitoring systems to collect data on local pollutant levels and integrate this data into logistics planning
- Logistics management systems to synchronize route planning, delivery schedules, fleet utilization, and key performance indicators
- Urban planning systems to integrate infrastructure, other urban structures, and demographic data to adapt logistics processes to urban conditions
- User (logistics planners): Logistics planners use the system as described above. However, city planners are not direct users who interact with the logistics planning system but are represented in their interests by the logistics planners.

The following Fig. 1 provides an abstract overview of the system environment of the planned software using a context diagram.

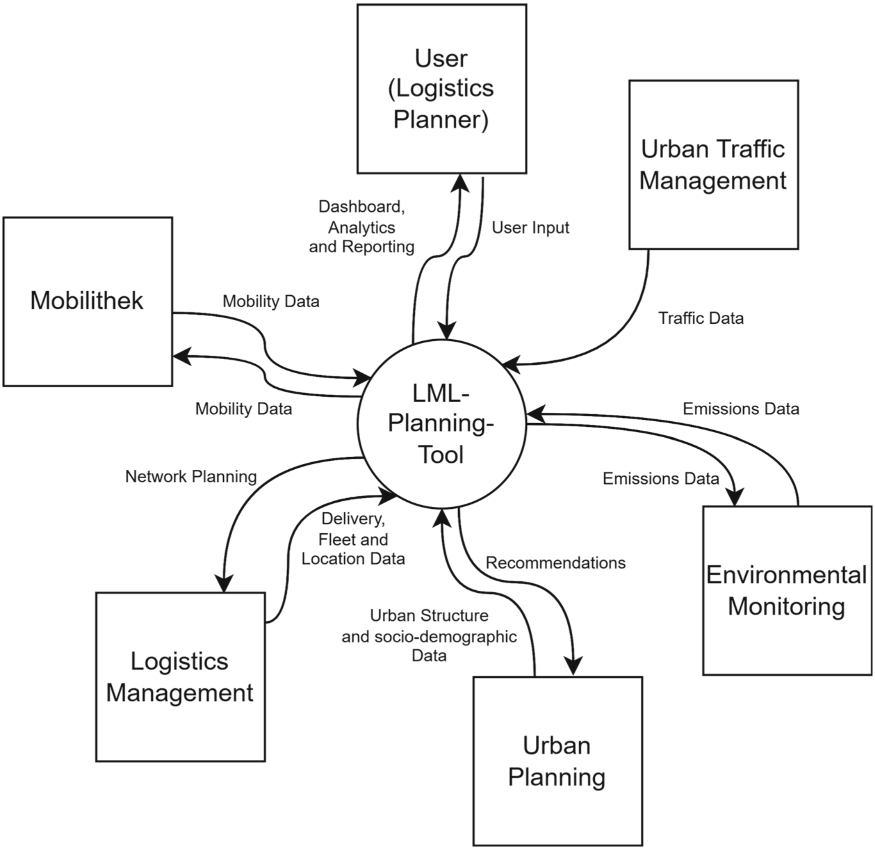



**Fig. 1** System-Context-Diagram highlighting the planned system's environment

The integration of different data sources and systems means that appropriate interfaces are required. Flexible data import/export APIs for the integration of traffic, environmental, and logistics data of different structures and formats, enabling the integration of data from different providers such as cities and logistics companies. The APIs support the export of optimized logistics plans and performance reports to external systems and the export as report documents. Dashboards and analysis tools enable intuitive user interactions with the planned software, complemented by customizable reporting tools to generate insights and recommendations for action. Security and compliance are ensured through secure interfaces, data confidentiality and integrity, and adherence to relevant data protection regulations and standards.

## 4.2 Assumptions and Dependencies

The successful implementation of the logistics planning software depends on factors such as data availability, stakeholder cooperation, and technological infrastructure. Assumptions about the compatibility of data sources, regulatory support for sustainable initiatives, and the scalability of the software can have an impact on project outcomes. The following assumptions and dependencies are particularly relevant:

- Availability of reliable data from logistics companies and urban planners.
- Integration with existing systems, databases, and data.
- Data from various sources will be of sufficient quality and cleanliness for analysis.
- Technological infrastructure and APIs will remain compatible with evolving systems.
- Logistics companies will adopt sustainable practices.
- External stakeholders will participate in evaluation and validation.
- Continued political and regulatory support for sustainable logistics initiatives.

## 4.3 Needs and Features

GDA aims to meet the needs of users by providing network optimization, data analytics, and sustainability tracking capabilities for decision support. These features enable users to assess the sustainability and efficiency of delivery operations, identify opportunities for improvement, and implement a specific logistics plan for zero-emission delivery to any city. The following Table 3 provides an overview of the needs and respective supporting functions:

The functions described above can be summarized and implemented using the following components:

- Optimization: Optimizing the placement and utilization of logistics locations (micro-hubs) and delivery areas for better distribution efficiency.
- Analysis: Advanced analysis of geographic demand distribution, traffic and demand forecasting, and performance tracking.
- Reporting: Compilation of analysis results and reports including recommended actions for urban planning.
- Sustainability Tracking: Tools for measuring and reporting carbon emissions and other sustainability metrics.
- User interface: Intuitive, user-friendly interface with data visualization and reporting capabilities.
- Data management: Manage geospatial, graph, time, transaction, analytics, and performance data and make it available for analysis and reporting.
- Integration capabilities: Flexible interfaces for integration into and from existing logistics and urban planning systems as well as data exchange.

**Table 3** Stakeholder needs and their supporting functions of the planned software

| Needs | Supporting Functions |
|---|---|
| Precise planning and execution of logistics operations [37] | Integration of extensive logistics performance data (logistics system and operations) with urban, traffic, and emissions parameters to facilitate detailed logistics planning |
| Demand and traffic forecasting | Forecasting features for both demand and traffic are driven by a combination of historical data, socio-demographic data, and machine learning |
| Location and route optimization [17, 20] | Optimization of micro-hub locations and delivery routes by integrating spatial structure, traffic, demand data, and logistics systems data and using heuristic algorithms such as Location Routing Problem (LRP) to minimize distances, reduce emissions, and enhance delivery efficiency |
| Greater sustainability in last mile logistics | Data-driven insights to optimize logistics, reduce emissions, and support carbon–neutral operations |
| Cost-effectiveness | Data analysis to ensure cost-effective logistics solutions that balance sustainability with economic considerations |
| Economic feasibility | A market-accepted software that considers the needs of shippers and consumers, thus ensuring economic viability |
| Comprehensive data integration [29, 33] | Access to a wide range of data sources, including urban planning, traffic patterns, and emissions statistics, for a holistic view of logistics operations |
| Customized logistics optimization strategies | Customizable settings to develop strategies based on specific operational requirements and constraints, ensuring tailored and effective logistics solutions |
| Informed decision-making [29, 33] | Advanced analysis and reporting tools to provide logistics companies, municipal authorities, and urban planners with actionable insights |
| Improved environmental impact | Tools to monitor and reduce local emissions (e.g., nitrogen oxides and particulate matter) and greenhouse gas emissions, aligning logistics operations with environmental goals |
| Analysis of historical data [19, 29] | Capability to analyze historical data to identify trends, improve future planning, and develop more efficient logistics strategies |
| Integration with existing systems | Compatibility with existing logistics management systems to ensure seamless integration and enhanced functionality |
| Scalability and flexibility | Scalable architecture to accommodate the growth of logistics operations and flexible enough to adapt to changing requirements and technologies |

The tool incorporates partial automation for logistics planning, relying on the intelligence derived from advanced analytics processes that can adapt to the specific needs of logistics service providers and city planners. By integrating their data as input to the model, the system enables automated location and route optimization and recommendations for city planning and sustainability reporting. The tool employs cloud-based data storage and analytics platforms (e.g., BigQuery) that support the integration of various types and formats of data via flexible APIs. Data cleansing methods, including outlier detection and imputation techniques, are applied to enhance the reliability of the input data and ensure data quality. The following Fig. 2 shows the various components at an abstract level within a three-tier software architecture.

The potential and roughly defined use cases that could be realized by the software for the different users are demonstrated in the following Fig. 3. In addition to the interfaces to the user, the interfaces to other systems or stakeholders are also shown, which could be realized manually by passing on data or connecting the systems. The diagram demonstrates how the logistics planner interacts with the planning software. The use cases for this user are planning the LML network, configuring settings, and monitoring KPIs. The logistics planning includes analyzing data to cluster and classify data and find patterns as a basis for the planning and optimization (Gutierrez-Franco, 2019) [37]. The process of optimizing the last mile logistics network regarding sustainability, demand, locations, fleet, and delivery areas is the

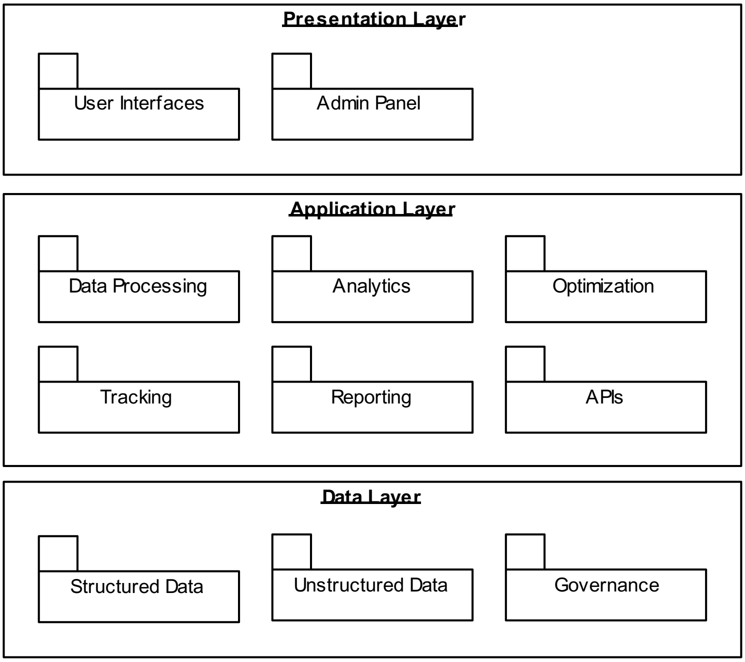



**Fig. 2** Components in a three-tier architecture that are needed to fulfill the needs and features of the planned tool

main use case of the software. Another use case for the logistics planner is reporting of planning results as well as the possibility of gathering recommendations for city planning purposes.

## 5 Summary and Future Outlook

The paper demonstrated the rationale for developing software to facilitate the analyses and planning of sustainable last mile logistics solutions. The planned tool of the Green Delivery Analytics research project is primarily intended to provide data-driven analyses and planning support, with a particular focus on enabling city hub structures for the use of electrically or hydrogen-powered delivery vehicles like cargo bikes.

The project integrates various datasets (e.g., spatial structure data, delivery data, traffic flow data, socio-demographic data) to design optimized combinations of micro-hubs, e-transporters, and cargo bikes for each city. An open data approach is utilized, allowing data to be accessible to the public and enabling cities and logistics companies to implement sustainable solutions that are both economically viable and environmentally beneficial.

A software vision based on the RUP process model was assessed to develop the tool in accordance with the specified requirements, thereby maximizing the positive impact on the advancement of sustainable last mile logistics. This includes a positioning statement, an analysis of the stakeholders and the user environment, as well as a product overview. The latter includes an analysis of interfaces to neighboring systems, a description of dependencies and assumptions of the project, and a detailed explanation of the needs and features of the product. The results of the vision were presented in this paper with the aid of a variety of diagrams, including a system context diagram, an architecture overview diagram, and a use case diagram.

The project will utilize the findings of the software vision to inform the development of the planned tool, thereby facilitating the implementation of sustainable last mile logistics. Nevertheless, the software vision can also be employed in conjunction with analogous projects within the domain of last mile logistics, where the implementation of data-driven software initiatives is a key objective.

As of the time of writing, no fully functional software tool has been developed in the Green Delivery Analytics project. The current work primarily involves the conceptual modeling and design phase, which lays the groundwork for future development efforts. In the upcoming phases of the project, the development of optimization and forecasting models will focus on refining algorithms such as clustering and regression models, which will be evaluated for their effectiveness in improving location and route planning as well as traffic and demand prediction. Additionally, various implementation technologies, including cloud platforms and machine learning frameworks, will be assessed to ensure scalability, performance, and compatibility with logistics and urban planning systems. Another project phase will focus on the actual implementation of the software, following the guidelines

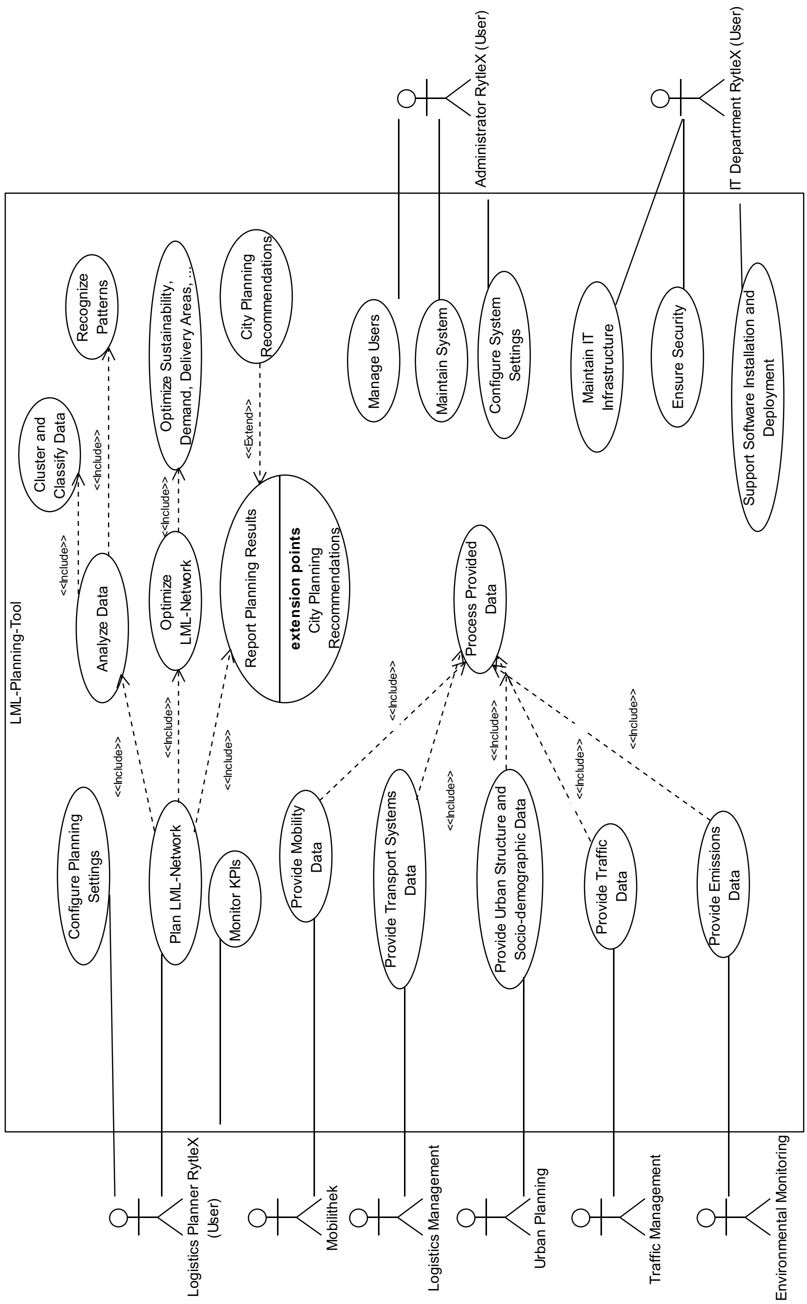



**Fig. 3** UML use case diagram to show the use cases the planned tool must cover

established in this vision document and subsequent requirements documentation. The value of this conceptual modeling exercise will be fully realized when the tool is developed and evaluated in real-world logistics environments.

The implementation of the vision will also be accompanied by future scientific studies, with a focus on the following aspects:

1. Which technologies are suitable for the implementation of such a tool?
2. Which optimization models are suitable for the planning tool?
3. How can business models be developed that are suitable within the three pillars of sustainability?
4. What is the quality and quantity of the available data, and how can deficiencies in the data set be addressed?
5. How do the shortcomings of data sources, including inconsistencies and low quality, affect the optimization methods?

**Acknowledgements** GDA is funded by the Bundesministerium für Digitales und Verkehr (BMDV, German Federal Ministry of Digital and Transport) as part of the mFUND program (project number 19F2262A) with a funding amount of around 2.1 Mio Euros. As part of mFUND the BMDV supports research development projects in the field of data-based and digital mobility innovations. Part of the project funding is the promotion of networking between the stakeholders in politics, business, administration, and research as well as the publication of open data on the Mobilithek portal.

## References

1. Wittowsky, D., Altenburg, S., Esser, K., Garde, J., Groth, S., Kurte, J.: KEP-Verkehre und die Letzte Meile: Anmerkungen zum potentiellen Umgang mit dem Boom der Lieferverkehre im Quartier. jmv. 1–12 (2020). https://doi.org/10.34647/jmv.nr5.id33
2. Gevaers, R., Van De Voorde, E., Vanelslander, T.: Characteristics and typology of last-mile logistics from an innovation perspective in an urban context. In: Macharis, C., Melo, S. (eds.) City Distribution and Urban Freight Transport. Edward Elgar Publishing (2011). https://doi.org/10.4337/9780857932754.00009
3. Bates, O., Friday, A., Allen, J., Cherrett, T., McLeod, F., Bektas, T., Nguyen, T., Piecyk, M., Piotrowska, M., Wise, S., Davies, N.: Transforming last-mile logistics: opportunities for more sustainable deliveries. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. pp. 1–14. ACM, Montreal QC Canada (2018). https://doi.org/10.1145/3173574.3174100
4. Giret, A., Carrascosa, C., Julian, V., Rebollo, M., Botti, V.: A crowdsourcing approach for sustainable last mile delivery. Sustainability. **10**, 4563 (2018). https://doi.org/10.3390/su10124563
5. Nadine Metzler, Hochschule Bremerhaven: Forschende der Hochschule Bremerhaven arbeiten an Software für klimafreundliche Warenzustellung, https://www.hs-bremerhaven.de/de/aktuelles/neuigkeiten-uebersicht/neuigkeiten-2024/green-delivery-analytics. Last accessed 14 August 2024
6. Hochschule Bremerhaven, RytleX: Anlage zum easy-Online Antragsformular Vorhabenbeschreibung - Verbundprojekt: Green Delivery Analytics - Datenbasierte Ausgestaltung einer CO2-neutralen Last-Mile-Logistik für die 80 größten deutschen Städte - GDA (2024, intern)

7. Bundesministerium für Digitales und Verkehr (BMDV): Datenbasierte Ausgestaltung einer CO2-neutralen Last-Mile-Logistik für die 80 größten deutschen Städte - GDA, https://bmdv.bund.de/SharedDocs/DE/Artikel/DG/mfund-projekte/green-delivery-analytics.html Last accessed 14 August 2024
8. Schnedlitz, P., Lienbacher, E., Waldegg-Lindl, B., Waldegg-Lindl, M.: Last Mile: Die letzten – und teuersten – Meter zum Kunden im B2C ECommerce. In: Crockford, G., Ritschel, F., and Schmieder, U.-M. (eds.) Handel in Theorie und Praxis. pp. 249–273. Springer Fachmedien Wiesbaden, Wiesbaden (2013). https://doi.org/10.1007/978-3-658-01986-0_10
9. Brabänder, C.: Die Letzte Meile: Definition, Prozess, Kostenrechnung und Gestaltungsfelder. Springer Fachmedien Wiesbaden, Wiesbaden (2020). https://doi.org/10.1007/978-3-658-29927-9
10. Klaus, P., Krieger, W., Krupp, M. eds: Gabler Lexikon Logistik. Gabler Verlag, Wiesbaden (2012). https://doi.org/10.1007/978-3-8349-7172-2
11. Zimmermann, F.M.: Was ist Nachhaltigkeit – eine Perspektivenfrage? In: Zimmermann, F.M. (ed.) Nachhaltigkeit wofür? pp. 1–24. Springer Berlin Heidelberg, Berlin, Heidelberg (2016)https://doi.org/10.1007/978-3-662-48191-2_1
12. Deloison, T., Hannon, E., Huber, A., Heid, B., Klink, C., Richa Sahay, Christoph Wolff: The Future of the Last-Mile Ecosystem. https://www3.weforum.org/docs/WEF_Future_of_the_last_mile_ecosystem.pdf (2020)
13. Deutsche Gesetzliche Unfallversicherung e.V. (DGUV): DGUV Information 208–055 „Sicher unterwegs mit dem Transport- und Lastenfahrrad". https://publikationen.dguv.de/widgets/pdf/download/article/3623 (2019). Last accessed 14 August 2024
14. Wellbrock, W., Ludin, D., Knezevic, I.: Letzte Meile 4.0: Potenziale innovativer Technologien für die Auslieferung im B2C-Bereich. Springer Fachmedien Wiesbaden, Wiesbaden (2022). https://doi.org/10.1007/978-3-658-37551-5
15. Lars Mauch, Fabio Weiss, Manuela Wohlhüter, Stephanie Wagner: Micro-Hubs für eine nachhaltige Citylogistik | Erfahrungen aus dem Pilotprojekt, https://publica-rest.fraunhofer.de/server/api/core/bitstreams/a2109866-283a-4662-ae74-18abcd4aecf7/content (2021). Last accessed 14 August 2024
16. Huang, Z., Huang, W., Guo, F.: Integrated sustainable planning of micro-hub network with mixed routing strategy. Comput. Ind. Eng. **149**, 106872 (2020). https://doi.org/10.1016/j.cie.2020.106872
17. Klose, A.: Standortplanung in Distributiven Systemen. Physica-Verlag HD, Heidelberg (2001). https://doi.org/10.1007/978-3-642-57582-2
18. Mattfeld, D., Vahrenkamp, R.: Logistiknetzwerke: Modelle für Standortwahl und Tourenplanung. Springer Fachmedien Wiesbaden, Wiesbaden (2014). https://doi.org/10.1007/978-3-8349-6912-5
19. Fleischmann, B., Kopfer, H.: Transport- und Tourenplanung. In: Tempelmeier, H. (ed.) Planung logistischer Systeme. pp. 71–98. Springer Berlin Heidelberg, Berlin, Heidelberg (2018). https://doi.org/10.1007/978-3-662-57782-0_3
20. Nadizadeh, A., Sahraeian, R., Zadeh, A.S., Homayouni, S.M.: Using greedy clustering method to solve capacitated location-routing problem. Afr. J. Bus. Manage. **5**, 8470–8477 (2011). https://doi.org/10.5897/AJBM11.424
21. Rupp, C.: SOPHIST-Gesellschaft für Innovatives Software-Engineering: Requirements-Engineering und -Management: das Handbuch für Anforderungen in jeder Situation. Hanser, München (2021)
22. Taft, D.K.: IBM Acquires Rational. https://www.eweek.com/pc-hardware/ibm-acquires-rational/. Last accessed 14 August 2024
23. Péraire, C., Edwards, M., Fernandes, A., Mancin, E., Carroll, K.: The IBM Rational Unified Process for System z. (2007).
24. Anwar, A.: Rational Unified Process. International Journal of Software Engineering. (2014).
25. IBM Redbooks: Appendix A. Catalog Manager Case Study—Inception Work Products: Catalog Manager—Vision [Template-Document] in The IBM Rational Unified Process for System z, https://www.redbooks.ibm.com/abstracts/www.redbooks.ibm.com/abstracts/sg247362.html (2007). Last accessed 14 August 2024

26. Bosona, T.: Urban freight last mile logistics—Challenges and opportunities to improve sustainability: A literature review. Sustainability. **12**, 8769 (2020). https://doi.org/10.3390/su12218769
27. Conway, A., Cheng, J., Kamga, C., Wan, D.: Cargo cycles for local delivery in New York City: Performance and impacts. Res. Transp. Bus. Manag. **24**, 90–100 (2017). https://doi.org/10.1016/j.rtbm.2017.07.001
28. Gruber, J.: Das E-Lastenrad als Alternative im städtischen Wirtschaftsverkehr. Determinanten der Nutzung eines „neuen alten" Fahrzeugkonzepts. https://doi.org/10.18452/22349 (2020)
29. Gutierrez-Franco, E., Mejia-Argueta, C., Rabelo, L.: Data-driven methodology to support long-lasting logistics and decision making for urban last-mile operations. Sustainability. **13**, 6230 (2021). https://doi.org/10.3390/su13116230
30. Ehmke, J.F.: Integration of Information and Optimization Models for Routing in City Logistics. Springer US, Boston, MA (2012). https://doi.org/10.1007/978-1-4614-3628-7
31. Taniguchi, E. ed: City logistics: network modelling and intelligent transport systems. Pergamon, Amsterdam; New York (2001)
32. Assmann, T., Bobeth, S., Fischer, E.: A conceptual framework for planning transhipment facilities for cargo bikes in last mile logistics. In: Nathanail, E.G. and Karakikes, I.D. (eds.) Data Analytics: Paving the Way to Sustainable Urban Mobility, pp. 575–582. Springer International Publishing, Cham (2019). https://doi.org/10.1007/978-3-030-02305-8_69
33. Cardenas, I., Borbon-Galvez, Y., Verlinden, T., Van De Voorde, E., Vanelslander, T., Dewulf, W.: City logistics, urban goods distribution and last mile delivery and collection. Competition and Regulation in Network Industries. **18**, 22–43 (2017). https://doi.org/10.1177/1783591717736505
34. Allaoui, H., Guo, Y., Sarkis, J.: Decision support for collaboration planning in sustainable supply chains. J. Clean. Prod. **229**, 761–774 (2019). https://doi.org/10.1016/j.jclepro.2019.04.367
35. Aldrighetti, R., Battini, D., Ivanov, D., Zennaro, I.: Costs of resilience and disruptions in supply chain network design models: A review and future research directions. Int. J. Prod. Econ. **235**, 108103 (2021). https://doi.org/10.1016/j.ijpe.2021.108103
36. Letnik, T., Farina, A., Mencinger, M., Lupi, M., Božičnik, S.: Dynamic management of loading bays for energy efficient urban freight deliveries. Energy **159**, 916–928 (2018). https://doi.org/10.1016/j.energy.2018.06.125
37. Gutierrez-Franco, E.: A Methodology For Data-Driven Decision-Making In Last- Mile Delivery Operations http://purl.fcla.edu/fcla/etd/CFE0007645 (2019)

# Climate and Ecosystems Modelling II

# Strategic Sensor Placement for the Identification of Disinfection Byproducts from Chlorinated Drinking Water: Case Study in the Water Distribution Network of Coimbra, Portugal

**Aristotelis Magklis and Andreas Kamilaris**

**Abstract** Providing clean drinking water to people worldwide is a challenging task due to the contaminants that tend to form in water distribution systems at any point in time, causing various health impacts to the consumers. An emerging contaminant with not well-known effects on human body is the family of disinfection byproducts (DBPs) occurring in chlorinated drinking water. Due to the complexity of their formation, it is very challenging to monitor and detect DBPs in a water network. By placing sensors throughout the distribution system, it becomes feasible to monitor certain environmental parameters that affect their formation and identify locations in the network which have high probabilities for the detection of DBPs. The problem is complicated due to the large number of possible DBPs and the constrained number of sensors available, thus a clever approach needs to be followed, especially in cases where prioritization of DBPs needs to be considered. In this work, we propose a methodology for detecting DBPs, considering a real-world case study of a calibrated water distribution network in Coimbra, Portugal. Different algorithms are investigated to satisfy different performance objectives, aiming to locate the best possible locations for sensor placement, toward maximizing possible detection and minimizing overall impact of DBPs.

**Keywords** Disinfection Byproducts · Sensor Placement · Algorithms · Water Distribution Network · Chlorinated Drinking Water

A. Magklis (✉) · A. Kamilaris
CYENS Center of Excellence, Nicosia, Cyprus
e-mail: a.magklis@cyens.org.cy

A. Kamilaris
University of Twente, Enschede, The Netherlands

 
V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_17

## 1 Introduction

Water quality is essential to civilian daily life. Detecting contaminants in water distribution networks before they reach into households is vital in order to reduce the number of people exposed to health risks due to polluted water. In order to detect these contaminants quickly, a common practice is to install sensors strategically inside a water distribution network (WDN), to monitor environmental parameters that can lead to the formation of these contaminants. However, due to very large WDNs and the complex nature of the problem, considering that numerous contaminants can be formed inside a WDN, the problem of detecting those contaminants using proper sensor placement is very difficult to solve.

The most common chemical disinfectant that is used for water treatment is chlorine. Disinfectant byproducts (DBPs) constitute an emerging category of contaminants, formed when chlorine is used for the disinfection of drinking water. When chlorine interacts with natural organic matter (NOM) [1], then it is likely for a variety of DBPs to be formed. More than 700 DBPs have been reported to date [2]. DBPs are generally grouped into three main categories: (i) aliphatic, (ii) alicyclic and (iii) aromatic [2]. Some examples of NOM include humic substances and fulvic acids, that can impact the color, taste and odor of drinking water. Furthermore, many NOMs are results of human activities, such as polycyclic and heterocyclic aromatic hydrocarbons which are formed due to industrial processes, agricultural runoff and wastewater discharge [3]. Apart from various types of NOMs, other parameters can affect the creation of DBPs, such as temperature, turbidity, pH, topology, temperature as well as the increasing human population, which leads to higher needs for purified drinking water [4]. Being able to detect these DBPs early would allow to minimize the human population exposed to contaminated drinking water. The minimization of the time of detection can lead to fast and drastic measures by water operations before the contaminated water reaches the consumer. The long-term impact of these DBPs is severe and can create significant health risks, such as genotoxicity and cytotoxicity [5]. Genotoxicity can damage the DNA, which can result in mutation and increased risk of cancer [6]. On the other hand, cytotoxicity can harm or kill cells, which lead to tissue damage and organ failure [7]. The precise effects of DBPs on human body are still under investigation.

Identifying DBPs formed in chlorinated drinking water using sensors is an issue that has not yet been well studied because of the complexity of the formation of DBPs, as well as the difficulty of obtaining an accurately calibrated model of a WDN. Moreover, as previously mentioned, having hundreds of DBP compounds makes it hard to prioritize certain DBPs over others without having historical data of monitored environmental parameters, to understand which DBPs may form at certain locations or conditions inside a water network. Another challenge is the fact that usually operators of WDNs have limited sensory equipment available, thus this equipment needs to be used as efficiently as possible. It is important to note the difference between the type of sensory equipment available: low-end sensors monitor only relevant environmental parameters, such as pH, temperature, conductivity and more. On the other hand, high-cost sensors can directly identify DBPs, such as electrochemical sensors [8], gas and liquid chromatography and mass spectrometry [9] and others. Being able to perform strategic sensor placement of both low-cost and high-cost sensors is a challenge in many WDN cases.

In this paper, a calibrated model of an existing WDN is utilized for addressing the problem of strategic placement of low-end sensors for the detection of DBPs formed in chlorinated drinking water. The contribution of this paper is the proposal of a novel methodology for addressing this problem, employing various techniques related to prioritization of DBPs, modeling of a WDN and use of algorithms for solving the optimization problem. The novelty of this paper lies in the proposed methodology to address a complex challenge which is still generally poorly investigated by the scientific community. By analyzing multiple scenarios using two different approaches, applied in real-world settings, attempting to combine multiple methodologies to define strategies for the identification of DBPs, in water distribution networks.

## 2 Related Work

Several studies have explored the impact of DBPs on human health as well as their formation, examining which environmental parameters are the most relevant. Most literature regarding chlorinated DBPs focuses on the categorization of the families of DBPs that can be identified in a WDN. In order to estimate the formation and fate of DBPs, predictive models have been developed and analyzed in scientific literature. These models can assist in the decisions that need to be taken by the drinking water industry [10]. The prediction of the concentration of DBPs in the water distribution networks is based on the effects of different water quality and operational parameters in controlling DBP formation, under different environmental conditions [11].

Literature regarding strategic sensor placement for the detection of DBPs formed in chlorinated drinking water is limited due to the complexity of the subject, as well as the difficulty to understand their formation. The optimal placement of water quality sensors has been extensively studied in [12] and [13], focusing on indirect detection of disinfection byproducts by means of monitoring relevant environmental parameters, instead of actual DBP concentration in the network.

Nevertheless, many studies have been conducted utilizing different algorithms for sensor placement in WDNs for detecting other contaminants such as microorganisms, inorganic and organic chemicals, etc. [14]. For example, strategic placement of sensors for the detection of Escherichia coli and organophosphates was investigated in [15] and [16], respectively. Some literature focuses on the intelligent sensor placement for the detection of leakages based on the changes found in the flow of the water in the network [17]. Moreover, multiple experiments were conducted using specific field sensors for the monitoring of environmental parameters that help in the formation of DBPs, such as chlorine [18], pH [19], water temperature [20] and more. The most important parameter that is hard to detect is NOM, however, advances have been made to steadily monitor and characterize it [21]. In some cases, the optimal sensor placement needs to take into consideration topological and connectivity features of the WDN [22]. The most common algorithms for addressing the problem of sensor placement include heuristic approaches, genetic algorithms [23], the NSGA-II algorithm [24] or mixed-integer programming techniques [25].

# 3 Methodology

The methodology for approaching the challenge of strategic sensor placement in WDNs for detection of DBPs formed in chlorinated drinking water is depicted in Fig. 1.

Foremost, the WDN under study needs to be modeled in a digital form. Then, the environmental parameters that affect the formation of DBPs need to be selected, and some or all of these parameters need to be monitored, to extract their values inside the WDN. Such data is valuable to assess the seasonal and spatial variation of the values of the relevant environmental parameters, which in turn affect the likelihood of forming DBPs inside the WDN. In the following sub-sections, the different steps of the general methodology proposed will be analyzed in more detail, such as prioritization of the families of disinfection byproducts depending on the available environmental parameters, algorithms that could be harnessed to solve the strategic placement problem and performance objectives tackled for the early detection of DBPs.

## *3.1 Step 1: Water Distribution Network Modeling*

Access to an accurately calibrated water distribution system is very important. Most research papers consider hypothetical networks and many assumptions [14, 24], and it is difficult to assess whether their proposed algorithms work in a real-world setting. Collecting available environmental data from water utility operators is vital, to understand the range of possible values, especially as the chlorine diffusion inside

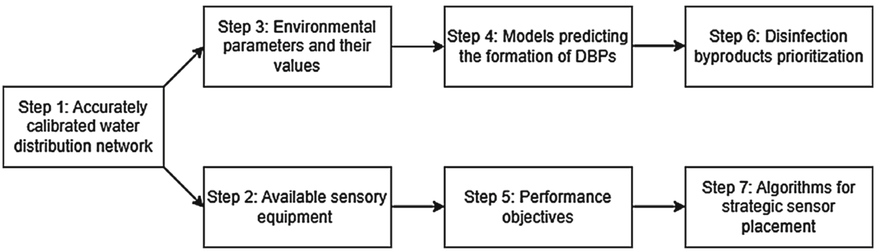



**Fig. 1** Methodology

the network and its concentration at each node/junction inside the pipe system (see also Step 3). For example, having higher doses of chlorine in the water network leads to the formation of trihalomethanes in a faster pace in the water distribution system [26]. If the samples contain smaller doses of chlorine, it is more likely that haloacetic acids and haloacetonitriles are present in the WDN [27].

### *3.2 Step 2: Sensory Equipment*

Afterward, it is important to consider the resources available for sensory equipment (number and type of sensors available: low-cost vs. high cost), examine the size and type of the network, the locations where sensory equipment can be placed inside the network, etc. There are many sensors that can assist in the monitoring of the environmental parameters, as well as analyzers that directly measure the DBPs found in the drinking water. However, depending on the nodes that must be covered, a decision needs to be made depending on the number of sensors that must be placed throughout the network and the cost of each sensor. Sensory equipment can be either low-end or high-end sensors: low-end sensors monitor only relevant environmental parameters (e.g. chlorine, pH, temperature, conductivity) while high-cost sensors can directly detect DBPs.

### *3.3 Step 3: Environmental Parameters and Their Values*

There are multiple environmental parameters that assist in the formation of DBPs. These parameters include: (i) bromide, (ii) dose of disinfectant, (iii) water temperature, (iv) potential of hydrogen (pH), (v) composition and concentrations of NOM and (vi) chlorine. As previously mentioned, the most common treatment for disinfecting water is by using disinfectants like chlorine. Depending on the amount of chlorine concentrated in the network, different families of DBPs are formed. Governing organizations have set safety standards for the acceptable amount of disinfectant dosage, to balance the amount of disinfection needed with the potential of harmful DBPs being formed. Regarding pH, a neutral range is vital (6.5 to 8.5), to minimize the formation of DBPs. Higher ranges (9.0–10.0) reduce the reaction of the chlorine in the system, while lower ranges (5.0–6.0) help in the formation of haloacetic acids [28]. The water temperature varies in values significantly in WDNs. Generally, higher temperatures in the water distribution system can increase the reaction rate between the disinfectant and the natural organic matter, leading to an increase in the formation of disinfection byproducts [29]. Lastly, turbidity is a measure of level of particles, such as sediment, or organic byproducts in the water [30]. The compliance for turbidity standards is below 1 Nephelometric Turbidity unit (NTU), in order for water utility operators to provide safe drinking water.

### 3.4 Step 4: Models Predicting Formation of DBPs

Probabilistic models use certain criteria to predict, for example, chlorine concentrations in the network, or DBP concentrations depending on the parameters available for input. For the development of these models, machine learning (ML) methods are used, such as back propagation artificial neural networks (BPANNs) [31], Random Forests (RFs) [32], support vector machines (SVMs) [33] and many more [34]. Prediction models can also be utilized for identifying concentrations of DBPs in the network. Depending on the DBP family under focus, different parameters that assist in the formation of relevant DBP compounds are used as inputs, such as pH, temperature, dissolved organic carbon and more.

### 3.5 Step 5: Performance Objectives

Performance metrics need to be selected wisely, and they directly relate to the decisions about sensor placement. Based on literature, there are two approaches to handle the strategic sensor placement: either with a single performance or multi-performance objective approach. The aim can be: (a) to maximize the coverage of the network based on the selected number of sensors, (b) to minimize the time of detection of the contaminants and (c) to reduce the infected water consumed by the consumers. This leads to a multi-performance objective approach, with the main focus being the minimization of time of detection, in order to effectively react fact to any potential dangers that might occur in the network. Most of the metrics mentioned benefit from one another, since maximizing the coverage allows the minimization of time of detection at the same time.

### 3.6 Step 6: Disinfection Byproducts Prioritization

To effectively place sensors for disinfection byproducts, it is of outmost importance to highlight which families of DBPs are targeted by the sensors. For example, if the WDN has higher dosages of chlorine, then trihalomethanes are the DBPs targeted for monitoring. Another factor to consider when prioritizing DBPs is the longevity of each contaminant in the network. For example, trihalomethanes can survive longer in the pipes of the network, while, on the other hand, haloacetonitriles have very low longevity [35]. Considering the longevity of the DBPs, compounds that have higher lifetime should be prioritized. Lastly, data for certain families of DBPs is available, depending on the region where the water distribution is located, and the legal restrictions that have been applied for maintaining an average value of chlorine concentration in the network. Moreover, each family of disinfection byproducts has different long-lasting effects on the human health, making it another factor to consider when prioritizing compounds for sensor placement [2]. Lastly, assuming it is possible to have data for each node at specific times throughout the day for the network, we categorize the possible families of DBPs that can be formed in the network (trihalomethanes, haloacetic acids, haloacetonitriles) and which precursors are dominant (e.g. total organic carbon, dissolved organic carbon, higher temperature of water). Kalita et al. [2] attempted a prioritization of DBPs based on a range of relevant factors.

### *3.7 Step 7: Algorithms for Strategic Sensor Placement*

After the identification of the performance metrics that need to be tackled, the selection of the algorithm which will solve the strategic placement problem needs to be decided. As previously mentioned, the most common algorithms for addressing the problem of sensor placement include heuristic approaches, genetic algorithms [23], the NSGA-II algorithm [24] or mixed-integer programming techniques [25]. There can be two approaches when examining the issue of strategic sensor placement. The empirical approach and the optimization-based approach. The empirical approach does not consider the hydraulic data of a network, instead it focuses on certain criteria that are vital for the placement of the sensors. For example, nodes with high numbers of consumers nearby, points in the network that can alter the quality of water instantly (source nodes) and accessible nodes where installing the equipment is feasible. On the other hand, the optimization-based approach focuses on the implemented performance objectives, such as time of detection, population exposed to contamination, maximum coverage, thus aiming to detect the most optimal locations for sensor deployment. This approach requires a calibrated model, to examine the network thoroughly using the hydraulic data provided and run different simulations and what-if scenarios to discover the best possible algorithm. Moreover, for smaller networks mixed-integer programming techniques are more effective due to their higher precision while, for larger networks, heuristic approaches can be used due to their faster execution, with small penalties in accuracy.

## 4 Implementation

We have implemented the methodology described in Sect. 3 using an accurately calibrated model of a WDN, at the civil parish Ameal e Arzila in Coimbra, Portugal (see Fig. 1).

The WDN provided by the utility operators is of moderate size, includes 227 nodes, a tank and a reservoir and is categorized as a dead-end network. Portugal is characterized by higher temperatures during the summer season, and relatively low temperatures during the winter season. Higher doses of chlorine are required to be added to the network when the temperatures are rising, so the formation of trihalomethanes in the network is favored. The environmental parameters that are monitored in the WDN under study are: the dose of disinfectant, water temperature, pH and turbidity. In our case study, the pH is in the neutral range between 6.5 and 8.5. Lastly, the turbidity value is steadily below 1.0, meaning that the water is purified, thus the levels of turbidity do not assist in the formation of harmful contaminants in the drinking water [30]. These ranges of values have been provided to us by the local operator.

Using the software tool EPANET [36], it was possible to examine the chlorine concentration over time at each node. One of the first tests for simulating DBPs in the network was by applying the 2R bulk-reaction model using the extension of EPANET, EPANET-MSX. The environmental parameters considered are: chlorine dosage added to the network through the tank and the reservoir, as well as water temperature. Without having ground-truth data, it was impossible to include more environmental parameters and characterize their values. The DBP families of trihalomethanes [37] and haloacetic acids [38] were considered in this simulation, due to the lack of available data for DBPs' historical occurrence from the water utility operators. The EPANET-MSX model, which simulates the concentration of DBPs in the network, has two reactants, one fast and one slow after the initial dosing of chlorine [39]. This simulation creates an output file which contains the ID of each node and the timestep at each step of the simulation (e.g. report of chlorine concentration and DBP concentration every 15 min).

The two algorithms utilized in this case study are: mixed-integer programming (MIP) and the greedy randomized adaptive search procedures (GRASP). For the MIP algorithm, a Python model has been developed based on two Python frameworks: (i) Chama [40], which specializes in the placement of the sensors and (ii) WNTR [41], a framework specialized for simulating and analyzing resilience of water distribution networks. Moreover, for the GRASP algorithm, the Threat Ensemble Vulnerability Assessment and Sensor Placement Optimization Tool (TEVA-SPOT) [42] was utilized, that allows the simulation of many different scenarios for any contaminant the user wants to create, strategic sensor placement, planning utility response to sensor detection and the design of the basis threat.

The performance metrics of maximum coverage and time of detection were selected, to analyze the trade-offs between these performance metrics, due to their correlation. Having a larger number of sensors in the network automatically decreases the time of detection, due to the coverage. The minimization of mass consumption of polluted water is also another performance metric considered. The prioritization of the DBPs was based on the region of the WDN and the data available. The region has a history of formation of trihalomethanes in most networks in Portugal, this is why trihalomethanes are better regulated than other families of DBPs. Then, the effects of each compound on the consumers as well as its longevity in the network were considered. Thus, the selected family we focused on was trihalomethanes, with a focus on the following compounds: chloroform (CF), Dibromochloromethane (DBCM) and their brominated counterparts. Our prioritization strategy is aligned with the prioritization exercise performed in [2].

Another important aspect of strategic placement is the type of sensors deployed, as previously mentioned. A range of sensors was considered during experimentation, assuming various scenarios of available resources and budget for sensory equipment by the water utility operator.

The equations selected for the DBP predictive models developed are listed in Table 1. The reasoning behind the selection of these equations is their high accuracy and the wealth of environmental parameters involved. The selection of these equations is based on an extensive literature review we performed. The equations selected are based on the family of DBPs and not on individual compounds (e.g. chloroform), since the simulations of the strategic placement focus on the selection of the major DBP families.

Through the simulation, the names of the nodes/junctions that have been selected for sensor placement are outputted to an excel file, in order to compare the results of the two algorithms.

Regarding the MIP algorithm, for each node, up to one sensor is assigned, with respect to the number of sensors available. Since this algorithm has a binary nature, each option is either chosen (1) or not (0). Based on the defined constraints, this algorithm aims to provide an optimal or near optimal solution regarding the strategic sensor placement on a provided WDN. To satisfy the constraints, the following formula was implemented:

$$\sum\nolimits_{a\in A}\alpha_a \sum\nolimits_{i\in L_a} d_{ai}x_{ai} \tag{1}$$

$$\sum\nolimits_{i\in L_a} x_{ai} = 1\forall a \in A \tag{2}$$

$$x_{ai} \le s_i \forall a \in A, i \in L_a \tag{3}$$

**Table 1** Equations for the detection of disinfection byproducts

| DBP | Example equation | $R^2$ | Reference |
|---|---|---|---|
| THMs | THM = $10 - ^{0.038} \times (Cl2)^{0.654} \times (pH)^{1.322} \times (time)^{0.174} \times (SUVA)^{0.712}$ | 0.88 | Uyak et al. [43] |
| HAA9 | HAA9 (μg / L) = $-345 + 1.695(\text{Temperature}) + 93.1(\text{pH}) - 226(\text{UVA254}) + 4.95(\text{Cl2}) + 5.66(\text{NO} - 2 - \text{N}) + 16.6(\text{DOC}) + 0.325(\text{NH} + 4 - \text{N}) - 0.0693(\text{Temperature})^2 - 6.41(\text{pH})^2 + 190{,}821(\text{UVA254})^2 - 1.73(\text{NO} - 2 - \text{N})^2 - 3.77(\text{DOC})^2 - 0.01663(\text{NH} + 4 - \text{N})^2$ | 0.811 | Okoji et al. [44] |
| HANs | T-HANs = $10^{-1.065} (\text{Br})^{0.346}(\text{DOC})^{0.369} \times (\text{Cl2/DOC})^{0.520} (t)^{0.238} (\text{Temp})^{0.373} \times (R^2 = 0.943, p < 0.0005, n = 36)$ | 0.943 | Hong et al. [45] |

$$\sum_{i \in L} c_i s_i \leq p \tag{4}$$

$$s_i \in \{0, 1\} \forall i \in L \tag{5}$$

$$0 \leq x_{ai} \leq 1 \forall a \in A, i \in L_a \tag{6}$$

The formulation is based on the p-median facility location problem and the goal is to minimize objective (1) subject to constraints (2 to 6), where $A$ is the set of all scenarios, $L$ is the set of all candidate locations, $L_a$ is the set of all scenarios capable of detecting scenario $\alpha$, $\alpha_a$ is the probability of the scenario $\alpha$ happening, $d_{ai}$ is the impact assessment, $x_{ai}$ is an indicator variable that will be 1 if sensor $i$ is installed, $s_i$ is a binary variable that will either be 1 if sensor $i$ is installed or 0 otherwise, $c_i$ is the cost of the sensor $i$ and lastly, $p$ is the sensor budget, meaning the number of sensors available for deployment.

Regarding the GRASP algorithm, the greedy randomized adaptive search procedure is used to solve optimization problems. It combines elements of greedy algorithms and randomization and through iterations it aims to identify the best possible solutions. A candidate list is constructed with all the best possible solutions (in our case, the selection of nodes depending on the performance objective defined) and the ranking of the list is based on the parameters defined (time, mass consumed, etc.). The selection of the solution however is randomized, to avoid selecting only local optima solutions (e.g. multiple nodes near the injection location). This process is repeated with respect to the provided budget (number of sensors) and each iteration starts with a new randomized solution.

## 5 Evaluation

### *5.1 Maximum Coverage*

For the performance objective of maximum coverage, both the MIP and the GRASP algorithm were used, to identify the optimal locations for the potential detection of DBPs. The Pareto front in Fig. 2 showcases the results of the maximum coverage objective, including the minimization of time of detection at the same time. After a certain percentage of the network has been covered (e.g. 30 to 35%) by sensors, the effectiveness of further increasing the number of sensors is very low. The same procedure was conducted using the GRASP algorithm to determine any major changes in the locations of the sensors between the two methods. Only eight nodes selected were different in the two approaches, without a big distance between the locations selected, as the next node in the network was selected instead. Thus, the two algorithms had similar behavior and results.

**Fig. 2** Water distribution network of civil parish Ameal e Arzila in Coimbra, Portugal

## *5.2 Minimization of Time of Detection and Mass Consumption*

A more realistic number of 10 sensors, assuming lower-cost sensors and considering the size of the network, was used for examining the minimization of time of detection and reduction of mass consumption. For the prioritization of the disinfection byproducts family, the diameter size of the pipes was used as a parameter for the sensor placement, to consider the spatial variation for the formation of haloacetic acids and haloacetonitriles [46]. The results are depicted in Fig. 4, with the left image showing the sensor placement prioritizing nodes with higher concentrations of chlorine, while the right image includes the diameter size of the pipes as a parameter for the placement.

For the selection of nodes where spatial variation is prioritized, locations closer to the dead-end nodes are preferred. These nodes have a lower concentration of chlorine, which as previously mentioned, benefits in the formation of haloacetic acids and haloacetonitriles.

### 5.3 Randomized Injections in the Network

To further examine optimal sensor placement, different simulations were conducted based on the possibility of injecting chlorine in different locations inside the WDN. For example, injections of chlorine were performed at the south side of the network and at the east side of the network at the same time. Figure 5 shows the differences in placement depending on the performance objective, the left image focusing on the time of detection and the right image focusing on the minimization of mass consumption.

Another scenario included a set number of sensors (10) and a completely randomized injection in the network, in order to compare the results of the two algorithms regarding the optimal solutions they provide for minimization of time of detection. In Fig. 6, the left image showcases the selection of 10 sensors utilizing the MIP algorithm and the right image shows the results based on the GRASP heuristic approach.

## 6 Discussion

The case study at Coimbra and the methodology followed helped to understand the number of sensors that should be placed inside the network, the best candidate locations, as well as the trade-offs between budget/cost (i.e. purchase of sensory equipment) and efficiency, based on the performance metrics used. Based on the results showcased, the optimal number of sensors for the performance metric of maximum coverage would be 80 sensors, to cover 35% of the network and 60 sensors to minimize effectively the time of detection of the contaminants, as shown in Fig. 3. Regarding these two objectives, the related work in [24] shows similar results following this optimization-based approach for sensor placement.

Due to the fact that the network provided is of moderate size, running the simulations alone is enough to provide high-confident results about chlorine concentrations in different nodes of the network, i.e. the chlorine diffusion algorithm works well.

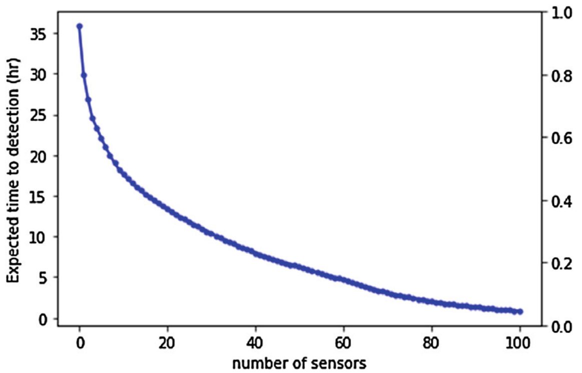



**Fig. 3** Pareto front for maximum coverage of the network and minimization of time of detection

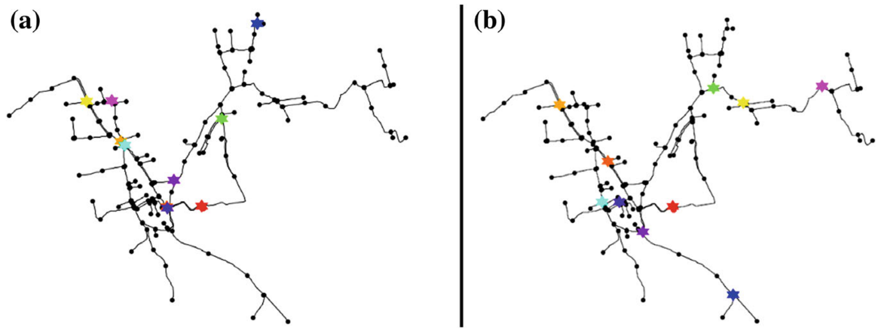



**Fig. 4** (a) Sensor placement of 10 sensors targeting trihalomethanes, (b) Sensor placement of 10 sensors focusing on haloacetonitriles and haloacetic acids. Injection of chlorine begins from tank and reservoir

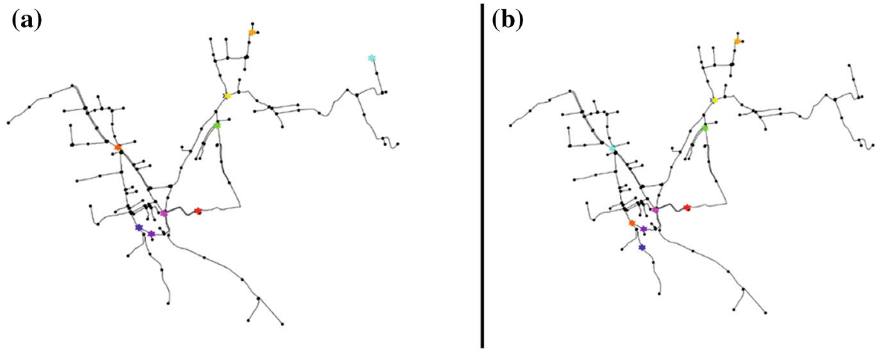



**Fig. 5** (a) Minimization of time of detection, (b) Minimization of mass consumption. Both scenarios are based on 10 sensors in the network

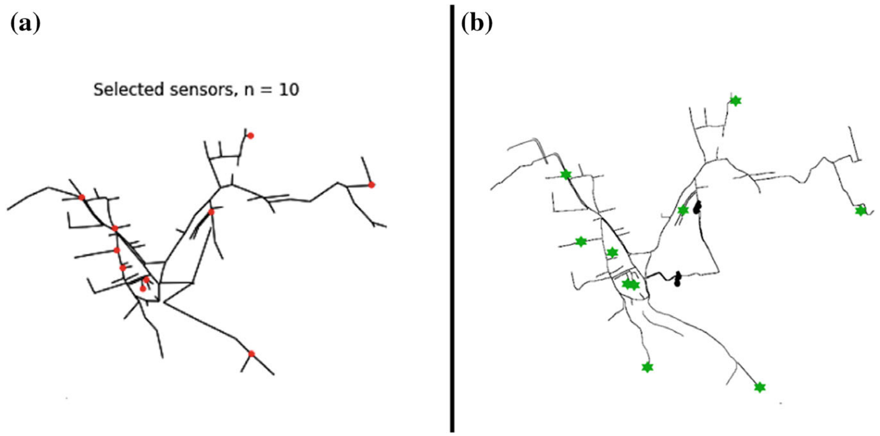



**Fig. 6** (a) Minimization of time of detection (MIP algorithm), (b) Minimization of time of detection (GRASP algorithm). Both algorithms consider the availability of 10 sensors

Additionally, having data related to the population of the network was very important in order to tackle the performance metric of minimizing the mass consumption effectively.

### *6.1 Limitations*

In order to achieve better results in such a complex subject, more data needs to be gathered from the water utility operators at a daily basis. The environmental parameters that relate to the formation of DBPs need to be monitored consistently, so methods such as machine learning can also be implemented to have a functioning model to predict more precisely the family that can be formed at each node of the network.

Furthermore, the installation of these sensors in the nodes in the system is sometimes a challenging task. The equipment needs to be located somewhere where it can be easily maintained and be safe from any dangers and to also have access to power and electricity. These restrictions make it almost impossible to place sensors at every node in the system. Most of the limitations relate to the fact that data is not available regarding the environmental parameters that can assist in the development of predictive models as well as create better simulations and what-if scenarios. Moreover, the use of only one calibrated water distribution network limits the possibilities of experimenting with different scenarios in real-world settings.

### *6.2 Future Work*

As future work, constant monitoring of the environmental parameters that assist in the formation of DBPs will be performed inside the pilot site, to extract useful data that will help to develop better prediction models. These data and models will then allow algorithms and simulations, such as the ones used in this paper, to achieve more realistic and optimized sensor placement. Regarding the prioritization of the DBPs in the network for the placement of the sensors, more research needs to be conducted regarding the longevity, the health risks associated at each DBP family or even at compound level. Having such data easily accessible can assist in the field of strategic sensor placement with diverse conditions in different water networks. Prioritization of DBPs needs to be embedded inside the constraint formulas when considering integer programming approaches, such as the ones we employ. The same goes for differentiating between low-end and high-end sensor devices, which is also an important aspect of future work.

Moreover, we plan to deploy both low- and high-end sensor devices inside the actual WDN of Coimbra, Portugal, using the approach and algorithms developed in this paper as guidance for strategic placement. This is part of the H20FORALL project [47], and the pilot will run for 12 months.

Finally, we intend to apply our methodology in other pilot sites as well, in the near future.

## 7 Conclusion

In this work, we proposed a methodology for strategic placement of sensors for disinfection byproducts formed from chlorinated drinking water, using a real-world water distribution network in Coimbra, Portugal as a case study. The results of our simulations varied depending on the family of disinfection byproducts prioritized, as well as on the number of sensory equipment available for placement. Different algorithms were utilized to compare different scenarios based on different performance objectives. Our results showed minimal changes to the placement problem when trying different algorithms or performance metrics. We have highlighted some of the challenges involved, such as the lack of data publicly available from actual water distribution networks, which could assist in the development of predictive models, as well as the shortage of accurately calibrated models. By having public data available, for both the water distribution networks and the impact of the disinfection byproducts on human health, it would be feasible to simulate more realistic scenarios in existing, water distribution networks, allowing higher accuracy in the placement based on the ground-truth data provided. This would facilitate better policy- and decision-making in the future, tackling successfully this emerging problem of potential contamination.

**Acknowledgements** Funded by the European Union, under the Grant Agreement GA101081953 attributed to the project H2OforAll Innovative Integrated Tools and Technologies to Protect and Treat Drinking Water from DBPs. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

## References

1. Plewa, M.J., et al.: Chemical and biological characterization of newly discovered iodoacid drinking water disinfection byproducts. Environ. Sci. Technol. **38**(18), 4713–4722 (2004)
2. Kalita, I., et al.: Assessing the health impact of disinfection byproducts in drinking water. ACS Es&t Water **4**, 1564–1578 (2024)
3. Tak, S., Vellanki, B.P.: Natural organic matter as precursor to disinfection byproducts and its removal using conventional and advanced processes: state of the art review. J. Water Health **16**(5), 681–703 (2018)
4. Moreno-Andrés, J., et al.: Operational and environmental factors affecting disinfection byproducts formation in ballast water treatment systems. Chemosphere **232**, 496–505 (2019)
5. Wagner, E.D., Plewa, M.J.: CHO cell cytotoxicity and genotoxicity analyses of disinfection by-products: An updated review. J Environ Sci (China). **58**, 64–76 (2017)
6. Lan, J., Lan, J., Rahman, S.M., Gou, N., Jiang, T., Plewa, M.J., Alshawabkeh, A.N., Gu, A.Z., Gu, A.Z.: Genotoxicity assessment of drinking water disinfection byproducts by DNA damage and repair pathway profiling analysis. Environ. Sci. Technol. **52**, 6565–6575 (2018)

7. Cortés, C., Marcos, R.: Genotoxicity of disinfection byproducts and disinfected waters: A review of recent literature. Mutat. Res-genet. Toxicol. Environ. Mutagen. **831**, 1–12 (2018)
8. Zhang, W., et al.: Recent advances on electrochemical sensors for the detection of organic disinfection byproducts in water. ACS Sensors. **4**(5), 1138–1150 (2019)
9. Yang, M., et al.: Current methods for analyzing drinking water disinfection byproducts. Current Opinion in Environmental Science & Health. **7**, 98–107 (2019)
10. Sadiq, R., Rodriguez, M.J.: Disinfection by-products (DBPs) in drinking water and predictive models for their occurrence: a review. Sci. Total. Environ. **321**, 1 (2004)
11. Chowdhury, S. et al.: Models for predicting disinfection byproduct (DBP) formation in drinking waters: a chronological review. Sci. Total Environ. **407**(14), 4189–4206 (2009)
12. Łangowski, R., et al.: An optimised placement of the hard quality sensors for a robust monitoring of the chlorine concentration in drinking water distribution systems. J. Process. Control. **68**, 52–63 (2018)
13. Waeytens, J., et al.: Adjoint-based numerical method using standard engineering software for the optimal placement of chlorine sensors in drinking water networks. Environ Model Softw. **92**, 229–238 (2017)
14. Rathi, S., Gupta, R.: Sensor placement methods for contamination detection in water distribution networks: A review. Procedia Engineering 89, 181–188 (2014)
15. Weickgenannt, M., et al.: Risk-based sensor placement for contaminant detection in water distribution systems. J. Water Resour. Plan. Manag. **136**(6), 629–636 (2010)
16. Ohar, Z., et al.: Optimal sensor placement for detecting organophosphate intrusions into water distribution systems. Water Res. **73**, 193–203 (2015)
17. Sophocleous, S. et al.: A Two-stage calibration for detection of leakage hotspots in a real water distribution network. Procedia Engineering **186**, 168–176 (2017)
18. Wilson, R., et al.: Continuous chlorine detection in drinking water and a review of new detection methods. Johns. Matthey Technol. Rev. **63**(2), 103–118 (2019)
19. Maineult, A. et al.: Detection of advected concentration and pH fronts from self-potential measurements. J. Geophys. Res. **110** (2005)
20. Qin, Y., et al.: Integrated water quality monitoring system with pH, free chlorine, and temperature sensors. Sensors and Actuators B-chemical. **255**, 781–790 (2018)
21. Chen, W., Yu, H.-Q.: Advances in the characterization and monitoring of natural organic matter using spectroscopic approaches. Water Res. **190**, 116759 (2021)
22. Perelman, L., Ostfeld, A.: Short communication: Topological clustering for water distribution systems analysis. Environ Model Softw. **26**(7), 969–972 (2011)
23. Casillas, M.V. et al.: Optimal sensor placement for leak location in water distribution networks using genetic algorithms. Sensors **13**, 11 (2013)
24. Diao, K. et al.: Sensor placement in water distribution networks using centrality-guided multi-objective optimization. J. Hydroinfo. **25**(6), 2291–2303 (2023)
25. Berry, J.W. et al.: Sensor placement in municipal water networks with temporal integer programming models. J. Water Resour. Plan. Manag. 132(4), 218–224 (2006)
26. Valdivia-Garcia, M. et al.: Climatic, geographic and operational determinants of trihalomethanes (THMs) in drinking water systems. Scientific Reports **6**(1), 35027 (2016)
27. Rodriguez, M.J. et al.: Formation and fate of haloacetic acids (HAAs) within the water treatment plant. Water Research **41**(18), 4222–4232 (2007)
28. Hung, Y.C., et al.: PH effect on the formation of THM and HAA disinfection byproducts and potential control strategies for food processing. J. Integr. Agric. **16**(12), 2914–2923 (2017)
29. Xue, R., et al.: Evaluation of thirteen haloacetic acids and ten trihalomethanes formation by peracetic acid and chlorine drinking water disinfection. Chemosphere **189**, 349–356 (2017)
30. Liu, Z., et al.: Enhanced coagulation of low-turbidity micro-polluted surface water: Properties and optimization. J. Environ. Manage. **233**, 739–747 (2019)
31. Pérez-Beltrán, C.H. et al.: Artificial intelligence and water quality: From drinking water to wastewater. TrAC Trends in Anal. Chem. **172** (2024)
32. Shin, H. et al.: Development of water quality prediction model for water treatment plant using artificial intelligence algorithms. Environ. Eng. Res. **29**(2), 117597 (2023)

33. Wang, D., et al.: Model predictive control for chlorine dosing of drinking water treatment based on support vector machine model. Desalin. Water Treat. **173**, 133–141 (2020)
34. Li, L., et al.: Recent advances in artificial intelligence and machine learning for nonlinear relationship analysis and process control in drinking water treatment: A review. Chem. Eng. J. **405**, 126673 (2021)
35. Pavelic, P. et al.: Fate of disinfection by-products in groundwater during aquifer storage and recovery with reclaimed water. J Contam Hydrol. **77**(4), 351–73 (2005)
36. Farina, G., et al.: Using EPANET for modelling water distribution systems with users along the pipes. Civ. Eng. Environ. Syst. **31**(1), 36–50 (2014)
37. Sathasivan, A., et al.: Trihalomethane species model for drinking water supply systems. Water Res. **184**, 116189 (2020)
38. Premarathna, S.M. et al.: Model for halo-acetic acids formation in bulk water of water supply systems. Sci. Total Environ. **857** (2022).
39. Fisher, I.H.: Integrated EPANET-MSX process models of chlorine and its by-products in drinking water distribution systems. Water Environ. Res. **95**, 12, e10949 (2023).
40. Overview—Chama 0.2.0 documentation. https://chama.readthedocs.io/en/latest/overview.html. Last accessed 29 August 2024
41. Haxton, T.: Case Study Applications of Water Network Tool for Resilience (WNTR). (2022).
42. USEPA/TEVA-SPOT-Toolkit. https://github.com/USEPA/TEVA-SPOT-Toolkit. Last accessed 29 August 2024
43. Uyak, V., et al.: Multiple linear regression modeling of disinfection by-products formation in Istanbul drinking water reservoirs. Sci. Total. Environ. **378**(3), 269–280 (2007)
44. Okoji, A.I. et al.: Performance evaluation of artificial intelligence with particle swarm optimization (PSO) to predict treatment water plant DBPs (haloacetic acids). Chemosphere **344**, 140238 (2023)
45. Hong, H., et al.: Using regression models to evaluate the formation of trihalomethanes and haloacetonitriles via chlorination of source water with low SUVA values in the Yangtze River Delta region. China. Environmental Geochemistry and Health. **38**(6), 1303–1312 (2016)
46. Soliyev, A.R., et al.: Spatio-temporal variability of halogenated disinfection by-products in a large-scale two-source water distribution system with enhanced chlorination. J. Hazard. Mater. **423**, 127113 (2022)
47. H2oforall. https://h2oforall.eu. Last accessed 10 September 2024

# Development of 5G Smart Farming Dashboard to Detect Wild Animals on Pasture by Using Convolutional Neural Network

**Ali Akyol, Rami Chahin, Jorge Marx Gómez, Hendrik Schwabe, Henrika Schwanke, and Nora Uderstadt**

**Abstract** Agriculture in Central Europe serves the purpose of food production but it also takes place in natural wildlife habitats. This means that disturbance of wild animals cannot be completely avoided. When it comes to mowing of grassland or harvesting, especially deer fawns are in danger because they naturally have no reflex of escaping. To address this, Germany has introduced legal regulations and support measures, including subsidies for drones equipped with thermal imaging cameras. These technologies are part of the "5G Smart Country" project, which focuses on integrating digital solutions into agriculture. This paper presents a drone-based system developed within the project to efficiently detect and protect fawns during agricultural activities, enhancing both animal welfare and farm productivity.

**Keywords** 5G · Fawn Protection · Smart Farming · Wildlife Conservation · Agricultural Digitalization · Grassland Mowing Safety

A. Akyol (✉) · R. Chahin · J. M. Gómez
University of Oldenburg, Ammerländer Heerstraße, Oldenburg, Germany
e-mail: ali.akyol@uol.de

R. Chahin
e-mail: rami.chahin@uol.de

J. M. Gómez
e-mail: jorge.marx.gomez@uol.de

H. Schwabe · H. Schwanke · N. Uderstadt
Chamber of Agriculture in Lower Saxony, Wunstorfer Landstr, Hannover, Germany
e-mail: hendrik.schwabe@lwk-niedersachsen.de

H. Schwanke
e-mail: henrika.schwanke@lwk-niedersachsen.de

N. Uderstadt
e-mail: nora.uderstadt@lwk-niedersachsen.de

 
V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_18

## 1 Introduction

Agriculture in Central Europe occupies a large area due to favorable cultivation conditions, thereby ensuring food production for the population. Simultaneously, agricultural lands are inhabited by various animal species, ranging from insects and small rodents to larger herbivores and carnivores. Since wild animals are part of the ecosystem and deserve protection, farmers are encouraged to preserve and even enhance biodiversity through their land management practices [1].

One of the agricultural duties includes the protection of fawns during grassland mowing. Deer naturally inhabit semi-open areas, making agriculturally used land ideal feeding grounds and preferred fawn-rearing sites. Research has shown that deer tend to choose bed-sites for fawns in dense vegetation and fields near forests. Both cultivated grasslands and taller cereal crops are suitable bed-sites. Fawns are especially likely to be found in areas where the surrounding fields have already been harvested, reducing the available protective cover [2, 3].

In Germany in 2023, an area of 16.4 million hectares was used for agriculture. The land use is divided into arable land with 11.7 million hectares and permanent grassland with 4.7 million hectares. Of the permanent grassland, 2 million hectares were used as meadows [4]. These areas are usually mown mechanically to produce fodder for livestock and are therefore potential fawn rescue areas.

In their first weeks of life, fawns lack a distinct scent and remain at the location chosen by their mother [5, 6]. Due to their effective camouflage, it is very difficult to detect resting fawns. Consequently, accidents regularly occur during mowing or harvesting, where fawns are caught in cutting machinery, leading to their death or severe injuries [6]. All parties involved aim to avoid such accidents as much as possible. These incidents are not only devastating for the animals and the deer population but can also psychologically burden the farmers. Additionally, the affected harvest becomes unusable, representing an economic loss for the farms. For these reasons, and due to the duty of care required from farmers, interest in efficient fawn protection has significantly increased in recent years.

To make fawn rescue more attractive, support measures for fawn rescue have already been launched in Germany. The purchase of drones with thermal imaging cameras is being subsidized [7]. Still, the workload is high resulting in the amount of people needed, the work in the early morning, often before sunrise, and the complicated communication between drone pilot and rescue team. The presented work in this paper provides implications for a more efficient model of fawn rescue that could be adopted in practice and enhance and facilitate fawn rescue. This model workflow of fawn rescue was developed in the context of the "5G Smart Country" project, which explores the possibilities of digitalization in agriculture. This paper provides an efficient solution for drone-based identification of fawns on agricultural lands which is presented below.

## 2 5G Smart Country Project

In the BMVI-funded project “5G Smart Country,” technological development and its application in rural areas are being explored in the districts of Helmstedt and Wolfenbüttel in Lower Saxony from December 2021 to December 2024 [8]. The project focuses on testing various 5G applications under real-world conditions to assess the potential of 5G mobile communication, especially in rural areas. The aim is to outline the opportunities and possibilities offered by 5G, providing a decision-making aid for potential users before their purchases, and thereby encouraging the adoption of 5G technology in the private sector [9].

The project is divided into two subprojects: “5G Smart Farming” and “5G Smart Forestry.“ Within the “Smart Farming” subproject, several use cases related to the digitalization of arable farming are investigated. It aims to leverage 5G technology to modernize and digitize agriculture in rural areas of Lower Saxony. It focuses on enhancing agricultural efficiency, sustainability, and competitiveness by developing and testing innovative applications and solutions. A promising application of 5G technology in this context is wildlife detection, specifically rescuing fawns from agricultural machinery. By integrating 5G technology, drones and other surveillance systems can be used more effectively for detecting and protecting wildlife.

## 3 State of the Art Wildlife Protection and Fawn Rescue in Farmed Grassland

Fawn rescue has long been a significant challenge in wildlife management, particularly during the mowing of meadows and fields. Finding fawns in the field remains a time-consuming task, requiring a well-coordinated process. Fawns are particularly vulnerable in their first weeks of life due to their natural protective strategies, such as remaining motionless in tall grass. The mother leaves the fawns alone for additional protection, returning only to nurse them. While this strategy protects them from natural predators, it makes them susceptible to agricultural machinery. Each year, it is estimated that several thousand fawns in Germany die or are maimed by mowers, leading to ethical, ecological, and economic issues [10, 11].

Before the mowing begins, the field must be inspected, which a farmer cannot accomplish alone. Assistance from local hunters and volunteers is essential. Weather and vegetation conditions are also crucial, allowing farmers to plan the mowing a few days in advance. Traditional fawn rescue methods are labor-intensive, requiring numerous helpers. In the early morning before mowing starts, helpers form a human chain to comb the fields on foot, often accompanied by trained dogs, to locate and secure the hidden fawns [6, 11, 12]. Although this traditional method is time-consuming and not always successful due to high vegetation and the fawns’ natural camouflage, modern alternatives, such as drones equipped with thermal imaging cameras, offer significant advantages.

The responsibility for locating fawns lies with farmers, who are legally required to actively protect wildlife as stipulated in the Animal Welfare Act §17. The German Animal Welfare Act (TierSchG) aims to protect animals from unnecessary suffering and pain. Particularly relevant for fawn rescue is § 17 of the TierSchG, which states:

"Anyone who kills a vertebrate without reasonable cause or inflicts significant pain or suffering on a vertebrate out of cruelty or causes it to endure significant pain or suffering over a longer period or repeatedly, shall be punished with imprisonment for up to three years or a fine." [1].

This paragraph clearly establishes that it is a criminal offense to kill vertebrates, including fawns, without reasonable cause or to inflict significant pain or suffering on them. As fawns are often overlooked during mowing and can suffer severely or die as a result, failing to take measures to prevent these deaths is considered a violation of this paragraph. Therefore, farmers are legally obligated to take all reasonable measures to prevent fawns from being harmed by agricultural machinery. This means either using traditional search methods or modern technologies such as drones with thermal imaging cameras to locate and secure the fawns before mowing.

## 3.1 5G Technology in Wildlife Detection

Wildlife detection benefits significantly from the advantages of 5G technology. The most important factors of data transmission via 5G mobile networks are:

- **Real-Time Data Transmission:** Drones equipped with thermal imaging cameras can transmit high-resolution images and videos in real time to control centers or mobile devices, enabling immediate response to detected animals.
- **Low Latency:** The low latency of 5G ensures that drone control and image processing occur without delay, crucial for precise navigation and quick decision-making during rescue operations.
- **AI Integration:** With 5G support, AI algorithms for image and video analysis can be executed in real time. These algorithms automatically detect fawns and other wildlife and send alerts to farmers and rescue teams on the ground.

The utilization of drone technology with thermal imaging cameras results in the generation of a substantial quantity of data at the point of field operation. This data should be processed digitally immediately. The deployment of 5G technology facilitates the transmission of the resulting data directly to data centers, where it is processed and subsequently transmitted to the user.

According to Martin Israel [13] the fawn rescue system with drones can be classified into distinct expansion stages, which are delineated as follows:

- First, a tool is needed that can detect fawns by identifying their relevant attributes on thermal images.
- The second phase is to accelerate the process steps due to georeferencing the detected position.

- The third phase provides an initial pattern recognition algorithm, which is applied to the analog video signal.
- In the fourth phase the georeferencing in particular is optimized, so that the process step of locating the site can be accelerated.

The third and fourth phases, in which a pattern recognition algorithm is applied to the analog video signal and georeferencing is optimized to localize the fawn, are an essential part of current scientific work.

### *3.2 Research Questions (RQ)*

1. What are the challenges and limitations of implementing 5G technology and drones in wildlife detection and rescue?
2. How effective are modern methods in fawn detection and rescue?
3. How can AI integration improve the efficiency and reliability of fawn rescue operations?

## 4 Results

### *4.1 Technical Concept*

The technical concept of the 5G project is the core framework where all project requirements are addressed. The two primary components of this concept are hardware and software. The hardware includes Unmanned Aerial Vehicles (UAVs), Unmanned Ground Vehicles (UGVs), and sensors, while the software encompasses the dashboard (see Fig. 1), backend, AI, and the database.

A user-friendly dashboard, containing essential information and interface elements, is crucial for farmers, who are the primary stakeholders of the project. The dashboard provides three types of information:

1. General information, such as weather data, crop yield analyses, soil conditions, wild animal intrusions, and weed management.
2. Live-streaming and live geoinformation from UAVs and UGVs.
3. Information about UAVs, UGVs, and sensors, including the ability to set work schedules and plans for activities such as harvesting, mowing, planting, sowing seeds, irrigation, and fertilization. Farmers can modify and reset this information, along with other mission parameters, as needed.

For example, the interaction dashboard pages, such as the machines page, enable farmers to manage their fieldwork by setting timers to start or stop specific actions in real time (see Fig. 2). These pages also provide essential information and detect potential threats to crops, including animals, diseases, and weeds, both in the field

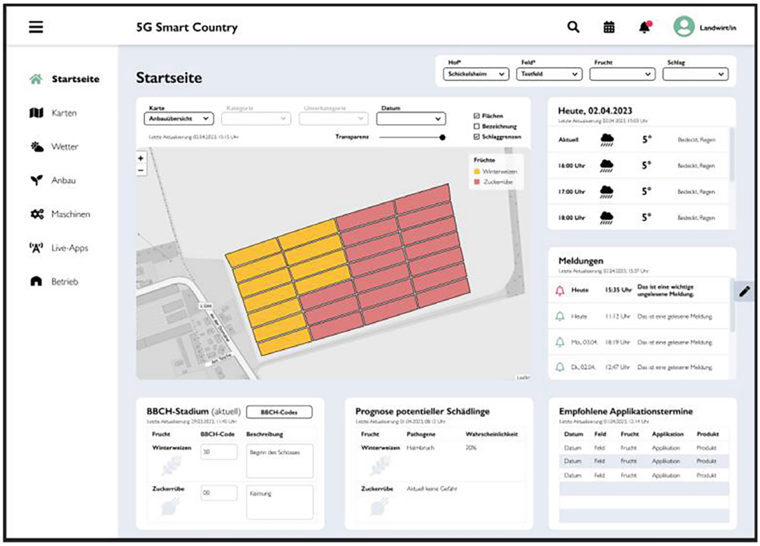



**Fig. 1** Developed concept for general information about the field (own source)

and in plantings. After presenting this initial information, the next step is to determine how to tackle the identified issues or challenges.

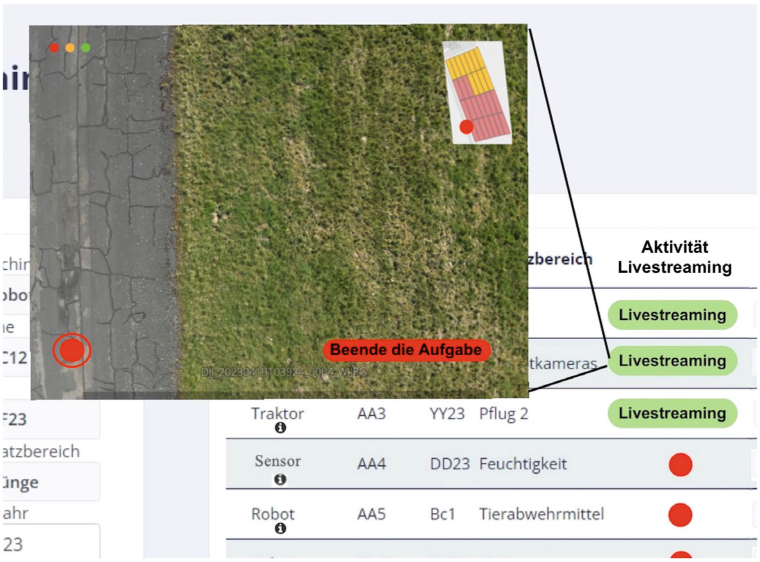



**Fig. 2** Live-streaming dashboard and machine status (own source)

## 4.2 Process Description

The UAVs are used to capture high-resolution images and live-stream videos of the agricultural fields. These UAVs provide valuable data for monitoring crop health, detecting wild animals, and assessing other field conditions. The MinIO, used to store the large volumes of data collected by the UAVs, is located at the German Aerospace Center (DLR). The aerial imagery and videos are uploaded to MinIO, where they are securely stored and easily accessible for further processing and analysis. Only authorized users, such as farmers and field managers, can access the data stored in MinIO through Keycloak, an open-source identity and access management solution. Keycloak provides single sign-on (SSO), user federation, and role-based access control.

In the backend service, FastAPI serves as the backend service that interacts with the UAV data and manages the application's business logic. It exposes RESTful APIs that allow users to query the UAV data, trigger processing tasks, and receive insights. FastAPI handles requests such as retrieving images from MinIO, processing data with machine learning models, or integrating with other services. The model used to analyze the photos is YOLO ("You Only Look Once"). Its mission is to detect wildlife in the field using an object detection algorithm designed to recognize and locate animals in images and videos in real time.

The reason for using the YOLO AI model is its different approach compared to traditional methods, which often involve multiple stages such as region proposal, feature extraction, and classification. YOLO divides the image into a grid and simultaneously predicts bounding boxes and class probabilities for each grid cell. This allows it to process images quickly and efficiently, making it suitable for applications that require fast object detection. YOLO is also ideal for this project due to its speed and accuracy in analyzing photos.

GeoServer and Leaflet are integrated into the existing system to add additional capabilities for geographic data management and visualization. GeoServer is a Java-based open-source server that allows for geospatial data to be edited, shared, and processed [A1]. The 5G project uses GeoServer to manage and serve geospatial data layers related to agricultural fields, particularly the geographic position data for wild animals in the field.

The interactive maps displayed by Leaflet, an open-source JavaScript library, provide an interface for displaying maps in web applications. Leaflet visualizes the geographic position of detected animals in real time for users [A2].

The MQTT broker (Message Queuing Telemetry Transport) serves as a crucial component for efficient data communication and integration between various system elements. MQTT is a lightweight, publish-subscribe messaging protocol designed for low-bandwidth and high-latency networks, making it ideal for Internet of Things (IoT) applications [A3]. Here's how the MQTT broker functions within the project's system:

Data collection and transmission through sensors mounted on UGVs, UAVs, and other equipment publish data (e.g., images, sensor readings, environmental data)

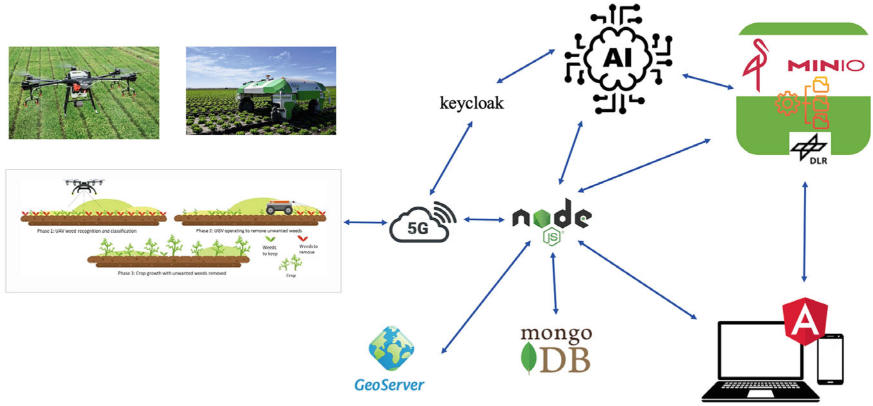



**Fig. 3** Relationship between the different elements of the project 5G (own source)

to the MQTT broker. The MQTT broker then handles the distribution of this data to the AI server. The lightweight nature of the MQTT protocol allows for real-time transmission of sensor data, which is crucial for timely analysis of wild animal presence and decision-making [A4]. The interactive maps displayed by Leaflet which is an open-source JavaScript library provide an interface for displaying maps in web applications. Leaflet visualizes the geographic position of the detected animals for users in real time [A2].

The relationship between the components in the 5G project for wild animal detection (see Fig. 3) can be summarized as follows:

- UAVs capture live-streamed data from the field.
- MinIO stores the data captured by the UAVs.
- GeoServer manages and serves geospatial data layers, showing the positions of detected wild animals.
- Leaflet provides interactive map visualization of the data, displaying wild animal positions on the map.
- FastAPI acts as the backend service for data access and processing, detecting wild animals in photos.
- Keycloak handles authentication and authorization.
- Angular is the web application framework where the dashboard is designed, integrating the backend, frontend, and database of the project. It also contains all important information and interface elements for the user.
- The MQTT Broker acts as a communication hub within the agricultural monitoring system, enabling efficient, scalable, and real-time data exchange between sensors, servers, external data sources, and user interfaces.
- MongoDB serves as a central data repository, providing a flexible, scalable, and robust solution for storing and managing the diverse data types generated by the system. It works in conjunction with the MQTT broker, application server, AI

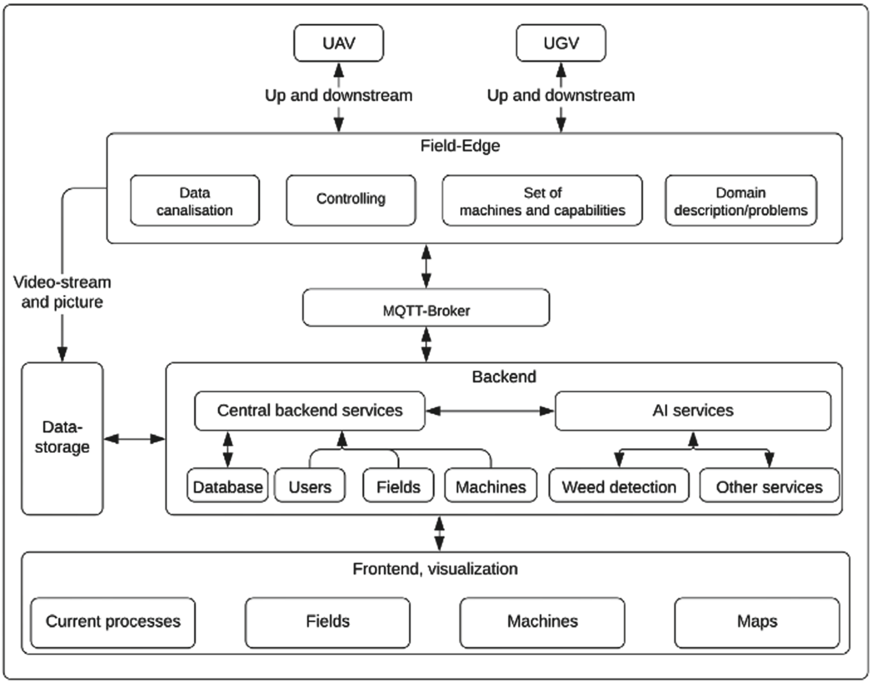



**Fig. 4** Technical concept (own source)

server, and frontend UI to enable efficient data handling, real-time insights, and comprehensive monitoring and management of agricultural operations.

Figure 4 shows the overview of the technical concept.

### 4.3 The Functionality of the AI Model and Its Accuracy

The AI model achieves a precision of 0.982, successfully detecting 98.2% of wild animals (see Fig. 5). Since this exceeds the target accuracy of 95%, this success criterion is met. A high recall indicates few false negatives, as recall measures the model's ability to detect all wild animals present without missing many. With a recall of 0.995, the model detects 99.5% of the wildlife that is actually present. This is an excellent result, suggesting that the model effectively minimizes false negatives. Furthermore, the F1 score of 0.988 demonstrates a strong balance between precision and recall, indicating that the model performs well in both minimizing false negatives and detecting wildlife. Therefore, this success criterion is also deemed fulfilled. The provision of real-time detection capabilities, with a processing time of less than one minute, depends on components that are still under development during the

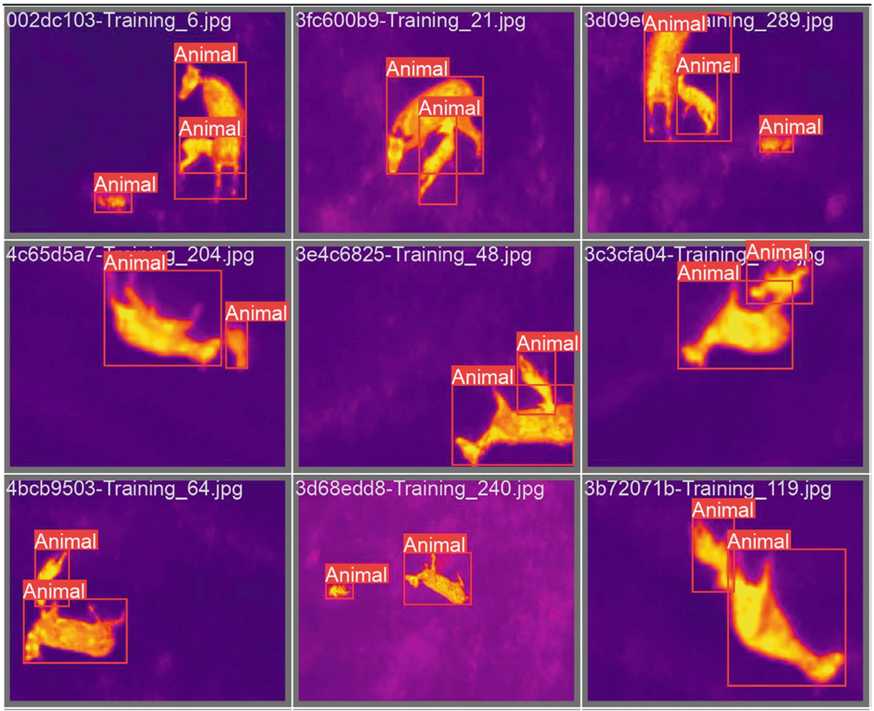



**Fig. 5** Example of predictions on test images (own source)

deployment phase. Consequently, the assessment of this criterion's fulfillment will be revisited at a later stage.

# 5 Implementation

## *5.1 Cross-Industry Standard Process for Data Mining (CRISP-DM)*

During the implementation, we followed the Cross-Industry Standard Process for Data Mining (CRISP-DM), which was introduced by Chapman et al. in 2000 [14]. This model facilitates the execution of data mining processes and comprises distinct phases that are represented within a project.

**Business Understanding.** During the Business Understanding phase, the goal is to comprehend and define the business objectives. This phase is divided into various sub-phases: understanding the fundamentals, identifying business objectives, and determining the criteria for business success [15]. In our case, the objective is to identify and rescue wild animals with the help of AI. The next step in Business Understanding is the "Situation Assessment." In this phase, the situation is evaluated, and the available resources are assessed. In this context, technological resources and an overview of them are presented. In this work, the overview of the resources used is detailed in Chapter 4.

**Data Understanding.** In the Data Understanding phase of the project aimed at detecting fawns using AI and drones, the primary objective is to collect, comprehend, and identify potential issues with the data gathered by the drones. This phase is methodically divided into several sub-phases: data collection, data description, data exploration, and data quality verification [15].

The initial step involves collecting data (Data Collection) from drones executing missions over agricultural fields. The drones capture real-time footage and transmit it directly to a Python-based AI server. This data comprises images intended for the identification of fawns. The next step in Data Understanding is Data Description. The data will be collected as images with GPS (Global Positioning System) coordinates and timestamps. The subsequent steps are Data Exploration and Data Quality assessment.

**Data Preparation.** To prepare data for modeling, several important steps must be undertaken. The data collected from the field should be cleaned, transformed, and structured using various algorithms to detect wild animals. These steps also include tasks such as feature selection and managing the final dataset. For the 5G project, this involves processing aerial images, annotating animal sightings, and integrating data from thermography cameras on UAVs and UGVs. Accurate data preparation is crucial for building effective detection models. This phase ensures that the data fed into the models is clean, relevant, and well-structured, which is essential for accurately identifying animals under different conditions.

A part of Data Preparation is Data Augmentation. Excluding standard drone footage has resulted in a smaller dataset, now comprising 37 thermal images that capture wildlife. These images were taken during various drone flights, ensuring that different ambient temperatures were considered. Due to the vegetation, such recordings can only be conducted in the spring, coinciding with the spring mowing, and it is not possible to expand the dataset with new recordings for this study.

To increase the dataset and enhance the success of model training, new data points were generated from the existing data using data augmentation techniques (see Fig. 6). Data augmentation involves artificially creating new data by modifying the original dataset. Various techniques were applied to the original images, including random rotations, transformations along the horizontal and vertical axes, cropping (either randomly positioned or centered), as well as random zoom levels, scaling, and resizing. This approach enables effective model training even with an initially small dataset.

For this purpose, the Python library Augmentor was utilized. Augmentor employs a stochastic approach, using building blocks that allow operations to be combined in a pipeline through which the images are processed [16]. Therefore, in this study, data augmentation was also performed using Augmentor.

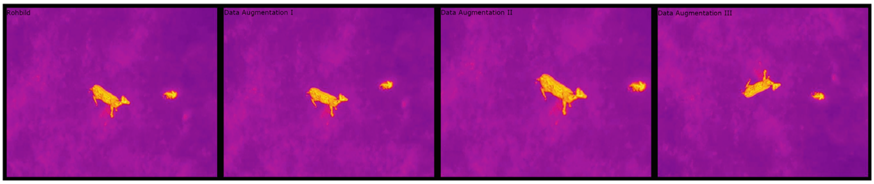



**Fig. 6** Example of Data augmentation (own source)

**Modeling.** For wildlife detection, YOLO (You Only Look Once) is used for object detection due to its efficiency and speed. YOLO, based on Convolutional Neural Networks (CNNs), allows for object detection with a single pass through the network, predicting multiple bounding boxes and class probabilities simultaneously. This architecture offers advantages over traditional methods by enabling the rapid identification of objects, making it suitable for images, videos, and live feeds. YOLO can process up to 45 images per second on a Titan X GPU, with potential speeds of up to 150 images per second with increased computational power.

Unlike other real-time systems like Faster R-CNN, which separate classification and regression outputs, YOLO uses a regression approach, resulting in fewer background errors. Since YOLOv5, the models have been fine-tuned to balance speed and accuracy, with various model scales (nano, small, medium, large, and extra-large) catering to different applications and hardware requirements. This flexibility allows YOLO to be optimized for edge devices and used in real-time detection for autonomous vehicles and precision agriculture [17–19].

**Evaluation.** The evaluation phase involves testing the detection models to ensure they meet the desired performance metrics. This includes validating the models with new data, testing in real field conditions, and measuring detection accuracy, false positives, and response times. Evaluation ties the modeling phase back to the original goals of the 5G project in improving agricultural practices. If the models don't detect animals accurately or in a timely manner, the project may need to revisit earlier phases to improve the data or models. Evaluation ensures that the system is reliable before it's deployed in the field.

**Deployment.** Deployment involves integrating the detection models into a real-world system. This includes implementing the models in UAVs and UGVs, deploying monitoring dashboards using Angular, and setting up automated alerts when wild animals are detected. Deployment is the culmination of all the previous phases, where the system is put into action to achieve the main goals of the 5G project. Once deployed, the system will actively monitor fields, detect animals, and help prevent accidents involving wild animals in the field.

Reflecting on the process of this work, it is crucial to evaluate whether all relevant factors have been adequately addressed and to what extent the insights gained can be applied to future projects. Initially, the project involved a thorough analysis of the requirements, using the goals and success criteria from the 5G project for wildlife detection. With the usage of AI and thermal cameras wild animals can be detected

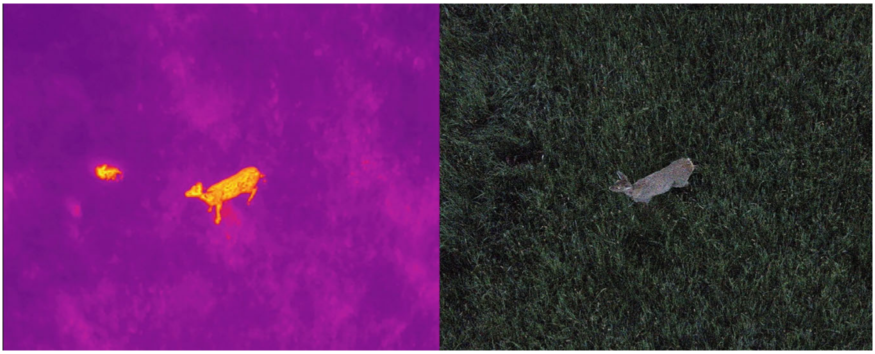


**Fig. 7** Difference between a thermal camera (left side) and a normal camera (right side) (own source) [20]

more efficiently (see Fig. 7). The fawn is in Fig. .7 on the image (normal camera) not visible but with thermal camera.

## 6 Discussion and Outlook

The use of 5G technology, drones with thermal cameras, and AI servers in the "5G Smart Country" project facilitates more effective and targeted collaboration among farmers, hunting associations, and conservationists to protect wildlife while enhancing agricultural production efficiency. This demonstrates how modern technology can contribute not only to productivity but also to animal and environmental protection. Additionally, the legal obligation underscores the necessity and importance of fawn rescue measures, placing responsibility firmly on farmers.

**RQ1**: What are the challenges and limitations of implementing 5G technology and drones in wildlife detection and rescue?

The implementation of 5G technology in wildlife rescue with drones offers several benefits that significantly enhance the efficiency, accuracy, and reliability of rescue operations. Higher data transmission rates, lower latency, more reliable connections, increased capacity for connected devices, improved edge computing utilization, and expanded deployment possibilities are key factors that can substantially improve the success of wildlife rescue operations with drones. Through the use of 5G, fawns and other wildlife can be protected more effectively from dangers, making an important contribution to animal welfare and biodiversity conservation.

Despite promising advances, several challenges hinder the widespread adoption of these technologies. These include the acquisition costs of drones and thermal cameras, the need for trained personnel to operate the drones, and regulatory hurdles regarding the use of drones in agriculture. Nonetheless, these technological innovations offer a promising perspective to significantly improve the efficiency and humaneness of fawn rescue, thus making a valuable contribution to wildlife protection.

**RQ2**: How effective are modern methods in fawn detection and rescue?

Drone-based rescue operations are significantly faster than traditional methods. In comparison to the current method, it shows a noteworthy improvement. Regarding manual drone flights, the act of flying in straight lines is often found to be challenging and less precise. In automated flight route planning, algorithms optimize the flight routes of the drone to ensure more efficient coverage of the area, which enhances the reliability of fawn rescue. This approach not only improves efficiency but also reduces the need for a large number of personnel.

Conversely, the provision of drone-based fawn rescue is only gradually being introduced in more regions. A crucial factor is the restricted timeframe for the application and the considerable expense associated with this modern technology. The objective of the optimization process is to maximize the area output, while ensuring that no fawn is overlooked [13]. It is therefore necessary to enhance the efficiency of the wildlife rescue system and reduce the cost of the technology in order to provide a greater incentive for its use in agricultural contexts.

**RQ3**: How can AI integration improve the efficiency and reliability of fawn rescue operations?

Manual flying in a straight line is challenging with a drone. Due to automated route creation, the drone flies more efficiently. With a planned flight mission, the processes are controlled and the parameters along the route are known. When fawn detection is indicated, it is important to achieve the highest possible level of precision to ensure that it is actually a fawn and not another heat source. This is where the AI helps, enabling targeted identification through learned patterns.

The incorporation of artificial intelligence into the field of fawn rescue operations, particularly in conjunction with drone technology, provides a notable enhancement in efficiency and reliability. In comparison to human operators, who require a longer period of time to interpret data or images, artificial intelligence enables the acceleration of decision-making processes and a higher degree of consistency in results. This is particularly advantageous in time-critical wild animal rescue operations. The deployment of AI mitigates the risk of human error, such as erroneous judgments about environmental conditions or misinterpretations of data. This is of crucial importance to ensure the safety and well-being of the fawns. Errors in identification may result in the failure to implement potential rescue measures or the infliction of harm upon the animals in question. The precision of AI systems enables the reliable identification of fawns.

The application of drones in combination with AI systems enables immediate data processing, for instance through thermal cameras, for the detection of fawns. The accelerated, automated recognition process reduces the time required for identification and improves the probability of a successful rescue (Fig. 7).

To further improve precision, the flight route could be adjusted during the mission based on the reliability of fawn detection. For instance, if detection accuracy falls below a certain threshold, adjustments such as using more powerful technology or lowering the flight altitude may be required. An algorithm could be designed to manage this optimization process. If necessary, the algorithm could generate an intermediate mission for more detailed imagery, ensuring reliable detection in a short time.

**Acknowledgements** This research was made possible by Federal Ministry for Digital and Transport. We thank our colleagues in the project (FKZ 45FGU117) who provided insight and expertise that greatly assisted this research.

## References

1. Animal Protection Act (TierSchG), as promulgated on May 18, 2006. Federal Law Gazette I p. 1206, 1313. Last amended by Article 2 of the Act of July 16, 2021. Federal Law Gazette I p. 2970. Available at: www.gesetze-im-internet.de/tierschg/BJNR012770972.html. Accessed 09 July 2024
2. Linnell, J.D.C., Nilsen, E.B., Andersen, R.: Selection of bed-sites by roe deer Capreolus capreolus fawns in an agricultural landscape. Acta Theriologica **49**, 103–111 (2004). [Online]. Available: https://doi.org/10.1007/BF03192512.
3. Baur, S., Kauffert, J., Hewison, A.J.M., Reinermann, S., König, A., Menzel, A., Peters, W.: Spatial scaling in bed-site selection by roe deer fawns: Implications for mitigating neonatal mortality during mowing. Ecology and Evolution **13**, e10729 (2023). [Online]. Available: https://doi.org/10.1002/ece3.10729.
4. RDC of the Federal Statistic Office and Statistical Office of the Federal States of Germany, 41271–0003
5. Linnell, J.D.C., Andersen, R.: Timing and synchrony of birth in a hider species, the roe deer Capreolus capreolus. J. Zool. **244**, 497–504 (1998)
6. Wimmer, T., Israel, M., Haschberger, P., Weimann, A.: Der Fliegende Wildretter in Aktion: DLR und BJV nutzen ferngesteuerte Flugplattform zur Rehkitzrettung. In Symposium des Landesjagdverbandes Bayern - Hege und Bejagung des Rehwildes, Augsburg (2013). ISBN 978–3–00–041796–2.
7. Bundesanstalt für Landwirtschaft und Ernährung (BLE) (2024). https://www.ble.de/DE/Projektfoerderung/Foerderungen-Auftraege/Rehkitz/Rettung_node.html
8. 5G-Smart-Country (2024). https://5g-smart-country.de/
9. Bundesministerium für Verkehr und digitale Infrastruktur: Bekanntmachung der Förderrichtlinie '5G-Umsetzungsförderung im Rahmen des 5G-Innovationsprogramms'. [Online]. Available: https://www.bav.bund.de/SharedDocs/Downloads/DE/5G/Foerderrichtlinie.pdf?__blob=publicationFile&v=3 (2020). Accessed 18 July 2024
10. FRANKONIA Handels GmbH&Co KG.. Fawn Rescue with Drones. The Little Hunter's Guide, Hunting and Knowledge, 2–13. Von https://www.frankonia.de: https://www.frankonia.de/blaetterkatalog/Jaegerfibel-Kitzrettung/ (2023, April 25). Accessed 09 July 2024
11. Kinser, A., Ganteför, S., Münchhausen, H., Brandt, A.A. (Eds.): Wildtierschutz bei der Mahd: Ein Ratgeber zur Rettung von Jungwild und Wiesenvögeln (3rd ed.), Deutsche Wildtier Stiftung (2023)
12. Cukor, J., Havránek,F., Vacek, Z., Bukovjan, K., Podrázský, V., Sharma, R.P.: Roe deer (Capreolus capreolus) mortality in relation to fodder harvest in agricultural landscape. Mammalia, **83**(5), 461–469 (2019) [Online]. Available: https://doi.org/10.1515/mammalia-2018-0002
13. Israel, M.: Entwicklung eines UAV-basierten Systems zur Rehkitzsuche und Methoden zur Detektion und Georeferenzierung von Rehkitzen in Thermalbildern. Der Fliegende Wildretter, pp. 50–51 (2016)
14. Chapman, P.: CRISP-DM 1.0: Step-by-step data mining guide (2000)

15. Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., Wirth, R.: The CRISP-DM user guide. Presented at the 4th CRISP-DM SIG Workshop, Brussels (1999).
16. Bloice, M.D.: Github repository: Augmentor [Online]. Available: https://github.com/mdbloice/Augmentor (2023). Accessed 12 July 2024
17. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You Only Look Once: Unified, Real-Time Object Detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788 (2016)
18. Chen, T., Girshick, R., Gu, S. et al.: YOLOv5: Real-Time object detection and localization. J. AI Res. **45**(3), 129–145 (2023)
19. Wang, C., Liao, M., Jiang, L., Smith, R.: YOLOv5 for precision agriculture: real-time object detection in autonomous systems. In: Proceedings of the International Conference on Precision Agriculture, pp. 98–105 (2020)
20. Akyol, A. University of Oldenburg. Bachelor thesis in 5G-Smart-Country Project (2023) (not public)

# Supply Chain and Sustainability Planning

# Overview, Evaluation and Comparison of Current CCS and CCU Technologies

**Josua Benz, Konstantin Blank, Silvio Hötzel, Jonas Just, Charlotte Lau, Franziska Schicks, Nicolas von der Ahe, and Jochen Wittmann**

**Abstract** This research paper comprehensively reviews current carbon capture and storage (CCS) and carbon capture and utilization (CCU) technologies. A thorough literature review served as the basis for compiling an overview of the prevailing technology landscape. Furthermore, the study examines the different pathways of the technologies in a life cycle analysis (LCA) concerning their efficiency in terms of carbon footprint. The analysis shows that some processes are highly energy-intensive, underlining the need for renewable electricity to minimize $CO_2$ emissions. However, the study also points out several challenges, including incomplete data and unknown variables that hinder the implementation and evaluation of these technologies. In addition, the criticisms and limitations associated with CCS and CCU stress the need for further research and development in this critical area.

J. Benz (✉) · K. Blank · S. Hötzel · J. Just · C. Lau · F. Schicks · N. Ahe · J. Wittmann
HTW Berlin - University of Applied Sciences, Berlin, Germany
e-mail: josua.benz@student.htw-berlin.de

K. Blank
e-mail: konstantin.blank@student.htw-berlin.de

S. Hötzel
e-mail: silvio.hoetzel@student.htw-berlin.de

J. Just
e-mail: jonas.just@student.htw-berlin.de

C. Lau
e-mail: charlotte.lau@student.htw-berlin.de

F. Schicks
e-mail: franziska.schicks@student.htw-berlin.de

N. Ahe
e-mail: nicolas.vonderahe@student.htw-berlin.de

J. Wittmann
e-mail: jochen.wittmann@htw-berlin.de

 
V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_19

**Keywords** Carbon capture and utilization (CCU) · Carbonation · Carbon footprint · Ex-situ · Harmonization · Life cycle analysis/assessment (LCA) · Mineralization · Negative emissions technologies (NETs)

## 1 Introduction

The Paris Agreement is a crucial milestone in global efforts to mitigate climate change and underlines the urgent need for nations to work together to reduce greenhouse gas emissions. Carbon Capture and Storage (CCS), the process of trapping and storing Carbondioxide ($CO_2$) emissions, and carbon capture and utilization (CCU), the transformation of $CO_2$ into valuable products, are crucial technologies for meeting the agreement's ambitious targets. In $CO_2$-intensive industries like the cement industry, capturing and mineralizing 1t of Carbondioxide equivalents (CO2-eq) could avoid over 1t of $CO_2$-eq emissions by substitution of conventional production [1]. However, the pressing issue is how effectively these technologies can fulfill current commitments and drive progress.

Understanding the opportunities and limitations of these technologies is essential to assess their potential to achieve the goals set out in the Paris Agreement. While there is extensive research on the individual aspects of these technologies, there still needs to be a notable gap in synthesizing this knowledge into a coherent framework. Therefore, our aim is twofold: to consolidate the existing research into a comprehensive overview and to assess the collective potential of CCS and CCU to meet global emission reduction targets in terms of their carbon footprint[1], evaluated by a Life Cycle Assessment (LCA). In addition, our analysis will highlight existing data gaps and areas that require further investigation, thereby contributing to the ongoing discourse on climate change mitigation strategies.

## 2 Literature Review

In recent years, multiple studies have conducted literature reviews on Life Cycle Assessment in mineral carbonation systems. [2] reviewed 27 LCAs to compare the environmental impacts of CCS and CCU systems. They found 16 studies focusing on CCU and eleven studies on CCS without a deeper explanation of the search method. Researchers carried out 15 of the 16 studies on CCU between 2008 and 2013 and one in 1999. [2] concluded that although the Global Warming Potential (GWP) of CCS (in-situ) is lower than in CCU (ex-situ), other environmental impacts are higher. We recalculated some studies using 1t of removed $CO_2$ as the functional unit to enhance comparability.

[1] total amount of greenhouse gases produced directly and indirectly by a certain activity, usually measured in metric tons of $CO_2$ equivalent

More recently [3] carried out a comprehensive bibliometric analysis of publications about LCA on Carbon Capture, Utilization and Storage (CCUS) technologies from 1995 to 2018. By using the search string below, they found 234 publications (after preprocessing). They conclude that the broader usage of LCA methods on this topic started from 2007 onwards.

```
('carbon␣capture' OR 'CO2␣capture' OR 'carbon␣dioxide␣capture')
AND ('life␣cycle␣assess*' OR 'life␣cycle␣analysis' OR 'lca')
```

A systematic literature review was conducted by [4], where they found 29 relevant articles about LCA regarding CCU. We searched the databases of Emerald Insight, JSTOR, RCS, ScienceDirect, Scopus, SpringerLink, Web of Science, and Wiley using the following keywords:

```
('Carbon␣capture' OR 'Carbon␣utilization'
 OR 'Carbon␣capture␣and␣utilization'
 OR 'CO2␣capture␣and␣utilization' OR 'CCU'
 OR 'Carbon␣capture␣and␣conversion'
 OR 'CCC' OR 'Carbon␣dioxide␣utilization'
 OR 'CO2␣utilization' OR 'CDU')
AND ('Carbonation' OR 'Mineralization')
AND ('Life␣cycle␣assessment' OR 'Life␣cycle␣analysis'
 OR 'LCA' OR 'Sustainability␣assessment'
 OR 'LCSA' OR 'Environmental␣impact*'
 OR 'Greenhouse␣gas␣emission*' OR 'GHG␣Emission*'
 OR 'Carbon␣footprint')
```

In conclusion, [4] points out various aspects of the LCAs, such as methodology and handling of multifunctional processes. Most studies have a shared focus on climate change or GWP as the assessed impact category. In terms of technology, they found that 17 out of the 29 studies were using direct aqueous carbonation as their carbonation route, with indirect aqueous carbonation being the next most frequent, with eight studies. Based on their findings, [4] were able to carry out a meta LCA to compare the different technologies.

Therefore, this paper will build on the work from [4], which examined the environmental impacts and technological advances of CCS and CCU. Research in this area has grown exponentially in recent years, reflecting the increasing global importance of climate change mitigation efforts. Thus, this paper considers more recent studies published between 2022 and 2024.

## 2.1 Methodology

Based on [5, 6], we conducted a systematic literature review. Due to time and access constraints, two licensed databases—ScienceDirect[2] and SpringerLink[3]—were searched. First, a keyword table was created based on the research question. In order to find new articles matching the work of [4], we selected similar keywords for the search string. The search string was slightly adapted for each data source to achieve better results. The search period was limited to 2022-2024, as the focus of this project was primarily on recent research developments .

*SpringerLink*

The search string below yielded 578 results. By excluding books to focus on the latest research, we narrowed down the results to 47 hits. After skimming the headings, we excluded 15 articles because they were not thematically relevant (focus on agriculture, human genetics, specialized steel processes, or thematically to general). After reading the summaries of the remaining 32 articles, we excluded 24 articles because they had a different focus (digital farming, bacteria for recycling concrete, methods for producing bio-hydrogen, anaerobic digestion of energy crops, hydrogen supply chain, or human and planetary health implications of negative emissions technologies).

```
('Carbon␣capture' OR'Carbon␣utilization')
AND 'LCA'
AND ('Mineral␣carbonation' OR 'Carbon␣mineralization')
AND ('carbon␣footprint' OR 'climate␣change')
```

*ScienceDirect*

The search string below returned 23 results. After skimming the headlines and reading the abstracts, we identified 13 relevant articles.

```
('Carbon␣capture' OR 'Carbon␣utilization' OR 'CCU')
AND ('Mineral␣carbonation' OR 'Carbon␣mineralization')
AND ('Life␣cycle␣assessment' OR 'Life␣cycle␣analysis' OR 'LCA')
AND 'ex-situ'
```

## 2.2 Literature Review Since 2022

As mentioned, we found 21 publications in the SpringerLink and ScienceDirect databases. We had to exclude two of these because of access limitations. We reviewed the remaining 19 results regarding the following aspects:

- Was an LCA or a similar method carried out?

[2] https://www.sciencedirect.com.

[3] https://link.springer.com

- Which CCUS-technologies are used?
- other information concerning technology, such as carbonation route, feedstock, $CO_2$ source, etc.
- other information concerning LCA, such as functional units, system boundaries, system products

Four publications were carried out or included an LCA or a similar method. One of these was a meta-LCA based on a literature review [7]. There have been 14 publications featuring the topic of mineral carbonation or carbon mineralization. Apart from that, at least twelve also featured other technologies, such as synthetic fuels, oil recovery, horticulture, biochar, etc. Overall, five publications consisted of or included some kind of systematic literature review. Since the number of results that included an LCA concerning mineral carbonation is limited, it does not seem suitable to analyze these publications in the same way as [4]. Instead, the most relevant new studies and their most important results will be summarized. Still, as a first result, the review of publications since 2022 indicates a need for up-to-date LCAs on mineral carbonation [8]. summarizes several LCAs concerning negative emission technologies (NETs), including mineral carbonation. Only three of the ten regarded LCAs conducted an impact assessment on factors other than climate change. The functional units differed between the mass of $CO_2$ removed, the amount of energy produced from $CO_2$ source (power plants), or the mass of slag (feedstock).

That also translates to the LCA on cement, paper, and rubber substitutes by [9] found in the systematic literature review as they only considered climate change as an impact category and 1t of the product as the functional unit. They have found that although paper fillers made of carbonated materials have the highest relative footprint reduction, the overall highest emission reduction potential lies in the cement industry because of its large market volume.

The other LCA found by [10] focuses on off-grid direct air carbon capture and storage systems (DACS). Note that this study considered in-situ carbonation, falling outside the scope of our work. We define 1t of removed $CO_2$ as the functional unit and assess 16 impact categories.

Reference [11] analyzed different pathways for producing liquefied biomethane by applying a life cycle climate impact assessment. They chose a well-to-wheel approach, and therefore, the functional unit was 1 km of transport with a 40-tonne long haulage truck. Utilization of $CO_2$ focused on methanol production, concrete curing, refrigeration, and $CO_2$ storage but not ex-situ mineral carbonation.

Finally, [7] reviewed studies that used the methods of LCA and concluded that the carbonation of alkaline solid wastes could bring environmental benefits.

### *2.3 Categorisation and Overview of Technology Landscape*

The main aim of the research project was to comprehensively map the technology landscape and enable a comparative analysis of different technologies. Various

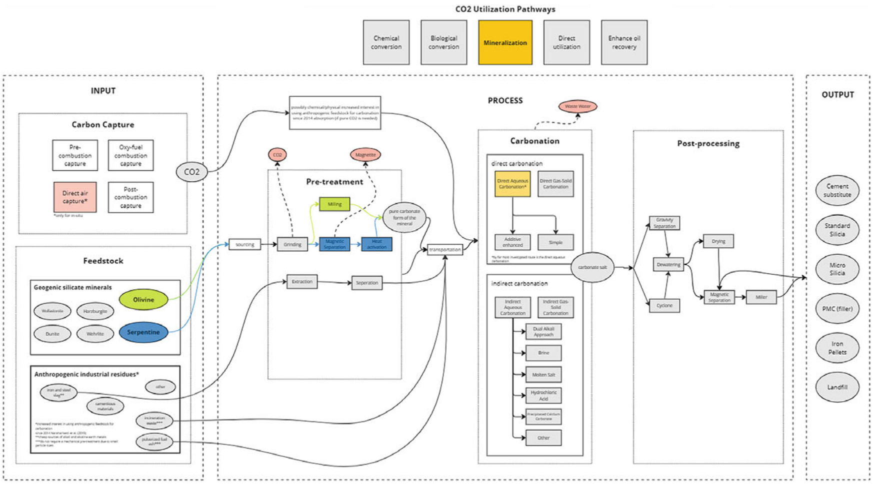



**Fig. 1** Overview of technology landscape (own illustration)

studies and sources were combined to create a complete picture of $CO_2$ utilization through mineralization. This summarized overview is presented visually in Fig. 1, divided into three main sections: Inputs, Processes and Outputs.

The figure shows inputs and outputs as ovals and the processes as rectangles. Most technologies have the same basic steps, including carbon capture, feedstock supply, pre-treatment, carbonization and post-treatment. However, a closer look at these process components reveals significant differences between technologies, highlighting the complexity and nuances in this area. The extensive overview attempts to summarize as many details as possible. Nevertheless, it is not nearly complete, but it offers a deeper insight into the technology landscape.

## 3 Meta-Model

A model is required to execute a LCA calculating the environmental impact of technologies and processes. LCAs of carbon mineralization face the challenge of modeling many different processes and materials simply and comparably. For this reason, we developed a simplified meta-model. The model should consider the complexity of the various processes and feedstocks involved in carbon mineralization.

**Problems creating a meta-model** Carbon mineralization involves various processes depending on the technology and feedstock used. These processes include $CO_2$ capture, the reaction of $CO_2$ with mineral materials, the transport of materials and products, the pre-treatment of feedstocks, the further processing of products, and the utilization of electric energy, heat, and water. Modeling these processes in a single comprehensive LCA model can be highly complex and resource-intensive. Some of the specific modeling issues we encountered during this study are:

**Technology pathways:** various technologies and approaches to carbon mineralization can differ greatly in their processes, materials, and environmental impacts. Modeling this diversity requires the development of a flexible model that can accommodate the different technologies without becoming too detailed.

**Feedstock diversity:** Feedstocks for carbon mineralization can vary widely and include natural minerals (like olivine or serpentine), waste products (like steel slag, ashes, etc.), and $CO_2$ from various sources. Each feedstock has different properties and requires different processes and conditions to optimally capture $CO_2$. The model must analyze the varying environmental impacts resulting from the different resources and energy amounts required by each feedstock.

**Interactions between processes:** The various carbon mineralization processes interact in complex ways that influence the system's overall performance and environmental impact. These interactions must be integrated and simplified into the model to enable a holistic assessment.

**Requirements for the meta-model** We developed a simplified meta-model to analyze the complexity of different carbon mineralization technologies and ensure comparability between different technologies and feedstocks. This meta-model should be an abstract model that simplifies the structure and behavior of the complex system by identifying the most critical factors and relationships. We intended the meta-model to have the following characteristics:

**Simplification:** The developed meta-model should reduce the complexity of carbon mineralization by eliminating unnecessary details and boundary effects. The goal was to abstract the different processes of the various pathways at a superordinate level and to summarise sub-processes.

**Comparability:** The meta-model should enable the comparability between different technologies and feedstocks by using consistent processes and a common functional unit as an assessment reference to quantify the relevant environmental impacts.

**Parameterization:** It should be possible to parameterize the meta-model to enable the variation of relevant variables such as energy consumption, input quantities, and outputs. In this way, the user can execute various scenarios and sensitivity analyses to investigate the effects of changes to the input parameters.

**Development of the simplified meta-model** Based on the literature review described in Chap. 1, we analyzed various existing LCAs concerning the technologies, feedstocks, and process steps described. The relevant processes were determined based on these LCAs and the described models. The goal was to identify the intersections between the different models and determine generally applicable processes for the various pathways. Carbonation itself and the necessary pretreatment are the main processes at the center, accompanied by feedstock supply and $CO_2$-supply-processes on the input side and the post-processing on the output side. A possible utilization of end-products is not covered by the proposed meta-model because no meaningful generalising model assumption can be made at this point due to the large number of usage options. The next step was to implement this model in the LCA-software Umberto to carry out the LCA for various carbon mineralization pathways.

## 4 Life Cycle Assessment

LCA is a method used to comprehensively analyze a product's or technology's environmental impact over its life cycle. The ISO 14040 [12] and ISO 14044 [13] standardize and describe the procedure and structure of an LCA. Despite being standardized, LCAs in the field of carbon mineralization are challenging to compare, as critical factors such as the functional unit, system boundaries, and individual processes can be selected differently. For this reason, a guideline for implementing LCAs for carbon capture was used as a basis for this study [14].

### *4.1 Goal and Scope*

The main objective of this LCA is to quantify and compare the environmental impact of different carbon mineralization processes with the developed meta-model. Therefore, testing the meta-model with actual data is another study objective. The assessment concentrates on the carbon foot-print caused by the different technologies. Such a comparison helps drive the development and implementation of environmentally friendly carbon capture technologies and, thus, significantly contribute to reducing global $CO_2$ emission. Carbon mineralization is an approach for permanently storing carbon by reacting and binding carbon dioxide in a stable mineral form [15]. For this reason, we defined the functional unit in this study as 1t of $CO_2$ bound by a carbonization process. This choice provides a direct and comparable benchmark for assessing different carbon capture technologies and processes. With this definition of the functional unit, we can cross-check the $CO_2$ emissions from all activities during the mineralization. If the emissions are less than 1t of $CO_2$ eq per tonne of bound $CO_2$, we can state that the emissions are net-negative, i.e., $CO_2$ the process removes from the atmosphere.

The system boundaries of this LCA were defined to consider the direct environmental impacts of the carboniza- tion processes. It was decided not to include the use of potential end products that could result from carbonization processes in this analysis. This focus allows a more accurate assessment of the carbonization technologies' environmental impacts without being distracted by variable application contexts, use scenarios and substitutions. The $CO_2$ supply was included in the system boundary, as the process of $CO_2$ supply can require a huge amount of energy and is, therefore, a decisive factor in determining the environmental impact of carbon mineralization. In this study we considered direct air capture and point sources, such as industrial facilities, as $CO_2$ sources.

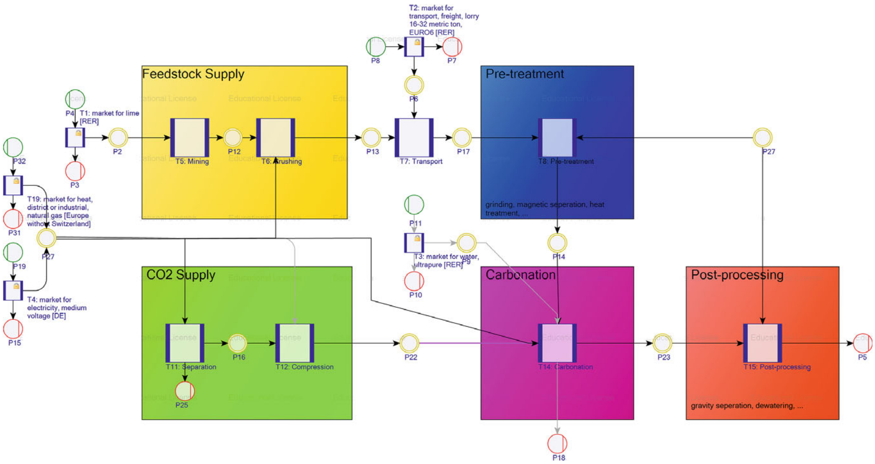



**Fig. 2** Screenshot of our developed model for LCA in Umberto

Umberto was used to implement the developed meta-model and calculate the environmental impact. In Umberto, the relevant process parameters and resource inputs were mapped in the model. Figure 2 shows the implemented Umberto model. To simplify the modeling in Umberto, we made the following assumptions:

- Identical transport (60km by lorry) for the feedstocks was assumed in all pathways
- Electricity generation based on the German electricity mix
- For comparison, label certified electricity from Switzerland with renewable energies
- Heat from natural gas
- Non-existent extraction of olivine and serpentine in Umberto was replaced by a comparable process of limestone extraction

## 4.2 *Life Cycle Inventory and Data Situation*

We had to obtain reliable data for different carbon mineralization pathways for the analyses on the developed meta-model. Own measurements or actual data from a company were not available. For this reason, we analyzed the data from LCAs found in the literature review to identify relevant data. This data research results in a collection of data presented in several tables, sorted according to the processes described in the meta-model. For further investigation, we found data on various technology pathways (direct and indirect). It was also important to consider data for different feedstocks, because they have a considerable influence on the mass balances of inputs and outputs. Data was found for olivine, serpentine, and various waste materials such as steel slack [1, 16, 17].

In addition to the data collected, various libraries and markets were used in Umberto to enable realistic modeling of the energy requirements [18]: market for transport, freight, lorry 16-32 metric ton, EURO6 [RER]; market for lime [RER], market for electricity, medium voltage [DE]; market for electricity, medium voltage, label-certified [CH]; market for heat, district or industrial, natural gas [Europe without Switzerland]; market for water, ultrapure [RER] (only for serpentine ); market for blast furnace slag [GLO] (rotary packed bed pathway).

During the data research, we identified several issues and challenges:

**Up-to-date data:** Up-to-date data is essential to correctly map technological developments and trends to obtain meaningful results in the LCA. Our data research revealed that in some cases, only older data (for example [16]) is available for individual processes and that this data formed the basis for various other LCAs found in the literature review [compare 1, 17, 19]. We must critically question whether this data is still meaningful today and reflects the current state of the art.

**Accessibility of the data:** In our data research, we have encountered problems with restricted access to specific datasets or data not being published in full. Reasons for this could be data protection, commercial interests, or other legal and administrative reasons. Due to the limited data available, we could not guarantee completeness across all carbon mineralization pathways in our LCA. We could only map and analyze the processes and technologies for which data was available.

**Accuracy:** The accuracy of the data is crucial, as incorrect or inaccurate information can lead to false conclusions. That applies, in particular, to data on the inputs and outputs of the individual processes and on energy consumption. As we have not measured any data, we have to rely on third-party information for the data we use. Note that some of the data come from experiments under laboratory conditions [20, 21, 22, 25]. For this reason, we can only make limited statements about our results for industrial and scaled applications where other conditions may exist.

**Consistency:** Data consistency is a critical factor in comparing different studies, technologies, and locations. As already described, there is the problem that sometimes only limited data can be retrieved, or data is only available for individual processes. For this reason, we combined our data from different sources. This results in a loss of consistency, as data generated under different conditions and for different purposes is correlated and summarized.

## *4.3 Life Cycle Impact Assessment*

The created meta-model enables the modeling of different CCS methods but builds on a limited data situation. In order to obtain meaningful results with the meta-model, we attempted to perform the impact analyses with data that was as coherent as possible. For this reason, this impact assessment focused on modeling the five pathways described in [1]. We implemented the meta-model with the modeling software Umberto LCA+.

Carbonation occurs in the direct pathway with a continuously stirred tank reactor (CSTR) without any intermediate steps. The pre-treatment and carbonation conditions depend on the feedstock. Data for olivine and serpentine were available in the study. Olivine is mined and prepared by grinding and milling in the pre-treatment stage. In the subsequent carbonization, the pulverized olivine reacts with water and $CO_2$ from the $CO_2$ supply. After that, the results undergo further processing in the post-processing stage. The procedure with serpentine is the same except for the pre-treatment stage. Magnetic separation isolates the iron, and heat treatment is required. The pathway OlivineCSTR100 is based on the study by [23]. The pathway SerpentineCSTR115 described in [1] referenced results from [16].

These sources also investigate the direct process using a rotary-packed bed reactor (RPB). This RPB often uses steel slack as a feedstock. Steel slack is a waste product in various industries, requiring no additional extraction process. After grinding, it can react with $CO_2$. The RPB process offers several advantages, including using off-gas containing 15–20% $CO_2$ instead of pure $CO_2$. Furthermore, the waste product steel slack is utilized as feedstock, resulting in possible cost and energy savings [1, 24].

In addition to these direct concepts, [1] describe two indirect pathways examined in this study using the meta-model. These are the Nottingham pathway and the AA pathway. Serpentine is usually the feedstock for both of these pathways. Pre-treatment and post-processing in the AA pathway correspond to the direct concepts, while it takes intermediate steps in the carbonation. The serpentine reacts with ammonium sulfate in a solid-solid reaction, and the actual reaction with $CO_2$ follows afterward Romao (2012). In the Nottingham pathway, the actual carbonation also takes place in two steps: first, an aqueous extraction and then an aqueous carbonation. In the Nottingham process, the feedstock supply and the pre-treatment stage are identical to the direct processes, as serpentine is also used as feedstock here. One exception is heat treatment during pre-treatment, which is unnecessary for the Nottingham Pathway. During the carbonation, the serpentine initially reacts with ammonium bisulfate in an aqueous reaction to generate a magnesium-rich solution that reacts with $CO_2$ in the second stage and binds the $CO_2$. The described Nottingham Pathway in [1] is based on [20].

With the modeling, we aimed to analyze which process is responsible for how much of the $CO_2$ emissions. The model was calculated once with the German electricity mix and heat from natural gas and once with a green electricity mix. Due to the data availability, which usually only contains energy consumption for a single process, we analyzed only $CO_2$ equivalence. However, this does not mean that other influences are irrelevant; instead, there is insufficient data to provide further information on other aspects.

Figure 3 shows the results for the German electricity mix. The $CO_2$ equivalence is positive in two methods, implying that the process produces more $CO_2$ than is stored. In the Rotary packed bed pathway, the $CO_2$ equivalence for the feedstock supply is negative, as it uses a blast furnace slag here. Umberto LCA+ rewards further use of this waste product with a negative $CO_2$ equivalence. For the Nottingham Pathway and Rotary packed bed pathway, no energy is required for the $CO_2$ supply, as emissions from point sources, such as industrial facilities, are used here. We then calculated

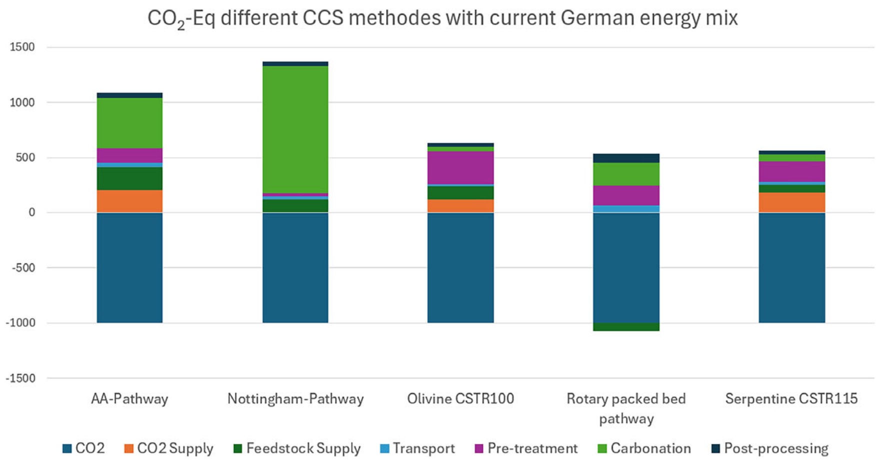



**Fig. 3** $CO_2$-eq for different mineralization pathways with German energy mix

all five models using green electricity. Due to the lack of a sustainable heat source in the Ecoinvent database, we also used green electricity as the heat source in these models. The results can be seen in Fig. 4.

All energy-intensive processes now have a significantly better $CO_2$ equivalence, and all methods store more $CO_2$ than they emit. On the other hand, the feedstock supply and transport processes still generate almost the same amount of $CO_2$ and are therefore responsible for a large proportion of $CO_2$ emissions. The latter happens because we include the pre-chains in these steps using data from Ecoinvent. We assume current rock extraction and transport conditions, not future conditions, that could decarbonize these steps. In all models, $CO_2$ is permanently bound in rock, i.e., long-term storage. We did not investigate other methods like producing e-fuels, which are burnt later in their lifecycle, emitting $CO_2$ again.

## 4.4 *Life Cycle Interpretation*

The results from diagrams Figs. 3 and 4 show that the energy-intensive processes are responsible for the $CO_2$-eq in particular. Using renewable electricity can avoid a large proportion of the $CO_2$ emissions caused by the storage of $CO_2$. Until this is the case, CSS only makes limited sense. The Nottingham pathway is particularly striking, as this process emits significantly more $CO_2$ than is stored under the current German electricity mix. That leads to the conclusion that CSS only makes sense if the electricity mix is entirely renewable.

It is worth noting that there is an issue with the data situation. For various process steps, data from different sources, including data obtained in laboratory situations,

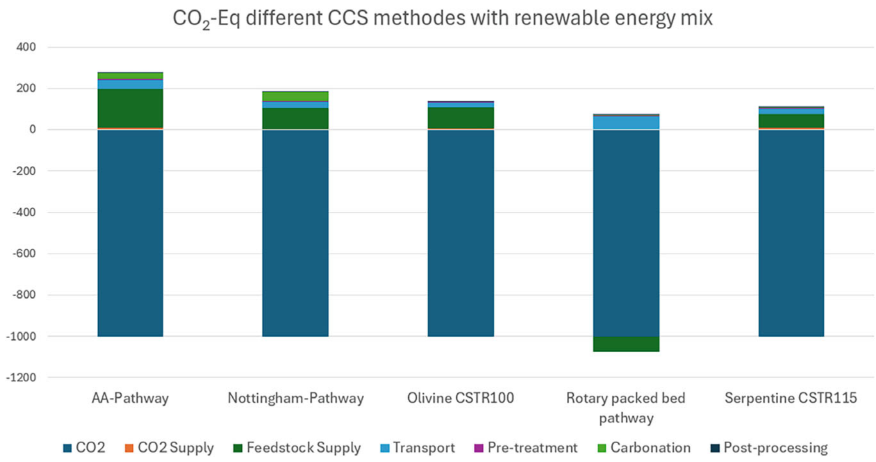



**Fig. 4** $CO_2$-eq for different mineralization pathways with renewable energy mix

were used in the analyzed studies. There needs to be empirical data on how energy use in large commercial systems will scale and develop.

# 5 Discussion

The LCA conducted in this study offers critical insights into the environmental impacts of CSS and CCU technologies. It underpins the urgent need for sustainable energy sources to power these technologies. Our findings highlight the challenges we must address to maximize the potential of CSS and CCU in effectively mitigating $CO_2$ emissions.

## 5.1 *Energy Intensity and the Need for Renewable Energy*

One of the most significant challenges highlighted by our analysis in Sect. 4.4 is the energy intensity of current CSS and CCU processes. The dependency on non-renewable energy sources not only undermines the overall carbon footprint reduction but also raises concerns about these technologies' sustainability and net environmental benefits. The transition to renewable energy sources is imperative to ensure that CSS and CCU technologies contribute positively to climate change mitigation efforts. This shift would align with the global push towards decarbonization and enhance the technologies' appeal from an environmental perspective.

## 5.2 Data Gaps and the Importance of Comprehensive Data Collection

As elaborated in Sect. 4.2, our study also reveals substantial gaps in the available data, particularly regarding up-to-date information on the energy consumption and environmental impacts of CSS and CCU processes. These gaps hinder the ability to make informed decisions and assess the technologies' viability and effectiveness. Therefore, there is a pressing need for standardized data collection methods and increased transparency in reporting to facilitate more robust and comprehensive LCAs. Our findings align with a consensus; most papers examined in the literature review share that collaboration among academia, industry, and regulatory bodies is essential to establish uniform data collection frameworks and databases.

## 5.3 Technological Innovation and Scalability

The tone of the examined papers in the detailed literature review also highlights the importance of technological innovation in improving the efficiency and scalability of CSS and CCU technologies. Advances in process optimization, material sciences, and system integration are critical to overcome current limitations and reducing costs. Furthermore, exploring novel $CO_2$ capture and conversion pathways could open up new avenues for carbon utilization, thereby expanding the potential applications and markets for CCU products. Continued investment in research and development is crucial to accelerating these innovations.

## 5.4 Policy Implications and the Role of Incentives

The findings from our study show a need for supportive regulatory frameworks and incentives to promote the adoption and development of CSS and CCU technologies. Policies aimed at internalizing the cost of carbon emissions, such as carbon pricing mechanisms, can enhance the economic viability of CSS and CCU. Additionally, targeted subsidies, tax incentives, and funding for research and development can accelerate further development.

## 6 Summary

This research report examines various studies on CCU and CSS. In addition to reviewing already completed literature analyses, we conducted a literature search for the years 2022–2024. Based on this, we created an overview of the prevailing technology landscape. This foundation enabled us to develop a meta-model and conduct life cycle assessments on various technologies. The LCA results show that the type of electricity used significantly affects the overall efficiency of the technologies, especially in reducing the carbon footprint.

We identified numerous challenges and points of criticism. These include incomplete or outdated data, methodological weaknesses, and a lack of research in certain areas. Nevertheless, the analysis shows that CSS and CCU could contribute to achieving the 1.5-degree target of the Paris Agreement. The paper underlines the importance of further research and development in this area to fully exploit the potential of CSS and CCU and achieve the goals of the Paris Agreement.

## 7 Outlook

Looking to the future of these technologies, it is clear that despite the progress made in developing CSS and CCU, there still needs to be more research. In particular, existing data gaps need to be closed to make informed decisions about implementing and scaling these technologies. A key focus of future research should be on improving the accuracy and completeness of available data. Achieving this demands increased collaboration between government agencies, research institutions, and industry partners to establish uniform standards for data collection and provide comprehensive data sets.

In addition, we need methodological improvements to enhance the robustness of life cycle assessments and better understand potential environmental, economic, and social impacts. New modeling and simulation approaches could help capture the systems' complexity and provide more accurate results.

Another important aspect is the research and development of cost- and energy-efficient $CO_2$ capture and utilization technologies. Novel materials, improved process technologies, and innovative business models could improve CSS and CCU's economic and ecological viability, thus enabling broader acceptance.

The future of CCS and CCU depends on the availability of high-quality data and continuous innovation.

## References


1. Ostovari, H., Sternberg, A., Bardow, A.: Rock "n" use of CO2: carbon footprint of carbon capture and utilization by mineralization. Sustain Energy Fuels **4**(9), 4482–4496 (2020). https://doi.org/10.1039/D0SE00190B
2. Cuéllar-Franca, R.M., Azapagic, A.: Carbon capture, storage and utilization technologies: a critical analysis and comparison of their life cycle environmental impacts. J. CO2 Utilizat. **9**, S. 82–102 (2015). https://doi.org/10.1016/j.jcou.2014.12.001
3. Da Cruz, T.T.: Life cycle assessment of carbon capture and storage/utilization: from current state to future research directions and opportunities. Int. J. Greenhouse Gas Control **108**, S. 103309 (2021). https://doi.org/10.1016/j.ijggc.2021.103309
4. Thonemann, N.: Environmental impacts of carbon capture and utilization by mineral carbonation: a systematic literature w and meta life cycle assessment. J. Cleaner Product. **332**, S. 130067 (2022). https://doi.org/10.1016/j.jclepro.2021.130067
5. Recker, J.: Scientific Research in Information Systems. Springer, Berlin (2013). https://doi.org/10.1007/978-3-642-30048-6
6. Fink, A.: Conducting Research Literature Reviews: From the Internet to Paper, 5th edn. SAGE, Los Angeles (2020)
7. Rahmanihanzaki, M., Hemmati, A.: A review of mineral carbonation by alkaline solidwaste. Int. J. Greenhouse Gas Control **121**, S. 103798 (2022). https://doi.org/10.1016/j.ijggc.2022.103798
8. Jeswani, H.K., Saharudin, D.M., Azapagic, A.: Environmental sustainability of negative emissions technologies: a review. Sustain. Product. Consumpt. **33**, 608–635 (2022). https://doi.org/10.1016/j.spc.2022.06.028
9. Kremer, D.: Separation of reaction products from ex-situ mineral carbonation and utilization as a substitute in cement, paper, and rubber applications. J. CO2 Utilizat. **62**, S. 102067 (2022). https://doi.org/10.1016/j.jcou.2022.102067
10. Gutsch, M., Leker, J.: Co-assessment of costs and environmental impacts for off-grid direct air carbon capture and storage systems. Commun. Engin. **3**(1), 1–14 (2024). https://doi.org/10.1038/s44172-023-00152-6
11. Gustafsson, M.: Climate performance of liquefied biomethane with carbon dioxide utilization or storage. Renew. Sustain. Energy Rev. **192**, S. 114239 (2024). https://doi.org/10.1016/j.rser.2023.114239
12. DIN EN ISO 14040:2021-02 Umweltmanagement - Ökobilanz - Grundsätze und Rahmenbedingungen, (2021a). https://doi.org/10.31030/3179655
13. DIN EN ISO 14044:2021-02 Umweltmanagement - Ökobilanz - Anforderungen und Anleitungen, (2021b). https://doi.org/10.31030/3179656
14. Müller, L.J. et al.: A guideline for life cycle assessment of carbon capture and utilization, (2020). https://doi.org/10.3389/fenrg.2020.00015
15. Stokreef, S. et al.: Mineral carbonation of ultramafic tailings: A review of reaction mechanisms and kinetics, industry case studies, and modellin, (2022). https://doi.org/10.1016/j.clet.2022.10049
16. Oconnor, W.: Aqueous mineral carbonation: mineral availability, pretreatment, reaction parametrics, and process studies. Last access: 2024-03-08 (2005). https://doi.org/10.13140/RG.2.2.23658.31684
17. Naraharisetti, P.K., Yeo, T.Y., Bu, J.: Factors influencing CO2 and energy penalties of CO2 mineralization processes. Chemphyschem: European J. Chem. Phys. Phys. Chem. 18(22), 3189–3202 (2017). https://doi.org/10.1002/cphc.201700565
18. Ecoinvent, ecoinvent Version 3.6, (2023). https://support.ecoinvent.org/ecoinvent-version-3.6
19. Kremer, D., Ex-situ mineral carbonation - a parameter study on carbon mineralisation in an autoclave as part of a large-scale utilisation process. J. CO2 Utilizat. **58**, S. 101928 (2022). https://doi.org/10.1016/j.jcou.2022.101928

20. Wang, X., Mercedes Maroto-Valer, M.: Integration of CO2 capture and mineral carbonation by using recyclable ammonium salts. ChemSusChem **4**(9), 1291–1300 (2011). https://doi.org/10.1002/cssc.201000441
21. Bodenan, F. et al.: Ex situ mineral carbonation for CO2 mitigation: Evaluation of mining waste resources, aqueous carbonation processability and life cycle assessment Carmex project, (2014). https://doi.org/10.1016/j.mineng.2014.01.011
22. Romao, I. et al.: CO2 fixation using magnesium silicate minerals. Part 2: Energy efficiency and integration with iron-and steelmaking, (2012). https://doi.org/10.1016/j.energy.2011.08.026
23. Eikeland, E.: Optimized carbonation of magnesium silicate mineral for CO2 storage. ACS Appl. Mater. Interfaces **7**(9), 5258–5264 (2015). https://doi.org/10.1021/am508432w.ss
24. Pan, S. et al.: High-gravity carbonation process for enhancing CO2 fixation and utilization exemplified by the steelmaking industry, (2015). https://doi.org/10.1021/acs.est.5b02210
25. Fabian, N. et al.: The influence of attrition milling on carbon dioxide sequestration on magnesium--iron silicate, (2010). https://doi.org/10.1016/j.mineng.2010.02.006

# Towards Sustainable Training Frameworks: Investigating Collaborative Learning in Virtual Reality

**Salsabil Lotfy Monib, Nada Nasser, and Nada Sharaf**

**Abstract** Virtual reality (VR) has gained widespread popularity and application across various domains, including industrial and organizational training. This study examines the efficacy of multi-user virtual reality (VR) technology in industrial training, focusing on the differential impacts of collaborative versus competitive VR environments. The study recruited undergraduate students from the German University in Cairo who participated in assembling a rapid cooling conveyor machine within these VR settings, supplemented by comprehensive pre- and post-training assessments. The findings reveal that while subjective measures of training effectiveness showed no significant differences between the two groups, competitive training participants were notably quicker in completing tasks during post-training evaluations. This suggests that competitive VR training may hold an advantage in scenarios where speed is critical. Both training approaches, however, significantly improved the participants' ability to perform the designated assembly tasks, underscoring the potential of VR as a transformative tool for industrial training. Despite the successes, the research was limited to a homogenous group of undergraduate students and required a simplification of VR scenarios to maintain network stability and user experience. Future research should extend to a more diverse participant pool and explore a wider array of VR tasks and settings to determine better the conditions under which collaborative or competitive training is most effective. Additional studies should consider multi-user setups involving more than two participants and include other training modalities like trainer-assisted scenarios. Longitudinal studies could further elucidate the long-term retention and application of skills acquired through VR training, enhancing the strategic use of VR in industrial and other professional training contexts.

S. L. Monib · N. Nasser (✉)
The German University in Cairo, New Cairo, Egypt
e-mail: nada.nasser@guc.edu.eg

S. L. Monib
e-mail: salsabil.bayoumi@student.guc.edu.eg

N. Sharaf
The German International University, New Administrative Capital, Egypt
e-mail: nada.hamed@giu-uni.de



V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_20

**Keywords** Sustainable learning · Collaborative learning · Virtual reality

## 1 Introduction

The rapid advancements in virtual reality (VR) technology have had a transformative impact across multiple sectors, especially in industrial training, where its ability to simulate complex machinery operations in a safe and controlled environment is highly valued. Traditional training methods in industrial settings are often constrained by high costs, safety risks, and the practical challenges of halting production lines for training purposes. Virtual reality offers a promising alternative, enabling immersive, risk-free environments where trainees can engage in hands-on learning without the associated hazards or disruptions to productivity. Moreover, VR training has the potential to significantly enhance the retention and speed of learning complex skills, making it an increasingly vital tool in industries where precision and efficiency are critical. Training through the use of 3D models and visual systems has consistently proven to be highly effective in general in different educational fields [7, 8].

The integration of sustainability into VR training is another notable benefit. By reducing the need for physical resources such as training materials, machinery, and energy consumption during operational downtime, VR promotes a more eco-friendly and cost-effective approach to workforce development. Additionally, reusable VR environments allow for continuous, scalable training without the environmental impact of traditional methods, aligning industrial training with broader sustainability goals. This eco-conscious dimension adds an important layer to VR's value proposition, positioning it as an efficient training tool and a sustainable one that minimizes the ecological footprint of industrial learning.

Despite the recognized benefits of multi-user VR training, there remains a significant gap in understanding how best to implement it-particularly when it comes to choosing between collaborative and competitive training modalities. Collaborative VR training, where participants work together to solve problems and complete tasks, has been shown to improve communication and teamwork skills, which are crucial in industries that rely on coordinated efforts. Conversely, competitive VR training introduces an element of rivalry, heightening engagement and motivation but potentially undermining cooperation and team cohesion. This study examined the effectiveness of both modalities in an industrial context and found that while participants in competitive VR training completed tasks more quickly, collaborative training fostered better teamwork and problem-solving skills. Both methods proved effective in skill development, but the choice of modality should be aligned with specific training goals, whether improving speed and efficiency or fostering collaboration.

This research highlights how VR training can be leveraged to meet productivity and sustainability goals in modern industries. It demonstrates that multi-user VR training facilitates effective skill acquisition and supports sustainable practices by reducing material waste and operational disruptions. Future studies can expand on these findings by exploring larger group dynamics, more diverse participant pools, and longer-term effects on skill retention. Ultimately, this work contributes to the development of optimized VR training programs that align with the evolving needs of industries focused on both efficiency and sustainability, offering practical guidelines for the broader implementation of VR in industrial training.

## 2 Virtual Reality in Industrial Training

In industrial settings, Virtual Reality provides a controlled and safe environment for training employees, particularly in high-risk industries such as manufacturing, construction, and emergency response. Virtual Reality simulations allow trainees to practice and hone their skills without the associated risks of real-world environments. For instance, Virtual Reality can simulate dangerous scenarios such as machinery operation or emergency responses, enabling trainees to learn and make mistakes in a safe space [6].

Research has shown that Virtual Reality training can improve learning outcomes by providing repetitive practice opportunities that are engaging and effective. For example, Virtual Reality assembly training has been found to enhance assembly time and reduce error rates compared to traditional training methods [3]. Additionally, Virtual Reality training can significantly reduce costs and risks associated with real-world training by providing a replicable, mistake-tolerant environment [5]. Furthermore, Virtual Reality is used to train workers in safety and risk assessment, which is crucial in high-risk occupations such as construction and emergency services [9].

A study on the use of Virtual Reality in industrial training demonstrated that Virtual Reality environments could simulate complex machinery and assembly tasks, allowing trainees to interact with virtual objects as they would with physical ones. This hands-on practice is essential for mastering intricate procedures without the risk of damaging expensive equipment or causing injury [3]. The study also highlighted that Virtual Reality training could be customized to replicate specific workplace conditions, providing a more relevant and practical learning experience [3].

Further, Virtual Reality technology has proven beneficial in enhancing training programs by integrating real-time feedback and performance tracking. Trainers can monitor trainee progress and provide instant feedback, ensuring that skills are correctly learned and applied. This capability not only enhances the learning curve but also helps in identifying and addressing knowledge gaps more effectively [6].

Moreover, Virtual Reality training modules can be continuously updated to include the latest industry practices and standards. This adaptability ensures that training programs remain current and relevant, providing employees with up-to-date skills and knowledge. For instance, in the automotive industry, Virtual Reality can be used to train employees on new car models and assembly techniques before they are introduced on the production line, thus reducing the time and cost associated with training [1].

Collaborative Virtual Reality training leverages the networked capabilities of Virtual Reality systems to allow multiple users to interact in a shared virtual space. This is particularly beneficial for tasks requiring teamwork and coordination, such as emergency response, medical procedures, and military operations. Collaborative Virtual Reality enables real-time communication and interaction among trainees, fostering a more interactive and supportive learning environment [10].

Studies have shown that collaborative Virtual Reality training can improve team performance and communication skills. For instance, Virtual Reality environments can simulate realistic scenarios that require trainees to work together to solve problems or complete tasks, enhancing their ability to collaborate effectively in real-world situations [2]. Moreover, collaborative Virtual Reality training provides a platform for distributed learning, where trainees from different locations can train together, thus promoting a more inclusive and extensive training program [1].

One notable example is the use of collaborative Virtual Reality for medical training. Medical professionals can practice complex surgical procedures together in a virtual operating room, communicating and coordinating their actions just as they would in a real surgical setting. This type of training is invaluable for improving team dynamics and ensuring that each member understands their role and responsibilities [5]. Furthermore, collaborative Virtual Reality can be used for disaster response training, where first responders from different regions can work together in a simulated disaster scenario, improving their coordination and response times [10].

The integration of Virtual Reality with other technologies, such as artificial intelligence (AI) and machine learning, further enhances collaborative training. Artificial Intelligence can provide adaptive learning experiences, tailoring scenarios to the specific needs and skill levels of trainees. This personalization ensures that each trainee receives the most relevant and challenging training, optimizing their learning experience. Machine learning algorithms can analyze trainee performance data to identify patterns and suggest improvements, making the training process more efficient and effective [9].

In the realm of industrial training, collaborative Virtual Reality has been used to train teams on complex assembly processes. For example, in the aerospace industry, Virtual Reality simulations can replicate the intricate assembly of aircraft components, allowing teams to practice and perfect their techniques in a virtual environment before working on actual aircraft. This approach not only improves team coordination but also reduces the risk of errors and rework, leading to significant cost savings [3].

Additionally, collaborative Virtual Reality training can play a crucial role in onboarding new employees. By immersing new hires in a virtual environment that replicates their future workplace, companies can provide a realistic and engaging introduction to their roles. This method has been shown to improve retention rates and accelerate the learning curve, as new employees can familiarize themselves with their tasks and the company culture in a risk-free setting [10].

## 3 CoLabXR

Upon launching the application, each participant's avatar representing him in the virtual environment, is automatically spawned into the designated Virtual Reality room. The initial few seconds post-launch are crucial for establishing a stable network connection. During this time, participants are directed toward a welcome panel in the virtual environment, which displays their connection status. Initially, this panel shows a "Connecting... "message, which, upon successful connection to the network, changes to "Joined ". This status change confirms that the participant is successfully integrated into the Virtual Reality training environment's network, ensuring they are ready to interact within the shared virtual space.

Once connected, participants are instructed to engage in a simple interactive task to confirm the presence and visibility of each other within the virtual space. They take turns waving at one another. This activity serves not only as a check for avatar visibility but also tests the real-time interaction capabilities of the network, ensuring that actions performed by one participant are immediately visible to the other, thereby verifying the system's responsiveness and synchronization. Following the visual confirmation, participants are asked to test the functionality of their Virtual Reality controllers. This involves performing various actions using the controllers, such as pointing, grabbing, and teleportation, which are crucial for the forthcoming assembly tasks. Participants confirm that the controllers are fully operational and that they can comfortably perform all required actions. This step is vital to ensure that hardware issues do not impede the training process. The last step before commencing the main training activity involves a direct interactive test between the participants. They are prompted to perform a virtual handshake with each other. This action tests the precision and timing of the networked interaction, ensuring that physical movements are accurately replicated and synchronized in the virtual environment. This check is crucial for activities requiring fine motor coordination and spatial accuracy, which are key components of the subsequent machine assembly task.

Upon successful completion of all preliminary checks, one of the participants is instructed to initiate the training session by pressing the 'Start' button, which activates the session timer and marks the beginning of the machine assembly task. This final action ensures that all network and interaction verifications are complete, and both participants are fully prepared to engage in the collaborative or competitive Virtual Reality training scenarios. This comprehensive setup and verification process is designed to minimize technical disruptions and ensure a smooth, uninterrupted training experience, which is essential for the effective learning and performance assessment of the participants in a virtual training environment. For further understanding of the flow of the setup and verification process refer to Fig. 1.

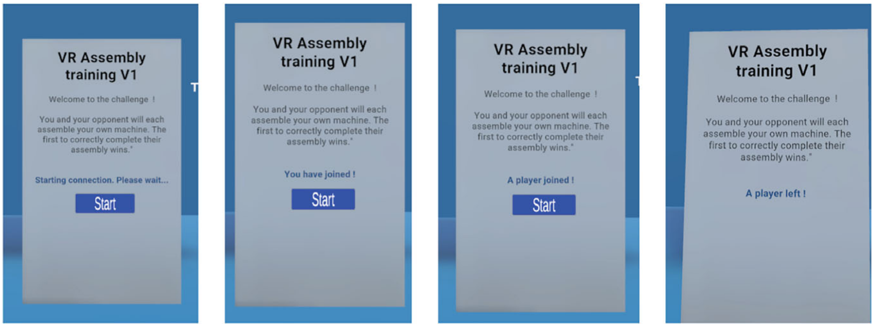



**Fig. 1** Setup and verification process

## *3.1 Collaborative Scene*

The collaborative scene consisted of an information panel that upon entry of the participants, gave them an introduction to the task. After checking the connection and testing their synchronization and calibration and their controls and ability to explore the scene making sure they were ready to start, one of the players should press the start button and the training session starts. When it starts they start seeing a synchronized timer counting up and a score counter for the number of pieces assembled so far, and they should work in collaboration by thinking and analyzing the machine and the pieces together and trying to finish the assembly as fast as they can. The scene the players see in the collaborative mode is displayed in Fig. 2.

**Fig. 2** Collaborative scene

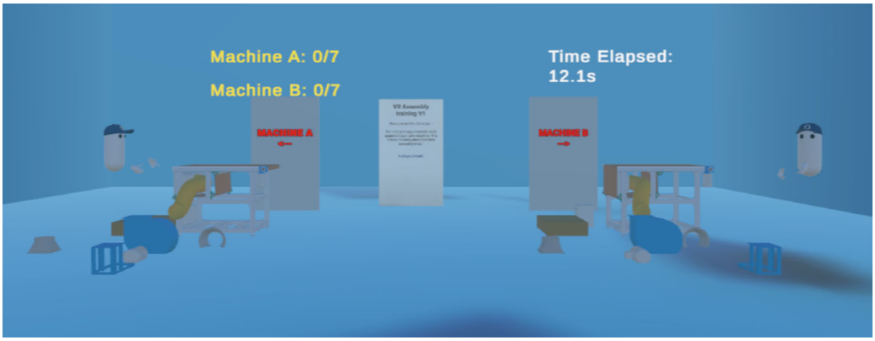



**Fig. 3** Competitive scene

## 3.2 Competitive Scene

The competitive scene was similar to the collaborative scene, in terms of having the same information panel and initial synchronization test process and having the same machine. However, in the competitive scene, there were 2 rapid cooling conveyor machines, one for each participant.

When the start button gets pressed and the training session starts, a synchronized timer counting up appears, and the scores of both participants are displayed to both of them. Each of the participants tries to finish his machine faster than the other one. Once the first participant completes the task, their finish time is displayed. However, the timer continues running as the second participant continues to try to finish their task. This is because the training experience requires both participants to complete the task, regardless of whether they win. The scene each player sees in the competitive mode is as displayed in Fig. 3.

## 3.3 Individual Scene

In addition to the collaborative and competitive modes, a third scene was implemented to serve as an individual post-training for each participant. This scene was crucial for evaluating the training effectiveness by measuring how quickly participants could complete the assembly task after undergoing collaborative or competitive training. The scene resembles the collaborative scene but with only one participant entering the training. The time taken to complete this post-training provided direct feedback on the efficiency and impact of the training received in the respective modes.

# 4 Experimental Design

The experimental design aims to investigate the effectiveness of collaborative and competitive virtual reality training modalities on learning outcomes in an industrial setting. Data were collected throughout the experiment, including machine assembly times during each training scenario and responses from the post-experiment surveys. Statistical analyses were then conducted to compare performance across training modalities and to correlate these outcomes with subjective measures from the questionnaires. The goal is to identify significant patterns that can inform best practices in multi-user training for industrial applications. Figure 4 illustrates the overall flow of the experimental design process.

The experiment involved 20 undergraduate students from the German University in Cairo, with varying skills relevant to the tasks involved in the training scenarios. The participants' demographic profile included 65% males and 35% females, all aged between 19 and 23 years. The participant's level of experience is presented in Fig. 5. This diverse group provided a solid basis for evaluating the training's impact on individuals with different levels of familiarity with technology.

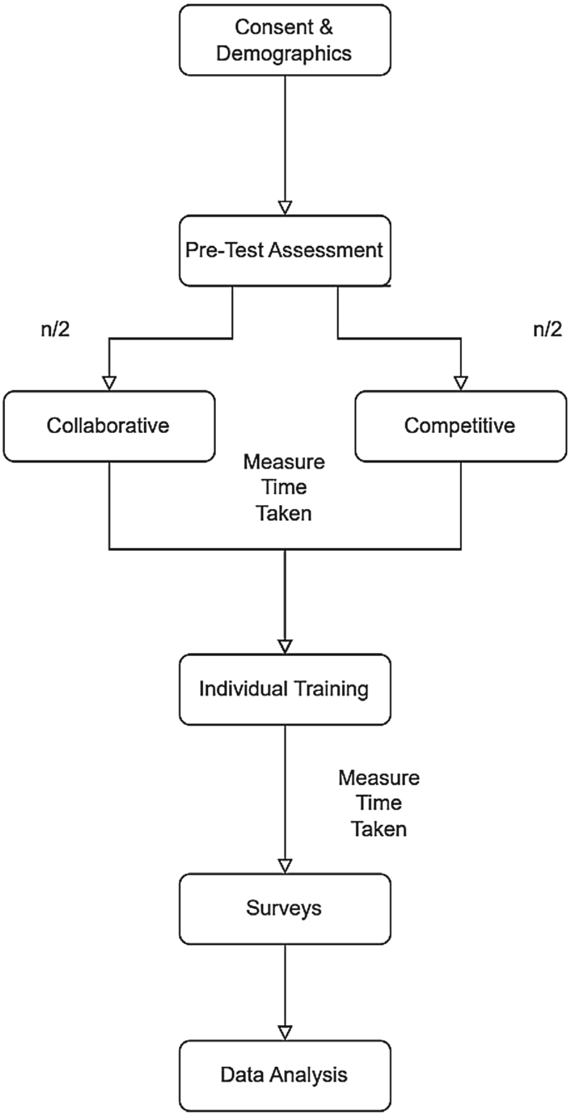
## 4.1 Pre-Training Assessment

Before the main training sessions, a pre-training was conducted individually with each participant to establish a baseline for their mechanical knowledge and familiarity with the equipment used in the experiment. Participants were presented with various component images from the machine they would be using and asked to match each part to its correct position. This task aimed to normalize the data for individual differences in initial knowledge and problem-solving ability, ensuring that any improvements observed could be attributed more directly to the training received. Figures 6 and 7 display the initial assessment given to the player before starting and the corresponding solution.

## 4.2 Training Sessions

Participants were divided into two groups of ten. One group underwent collaborative training, where pairs of participants worked together to complete the assembly tasks in a cooperative environment. The other group engaged in competitive training, where individuals competed against each other to see who could complete the same tasks faster and more accurately. Figure 8 shows the participants during the training session.

**Fig. 4** Flowchart of the experimental design process




## *4.3 Individual Training*

Following the group training sessions, all participants were asked to perform the assembly task individually. This session was designed to assess how well skills and knowledge from the earlier training were understood by each participant. The completion times were recorded to measure individual performance post-training. Figure 9 shows the participant during individual training.

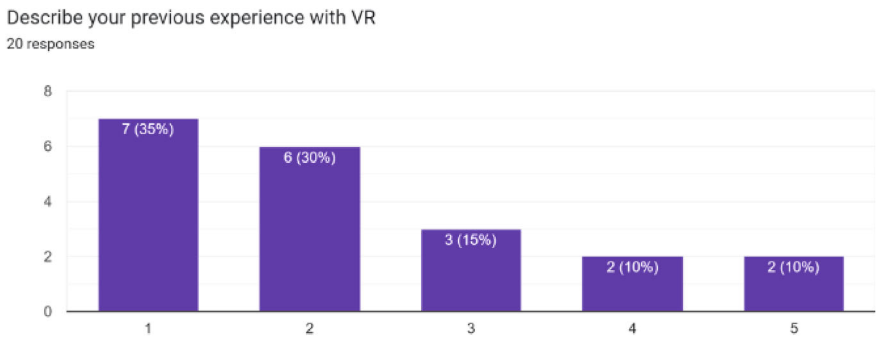



**Fig. 5** Participants' VR experience level

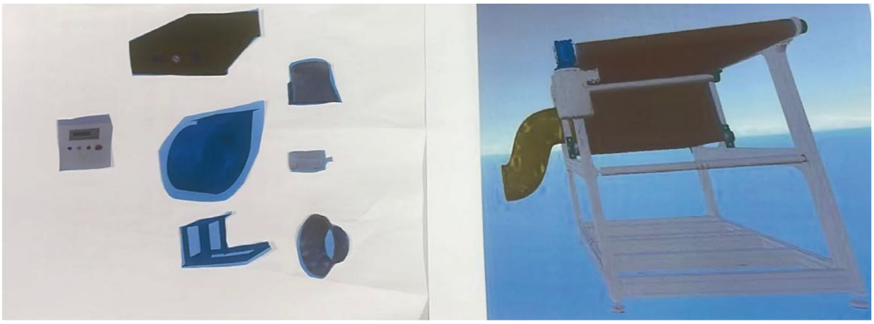


**Fig. 6** Pre-training assessment

## *4.4 Questionnaires and Surveys*

At the end of the training sessions, participants were required to complete several questionnaires designed to assess various aspects of their training experience. These questionnaires were:

- **Demographic Questionnaire**: Collected basic information about participants, including their details, educational background, and prior experience with. See Appendix B for the complete questionnaire.
- **NASA Task Workload Index**: Determined the mental, physical, and temporal demands placed on participants during the training. Check Appendix C for the full questionnaire.
- **System Usability Scale (SUS)**: Evaluated the usability and intuitive nature of the training system. Check Appendix D for the full questionnaire.
- **Telepresence & Immersion Questionnaire**: Measured the depth of immersion participants experienced, and how effectively the environment captured their attention. Check Appendix E for the full questionnaire.

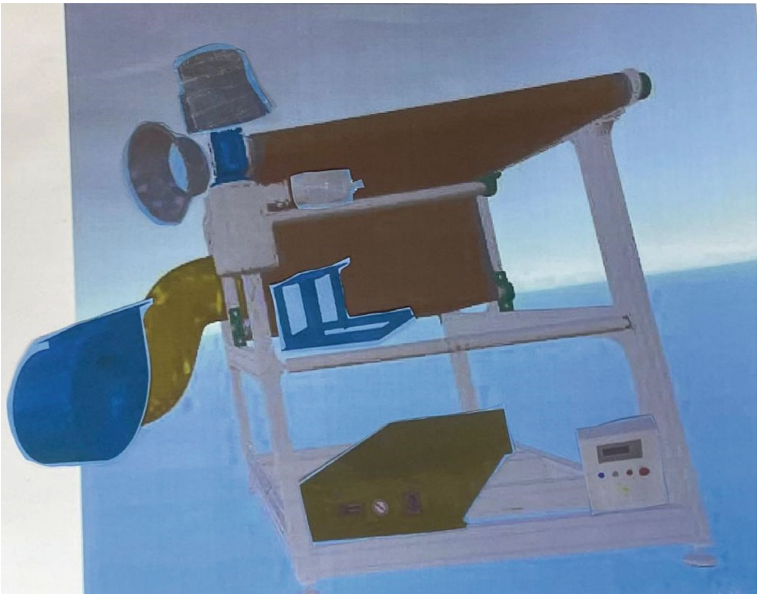


**Fig. 7** Sample pre-training assessment result

**Fig. 8** Sample multi-user training session


**Fig. 9** Sample individual post-training

- **Enjoyment & Engagement Questionnaire**: Investigated the level of enjoyment, engagement, and interest participants felt during the training. Check Appendix F for the full questionnaire.
- **Co-Presence Questionnaire**: Measured the participants' sense of being together with others in the virtual environment. Check Appendix G for the full questionnaire.
- **Social Presence Questionnaire**: Assessed how participants felt about the social interactions within the environment. Check Appendix H for the full questionnaire.
- **Test-Related Questionnaire**: To record participants' scores for subsequent evaluation and analysis. Check Appendix I.

These questionnaires were chosen to provide a comprehensive understanding of the training impact from both a technical and experiential standpoint. The data collected from these surveys was then analyzed to derive insights into how different training modalities influence learning effectiveness, engagement levels, workload management, and overall user experience.

# 5 Results

The pre-training assessment highlighted a uniform lack of familiarity with the task among all participants. Before undergoing the virtual reality training, participants demonstrated minimal knowledge and ability regarding the assembly process. The assessment results were strikingly similar across the board, with the majority of participants unable to correctly place any of the machine parts. A very small number managed to place at most one out of the seven parts correctly. This uniformity in performance indicated that all participants started the training with a comparable baseline, having little to no prior experience or understanding of the task at hand. This consistency in starting points ensures that the subsequent impacts of the VR training could be attributed more directly to the training itself rather than varying initial skill levels.

## *5.1 Co-presence*

Both the competitive and the collaborative groups had the same mean value. Similarly, their p-value ($p = 1.000$) implies that there is no statistically significant difference between the two groups in terms of co-presence. Both groups experienced a similar level of co-presence during the training sessions.

## *5.2 Engagement and Enjoyment*

The engagement and enjoyment questionnaire showed a p-value greater than 0.05 ($p = 0.247$) which shows that there is no statistically significant difference in engagement scores between the two groups. The slight mean difference is not enough to suggest a variance in engagement levels.

**Social Presence** The competitive group had a higher mean score of 61.50 compared to 58.00 for the collaborative group in social presence. Nonetheless, the two-sided p-value of 0.368 implies no statistically significant difference, indicating that both groups experienced a similar level of social presence.

**Workload** Participants in the competitive group reported a mean workload score of 19.00, slightly higher than the 18.50 reported by the collaborative group. The two-sided p-value of 0.663 reflects no statistically significant difference in the perceived workload between the two groups during the sessions.

**Immersion** With mean scores of 20.80 for the competitive group and 20.40 for the collaborative group, the immersion levels were nearly identical. The two-sided p-value of 0.749 suggests no significant difference in how immersed participants felt during the sessions.

**System Usability Scale** The SUS scores were 87.00 for the competitive group and 82.75 for the collaborative group. The two-sided p-value of 0.316 indicates no significant difference in system usability assessments by the participants across both training types.

**Training Duration** The p-value of the training duration was greater than 0.05 ($p=0.705$) which indicates that there is no statistically significant difference between the durations of the first training session of the competitive and collaborative groups.

**Post-training Test Duration** As the p-value of the post-training session duration yielded a value smaller than 0.05 ($p = 0.022$), a statistically significant difference in the second trial durations between the competitive and collaborative groups is indicated. This suggests that the competitive group completed their tasks significantly faster in the second trial than the collaborative group.

### *5.3 Duration Difference*

The p-value, greater than 0.05 ($p = 0.085$), suggests no statistically significant difference in the rate of change in durations between the first and second training sessions across both groups. Both groups showed improvements in the post-training, but the extent of improvement was not significantly different when comparing the two different types of training.

### *5.4 Summary of Findings*

The analysis revealed that throughout the training, no statistically significant differences were observed between the groups in terms of co-presence, engagement, social presence, workload, immersion, and system usability, suggesting that both training modalities were equally effective. However, during the post-training test, the competitive group completed tasks significantly faster than the collaborative group, highlighting a potential advantage of competitive training in enhancing task completion speed. Despite this, overall improvements from the first to the second training sessions were similar across both groups, indicating that both collaborative and competitive training methods effectively enhance performance over time.

During the training sessions, several behavioural patterns were drawn corresponding to the type of training. Participants in the competitive environment displayed more excitement and a higher level of engagement, driven by the challenges of independently managing the assembly tasks. Conversely, those in the collaborative setting exhibited more relaxed behaviors, with enhanced communication, particularly noticeable among participants who were friends or had previous interactions. This familiarity fostered more effective teamwork, although individual task remembrance was somewhat lower compared to the competitive group. Specifically, most participants in collaborative settings tended to remember parts they discussed and handled directly better than those managed by their partners. Despite the increased workload in competitive scenarios, the stimulating challenge appeared to promote task completion, suggesting that competitive settings might not only heighten learning but also increase efficiency under pressure.

## 6 Conclusions and Future Work

This paper investigated the impact of multi-user virtual reality (VR) technology on industrial training, with a particular focus on sustainability, by comparing collaborative and competitive VR environments. Participants, consisting of undergraduate students, were engaged in either collaborative or competitive training setups to assemble a rapid cooling conveyor machine. Pre- and post-training assessments were used to evaluate the effectiveness of each method in enhancing performance while considering sustainable training practices that minimize resource consumption and promote long-term efficiency.

The results demonstrated no significant differences in subjective measures between the collaborative and competitive groups. However, participants in the competitive setup completed tasks faster in post-training assessments, suggesting that competitive training may offer advantages in time-sensitive scenarios. Both methods, however, significantly improved participants' skills, confirming the effectiveness of VR training in an industrial context. Importantly, the VR-based training approach aligns with sustainability goals by reducing the need for physical resources and providing a reusable, scalable platform for skill development.

Future studies should include a more diverse participant base and explore a wider variety of tasks and virtual environments. Additionally, expanding VR training to larger, multi-user groups could provide deeper insights into collaborative dynamics and their impact on both learning and sustainability. Longitudinal research is also necessary to assess the long-term retention of skills and the broader sustainability benefits of VR training, such as reducing material waste and energy consumption. The use of mixed reality settings along with collaborative training scenarios should be also investigated [4]. These future directions should help refine training strategies for industrial use, ensuring both effective learning and a sustainable approach to workforce development.

## References

1. An, S.-B., Lim, K.-M., Go, H.-E., Jung, G.-Y., Ma, B.-C.: A study on development of multi-user training contents for response to chemical accidents based on virtual reality. J. Digit. Contents Soc. **21**(1), 1–10 (2020)
2. Bui, D.T., Barnett, T., Hoang, H., Chinthammit, W.: Tele-mentoring using augmented reality technology in healthcare: a systematic review. Australasian J. Educat. Technol. **37**(4), 68: Centre for Rural Health, School of Health Sciences, College of Health and Medicine.

University of Tasmania; Human Interface Technology Laboratory, School of Information and Communications Technology, University of Tasmania (2021)
3. Cecil, J., Gupta, A., Pirela-Cruz, M., Ramanathan, P.: A network-based virtual reality simulation training approach for orthopedic surgery. ACM Trans. Multimedia Comput., Commun., Appl. **14**(3) (2018)
4. Farouk, P., Faransawy, N., Sharaf, N.: Using hololens for remote collaboration in extended data visualization. In: Banissi, E., Ursyn, A., Bannatyne, M.W.M., Pires, J.M., Datia, N., Nazemi, K., Kovalerchuk, B., Andonie, R., Nakayama, M., Sciarrone, F., Huang, W., Nguyen, Q.V., Mabakane, M.S., Rusu, A., Temperini, M., Cvek, U., Trutschl, M., Müller, H., Siirtola, H., Woo, W.L., Francese, R., Rossano, V., Di Mascio, T., Bouali, F., Venturini, G., Kernbach, S., Malandrino, D., Zaccagnino, R., Zhang, J.J., Yang, X., Geroimenko, V., (eds.), 26th International Conference Information Visualisation, IV 2022, Vienna, Austria, July 19–22, 2022, pp. 209–214. IEEE (2022)
5. Lerner, D., Mohr, S., Schild, J., Göring, M., Luiz, T.: An immersive multi-user virtual reality for emergency simulation training: usability study. JMIR Serious Games **8**(3), e18822 (2020)
6. Philippe, S., Souchet, A.D., Lameras, P., Petridis, P., Caporal, J., Coldeboeuf, G., Duzan, H.: Multimodal teaching, learning and training in virtual reality: a review and case study. Virtual Reality Intell. Hardware **2**(5), 421–442 (2020)
7. Sharaf, N., Abdennadher, S., Frühwirth, T.W.: A rule-based approach for animating java algorithms. In: Banissi, E., McK, M., Bannatyne, W., Bouali, F., Burkhard, R., Counsell, J., Cvek, U., Eppler, M.J., Grinstein, G.G., Huang, W., Kernbach, S., Lin, C.-C., Lin, F., Marchese, F.T., Pun, C.-M., Sarfraz, M., Trutschl, M., Ursyn, A., Venturini, G., Wyeld, T.G., Zhang, J.J., (eds.), 20th International Conference Information Visualisation, IV 2016, Lisbon, Portugal, July 19–22, 2016, pp. 141–145. IEEE Computer Society (2016)
8. Sharaf, N., Ahmed, G., Ihab, S.: Virtual/mixed reality control of a game through scratch. In: Banissi, E., Khosrow-shahi, F., Ursyn, A., Bannatyne, M.W.M., Pires, J.M., Datia, N., Nazemi, K., Kovalerchuk, B., Counsell, J., Agapiou, A., Vrcelj, Z., Chau, H.-W., Li, M., Nagy, G., Laing, R., Francese, R., Sarfraz, M., Bouali, F., Venturini, G., Trutschl, M., Cvek, U., Müller, H., Nakayama, M., Temperini, M., Di Mascio, T., Sciarrone, F., Rossano, V., Dörner, R., Caruccio, L., Vitiello, A., Huang, W., Risi, M., Erra, U., Andonie, R., Ahmad, M.A., Figueiras, A., Cuzzocrea, A., Mabakane, M.S., (eds.), 24th International Conference on Information Visualisation, IV 2020, Melbourne, Australia, September 7–11, 2020, pp. 689–693. IEEE (2020)
9. Xie, B., Liu, H., Alghofaili, R., Zhang, Y., Jiang, Y., Lobo, F.D., Li, C., Li, W., Huang, H., Akdere, M., Mousas, C., Yu, L.-F.: A review on virtual reality skill training applications. Front. Virtual Reality 645153 (2021). Published: 30 April 2021
10. Yildiz, E., Møller, C., Melo, M., Bessa, M.: Designing collaborative and coordinated virtual reality training integrated with virtual and physical factories. In: Proceedings of the IEEE Conference on Industrial Production, Aalborg, Denmark, 2020. IEEE. Accessed: July 02, 2020

# Analytics for Environmental Sustainability II

# Mobile Application for Citizens for Information About Low Water and Drought

**Fabian Bressel, Frank Fuchs-Kittowski, Tim Seegert, Maximillian Deharde, Ruben Müller, Bernd Pfützner, Moritz Zemann, and Andreas Abecker**

**Abstract** Low water and drought have occurred more frequently in recent years and have negative effects on the economy, ecosystems, and people. As it is predicted that such events will occur more frequently and for longer periods in the future, it is necessary to sensitize the population to this issue and to raise awareness of the problem among the population. To meet this challenge, a mobile application was developed for citizens to inform them about low water and drought. This paper presents and discusses the requirements, the concept, the implementation, and the evaluation of this app.

**Keywords** Low water · Drought · Mobile app · Citizen

F. Bressel · F. Fuchs-Kittowski (✉) · T. Seegert · M. Deharde
Hochschule Für Technik Und Wirtschaft Berlin, Berlin, Germany
e-mail: Frank.Fuchs-Kittowski@htw-berlin.de

F. Bressel
e-mail: Fabian.Bressel@htw-berlin.de

T. Seegert
e-mail: Tim.Seegert@htw-berlin.de

M. Deharde
e-mail: Maximilian.Deharde@htw-berlin.de

R. Müller · B. Pfützner
Büro Für Angewandte Hydrologie GmbH, Berlin, Germany
e-mail: Ruben.Mueller@bah-berlin.de

B. Pfützner
e-mail: Bernd.Pfuetzner@bah-berlin.de

M. Zemann · A. Abecker
Disy Informationssysteme GmbH, Karlsruhe, Germany
e-mail: Moritz.Zemann@disy.net

A. Abecker
e-mail: Andreas.Abecker@disy.net



V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_21

# 1 Introduction

Critical drought and low water phases occurred in Germany in 2003, 2018, 2019, 2020, and 2022 [1, 2]. 2018 and 2022 were also the warmest summers in Germany since temperature records began and an increase in drought periods is expected in Germany [2, 3]. Overall, the last eight years have been the warmest years in Europe since records began in 1950 [4]. Periods of drought affect many different sectors such as agriculture and forestry, inland waterway transport, and large industrial consumers, but in severe cases also the public water supply [2]. One problem is that the population is not yet sufficiently aware of the issue. There is also a lack of information on the topic, which means that the population does not feel sufficiently informed and cannot form an awareness, although there is great interest in information on low water and drought among the population in many subject areas [5].

In order to overcome this problem, it is necessary that measures are taken by the government, but these must also be accepted by the population. To meet this challenge, a mobile application was developed for citizens to inform them about low water and drought. By informing and sensitizing people to the issues of low water and drought and by presenting the data in a generally understandable way, acceptance and understanding of the measures taken should be increased and the issue made comprehensible to the population. This paper presents and discusses the requirements, the concept, the implementation, and the evaluation of such a mobile app.

This paper is structured as follows: Section 2 presents the current state of research and what comparable applications exist in related areas. The requirements for an app for low water and drought are then presented and how these requirements were developed. Section 4 presents a concept for such an app based on the requirements. The following section explains the technical implementation of the concept and the difficulties encountered. Section 6 shows how the implemented app was evaluated and the results of the evaluation. Finally, an outlook on possible further functions is presented.

# 2 State of Research and Related Work

## *2.1 Low Water and Drought*

Drought and the low water levels that sometimes accompany it are a natural phenomenon. Drought describes water shortage or water scarcity in its various forms. Drought can be categorized into different types depending on its duration and effects.

- A meteorological drought is said to occur when there is below-average precipitation over a period of time. The period must be defined individually depending on the geographical area under consideration.

- An extended meteorological drought leads to an agricultural drought. A higher evapotranspiration compared to the precipitation supply leads over time to a decrease in plant-available water and thus to water stress of the vegetation with an accompanying decrease in yield. Depending on the storage capacity of the prevailing soils, geographical areas are susceptible to agricultural drought to varying degrees.
- A hydrological drought occurs when water supplies in surface waters or groundwater fall to below-average levels. A hydrological drought usually occurs with a longer time lag to meteorological and agricultural droughts and is defined for surface or underground catchment areas (river basins or aquifers).
- Economic drought refers to the effects of meteorological, agricultural, and hydrological drought on society, i.e. it compares the water demand of a society with the actual demand [2, 6, 7].

The frequency and duration of dry periods have increased in recent years and, according to forecasts, will continue to rise [2, 8]. Hydrological droughts already occurred in Germany in 2018, 2019, and 2022 [1, 9].

Human activities have significantly altered the water balance in the Berlin-Brandenburg region. Intensive groundwater extraction for drinking water, industry, agriculture, and energy has led to falling water tables, particularly in recharge areas. The lignite mining industry further disrupted groundwater levels, and post-mining water recovery has reduced surface water flow in rivers like the Spree. Extensive artificial drainage systems, which account for over 80% of the water network, have reduced groundwater recharge and increased surface runoff. Climate change exacerbates these issues with higher temperatures, longer droughts, and reduced rainfall, leading to greater evaporation and strain on already pressured water resources. Water management strategies are essential to stabilize the water balance and protect resources in the future [10].

To categorize the current situation with regard to low water in the state of Brandenburg, the Brandenburg Ministry of Agriculture, Environment and Climate Protection published the Landesniedrigwasserkonzept Brandenburg (Brandenburg State Low Water Concept). This presents the concept of the low water traffic light. For the monitoring of Brandenburg's river basins, 25 control gauges were defined with which the outflows and inflows from and into the river basins can be monitored. Based on the ecohydrological minimum discharge, two warning levels were defined for the gauges, a pre-warning level and a warning level. In addition, the calculation basis for the ecohydrological minimum discharge and the two threshold values was provided [11]. Therefore, only the seven-day average flow is relevant for determining whether there is low water at the current time, in comparison with the previously determined threshold values.

### 2.2 Awareness of Low Water and Drought

Personal involvement is a decisive factor for interest and the resulting awareness of an issue [12]. Because low humidity and drought are often perceived as seasonal phenomena of the summer months, they are associated with the high temperatures and low rainfall during these months. However, many people do not realize that drought can already be caused by low precipitation in spring, when only the lack of evapotranspiration prevents low water levels [2].

The fact that the reduction in groundwater and surface water reservoirs due to severe droughts such as 2018 greatly reduces the ability of the environment to mitigate another dry period is also often not considered in public perception. This defines the end of an excessively dry period with sufficient short-term precipitation, as this greatly reduces the impact on agriculture and general flora. The time it takes to replenish the natural and artificial water reservoirs can last for years and the replenishment can only take place at a reduced rate or not at all, especially in consecutive years with severe drought [13].

The basic awareness of the problem already exists in the population, but there is too little information available to meet the need for information that can be used to sensitize people to the issue [5].

### 2.3 Providing Knowledge and Information Through Mobile Apps

Mobile apps, i.e. software applications for mobile devices such as smartphones and tablets, have become part of everyday life and are increasingly being used to impart knowledge and information or for teaching and learning processes (mobile learning) [14]. Mobile apps are also already being used successfully to impart knowledge and information in the areas of climate protection and adaptation to climate change [15, 16]. The use of apps to impart knowledge has the advantage that most people always have their smartphone with them and can therefore always access the app. This also means that the app can always be used when time is available, allowing users to decide for themselves when and how much they want to engage with the app. The price of smartphones has fallen sharply in recent years, which means that almost everyone now owns one. This also makes it possible to experience the app together, as anyone with a smartphone can access the app and share content [17]. However, the use of apps also has disadvantages. For example, smartphones often only have a small display, which makes it difficult to display content. Tasks carried out on a smartphone are also perceived as more complex than on a desktop computer [18].

### 2.4 Comparable Applications

As far as the authors are aware, there are no apps available for the topics of low water and drought in Germany. However, there are many apps for the related topic of floods, such as ‘Meine Pegel’ (My Gauges), ‘HochwassergefahrST’ (Flood DangerST), or ‘Starkregen App’ (Heavy rain app). These usually visualize the potential flood risk of Germany or a federal state in the form of a flood risk map. They also make it possible to get an idea of the current flood situation by displaying current water levels (with alarm levels if necessary) for individual flood warning gauges and visualizing historical data as a hydrograph (e.g. water level and flow rate). Flood warnings for selected flood gauges are also displayed by the apps. The ‘Meine Pegel’ app also displays a threshold value for ‘mean low water’ for selected gauges. However, the use of the app for low water information is not advertised or described in any way.

The Brandenburg low water portal [11] is available as a website. This visualizes the 25 control gauges and provides historical data for the gauging stations. Threshold values for warning and advance warning levels as well as for the ecohydrological minimum discharge are specified for the gauging stations.

There are also apps that visualize other types of environmental data. In the field of air quality applications, the Federal Environment Agency has the ‘Luftqualität’ (Air quality) app and Plume Labs has ‘Plume Labs’. These apps visualize the topic very well through the coloring of the elements in the app and the general structure and make it possible to get a simple overview of the current situation without prior knowledge of the topic of air quality. The Federal Environment Agency’s app also provides further information on the topic of air quality to enable a better understanding.

## 3 Requirement Analysis

This section presents the requirements for an app for informing the public about low water and drought, which were developed on the basis of a literature analysis, workshops with stakeholders from public authorities, and an online survey of potential users.

### 3.1 Methodology

A literature review on the effects of low water and drought on various sectors was caried out to determine the requirements. Existing approaches and solutions for mobile apps to inform the population were also analyzed. An online survey was also conducted on citizens’ information needs regarding low water and drought [5]. A stakeholder survey was then conducted in workshops with representatives of the Brandenburg Ministry of Agriculture, Environment and Climate Protection (MLUK) and the Brandenburg State Office for the Environment (LfU) in order to develop further requirements. On this basis, a GUI (Graphical User Interface) concept and mock-ups for the mobile app as well as a concept for the app were developed. This concept was discussed and adapted with the stakeholders (MLUK, LfU, Low Water Steering Committee).

## 3.2 Objective and Target Group of the App

The app is aimed at all residents of the state of Brandenburg. The aim of the app is to sensitize and raise awareness of the problem of low water and drought among the population and, on this basis, to increase acceptance among the population for future, necessary official and institutional measures in the event of low water and drought.

## 3.3 Requirements

As part of the requirements analysis, many ideas and requirements for a mobile app were collected and developed. The most important requirements for the app, which formed the focal points in the development of the concept, are presented below.

**Presentation of Data.** The public should be shown current, historical, and forecast hydrological data that characterize low water and drought (water level and flow rate in rivers, water available to plants, etc.). By providing up-to-date data, it should be easy to recognize whether and to what extent low water and drought are present. The population should thus be able to assess the current situation and be sensitized. Historical data should make it possible to compare and categorize current data and information. Forecast data should also make it possible to sensitize the population to future developments so that they can draw conclusions for their own situation and actions.

**Visualization of the Data.** The available historical, current, and forecast hydrological data should be processed and visualized. This should be done in a way that is easy to understand, so that the data can be viewed, understood, and interpreted even without a deep understanding of the values and the underlying low water and drought problems. This should make it possible for any adult to use the app without prior knowledge. This is achieved through the use of familiar colors and familiar symbols.

**Simplification of the Topic.** The topic of low water and drought should be presented in an understandable and comprehensible way. Complex terminology should be replaced or explained using simple language so that the app can be used by anyone, regardless of prior knowledge. Understanding and comparability of the data: The hydrological data should be easy to compare and a basic understanding of the data presented should be created. This makes it possible for everyone to interpret the data and form their own opinion. The comparability of the data also enables people to make their own assessment of the current situation regarding low water and drought and should therefore increase the acceptance of measures.

**Availability of the App.** The app should reach the entire population. For this reason, the app should be multilingual. German and English should be included as a minimum. In addition, the app should be installable on smartphones with Android and iOS operating systems in order to reach as many mobile devices as possible.

Several ideas and requirements were discarded during the requirements analysis, partly due to hurdles during implementation (e.g. expert functions, messaging functions, games, etc.). Some of these requirements are to be implemented in the future, as described in Chapter 6.

## 4 Concept

The designed app consists of several interconnected parts (screens), which are described in more detail in the following sections.

### *4.1 Overview of the Screens*

The **map screen** displays all data in a geographically localized manner. Together with the coloring of the individual elements (Fig. 1), it is thus possible to get an idea of the current situation without looking at the individual values using the map. The map's customization options (Fig. 4) make it possible to view only the data that is of interest. The integration of the forecast values makes it possible to get an idea of the further course of the value. The option of displaying all values on the map for a different time period using the history and forecast settings (Fig. 3) helps to compare the current situation with values from past crises.

The **list screen** is suitable if you want to search for specific values or obtain information about several water levels in one water body. Here, all gauging stations and river basins are listed again.

The **detail screen** is intended for a closer look at a value and offers various options for visualizing the course of the value and linking it to other data in order to make the value more comparable and easier to understand.

The **favorites screen** makes it possible to quickly access the detailed view of favored elements, but also to get a quick overview via the detailed display within the favorites screen.

The **glossary screen** provides information about shown values, terminology, and how the model providing the data works.

The **chatbot screen** helps users with questions that are not included in the glossary as well as with more general explanations of the topic or the explanation of contexts.

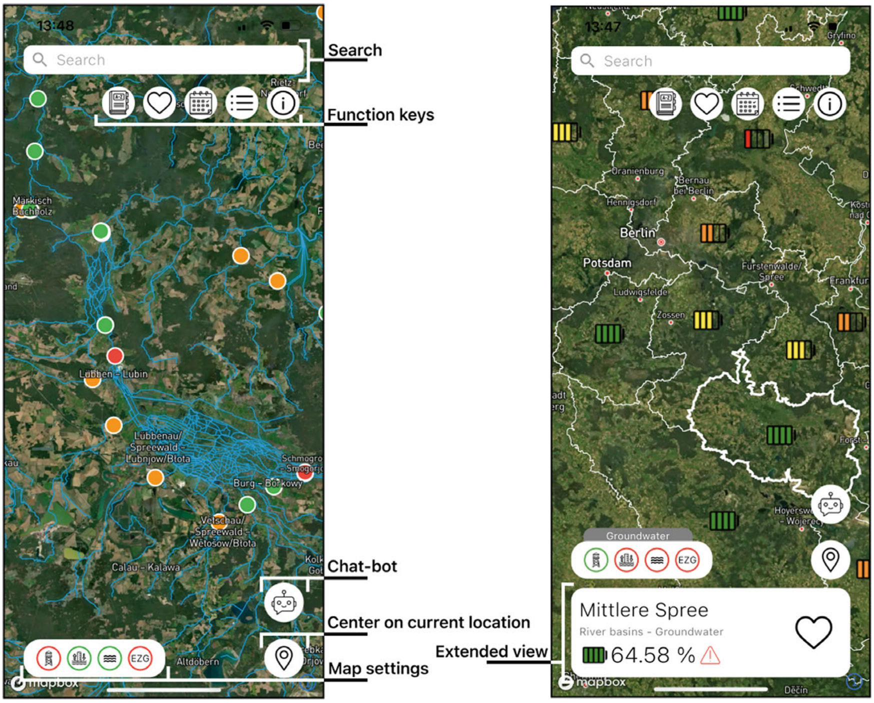



**Fig. 1** Map view without (left) and with (right) selection of a water reservoir

The settings screen can be used to customize the language and notifications and to view information on data protection and the NieTro_2 project.

## *4.2 Map Screen*

**Usage Scenario.** This view is the entry point to the app. All other views of the app can be accessed from here, with the exception of the settings view. Here you can search for specific elements (gauging stations, bodies of water, catchment area, or river basin), which are then displayed on the map. The list view and the favorites view can be accessed via the function buttons and the chatbot via the corresponding button. Below this button is also the button for centering the map on the user position. The map legend (Fig. 2) and the option to color the entire map according to a historical or forecast date (Fig. 3) can also be accessed from here. The map displays the layers defined in the map settings (Fig. 4) together with the associated data. The data on the map always corresponds to the current values of the elements, provided no changes have been made to the settings for history and forecast values. If changes have been made, the data at the selected time is displayed on the map. If an element is tapped on the map, the extended view of the element opens, allowing the name and current value to be viewed and the element to be favored in order to add it to the favorites (Sect. 4.5). In the expanded view, an arrow is also displayed, which indicates by its direction how the value will behave in the forecast next 14 days.

**Displayed.** The map view is largely taken up by the map. This always represents the base map. This contains a satellite image of Germany as well as designations for places and bodies of water. The level of detail of the designations depends on the zoom level and increases with the zoom level. Additional layers can be displayed above the base map. All layers except for river basins and catchment areas can be displayed together. For water bodies, these are highlighted on the map and are now interactive; the extended display for the water bodies can now be called up by tapping on them. In the case of the gauging stations layer, these are also shown on the map. Gauging stations are represented by a circle. The circle is colored according to the classification of the value of the gauging station according to the low water traffic light (see Fig. 1, left view). For river basins there are battery icons to display the filling level. The icon depends on the percentage of water currently available compared to

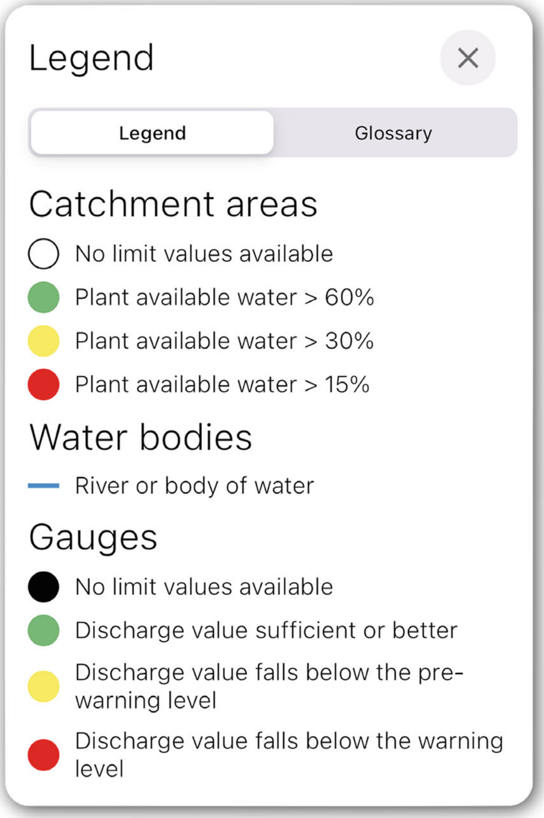



**Fig. 2** Legend of the map view

**Fig. 3** Time selection of the map view

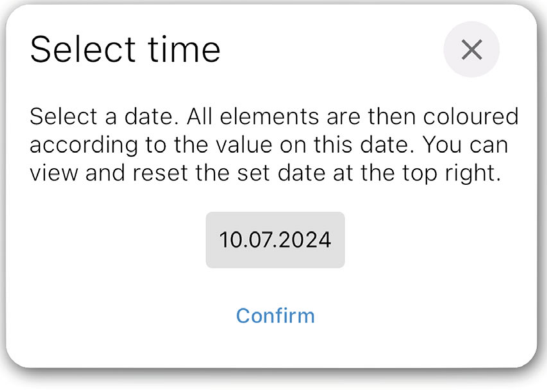
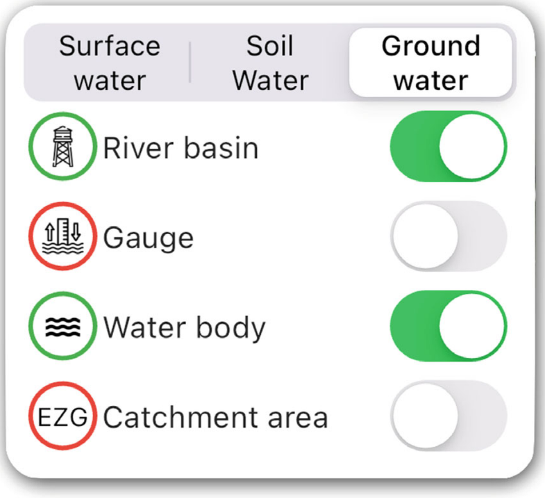



**Fig. 4** Map settings of the map view




the long-term average (see Fig. 1, right view). For the coloring of the catchment areas the low water traffic light colors are used again, but the color is determined by the usable field capacity. As can be seen in Fig. 1, the displayed elements of the map settings are outlined in green for a simple overview, the hidden ones in red. In the expanded view (see Fig. 1, right view bottom), the name and category of the tapped element is displayed together with the current value, its classification. For bodies of water is no classification available, since there are no reference values. An arrow is indicating whether the value is predicted to increase or decrease. The red warning symbol shows the message that the displayed value is a modeled value when tapped. The heart icon shows whether the element is favored, if this is the case it is red and filled in (Fig. 5).

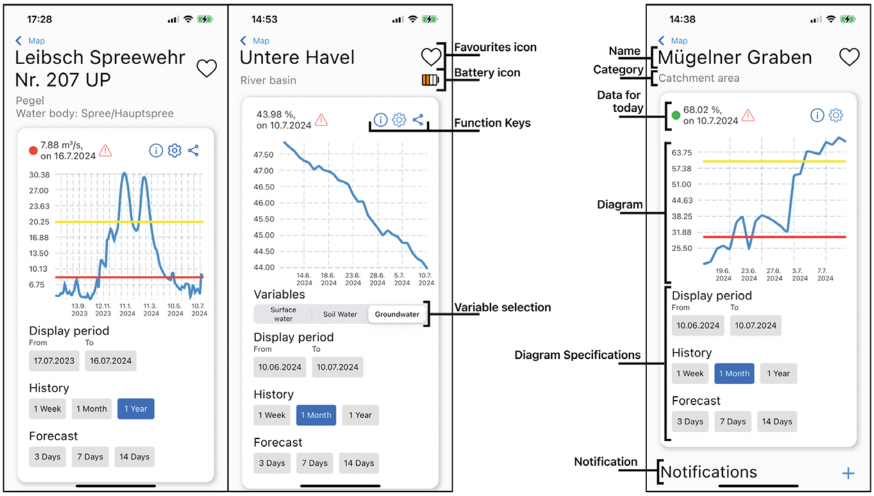



**Fig. 5** Detail screens with labeling; (l) gauge, (c) river basin, (r) catchment area

## 4.3 *Detail Screen*

**Usage Scenario.** All information on an element can be viewed in this view. The current value displayed is categorized by a colored or symbolic marker, with the exception of water bodies, since there are no limits here. The element can be added to the favorites by tapping the heart symbol. In this view, the current value of the element can be compared with historical values. The value can also be easily categorized by the warning and pre-warning marks in the diagram. The chart can also be customized using the chart specifications to determine the period for which the chart is to be created. At least 50 years of historical values and up to 14 days of forecast values can be displayed. The time period can be freely selected or selected by tapping one of the presets below the diagram. Additional data can be added to the diagram for better comparability and comprehensibility of the data. Weather data such as temperature or precipitation for the period can also be displayed in the diagram and hydrographs for the same time of year from other years can also be displayed. Furthermore, there is the option of displaying scenarios, with a red line for a very dry scenario displayed at the same time. The detailed view can also be shared with others, whereby they receive the name, current value, and an image of the diagram. Notifications can also be set for all elements if the value falls below a certain self-selected level. To help with the selection, the warning and pre-warning levels are also displayed when setting a notification.

**Displayed.** For each element, the name and the category are displayed as well as a heart, which symbolizes whether the element is favored. In the 'Data for today' area, the current value is displayed. In addition, the value is classified by the color coding, if available. Next to the current value is an info icon to display the legend, a cogwheel icon for the diagram settings, and a share icon to share the current view. When opening the detailed view for an element, the data for the last month is initially displayed. Notifications that have already been set are displayed in the Notifications area. When creating a notification, the warning and pre-warning levels are also displayed, to help with the selection of a fitting value (Figs. 6, 7)

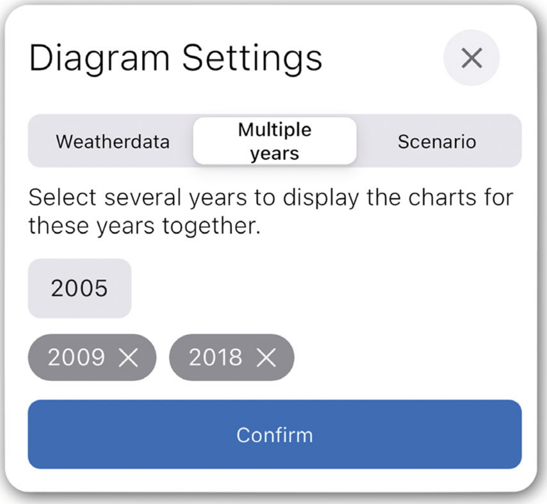



**Fig. 6** Settings for the diagram with option 'Multiple Years' chosen

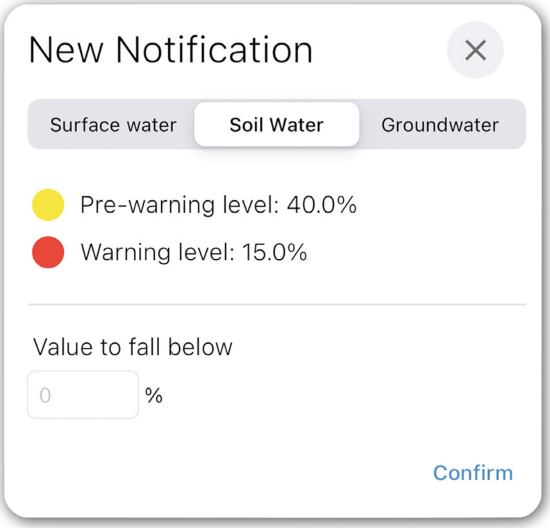



**Fig. 7** Pop-up to set a new notification for a river basin

## 4.4 Favorites Screen and List Screen

**Usage Scenario.** The favorites screen and list screen both provide a quick overview of elements in a list view. While the favorites screen displays more information for user chosen elements, the list screen displays all gauges and river basins to easily search through them. Tapping on one of the elements opens the detailed view. With the gear function key at the top left in the list screen, it is possible to change to the settings screen.

**Displayed.** The individual entries in the list screen only show the name of the element and a color coding of the current value for gauges and the fitting battery symbol for river basins. This allows as many elements as possible to be displayed clearly. The favorites screen is very similar to the list screen. However, the individual entries for the respective elements are much more detailed here than in the list view. The current value and its classification in the color scale as well as the forecast of the value can also be viewed here (Fig. 8).

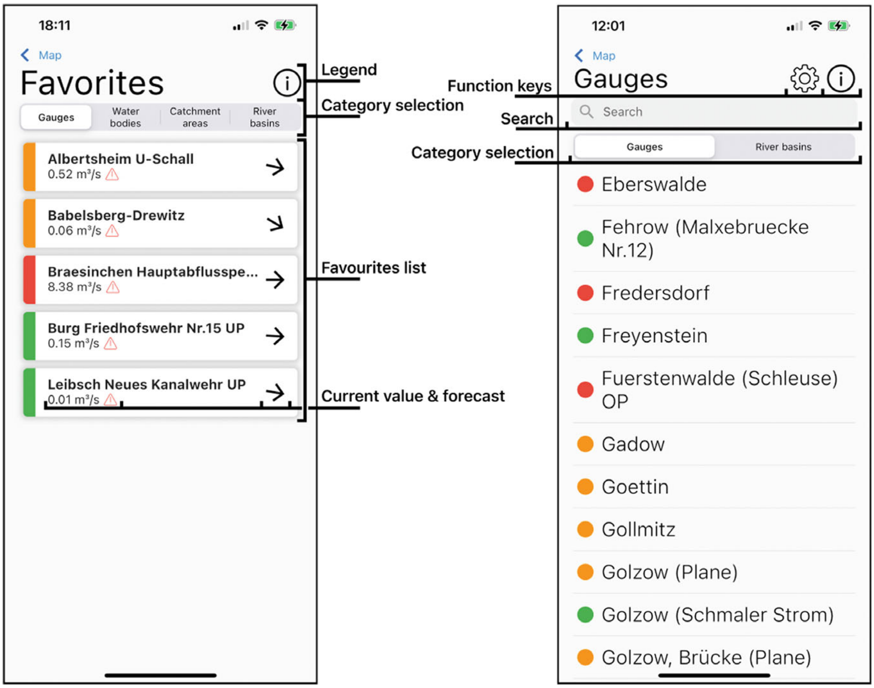



**Fig. 8** Favorites screen and list with labeling

## 4.5 Glossary Screen

**Usage Scenario.** The overview of all glossary entries (Fig. 9l) can be viewed. When the info icon is tapped, a pop-up appears showing the sources for the individual entries. The search bar can be used to search for the titles of all entries. The list of glossary entries is then filtered according to the search. If one of the entries is tapped, it is displayed in full length (Fig. 9r). If the tapped entry has subentries, these are also displayed.

**Displayed.** In the overview of all glossary entries (Fig. 9l), an info icon is displayed at the top for the sources. Below this is a search bar for searching all entries. In the list of entries, the main topics are displayed, with the corresponding subtopics indented below. The main topics are sorted alphabetically. If an entry has been selected (Fig. 9r), the title is displayed at the top and the associated text below. If there is an image for this entry, it is displayed at the end of the entry.

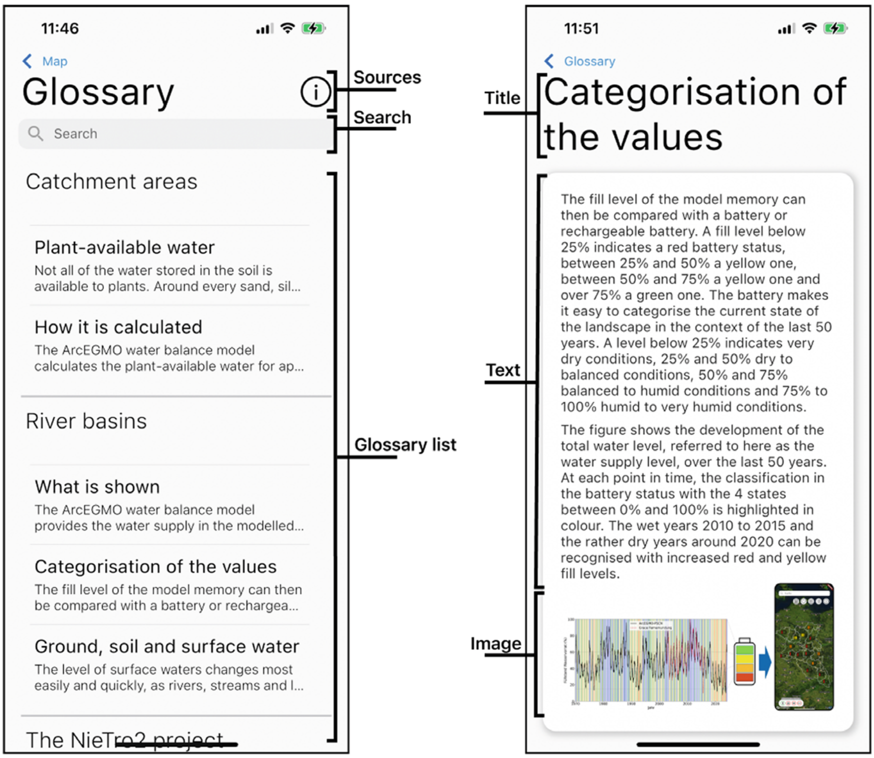



**Fig. 9** Glossary overview (l) and a glossary entry (r) with labeling

### 4.6 Chatbot Screen

**Usage Scenario.** The chatbot screen (Fig. 10) helps the user to answer questions in an interactive way. The chatbot only gives answers to low water and drought questions and tries to remain as scientific as possible. A chat can be conducted over several sessions and context from previous messages is used. If the user requests information about an element in the database, the chatbot is able to retrieve this information via actions and integrate it into the response. A new chat can be started by tapping the ‘New chat’ button. All previous chats can be called up with the ‘List all chats’ button and it is possible to switch between the chats.

**Displayed.** At the top right there are buttons to list all chats and to create a new chat. The chat history is visible below this. The user’s questions are in blue and the bot’s answers are in grey. The bot messages are also labeled with a robot icon and you can ask and send questions at the bottom of the screen.

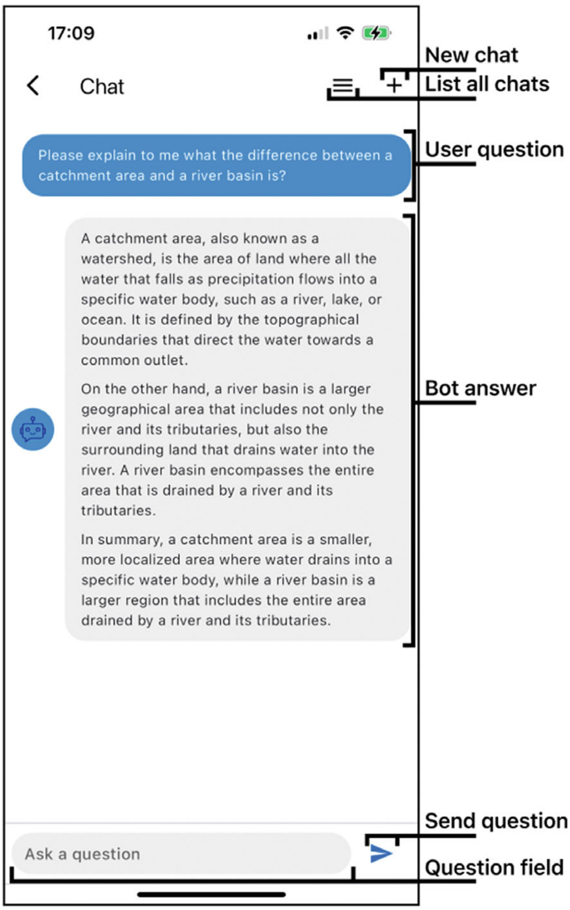



**Fig. 10** Chatbot screen with labeling

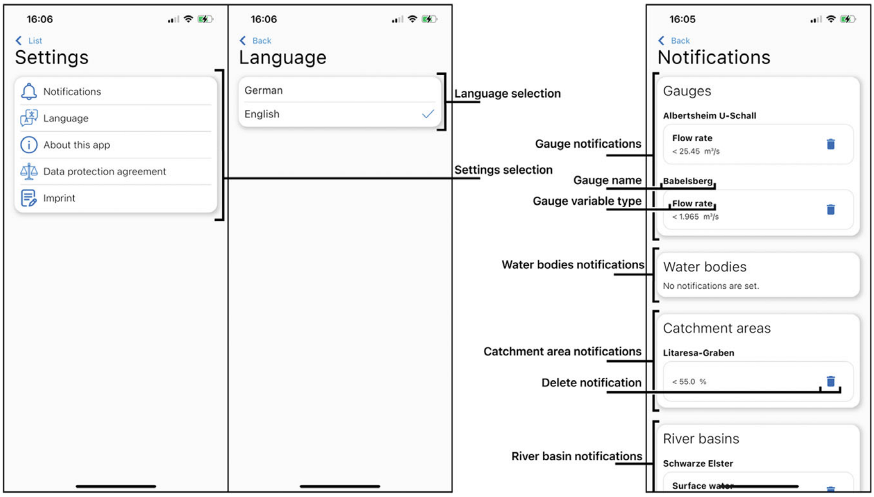



**Fig. 11** Settings screens with labeling: (l) settings overview, (c) language selection, (r) notification overview

### *4.7 Settings Screen*

**Usage Scenario.** The settings overview (Fig. 11l) is used to navigate to the various areas of the settings. In the language selection (Fig. 11c), the desired language for the app can be selected by tapping on ‘German’ or ‘English’. Under Notifications (Fig. 11r), all currently set notifications are displayed, ordered by the element type, and can be viewed. Tapping on the bin icon deletes the notification.

**Displayed.** The language settings are similar to the structure in Fig. 11, except that the two languages are offered there. The notifications are listed one below the other and can be edited or deleted by tapping on them. The data protection agreement and the ‘About us’ section only display text.

## 5 Implementation

The app was implemented using the Flutter framework (flutter.dev), which allows the creation of applications from a single codebase for nearly all common devices (smartphones, PCs, browsers). The necessary libraries were sourced through Flutter’s package management service, pub.dev (pub.dev).

The map is displayed using a ported version of Mapbox GL (github.com/mapbox/flutter- mapbox-gl) from JavaScript, with map data created and configured in Mapbox Studio. Mapbox GL supports the vector tile functionality required for this application and can generate maps from locally stored data, enabling offline use. One disadvantage of this library is that it can only be used on mobile devices with Android or iOS/

iPadOS. However, it is also possible to run the application on other devices using emulators, as the GUI adapts dynamically to the screen. For app localization, the 'Flutter Localization' library was used, selecting the appropriate language from a library of translated character sets. The fl_chart library (pub.dev/packages/fl chart) was utilized to create charts, known for its minimalist design and support for multiple chart types within a single diagram. Data storage is managed by the Isar database (isar.dev), specifically developed for Flutter and optimized for mobile devices. Comparisons show that Isar uses less storage and is faster than similar databases [19]. Given the app's offline capabilities, Isar was chosen for its ability to store large amounts of data in a compressed and quickly accessible format. The chatbot is based on OpenAI's cost-effective yet capable Large-Language-Model (LLM) GPT-4o-mini model (https://platform.openai.com/docs/models/gpt-4o), designed to understand complex topics and provide scientific responses specifically related to low water and drought. The model operates via a REST API, with the neural network running on OpenAI servers. Custom logic can be integrated into API calls, allowing the system to incorporate data from the Isar database based on user prompts, enhancing the relevance of its responses. For example, it can retrieve and display specific elements requested by the user. Additionally, the chatbot can access provided documents, offering information from them with proper sourcing.

The backend was realized with the Web-Framework Express (expressjs.com). It is used to make the required data available to the application and to store all the required data in the PostgreSQL (postgresql.org) database provided for this purpose.

The app communicates with the backend via a REST API. The backend provides interfaces that filter the data according to the desired criteria and only make the required data records available.

**Data Connection.** The Express.js backend obtains all the data required from the BAH Berlin hydrological model via the Disy data warehouse. The daily updated data, new forecasts, and scenario data are provided by BAH Berlin in a daily updated staging database, which is requested daily by the data warehouse.

**Visualization of the Data on a Map.** A combination of vector and raster tiles was used to display the map. Raster tiles are image files that form the map and must be pre-rendered for all zoom levels. They are ideal for displaying complex images like satellite views, which are used for the base map. Since the map is static, the same raster tiles can be reused. In contrast, vector tiles, which represent map elements such as place names, the water network, river basins, catchment areas, and gauges, use vectors that allow scalable tiles with less memory usage. Unlike raster tiles, vector tiles are generated at runtime, requiring more resources but enabling dynamic style changes. This combination allows the display of a satellite image with runtime color adjustments without needing a map server to recalculate the map data.

**Use Without an Internet Connection.** The app is designed for offline use, with map data around Brandenburg stored on the device. This includes raster tiles for satellite images and vector tiles for other map elements. Style file adjustments allow color and symbol changes. While real-time access to map elements requires an internet connection, all viewed data is stored in the Isar database and remains accessible offline. When the app starts, the entire map is loaded, caching data for all elements.

Detailed views trigger the retrieval and storage of forecast or historical data for specific elements, ensuring comparisons can be made. Only data for elements viewed in detail are saved, keeping the app size manageable. If storage becomes an issue, users can delete cached data in the settings.

## 6 Evaluation

For the evaluation, the application was evaluated with two external test groups according to the System Usability Score (SUS) [20]. This metric objectively assesses the user-friendliness of systems. As part of the evaluation, the testers were given a 10-min presentation with a quick overview on the content and functions of the app. Afterward, they were given 10 min to try out the application and complete set tasks, such as setting a notification or favoriting an element. These tasks served to familiarize the participants with the basic functions of the app. The participants then answered questions about the evaluation according to the SUS using an online form.

The result of an evaluation according to the SUS gives a value from zero to one hundred, with the value one hundred representing the best possible user-friendliness.

In the first evaluation with seven evaluators, a score of 81 points was achieved, which is good for the first evaluation of an application. The lowest number of points was achieved for the question 'I think that I would like to use the mobile app frequently'. This is because there should already be a prior interest in the low water and drought topic in order to be interested in the application, otherwise the content of the app may appear irrelevant and therefore there is no interest in using the app. During the evaluation some minor bugs appeared which were fixed before the second evaluation. In the second evaluation with nine participants, a score of 76 points was achieved. The fact that a lower score was achieved in the second evaluation despite improvements to the app shows how different people perceive user-friendliness. Overall, SUS scores ranging from 65 to 95 points were awarded by individual participants in the evaluation. The large differences in the results can be attributed to various things. For example, as already mentioned in the previous paragraph, some questions are also heavily dependent on the topic of the application and the participants' interest in it. However, Mclellan et al. also showed that users give a higher SUS rating if they are already familiar with the system [21]. Furthermore, Kortum et al. confirmed to some extent that the more often an application is used, the higher the SUS assigned by the users [22]. As the app is based on Apple's design guidelines [23], Android users may give lower scores as applications for Android are developed according to different design guidelines. As the answers were anonymized, no statement can be made afterward about the devices used during the evaluation.

## 7 Summary and Outlook

Low water and drought have occurred more frequently in recent years and have negative effects on the economy, ecosystems, and people. As it is predicted that such events will occur more frequently and for longer periods in the future, it is necessary to sensitize the population to this issue and raise awareness of the problem among the population. To meet this challenge, a mobile application was developed for citizens to provide information about low water and drought. The aim of the app is to sensitize and raise awareness of the issue of low water and drought among the population and to increase acceptance of the necessary future measures. This goal is to be achieved by processing and visualizing hydrological data on the current, historical, and predicted low water situation and presenting it in an understandable and comparable way. To fulfill these requirements, the app displays historical, current, and forecast data on water levels; flow rates; plant-available water; and the amount of water available in a river basin. This data is displayed on a map, with data on water bodies, gauging stations, catchment areas, and river basins available for retrieval. The app was implemented using the Flutter framework and the map display was realized using a Mapbox library. The data for the app is provided by an Express.js backend server. Beforehand, the data is determined by the hydrological model ArcEGMO-PSCN and imported into the backend server. The app is to be expanded in terms of content with a guide, with data to further enhance the diagrams and additional interfaces for the chatbot so that it can provide better support:

- The app should be expanded to include a guide to help people deal with problems caused by drought or low water levels and to help them reduce their own water consumption if necessary.
- Further data such as precipitation and temperature can improve our understanding of the development of the values. But other low water-related data such as the usable field capacity at a depth of 1.5 m can also help to better understand low water and drought and the associated problems.
- By setting up additional suitable APIs to the database, weather data, and other information portals on low water and drought, real-time information is to be made available to the LLM so that it can provide the best possible support to users. Furthermore, the chatbot should be provided with current literature on low water and drought. In this way, information on current topics can be retrieved and the data from the hydrological model can be analyzed and information about it presented.

Furthermore, it should be evaluated whether the application could be used for other federal states in Germany or even other countries. The BAH already operates a model for Saxony-Anhalt, which only needs to be calibrated for low water. The architecture of the application would easily allow further data to be fed in if it is available; at present, the Brandenburg water balance model is the only one in Germany that is calibrated for low water. At the time of publication, the range of functions described in this article will have been implemented. At this point, the app will be tested and evaluated by several partners. In addition, any changes and necessary optimizations identified in the evaluation will be implemented. On this basis, a stable version of the app for various mobile platforms will be available in October 2024, which will then be published in the corresponding app stores.

**Acknowledgements** The authors would like to thank the Federal Ministry of Transport and Digital Infrastructure (BMVI) and its mFUND research initiative for financial support as part of the 'NieTro_2' research project (funding code 19FS2018A/B/C).

## References

1. Marx, A.: Dürren in Deutschland (2023). https://www.ufz.de/index.php?de=47252. Accessed on 24 June 2024
2. Riedel, T., Nolte, C., aus der Beek, T., Liedtke, J., Sures, B., Grabner, D.: Niedrigwasser, Dürre und Grundwasserneubildung – Bestandsaufnahme zur gegenwärtigen Situation in Deutschland, den Klimaprojektionen und den existierenden Maßnahmen und Strategien (2021). https://www.umweltbundesamt.de/publikationen/niedrigwasser-duerre-grundwasserneubildung. Accessed on 24 June 2024
3. Deutscher Wetterdienst: Deutschlandwetter im Jahr 2022 (2022). https://www.dwd.de/DE/presse/pressemitteilungen/DE/2022/20221230deutschlandwetter jahr2022 news.html?nn=495078. Accessed on 24 June 2024
4. World Meteorological Organisation: Past eight years confirmed to be the eight warmest on record (2023). https://wmo.int/news/media-centre/past-eight-years-confirmed-be-eight-warmest-record. Accessed on 24 June 2024
5. Wein, E., Bressel, F., Berndt, J., Siewert, H., Bähne, R., Krüger, R., Fuchs-Kittowski, F.: Informationsbedarf von Bürger*innen bezüglich Niedrigwasser, Trockenheit und Dürre - Ergebnisse einer Umfrage. In: INFORMATIK 2023 - Designing Futures: Zukünfte gestalten, pp. 1221–1242. Gesellschaft für Informatik e.V., Bonn (2023). https://doi.org/10.18420/inf2023 132
6. Van Loon, A.F.: Hydrological drought explained. WIREs Water **2**(4), 359–392 (2015). https://doi.org/10.1002/wat2.1085
7. Wilhite, D.A., Glantz, M.H.: Understanding the drought phenomenon: The role of definitions. Water International **10**(3), 111–120 (1985)
8. Thober, S., Marx, A., Boeing, F.: Auswirkungen der globalen Erwärmung auf hydrologische und agrarische Dürren und Hochwasser in Deutschland. Tech. rep., Helmholtz-Zentrum für Umweltforschung GmbH – UFZ (2018).
9. Bathiany, S., Ney, P., Belleflamme, A., El Zohbi, J., Goergen, K., Rechid, D.: Entwicklung von Dürren in Deutschland, Europa und weltweit (2021). https://doi.org/10.25592/uhhfdm.9989
10. Merz, C., Pekdeger, A.: Anthropogenic changes in the landscape hydrology of the berlin-brandenburg region. DIE ERDE—J. Geograp. Soc. Berlin **142**(1–2), 21–39 (2011). https://www.die-erde.org/index.php/die-erde/article/view/41
11. Ministerium für Landwirtschaft, Umwelt und Klimaschutz Brandenburg: Landesniedrigwasserkonzept Brandenburg. Technischer Bericht, Ministerium für Landwirtschaft, Umwelt und Klimaschutz des Landes Brandenburg (2021). https://mluk.brandenburg.de/sixcms/media.php/9/Landesniedrigwasserkonzept-Brandenburg.pdf. Zugriff am: 24 June 2024
12. Oschatz, C.: Wissen im Wandel, Der Einfluss der Klimaberichterstattung auf den klimapolitischen Wissenserwerb. Springer VS, Mainz (2017).
13. Fliß, R., Baumeister, C., Gudera, T., Hergesell, M., Kopp, B., Neumann, J., Posselt, M.: Auswirkungen des Klimawandels auf das Grundwasser und die Wasserversorgung in

Süddeutschland. Grundwasser - Zeitschrift der Fachsektion Hydrogeologie **26**, 33–45 (2021). https://doi.org/10.1007/s00767-021-00477-z
14. Crompton, H., Burke, D.: The use of mobile learning in higher education: A systematic review. Comput. Educ. **123**, 53–64 (2018). https://doi.org/10.1016/j.compedu.2018.04.007
15. Leuzinger, S., Borrelle, S.B., Jarvis, R.M.: Improving climate-change literacy and science communication through smart device apps. Frontiers Educ. **4** (2019). https://doi.org/10.3389/feduc.2019.00138
16. Lude, A., Schaal, S., Bullinger, M., Bleck, S.: Mobiles, ortsbezogenes Lernen in der Umweltbildung und Bildung für nachhaltige Entwicklung. Der erfolgreiche Einsatz von Smartphone und Co. in Bildungsangeboten in der Natur. Schneider Verlag Hohengehren GmbH, Baltmannsweiler (2013).
17. Gerlicher, A., Jordine, T.: Mobile Learning und Mobile Game-based Learning, pp. 161–176. Springer Fachmedien, Wiesbaden (2018). https://doi.org/10.1007/978-3-658-19123-8 9
18. Parsons, D., Ryu, H., Cranshaw, M.: A study of design requirements for mobile learning environments. In: Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06), pp. 96–100 (2006). https://doi.org/10.1109/ICALT.2006.1652376
19. Choi, S.: Isar 3.1.0, benchmarks (2023). https://pub.dev/packages/isar#benchmarks. Accessed on 24 June 2024
20. Brooke, J.: Sus - A quick and dirty usability scale. In: Usability Evaluation in Industry, pp. 189–194. CRC Press Taylor and Francis (1996).
21. Mclellan, S., Muddimer, A., Peres, S.: The effect of experience on system usability scale ratings. J. Usability Stud. **7** (2011).
22. Kortum, P., Johnson, M.: The relationship between levels of user experience with a product and perceived system usability. Proc. Hum. Factors Ergon. Soc. Annu. Meet. **57**, 197–201 (2013). https://doi.org/10.1177/1541931213571044
23. Apple Inc.: Human interface guidelines (2024). https://developer.apple.com/design/human-interface-guidelines. Accessed on 11 July 2024

# Eco-Innovation and Circular Economy Performance in European Union Countries

**Pawel Bartoszczuk**

**Abstract** To date no uniform proposal for measuring and analysing eco-innovation and circular economy in the literature has been emerged. The main challenges in this regard are the need to establish indicators at micro level, taking into account the whole life cycle. Both eco-innovation and circular economy performance in the EU are not satisfactory. The effects of eco-innovation can be measured on a micro scale (for product or service, process, enterprise), meso (sector) and macro. Conversely, the effects of eco-innovation can be not always positive. Notwithstanding the high investment of companies, eco-innovation activities are producing results in the longer term on a micro scale by reducing the negative impact of its activities on the environment, while at the same time leading to the impact of eco-innovation on the company's performance in the form of lower environmental pollution charges and lower energy costs.

**Keywords** Circular economy · Eco-innovation · Eco-innovation activities · Eco-innovation observatory

## 1 Introduction

Eco-innovation contributes to the company's sustainability in reducing costs. In addition, consumers are more interested in purchasing environmentally friendly products supplied by companies using eco-innovative technologies. The positive relationship between eco-innovations and economic growth has been so widely acknowledged both in the field of evolutionary economics [1], and management theories [2], that it is currently taken for granted. Arising the "Porter hypothesis" eco-innovations are seen as instruments for enhancing economic performance [3], profitability [4, 5], market position [6], key engine of growth [7] and competitiveness [8]. As entrepreneurs demonstrate, employees engage in believing in company missions

P. Bartoszczuk (✉)
Warsaw School of Economics, Warszawa, Poland
e-mail: pbarto1@sgh.waw.pl

 
V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_22

and the deeper understanding of their own work in modern eco-innovative organisations [9]. The European Commission members countries pay particular attention to the factors determining the success of eco-innovation, namely the acquisition of business benefits and environmental protection. One is to reduce energy efficiency to reduce carbon emissions in order to meet the commitments contained in the Kyoto Protocol. Comprehensive solutions, including the Eco-IS methodology, as well as tools developed by "*the Community Innovation Survey* (CIS)" in the European Union and proprietary proposals for eco-innovation assessment, are noteworthy.

European eco-innovation research was carried out systematically by Eco-Innovation Observatory, and was limited to sources of eco-innovation determinants, barriers in selected industrial sectors in European countries, mainly Germany. It resulted in a large shortage of eco-innovation research introduced by companies operating in Central Europe, especially in Poland. Only selected aspects of the activity have so far been the subject of comprehensive analyses carried out by the Central Statistical Office, PSDB in cooperation with GfK Polonia on behalf of PARP [10], the Environmental Partnership Foundation, as well as the Gallup Institute as part of the Eurobarometer survey [11].

Circular economy is an economic system based on the reuse and regeneration of materials or products, especially as a means of continuing production in a sustainable way. With a circularity rate of 11.5% in 2022, Europe consumes a higher proportion of recycled materials than other world regions. Nevertheless, progress in the EU has been unsatisfactory, and EU is still far away from the ambition to double the Union's circularity rate by 2030. Assessing progress towards current circular ambitions, the EEA report states that there is a low or moderate likelihood for them to be achieved in the coming years [12].

The EEA report explains, however, that many circular economy policies are still relatively new and some have not yet been fully put in place at national level. Added to this, the impact of these measures takes time to filter down to changes in business models, consumption patterns and, ultimately, our patterns of resource use. However, in addition to the implementation of existing policies, more can be done (EEA, 2023) [13]. European Policy towards Green Deal is shown in Fig. 1.

## 2 Methodologies and Goal

The main goal of the article is to demonstrate that eco-innovation and circular economy, which allow European countries' economies to increase their performance, are not fully implemented yet in many countries. Another aim is to analyse eco-innovation scoreboard data and to expose the low index value causes for Central Europe region countries. To gain so declared goals, a research question was formulated as follows: Do countries with less developed economies have lower eco-innovation and circular performance or not? The performance measurement was based on secondary data analysis obtained from EU Eco-Innovation Observatory. These basic indicators and trends described European countries' eco-innovation


**Fig. 1** Policy objectives related to Green Deal ( *Source* own elaboration based on [12])

performance. The analysis is based on comparative research and available data. Paper identified a problem with the establishment of measures for the effects of eco-innovations. In the further part of article, European countries were indicated with the highest, medium and lower eco-innovation performance. Hence, the problem of the economy transformation towards the eco-innovation was addressed. The search in the main scientific databases and a review of the literature revealed an unfocused approach of the relationship between eco-innovations indicators performance, which shows a research gap regarding this subject. We decided to analyse the data, by assuming that eco-innovations scoreboard's factors are capable of explaining the companies' slow and moderate performance.

## 3 Measuring the Effect of Eco-Innovation and Circular Economy

Measuring the effects of eco-innovation and circural economy is very challenging to carry out throughout the eco-innovation chain (including R&D, the dissemination of innovative products and the use of innovative processes or organisational methods). Furthermore, the value of these innovations makes it tough for companies to access data due to companies not identifying such solutions as eco-innovation. It is also impossible to obtain them through the national accounts of the sector's production, as they occur in different areas of the economy. The main challenges in creating and operationalising eco-innovation concepts are: defining indicators at micro level, taking into account the life cycle and distinguishing eco-efficiency from the economic category of eco-innovation, naming eco-innovation analysis levels and establishing an operational approach to eco-innovation analysis in order to comprehend the effect of eco-innovation. The possible impact of eco-innovation on the competitiveness of the company is also emphasised. This, in turn, seeks to active information on the factors (the condition of competitors, the type of competing products or services) that will affect its activities over the next years, which will enable it to survive in the longer term on the market [14, 15]. Considering the traditional theory of innovation, Wagner [16] concluded that innovation considerably improves companies' social performance.

Community Innovation Survey (CIS) is steadily conducted in the European Union countries by Eurostat and based on the methodology described in *Oslo Manual* [17]. From the perspective of this article, the 2008 CIS study, which first time contained a module on eco-innovation, is noteworthy. That study covered innovations introduced between 2006 and 2008. The environmental benefits of implementing innovation are divided into:

1. environmental benefits that may arise during the period of manufacture of the product or service:
   a. reduction in material consumption per unit of product,
   b. reduction in energy consumption per unit of product,
   c. reduction of carbon dioxide emissions by the company,
   d. the use of less polluting or environmentally hazardous materials and
   e. reuse (recycling) of waste, water or materials,
2. environmental benefits that may arise during the useful life of the purchased product or the use of the service by end-users:
a. reducing energy consumption,
b. reducing air, water, soil or noise pollution and
c. improving the reusability (recycling) of the product after its useful life.

In this context, however, the establishment of measures for the effects of eco-innovation is a challenge. The most well-known initiative to consider eco-innovation in the European Union is a study by the European Observatory on Eco-Innovation (EIO). On the basis of numerous data collected from European countries, for individual countries *eco-innovation scoreboard* has been developed [18, 19]. This index consists of 16 indicators divided into 5 groups. Three of them directly relate to eco-innovation: expenditure on eco-innovation, eco-innovation activities and eco-innovation results. The other two groups of indicators are the effects resulting from the introduction of eco-innovation—environmental and socio-economic (development of the "eco-industry" of the economy). Their detailed description is specified in the following list [20]:

1. *inputs*—the share of government resources and expenditure spent on environmental and energy R&D in relation to GDP, the share of R&D participants in the number of all employees, the volume of green investments under PE/VC funds;

2. activities—*participation* of eco-innovation companies improving material and energy efficiency in the number of companies in the country in total, the number of organisations holding ISO14001 environmental management certificates (in relation to the number of citizens in a given country);
3. outputs—*number* of patents, scientific publications (in relation to the number of citizens in a given country), information in the media on eco-innovation (in relation to the number of electronic media available);
4. *environmental outcomes—efficiency* of energy, raw materials, water and greenhouse gas emission factors and
5. *socio-economics outcomes*—the share of environmental exports in general exports, the share of workers employed in the environmental sector in the total number of employees and the overall turnover generated by the environmental industry.

There are some indicators of circularity, one of them is presented in the first edition of the global Circularity Gap Report, in 2018. Circle Economy launched the Circularity Metric for the global economy. That document adapted the Metric to country profile. This shows that only 10.2% of the resources Poland uses are cycled back into the economy after use. Unfortunately that the vast majority—nearly 90%—of Poland's resources are not reused. While the country's yearly per capita material footprint of 13.8 tonnes may seem rather modest in comparison to other European nations, it still surpasses the global average of 11.9 tonnes—and is nearly double what is measured as a "sustainable" level of consumption.

## 4 Results

The composite index relating to both eco-innovation and innovation in general reflects the disadvantage of Polish economy belonging to non-innovative and non-ecoinnovative European countries (Fig. 2). Poland is in the low eco-innovation performance group of countries, as are majority of Central and Eastern European countries and some Western countries (such as Malta). Even against the background of these countries, there is a great distance of Poland to other so-called "the New Member States", e.g. the Czech Republic and Estonia. The results of these indices specify insufficient implementation of eco-innovation in enterprises in Poland. Currently available data presented by the European Eco-Innovation Observatory confirm a significantly insufficient degree of eco-innovation of the Polish economy [20–23].

The negative assessment covers almost each of the five sub-elements of the index, except for the socio-economic effects that have been positively assessed. The level of involvement of Polish companies in the implementation of eco-innovation meant that Poland, together with Bulgaria, took the last place in the 2011 ranking with an index of only 41 points (while in 2013 and 2015 it was the penultimate place and 22nd place with 59 points in the 2019 ranking). Other countries that joined the EU at the same time as Poland achieved much higher index values in 2011,

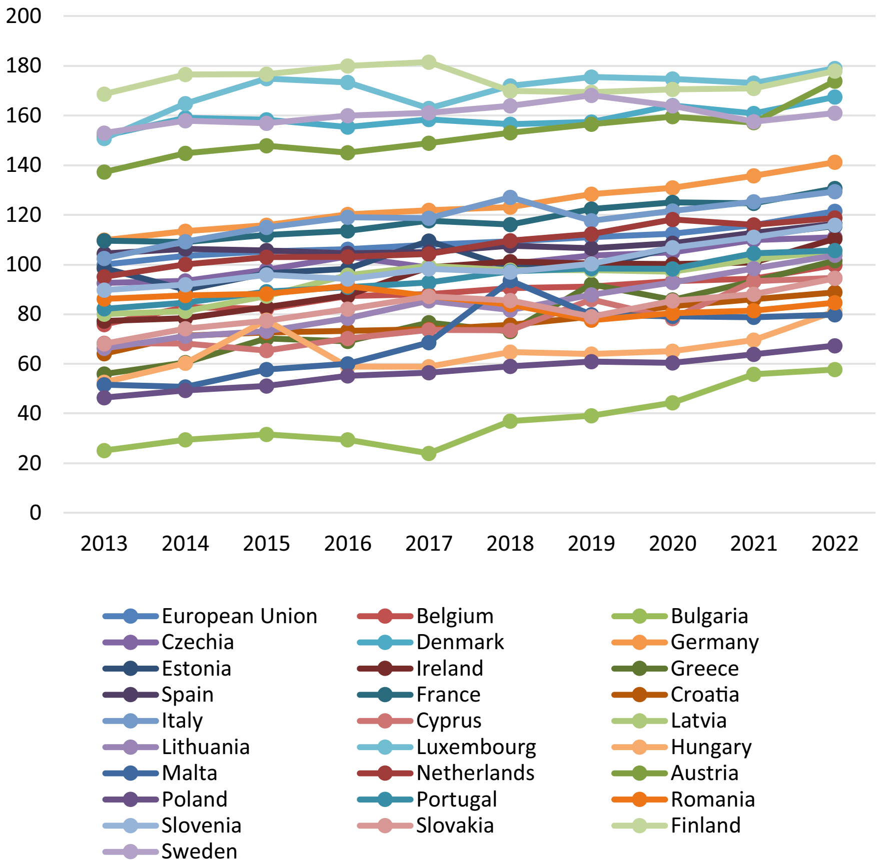



**Fig. 2** EU eco-innovation indicator (EU average-27 = 100) (*Source* own study based on Eco-innovation at the heart of European policies, European Commission, 2020, https://green-business.ec.europa.eu/eco-innovation_en [19. 08. 2024])

e.g. The Netherlands. Czech Republic—171 points (in 2019—96 points (Fig. 2) and Slovenia—119 points (in 2019—94 points). These results verify that Poland is in a rather unfavourable climate and that the interest of companies in introducing eco-innovation is insufficient.

Compared to 2014, Poland moved from 3rd last place to 2nd last place in the eco-innovation ranking in 2020, which gave it a performance index score (64) better than the one previously obtained (25) (Fig. 2). Poland is still ahead of countries such as Bulgaria. Luxembourg came first in the eco-innovation ranking (2020), with a score of 162 points, ahead of Denmark by 18 points. Bulgaria, on the other hand, came last in the ranking of European countries.

In the Polish case, there is a considerable distance within the five groups of indicators of the complex eco-innovation index compared to the Germany and even the Czech Republic (Fig. 3). Only in terms of socio-economic performance, Poland leads the ranking, but for all others it ranks below the EU average. These data should be treated with a degree of caution, as not always, for example, higher employment in eco-services contributes to greater eco-innovation and results. Three component indicators on socio-economic outcomes were highlighted:

- export of products from the eco-industry (% of total exports—Eurostat), [28]
- employment in the eco-sector and circular economy (% of total employment in all enterprises—Orbis base), profits in the eco-sector and circular economy (share % of total employment in all enterprises—Orbis base) (Fig. 4). [29]

In terms of indicators reflecting expenditure on eco-innovation and performance, Poland did not even reach half the average value of European countries (as it scored only 40 points). We have to note that the eco-innovation index has some drawbacks. For example, the definition of eco-industry adopted by Eurostat and Ecorys, on the basis of which socio-economic indicators were created, refers to a large range of green activities (e.g. recycling, wastewater treatment, the introduction of renewable

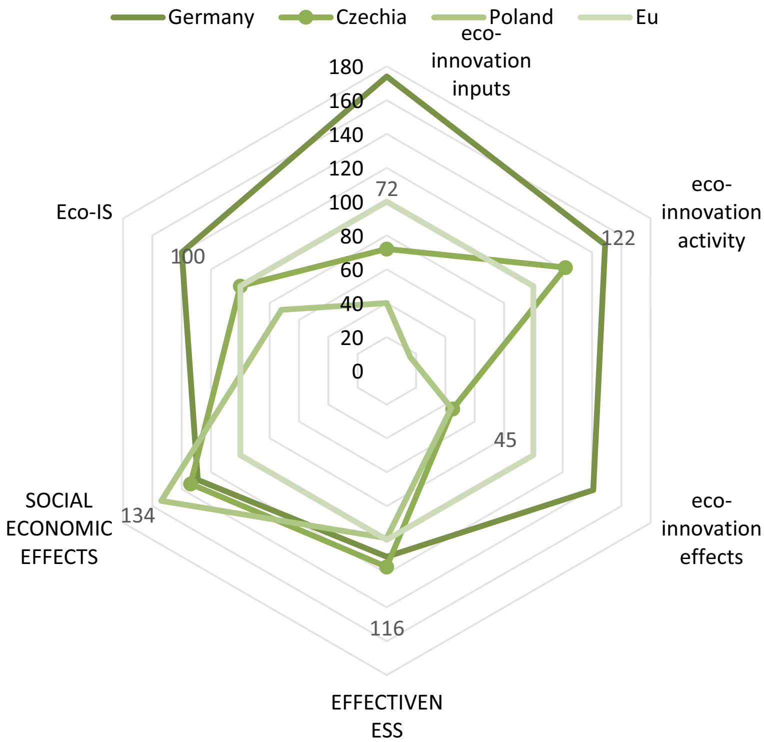



**Fig. 3** Eco-innovation gap in five themes Polish (relative to the EU average between Poland and the Czechia and eco-innovation leader Germany, EU = 100) (*Source* own study based on[21, 24–27] [16. 08. 2024])

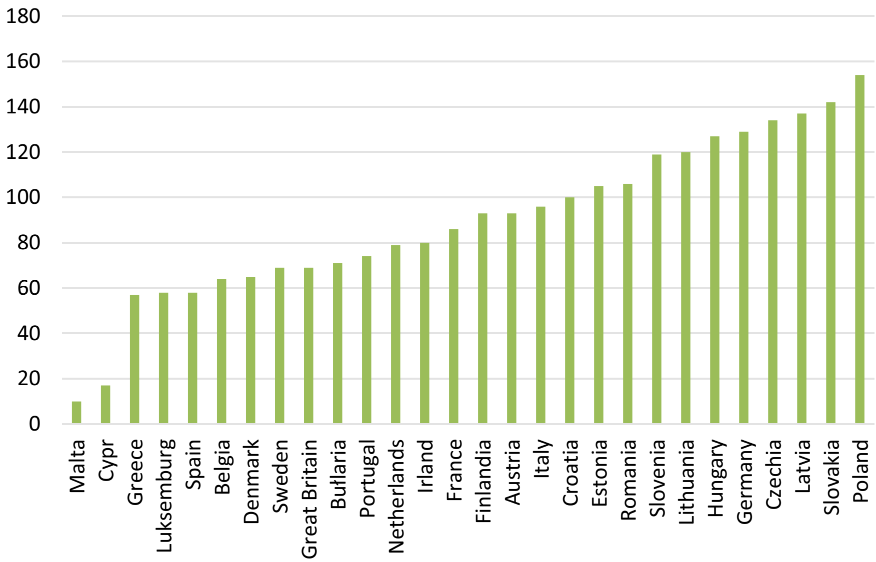



**Fig. 4** Indicator for social economic results of eco-innovation for European Union countries (*Source* own study on the basis of: *Eco-innovation at the heart of European policies*, European Commission, 2016, other data were not available https://ec.europa.eu/environment/ecoap/scoreboard_en [16.08.2024])

energy sources), with competitiveness and employment levels in some of them not necessarily directly derived from eco-innovation (e.g. hydropower plants). At the same time, an important part of the economic impact of the implementation of eco-innovation is becoming apparent in many other sectors where companies that have introduced eco-innovation were not motivated by eco-innovation, but by the need to improve productivity and reduce costs. Therefore, performance indicators do not differ between the EU-27 countries to the same extent as those directly related to eco-innovation, as is also evident in the case of Polish. In addition, performance indicators are poorly correlated with indicators of the intensity of the development and implementation of eco-innovation. The environmental effects are measured by static indicators, which depend to a large extent on the historical development paths of economies and not necessarily on eco-innovation introduced in the last dozen years[30, 31].

Contrary to expectations, countries with a higher resource of scientific and technical personnel staff do not necessarily achieve better results in the implementation of eco-innovation. Countries such as France, Italy, Spain and Poland have a high share of research personnel in general, but it is much less efficient than countries that have achieved a lower eco-innovation rate (e.g. Denmark) (Fig. 5).

The circular economy incorporates a regenerative system that minimises the entry and waste of resources, emissions, and expenditure of energy through slowing down, closing, and straightening material and energy circuits. Companies must reshape their

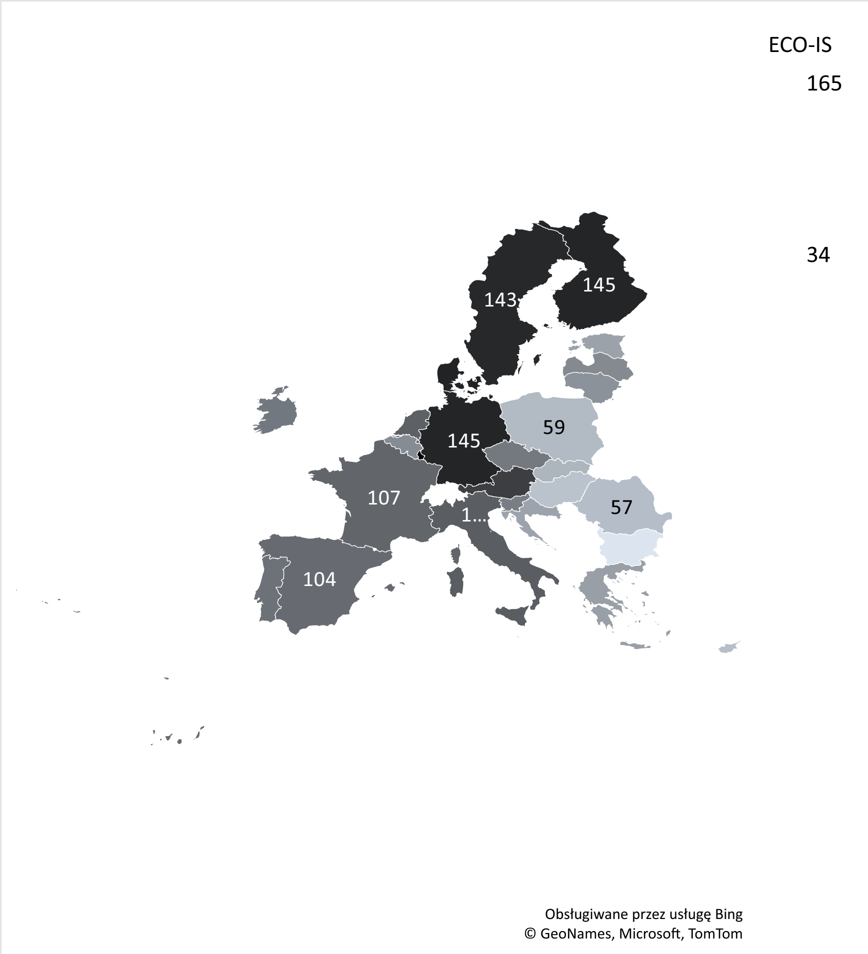



**Fig. 5** Eco-innovation Scoreboard indicator for European Union countries (*Source* Own elaboration based on Eco-innovation data https://ec.europa.eu/environment/ecoap/indicators/index_en https://green-business.ec.europa.eu/eco-innovation_en 19.08.2024)

business models and the ways they propose value to their clients while simultaneously considering environmental and social facets[32]. Poland has a material footprint of 13.8 tonnes per capita—definitely exceeding the global average of 11.9 tonnes per capita. However, this is well-beaten by Norway's footprint, which comes in at a huge 44.3 tonnes per capita, a level far beyond planetary boundary limits. In terms of the circularity, Poland is around four times more circular than Norway (10.2%) [33].

Better performance of Poland in comparison with Norway can be explained due to Polish consumption of more cycled materials than Norway's, compared to their material consumption (Table 1). However, it's worth strongly underlying that the

**Table 1** Comparison of Poland and Norway economy and circular economy performance

| | Poland | Norway |
|---|---|---|
| Population (mln) 2021 yr | 37,8 | 5,4 |
| Population density (persons per sq.km) | 123 | 15 |
| GDP per capita | 15,165 | 140,000 |
| GDP distribution (% of GDP) | Agriculture = 2,4<br>Industry = 29,3<br>Service = 55,6 | Agriculture = 1,6<br>Industry = 35,6<br>Service = 52,5[33] |
| Unemployment % of total labour force | 2,6 | 2,9 |
| Material footprint (tonnes per capita per year | 16,7 | 63 |

Circularity Metric cannot be directly compared between countries: the methodology used for Poland's analysis is a newer, more sophisticated version than that used for the Circularity Gap Report Norway. It's not known to which extent applying the new methodology to Norway's economy would change its Metric—but we would expect a significant impact. Construction, agriculture and extractive industries contribute most to the material footprint in both Poland and Norway, with the impact of extractive industries being seen by domestic extraction levels far exceeding the global average of 12,3 tonnes per capita per year in both cases. However, potential for reducing impact is also present in other sectors, such as the mobility and energy sectors.

According to [34], the nations at the forefront of implementing operations corresponding to circular economy principles are Germany, Belgium, Spain, France, Italy, the Netherlands, and the United Kingdom. The second pole encompasses European Union member states located in Central and Eastern Europe, as well as those situated in the southern region of Europe (data from 2018).

## 5 Conclusion

The considerable concerns related to the development of eco-innovativeness and circularity are observed in many European countries, particularly Central European countries. First, obsolete industries representing, linear economy are not interested in development of eco-innovations. On the one hand, there is no or limited pressure on research in new eco-innovative solutions. Furthermore, academic R&D centres are unable to satisfy the demands of the industry. It is of key importance that eco-innovations are perceived as the driving factor of the third transformation of EU New Members (following system transformation and the EU accession). The existence of a "two-speed Europe" was identified in terms of EU countries' advancement towards circular economy and eco-innovation. Countries with less developed economies have lower eco-innovation but not necessarily circular performance. It can be explained by methodologies of gathering data, but also different levels of economic development. Because less developed countries can use more recycled and used products.

This research definitely has some caveats. For example, poor methodologies of calculating eco-innovation scoreboard, deficit of statistical data at National Statistical Institutes, many data were available only from the earlier periods. We observed shortcomings of the sub-indicators that are part of the index, e.g. expenditures. This is particularly true under the heading government outlays and expenditure on research and development in the field of environment and energy. It only includes part of the expenditure and does not reflect the expenditure incurred by organisations from their resources or from bank loans, EU funds or other sources (e.g. in Poland most innovations are financed by non-governmental funds). In situation, when the structure of eco-innovation funding sources were similar in all EU countries, this would not affect the ranking. However, if this were not the case, it could turn out that the position of countries in terms of R&D expenditure is different from that of government expenditure only [35, 36]. It seems that more attention should be paid to this issue in further planned research.

## References


1. Rennings, K.: Towards a Theory and Policy of Eco-Innovation—Neoclassical and (Co-) Evolutionary PerspectivesRe. (1998)
2. Marín-Vinuesa, L.M., Scarpellini, S., Portillo-Tarragona, P., Moneva, J.M.: The impact of eco-innovation on performance through the measurement of financial resources and green patents. Org. Envi. **33**, 285–310 (2020). https://doi.org/10.1177/1086026618819103
3. Shuwaikh, F., Benkraiem, R., Dubocage, E.: Investment in green innovation: how does it contribute to environmental and financial performance? J. Innov. Econ. Manag. **41**, 107–149 (2023). https://doi.org/10.3917/jie.pr1.0137
4. Kemp, R., Pearson, P.: Final report MEI project about measuring eco-innovation. UM Merit, Maastricht. **10** (2007)
5. Zhang, R., Yuan, J.: Enhanced author bibliographic coupling analysis using semantic and syntactic citation information. Scientometrics **127**, 7681–7706 (2022)
6. Lin, R.-J., Tan, K.-H., Geng, Y.: Market demand, green product innovation, and firm performance: evidence from Vietnam motorcycle industry. J. Cleaner Prod. **40**, 101–107 (2013)
7. Leoncini, R., Marzucchi, A., Montresor, S., Rentocchini, F., Rizzo, U.: 'Better late than never': the interplay between green technology and age for firm growth. Small Business Economics. **52** (2019). https://doi.org/10.1007/s11187-017-9939-6
8. Oncioiu, I., Căpuşneanu, S., Türkeş, M.C., Topor, D.I., Constantin, D.-M.O., Marin-Pantelescu, A., Ștefan Hint, M.: The sustainability of romanian smes and their involvement in the circular economy. Sustainability. **10**, 2761 (2018). https://doi.org/10.3390/su10082761
9. Rutkowska, M., Kamińska, A.: Turquoise Management model -teal organizations. In: Education Excellence and Innovation Management: A 2025 Vision to Sustain Economic Development during Global Challenges (2020)
10. Kornecki, J., Michaliszyn, B., Krupanek, J., Ratman-Kłosińska, I., Banasiek, M., Pochroń, A., Kondrat, P., Jagusztyn-Krynicki, T., Czyż, P., Wolińska, I., Pylak, K.: Potencjał małych i średnich przedsiębiorstw w dziedzinie kreowania nowych produktów innowacyjnych – rozwiązania proekologiczne, Warszawa (2010)
11. Gallup: Attitudes of European entrepreneurs towards eco-innovation. Analytical report, Flash Eurobarometer. **315** (2011)
12. European Environment Agency.: Accelerating the circular economy in Europe: State and outlook 2024. Publications Office, LU (2024)

13. Circular economy: https://www.eea.europa.eu/en/topics/in-depth/circular-economy. Last accessed 10 May 2024
14. Woźniak, L., Strojny, J., Wojnicka, E.: Ekoinnowacyjność dziś i jutro - wyzwania, bariery rozwoju oraz instrumenty wsparcia. Polska Agencja Przedsiębiorczości, Rzeszów (2010)
15. Tutaj, J., Rutkowska, M., Sulich, A.: Eco-innovation as an Element of Business Value and Performance Management. Presented at the , Conference: Hradec Economic Days 2019At: Hradec/ Czech Republic February 5 (2019)
16. Wagner, M.: Empirical influence of environmental management on innovation: Evidence from Europe. Ecological Econo. **66**, 392–402 (2008). https://doi.org/10.1016/j.ecolecon.2007.10.001
17. Development, O. for E.C. and, Communities, S.O. of the E. eds: Oslo manual: guidelines for collecting and interpreting innovation data. Organisation for Economic Co-operation and Development : Statistical Office of the European Communities, Paris (2005)
18. O'Brien, C.: Sustainable production—A new paradigm for a new millennium. Int. J. Prod. Econ. **60–61**, 1–7 (1999)
19. O'Brien, M., Miedziński, M.: EIO_Annual_Report_2012 (2012)
20. Eco Innovation Observatory—Annual Reports—Annual Reports, http://www.eco-innovation.eu/index.php/reports/annual-reports
21. Bleischwitz, R., Giljum, S., Kuhndt, M., Schmidt-Bleek, F., Wuppertal-Institut für Klima, U.: Eco-innovation—Putting the EU on the path to a resource and energy efficient economy. Wuppertal Inst. for Climate, Environment and Energy, Wuppertal (2009)
22. Doranova, A., Roman, L.: Policies and Practices for Eco-Innovation Up- take and Circular Economy Transition. **88** (2016)
23. Eco-innovation Action Plan—European Commission, https://ec.europa.eu/environment/ecoap/frontpage_en, last accessed 12 February 2018
24. The Eco-Innovation Scoreboard and the Eco-Innovation Index—Eco-innovation Action Plan - European Commission, https://ec.europa.eu/environment/ecoap/indicators/index_en
25. European Comission: The Eco-Innovation Scoreboard and the Eco-Innovation Index, https://ec.europa.eu/environment/ecoap/indicators/index_en. Last accessed 29 January 2021
26. Foxon, T., Andersen, M.M.: The greening of innovation systems for eco-innovation–towards an evolutionary climate mitigation policy. In: DRUID Summer Conference-Innovation, Strategy and Knowledge (2009)
27. Eco-innovation at the heart of European policies, European Commission, 2016. European Comission (2016)
28. Database—Eurostat. https://ec.europa.eu/eurostat/data/database. Last accessed 17 December 2020
29. International Company Information Database | Bureau van Dijk, https://www.bvdinfo.com/en-gb/our-products/data/international. Last accessed 17 December 2020
30. Park, M., Bleischwitz, R., Han, K., Jang, E., Joo, J.: Eco-Innovation indices as tools for measuring eco-innovation. Sustainability **9**, 2206 (2017)
31. Bilsen, R.: Study on the competitiveness of the eu eco-industry. within the framework contract of sectoral competitiveness studies. Final Report—Part 1; ECORYS: Rotterdam, The Netherlands, ECORYS, Rotterdam (2009)
32. Suchek, N., Fernandes, C.I., Kraus, S., Filser, M., Sjögrén, H.: Innovation and the circular economy: A systematic literature review. Bus Strat Env. **30**, 3686–3702 (2021). https://doi.org/10.1002/bse.2834
33. CGR Poland, https://www.circularity-gap.world/poland. Last accessed 18 August 2024
34. Mazur-Wierzbicka, E.: Circular economy: advancement of European Union countries. Environ. Sci. Eur. **33**, 111 (2021). https://doi.org/10.1186/s12302-021-00549-0
35. Pakulska, J.: The eco-innovation versus economic development on the EU example. ERSJ. **XXIV**, 999–1008 (2021). https://doi.org/10.35808/ersj/2007
36. Pakulska, J.: Nakłady na innowacje ekologiczne w krajach Unii Europejskiej w latach 2010–2018. Studia Ecologiae et Bioethicae, 57–70 (2020). https://doi.org/10.21697/seb.2020.2.06

# Empowering Sustainable Development: The Role of E-Commerce Adoption in Advancing SDGs Among Youth

**Noha Tarek Amer, Hany Abdelghaffar, and Laila Amr**

**Abstract** E-commerce has been growing exponentially impacting various industries, one of which is the fast-food sector. The purpose of this study is to examine the factors that influence college students to adopt fast-food e-commerce platforms by extending the Technology Acceptance Model (TAM). While substantial research exists on e-commerce adoption, niche markets like fast-food e-commerce platforms have received comparatively little attention. This research addresses this gap and highlights the critical role of e-commerce in advancing Sustainable Development Goals (SDGs). Using a survey method for data collection, the results reveal that perceived ease of use, user experience, and product appeal significantly contribute to the prediction of using fast-food e-commerce platforms among college students. These findings highlight a shift toward increased online ordering among youth, which has implications for sustainability. By focusing on these factors, e-commerce decision-makers can enhance platform adoption ultimately resulting in promoting economic growth (SDG 8), enhancing sustainable and consumption patterns (SDG 12), and reducing carbon emissions (SDG 13).

**Keywords** Adoption · E-commerce · TAM · Sustainability · SDG 8 · SDG 12 · SDG 13

## 1 Introduction

Electronic commerce (E-commerce), especially Business-to-Consumer (B2C) platforms, has become crucial in today's business landscape due to the growing reliance on the internet [1]. E-commerce enables digital exchanges between buyers and sellers through the use of the internet and web. The huge growth of e-commerce platforms can be traced back to the unique characteristics of the internet such as its widespread availability, global reach, standardization, abundance of information,

N. T. Amer (✉) · H. Abdelghaffar · L. Amr
German University in Cairo, Cairo, Egypt
e-mail: noha.amin@guc.edu.eg



V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-031-85284-8_23

interactive features, rich media, and ability to personalize and customize. These platforms contribute to Sustainability Development Goals (SDGs): supporting economic growth (SDG 8) through the creation of new economic opportunities, enhancing responsible consumption and production (SDG 12) through sustainable consumer behavior, and promoting climate action (SDG 13) through reducing carbon emissions [2].

The fast-food e-commerce industry has experienced steady growth in recent years, with college students emerging as the largest group of online shoppers due to convenience, appealing offers, and the rapid advancement of e-commerce [3]. The literature on e-commerce adoption, particularly in the fast-food industry, extensively utilizes the Technology Acceptance Model (TAM) to understand consumer behavior and preferences. A study by Ye Zihan et al. [4] examines the adoption factors of perceived ease of use, time savings, and privacy protection as influencing toward online food delivery services during crises like the COVID-19 pandemic. However, this model doesn't take into account the specific characteristics of college student's niche and may not fully address some notable factors influencing adoption behavior toward online food delivery services.

Understanding consumer's adoption behavior in e-commerce can be a powerful tool helping decision-makers develop a superior strategy to compete in the digital age. Additionally, promoting e-commerce platforms can also support macro-economic goals of supporting economic growth, enhancing responsible consumption and production, and promoting climate action [5]. Therefore, our study aims to dive into the critical elements that influence the adoption of fast-food e-commerce platforms answering the research question of:

> Which factors do fast-food e-commerce decision-makers need to consider to promote the adoption of their platforms among college students in order to achieve Sustainable Development Goals (SDGs)?

The structure of this paper is as follows: Sect. 2 presents a comprehensive literature review and proposed framework, Sect. 3 outlines the research methodology, Sect. 4 presents and discusses the findings, and Sect. 5 offers conclusions and implications.

## 2 Literature Review and Proposed Framework

The 2030 Agenda for Sustainable Development, ratified by 193 member states of the United Nations, encompasses 17 SDGs that seek to enhance the quality of life while promoting sustainable consumption and production modalities. E-commerce contributes to achieving SDGs by fostering economic growth, innovation, and enhancing legal frameworks for sustainable development. A study by Revinova [2] concludes that e-commerce significantly contributes to achieving 10 out of the 17 SDGs. Another study by Ju et al. [6] found that e-commerce contributes favorably to advancing sustainable development in emerging and developing countries in Asia.

The Technology Acceptance Model (TAM) is one of the most widely cited models for understanding user's adoption of a certain technology. Davis [7] found that TAM posits two variables, perceived usefulness (PU) and perceived ease of use (PEOU) that are associated with an individual's intention and attitude to accept a specific technology. TAM suggests that higher PU and PEOU contribute to greater acceptance and adoption of the technology.

Studies utilizing TAM to investigate technology adoption have yielded significant findings. Nihayah and Purnama [8] identified perceived usefulness and ease of use as pivotal factors affecting users' intention to adopt digital banking services. While TAM has been widely utilized in studying technology adoption, it is criticized for its tendency to overlook critical factors [9]. A study by Chen [10] extended the TAM model to understand blockchain technology adoption by incorporating factors such as strategic management and social influence at the firm level, and individual innovation and self-efficacy at the individual level. Another study investigates mobile application adoption extended TAM to include mobile skillfulness-related factors to better fit the unique characteristics of mobile applications [11].

This study advances the literature on adoption behavior among college students, with a particular focus on e-commerce platforms in the fast-food industry. The widely used Technology Acceptance Model (TAM) often does not account for the specific attributes of different industries or the potential variation in adoption. Researchers have expanded its applicability by incorporating other influential factors that can significantly impact technology acceptance. For instance, the presence of engaging user interface and product recommendations has been shown to enhance students' willingness to embrace technologies [12, 13]. This suggests that a more refined approach which considers industry-specific dynamics may be essential for effectively promoting technology adoption.

To address this gap, this paper proposes a new framework (see Fig. 1) designed to enhance our understanding of the critical factors influencing college students' adoption of fast-food e-commerce platforms.

### *2.1 Perceived Ease of Use*

Perceived ease of use is defined as the degree a person believes that using a certain technology is easy to use and trouble free [14]. It is conceptualized through ease of learning, adaptability, and feedback mechanism [15–17]. Ease of learning refers to how easy users find it to use the system, adaptability refers to how users can adapt to using the system, and feedback mechanism refers to how the technology provides clear feedback on the user's actions.

Perceived ease of use contributes to SDG 12 (responsible consumption and production) by making it simpler for consumers to access and choose products online. Numerous academic studies have established that the perceived ease of use (PEOU) has a substantial impact on the adoption of technology. Wafiyyah et al. [17] highlight

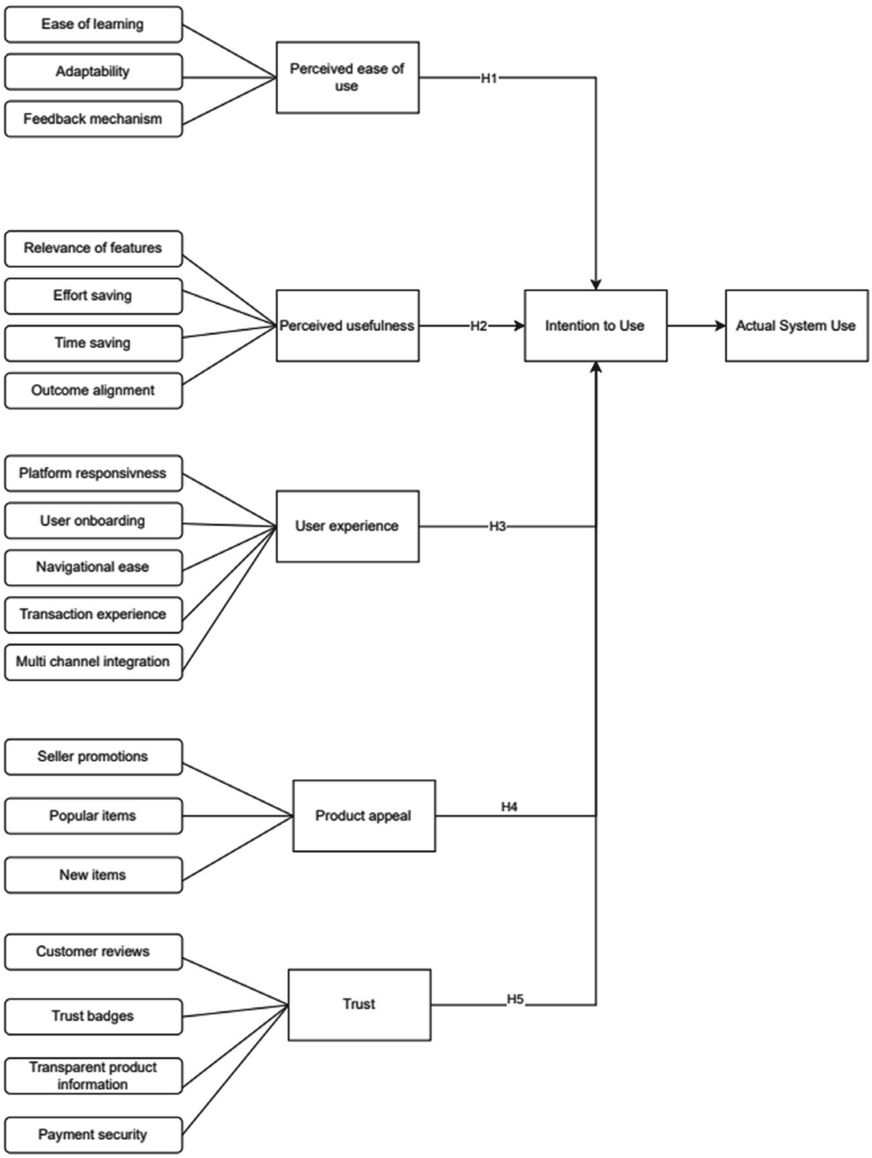



**Fig. 1** Proposed framework

the crucial role played by perceived ease of use in repurchase intentions, contributing to the successful adoption of various technologies in e-commerce, hence:

**H1:** Perceived ease of use has a positive effect on fast-food e-commerce adoption among college students.

### 2.2 Perceived Usefulness

Perceived usefulness is defined as the degree to which a person believes that using a certain system or platform would improve their job performance. It is conceptualized through relevance of features, effort saving, time saving, and outcome alignment [18, 19]. Relevance of features refers to how the users find the features relevant to their needs, effort saving and time saving refer to how the system saves the users effort and time, and outcome alignment refers to how the system allows the user to achieve their goal.

Perceived usefulness of e-commerce platforms contributes to SDG 13 (climate action) by reducing the need for physical travel and lowering carbon emissions [20]. Various research papers confirm that perceived usefulness has a significant impact on technology adoption. A study on large language models (LLMs) among students emphasized the importance of perceived usefulness [21]. Additionally, research on E-CRM systems in communication companies emphasized the impact of perceived usefulness on employee performance and user satisfaction [22]. Overall, these studies emphasize the key role of perceived usefulness in technology adoption, hence:

**H2:** Perceived usefulness has a positive effect on fast-food e-commerce adoption among college students.

### 2.3 User Experience

User Experience (UX) is defined as the totality of a person's involvement with the system (Tovar 2022). UX is conceptualized through various factors [23]. Platform responsiveness refers to how responsive the system is to the user's actions, user onboarding refers to how easy the initial setup of the system is, navigational ease refers to how easy the user finds it to navigate the system, e-commerce transaction experience refers to how streamlined the checkout process is, and multichannel integration assess the ability to access the system from across multiple channels.

User experience (UX) supports SDG 12 by ensuring that e-commerce platforms are intuitive and efficient which encourages consumers to make responsible consumption choices. Several studies have established that the user experience plays a crucial role in influencing the adoption of technology. For example, an examination of the UX impact on m-commerce suggests that UX should be a top priority for website designers and developers in order to enhance customer satisfaction and loyalty [24]. This indicates that a positive user experience positively impacts the adoption of diverse technologies, hence:

**H3:** User experience has a positive effect on fast-food e-commerce adoption among college students.

## 2.4 Product Appeal

Product appeal refers to the attractiveness of a product as perceived by consumers. It is conceptualized through seller promotions, popular items, and presence of new items [25, 26].

Product appeal enhances SDG 12 by showcasing products in a way that makes them more attractive thereby promoting environmentally responsible shopping behaviors. Notable articles have explored the role of product appeal on e-commerce. For example, a study by Liu et al. [27] found that product appeal enhances consumer's perceptions of the products offered. Another study by Peng et al. [28] highlights the importance of product appeal in positively influencing the consumer's perception of e-vendor's website. These studies highlight the critical role that product appeal plays in fostering the adoption of various technologies, hence:

**H4:** Product appeal has a positive effect on fast-food e-commerce adoption among college students.

## 2.5 Trust

Trust is an attitude that an agent (the trustor) has toward an entity (the trustee) where the trustor relies on the trustee to act in a way that aligns with the trustor's goals and interests. In the context of e-commerce, trust encompasses the confidence that consumers have in the reliability, security, and integrity of online merchants and their platforms. Trust is conceptualized through customer reviews, trust badges, transparent product information, and payment security [29, 30].

Trust in e-commerce platforms is critical for SDG 8: fostering a healthy digital economy. When users trust an e-commerce platform, they are more likely to engage in transactions which drives economic activity and contributes to job creation. Several studies have investigated the role of trust in relation to technology adoption. For example, in India, trust has a substantial impact on the behavioral intention to utilize electronic payments [31]. Another study revealed that trust, along with government support, significantly impacts the adoption of fintech services [32]. These studies highlight that trust plays an important role in the adoption of various technologies, hence:

**H5:** Trust has a positive effect on fast-food e-commerce adoption among college students.

# 3 Research Methodology

## *3.1 Data Collection*

The central point of this study is college students. This is mainly because college students became the largest users of online shopping due to their frequent internet usage making them an important population to study [33]. Additionally, for convenience of data collection and control of cultural factors, this study will only focus on Egypt [34]. This research is categorized as quantitative in nature utilizing stratified random sampling. This approach involves dividing the target population into distinct subgroups based on characteristics like academic major or campus location and then randomly selecting participants from each strata. This method ensures accurate representation within the demographic group [35].

## *3.2 Instrument Design*

This study utilizes a survey instrument as it helps obtain data from a large and widely distributed group of college students, allows for the examination of multiple constructs at the same time, and provides participants with anonymity resulting in honest responses.

The questionnaire items are organized according to each variable. All but two variables are assessed using a Likert scale, where a response of 1 signifies strong disagreement and a response of 5 signifies strong agreement. The remaining two variables were assessed using the yes/no method [36].

## *3.3 Respondent Characteristics*

A total of 287 valid responses were analyzed. Table 1 exhibits respondent characteristics where most of the respondents are between the ages of 19 and 25 representing 88% of the respondents. Of all participants, 58.5% are females and 41.5% are males. The age distribution reveals that a significant majority of the sample falls within the 19–25 age group, indicating a predominantly young demographic. The other age groups comprise much smaller proportions.

**Table 1** Respondent characteristics

| | Category | Percentage (%) |
|---|---|---|
| **Age** | 13–18 | 9.1 |
| | 19–25 | 88.1 |
| | 26–30 | 2.8 |
| **Gender** | Male | 41.5 |
| | Female | 58.5 |

## 4 Results and Discussion

Cronbach's alpha and the composite reliability serve as a tool for assessing the internal consistency and reliability of measurements by examining the relationships among observed item variables. Our results show a Cronbach alpha value of 0.814 which is above the acceptable threshold of 0.700 [37]. This finding implies that the test items are effectively capturing the same latent construct, thus establishing the reliability of the test.

Multicollinearity can negatively impact the predictive ability of the framework as it indicates that two or more independent variables are related making it complicated to find relationships between independent and dependent variables. This can be determined by calculating the variance inflation factor (VIF) for each variable [38]. In cases where the VIF value of an independent variable surpasses 10, it signifies high collinearity and warrants consideration for removal from the regression model [38]. Findings show that VIF values are ranging from 1.007 to 1.672 concluding that multicollinearity is not a concern within this dataset.

Convergent validity, which measures the amount of variance captured by the construct in relation to the amount of variance attributable to measurement error, requires an average variance extracted (AVE) no less than 0.50 [39]. An AVE below this threshold, as seen here with 0.180, points to significant unexplained variance, which may be due to measurement errors, irrelevant variables, or weak item loadings. It is crucial to recognize that a low Average Variance Extracted (AVE) does not necessarily imply that a construct fails to capture relevant material. Constructs in business studies such as perceived ease of use and trust often encompass multiple dimensions which can complicate their measurement and validation processes [39]. Despite the low AVE, there is still a possibility that the construct captures material beyond its fundamental concept.

The multiple regression analysis is used to test the hypotheses. The purpose of the regression analysis is to relate the dependent variable to a set of independent variables. To determine the relationships among the variables, R-squared measures the proportion of the variance in the dependent variable that is explained by the independent variables in the regression model. It was found that perceived ease of use, user experience (UX), and product appeal significantly contribute to the adoption of fast-food e-commerce. Whereas perceived usefulness and trust insignificantly contribute to the adoption of fast-food e-commerce (Table 2).

**Table 2** Hypothesis results

| Hypothesis | Factor | Significance | R-Squared | Outcome |
|---|---|---|---|---|
| H1 | Perceived ease of use | 0.048 | 0.027 | Accepted |
| H2 | Perceived usefulness | 0.612 | 0.009 | Rejected |
| H4 | User experience (UX) | < 0.001 | 0.080 | Accepted |
| H3 | Product appeal | < 0.001 | 0.866 | Accepted |
| H5 | Trust | 0.797 | 0.006 | Rejected |

The study found that perceived ease of use (H1, R-squared = 0.027, *p*-value = 0.048) significantly affects the intention to adopt fast-food e-commerce among college students. When users perceive the platform as easy to navigate, they are more inclined to embrace it [40]. The ease with which users engage in e-commerce enhances sustainable consumption (SDG 12) as it ensures sustainable consumption and production patterns. Despite its statistical significance, the modest R-squared value implies that while ease of use is crucial, it only accounts for a small portion of the variance in adoption, suggesting that other significant factors may also influence adoption decisions.

Perceived usefulness (H2, R-squared = 0.0009, *p*-value = 0.612) does not significantly influence the adoption of fast-food e-commerce platforms among college students. This discovery challenges the conventional notion that perceived usefulness serves as a primary driver for the adoption of e-commerce platforms, suggesting that other variables may yield greater influence in this context. This could suggest that consumers do not immediately recognize the broader sustainability benefits (e.g., reducing carbon emissions, SDG 13) through their e-commerce actions, even though those benefits exist. This might draw the attention of government agencies to provide awareness campaigns to highlight the benefits of e-commerce on sustainability.

User Experience (H3, R-squared = 0.080, *p*-value = 0.001) has surfaced as a significant element. This result indicates the critical nature of a positive user experience in relation to fast-food e-commerce platforms. An effective UX, distinguished by website responsiveness, seamless navigation, efficient checkout process, and visual appeal, plays a substantial role in shaping students' inclination toward adopting the platform [13]. UX ultimately enhances sustainable consumption (SDG 12) as users are more likely to engage in sustainable behaviors when platforms are user-friendly and provide positive experiences. Despite the indication from the R-squared value that UX accounts for only a limited percentage of the variance in adoption, the importance it holds highlights the necessity for platforms to prioritize it.

Product appeal (H4, R-squared = 0.866, *p*-value = 0.001) yielded a highly significant outcome supporting the hypothesis. This strong association highlights the critical role that product appeal plays in the adoption of fast-food e-commerce platforms. Students are notably motivated by promotions, trending items, and new items in relation to their intention to use a fast-food e-commerce platform [12]. Product appeal ultimately enhances responsible consumption (SDG 12) as users are more likely to engage in sustainable behaviors when they are attracted by promotions and trending items. This result stresses the significance of integrating product appeal techniques into e-commerce platforms to increase the adoption among college students.

Trust (H5, R-squared = 0.006, $p$-value = 0.797) has been identified as statistically insignificant. This observation suggests that students likely uphold a basic level of trust in e-commerce platforms potentially due to the inherently low-risk nature of e-commerce transactions [41]. Additionally, this suggests that users' adoption of e-commerce for sustainability may not rely heavily on trust, but more on ease of use and product appeal, as these factors are stronger drivers toward achieving SDGs like SDG 13 (Climate Action) and SDG 8 (Job Creation). Consequently, this discovery prompts a deeper exploration into the perception of trust within this setting.

## 5 Conclusion

This study offers valuable insights into the factors impacting the adoption of fast-food e-commerce platforms among college students and their implications for sustainability development goals. The research reveals that perceived ease of use, user experience (UX), and product appeal are crucial in positively impacting user adoption. Specifically, perceived ease of use, UX, and product appeal were found to significantly influence sustainable consumption (SDG 12) as users are more likely to engage in sustainable behaviors when products are attractive and platforms are user-friendly. The study also challenges some traditional assumptions by showing the limited impact of perceived usefulness and trust. Despite their prominence in existing models, these factors showed insignificant influence in this context. This could indicate a shift in what drives user engagement, suggesting a need to re-evaluate and perhaps refine these constructs in e-commerce adoption models, considering the evolving nature of user expectations and experiences.

To enhance the adoption of fast-food e-commerce services, it is recommended that decision-makers focus on developing platforms that prioritize ease of use, UX, and product appeal rather than emphasizing perceived usefulness or trust. Practical suggestions include organizing regular UX audits, encouraging user feedback, and integrating design thinking to improve the platform's interface. Additionally, as the competitive landscape of fast-food e-commerce continues to evolve, integrating advanced technologies such as artificial intelligence and machine learning can further enhance user experience by personalizing service offerings. For instance, utilizing data analytics to understand consumer preferences could lead to tailored promotions that resonate more deeply with customers, thereby increasing engagement and loyalty.

There are some limitations in the data collected for this research. First, this study is conducted only in Egypt and caution should be made when generalizing the results to other contexts. In addition, this study does not account for the impact of digital marketing which plays a crucial role in shaping consumer preferences. To address these limitations, future research should consider supplementing quantitative data with qualitative methods, such as interviews or focus groups, to gain a deeper understanding of the underlying factors that influence the adoption of fast-food e-commerce platforms among college students.

## References

1. Rajput, S.: E-commerce the world of online shopping. Poonam Shodh Rachna. **3**(4), 71–74 (2024). https://doi.org/10.56642/brdu.v03i04.014
2. Revinova, S.: E-commerce effects for the sustainable development goals. SHS Web of Conferences. **114**, 01013 (2021). https://doi.org/10.1051/shsconf/202111401013
3. Papaioannou, E.: Factors affecting customers' perceptions and firms' decisions concerning online fast food ordering. Int. J. Agri. Environ. Info. Sys. (IJAEIS). **6**(1), 48–78 (2015)
4. Ye Zihan, Chee Wei Ming: Consumer E-commerce adoption and behavior: A perspective of technology acceptance model (TAM). J. Digitainability Realism & Mastery (DREAM). 2, 10, 11–33 (2023). https://doi.org/10.56982/dream.v2i10.159
5. Georgiou, M.N.: E-commerce has a positive impact on economic growth: A panel data analysis for Western Europe. SSRN Electron. J. **18**, 3 (2009). https://doi.org/10.2139/ssrn.1484687
6. Ju, C., et al.: Green logistics of fossil fuels and E-commerce: Implications for sustainable economic development. Res. Policy **85**, 103991–103991 (2023). https://doi.org/10.1016/j.resourpol.2023.103991
7. Davis, F.D.: Technology acceptance model: TAM. Info. Seeking Behavior Tech. Adop. **205**, 219 (1989)
8. Nihayah, N., Purnama, N.: Evaluation of digital banking application adoption based on the technology acceptance model (Tam). Int. J. Sci. Tech. Manag. **5**(2), 424–430 (2024). https://doi.org/10.46729/ijstm.v5i2.1083
9. Bryan, J.D., Zuva, T.: A review on TAM and TOE framework progression and how these models integrate. Adv. Sci. Technol. Eng. Syst. **6**, 3, 137–145 (2021). https://doi.org/10.25046/aj060316
10. Chen, C.-H.: Extending the technology acceptance model: A new perspective on the adoption of blockchain technology. Hum. Behav. Emerg. Techno. **2023**(1), 1–14 (2023). https://doi.org/10.1155/2023/4835896
11. Dah, J., Hussin, N.: A conceptual framework of a streamlined extended technology acceptance model for mobile application adoption. Int. J. Soc. Sci. Res. **9**(2), 42–42 (2021). https://doi.org/10.5296/ijssr.v9i2.18649
12. Carvajal-Trujillo, E.: Online recommendation systems: factors influencing use in e-commerce. Sustainability. **12**(21), 1–15 (2020)
13. Mailangkay, A. et al.: Analyzing factors influencing adoption of ai in e-commerce ui/ux design using TAM. In: IEEE 3rd International Conference on Social Sciences and Intelligence Management (SSIM). (2023). https://doi.org/10.1109/ssim59263.2023.10468833
14. Raza, S.A., Khan, K.A.: Corona fear and e-commerce adoption in an emerging economy: paradigm shift of consumer intention. Foresight. **24**(2), 195–209 (2021). https://doi.org/10.1108/fs-02-2021-0034
15. Kahar, A. et al.: The influence of perceived of usefulness, perceived ease of use, and perceived security on repurchase intention at tokopedia.com. In: Proceedings of the 2nd Padang International Conference on Education, Economics, Business and Accounting (PICEEBA-2 2018). (2019). https://doi.org/10.2991/piceeba2-18.2019.20.
16. Marianus, S., Ali, S.: Factors determining the perceived security dimensions in B2C electronic commerce website usage: An indonesian study. J. Account. Invest. **22**(1), 104–132 (2021). https://doi.org/10.18196/jai.v22i1.8171

17. Wafiyyah, R. et al.: The effect of perceived usefulness, perceived ease of use, and trust on repurchase intention on e- commerce shopee. IJISET–Int. J. Innov. Sci. Engin. Tech. **8**(7), (2021)
18. Butt, I.: Using technology acceptance model to study adoption of online shopping in an emerging economy. J. Internet Bank. Commer. **21**, 2 (2016)
19. Hussein, L., et al.: B2B e-commerce technology factors with mediating effect perceived usefulness in Jordanian manufacturing smes. J. Eng. Sci. Technol. **14**(1), 411–429 (2019)
20. Xie, H., et al.: The impact of e-commerce on environmental sustainability targets in selected European countries. Economic Research-Ekonomska Istraživanja. **36**(1), 230–242 (2022). https://doi.org/10.1080/1331677x.2022.2117718
21. Rasnayaka, S. et al.: An Empirical Study on Usage and Perceptions of LLMs in a Software Engineering Project. (2024)
22. Anaam, E.A. et al.: Analysis of perceived usefulness and perceived ease of use in relation to employee performance. Int. J. Membrane Sci. Tech. **10**, 2 (2023).
23. Savila, I.D. et al.: The role of multichannel integration, trust and offline-to-online customer loyalty towards repurchase intention: An Empirical Study in Online-to-Offline (O2O) e-commerce. Procedia Comp. Sci. **161**, 859–866 (2019)
24. Kumar, V. et al.: The impact of user experience design on customer satisfaction in e-commerce websites. Int. J. Sci. Techno. Eng. **11**(5), 4571–4575 (2023). https://doi.org/10.22214/ijraset.2023.52580
25. Purnomo, Y.J.: Digital marketing strategy to increase sales conversion on e-commerce platforms. J. Contemporary Admin. Manag. (ADMAN). **1**(2), 54–62 (2023). https://doi.org/10.61100/adman.v1i2.23
26. Suvattanadilok, M.: Factors influencing consumer behaviours via web personalization and information content on social media. Afr. J. Hospitality Tour. and Leisure **9**(1), (2020)
27. Liu, F., et al.: The art of appeal in electronic commerce. Internet Res. **27**(4), 752–771 (2017). https://doi.org/10.1108/intr-09-2016-0280
28. Peng, X., et al.: The effect of product aesthetics information on website appeal in online shopping. Nankai Bus. Rev. Int. **8**(2), 190–209 (2017). https://doi.org/10.1108/nbri-11-2016-0038
29. Liu, G., et al.: An empirical study on response to online customer reviews and e-commerce sales: from the mobile information system perspective. Mob. Inf. Syst. **2020**(1), 1–12 (2020). https://doi.org/10.1155/2020/8864764
30. Liu, Y., Tang, X.: The effects of online trust-building mechanisms on trust and repurchase intentions: An empirical study on eBay | Emerald Insight. Inf. Technol. PeoplePeople **31**(3), 666–687 (2016). https://doi.org/10.1108/itp-10-2016-0242
31. Kanojia, P., Lal, M.: Impact of trust on customer adoption of digital payment systems. In: Impact of Mobile Payment Applications and Transfers on Business. pp. 16–42. IGI Global (2020). https://doi.org/10.4018/978-1-7998-2398-8.ch002
32. Balaskas, S., et al.: FinTech services adoption in greece: the roles of trust, government support, and technology acceptance factors. FinTech. **3**(1), 83–101 (2024). https://doi.org/10.3390/fintech3010006
33. Junaidi, M.H., Miralam , M.S.: Online buying behavior among university students: A cross cultural empirical analysis. J. Info. Techno. Manag. **12**(2), (2020). https://doi.org/10.22059/jitm.2020.75789
34. Pepinsky, T.B.: The return of the single-country study. Annu. Rev. Polit. Sci. **22**(1), 187–203 (2019). https://doi.org/10.1146/annurev-polisci-051017-113314
35. De Vries, P.G.: Stratified random sampling. In: Sampling Theory for Forest Inventory: A Teach-Yourself Course. pp. 31–55 Springer Nature (1986). https://doi.org/10.1007/978-3-642-71581-5_2.
36. Molenaar, N.J., Smit, J.H.: Asking and answering yes/no-questions in survey interviews: a conversational approach. Qual. Quant. **30**(2), 115–136 (1996). https://doi.org/10.1007/bf00153983

37. Bernardi, R.A.: Validating research results when Cronbach'S Alpha is below .70: A methodological. Procedure. Educ Psychol Meas. **54**, 766–775 (1994). https://doi.org/10.1177/0013164494054003023. Last accessed 2024/08/01.
38. Wetherill, G.B. et al.: Multicollinearity. In: Regression Analysis with Applications. pp. 82–107 Springer Nature (1986). https://doi.org/10.1007/978-94-009-4105-2_4
39. MacKenzie, S.B.: The dangers of poor construct conceptualization. J. Acad. Market. Sci. **31**(3), 323–326 (2003). https://doi.org/10.1177/0092070303031003011?download=true&journalCode=jama. Last accessed 31 July 2024
40. Sathyanarayana, S. et al.: The effectiveness of e-recruitment adoption: A TAM approach examining user perspectives. SDMIMD J. Manag. **15**, 87–99 (2024). https://doi.org/10.18311/sdmimd/2024/41928
41. Choudhury, A., Shamszare, H.: Investigating the impact of user trust on the adoption and use of ChatGPT: Survey analysis. J. Med. Internet Res. **25**, e47184–e47184 (2023). https://doi.org/10.2196/47184