UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA CIVIL E AMBIENTAL
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL

JOSIVAN LEITE ALVES

BIM-DRIVEN ARTIFICIAL INTELLIGENCE SOLUTIONS FOR ENERGY EFFICIENCY: FROM THEORETICAL FRAMEWORK DEVELOPMENT TO PHOTOVOLTAIC ENERGY INTEGRATED PLANNING

Recife
2025

JOSIVAN LEITE ALVES

Doctoral thesis submitted to the postgraduate program in Civil Engineering at the Federal University of Pernambuco, Center for Technology and Geosciences, as a partial requirement for obtaining the title of Doctor in Civil Engineering, Area of Civil Construction.

Supervisor: Profa. Dra. Rachel Perez Palha

Co-advisor: Prof. Dr. Adiel Teixeira de Almeida Filho

Recife

2025

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Alves, Josivan Leite.

BIM-driven artificial intelligence solutions for energy efficiency: from theoretical framework development to photovoltaic energy integrated planning / Josivan Leite Alves. - Recife, 2025.

163f.: il.

Tese (Doutorado) - Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, Pós-Graduação em Engenharia Civil, 2025.

Orientação: Rachel Perez Palha.

Coorientação: Adiel Teixeira de Almeida Filho.

Inclui referências e apêndices.

1. Building Information Modeling; 2. Artificial Intelligence; 3. Solar Energy; 4. Sustainability; 5. Times series. I. Palha, Rachel Perez. II. Almeida Filho, Adiel Teixeira de. III. Título.

UFPE-Biblioteca Central

JOSIVAN LEITE ALVES

BIM-DRIVEN ARTIFICIAL INTELLIGENCE SOLUTIONS FOR ENERGY EFFICIENCY: FROM THEORETICAL FRAMEWORK DEVELOPMENT TO PHOTOVOLTAIC ENERGY INTEGRATED PLANNING

Tese apresentada ao Programa de Pós-Graduação em Engenharia Civil da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para obtenção do título de Doutor em Engenharia Civil, Área de Construção Civil.

Aprovada em 16/04/2025

Orientadora: Prof.ª Dr.ª Rachel Perez Palha, UFPE
Coorientador: Prof. Dr. Adiel Teixeira de Almeida Filho, UFPE

BANCA EXAMINADORA

participação por videoconferência
Prof. Dr. Maurício Oliveira de Andrade (examinador interno)
Universidade Federal de Pernambuco

participação por videoconferência
Prof.ª Dr.ª Vanessa Ribeiro Campos (examinadora externa)
Universidade Federal do Ceará

participação por videoconferência
Prof.ª Dr.ª Marcele Elisa Fontana (examinadora externa)
Universidade Federal de Pernambuco

participação por videoconferência
Prof. Dr. Fabrício Bradaschia (examinador externo)
Universidade Federal de Pernambuco

participação por videoconferência
Prof.ª Dr.ª Maria de Lourdes Florencio dos Santos (examinadora interna)
Universidade Federal de Pernambuco

AGRADECIMENTOS

A minha jornada durante o doutorado foi repleta de aprendizados, desafios e conquistas, e seria impossível chegar até aqui sem o apoio de pessoas importantes que mudaram minha vida e contribuíram para minha carreira.

À minha orientadora, Profª Rachel, a quem detenho grande admiração, expresso minha profunda gratidão pela orientação, pelas palavras encorajadoras e pelo exemplo de compromisso com a pesquisa acadêmica.

Ao Prof. Adiel, por suas palavras, me fizeram refletir sobre quem eu sou, o que almejo para o meu futuro e de que forma posso contribuir para a ciência brasileira. Ele me ajudou a compreender o impacto do meu trabalho na comunidade acadêmica e foi essencial para o meu amadurecimento.

À minha mãe, Sebastiana Leite, pelo amor imensurável, pelo exemplo de força e por nunca deixar de acreditar em mim. Aos meus irmãos, Eduardo e Maria, pelo companheirismo e pela fé.

Ao meu companheiro de vida, Selton, por estar ao meu lado em cada etapa desta jornada. Pela paciência, pelo carinho, pela atenção aos detalhes e, principalmente, por me apoiar nos momentos em que tudo parecia impossível.

Às minhas tias Maria e Zulmira, por serem exemplos de mulheres fortes, dedicadas e generosas. Levo comigo a força que aprendi com vocês.

Aos meus amigos da UFPE, Bruna, Evandro, Fernanda, Jéssica, Juan, Nakita e Thales, que me acolheram e me fizeram sentir em casa. Tudo foi mais leve com vocês: as risadas, o companheirismo e as conversas tornaram essa caminhada muito mais especial. Sou grato por cada momento vivido ao lado de vocês.

Aos meus amigos Alinny, Jair, Caio, Igor, Marcos e Jairo, que acompanham essa jornada desde cedo, quando tudo ainda era um sonho cheio de incertezas, conquistas por vir e histórias por viver. Pessoas como vocês me ajudaram a me tornar quem eu sou hoje, e sou imensamente grato por isso.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), ao Programa de Pós-Graduação em Engenharia Civil (PPGEC) e à própria UFPE. O ensino público de qualidade molda realidades, transforma vidas e eu sou o exemplo vivo disso.

"*Se vi mais longe, foi por estar sobre os ombros de gigantes*".

Isaac Newton, em uma carta para Robert Hooke (1676).

RESUMO

O setor de Arquitetura, Engenharia, Construção e Operações (AECO) tem obtido vantagens com a geração e o gerenciamento de dados por meio do Building Information Modeling (BIM). Esse aumento de dados disponíveis pode ser fundamental para impulsionar a inovação ao ser analisado com modelos de Inteligência Artificial (IA). Nesse cenário, a busca por projetos sustentáveis inteligentes tem se consolidado como estratégia para promover o desenvolvimento sustentável e reduzir a dependência de recursos não renováveis. A relevância do tema se evidencia pelo fato de os edifícios consumirem cerca de 40% da energia global e serem responsáveis por 33% das emissões de gases de efeito estufa, em grande parte devido à baixa eficiência energética. Diante desse problema, esta pesquisa tem como objetivo desenvolver soluções automatizadas BIM orientadas por IA para apoiar o planejamento energético de edifícios sustentáveis, com foco na simulação e otimização de sistemas fotovoltaicos tanto em projetos novos quanto em retrofit. A investigação adota uma abordagem multimétodo, combinando revisão sistemática de literatura, desenvolvimento de algoritmos de aprendizado profundo e implementação de processos automatizados de modelagem em BIM. Primeiramente, foi realizado um mapeamento da integração entre BIM e IA, identificando domínios de aplicação, problemas abordados, resultados alcançados e capacidades fundamentais de ambas as tecnologias. Em seguida, foi proposto e testado um algoritmo de aprendizado profundo orientado por BIM para estimar a produção de energia fotovoltaica, relacionando séries temporais de irradiação solar com dados extraídos automaticamente de modelos BIM. Os resultados demonstram que a abordagem proposta (denominada EnergyBIM.AI) possibilita a quantificação automática da produção de energia solar e da redução potencial de emissões de CO_2. Além disso, a pesquisa propõe um processo para apoiar a seleção e posicionamento de painéis solares em modelos BIM. Assim, esta tese evidencia que a integração entre BIM e IA pode transformar o planejamento energético no setor AECO, oferecendo processos automatizados capazes de aumentar a eficiência energética, reduzir emissões e apoiar o design e retrofit de edifícios sustentáveis. As principais contribuições desta são: (i) oferecer um framework de integração entre BIM e IA aplicado a projetos inteligentes; (ii) propor uma abordagem para previsão energética baseada em aprendizado profundo e automação em BIM; e (iii) fornecer evidências para apoiar o design e o retrofit de edifícios sustentáveis.

Palavras-chave: Modelagem da Informação da Construção. Inteligência Artificial. Energia Solar. Sustentabilidade. Séries temporais.

ABSTRACT

The Architecture, Engineering, Construction, and Operations (AECO) sector has benefited from data generation and management through Building Information Modeling (BIM). This increased availability of data can be fundamental for driving innovation when analyzed with Artificial Intelligence (AI) models. In this context, the pursuit of smart sustainable projects has become a strategy to promote sustainable development and reduce dependence on non-renewable resources. The relevance of this topic is evident in the fact that buildings consume approximately 40% of global energy and are responsible for 33% of greenhouse gas emissions, largely due to low energy efficiency. Given this problem, this research aims to develop automated BIM-driven AI solutions to support the energy planning of sustainable buildings, focusing on the simulation and optimization of photovoltaic systems in both new and retrofit projects. The investigation adopts a multi-method approach, combining a systematic literature review, the development of deep learning algorithms, and the implementation of automated BIM modeling processes. First, a mapping of the integration between BIM and AI was performed, identifying application domains, problems addressed, results achieved, and fundamental capabilities of both technologies. Next, a BIM-driven deep learning algorithm was proposed and tested to estimate photovoltaic energy production, relating solar irradiation time series with data automatically extracted from BIM models. The results demonstrate that the proposed approach (EnergyBIM.AI) enables the automatic quantification of solar energy production and the potential reduction of CO_2 emissions. Furthermore, the research proposes a process to support the selection and placement of solar panels in BIM models. Thus, this thesis demonstrates that the integration of BIM and AI can transform energy planning in the AECO sector, offering automated processes capable of increasing energy efficiency, reducing emissions, and supporting the design and retrofit of sustainable buildings. The main contributions of this thesis are: (i) offering a framework for integrating BIM and AI applied to smart projects; (ii) proposing an approach for energy forecasting based on deep learning and automation in BIM; and (iii) provide evidence to support the design and retrofit of sustainable buildings.

Keywords: Building Information Modeling. Artificial Intelligence. Solar Energy. Sustainability. Times series.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AI	Artificial Intelligence
BIM	Building Information Modeling
FNN	Feedforward neural networks
GIS	Geographic Information System
HVAC	Heating, Ventilating and Air Conditioning
IBGE	Brazilian Institute of Geography and Statistics
IFC	Industry Foundation Classes
INMET	National Institute of Meteorology
IoT	Internet of Things
LiDAR	Light Detection and Ranging
LoD	Level of Development
LSTM	Long Short-term Memory
MEP	Mechanical, Electrical, Plumbing
MLP	Multilayer Perceptron
PV	Photovoltaic
SLR	Systematic Literature Review
XGBoost	Extreme Gradient Boosting

SUMMARY

1 INTRODUCTION

Urbanization and population growth rates have increased global energy consumption and emissions. Globally, the most recent reports identified that the main areas of energy consumption are industry, transport, and buildings (Di Giovanni *et al.*, 2024). Specifically for buildings, the need arises to develop sustainable alternatives that benefit the environment, economy, and society, as the architecture, engineering, construction, and operations (AECO) industry adapts to emerging technologies. Therefore, this research promotes a sustainability-oriented approach by discussing the development of an energy-oriented framework through the promotion of analyses and simulations of photovoltaic (PV) energy production in buildings (Olu-Ajayi *et al.*, 2022; Shao *et al.*, 2021; Wang *et al.*, 2023). This thesis proposes an integrated approach to the AECO industry through three research methods.

First, the AECO industry's practitioners and academic community recognize and apply the potential benefits generated using Building Information Modeling (BIM). BIM is considered a data-driven modeling, simulation, and information management environment for projects. Digital representations provide essential information for the various phases of projects, particularly useful in the design phase for simulations and building performance predictions (Alves *et al.*, 2025). 3D BIM models incorporate a series of geometric data, and these models evolve to other dimensions as new information is integrated. Particularly in the design phase, BIM offers the possibility of implementing, in addition to 3D modeling, scheduling (4D), cost estimation (5D), as well as aspects of sustainability assessment (6D). Design tasks that can be integrated into BIM include energy performance analysis, CO_2 emission analysis, solar and light simulation, thermal comfort analysis, and waste management (Scherz *et al.*, 2022). Therefore, it is recognized that BIM models can form databases that can be used proactively for performance analysis and performance measurement of buildings.

Second, the digital representation of buildings, a central feature of BIM, enables the collection and storage of essential data for automated design processes in the AECO industry. With the use of Artificial Intelligence (AI), particularly through deep learning techniques applied to time series, it is possible to identify historical performance patterns, predict failures, and optimize the operation of building systems. Models such as recurrent neural networks (RNNs) and long short-term memory networks (LSTMs) are examples for dealing with sequential and temporal data, facilitating, for example, the prediction of energy consumption or predictive maintenance of systems. In a BIM environment, these models can integrate information throughout the building's life cycle, optimizing design and operation phases (Baduge *et al.*, 2022; Pan; Zhang, 2021; Tao *et al.*, 2024).

Third, this thesis identified a research trend in sustainability-oriented BIM-AI applications and energy simulations (see Alves *et al.*, 2025). Studies evaluate applications of BIM-AI models to optimize building parameters such as orientation and material properties to reduce energy costs and emissions. Furthermore, the use of AI has expanded to many areas, including predicting building energy consumption. Despite extensive research on the application of AI in buildings, little research explores the combined use of BIM models with AI to optimize the use of renewable energy (Mehraban *et al.*, 2024). For example, previous research has advanced the body of knowledge, seeking automated alternatives to improve the energy efficiency of buildings. Alawi *et al.* (2024) developed predictive simulations for residential buildings' annual heating and cooling loads. Olu-Ajayi *et al.* (2022) seek to predict energy consumption in the building design phase. Chou *et al.* (2017) developed a BIM data fusion process on energy consumption datasets collected. Li *et al.* (2024) propose an adaptive sea lion-optimized genetic adversarial to predict renewable energy sources. Tao *et al.* (2024) apply tree-based, linear, and non-linear regression techniques to predict the energy and exergy efficiency of Parabolic Trough Solar Collectors using oil-based nanofluids. However, Tian *et al.* (2023) highlight a growing field in the literature related to predicting solar radiation and PV energy production using time series data. Also, there is a gap in the literature addressing BIM-driven renewable energy solutions, especially with the association of AI algorithms.

Thus, this thesis applies a multimethod approach to develop energy BIM-AI solutions for the AECO industry. In the first stage, based on an extensive literature review, the thesis presents an integrative research model, which explores the main insights of the association between BIM and AI in AECO projects (Chapter 4). The research design is a systematic literature review that applies bibliometrics and content analysis with the assistance of Bibliometrix and Mendeley software. The main topics, thematic evolution, and concept maps are covered. Finally, using a coding scheme in the content analysis stage, the article explores the relationships between AI and BIM applications, essential capabilities for this integration, and potential benefits for the sector by proposing a framework (see Alves *et al.*, 2025).

In the second stage, this thesis explores the application of a BIM-driven deep learning algorithm to estimate PV energy production, associating solar irradiation data and automated extraction of information in BIM models (Chapter 5). This paper quantifies the energy produced and CO_2 based on the predicted values of the implemented algorithm, using a routine in Dynamo that extracts the information from a BIM model. For these applications, the paper tests three deep learning algorithms: Long Short-Term Memory (LSTM), Extreme Gradient Boosting (XGBoost), and Feedforward Neural Network (FNN). It also offers general guidelines for BIM modelers to consider implementing PV modules as early as the Design phase of buildings, establishing a means to maximize efficiency in PV energy production and support feasibility studies for implementing PV modules.

The third stage aims to develop an automated allocation of PV modules to maximize PV energy production (kWh/day) while minimizing implementation costs (Chapter 6). This research integrates visual programming in Dynamo with programming in Python to analyze different combinations of PV modules, considering the dimensions of 21 PV modules from 4 brands for allocation on the roof of a building. The algorithm identifies the most efficient configuration of PV cost-production. It uses Dynamo to extract information on the families of PV modules and the available roof area from a BIM model in Revit. Finally, the model automatically allocates the best arrangement of PV modules directly in the Revit model.

1.1 Thesis Justification and Motivations

Buildings consume 40% of global energy and 33% of greenhouse gas emissions (Asif *et al.*, 2024). Furthermore, the low energy efficiency of buildings is one of the main factors contributing to high global energy consumption and greenhouse gas emissions. In recent decades, the growing demand for energy in buildings has been driven by population growth and the rapid expansion of urban areas (Olu-Ajayi *et al.*, 2022). In this context, buildings are strategic ventures for the global transition towards sustainability, against climate change, and in favor of the zero-carbon energy transition, especially in the planning and design phases. To this, technologies such as BIM and AI, which enable the collection and analysis of data to improve the physical environment and project modeling, offer new alternatives for the sustainable digital revolution (Asif; Naeem; Khalid, 2024). Therefore, this thesis argues that using PV energy provides fertile paths to promoting sustainability in buildings, and the available technologies can enhance the planning of sustainable strategies in the design phase of BIM models.

As a primary renewable energy source, PV energy is used due to its potential to meet the growing demand for energy and the limited fossil fuel resources on the planet emissions (Asif et al., 2024). With the support of national policies, renewable energy targets, and the falling costs of PV modules, the PV energy market has experienced growth. Consequently, an increase in the number of buildings integrated with PV systems has been observed (WANG *et* *al.*, 2023). In this context, the forecast of PV energy production is one of the strategies in the design phase for developing sustainable projects (Tian *et al.*, 2023). Furthermore, energy simulation models must be able to process a large amount of data for different climate conditions. This allows designers to adopt alternative design solutions for bioclimatic building elements.

In the context of the AECO industry, emerging BIM tools and technologies have transformed the generation, storage, and exchange of project information. The development of Industry Foundation Classes (IFCs) has facilitated more integrated methods for exchanging construction data, thereby enhancing interoperability across various systems. However, with the rapid advancement of digital technologies such as AI, including data analytics, machine learning, and deep learning, the imperative for integrating BIM with AI has become a growing topic in the research field. This convergence promises to optimize processes further, enhance decision-making, and drive innovation in construction practices. AI offers the ability to process large volumes of data, identify patterns, and generate accurate predictions, extending the capabilities of BIM in areas such as process automation, generative design, and resource optimization (Boje *et al.*, 2020). In this context, literature opens new paths to develop automated systems, processes, and models based on BIM, taking advantage of digital advances in AI to improve efficiency, innovation, and quality in the AECO industry.

Due to the vast amount of data generated by BIM models, the storage volume and complexity of this data have brought new challenges to the design review process, resulting in a reliance on manual methods to deal with the current intricate structure of models. This approach creates problems such as high demand for specialized professional skills and low process efficiency (Li *et al.*, 2024). Moreover, in many cases, analyzing large volumes of data and recognizing patterns using traditional programs or rule-based methods has proven to be impractical. In this context, AI has the capacity to process vast amounts of data, identify complex patterns, and construct large-scale statistical models (Baduge *et al.*, 2022).

Recent research shows that this combination has the potential to transform the way construction schedules, costs, quality, and cybersecurity are managed. AI can explore and learn from BIM data and provide solutions for architects, engineers, manufacturers, and other stakeholders. Furthermore, AI can support the generative design process by creating multiple alternatives based on defined parameters and constraints. For manufacturers, the integration of AI and BIM makes it possible to digitize their product catalogs, predict market demands and trends, optimize production processes, and offer customized solutions to customers (Zawada *et al.*, 2024).

Therefore, research applied to the AECO industry can lead to actions to address the consequences of climate change through sustainability initiatives in smart projects. The focus is to seek greater energy and resource efficiency through digitalizing and automating projects. In this context, the research community is actively studying the applications of digital technologies to improve the sustainability of buildings. The digital revolution can increase energy efficiency through technologies that collect and analyze data to improve the physical environment (Asif; Naeem; Khalid, 2024).

This research seeks to facilitate sustainability assessments through clean energy production analyses, specifically in the field of BIM and AI integration. While BIM is used as the input data for energy assessments, AI is applied to predictive modeling oriented towards energy management.

While the body of knowledge related to photovoltaic projects in the AECO industry explores alternatives to enhance PV energy production through the application of different materials for PV modules (Myint *et al.*, 2025; Serat *et al.*, 2025; Zhi *et al.*, 2023), the literature on BIM in sustainable projects seeks to propose solutions in the various dimensions of the framework (Cao; Huang, 2023; Cassandro *et al.*, 2024; Lins *et al.*, 2024; Mandičák *et al.*, 2024). However, there is a gap in the literature related to integrating PV projects and BIM technologies for simulations in three-dimensional models during the project design phase. For example, Myint *et al.* (2025) discusses the importance of research on using technologies such as BIM to drive the development of automated projects that evaluate different PV design options. Myint *et al.* (2025) argues that articles about BIM-PV can guide the development of low-carbon buildings. However, research in the field of PV systems, such as Serat *et al.* (2025) and Zhi *et al.* (2023), has advanced mainly in simulations of solar energy production, but often without considering the direct integration of photovoltaic systems into the building envelope, as proposed by building PV systems.

This gap is also reflected in the lack of integration between PV simulation tools and BIM-based building design processes, causing interoperability problems and limiting automated and data-driven design approaches (Di Giovanni *et al.*, 2024; Palha *et al.*, 2024; Zhi; *et al.*, 2023). In addition, although recent research in BIM has focused on automated solutions that address sustainability requirements (Cao; Huang, 2023; Cassandro *et al.*, 2024), there is still a lack of studies proposing comprehensive workflows that integrate BIM-PV design, energy simulations, and BIM technologies from the early design stages (Lu *et al.*, 2022; Mandičák *et al.*, 2024).

The AECO can contribute to the achievement of the United Nations' 2030 Agenda, given its direct impact on global energy consumption and greenhouse gas emissions. In this way, the motivation for this research is also the alignment with three Sustainable Development Goals (SDGs). Affordable and Clean Energy (SDG 7) drives the development of technological solutions that expand renewable energy generation through the integration of photovoltaic systems in buildings. Sustainable Cities and Communities (SDG 11) guides the creation of innovative processes that contribute to more resilient, efficient, and environmentally responsible urban environments. Climate Action (SDG 13) reinforces the urgency of reducing CO_2 emissions, a target directly addressed through the use of AI and BIM for energy simulation in buildings. Therefore, the motivation of this thesis arises from the need to align digital transformation and sustainability, exploring how the integration of BIM and AI can support intelligent energy planning, reduce environmental impacts, and accelerate the transition toward low-carbon buildings.

1.2 Research Questions and Hypothesis

- **(#RQ1)** What are the essential BIM and AI capabilities for smart AECO projects? **H1:** The integration of BIM capabilities with AI capabilities impacts the development efficiency of smart projects in the AECO sector.
- **(#RQ2)** What are the main benefits of the connection between BIM and AI for developing smart AECO projects? **H2:** Smart AECO projects can be developed through integrated BIM and AI capabilities due to the large volume of data that requires higher levels of automation in decision-making than traditional projects.
- **(#RQ3)** What is the potential of integrating deep learning models and BIM automation tools to support early-stage PV system planning and energy performance assessment? **H3:** Integrating deep learning algorithms with automatic data extraction routines in BIM models can estimate photovoltaic energy production and avoid CO_2 emissions even at the design stage.
- **(#RQ4)** How to automate the allocation of photovoltaic panels to maximize energy production and minimize building implementation costs? **H4:** The application of a process model integrated with BIM can automatically identify photovoltaic module configurations with greater energy efficiency and lower implementation costs, considering the roof's physical restrictions and the modules' technical characteristics.

1.3 Objectives

1.3.1 Main Objective

This thesis aims to investigate how the integration of BIM and AI applications can foster a BIM-driven design process to support the planning and retrofit of sustainable buildings, with a specific focus on photovoltaic systems in the AECO sector.

1.3.2 Specific Objectives

a) Investigate how BIM and AI capabilities can improve the development of smart architecture, engineering, construction, and operation projects.
b) Explore the application of a BIM-driven deep learning algorithm to estimate PV energy production, associating solar radiation time series and automated extraction of information in BIM models.
c) Develop an automated process model for allocating photovoltaic modules to maximize photovoltaic energy production (kWh/day) while minimizing implementation costs.

1.4 Relevance of the Research Problem to the AECO Industry

The chapters of this thesis guide the main contributions of integrating BIM and AI. The findings show how these emerging technologies evolve the AECO sector and how the industry can position itself in the face of digital transformation. In addition, the results guide the development of sustainable projects and PV strategies to establish theoretical frameworks and automated processes. Thus, the results contribute in three ways.

1.4.1 Contribution to Academic Field

First, the thesis advances the theoretical field by articulating BIM and AI applied to developing smart and sustainable projects in the AECO sector. Through mapping organizational capabilities, the thesis structures a conceptual framework that integrates BIM and AI as essential dynamic capabilities for generating value throughout the project life cycle. This theoretical framework fills gaps in literature by operationalizing how these technologies complement each other in the production of data, automation of processes, and generation of solutions oriented to sustainability.

In addition, the thesis contributes to advancing the discussion on BIM 6D by presenting a predictive and automated process to estimating renewable energy generation and carbon footprint emissions while still in the project design phase. This theoretical approach broadens the understanding of the role of BIM in data-driven sustainable projects. In addition, a process is proposed to measure energy efficiency relating to PV energy production and implementation cost for selecting PV module manufacturers and models.

1.4.2 Contribution to Managers and Professionals

Second, from a practical perspective, the thesis offers solutions for AECO professionals working with sustainable projects, renewable energy, and digitalization. First, it guides the identification and development of essential organizational capabilities, encouraging investments in data science and multidisciplinary practices. It develops an automated process for forecasting solar energy generation and calculating avoided CO_2, using time series and deep learning, applicable to different regions and climate scenarios, which supports energy and environmental feasibility analyses in new projects and retrofits. By automating the allocation of photovoltaic modules in Revit and proposing comparisons between different layouts and brands, the thesis offers designers, engineers, and managers a replicable method for technical and economic decision-making. This contributes to preparing technical documents supporting credit lines and tax incentives for sustainable buildings.

1.4.3 Contribution to Policy Makers

Third, the growing pressure for decarbonization, energy efficiency, and digital transformation has redefined the priorities of the AECO industry and how society attributes value to urban development. In this scenario, integrating BIM and AI becomes a strategic necessity. However, despite the theoretical recognition of the potential of these technologies, there is still a gap between their availability and their effective application in design practices. This research addresses this challenge by focusing on one of the urgent demands of the sector: designing efficient and functional buildings capable of generating renewable energy and reducing environmental impacts from the design phase. By proposing the integration of BIM and photovoltaic systems, the research offers practical answers to global challenges, such as the goal of near-zero energy buildings, environmental certifications, and the construction of climate-resilient infrastructure.

Thus, by demonstrating the feasibility of integrating BIM and AI for energy performance prediction, the research provides a technical basis to encourage the adoption of digital technologies in public and private projects. The possibility of quantifying renewable energy generated and carbon footprint, even at the design stage, creates opportunities for developing incentive mechanisms, such as credit lines, national environmental certifications, and tax exemptions for projects with proven energy efficiency. In addition, the methodological approach can serve as a reference for regulatory guidelines that promote the use of digital tools in urban planning and public infrastructure, contributing to the achievement of climate goals, such as those set out in the 2030 Agenda and Brazil's energy transition commitments.

1.5 Thesis Structure

Figure 1 shows the thesis structure, which was developed through a collection of papers. This structure seeks to contribute to the research field while receiving reviews from target journals during the development process. This ensures that the thesis has peer-reviewed contributions and is validated through phased contributions.

Chapter 1 introduced the general aspects of the Thesis contextualization and motivation, as well as the objectives, hypotheses, and research questions.

Chapter 2 provides the basis for understanding the theorizing process. It discussed the problems and characteristics of the AECO industry and how innovation has transformed decision-making processes. It then delves deeper into the context of BIM and AI, focusing specifically on the context of implementing solar energy solutions.

Chapter 3 presents an overview of the Thesis's development. This chapter only explores in general terms how the subsequent chapters were structured and how their contributions are interconnected.

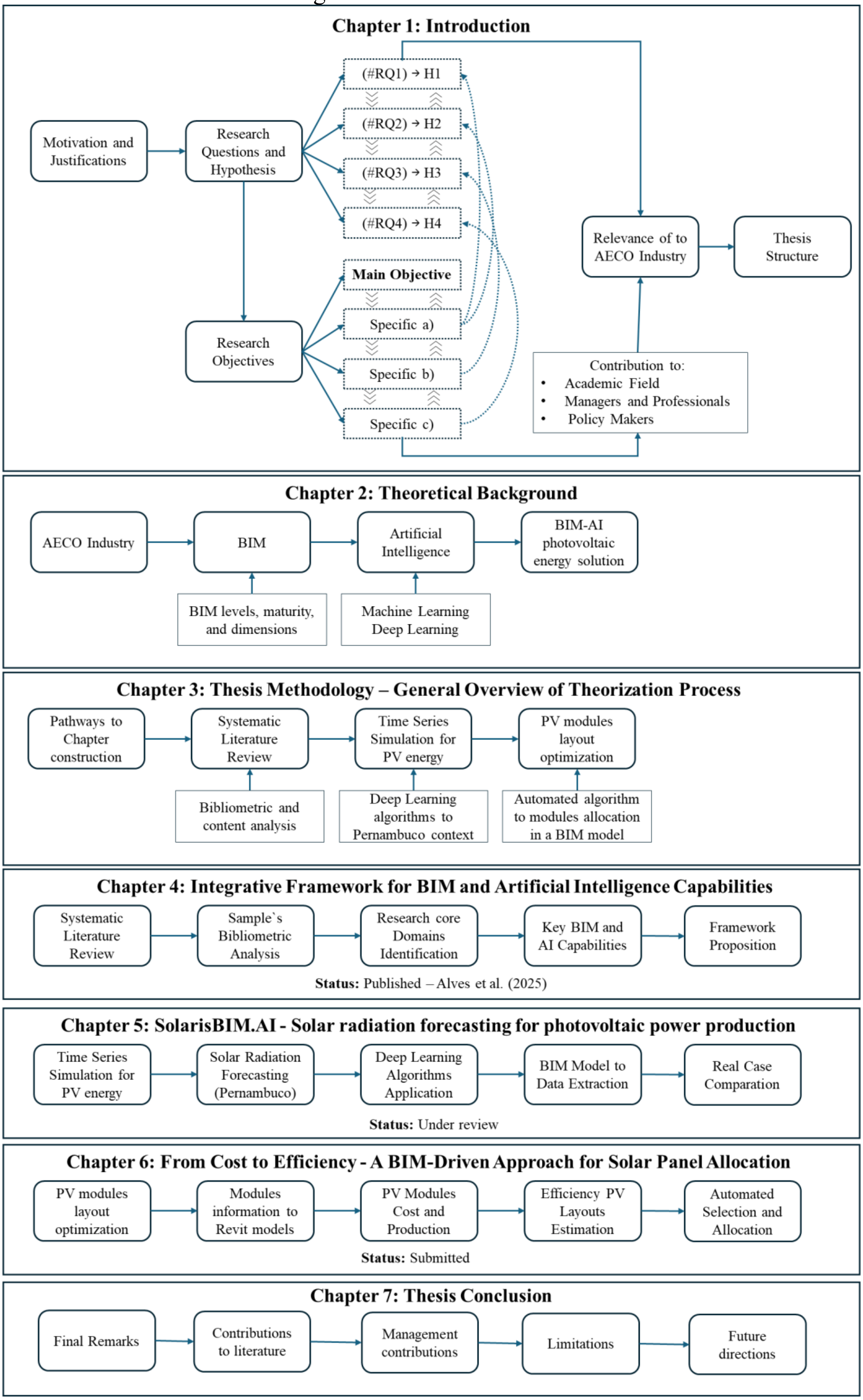
Chapter 4 discusses the advances in integrating BIM and AI through a Systematic Literature Review. This chapter identifies seven key core domains of this integration. Based on these key integration domains, this chapter attests that a research trend related to developing research oriented toward sustainability and renewable energies exists. Thus, the following chapters focus specifically on the context of PV energy production in buildings. This chapter generated the first article of the thesis, **published** in the Automation in Construction (Impact Factor = 9.6) (Alves *et al.*, 2025).

Chapter 5 develops a process for predicting PV energy production. This process is called SolarisBIM.AI, and deep learning algorithms are applied to extract information from a BIM model in Revit. This chapter generated article 2, which was **under review** in the Journal of Construction Engineering and Management (Impact Factor = 5.1).

Chapter 6 seeks to improve SolarisBIM.AI. This chapter focuses on selecting the best alternative for allocating solar modules, taking the cost and PV energy production of different PV module manufacturers as variables. This chapter generated the third article of the thesis, which was **submitted** to Building and Environment (Impact Factor = 7.1)

Finally, **Chapter 7** and **Chapter 8** highlight the main contributions of the thesis chapters and explore their relationship.

Figure 1 - Thesis Structure



2 THEORETICAL BACKGROUND

2.1 Architecture, Engineering, Construction, and Operations Industry

The academic research field of the Architecture, Engineering, Construction, and Operations (AECO) industry develops advanced techniques and computing technologies to address industry challenges such as low productivity and project delivery, poor performance, and ineffective resource management (Rangasamy; Yang, 2024).

Historically, the AECO industry has been quite resistant to change due to the difficulty of adopting new technologies. In addition, the industry faces challenges such as uncertainty, complexity, fragmented supply chains, and cultural barriers. The high level of project uncertainty compounds this complexity due to the unpredictable project environment (Wang, 2024). Furthermore, buildings and structures differ in types of use (e.g., residential, commercial, municipal, infrastructural), in age (e.g., new, existing, heritage), and in ownership (e.g., private owner, housing association, public owner, universities). These different framework conditions influence the application of BIM, its level of detail (LoD), and its supporting functionalities in the design, construction, maintenance, and deconstruction processes due to stakeholder requirements (Volk; Stengel; Schultmann, 2014).

In the AECO sector, building performance involves a variety of functions that buildings need to meet, such as energy efficiency, indoor environmental quality, thermal and visual comfort, and well-structured strategies for ongoing building maintenance. Several software tools are available for performance analysis and simulation, allowing these assessments to be carried out at the design stage. However, most of these tools require detailed definitions of aspects such as the structure's geometry, the layout of spaces, furniture, and mechanical, electrical, and plumbing (MEP) systems. These elements are accompanied by domain-specific parameters that influence their functionality, such as the thermal properties of building components, use and operation schedules, and internal lighting and equipment loads (Utkucu *et al.*, 2024).

Researchers have developed studies for the integration of AI and BIM in the AECO industry to solve problems that cover asset management and maintenance, construction problem solving, construction management, sustainable development, safety enhancement, building information management, and infrastructure management (Wang, 2024).

2.2 Building Information Modeling

The various projects in the AECO industry require data to be exchanged between interdisciplinary teams throughout the project lifecycle, such as updates, verification, and validation. In this context, BIM, as an element for data-driven element management, includes geometric and semantic data and is an essential tool (Rogage; Doukari, 2024). BIM can be defined as a process that involves the creation and management of a digital representation of the physical and functional characteristics of a building or infrastructure, serving as a shared source of information that supports decision-making throughout the project lifecycle, from conception to demolition. It integrates designing, constructing, or operating a building or infrastructure asset using electronic object-based information (Huang; Ninić; Zhang, 2021).

In this thesis, BIM is conceptualized as Building Information Model (the artifact) and Building Information Modeling (the process), as the literature and software developers recommended. The distinction between Building Information Model and Building Information Modeling is that the first refers to an intelligent, parametric, object-oriented, data-rich digital representation of a facility. In contrast, Building Information Modeling is a process that involves the creation and management (development and use) of this digital representation of the physical and functional characteristics of a building throughout its life cycle. BIM emerged from the technique of parametric object-oriented modeling, where the term "parametric" describes a process in which the modification of one element results in the automatic adjustment of adjacent or dependent elements (such as a door attached to a wall) to maintain a previously established relationship (Volk; Stengel; Schultmann, 2014).

BIM allows professionals to perform various analyses, such as design option comparison, cost estimation, automated code checking, clash detection, safety management, and construction simulations. Recent integration with AI has fueled research into using and operating data-driven BIM software. Technologies such as Generative Adversarial Networks (GANs) and generative language and image models, including DALL·E 2, Stable Diffusion, and Midjourney, have been applied to explore initial design alternatives, generate design images and floor plans, and predict building performance. Specifically, technology that enables the creation of schematic images from natural language commands has reached a commercial level with image diffusion-based approaches, such as those offered by Veras and SketchUp Diffusion tools, which can be directly integrated with BIM solutions (Jang *et al.*, 2024).

In this context, information exchange is done through open data schemas, mainly IFC. Industry Foundation Classes (IFC) are an open and standardized data format created to facilitate interoperability and information sharing between systems and software in the AECO industry. Developed by the buildingSMART organization, IFC is adopted in BIM to represent information related to the physical and functional aspects of a building throughout its life cycle. It allows different professionals and platforms to exchange and use project data, regardless of which tools they are using (Boje *et al.*, 2020; Huang; Ninić; Zhang, 2021).

2.2.1 BIM levels, maturity, and dimensions

BIM functionalities depend on the accuracy, richness, and timeliness of the underlying data to achieve their goals. A widely used concept to describe the amount of information contained in BIM objects is the "Level of Detail" or "Level of Development" (LoD). The LoD specifies a model component's geometric and non-geometric attributes, often associated with a specific point in time or a contractual responsibility. In the case of buildings, for example, for analysis and programming, it is necessary to define the appropriate LoD for the attributes and relationships of the objects, including durations, dependencies, and precedence information. In the literature, different levels of LoD vary in terms of geometric accuracy, quality, and completeness of the semantic information provided (Volk; Stengel; Schultmann, 2014).

In new construction projects, the LoD increases as the project progresses, from the initial stages to the production/construction phase, following the requirements and refinements throughout the process. The literature defines different LoDs for various functionalities, such as general modeling, 3D imaging, and energy performance. In the case of existing buildings, the necessary LoD is determined by the required functionality, which directly influences the costs and effort involved in creating the BIM model. For maintenance operations, the Construction Operations Building Information Exchange (COBie) standard defines a specific LoD for technical equipment, covering information such as type and location, make, model, serial numbers, label, installation date, warranties, and scheduled maintenance requirements. However, no adequate LoD is established for functionalities related to deconstruction and waste management (Volk; Stengel; Schultmann, 2014).

BIM can be classified into different levels, from Level 0 to Level 3, according to the degree of maturity of its development. Level 0 refers to unmanaged CAD, usually in 2D, with the exchange of information occurring through paper or electronic documents without following standardized processes. At Level 1, CAD is managed and can be in 2D or 3D, using collaboration tools that provide a shared data environment for electronic information sharing.

Much of the market still operates in Level 1 processes, while industry leaders are already beginning to reap the benefits of implementing Level 2 (Gao; Koch; Wu, 2019).

Level 2 BIM is characterized by a managed 3D environment, where data is maintained in separate BIM tools, with information attached to the models. This level is distinguished by collaboration between stakeholders, who share information through a standard file format, although each still works on its own CAD models. Finally, Level 3, known as Integrated BIM (iBIM), represents a complete integration of data and processes, enabled by web services compatible with emerging standards such as IFC, and managed by a collaborative model server. Level 3 creates a fully integrated design environment where all disciplines collaborate using a single shared project model, maintained in a centralized repository that allows access and modification of the model by all stakeholders (Gao; Koch; Wu, 2019).

In addition to the level of detail and maturity, BIM is classified into dimensions. The third dimension (3D) refers to geometric modeling, which is the basis of the process, digitally representing the physical form of a building. This includes creating detailed three-dimensional models of the project's structural and architectural components. From this 3D base, the subsequent dimensions — 4D (time), 5D (costs), 6D (sustainability), and 7D (facilities management), are integrated to provide analysis throughout the life cycle of the building. The fourth dimension (4D) introduces time, allowing the visualization and analysis of aspects of the project, such as graphical models, management, costs, resources, and safety — throughout the schedule. The fifth dimension (5D) integrates cost estimation, providing detailed financial control throughout the project's life cycle. The sixth dimension (6D) is related to sustainability, allowing the analysis of energy performance and the environmental impact of buildings. The seventh dimension (7D) focuses on facilities management, using model data to optimize operation, maintenance, and asset management throughout the building's lifespan (Alzara *et al.*, 2023; Boje *et al.*, 2020; Pan; Zhang, 2023).

These theoretical classifications guide researchers and professionals in following the market evolution, research, and companies. The main BIM applications include optimizing the orientation of buildings to reduce energy consumption; analyzing the mass of the building to evaluate its shape and optimize the envelope, such as the proportion of transparency; analyzing natural lighting; assessing the potential for water collection; energy simulation using tools such as Green Building Studio; evaluating the suitability of sustainable materials, with an emphasis on minimizing use and adopting recycled materials; and planning logistics and site management to reduce waste and carbon emissions (Gao; Koch; Wu, 2019).

For example, Azhar *et al.* (2011) subdivided BIM applications in the design phase into three main stages: in schematic design, BIM enables comparative analysis of various design options and integration of photorealistic images with existing conditions; in detailed design, BIM tools allow the creation of internal and external 3D models, walk-through animations, and energy performance simulations; and in construction detailing, BIM assists in 4D phasing and scheduling planning, detection of construction system conflicts, and generation of manufacturing drawings (Azhar, 2011). These applications benefit all key project team members, including the project manager, architect, structural engineer, mechanical engineer, electrical engineer, civil engineer, quantity surveyor, and construction manager (Gao; Koch; Wu, 2019).

2.3 Artificial Intelligence

Artificial intelligence (AI) is the intelligence demonstrated by machines in contrast to the natural intelligence exhibited by humans. Founded as an academic discipline in 1956, AI has experienced rapid growth in recent years due to advanced computing technologies, extensive data capabilities, and theoretical understanding. The field of AI draws on computer science, mathematics, psychology, linguistics, philosophy, and many other fields (Song *et al.*, 2022). It is one of the emerging research topics in various scientific fields (Zhu *et al.*, 2022).

AI is the science that seeks to develop machines or computer programs capable of replicating human intelligence. AI has advanced in computer vision, robotics, autonomous vehicles, language translation, gaming, and medicine (Baduge *et al.*, 2022). In computer science, AI research is defined as the study of intelligent agents, which are systems capable of perceiving their surroundings and performing actions to maximize the probability of achieving their goals effectively. Traditionally, AI research has focused on computational technologies that allow machines to mimic human cognitive functions, such as learning and problem-solving, seeking to replicate or enhance typical human mental capabilities (Song; Xu; Zhao, 2022).

Unlike computer scientists, who focus primarily on developing and improving AI techniques, information systems researchers apply these techniques to solve business problems and improve management. AI has been widely used in various areas, including decision support and expert systems, knowledge management, financial forecasting, systems design, technological innovation, supply chain management, and big data analytics. Studies have demonstrated the benefits of AI applications in finance, healthcare, and marketing industries, highlighting their positive impact in multiple contexts (Song; Xu; Zhao, 2022).

Data science is a broad field that involves collecting, analyzing, and interpreting large volumes of data to solve complex problems and inform decisions. Within this context, AI is a subfield that focuses on creating systems capable of performing tasks that would normally require human intelligence, such as pattern recognition and decision-making. Machine learning, in turn, is a subset of AI that allows systems to learn from data without being explicitly programmed, developing predictive or classification models. Deep learning is a subcategory of machine learning that uses artificial neural networks with multiple layers to analyze large volumes of data and solve more complex problems, such as image and natural language processing (Baduge *et al.*, 2022; Pan; Zhang, 2021).

2.3.1 Machine Learning

Machine Learning (ML) is an area of AI in which computers analyze a set of data and create models based on that data, which can then be used to solve problems. Unlike traditional programming, where rules are explicitly coded in a computer language without learning from the data, ML excels at generating predictive models from available data. These models are then applied to make predictions on new data, allowing the system to learn and adapt (Baduge *et al.*, 2022).

ML methods can be classified in different ways, one of the main ways being the amount of supervision received during the training process. Based on this, machine learning models are generally divided into supervised and unsupervised learning. In supervised learning, the dataset includes both predictors and outcomes, called “labels”. The model is initially trained on this labeled data, allowing it to make predictions on new, unlabeled data. The two most common tasks in supervised learning are classification and regression. The classification task predicts discrete class labels, while regression predicts continuous values. Popular supervised learning algorithms include k-nearest neighbors, support vector machines (SVMs), logistic regression, linear regression, and neural networks (Baduge *et al.*, 2022).

In unsupervised learning, an unlabeled dataset is used to determine hidden patterns or intrinsic structures in data. It is used for tasks such as clustering, anomaly detection, novelty detection, visualization, and dimensionality reduction (Baduge *et al.*, 2022).

2.3.2. Deep Learning

Deep Learning (DL) is a machine learning subfield that focuses on studying artificial neural networks and related algorithms, which contain more than one hidden layer. As a result, the computational process in a deep learning algorithm involves multiple steps from input to output. DL algorithms are particularly effective when working with high-dimensional data, such as images, video, and audio, compared to traditional machine learning algorithms due to their complex computational paths. Some of the most widely used DL algorithms in the construction and building industry are briefly introduced in the following paragraphs (Baduge *et al.*, 2022; Pan; Zhang, 2021).

Feedforward neural networks (FNN), also known as "Multi-Layer Perceptrons" (MLPs), are a commonly used deep learning algorithm in which information flows exclusively in one direction, from input to output, without any feedback loops. FNNs consist of multiple interconnected layers of neurons, where input data is processed by the input layer, passes through several hidden layers, and then the output layer produces the result. In the hidden layers, each neuron receives input from the previous layer, computes a weighted sum of the inputs (wixi), adds a bias term (b), and then applies a nonlinear activation function to produce the output (Baduge *et al.*, 2022).

The neural network is trained using a dataset, during which the output produced by the network's output layers is compared to the expected real values, and the resulting error (or loss) is calculated. There are various methods for calculating loss, such as mean squared error, mean absolute error, and binary cross-entropy. By summing up the losses across the entire training dataset and adding any regularization terms to prevent overfitting, the cost function is determined. The objective is to minimize this cost function by adjusting the network's weights through a process called backpropagation. Backpropagation computes the gradient between the error and the weights. Using this gradient, optimization algorithms like Adam, NAdam, Adadelta, and gradient descent adjust the weights to minimize the loss. The dataset is processed multiple times to fine-tune these weights, resulting in a trained model with minimized error. The trained model retains the adjusted weights for each input at every neuron, with these weights reflecting the importance of each input for the output results. Finally, the model can be used to predict outputs based on new, unseen data (Baduge *et al.*, 2022).

A Convolutional Neural Network (CNN) is a specialized type of Artificial Neural Network (ANN) designed to process data with a grid-like structure, making it particularly effective for image classification and computer vision tasks. The architecture of CNNs typically consists of three main types of layers: convolutional layers, pooling layers, and fully connected (FC) layers. In a standard CNN, convolutional layers are followed by pooling layers or additional convolutional layers, with the fully connected layer positioned at the end. The input layer of a CNN receives the image data. The convolutional layer, which is the fundamental component of the CNN, utilizes elements such as filters (also known as kernels or feature detectors) and generates a feature map. A filter is a 2D array of weights smaller than the image, and a dot product is computed between the image's pixel values and the filter's weights, with the result forming an output array. This operation, called convolution, is repeated as the filter moves across the entire image to identify features. Unlike traditional neural networks, neurons in one layer of a CNN are not fully connected to the neurons in the next layer, enhancing the network's efficiency and focus on local patterns (Baduge *et al.*, 2022; Pan; Zhang, 2021).

A Generative Adversarial Network (GAN) is a deep learning algorithm that focuses on generative modeling, enabling the creation of new images, videos, or audio that closely resemble the data from the training set. GANs consist of two neural networks, the 'generator' and the 'discriminator,' which work together in a competitive framework. The generator's role is to produce new data that mimics the characteristics of the training data, using feedback from the discriminator to improve the quality of its output. Meanwhile, the discriminator's job is to distinguish between real data from the training set and the synthetic data produced by the generator, providing feedback on how realistic the generated data is. Initially, the generator produces clearly fake outputs, which the discriminator can easily identify. However, as training progresses, the generator becomes more adept at producing outputs that can deceive the discriminator. When training is successful, the generator creates data that the discriminator increasingly classifies as real, reducing the discriminator's accuracy in distinguishing between real and generated data (Baduge *et al.*, 2022).

Thus, deep learning has become the dominant approach in computer vision, surpassing traditional statistical models due to its superior ability to capture contextual information from images, achieving state-of-the-art results. Deep learning-based methods primarily focus on three tasks: image classification, object detection, and semantic segmentation.

Image classification involves understanding an entire image by assigning a specific label to it. This task is commonly performed using convolutional neural networks (CNNs), which employ three main types of layers: convolutional layers, responsible for generating feature maps; pooling layers, which reduce the spatial dimensions of the inputs; and fully connected layers, which create one-dimensional feature vectors for classification. More advanced models, such as AlexNet, VGGNet, and ResNet, have been developed using CNNs as their backbone architecture (Pan; Zhang, 2021).

Object detection identifies and locates one or more conditions of interest within an image by drawing bounding boxes around each object and assigning appropriate labels. Region-based convolutional Neural Networks (R-CNN) serve as the foundational algorithm, combining rectangular region proposals with the features extracted by convolutional networks. However, variations such as Fast R-CNN, Faster R-CNN, and Mask R-CNN were introduced to improve R-CNN's computational efficiency. Additionally, the "You Only Look Once" (YOLO) family of algorithms plays a role, employing a single convolutional network trained end-to-end to predict both bounding boxes and class probabilities simultaneously (Pan; Zhang, 2021).

On the other hand, semantic segmentation aims to semantically interpret each pixel in an image by assigning a label to each one, precisely identifying the location and shape of objects or damage. This task heavily relies on Fully Convolutional Networks (FCNs), an extension of classical CNNs that replace fully connected layers with fully convolutional layers. FCNs have proven effective in learning pixel-to-pixel mappings and making predictions for varying-sized inputs (Pan; Zhang, 2021).

2.4 Thesis Main Focus: BIM-AI photovoltaic energy solution

Solar energy, as a primary renewable energy source, is used due to its great potential to meet the growing demand for energy and the limited fossil fuel resources on the planet. With the support of national policies, renewable energy targets, and the falling costs of photovoltaic (PV) modules, the solar energy market has experienced growth. Consequently, an increase in the number of buildings integrated with PV systems has been observed. In this context, the forecast of solar energy production is one of the strategies in the design phase for the development of sustainable projects (Tian; Ooka; Lee, 2023). Furthermore, energy simulation models must be able to process a large amount of data for different climate conditions. This allows designers to adopt alternative design solutions for bioclimatic building elements.

In this context, deep learning algorithms can extract features from non-linear data by identifying complex patterns in large data sets. They are suitable for energy simulations, such as estimating solar radiation and predicting PV energy production. When associated with BIM, specifically in the design and planning phase of buildings, quantifying PV energy production through predictive solar energy algorithms becomes a proactive strategy for sustainable projects (Olu-Ajayi *et al.*, 2022; Shao *et al.*, 2021; Wang *et al.*, 2023).

Previous research has advanced the body of knowledge, seeking automated alternatives to improve the energy efficiency of buildings. Alawi *et al.* (2024) developed predictive simulations for residential buildings' annual heating and cooling loads. Olu-Ajayi *et al.* (2022) seek to predict energy consumption in the building design phase. Chou *et al.* (2017) developed a BIM data fusion process on energy consumption datasets collected. Li *et al.* (2024) proposes an adaptive sea lion-optimized genetic adversarial to predict renewable energy sources. Tao *et al.* (2024) apply tree-based, linear, and non-linear regression techniques to predict the energy and exergy efficiency of Parabolic Trough Solar Collectors using oil-based nanofluids. There is a gap in the literature addressing BIM-driven solutions with PV studies, especially with the association of AI algorithms.

Furthermore, the design process for PV systems is carried out in different phases, such as architectural design and photovoltaic system design, with each phase being conducted by specialists or engineers using tools specific to their respective fields. For example, while an architect is responsible for the design of the building, an electrical engineer designs the PV system. This results in using different software tools in each phase, even for the same project, which can lead to different approaches to representing the same concept. In the architectural design phase, programs such as AutoCAD, MyArchiCAD and SketchUp are commonly used. In the photovoltaic design phase, several tools are applied, ranging from more general options, such as PVsyst and Retscreen, to specific solutions, such as models for analyzing partial shading or optimizing the connection of PV modules. This use of proprietary models in each phase makes it difficult to seamlessly integrate the PV system, since the models created in one phase are not automatically compatible with those in the next phase. Therefore, changes made in one phase cannot be easily synchronized or adapted in the others without a system of integration and continuous updating (Ning *et al.*, 2018).

BIM can solve this problem in the PV design phase by offering a collaborative approach to project development. Through 3D modeling and the creation of a centralized model, BIM allows all parties involved in the project, from energy engineers to architects and builders, to work with the same information in real-time. This ensures that changes made in one phase are automatically reflected in all others, avoiding disconnection between the different models and tools used. The use of the IFC format and BIM projects offers an open standard that facilitates the exchange of information between different software and platforms, ensuring compatibility between the models created in the different phases of the project.

In this context, the large amounts of information from BIM models can be associated with AI algorithms. Specifically in the field of AI, the literature highlights that deep learning algorithms provide advanced techniques to achieve improved modeling and better prediction performance. DL uses deep architecture or multi-layer architectures (Olu-Ajayi *et al.*, 2022).

Numerous AI-based methods, including neural networks (NN), support vector regression (SVR), long short-term memory (LSTM), gated recurrent units (GRU), convolutional neural networks (CNN), and multi-layer perceptrons (MLP), have been developed by researchers for solar irradiance prediction. More recently, hybrid approaches, such as extreme gradient boosting trees (XGBT) and deep neural networks (DNN), have been applied to predict hourly irradiance levels. Given that photovoltaic generation is dependent on the variability of solar radiation, managing the power grid's operation becomes a complex task (Sammar *et al.*, 2024). Also, early research has investigated the implementation of an optimized Long Short-Term Memory (LSTM) deep learning network and compared it with two different algorithms, genetic algorithm (GA) and particle swarm optimization (PSO), to predict electric loads. The optimized LSTM networks revealed better results than tree-based ensemble models, Support Vector Regression, and artificial neural networks through an extensive comparison (Alawi; Kamar; Yaseen, 2024; Sammar *et al.*, 2024).

Yan *et al.* (2020) presents a hybrid deep learning model that combines a neural network based on recurrent units with an attention mechanism to predict solar irradiance variations, extracting features from the data via Inception and ResNet-NN and subsequently processing them in a recurrent neural network (Yan *et al.*, 2020). Mutavhatsindi *et al.* (2020) focus on hourly solar energy forecasting, where different machine learning methods were compared, such as recurrent neural networks, LSTM, and Feed-forward neural networks (Mutavhatsindi; Sigauke; Mbuvha, 2020). Brahma and Wadhvani (2020) evaluated DL techniques to predict daily solar radiation, using regional data and techniques such as bidirectional LSTM and attention-based networks, which showed good performance in metrics such as MSE and RMSE (Brahma; Wadhvani, 2020; Kumari; Toshniwal, 2021).

However, little research in the AECO sector has investigated the integration between BIM and AI. This paper recognizes that buildings are strategic projects for implementing sustainable actions. It is argued that an integrated BIM and AI approach considering the use of solar energy in the design phase enhances the gains related to energy production. Information about the geometry and geographic positioning of the building can be considered during the planning of the building. Therefore, BIM modelers can develop project designs that enhance and integrate the gains of photovoltaic plants, increasing solar energy production. In an ideal scenario, buildings could be self-sufficient regarding energy demands.

Also, literature has shown that it is possible to develop BIM-PV integration even in the design phase. For example, (Aksoy Tirmikçi *et al.*, 2025) propose simulating a PV system installed in a near-zero energy building (NZEB) with PVSOL. The authors developed a machine-learning model based on local climate data. The study presented a performance ratio (PR) of 81.9% in the first year, representing the PV system's operational efficiency compared to its theoretical energy generation potential. The initial investment of USD 435,600 has a predicted payback period of 11.42 years, while PVSOL is estimated to be 14.9 years.

Zhang *et al.* (2025) analyze the solar potential, architectural modeling, financial feasibility, and environmental impacts of projects to simulate distributed PV systems on the roofs of a community. The results indicate that the installation of 79 units generates 1328.74 MWh annually. This meets residents' energy needs and provides a surplus to the electricity grid. The use of light-colored PV modules and elevated pavilion-type structures meet the aesthetic standards of the projects in terms of local architecture while maximizing energy efficiency and rooftop utilization. The life cycle assessment confirmed the project's economic viability, presenting an internal rate of return of 5.82% and a discounted payback period of 15.31 years, considering additional architectural integration costs. In addition, the installation reduced 24,754.77 tons of CO2 over 25 years.

Although PV technology can potentially drive energy transition and decarbonization, its adoption in the AECO industry still faces regulatory, economic, and technical challenges (Chen *et al.*, 2022). The literature highlights that BIM-PV integration enhances standardization and design optimization and improves collaboration among AEC stakeholders. Research shows that BIM can strengthen the integration of PV systems in buildings, from solar potential analysis and parametric modeling to cost-benefit and life cycle assessments. Although BIM-PV integration is still primarily focused on solar mapping and automating panel positioning, there is a vast field of research and innovation, especially in developing solutions at the design stage (Tian *et al.*, 2023).

While the body of knowledge related to photovoltaic projects in the AECO industry explores alternatives to enhance solar energy production through the application of different materials for solar modules (Lins *et al.*, 2024; Myint *et al.*, 2025; Serat *et al.*, 2025; Zhi *et al.*, 2023), the literature on BIM in sustainable projects seeks to propose solutions in the various dimensions of the framework (Cassandro *et al.*, 2024; Lins *et al.*, 2024; Mandičák *et al.*, 2024; Myint *et al.*, 2025). However, there is a gap in the literature related to integrating photovoltaic projects and BIM technologies for simulations in three-dimensional models during the project design phase. For example, Myint *et al.* (2025) discuss the importance of research on using technologies such as BIM to drive the development of automated projects that evaluate different BIPV design options. Myint *et al.* (2025) argue that articles about BIM-PV can guide the development of low-carbon buildings. However, research in the field of PV systems, such as (Serat *et al.*, 2025; Zhi *et al.*, 2023), has advanced mainly in simulations of solar energy production, but often without considering the direct integration of PV systems into the building envelope, as proposed by building-integrated photovoltaic systems. This gap is also reflected in the lack of integration between PV simulation tools and BIM-based building design processes, causing interoperability problems and limiting automated and data-driven design approaches (Palha *et al.*, 2024; Zhi *et al.*, 2023). In addition, although recent research in BIM has focused on automated solutions that address sustainability requirements (Cao; Huang, 2023; Cassandro *et al.*, 2024), there is still a lack of studies proposing comprehensive workflows that integrate BIPV design, energy simulations, and BIM technologies from the early design stages (Mandičák *et al.*, 2024; Nascimento *et al.*, 2023).

3 THESIS METHODOLOGIES

This thesis adopts a multi-method approach. The three subsequent chapters apply different methodologies to address each of the research objectives introduced. The literature review showed the trends related to the research field and how this field impacts the AECO industry. The application of deep learning algorithms is a contribution oriented to the planning of photovoltaic projects integrated into buildings. Specifically, this chapter is an excerpt from the key areas identified in the literature review. Finally, the thesis contributes by establishing a process for selecting solar module manufacturers and models based on cost and total PV energy production. Therefore, the thesis has three stages, as shown in Table 1.

Table 1 - Methods and objectives of the thesis, with the software used and main results

Main Objective: Investigate how the integration of BIM and AI applications can foster a BIM-driven design process to support the planning and retrofit of sustainable buildings, with a specific focus on photovoltaic systems in the AECO sector.				
Stage	**Specific Objective**	**Method**	**Software**	**Key Results**
Chapter 4 (Paper 1)	a) Investigate how BIM and AI capabilities can improve the development of smart architecture, engineering, construction, and operation projects	Systematic Literature Review	Rstudio Biblioshiny	Maps 14 BIM, 16 AI capabilities, BIM-AI benefits, and a framework
Chapter 5 (Paper 2)	b) Explore the application of a BIM-driven deep learning algorithm to estimate PV energy production, associating solar radiation series and automated extraction of information in BIM models	Times series simulation with DL	Revit/Dynamo Solarius PV Google Collab/Python	Automated process for solar energy estimation
Chapter 6 (Paper 3)	c) Develop an automated process model for allocating solar modules, seeking to maximize photovoltaic energy production (kWh/day) while minimizing implementation costs	Optimization algorithm	Revit/Dynamo Google Collab/Python	Optimized layout of solar modules

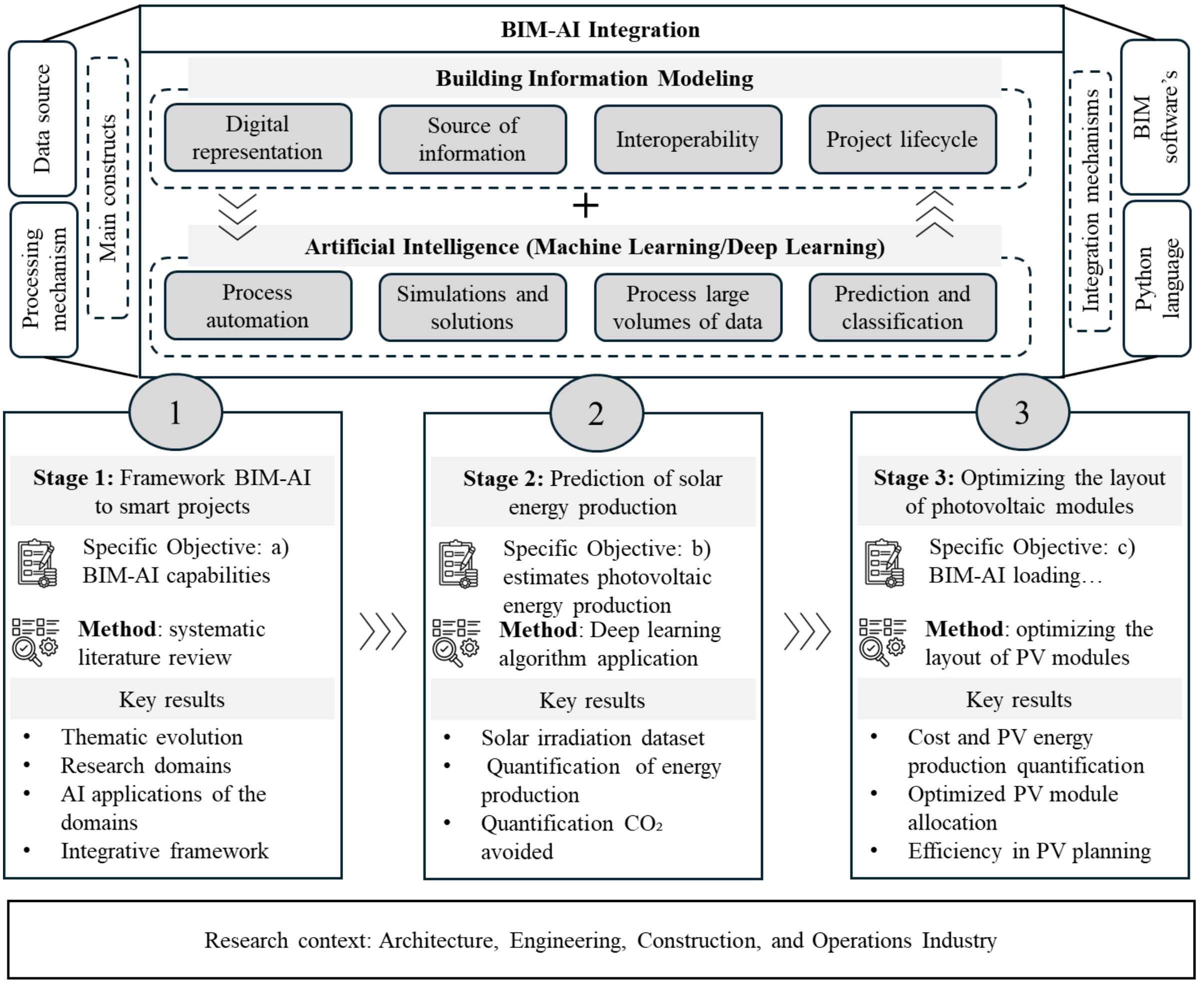
In the first Thesis stage, it is argued that the AECO sector gains advantages through generating and effectively managing BIM data. This increased available data can be fundamental in deriving innovative advances by processing them through AI models. In this context, this stage investigates how BIM and AI capabilities can benefit the development of smart AECO projects. The research design is a systematic literature review, applying bibliometric and content analysis. First, the chapter explores the relationships between the topics of AI and BIM applications, identifies seven core domains of BIM and AI finds application, explores research contributions, the problems addressed, and their primary outcomes. Second, the paper maps out 14 BIM and 16 AI capabilities fundamental to developing smart projects. Third, three propositions that sustain an integrative framework are suggested. The chapter also suggests that practitioners identify critical organizational capabilities to be built and strengthened.

In the second stage, it argued that photovoltaic energy is a renewable source that offers the potential to meet the growing demand in buildings. Through stage 1, it is observed that solar energy use in buildings opens the potential for research in solar radiation forecasting and photovoltaic energy production with application in BIM. In this context, this stage aims to explore the application of a BIM-driven deep learning algorithm to estimate PV energy production. This stage quantifies the energy produced and CO_2 emissions avoided based on the predicted values of the implemented algorithm, using a routine in Dynamo that extracts the information from a BIM model. Thus, this stage uses solar radiation time series and automatic BIM data extraction to establish an automated process in the design phase, called SolarisBIM.AI, for quantifying solar energy production and avoided CO_2. This strategy offers designers, engineers, and managers another way to analyze buildings' energy efficiency and sustainability. The stage provides an automated process that can serve as a strategy to predictively quantify the sustainable actions of the project during the design phase.

In the third stage, it is argued that to determine the efficiency of PV systems, aspects such as solar radiation (to estimate total energy production), available area on the building roof, PV module brands, models, and costs must be considered. The traditional project process for applying photovoltaic solutions is still limited to two-dimensional data without considering all structural elements of buildings' roofs. This stage aims to develop an automated process for allocating solar modules to maximize PV energy production while minimizing implementation costs. This stage integrates visual programming in Dynamo with programming in Python to analyze different combinations of PV modules, considering the dimensions of 21 PV modules from 4 brands for allocation on the roof of a building. The algorithm identifies the most efficient configuration of photovoltaic cost-production. It uses Dynamo to extract information on the families of PV modules and the available roof area from a BIM model in Revit. Finally, the model automatically allocates the best arrangement of PV modules directly in the Revit model. Experiment 1 used 721 photovoltaic modules and obtained the highest daily energy production, with 2723.63 kWh/day. However, this solution also presented the highest total cost, reaching USD $ 387,575.45. Experiment 2 was the one that used the greatest number of panels, totaling 969 units. Energy production was lower than in Experiment 1, with 2630.09 kWh/day. On the other hand, this layout presented a lower cost than Experiment 1, totaling USD $ 370,504.08. This study also compares the results with a case study.

Thus, Figure 2 shows the schematic structure of the thesis.

Figure 2 - Thesis schematic structure



4 TOWARDS AN INTEGRATIVE FRAMEWORK FOR BIM AND ARTIFICIAL INTELLIGENCE CAPABILITIES IN SMART ARCHITECTURE, ENGINEERING, CONSTRUCTION, AND OPERATIONS PROJECTS

4.1 Chapter Introduction

The digitalization domain of the Architecture, Engineering, Construction, and Operations (AECO) industry is advanced by BIM, a fundamental approach to dealing with the growing volume of information and data generated throughout the lifecycle of construction projects (Zhang *et al.*, 2022). Adopting BIM provides capabilities to the AECO sector, covering technological, organizational, and procedural levels aimed at innovation in the industry (Alzara *et al.*, 2023b). The existing limitations in processing data from BIM models have led to academic research exploring the application of artificial intelligence (AI) algorithms, specifically Machine Learning (ML) and Deep Learning (DL) models (LI *et al.*, 2024; Padala; Skanda, 2024). Integrating BIM and AI holds promising advantages, opening research areas to explore how capabilities for advancing smart projects can be effectively applied in the realm of AECO in two ways. Thus, many professions currently face the prospect of evolution and change, and AECO is no different. The adoption of BIM and AI in the field is growing fast and becoming a reality in industry; therefore, future research shall address key points to enable the evolution of the AECO field.

First, BIM-based projects have 3D parametric, object-based, and data-rich information associated with attributes and datasets. In AI, these attributes of BIM models are applied for automation, prediction, and various forms of learning in the AECO industry (Abdulfattah *et al.*, 2023). The BIM modeling process encompasses tools and technologies applied throughout the entire life cycle of buildings, digitally documenting requirements related to the performance, planning, construction, and operation of projects (Bloch; Sacks, 2020; Ying *et al.*, 2023). This process enables the creation of accurate information models and aids design analysis, simulation, and interpretation for enhanced utility (Li *et al.*, 2024; Padala; Skanda, 2024). Due to this digital transformation, BIM data can be used to create innovative architectural and structural designs, improve construction safety, reduce operational costs, increase construction speed, and enhance sustainable solutions (Li *et al.*, 2024).

In this sense, this chapter theorizes that as the AECO industry incorporates the demands of BIM data-driven digitalization and process automation, researchers and practitioners will be challenged to deal with large volumes of data and understand object-data relationships and interdependencies (Alzara *et al.*, 2023b; Li *et al.*, 2024). Data management, processing, and interpretation capabilities will be necessary in several domains of the AECO industry. In this context, by employing AI, especially ML and DL, this data can be transformed into knowledge for various applications (Abdulfattah *et al.*, 2023). While ML methods develop models that predict future outcomes based on historical data, DL methods adopt a more automatic neural pipeline that does not require complex feature learning and detection. Both DL and ML are subfields of AI (Abdulfattah *et al.*, 2023; Wang; Gan, 2023a).

Second, it is essential to establish BIM capabilities for data management and AI capabilities for processing and interpreting that data. In this Thesis, the capability construct encompasses the knowledge, skills, and experience professionals need to understand, evaluate, implement, and promote the development of smart projects (Chen *et al.*, 2023a). Previous research investigated the development of BIM and AI from different perspectives. Some of them aim to explore the evolutionary development of the BIM research area and AI applications, showing that this combined approach gradually grows in the field of cost management (Naderi *et al.*, 2024), point cloud applications (Tavolare *et al.*, 2023), building renovation (Mulero-Palencia *et al.*, 2021), energy efficiency (Ratajczak *et al.*, 2023) and Historical Building Information Modeling (HBIM) (Garcia-Gago *et al.*, 2022). Current literature seeks to formulate strategies that aim to automate processes and simulations in intelligent projects through data generated in BIM models in different phases of the project's life cycle (Abdulfattah *et al.*, 2023; Bloch; Sacks, 2020; Li *et al.*, 2024; Padala; Skanda, 2024; Ying *et al.*, 2023). Regardless, there is a gap in the literature in pointing out which BIM and AI capabilities are essential for developing smart projects. Furthermore, discussing the potential benefits of combining BIM and AI capabilities in the AECO sector is essential.

Thus, this chapter investigates how BIM and AI capabilities can improve the development of smart architecture, engineering, construction, and operation projects. To support this research objective, the paper answers the following research questions: (#RQ1) What are the essential BIM and AI capabilities for smart AECO projects? (#RQ2) What are the main benefits of the connection between BIM and AI for developing smart AECO projects?

The chapter presents an integrative research model, which explores the main contributions of the association between BIM and AI in AECO projects. The research design is a systematic literature review that applies bibliometrics and content analysis with the assistance of Bibliometrix and Mendeley software. The main topics, thematic evolution, and concept maps are covered. Finally, using a coding scheme in the content analysis stage, the article explores the relationships between AI and BIM applications, essential capabilities for this integration, and potential benefits for the sector. To answer research question 1 (#RQ1), this paper maps two types of capabilities oriented to data management and application. As for research question 2 (#RQ2), this paper covers five core benefits domains in the AECO sector.

Previous literature reviews have presented maps and quantitatively analyzed the literature on AI applications in the AECO domain. These researches provide directions for the AECO literature by visualizing and understanding trends, identifying research topics, journals, institutions, key authors, and countries, and providing directions for future research (Darko *et al.*, 2020; Heidari *et al.*, 2023; Pan; Zhang, 2021b, 2023b; Zabin *et al.*, 2022). However, this study suggests an approach that goes beyond mapping recent literature on the application of AI in the field of AECO. This chapter presents a theoretical framework grounded in three propositions, two essential categories of technological capabilities, and five potential benefits of integrating BIM and AI. The framework systematically incorporates recent literature contributions by organizing them into defined categories of capabilities and associated benefits. This structured approach offers an understanding of how BIM and AI can be effectively combined to enhance various aspects of AECO projects, providing a comprehensive foundation for both theoretical exploration and practical application.

This research also highlights the global relevance of the BIM-AI subject in terms of collaboration between countries, recent research topics, and their distribution over the years, as well as the main problems proposed by the literature and their leading solutions for the evolution of the AECO industry. In this context, it is observed that recent research seeks to improve predictive maintenance and facility management (Cheng *et al.*, 2020; Marzouk; Zaher, 2020a; Palha *et al.*, 2024), proposes to assess and reduce the environmental impact of buildings through life cycle analysis and CO_2 emissions monitoring (Arsiwala *et al.*, 2023; Wu; Maalek, 2023) and develops intelligent systems that can optimize energy consumption and thermal comfort in buildings (Erişen, 2023; Hou *et al.*, 2022). They are still advancing in the theoretical and practical fields with the use of AI to facilitate the documentation, preservation, and restoration of historic buildings, employing computer vision and machine learning techniques for 3D reconstruction (Croce *et al.*, 2023; JIANG *et al.*, 2022) and applying AI in generative design to automate design processes, such as the generation of multiple layout options (Abdirad; Mathur, 2021; Wang *et al.*, 2023).

This chapter makes contributions to the AECO field from this research. It identifies seven core domains and maps out 14 BIM and 16 AI capabilities fundamental to developing smart projects. The discussion of these capabilities lies in understanding their inherent nature, their diverse applications, and the potential benefits they bring. These analyses suggest three propositions, and the paper explores the potential benefits of BIM and AI capabilities.

4.2 Research Methodology

The chapter applies a Systematic Literature Review (SLR) because it is a structured, replicable, and transparent method essential in developing frameworks based on scientific evidence (G. Alves *et al.*, 2024). The SLR is conducted through a bibliometric analysis and content analysis to identify research topics, approaches, and integrated applications of BIM and the AI domain (Tranfield *et al.*, 2003).

Given the rapid technological advancements and the increasing complexity of AECO projects, an SLR is needed to map out which capabilities are essential to integrate BIM and AI efficiently. Furthermore, what are the potential benefits that such integration generates? This review goes beyond previous studies by identifying, categorizing, and analyzing the essential capabilities of BIM and AI; the paper proposes an integrative framework that articulates the capabilities of BIM and AI, highlighting the potential benefits of this integration. Therefore, this paper provides theoretical and practical guidance through SLR to support future research and adopt innovative solutions in the AECO industry.

Furthermore, this paper follows a well-defined protocol that ensures comprehensive coverage of the relevant literature. Given the novelty and complexity of BIM and AI integration and the rapidly evolving technological landscape, an SLR provides the methodological approach to map essential capabilities required for smart AECO projects. Furthermore, the systematic process allows for identifying gaps in existing research and supports the development of a theoretically grounded and evidence-based integrative framework.

This research introduces two questions to be answered (#RQ1 and #RQ2). The first question seeks to advance the body of knowledge by identifying and categorizing BIM and AI capabilities and functionalities for the development of smart projects within the AECO industry. The goal is to understand which technologies enable smart project design, construction, and operation. The second question seeks to elucidate how BIM and AI improve project efficiency, sustainability, cost-effectiveness, and innovation. Both questions arise from the need to know the fundamental technological capabilities for developing smart projects, aiming at automating the design process through BIM and AI. In this context, SLR was essential to map the necessary skills based on scientific evidence and to understand the main applications of BIM and AI that drive the technological development of the AECO sector.

This section describes the SLR methodology for data composition and analysis. The methodological procedures adopted in developing the SLR are detailed to ensure the transparency and reproducibility of the sample composition and results. Thus, the following subsections present the databases used and the search strings. Furthermore, the paper describes the inclusion and exclusion criteria for the sample composition to ensure the quality and relevance of the articles analyzed. Finally, the bibliometric and content analysis steps are described.

4.2.1 Sampling Process

4.2.1.1 Databases selection

The sample was collected from the Scopus and Web of Science (WoS) databases. Both databases provide comprehensive and up-to-date coverage of scientific publications, enabling researchers to access recent and relevant research and ensuring that the literature review accurately reflects the current state of the art (Alves; De Carvalho, 2023). The SLR adheres to the recommended methods proposed by Tranfield *et al.* (2003), ensuring a systematic approach divided into three stages, as described in the following subsections.

4.2.1.2 Search strings

First, the SLR scope was defined, including the objectives, research questions, and inclusion and exclusion criteria. At this stage, it was also formulated search strings: *[("Building information model*" OR BIM) AND ("machine learning" OR "deep learning" OR "artificial intelligence")*], to be applied to titles, abstracts, or keywords up to the year 2023. The formulation of the selected search strings seeks to capture scientific research at the intersection between BIM and AI, including its subareas, machine learning, and deep learning. This search was conducted on March 27, 2024, resulting in 1338 documents in Scopus and 729 in Web of Science.

4.2.1.3 Scientific document type filtering

Second, the sample was refined to include only articles, reviews, and early access publications in English, resulting in 511 articles in Scopus and 540 articles in Web of Science.

The paper specifically opted for articles due to their typical submission to a peer-review process, ensuring a certain standard of quality, reliability, and methodological rigor.

4.2.1.4 Duplicate Article Exclusion

The metadata files (savedrecs.bib for WoS and scopus.bib for Scopus) were exported from both databases in the '.bib' format and employed the RStudio computing environment (version 2023.9) to remove duplicates. Using a code to convert BibTeX files to a data frame, the two files were merged and identified duplicates, resulting in the identification of 397 duplicate articles. Therefore, the combined dataset from both databases amounted to 654 unique articles.

4.2.2 Inclusion and Exclusions criteria

After combining the metadata from the two databases, the articles were then evaluated using inclusion and exclusion criteria to compose the final sample, which are discussed in the two following subsections.

4.2.2.1 Title and abstract reading

The dataset was exported to ".xlsx" format with all the article information. First, the articles that remained duplicated were removed. Then, all the titles and abstracts of the articles were read. In this stage, 184 articles that were outside the scope of the study or that remained duplicates were excluded. In this phase, two exclusion criteria were applied: 1) articles that use the acronym 'BIM' (related diseases) in the healthcare domain; 2) articles that only concentrate on BIM, treating AI as a subject for future research without exploring direct correlations between the two technologies. These articles were downloaded and read before being definitively excluded from the sample.

The exclusion criteria adopted at this stage ensure the relevance and specificity of the studies analyzed in relation to the main objective of the research. First, excluding articles that use the acronym 'BIM' in the context of diseases in the health domain avoids confusion with the term Building Information Modeling, which is the focus of this research. Second, the exclusion of articles that treat BIM in isolation or consider AI only as a topic for future research without exploring the direct relationship between these two technologies is important to focus on the current and practical integration of BIM with AI.

4.2.2.2 Content analysis

At this stage, all remaining articles were downloaded so that authors could read them in full. Even at this advanced stage of sample composition, 18 downloaded documents were not of the article type or only had the title and abstract in English (with the rest of the manuscript in another language). Although many articles have titles and abstracts available in English, the whole body of the document may be in another language, which may not be immediately apparent in the initial screening phase based on titles and abstracts. In addition, document type classification may be inaccurate or incomplete in database indexing information.

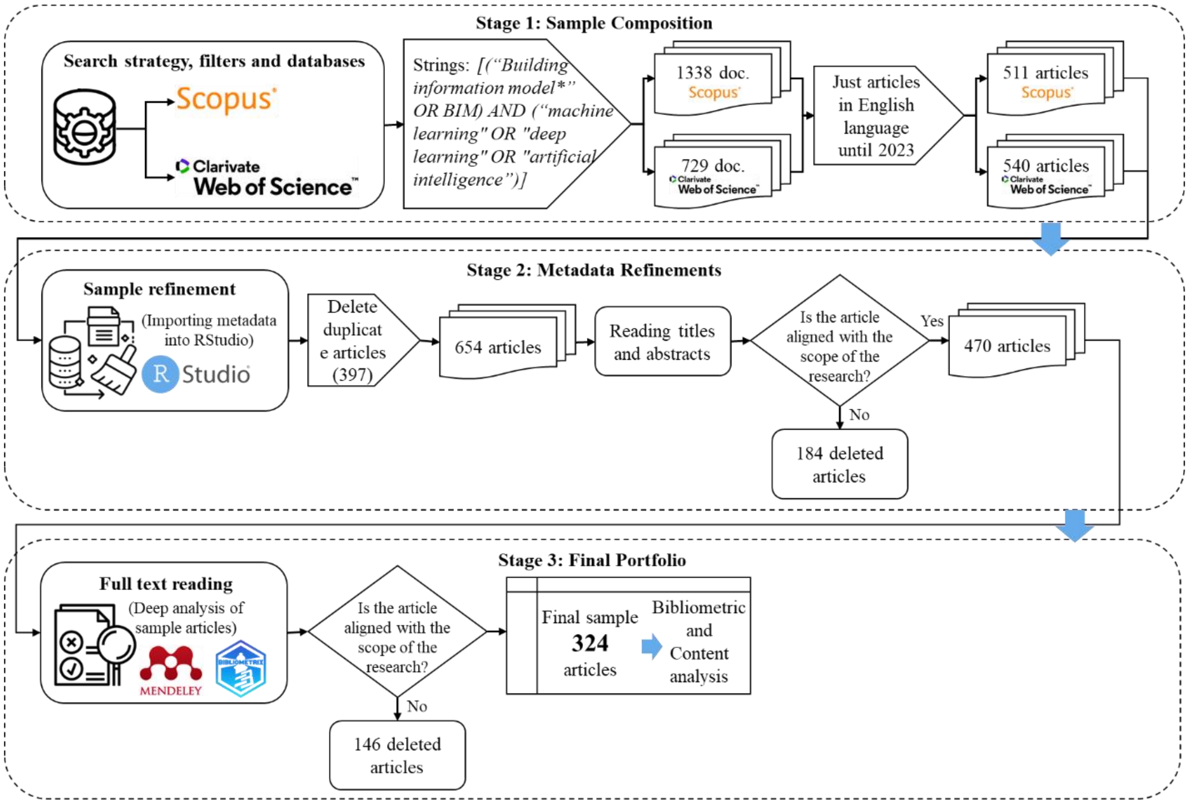
After reading all manuscripts in full, 128 articles are excluded. The exclusion criteria were: 1) articles that primarily focused on analyses limited to the quantification of research on the topic; 2) lacked methodological rigor, provided superficial accounts, or lacked a connection between BIM and AI; 3) articles that discussed only the importance of addressing the AI and BIM theme in undergraduate curriculum. Articles that limited themselves to quantifying research on the topic were excluded because this article aims to understand the applications and theoretical advances between BIM and AI. In addition, the established exclusion criteria aim to ensure that the analysis is based on robust studies that offer contributions to the field. Finally, the review is focused on research that effectively explored the synergies between BIM and AI.

Hence, the final sample yielded 324 articles (see Appendix A).

4.2.2.3 Data Analysis

Given the final sample of 324 articles, the data analysis was performed through bibliometric and content analysis techniques. This phase aimed to identify patterns, trends, and relationships within the examined literature. The entire sample composition process is shown in Figure 3.

Figure 3 - Sample composition process



Bibliometric analysis enables a quantitative assessment of elements, such as the frequency of publications, most cited authors, and temporal evolution of research (Araújo *et al.*, 2020a). To accomplish this, RStudio was utilized to access the Bibliometrix package, and the unified data frame of the final sample in the Biblioshiny environment was subsequently analyzed. Bibliometrix, a specialized R library, is designed to extract detailed bibliometric information from datasets, offering a broad range of quantitative metrics. It facilitates a analysis of publication patterns, prominent authors, co-authorship networks, and the temporal evolution of topics (Aria; Cuccurullo, 2017). Conversely, Biblioshiny is an interactive graphical interface built on Shiny, a framework for developing web applications in R. This tool provides a dynamic visual experience, enabling researchers to explore and interpret bibliometric data intuitively. The integration of these tools offers a approach to bibliometric analysis, combining the analytical robustness of Bibliometrix with the accessibility and interactivity provided by Biblioshiny, thereby enhancing understanding and utility in investigating bibliographic patterns and research dynamics (Aria; Cuccurullo, 2017).

Furthermore, content analysis deepens qualitative understanding by exploring the specific nuances and characteristics addressed in the articles (Almeida-Filho *et al.*, 2021; Alves; De Carvalho, 2023). In this case, all articles were downloaded and managed using Mendeley software (version 1.19.8). In addition to its primary functionality as a reference manager, Mendeley offers features that facilitate the organization and categorization of documents, essential elements for content analysis.

In this sense, it was possible to structure and classify the articles according to categories that reflect the coding of the sample, identification of patterns, and emerging topics. During coding at this stage, words or phrases were considered units of meaning to be assigned to specific categories or codes, reflecting emerging concepts or themes (Alves; De Carvalho, 2023). It adopts an inductive approach, allowing categories to appear throughout the process until sample saturation.

Therefore, this chapter first categorized the sample into disciplines and the applications of AI and BIM. At this stage, the problems to be solved and the results obtained through the analyses carried out by the authors are identified. This chapter then codifies the capabilities needed for BIM and AI applications. Based on these capabilities, this paper maps the benefits generated by the integrated application of these technologies.

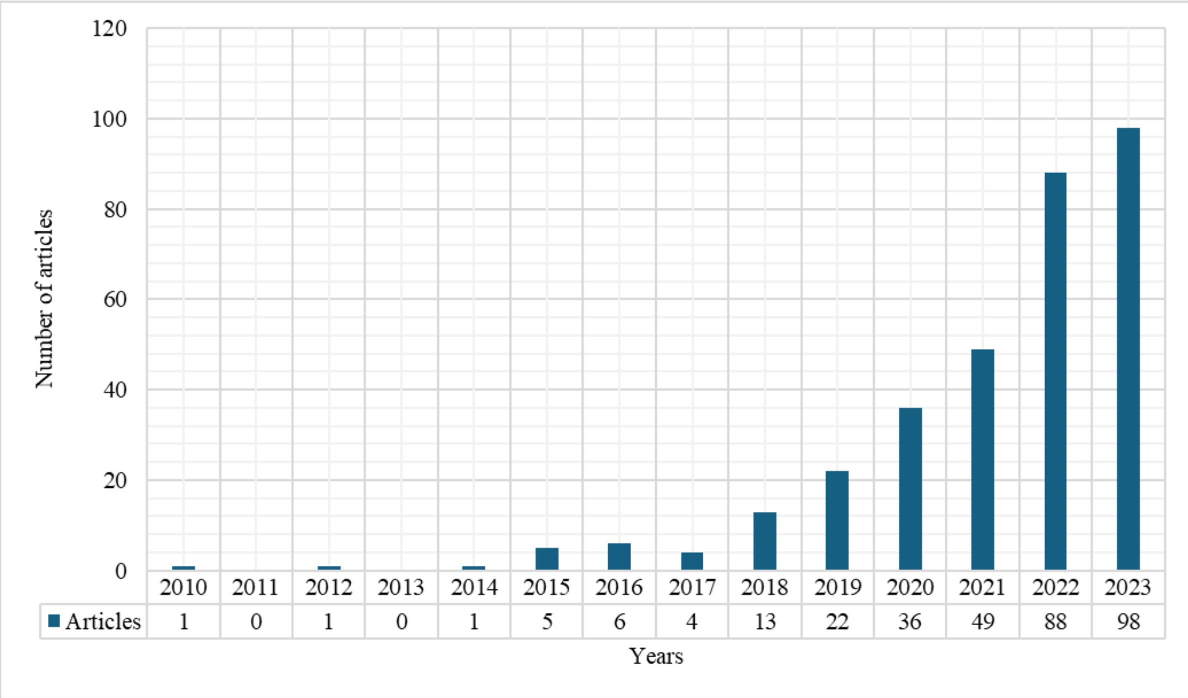
4.3 Systematic Literature Review Results

The results section initially presents the bibliometric findings through conceptual and correlation maps generated from analyzing the sample metadata. Thus, the core domains of the research are identified through content analysis. Each of these core domains is defined, and then the capabilities related to BIM and AI are presented, as well as the benefits arising from the integration between BIM and AI.

4.3.1 Bibliometric Analysis for Thematic Maps

This section presents the results of the bibliometric analysis in Biblioshiny (utilizing Bibliometrix) for the 324 articles comprising the sample in this research. Figure 4 shows an increasing trend in academic production over the years, with an increase in articles published from 2018 onwards. Between 2010 and 2017, the number of publications was low, with an average of 2 articles per year. However, from 2018 onwards, scientific production increased, reaching 13 articles in 2018 and a peak of 98 articles in 2023. This growth in recent years suggests increased interest and relevance of the topic researched, possibly driven by technological advances and a greater demand for innovation in the field. The average annual production for the entire period (2010-2023) is approximately 26 articles per year.

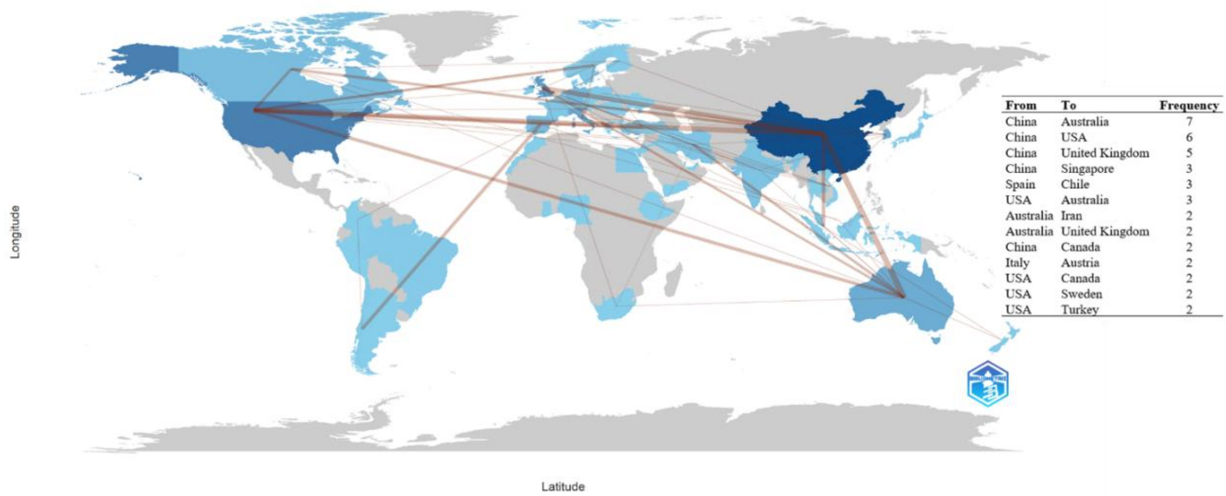
Figure 4 - Annual Production



The analysis of the geographical distribution of scientific production in this sample, which includes 58 countries, shows that China leads with approximately 26% of publications. The United States follows with approximately 13%, and Australia and the United Kingdom each contribute approximately 7%. South Korea contributes approximately 6% of publications, and Italy approximately 5%. The remaining countries contribute 3%. Figure 5 reflects the nationality of all the authors in the collection, with the intensity of the color corresponding to the number of publications.

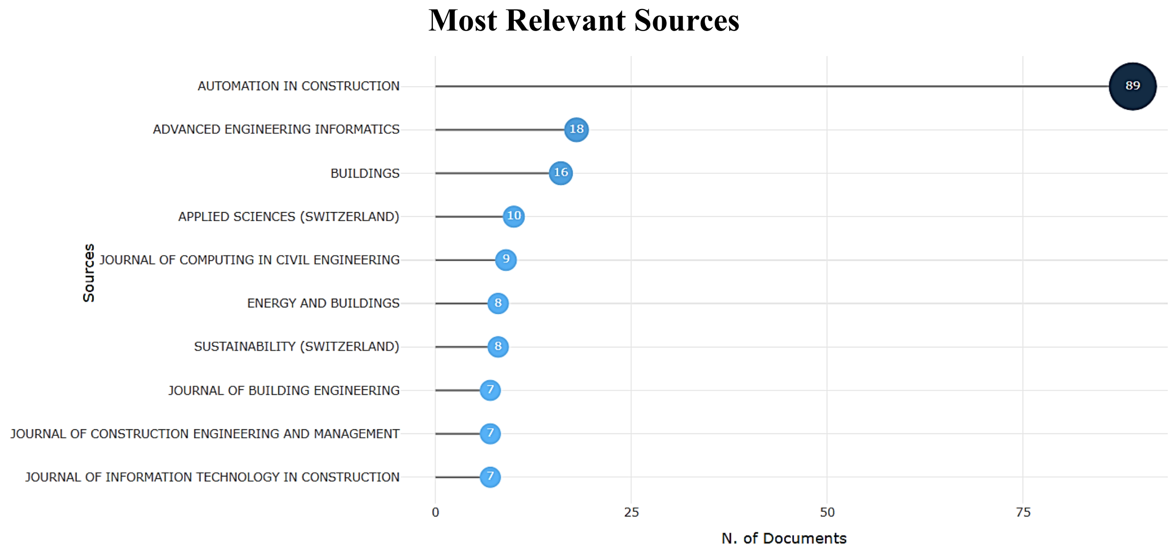
The data also indicates 61 international collaborations across various countries, with notable partnerships between China, Australia, the USA, the United Kingdom, Singapore, and Chile. Figure 5 illustrates the geographical distribution of the research sample, specifically focusing on collaborations that occurred at least twice.

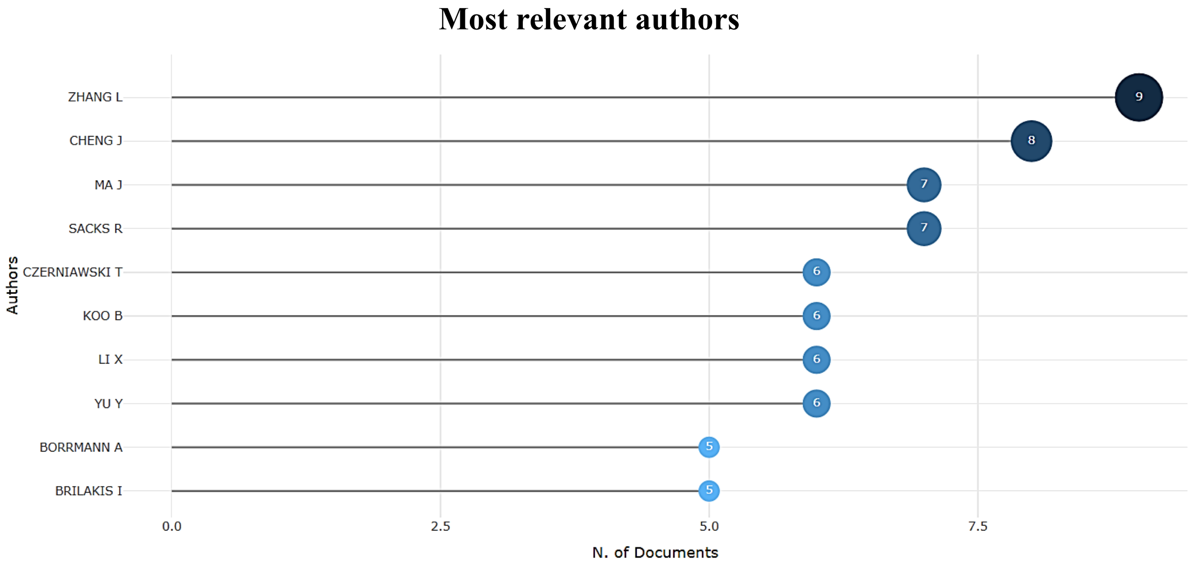
Figure 5 - Country Collaboration Map



The journal Automation in Construction leads the sample with 89 papers (see Figure 6), consolidating itself as the primary source of research on automation and technological innovation in construction. The two main authors of the sample, Zhang L and Cheng J, focus on the application of AI and BIM in the AECO sector, but with specific emphasis. Zhang L focuses on improving smart construction management, including mining BIM logs to predict design commands and detect logical relationships in mechanical, electrical, and plumbing (MEP) systems. He also investigates the assessment of safety risks in complex projects, such as tunnels, using explainable tree-based optimization methods (Lin *et al.*, 2023; Wang *et al.*, 2022). Cheng J focuses on areas such as using BIM and IoT for predictive maintenance of MEP components and optimization methods for building surveillance and fire evacuation. In addition, Cheng J explores the automatic segmentation of industrial point clouds using neural networks, demonstrating a continued interest in applying machine learning for spatial and safety data analysis in buildings (Cheng *et al.*, 2020; Yin *et al.*, 2021).

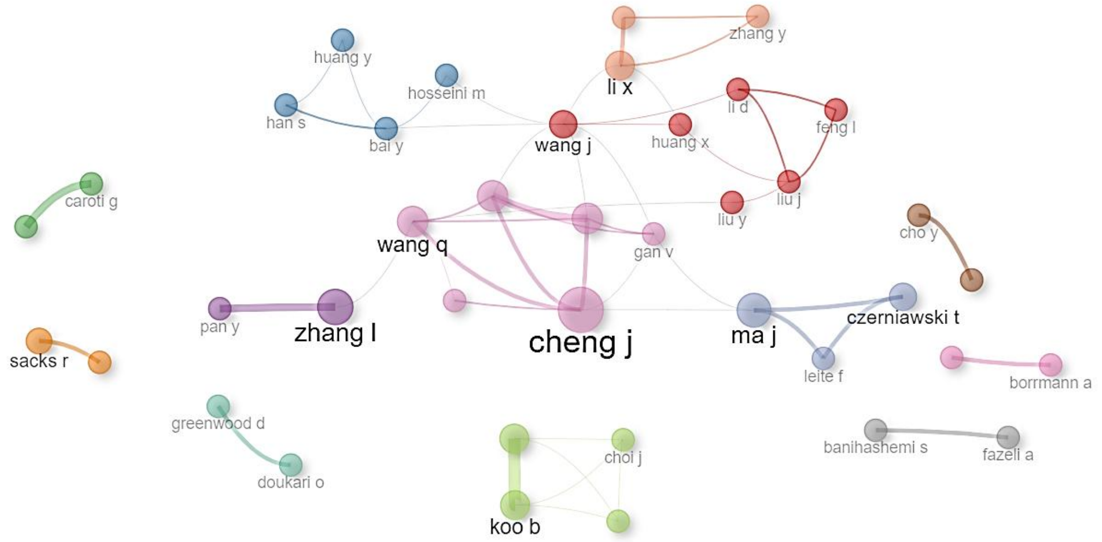
Figure 6 - Most Relevant Sources and Authors





In the collaboration network (Figure 7), Cheng J also is highlighted by its research partnerships. Wang J investigates the application of graph neural networks (GNNs) in construction, with a focus on point cloud segmentation. Li X is notable for developing IoT-enabled BIM platforms and applying parallel computing and big data to assembly services in modular construction; the author's research includes generating BIM models from 3D point clouds and assessing environmental satisfaction through energy digital twins. Bai Y explores automated modeling of historic high-rise building facades with drones and AI and uses deep learning to develop as-built models. Ma J, on the other hand, invests in automated processes for the digital fabrication of industrialized buildings, combining BIM with computational design and digital fabrication. Ma J also uses deep learning for semantic segmentation of indoor point clouds and applies image retrieval systems to facilitate facility management.

Figure 7 - Collaboration Network



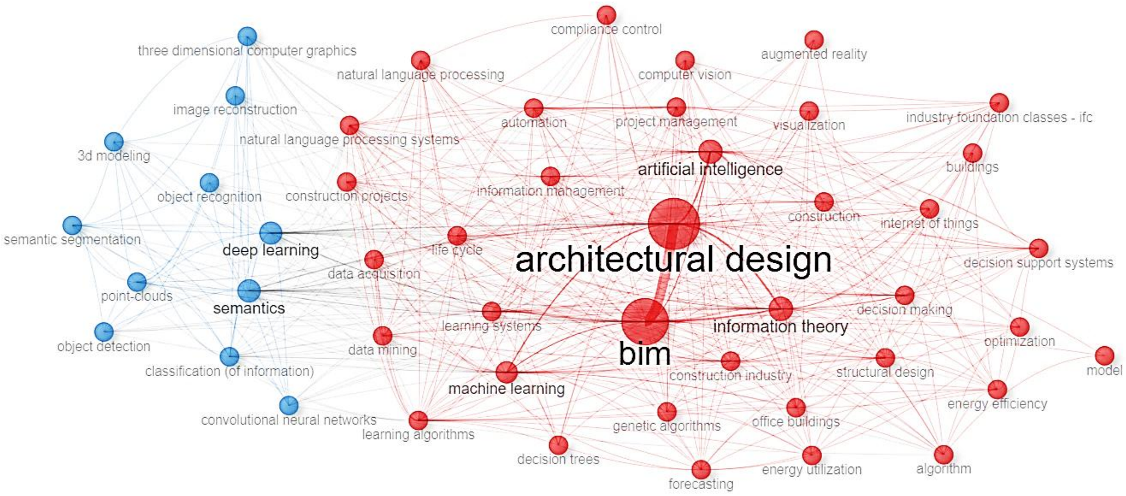
Thus, the initial analysis of the sample reveals a broad diversity of topics and applications associated with integrating the BIM and AI AECO sectors. The applications presented by the sample integrate analyses based on ML and DL for modeling construction elements, such as structural analyses, facility maintenance, performance analyses, and environmental assessments (Shen; Pan, 2023; Wang; Gan, 2023a; Wei; Akinci, 2019). The primary source of data that connects BIM and AI is modeling 3D point clouds of built environments through the development of automated model generation methods with the extraction of geometric information. In this context, the authors use AI computational models to classify, segment, or generate elements to model the geometry of BIM elements (Shu *et al.*, 2023).

In the data science domain, the sample articles develop capabilities to process complex and multidimensional data from BIM models (Kim; Kim, 2021). While most authors use point clouds as a central source of BIM data, other authors use data sources such as text (such as work order requests for maintenance), audio (in the field of architecture to incorporate customer requests), IFC libraries, time series (mainly related to energy consumption), images and videos (both for real-time monitoring and for generating 3D models) (Kayhani *et al.*, 2023; Ma; Leite, 2022; Zhang; El-Gohary, 2023; Zhou *et al.*, 2022).

Thus, the variety of AI techniques presented in Figure 6 are employed to process this diversity of data. For example, CNN can process and analyze images and videos, and natural language processing can be applied to analyzing textual data such as work order requests. On the other hand, genetic algorithms can be used to optimize design elements and computer vision can be used in 3D reconstruction from multiple images (photogrammetry) to recognize specific image patterns (Chen *et al.*, 2023b; Matrone; Martini, 2021).

Figure 8 is generated through graph theory and is used to model relationships between pairs of objects. A graph consists of nodes that are interconnected by edges. In the context of scientific mapping, network graphs are used to represent co-occurrences between bibliographic metadata. The basis for this representation is a co-occurrence matrix, where non-diagonal elements indicate how often two items, such as words, occur together in the same corpus (such as keyword lists, titles, or abstracts). Diagonal elements reflect the frequency of each item within the document collection. In this way, the network is organized into colors. The colors in the graph indicate the clusters to which each word is associated, with each cluster representing a research field within the analysis (Aria; Cuccurullo, 2017).

Figure 8 - Co-occurrence of keywords

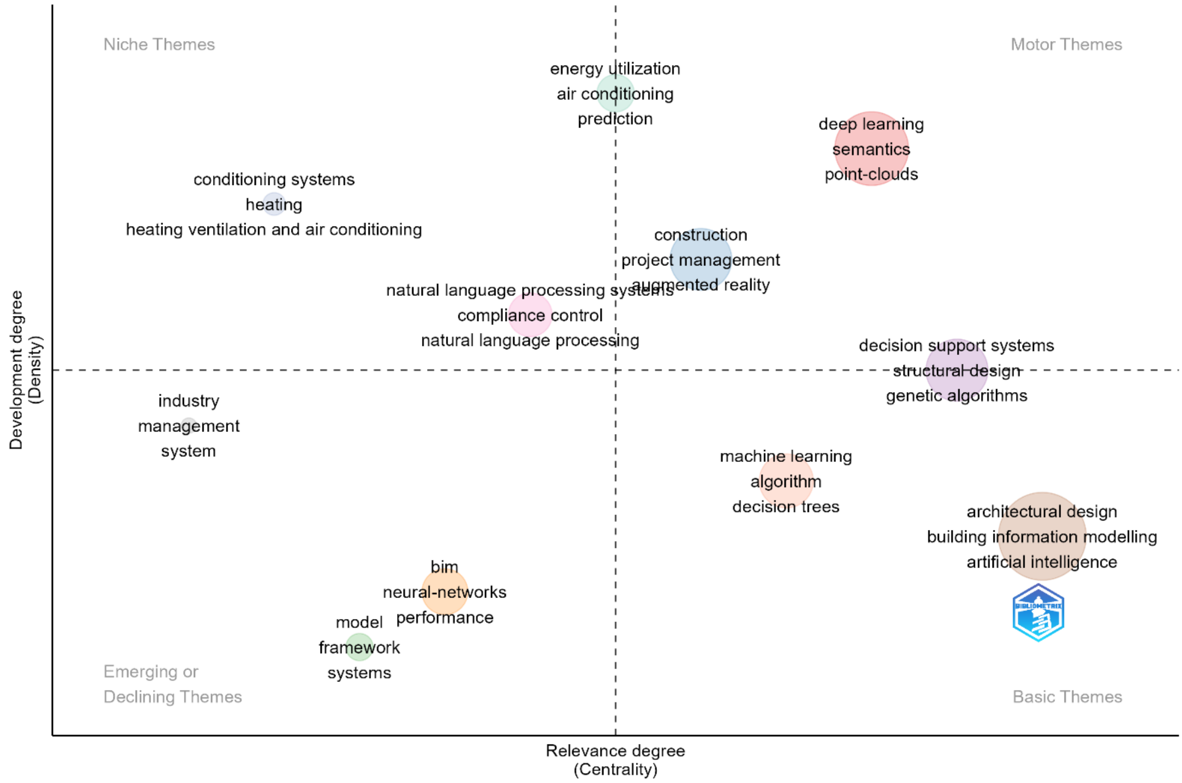


These data science topics and approaches are integrated into the AECO industry to establish mechanisms for processing automation and, consequently, the development of intelligent projects. Thus, research topics related to research techniques that can be used in different disciplines of the AECO industry arise, as shown in Figure 7.

The Thematic Map in Figure 7, generated by applying a clustering algorithm to a keyword network, outlines the themes in the BIM and AI domains. This algorithm groups keywords based on their associations and co-occurrences. In this way, distinct themes or clusters emerge. Centrality indicates the relevance of a theme within the broader field, and Density reflects the internal strength and cohesion of the theme, suggesting its development and the maturity of research within that cluster. The main limitation of this approach is that each keyword is associated with only one topic. Furthermore, the indicated topics must be analyzed in depth through content analysis to verify whether the thematic indications fully reflect the topics of the contributions (Aria; Cuccurullo, 2017).

Thus, Figure 9 has thematic groups related to architectural design, mainly about treating point clouds as a data source. Specifically, the algorithms seek to optimize separating spatial-spectral attributes into their constituents in cloud segmentation. Point cloud algorithms segmentation, for example, enables the capability to extract relationships between neighborhoods, graphs, and topology, allowing the transition from subsymbolic to symbolic 3D data analysis (Ma; Leite, 2022; Poux *et al.*, 2022).

Figure 9 - Thematic Map for BIM e AI in Smart Construction Projects



4.3.2 Research core domains

An in-depth analysis of all articles in the sample was conducted to enhance the results from Figures 8 and 9. This stage involves categorizing the articles into central topics correlating data science techniques and BIM. These research topics were identified and tabulated in Table 2. Seven application domains of BIM and AI in the AECO industry were identified. These applications guide recent contributions to literature and integrate different ways AI and BIM are associated with generating essential capabilities for data processing. The workflow generally involves the Scan-to-BIM process, which comprises point cloud-based data processing.

To identify the seven domains presented in Table 2, this paper coded the findings of the sample through content analysis, starting from the analysis of the co-occurrence of keywords to group the articles into related themes. Through this iterative process, seven key domains emerged that presented greater frequency in the analyzed publications, representing the most prominent BIM and AI integration areas. To this, the paper initially thoroughly read all the selected articles, extracting key concepts, recurring themes, and specific BIM-AI applications. These concepts and themes were transformed into initial codes, representing information units. Thus, the articles are grouped into domains based on similarities and thematic relationships, forming categories. The codes and categories were constantly reviewed and adjusted through an iterative refinement process.

This chapter argues that literature advances mainly in seven domains to create data processing and interpretation capabilities. This research further distributes these seven key research domains over the years in Table 3. The following subsections detail this evolution with current contributions on the topic.

Table 2 - Seven key research domains and applications

Domains	BIM and AI integration in construction phases	Problem to solve	Output	Source ID
Facility management	GA, CV, and RF frameworks are applied to acquiring and processing point cloud data, converting BIM models into real-world images, and focusing on operations and building maintenance. The sample uses a Scan-to-BIM workflow to monitor and forecast anomalies in the 3D model.	Provide diverse data collection and analysis methods to automate the generation of various scenarios and as-built projects.	Real-time object detection, predictive clash models, visualization of anticipated facility anomalies, and generation of automatic as-built models.	19,20,54,95,102,112,113,175, 178,190,197,212,215,276,295.
Sustainability	A multi-criteria decision support system, semantic classification, GA, ANN, and support vector machines are employed to assess a building's energy consumption in operation. Data can be extracted with Dynamo and point clouds, applied in the planning phase for BIM-based energy simulation, and a dashboard for circular building design has been proposed.	Optimizing energy performance, designing sustainable buildings, identifying ideal scenarios, automating material classification, and developing digital twins for predictive monitoring.	Assessing energy performance, creating a decision support system for sustainable design, identifying optimal scenarios, classifying building materials, and developing a digital twin.	5,6,8,25,44,61,65,80,84,137, 150,160, 167,177,237,239,284, 298.
Energy Solution / Consumption	The utilization of CNN, LR, RF, Support Vector Regression, Decision Tree, and ANN is directed toward analyzing photogrammetry, point cloud data, and time series data related to energy consumption, volume, and floor area size. This is applied in the planning phase to explore alternative solutions by developing a 3D model for energy analysis.	Alternative energy consumption in buildings enhances energy efficiency, integrates renewable systems, and enables continuous monitoring for informed, sustainable decision-making.	Energy forecasts, automatic generation of renovation scenarios based on user preferences, and better economic solutions for energy consumption	27,85,138,139,140,146,149, 152,168, 170,186,200,241, 245,246,263,268,291,310,314, 317.
Heritage BIM (H-BIM)	SS (with RF, CNN, and Pixelwise) is applied in point cloud scanning, space grid structures, photogrammetry, and Scan4Façade. Use Scan-to-BIM workflow to provide accurate information for BIM feature creation in Revit.	Optimize the analysis of data generated from historic buildings, including data from point cloud clustering methods.	Reconstruction of template geometries of classes of architectural elements.	14,21,39,68, 71,78,104, 130,133,134, 135,154,159, 162,206,308,311.
Time and Cost Optimization	GA and KNN techniques are employed in Planning, specifically for creating a 4D BIM model integrating bill of quantities, 3D BIM model, and productivity factors (labor, equipment). A 5D BIM model cost database is developed by incorporating cost data such as equipment, person-hours, materials, and overhead costs into the model, using time series information.	Inefficient schedules and budgets. The authors provide real-time schedules, highlighting the need to improve project status awareness to generate more realistic schedules.	The plugin reduces project costs and time, provides a cash flow dashboard with 5D BIM mode, and provides a dashboard for cost estimation.	1,11,12,24,26, 58,67,72,76,98,101, 105,155,157,181, 189,203,204,234,265,270,280, 283,322.
Object-based computer modeling	CV is applied with augmented reality and virtual reality data, contributing to design, planning, monitoring, and control activities. The integration of the Internet of Things is utilized to manage diverse tasks and activities across various construction and operational phases, supported by BIM-GIS integration.	Improve the generation, control, and optimization of digital models to accurately reflect structures or systems' physical and functional state.	Provides a real-time visualization of the building in an interactive 3D map connected to analytical dashboards for management support. Data-centric management.	36,38,46,77,83,106,141,151, 184,188,191,207,222,224, 231,240,264,324.
Generative Design (Architectural and structural drawings)	Employing techniques such as the Mask R-CNN algorithm, CNN, and ANN, this domain enhances the design process for concrete and steel buildings. A BIM generation model is created by integrating architectural and structural blueprints with 3D scan data. This model enables a seamless BIM-to-BIM workflow, automating the generation of optimized BIM models.	Optimizing the process of generating design alternatives in architectural projects, structural analyses to ensure stability, and efficiently implementing automated systems for creating libraries and object recommendations based on images and point clouds.	Generation of design alternatives in architectural projects, optimizing analyses for structural projects with a focus on stability. Additionally, it investigates the automated creation of libraries based on images and object recommendation systems.	13,23,28,29,31,32,33,37,43,51, 52,54, 59,63,70,73,75,90,93,99, 107,114,115,116,136,143,180, 183,192,196,201,208,209,211, 214,219,220,221,227,232,275, 296,297,304,309,318.

Note: Computer Vision (CV), Genetic Algorithms (GA), Random Forest (RF), Artificial Neural Network (ANN), Convolutional Neural Networks (CNN), Linear Regression (LR), Deep Neural Network (DNN), K-nearest neighbor (KNN), Semantic Segmentation (SS).

Table 3 - Distribution of key research domains over the years

Topic	2010-2011	2012 - 2014	2015 - 2017	2018 - 2020	2021 - 2023
Facility management				- BIM visualization for maintenance and identification of new MEP elements through image classification; - Localization of facility components within a building and associate them with their digital twins; - Predictive maintenance planning	- Organization and retriever photos; - Reconstruction from 3D LiDAR point clouds for MEP scenes; - Anomaly prediction using an IoT sensor; - Automatic detection of small objects
Sustainability			- Decision support system for selection of sustainable materials, low-carbon building measures selection, green building envelope design considering energy and maintainability		
					- Regional simulations for predicting natural hazard; - Life-cycle perspective for calculates design trade-offs; - Predictive monitoring of CO2 equivalent from existing buildings
Energy Solution / Consumption				- Energy prediction and simulations with EnergyPlus; - Energy Efficiency and Daylight Performance; - Generation of alternatives and evaluations of energy performance by analyzing the shape of the envelope; - Opportunities and costs of energy saving; - Learning Models for Thermal Comfort	
Heritage BIM (H-BIM)				- Application of performance enhancement techniques for deep semantic segmentation point clouds	- Three-dimensional (3D) reconstruction; - BIM and geographic information system (GIS) for operation and maintenance; - Automated vision-based construction progress monitoring
Time and Cost Optimization		- Estimation activity; - Cost analysis for predictive models		- 4D BIM model for the optimization of material layout of task scheduling	- Schedule optimization
Object-based computer modeling				- Automated data acquisition for digital twin information systems and IA	- Combination of computer vision with semantic analysis
Generative Design	- Automated generation of parametric BIMs based on hybrid video and laser scanning data				
				- BIM object classification, clash detection, automating design options; - Rule mining for construction detailing; - Natural language processing for predicting design commands; - Pipelines to extract 2D digital information from floorplans - Reinforcing details construction design; -Generation of BIM model from structural and architectural plans; - Automated building layout generation	

4.3.2.1 Facility Management

In Facility Management, the focus has shifted from BIM visualization for maintenance and identification of new MEP elements to the integration of digital twins and IoT sensors. This has enabled the creation of predictive models for facility maintenance and automatic anomaly detection. The research advances in generating and capturing different scenarios for automatically generating as-built projects (Chen *et al.*, 2023b). Similar processes occur in the HBIM domain, which uses laser scanning or photogrammetry techniques to digitally represent buildings (Mohammadi *et al.*, 2023; Tavolare *et al.*, 2023).

Marzouk and Zaher (2020) advanced the use of AI in FM by employing a deep learning model for image classification of MEP elements and an expert system integrated with an Android application to identify required maintenance tasks. Villa *et al.* (2022) introduced anomaly prediction models for sustainable building maintenance using IoT sensor networks and BIM models, demonstrating the ability to predict faults in HVAC systems and visualize data in 3D building models in real-time. Cheng *et al.* (2020) proposed a predictive maintenance planning framework combining BIM and IoT, featuring information and application layers that integrate data and apply ML algorithms to forecast future conditions of MEP components.

4.3.2.2 Sustainability

In Sustainability, research has evolved from decision support systems for selecting sustainable materials and low-carbon measures to regional simulations that predict natural hazards and life cycle assessments that calculate design trade-offs, as well as predictive monitoring of CO_2 emissions. Data extraction is generally done using Dynamo and point clouds, applied in the planning phase for BIM-based energy simulation and the development of a control panel for sustainable building design (Caterino *et al.*, 2021; Garcia-Gago *et al.*, 2022; He *et al.*, 2021).

The findings of B. Wu and Maalek (2023), Arsiwala *et al.* (2023) and C. Wang *et al.*, (2021) integrate digital technologies into renovation projects and asset management. B. Wu and Maalek (2023) developed an intelligent decision support framework for aging buildings, considering sustainability throughout the life cycle. They integrated BIM, point cloud processing and structural optimization, assessing the environmental impact of renovation or demolition and requalification solutions, focusing on cost, energy consumption and carbon emissions. Arsiwala *et al.* (2023) presented a digital twin solution that automates the monitoring and control of CO_2 equivalent emissions in existing assets, combining IoT, BIM and AI. The research showed that these technologies are essential for facility management, enabling the visualization of critical spatial information and prediction of carbon emissions through an AI-supported system, displaying the results in an interactive dashboard that facilitates the implementation of data-driven retrofit strategies. C. Wang *et al.* (2020) introduced a framework for generating and collecting information on a regional scale for risk analysis, using deep learning to extract building data from satellite and street images, contributing to the creation of semantic profiles of buildings in a city.

4.3.2.3 Energy Solution / Consumption

In Energy Solutions, developments have been marked by simulations of energy performance and thermal comfort, exploring opportunities for energy savings, and learning models for energy efficiency, such as studies of Erişen (2023) and Hou *et al.* (2022).

Erisen (2023) highlights the use of energy-efficient automated systems, combining BIM, IoT, and optimization algorithms to monitor and control thermal comfort parameters, such as natural ventilation. The research applies machine learning models to optimize the operation of these systems and deep learning models to predict user activities and thermal comfort levels, resulting in the optimization of energy use in smart buildings. Hou *et al.* (2022) propose a prediction and optimization framework to balance thermal comfort, indoor air quality, and energy consumption in HVAC systems, using BIM, simulations, and the Extreme Learning Machine (ELM) model optimized by the Grey Wolf Optimizer algorithm. The research showed that after optimization, the average CO_2 concentration was reduced and thermal comfort was kept within acceptable limits, resulting in energy savings of 14.34%.

4.3.2.4 Heritage BIM (H-BIM)

In Heritage BIM (H-BIM), advances range from semantic segmentation of point clouds to 3D reconstruction integrated with GIS, facilitating the operation, maintenance, and monitoring of progress in historic buildings. Jiang *et al.* (2022), Croce *et al.* (2023), and Pierdicca *et al.* (2020) propose methods for the digital modeling and reconstruction of historic buildings, using drones, photogrammetry, and machine learning.

Jiang *et al.* (2022) present the Scan4Façade method, which uses drones to capture images of historic building facades and employs photogrammetry to create 3D models. An AI model (U-net) segments the generated orthoimages, and a clustering algorithm extracts dimensions and coordinates of facade elements, achieving high accuracy in window extraction. Croce *et al.* (2023) propose an approach for automated Scan-to-BIM reconstruction, using semantic segmentation with Random Forest and geometric reconstruction of architectural elements. Pierdicca *et al.* (2020) develop a Deep Learning framework for point cloud segmentation in the ArCH Dataset, which includes point clouds of architectural heritage, both internal and external.

4.3.2.5 Time and Cost Optimization

In Time and Cost Optimization, innovations ranged from predictive cost analysis to schedule and layout optimization with 4D BIM. Peiman *et al.* (2023) and Huang and Hsieh (2020) explore approaches for schedule and cost forecasting in construction projects using machine learning models and data mining techniques. Peiman *et al.* (2023) developed a gradient boosting ensemble model to estimate the completion duration of construction projects using legal and institutional variables. They used data from 30 projects of different building types and 426 follow-up periods to train and test the models using 17 dimensional variables, including EVM performance indices. Huang and Hsieh (2020) proposed a hybrid methodology based on CRISP-DM, combining Random Forest (RF) and Simple Linear Regression to improve the accuracy in forecasting labor costs in BIM projects in the construction phase. Based on case studies of 19 completed BIM projects in Taiwan, they developed a cost decomposition framework to train machine learning models and proposed using effective area instead of gross area as an input variable, improving model performance with clustering analysis.

4.3.2.6 Object-based computer modeling

Object-Based Computational Modeling combined computer vision with semantic analysis for automated data acquisition in digital twin systems. Meschini *et al.* (2022), H. Wu *et al.* (2021) and T. Wang and Gan (2023)explore approaches to object-based computational modeling, focusing on improving building management and safety through technologies such as BIM, GIS, computer vision and machine learning. Meschini *et al.* (2022) develop a BIM-GIS platform for the Operation and Maintenance phase of an Italian university campus, integrating spatial and functional data to create an interactive 3D visualization of assets, which aids decision-making and safety in emergencies such as fires. Wu *et al.* (2021) present a framework that combines computer vision and ontology to manage safety on construction sites, detecting hazards through images and inferring mitigators based on rules defined in semantic language. T. Wang and Gan (2023) propose an automated computer vision approach for 3D reconstruction and visual inspection of buildings, using transfer learning to identify surface defects in 3D reconstructed scenes, applying advanced models such as ResNet-50 and Grad-CAM techniques.

4.3.2.7 Generative Design

Generative Design, there was a leap from automated generation of parametric BIMs to advanced techniques such as clash detection, rule mining for detailing, and automated generation of layouts and BIM models from structural and architectural plans. The central focus is to optimize the process of generating design alternatives in architectural projects, performing structural analyses to ensure stability, and implementing automated systems efficiently to create libraries and object recommendations based on images and point clouds (Garcia-Gago *et al.*, 2022; Van Der Zwaag *et al.*, 2023).

Recent studies by Abdirad and Mathur (2021), Leon-Garza *et al.* (2022), Frías *et al.* (2022), Urbieta *et al.* (2023), and L. Wang *et al.* (2023) illustrate the advances in generative design. Abdirad and Mathur (2021) developed a BIM content recommendation system that improves the accuracy in predicting content needs using unsupervised machine learning and association rule mining. Leon-Garza *et al.* (2022) presented an innovative approach that transforms 2D floor plans into 3D BIM models through type-2 fuzzy logic, increasing the interpretability and adjustability of the processes. On the other hand, Frías *et al.* (2022) proposed a deep learning framework for object classification in point clouds with high accuracy, using synthetic data and orthographic projections for training. Urbieta *et al.* (2023) advanced the automation of BIM model creation from architectural drawings using Mask R-CNN, facilitating the integration of diverse geometric representations. Finally, L. Wang *et al.* (2023) developed a framework for automatic generation of building layouts, demonstrating the ability to produce accurate and visually verifiable layouts with the help of the U-Net network.

4.3.3 BIM and AI Capabilities for Smart Buildings

For organizations in the AECO sector to proactively benefit from the advantages of BIM-AI integration, this paper argues that essential BIM and AI capabilities are needed in smart building management. First, BIM capabilities refer to the skills and functionalities that a BIM platform or system offers to improve the design, construction, and management process of buildings and infrastructures (Munianday *et al.*, 2022; Yilmaz *et al.*, 2023). Second, AI capabilities refer to the skills and functionalities embedded in systems or platforms that use AI techniques to improve processes, make automated decisions, and learn from data (Mikalef; Gupta, 2021). These capabilities become strategic artifacts to select, orchestrate, and leverage their specific capabilities from BIM and AI models. To answer research question 1 (#RQ1), this paper maps two types of capabilities oriented to data management and application.

This paper maps BIM capabilities in Table 4 and argues that these capabilities underline the evolution of organizations in the AECO sector, providing process automation and project optimization. First, BIM platforms serve as a centralized and accessible source of project-related information. Additionally, BIM models offer details about project geometry, materials, and components (Tixier *et al.*, 2016; Zheng; Fischer, 2023). This wealth of information becomes essential for AI algorithms, especially when considering performance data such as energy consumption and structural analysis. In the context of computer vision, BIM provides a detailed visual representation of the project and allows the integration of real-world data such as point clouds. This ability to connect virtual to the real enriches AI analysis even further. The feasibility of this integration is enhanced by standardization in the use of BIM, facilitating interoperability between diverse AI systems and tools and establishing a solid foundation for advanced analytics (Fenz *et al.*, 2023; Garcia-Gago *et al.*, 2022). Furthermore, efficient data management in BIM environments provides a robust framework for dealing with complex sets of information, a fundamental necessity when training AI models with large volumes of data (Tavolare *et al.*, 2023; Tixier *et al.*, 2016).

Table 4 - Key BIM Capabilities for Smart Constructions and AI

Code	Capability	Source ID
BIM1	Continuous planning and monitoring of components	3,17,19,20,49,69,74,80,119,128,134,148,151,166,174,176, 179,212,234.
BIM2	Continuous utilization of data throughout the project lifecycle	53,72,73,77,84,110,115,131,149,155,157,163,167,168.
BIM3	Creation of BIM-based decision support system	5,6,8,15,21,25,41,43,48,51,55,59,61,75,106,123,132,137, 148,150,152,155,157,161,288.
BIM4	Data Management	26,40,93,116,150,166,167,126,170,176,190,257.
BIM5	Digital representation and data integration	29,31,33,38,39,59,70,78,81,89,91,93,96,99,104,107,115, 129,136,137,138,141,142,143,149,118,126,127,176,190.
BIM6	Digital representation of existing ventures	2,3,14,21,31,39,55,57,71,78,93,104,116,121,125,130,132, 133,135,153,154,159,162,206,231,308,311.
BIM7	Enhanced project visualization	19,33,53,72,95,110,113,114,127,141,151,166,169,316.
BIM8	Data control for point cloud processing	3,39,68,71,78,96,114,153,154,156,159,162,170,192,204,206, 222,224,226,231,236,243,253,269,276,281,300,323.

BIM9	Multidimensional Modeling (3D, 4D and 5D)	24,26,30,33,42,45,50,64,73,94,96,110,114,116,125,127,135, 143,127,151,155,156,157,162,163,167,175,179,188,189,192, 208,222,224,227,243,259,262,265,275,287,297,308,319,324.
BIM10	Multidisciplinary Integration	28,35,54,62,65,66,72,74,146,112,167,168,172
BIM11	Accurate Performance Analysis	14,27,58,69,82,126,138,139,140,152,155,157,161,166,168, 169,173,177,179,246,249,299,303,314.
BIM12	Time dashboards development	6,17,20,63,134,141,151,161,164,167.
BIM13	Simulation and Dynamic Analysis	6,8,11,15,28,44,55,67,69,85,86,87,140,149,150,152,155,162, 168,173,179,289,316.
BIM14	Standardization with IFC schema	3,9,33,37,57,90,92,112,114,118,132,139,152,160,164,166, 174,181,187,254,261,313.

Integrating data science with BIM establishes mechanisms for creating and enhancing new capabilities in the AECO industry. The digital representation of buildings generates complex and multidimensional databases, opening a wide field of research for literature. This digital representation goes from existing buildings' planning, monitoring, and reproduction phases (as happens in the context of HBIM). Data management through BIM capabilities improves the use of data by AI by continuously storing information throughout the life cycle of projects (X. Chen *et al.*, 2023; Marzouk & Zaher, 2020). This data can then be processed to improve different forms of project representation, from incorporating specific customer requirements to integrating sustainability-oriented alternatives. These capabilities can nurture data-based decision-making systems that can be generated from performance, safety, comfort, and energy simulations. However, all of this is only possible through the ability of BIM models to standardize their data across IFC schemas (Van Der Zwaag *et al.*, 2023). In this sense, this paper suggests the following research proposition:

Proposition 1: BIM capabilities can be positively related to the development of smart projects when associated with AI techniques

Table 5 identifies the key AI capabilities in the sample, highlighting their importance for driving innovation and efficiency in the AECO industry. AI capabilities incorporate technical skills to handle data and implement AI techniques, as well as managerial skills to understand how each available AI algorithm can be applied in each discipline or construction stage. For example, data mining capabilities enhance innovation processes, while predictive algorithms enable anticipating results and trends (Lin *et al.*, 2022; Muhammad *et al.*, 2021). Furthermore, AI offers interface customization, real-time control, and risk prediction, promoting a proactive approach to management. The ability to process large volumes of data quickly and efficiently enables more accurate and predictive analyses, contributing to the optimization of construction processes (Zhang *et al.*, 2022). The application of advanced algorithms can improve operational efficiency, reduce costs, and mitigate risks by anticipating potential problems during the planning and execution phases of projects (Caterino *et al.*, 2021). The ability to perform simulations and dynamic analyses also contributes to developing more sustainable and energy-efficient projects (Villa *et al.*, 2022; P. Wu *et al.*, 2023).

Table 5 - Key AI capabilities for Smart Constructions and BIM

Code	Capability	Source ID
AI1	Advanced Cloud Services	46,88,91,106,108,117,120,122,132,134,135,144,147,127.
AI2	Equipment for point cloud acquisition for building reconstruction	39,71,104,121,125,130,133,135,153,154,159,162,276,271, 309,319,324.
AI3	Capability to develop algorithms for recommendation of multiple design options	5,11,18,28,29,89,116,136,138,140,113,160,161,171,178,198, 211,214,275,301.
AI4	Automated point cloud processing	3,7,22,38,39,68,78,96,100,104,116,129,130,135,137,142,147, 153,154,156,159,162,170,174,178,192,204,206,207,210,222, 223,224,226,231,233,236,243,253,269,276,278,281,282,286, 289,293,300,312,320,323.
AI5	Automatic BIM models Generation	1,7,26,32,37,52,57,95,97,98,114,116,129,141,115,156,160, 161,167,169,171,180,243,286,296.
AI6	Automatic Object detection	7,60,71,90,100,107,121,135,138,153 170,115. 174,175,187.
AI7	Customization of project elements (objects, libraries, structural elements)	1,5,29,32,52,63,73,114,129,136,137,114,153,162,167,169
AI8	Establishment of standardized routines for data analysis	10,40,56,80,90,91,111,113,118,126,140,157,159,168,169, 171,175,179,183.
AI9	Development of probabilistic analysis	59,150,157,165,166
AI10	Natural language Processing	10,34,76,92,103,109,111,145,182,194,202,255.
AI11	Performance Evaluation	11,48,49,58,82,85,127,139,140,149,162,166,167,169,170,171, 188,246,249,299,303,314.
AI12	Predictive modeling	12,19,23,24,27,63,67,76,102,103,106,109,113,124,149,123,126, 151,155,157,162,167,168,172,173,178,200,203,212,245,252, 267,275,280,283,304.
AI13	Process unstructured data (sensors, IoT, cameras)	7,16,17,22,79,83,86,88,89,99,113,129,133,134,135,139,146, 149,152,153,154,156,158,162,164,165,167,175,258,285.
AI14	Real-time monitoring	31,70,74,100,101,127,128,147,151,156,162,167,173
AI15	Scalable data storage infrastructures	4,142,113,153,156,166,169,173.
AI16	Semantic analyzes	13,36,47,114,118,137,143,145,153,154,156,159,164,182,185, 185,197,204,222,230,231,240,266,284,293,307,313.

In this context, the AI capabilities mapped in Table 4 reveal competencies and processes for implementing BIM-AI in the AECO sector. Advanced cloud capabilities indicate the exploration of robust and scalable processing environments. Algorithms aimed at reconstructing buildings demonstrate the ability to create accurate and up-to-date digital representations of existing structures (Wu *et al.*, 2023). Automatic model generation and automatic object detection point to a practical approach to creating and analyzing digital representations, while data analysis and the development of probabilistic analyses provide a deeper understanding of project performance (Marroquin *et al.*, 2018; Marzouk & Zaher, 2020). Additionally, applying natural language processing suggests more intuitive interfaces, while performance assessment and predictive modeling contribute to proactive strategies and continuous optimization (Pan; Zhang, 2021b; Van Der Zwaag *et al.*, 2023). Unstructured data processing, real-time monitoring, scalable data storage infrastructures, and semantic analytics complement a set of capabilities, enabling effective integration of AI into construction and design processes (Novembri; Rossini, 2020; Rafsanjani; Nabizadeh, 2023).

These capabilities stand out as essential catalysts for driving innovation, which range from exploring advanced cloud environments to automating model generation and data analysis (Chen *et al.*, 2023a; Fenz *et al.*, 2023). The strategic combination of these capabilities with BIM models creates a foundation for building smart and efficient projects, promoting a proactive approach to management, agile decision-making, and continuous performance optimization (Çetin *et al.*, 2022; Villa *et al.*, 2022). Thus, it is suggested the following proposition:

Proposition 2: AI capabilities can provide subsidies for data processing from BIM models for smart buildings.

4.3.4 Benefits of BIM-based and AI Capabilities for Smart Construction

AI and BIM capabilities are applied in architectural design to automate smart layouts, reproduce existing buildings and simulations oriented towards energy efficiency, and model budgets and schedules. These capabilities generate potential benefits that organizations in the sector must explore. The capabilities enhance financial gains and proactively integrate data management that reflects the unique organizational context, enables project optimization, and helps preserve historic buildings. In this sense, Table 6 maps the five dimensions of benefits that AI and BIM capabilities generate for the AECO sector, answering research question 2 (#RQ2).

Regarding point cloud processing, (Shu *et al.*, 2023) highlight that AI techniques, specifically deep learning, have a strong learning capacity and advantages in completing object recognition tasks, which can use the Scan2BIM-NET scheme to semantically segment the construction of point clouds into structural, architectural subcomponents and mechanics. Other models that can be used in this process are PointNet, PointCNN, and Dynamic Graph Convolutional Neural Network to classify point clouds of different bridge components (SHU *et al.*, 2023). Machine Learning models can also automate the creation of BIM models from the point cloud, simplifying the Scan-to-BIM process. This includes three-dimensional reconstruction of architectural elements based on captured data (Garcia-Gago *et al.*, 2022).

Another fundamental application is to compare different point clouds over time to identify changes, updates, or progress in construction projects, which is valuable for managing as-built projects (Caterino *et al.*, 2021; Hsu *et al.*, 2020).

The authors who opt for the PointNets approach apply the analysis approach proposed by Qi *et al.* (2017). This architecture uses point clouds directly as input, avoiding the irregularities and complexities of meshes, which facilitates the learning process. PointNet is a unified framework with point clouds as input, generating class labels for the entire input or point segment/part-specific labels for each point in the input. In object classification tasks, the input point cloud is directly sampled from a shape or pre-segmented from a scene point cloud (Qi *et al.*, 2017). The proposed deep network produces scores for all candidate classes, approximating a general function defined on a set of points by applying a symmetric function to the transformed elements of the set, as shown in Equation (1).

$$f(\{x_1, \dots, x_n\}) \approx g\big((x_1), \dots, h(x_n)\big)$$

where,

$$f : 2^{\mathbb{R}^N} \rightarrow \mathbb{R}, h : \mathbb{R}^N \rightarrow \mathbb{R}^K \; and \quad (1)$$
$$g : \mathbb{R}^K \times \dots \times \mathbb{R}^K \rightarrow \mathbb{R} \; is\ a\ symmetric\ function$$

The PointNet approach approximates h by a multi-layer perceptron network and g by a composition of a single variable function and a max pooling function. By collecting h, the model learns several f's to capture different properties of the set (Qi *et al.*, 2017).

Hence, the capabilities can then be used for design customization, whether to analyze data such as point clouds and design information to generate detailed BIM models automatically or to enhance clash detection algorithms by analyzing large data sets and recognizing complex patterns that may go unnoticed (Caterino *et al.*, 2021; He *et al.*, 2021). Some construction companies need specific objects and families in their BIM models. In this sense, AI algorithms can learn from existing examples and automatically generate personalized objects and families, saving time and ensuring consistency in design. Furthermore, AI can apply natural language processing and machine learning techniques to enrich BIM data with semantic information, facilitating the analysis and interpretation of data by professionals involved in the project (Marroquin *et al.*, 2018; Marzouk & Zaher, 2020; Pan; Zhang, 2021).

The simulations carried out with AI reflect the integration of BIM into different phases of a building's life cycle. Artificial lighting simulation allows for the prior assessment of the lighting environment, providing energy efficiency and improving user experience (Carreira *et al.*, 2018). The development of individualized designs, driven by AI, represents an evolution in customizing architectural spaces to meet specific user needs. Optimizing indoor air quality, simulating energy consumption, and thermal comfort highlight the role of AI in promoting sustainable and healthy environments (Carreira *et al.*, 2018; Chen *et al.*, 2023b). Making smart decisions about sustainable building materials highlights the potential of AI in aligning architectural designs with environmental principles (Garcia-Gago *et al.*, 2022; Van Der Zwaag *et al.*, 2023).

Finally, the benefits of cost and schedule modeling range from controlling estimates to automating the detection of potential delays in the project. Capabilities are applied in these disciplines through the continuous analysis of databases from previous projects to suggest alternative methodologies for cost and time management (Alzara *et al.*, 2023; Peiman *et al.*, 2023). The focus is to predict scenarios to avoid delays and reduce expenses throughout the life of the projects. Thus, it is suggested the following proposition:

Proposition 3: Integrating BIM and AI can promote efficiency in design and construction processes, resulting in tangible benefits such as improved point cloud processing, customization, simulation, cost, and schedule management.

Table 6 - Smart Construction Benefits of AI-BIM-based Capabilities

Smart Benefits	Code	Description	Source ID
Smart Point Cloud Processing	PCP1	Control for Data Processing	14,21,104,130,133,154,182,288.
	PCP2	Feature Extraction	39,51,78,130,136,175,196,198, 210,222,235,244,278,297.
	PCP3	Geospatial accuracy for component identification	33,38,75,78,104,130,179,190, 197.
	PCP4	Identification of Changes in As-Built Projects	4,14,43,51,68,122,135,175,224, 233,286,304,306,309.
	PCP5	Model View Definition	14,78,104,134,162,185,216,294.
	PCP6	Object Classification	39,54,63,71,104,116,187,194,278.
	PCP7	Optimization for Processing	22,68,78,104,121,130, 137,159.
	PCP8	Point Cloud Segmentation	31,71,78,107,133,154,159,206,207, 222,223,236,249,253,269,272,282, 293,312,320.
	PCP9	Scan-to-BIM reconstruction	99,114,130,133,134,135,159,180, 192,208,244,269,306.
	PCP10	Automated inspection for parametric modeling	14,31,33,59,63,68,71,130,134, 154,159,324.
Smart Design Customization	DC1	Application programming interface (API)	4,41,51,52,68,95,143,145,215, 228, 245,250,287,316,321.
	DC2	Automatic generation of BIM model	37,43,59,73,75,93,99,180,224,243, 249,296,312.
	DC3	Clash detection	23,33,43,50,51,127,134,136, 184, 195,209,250.
	DC4	Custom creation of objects and families	90,93,107,116,146.
	DC5	Data Format Compatibility	32,33,37,43,73,75,114,265,317.

	DC6	Semantic enrichment	13,14,43,47,75,90,102,103,120, 131,143,164,251,254,307.
Smart Design Simulations	DS1	Artificial lighting simulation	25,27,80,84,241.
	DS2	Development of individualized design	5,6,80,150,167,271,289.
	DS3	FM-compliant planning	19,20,84,54,95,102,112,113. 150
	DS4	Optimization of indoor air quality	8,27,80,84,160,190,197.
	DS5	Simulation of energy consumption	5,25,27,61,80,138,139,140,160, 177,241,245,246,291,317.
	DS6	Simulation of thermal comfort	25,27,84,152,160,241,242,245,258, 263.
	DS7	Smart decision of sustainable building materials	5,6,25,61,137,167,183,265.
Smart Cost Modeling	CM1	Accurate cost estimates	1,67,98,127,155,203,265,280,283.
	CM2	Control of delivery of contractual milestones to avoid fines	67,94,105,126.
	CM3	Cost savings on resources (reduction of expenses associated with the use of materials, labor, and equipment)	5,12,98,105,126,157,216,245,298.
	CM4	Cost savings on rework reduction (decreased expenses related to corrections and adjustments required due to errors or changes during the construction process related to planning, execution, and coordination between different teams and disciplines)	1,72,94,119,155,271,281.
	CM5	Reduction of operational costs (Reduced preventive and corrective maintenance costs due to better monitoring and predictability of failures)	11,72,98,113,155,228,309.
	CM6	Variation detection on the project	11,72,105,126,157,183,218,233,253 280,292.
Smart Schedule Modeling	SM1	Automatic adjustment of schedules based on available resources	24,58,94,113,148,172,213,256,282.
	SM2	Automatic detection of construction delays	67,72,76,89,101,126,209,268,282, 306.
	SM3	Automatic update of project progress	1,72,76,101,112,181,189,204,321.
	SM4	Scenario Analysis (Running simulations to evaluate different scenarios and their possible implications on project deadlines)	1,24,76,94,117.
	SM5	Task overlap detection	58,72,94,101,112.

4.4 Discussion

This chapter explores the main application domains of BIM and AI in the development of smart projects in the AECO industry. It maps out the essential capabilities to be developed for this integration. This article's contributions advance the understanding of how these capabilities are applied and their potential benefits.

This article advances the body of knowledge in the AECO industry by arguing that BIM has evolved from an information management model throughout the project life cycle to application in key domains. With the international Industry Foundation Class (IFC) standard, data scientists and engineers can share and store information from different software providers and transform this BIM data into semantic elements such as XML and RDF format (Bloch; Sacks, 2020).

However, the complexity and multidimensionality of BIM data increase computational processing capacity and make manual data processing unfeasible. To overcome this challenge, the researchers in this article's sample propose strategies that exploit machine learning capabilities for understanding semantic relationships, performance simulations, real-time monitoring, and predictions based on time series.

The interconnected capabilities of BIM and AI in data science provide the necessary mechanisms for representing and analyzing tangible results from predictions and contributions in real-time. For example, in the planning phase, algorithms can use semantic information from previous project data repositories to optimize the development of projects that contain similar requirements from the owner and architects. This can be done through algorithms that use mining rules based on natural language processing to create a library or layout recommendation systems within software such as Revit. During construction, the trained models can be used in real-time planning for necessary work sequence adjustments, conflict detection, and site layout planning. In the project finalization phase, the algorithms can generate as-built models using automated approaches such as image detection, point cloud generation and processing, or laser scanning based on 3D reconstruction (Mousavi *et al.*, 2022).

Thus, through Proposition 1, this article argues that the data generated by BIM models can be considered centralized and multinational information repositories. This repository comprises data acquisition at different phases of construction. Thus, machine learning and deep learning techniques adopt an automated pipeline for processing data by extracting features and patterns in the set. Therefore, these AI models establish solid analytical capabilities for data information, enabling them to analyze complex patterns with multiple variables. Data scientists can adjust the necessary parameters of the algorithms based on the experiences of professionals in the AECO industry according to the proposed requirements (Mohammadi *et al.*, 2023; Tavolare *et al.*, 2023).

Proposition 2 endorses the wealth of information in BIM models that can be analyzed to generate smart projects in the AECO sector. The algorithms used by AI process this information to represent existing buildings, create a greater diversity of layout options, objects, and families for architectural projects, and optimize the customization of objects to meet the specific needs of a given organization. This data can be structured or unstructured. In the case of structured data, information is organized in a formal and predefined way. They are organized in a specific format, like tables in relational databases, where each field has a predefined data type. This is the case with databases that store information about construction projects, including details about the parties involved, schedules, costs, and materials (Zhou *et al.*, 2019). Another example is 3D digital models that contain detailed information about the geometry, materials, and components of a building, such as a BIM model in software like Revit, which has distinct categories for walls, floors, HVAC systems, etc. Unstructured data is information that does not have a predefined form or rigorous organization. They do not fit easily into a tabular format or have a fixed schema. Unformatted texts, images, videos, audio, and documents in natural language can be cited as examples of unstructured data (Hong *et al.*, 2021). These data types generally do not have an organized structure in tables or fields.

In this sense, Proposition 3 suggests that integrating BIM and AI can drive efficiency in design and construction processes, resulting in tangible benefits. Specifically, the paper highlights improvements in point cloud processing, customization, simulation, cost, and schedule management. The benefits range from the ability to process and analyze large data sets to adaptive customization and more accurate simulations.

It is through capability mapping, proposition suggestion, and knowledge of BIM and AI application domains that a theoretical framework is suggested, as presented in Figure 10. The Framework develops an approach for analyzing and processing data from BIM models and their respective contributions to domains in the AECO sector, which generate potential benefits. First, it is endorsed that the information in BIM models can be collected by defining the model and updating it whenever necessary. The structure suggests identifying the type of data (whether it is structured or not) so that this data can be collected to consolidate a database.

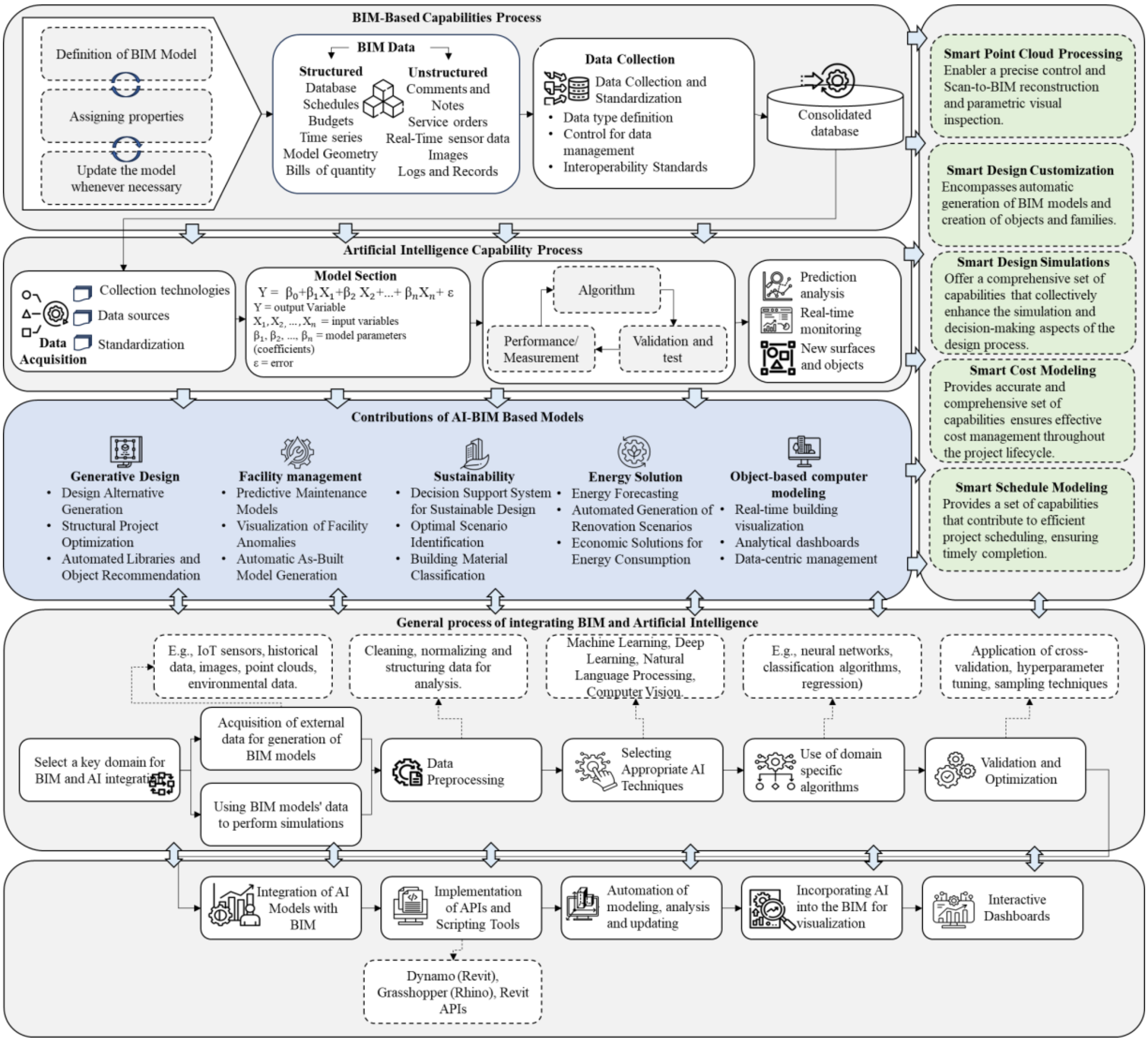
This database is the input for analysis in AI models. These analyses are based on full knowledge of the type of data to be processed to select the AI model to be used. Based on this selection, the algorithm processes the information through validations and tests to generate performance measurement parameters, such as accuracy. In construction projects, this process creates forecast analyses, models for actual monitoring of construction sites or existing buildings, and the creation of new designs or objects to compose BIM libraries. These processes are then applied in different disciplines, such as generative design, FM, sustainability, energy solutions, and object-based modeling, as mapped in this study ;(Marzouk Zaher, 2020; Villa *et al.*, 2022; Xia; Gong, 2022). The main output of these processes and capabilities are smart benefits, such as point cloud processing, design customization, design simulations, cost modeling, and schedule modeling.

For the implementation of the proposed framework, researchers and professionals must understand that each of the seven highlighted domains has a specific approach to data creation and analysis. For example, facilities management begins with data collection and preparation. This involves extracting information from BIM models, such as assets, MEP systems, maintenance history, and equipment specifications. Simultaneously, integration with IoT devices allows real-time monitoring of system performance, environmental conditions, and space usage. Data preprocessing includes cleaning, normalization, and proper structuring in databases. Then, AI models are developed using machine learning techniques for predictive maintenance. Algorithms such as Random Forest, XGBoost, and Artificial Neural Networks help predict equipment failures. Natural language processing (NLP), with tools such as NLTK and SpaCy, analyzes service requests and user feedback, identifying patterns and prioritizing tasks. The implementation of recommendation systems uses association rule mining (Apriori algorithms, FP-Growth) to identify relationships between failures and environmental conditions, as well as unsupervised learning (K-Means, DBSCAN) to group similar assets and optimize maintenance strategies. Visualization and integration with BIM are performed through tools such as Autodesk Forge and Revit APIs, allowing the incorporation of the results of AI models into the BIM model for 3D visualization and contextual analysis. The main tools include programming languages such as Python and R, machine learning platforms such as TensorFlow and Scikit-learn, databases such as SQL and MongoDB, and BIM software such as Autodesk Revit and Dynamo.

In the sustainability domain, BIM can be integrated with environmental impact databases, using tools such as SimaPro and OpenLCA. AI models apply machine learning to predict environmental impacts based on project parameters. Multicriteria analysis can be applied through genetic and evolutionary algorithms to balance cost, performance and environmental impact, with simulations performed in tools such as EnergyPlus and DesignBuilder. Specifically in the field of H-BIM, it begins with the acquisition of data through 3D surveys using laser scanning (LiDAR) and photogrammetry, including the use of drones for hard-to-reach areas. Initial processing involves the registration and alignment of point clouds to form a complete model, followed by filtering and cleaning to remove noise. Segmentation and classification use neural networks for point clouds (PointNet, PointCNN) to classify architectural and structural elements, in addition to traditional machine learning algorithms (Random Forest, SVM) for semantic segmentation. BIM reconstruction and modeling are automated with plugins such as Scan-to-BIM, converting point clouds into parametric BIM objects.

Finally, in the topic of computer vision, object detection and recognition are performed with computer vision, using models such as Mask R-CNN and YOLOv5 to identify components in images or point clouds. Furthermore, in the context of generative design, semantic analysis and ontology creation use tools to define relationships between objects and attributes. Automation of tasks in these two topics is achieved through scripting APIs with tools such as Dynamo and Grasshopper, automating modeling and analysis processes. BIM integration and updating include data synchronization to update the model with information from AI. The main tools and techniques are programming languages (Python, C#), computer vision frameworks (OpenCV, TensorFlow), and BIM tools (Revit, Rhino).

Figure 10 - Integrative Framework for BIM and Artificial Intelligence Capabilities in Smart Architecture, Engineering, Construction, and Operation Projects



The process suggested by the Framework proposed in Figure 8 establishes that BIM and AI capabilities are interconnected from data generation and processing until completion with the potential benefits in the different domains mapped in this article. In this context, BIM emerges as a data supply and centralization model, expanding its data management and standardization capabilities. This data ranges from semantic information through point clouds to control data on costs, schedules, and energy consumption. In this context, AI capabilities require professionals working in the AECO industry to know the area of data science, ranging from basic programming knowledge to advanced machine learning and deep learning models.

Figure 10 establishes that it is necessary to recognize which BIM and AI capabilities are essential for the data-driven technological development that AECO industry organizations must establish. This process guides researchers and practitioners to recognize what is technically feasible, assess the readiness of these technologies for implementation, and identify areas where further development or innovation is needed within the organization, aiming for long-term growth. By understanding the benefits of BIM and AI integration, tangible and intangible advantages are articulated. In addition, the interaction between capabilities and benefits facilitates strategic planning. This allows prioritizing specific capabilities that generate potential benefits for the organization so that resource allocation is guided by organizational strategy.

Given the complexity of the AECO industry, process automation essentially comes from the multidisciplinary knowledge and skills of the data science and AECO areas. AECO professionals have specific knowledge about project needs and their technical and functional requirements. Data science professionals know the fundamental means and techniques for developing intelligent design-oriented models based on your requirements. Therefore, a professional who has mastered these two fields of knowledge is equipped with skills that can overcome several difficulties in the AECO industry. Thus, the framework presented in Figure 10, in addition to presenting a workflow and its benefits, can suggest new directions for future research on how engineering and architecture schools can integrate data science disciplines to train an entire generation of professionals focused on smart projects and process automation.

Although it is not the focus of this article, it should be noted that investments in materials are necessary for the AECO industry to evolve towards the practical application of AI techniques. All machinery, such as scanners, computers, drones, and servers, require monetary investments. Furthermore, despite the gains related to time due to the automation of algorithms' analysis, there is time to be dedicated to training qualified professionals with the capabilities for these tasks. Finally, professionals in the AECO industry can implement the framework presented in Figure 10 by establishing essential data management and analysis skills. By mapping essential capabilities, professionals can use them as a guide to good practices and fundamental elements to effectively implement AI techniques in their organizations. Given these capabilities, professionals can compare the benefits generated with those mapped in this research.

4.4.1 Future Directions

When utilized as a data source for AI, BIM capabilities facilitate simulations designed to optimize the development of smart projects. The process initiates with the definition of the criteria or variables to be integrated into the models. Subsequently, the most promising algorithms are selected, and their fundamental characteristics are detailed. The models, whether they are classificatory or predictive, are then trained using these selected algorithms. To enhance the robustness of the models, cross-validation methods and sampling techniques are employed. Finally, tests are conducted to identify the algorithm that performs best, which is subsequently applied in specific case studies.

Wang and Gan (2023) develop a model for crack detection in buildings, setting the stage for future research to train models with more diverse datasets. These datasets could include various types of anomalies encountered in built environments, such as advanced structural wear, corrosion, or water damage. Literature underscores the necessity to create 3D models that incorporate a rich semantic representation of objects. Future research could explore fully machine-learning-based methods to generate 3D models that encode pixel characteristics into high-level representations, enhancing the depth and utility of the generated models. To overcome the limitations in generating 3D models, researchers such as Su *et al.* (2023) suggest the use of Zero-reference Deep learning model for the low-light image Enhancement for underground utilities 3D reconstruction (ZDE3D). Furthermore, much effort can still be directed towards automating the reconstruction of as-built BIM models from the point cloud. Another approach would be investments related to real-time monitoring systems for detecting changes in the pipeline structure, such as wear and obstructions.

Following the logic of applying AI capabilities to smart projects, there exists an opportunity for developing algorithms capable of detecting building materials and geometries across various scenarios in the AECO industry, both during the design phase and in H-BIM. The exploration of point cloud applications could enhance models' ability to identify different materials and detect structural anomalies. Additionally, incorporating machine learning algorithms that can adapt to dynamic building environments and learn from incremental data inputs may greatly improve the robustness and accuracy of assessments. Another direction is the exploration of semi-supervised or unsupervised learning techniques, which can reduce reliance on large volumes of labeled data, this is particularly beneficial in scenarios where training data are scarce or difficult to gather.

There also remains a vast field of research linking digital architecture with energy efficiency. Future research simulations could aim to precisely simulate and quantify the actual reductions in energy consumption, CO_2 emissions, and thermal discomfort in buildings, considering the entire chain of building materials used. Another promising avenue could involve deepening the integration of generative design into sustainable architecture, creating automated systems that combine energy prediction and optimization with parametric design tools like Dynamo and Grasshopper. For instance, as discussed by Erisen (2023) and Hou *et al.* (2022), the integration of AI into building energy management and sustainability can optimize controls over energy consumption and thermal comfort through intelligent systems that actively analyze and adjust to the built environment in real-time. These also enables to build a portfolio selection approach (Nascimento *et al.*, 2023) as new alternatives may be considered with a reduced development cost.

Future research should prioritize expanding the datasets used in BIM applications integrated with AI to overcome the current limitations related to the quantity and quality of available data. The acquisition of data on a larger scale and its adequate processing and labeling are fundamental steps to strengthen the training base for deep learning models. For example, Ma *et al.* (2020) faced difficulties constructing datasets for BIM-AI applications and encouraged sharing these databases to improve this field of research. Future research can use datasets available in academic repositories, such as Zenodo, Figshare, IEEE DataPort, and Open Science Framework, to analyze data from BIM models through AI. These repositories offer open access to datasets, such as point clouds, parametric models, and sensor information, which are essential for the development and validation of machine learning and deep learning algorithms applied to the modeling, analysis, and automation of processes in architecture, engineering, construction, and operation projects. By exploring these datasets, researchers can advance studies on semantic segmentation, automated fault detection, regulatory compliance verification, and the creation of Digital Twins. Using these scientific datasets also ensures the reproducibility of experiments and encourages the evolution of AI-based solutions for the intelligent management of the built environment. Expanding the datasets will allow the exploration of more complex tasks, such as point cloud segmentation, which are currently limited by the scarcity of labeled data and the complexity of the algorithms.

Furthermore, the generation of BIM models from point clouds of historical buildings represents one of the most challenges in AECO literature. The difficulties arise from the complexity of architectural forms and the limitations of current techniques for segmenting and classifying building elements. The potential of convolutional neural networks for automating the semantic segmentation of complex architectural elements stands out, as highlighted in this paper. To improve classification accuracy and simplify the modeling workflow, specialized architectures, such as PointNet, PointCNN, and voxel-based networks, can improve classification accuracy and the modeling workflow.

Finally, it is essential to consider that the construction environment is becoming increasingly permeated by IoT devices, cameras, and sensors that collect data in real-time. Despite this data's increasing availability, the heterogeneity of the hardware and software systems challenges the integration and flow of information. Future work can focus on developing adaptive networks and data management systems capable of handling these non-linear interactions, ensuring interoperability between devices and platforms.

5 SOLARISBIM.AI: SMART SUSTAINABLE BUILDING PLANNING WITH SOLAR RADIATION FORECASTING FOR PHOTOVOLTAIC ENERGY PRODUCTION WITH A BIM-DRIVEN DEEP LEARNING MODEL

5.1 Chapter Introduction

The photovoltaic sector in AECO building projects has seen unprecedented growth. In 2024, global renewable energy capacity reached 507 GW, with solar PV comprising approximately 75% of new installations. Forecasts suggest that by 2050, global PV generation could exceed 4,600 TW, largely driven by China and India (Wang; Wang; Zhao, 2025). This growth reflects a convergence of supportive policies, technological innovation, and increasing awareness of climate imperatives. Solar energy, in particular, stands out for its versatility and accessibility in both urban and rural contexts (López; Olivieri, 2025). Several approaches have been explored to address challenges in predicting PV energy production and optimizing building design for energy efficiency. Traditional statistical methods and physical-spatial modeling techniques are commonly employed to predict power generation and assess building performance (Barbosa *et al.*, 2024; Dong; Zhong, 2025). Furthermore, simulation-based tools allow designers to test different architectural strategies and integrate renewable energy systems to maximize energy production (Zhang *et al.*, 2025). However, as the complexity and volume of climate and energy data increase, more advanced computational methods are needed to improve forecast accuracy and support decision-making processes (Palha *et al.*, 2024). Thus, previous research has applied deep learning algorithms to process complex time series data sets, specifically in energy production and consumption forecasts (Zhang *et al.*, 2025).

In this context, deep learning algorithms can extract features from non-linear data by identifying complex patterns in large data sets. They are suitable for energy simulations, such as estimating solar radiation and predicting photovoltaic energy production. When associated with BIM, specifically in the design and planning phase of buildings, quantifying PV energy production through predictive solar energy algorithms becomes an initiative-taking strategy for sustainable projects (Olu-Ajayi *et al.*, 2022; Shao *et al.*, 2021; Wang *et al.*, 2023).

This research addresses that need by proposing an innovative approach: the integration of deep learning algorithms into a BIM-based workflow to support automated PV system design and solar energy prediction. Unlike traditional design processes, this method, called SolarisBIM.AI, leverages time series data of solar radiation and automates information extraction from BIM models to enable data-driven decisions during the planning and retrofit of solar systems in buildings.

Previous research has advanced the body of knowledge, seeking automated alternatives to improve the energy efficiency of buildings. Alawi *et al.* (2024) developed predictive simulations for residential buildings' annual heating and cooling loads. Olu-Ajayi *et al.* (2022) seek to predict energy consumption in the building design phase. Chou *et al.* (2017) integrate data from BIM with power consumption datasets to enhance the visualization of analysis results. Additionally, they implement a spatiotemporal analysis mechanism to assist residents in identifying energy-saving opportunities. Li *et al.* (2024) propose an adaptive sea lion-optimized genetic adversarial to predict renewable energy sources. Tao *et al.* (2024) apply tree-based, linear, and non-linear regression techniques to predict the energy and exergy efficiency of Parabolic Trough Solar Collectors using oil-based nanofluids.

Building upon these advancements, this study proposes a novel BIM-driven deep learning based on radiation forecasting to PV energy prediction. While previous research has focused on energy consumption estimation (Olu-Ajayi *et al.*, 2022), heating and cooling load simulations (Alawi *et al.*, 2024), and energy visualization through BIM integration (Chou *et al.*, 2017), these approaches primarily emphasize building energy efficiency rather than strategic PV energy planning. Moreover, although AI techniques have been leveraged for renewable energy forecasting (Li *et al.*, 2024; Tao *et al.*, 2024), there remains a gap in integrating BIM with predictive models optimized to PV applications. Addressing this gap, the proposed SolarisBIM.AI method utilizes deep learning algorithms to enhance solar radiation forecasting within BIM-based modeling for data-driven decision-making in PV system design.

Thus, this chapter aims to explore the application of a BIM-driven deep learning algorithm to estimate PV energy production, associating solar radiation time series and automated extraction of information in BIM models. This chapter quantifies the energy produced and CO_2 emissions avoided based on the predicted values of the implemented algorithm, using a routine in Dynamo that extracts the information from a BIM model. To support this research objective, this chapter also answers the following research question:

(#RQ3): What is the potential of integrating deep learning models and BIM automation tools to support early-stage PV system planning and energy performance assessment?

This chapter makes contributions to the field of AECO industry research. First, it highlights that the design phase in BIM projects is essential for establishing efficient PV plants. Associated with predicting solar energy production, PV design can be justified with data-based documentation of the potential economic benefits generated by clean energy production and the environmental benefits from avoided CO_2 emissions. In addition, engineers and architects can plan the physical layout that best accommodates solar modules according to the available physical space. In summary, this paper uses solar radiation time series and automatic BIM data extraction to establish an automated design process called SolarisBIM.AI for quantifying solar energy production and avoiding CO_2, highlighting the role of BIM modelers in the design phase.

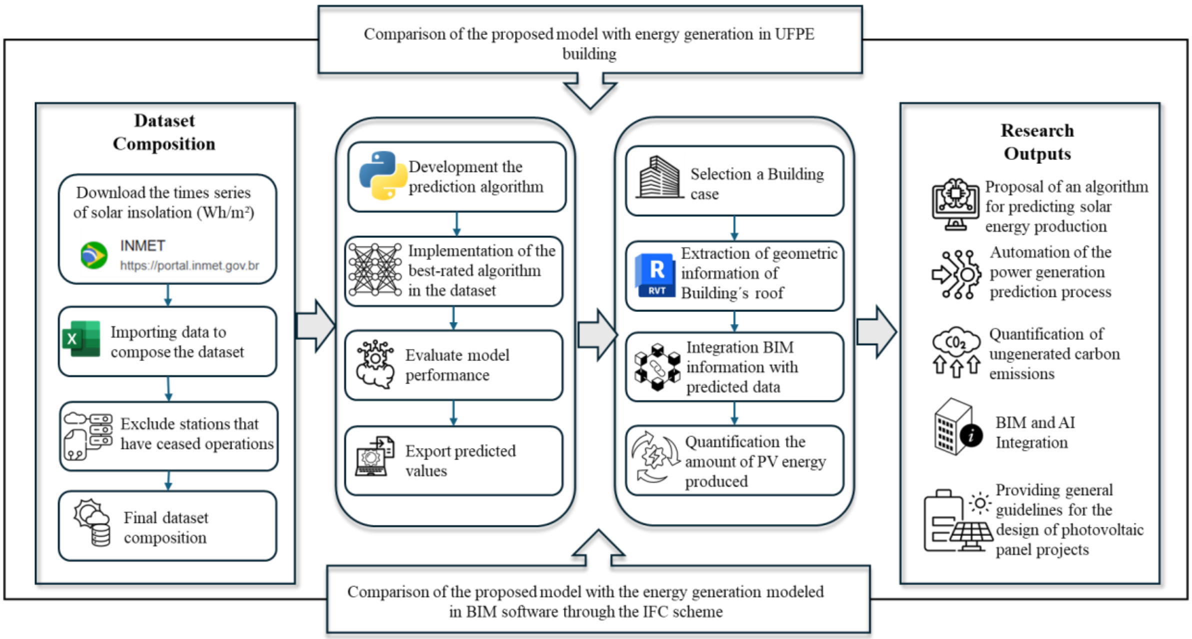
5.2 SolarisBIM.AI Methodology

In this research, Long Short-Term Memory (LSTM), Extreme Gradient Boosting (XGBoost), and Feedforward Neural Network (FNN) are applied to model a time series of solar radiation. LSTM was selected for its recurrent network architecture, which uses memory cell mechanisms and control gates. LSTM allows for capturing long-term dependencies and seasonal patterns common in time series. XGBoost, based on gradient boosting, is efficient in boosting structured data and can be adapted to time series by transforming temporal observations into a supervised learning problem. FNN, on the other hand, considers more complex temporal dependencies ideally suited to non-linear patterns in the time series (Tian *et al.*, 2023). These algorithms are widely used in forecasting solar radiation and PV energy from a multiscale perspective, including mesoscale, microscale, and building scale forecasting in urban environments, as pointed out in the systematic literature review by Tian *et al.* (2023).

The performance of each algorithm was evaluated, and the one with the best performance on the dataset was selected for implementation in this article's experiment. The predictions made in the model feed the calculation of PV energy production, which converts radiation data into energy generation estimates. Thus, this paper develops a database of daily solar radiation with data from municipalities in Pernambuco, Brazil. The research flowchart is shown in Figure 11, and each research step is described in the subsequent sections.

This research used Python programming to develop the algorithm, leveraging the scikit-learn library for predictive data analysis tools. The 3D model was created using Revit software, and building geometry data was extracted with Dynamo. The IFC data standard was employed to import the 3D model into Solarius PV software. Additionally, QGIS software was used to generate maps of the mesoregions for the study area.

Figure 11 - Research workflow



5.2.1 Dataset Composition

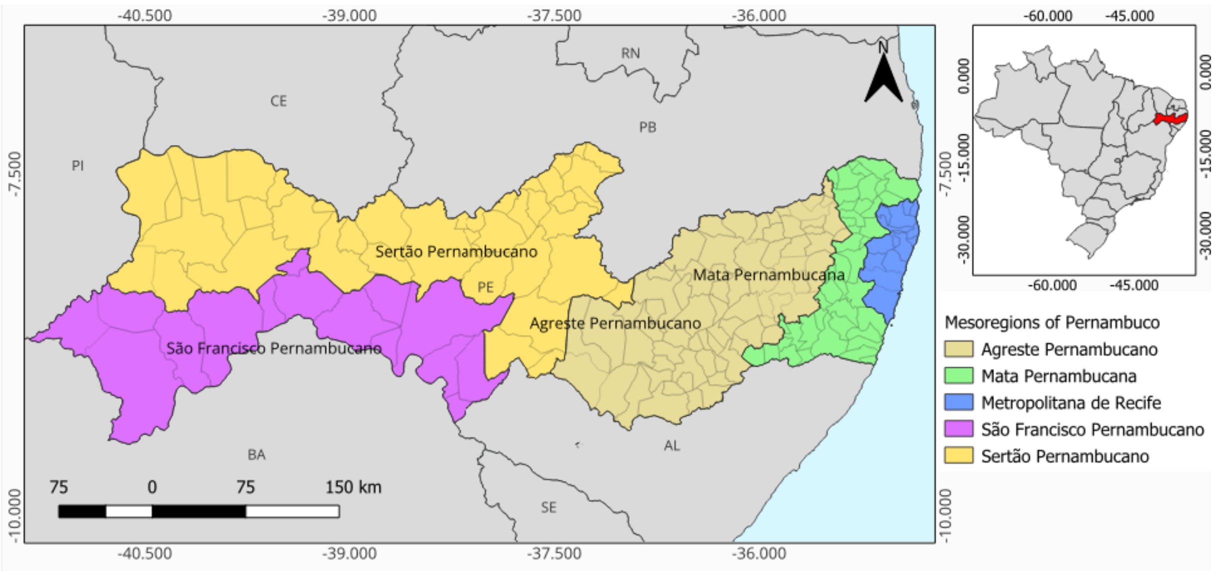
The solar radiation data was obtained from the National Institute of Meteorology (INMET), an agency of the Brazilian Ministry of Agriculture and Livestock. INMET is the principal meteorological agency in Brazil, responsible for collecting, monitoring, and providing meteorological data on a national scale (INMET, 2023). Solar radiation is the quantity used to express the solar energy that falls on a given flat surface area over a given time interval. Solar radiation is given in kWh/m² (kilowatt-hours per square meter).

Pernambuco, located in the northeast region of Brazil, was selected for the study of PV energy generation due to the state's favorable climate conditions. Pernambuco has high levels of sunlight throughout the year, which is essential for the performance of solar systems. In addition, public institutions in Pernambuco, such as the Universidade Federal de Pernambuco (UFPE), have shown a growing interest in renewable energy and seek to diversify their energy matrix, reducing dependence on polluting sources. In addition, this research in the state can contribute to technological advancement and efforts to mitigate CO_2 emissions.

Pernambuco is in the center-east of the Northeast region of Brazil, with a total area of 98,067,877 km². It borders Ceará and Paraíba to the north, Piauí to the west, Bahia and Alagoas to the south, and is bathed by the Atlantic Ocean to the east. According to the Brazilian Institute of Geography and Statistics (IBGE), Pernambuco is divided into five distinct mesoregions that reflect geographic, economic, and cultural variations, as shown in Figure 12. The Sertão

Pernambucano, to the west, is characterized by a semiarid climate and an economy focused primarily on livestock farming and agricultural activities adapted to drought conditions. The Agreste Pernambucano, located between the Sertão and the Zona da Mata, has a climate that favors agriculture and the development of regional industrial hubs. The Metropolitan Region of Recife, to the east, has the highest level of urbanization, being the economic and administrative center of the state, with emphasis on the services, commerce, and industry sectors. The Mata Pernambucana, adjacent to the Metropolitan Region, is traditionally a sugarcane growing area that has historically been integrated into the sugar economy. Finally, the mesoregion of Saint Francisco Pernambucano, located in the extreme west and bathed by the Saint Francisco River, is fundamental for irrigation and agricultural development of irrigated fruit growing, with emphasis on the production of fruits for export (IBGE, 2024).

Figure 12 - Mesoregions of Pernambuco, Brazil



Regarding the climate, the Sertão Pernambucano is predominantly semiarid, characterized by high temperatures and low precipitation, with strong sunlight throughout the year, contributing to the region's aridity. In the Agreste Pernambucano, the climate is marked by higher relative humidity and precipitation during the winter, with moderate sunlight that favors crops adapted to the more fertile soil. With a humid tropical climate, the Metropolitan Region of Recife experiences high annual precipitation rates and less thermal variability, with intense sunlight in the drier periods, highlighting the region's summer characteristics. In the Mata Pernambucana, the influence of the humid tropical climate maintains dense and green vegetation, with high humidity and rainfall distribution throughout the year, although with high sunlight during intermittent dry periods. In the São Francisco Pernambucano, the semiarid climate is mitigated by the presence of the Saint Francisco River, and the high sunlight incidence becomes advantageous for irrigated fruit-growing practices (IBGE, 2024).

In Pernambuco, INMET has 16 stations collecting sunlight data, as shown in Table 13. However, three stations have been on record for over 10 years, so they were excluded from the analysis. Seven stations were deactivated and have no records. Therefore, only six stations (Arcoverde, Cabrobó, Garanhuns, Petrolina, Recife, and Surubim, names in bold in Table 13) have records for available use. To encompass a more extensive historical series between the six municipalities, the dataset has records from 01/01/1975 to 12/31/2023, corresponding to 17,896 days (49 years).

The preliminary analysis also revealed some days without sunlight records. To get around this, the cells in the dataset with “null” values were filled with the average sunlight of the stations on that day.

Figure 13 - Solar Insolation data collection stations in Pernambuco, Brazil

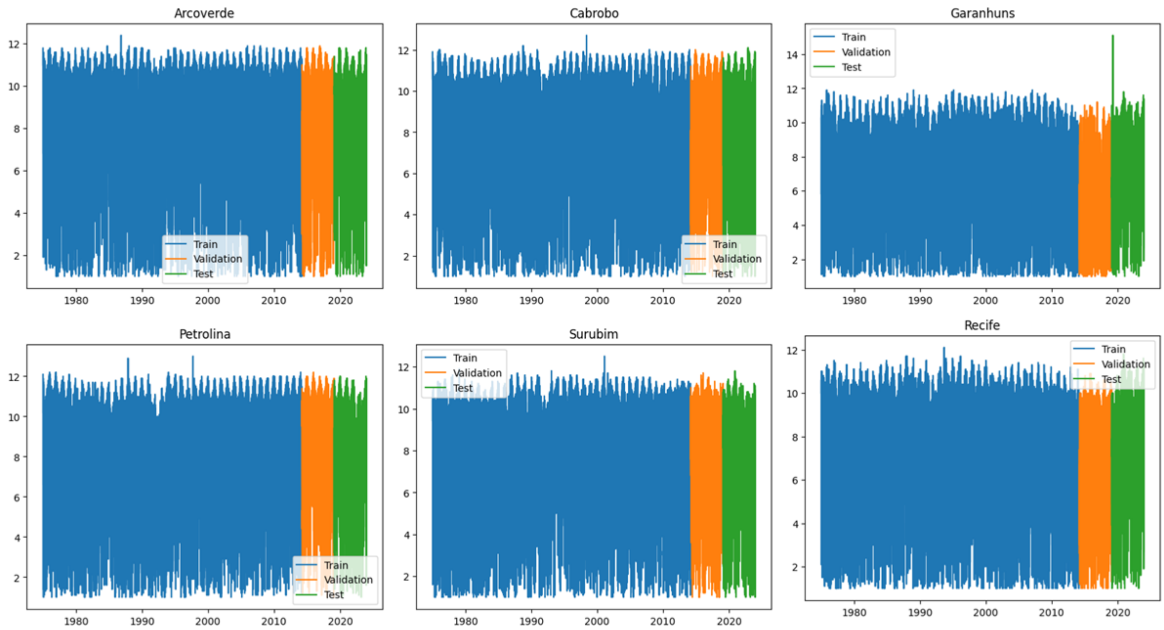
City	Code	Latitude	Longitude	Altitude	First measurement	Zone
Arcoverde	82890	-8.4336111	-37.05527777	683.91	1973-01-31	Sertão Pernambucano
Ouricuri*	82753	-7.87944443	-40.09194444	462.01	1975-09-15	
Triunfo**	82789	-7.82972221	-38.12222221	1105	1953-05-31	
Garanhuns	82893	-8.91083333	-36.49333333	827.78	1913-01-31	Agreste Pernambucano
Surubim	82797	-7.839628	-35.801056	421.44	1929-09-30	
Pesqueira*	82892	-8.370701	-36.707812	643.38	1911-08-31	
Caruaru**	82895	-8.28	-35.97	537	1928-12-31	
Recife	82900	-8.05916666	-34.95916666	11.3	1961-07-06	Metropolitana de recife
São Lourenço da Mata (Tapacurá)**	82897	-8.17	-35.18	102	1918-12-31	
Olinda**	82898	-8.02	-34.85	55	1921-12-31	
Nazaré da Mata**	82781	-7.73	-35.25	87	1908-12-31	Mata Pernambucana
Goiana**	82799	-7.55	-34.98	11	1910-12-31	
Cabrobó	82886	-8.50388888	-39.31527777	342.78	1927-10-16	São Francisco Pernambucano
Petrolina	82983	-9.3886111	-40.52333332	372.54	1940-12-31	
Floresta*	82887	-8.6	-38.56999999	309	1952-08-31	
Petrolândia**	82987	-9.07	-38.32	286	1937-12-31	

* Station without measurement data in more than 10 years
** Disabled

5.2.2 Deep Learning Algorithms

In all algorithms described in the following subsections, the data is divided according to these indices: the first 80% of the data is assigned to the training set, the data between 80% and 90% is allocated to the validation set, and the last 10% of the data goes to the test set, as shown in Figure 14.

Figure 14 - Division of data into training, testing, and validation



5.2.2.1 LSTM network

The LSTM network was developed to overcome the limitation of original RNNs in handling long-term dependencies, i.e., situations where current predictions need to access previously stored information. Unlike the traditional RNN model, the LSTM includes an enhanced memory block that holds the weights and thresholds of all previous learning samples. This memory block regulates the flow of input and output data through input, forget, and output gates. At a given time t, the LSTM takes three inputs (the current input value x_t, the previous cell hidden state h_{t-1}, and the previous cell state C_{t-1}) and generates two outputs (the current hidden state h_t and the current cell state C_{t-1}) (Kumari; Toshniwal, 2021; Long et al., 2023).

- Forget Gate: Acting as one of the components that regulates the state of cell Ct-1 in the LSTM network, the forget gate ft decides whether the state of cell C_t from the previous instant will be maintained until the current instant, defined as in Equation (2) (LONG *et al.*, 2023).

$$f_t = \text{sigmoid} \times (W_f\,[h_{t-1}, x_t] + b_f)$$

where (2)

$$[w_f]\begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix} = [w_{fh}\; w_{fx}]\begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix} = w_{fh}\,h_{t-1} + w_{fx}\,x_t$$

The f_t is the forget gate activation that determines how much of the past memory (C_{t-1}) will be kept in the cell. The sigmoid function generates values between 0 and 1 by controlling the proportion of information that will be forgotten. W_f is the forget gate weight matrix associated with the forget gate, which modulates the influence of h_{t-1} and x_t. h_{t-1} is the previous hidden state of the previous stage, which carries information from the past. x_t is current input vector in the stage t. The b_f forget gate adjusts the output of the sigmoid activation. W_{fh} is the hidden state weight matrix connect the previous hidden state h_{t-1} at the door of oblivion. W_{fx} is the input weight matrix connect the input x_t at the door of oblivion.

- Input gate: Another essential control component in this algorithm is the input gate, which is responsible for deciding whether the current network input x_t will be stored in the cell state C_t. This can be expressed by Equation (3), where w_i and w_c are the weight matrices corresponding to the input data and the cell state value at the input gate, while b_c and b_f represent the thresholds. With this information, the current cell state c_t is obtained (Long et al., 2023).

$$\begin{gathered} i_t = \text{sigmoid} \times (w_i\,[h_{t-1}, x_t] + b_i) \\ C'_t = \tanh\,(w_c\,[h_{t-1}, x_t] + b_c) \\ C_t = f_t * C_{t-1} + it * C'_t \end{gathered} \tag{3}$$

The candidate cell state (C'_t) represents a potential update to the LSTM memory, regulated by the tanh activation function, which scales values between -1 and 1. This state is influenced by the candidate cell state weight matrix (W_c) and its corresponding bias (b_c). The cell state (C_t) is then updated by integrating the previous cell state (C_{t-1}), modulated by the forget gate activation (f_t), with the candidate cell state (C'_t), weighted by the input gate activation (i_t). This selective memory update mechanism ensures that information is retained while unnecessary data is discarded, allowing the LSTM to effectively capture long-term dependencies in sequential data.

- Output Gate: As the last control component of this algorithm, the LSTM uses the output gate to regulate the cell state C_t, which is passed to the current output value h_t. This process can be represented by Equation (4), where O_t, W_o, and b_o corresponds to the output gate value, weight matrices, and thresholds at the output gate, respectively (Long *et al.*, 2023).

(4)

$$O_t = \text{sigmoid} \times (W_o [h_{t-1}, x_t] + b_o)$$
$$ht = O_t \tanh (C_t)$$

The output gate activation (O_t) regulates how much of the updated cell state (C_t) will contribute to the final hidden state (h_t). It is computed using a sigmoid activation function, which ensures values remain between 0 and 1. The output gate weight matrix (W_o) and its bias (b_o) determine the influence of the previous hidden state (h_{t-1}) and the current input (x_t). The final hidden state (h_t) is obtained by applying the tanh activation function to the cell state (C_t) and modulating it by the output gate activation (O_t), allowing the LSTM to expose information to the next step selectively. The three gates act together to generate the cell state C_t and the output value h_t, allowing the LSTM to solve the vanishing gradient problem caused by long-term dependencies in the original RNN model (Long *et al.*, 2023).

In the logical structuring of the LSTM algorithm, the first step was to transform the time series data into a supervised learning problem with input and output patterns, as shown in Algorithm 1. A main function created a DataFrame containing the data windows with a specified length parameter n. The function receives a date range and extracts subsets from the original DataFrame, so that each subset contains *n* consecutive observations and a target date. For each valid window, the x values are separated into variables X (for historical data) and Y (for the value to be predicted). Finally, a DataFrame is created with the Target Date column containing the target dates, the Target-X columns representing the window data, and the Target column for predicting the value. Table 7 presents this application logic.

The logic is to evaluate the network's performance for predicting multiple points in time, that is, given that the dataset is composed of a daily historical series, the LSTM algorithm performs the accuracy considering three points of the historical series. For example, given one day as input, the model predicts the next three days.

Table 7 - Structuring the dataset in DateTime

	Target Date	**Target-3**	**Target-2**	**Target-1**	**Target**
0	1975-01-04	7.3	10.4	10.6	10.0
1	1975-01-05	10.4	10.6	10.0	10.6
2	1975-01-06	10.6	10.0	10.0	8.7
3	1975-01-07	10.0	10.6	8.7	2.1
4	1975-01-08	10.6	8.7	2.1	20.4
…	…	…	…	…	…
17889	2023-12-27	7.6	9.7	10.2	11.1
17890	2023-12-28	9.7	10.2	11.1	11.3
17891	2023-12-29	10.2	11.1	11.3	10.8
17892	2023-12-30	11.1	11.3	10.8	10.8
17893	2023-12-31	11.3	10.8	10.8	8.6

17894 rows × 5 columns. Results of Recife

Algorithm 1: Transforming Time Series into Data Windows for Predictive Modeling
Step 1. Define function to transform string column into date format Receive string in format YYYY-MM-DD Convert string to datetime **Step 2.** Define Function to receive data Windows Receive a DataFrame, the first date, the last date, and the window size n Convert first date and last date to datetime **Step 3.** Create data Windows Extract last n+1 records until target_date ***If*** window size is n+1, separate data X (last n values) and Y (last n values): *X as the first n values*: $X = [x_1, x_2, \ldots, x_n] = \text{df\_subset}[:-1]$ *Y as the value of the last element*: $\text{Y} = [x_{n+1}] = \text{df\_subset}[-1]$ *Add* target_date, X, and Y to the respective vectors *Advance* target_date by 1 day **Step 4**. Construct Return DataFrame *Add* dates in the Target Date column ***For*** each i in the range 0 to n-1: Add X[:, i] to column Target-(n-i) *Add* Y to column Target return [Target] = Y **Step 5**. Return with structured windows for modeling Store result in Dataframe End code.

The DataFrame is then converted to a NumPy array. The dates are extracted from the first column and stored. The intermediate columns, corresponding to the input data (X), are extracted and reshaped to include an extra dimension, preparing the data for models that require three-dimensional inputs. Finally, the last column containing the output values (y) is extracted and converted to the *np.float32* type. The function returns the dates, inputs (X), and outputs (y).

Algorithm 2 then defines the LSTM using Keras, to predict a continuous value based on time sequences. The first part of the code imports essential libraries for building the model, such as Sequential (to define the sequential model), Adam (the optimizer used to adjust the weights during training), layers (which contains the network layers, such as LSTM and Dense), and ReduceLROnPlateau (a callback that adjusts the learning rate if the model's performance stagnates).

In the model definition, it is specified that the model will be of the sequential type, where the layers are stacked linearly. The first layer is an input layer that receives sequences of 3 timesteps, with one value per timestep. Then, the LSTM layer is added. The number of neurons in the LSTM is set at 256. After the LSTM, two dense layers with 64 neurons each are added, using the ReLU activation function to introduce non-linearities into the model and allow it to learn more complex representations of the data. The last layer of the model is a dense layer with a single neuron, which will generate the final prediction of the model.

After defining the architecture, the model is compiled using the MSE (mean squared error) loss function. The optimizer used is Adam, which is known for dynamically adapting the learning rate during training.

Algorithm 2: Defining and Training the LSTM Model
Step 1. Model Definition **Input layer:** Receives sequences of size 3, with one value per timestep: The input has the form (n,3,1), where n is the number of samples per timestep **Model Layers:** Uses an LSTM layer with 256 units (neurons) Two dense layers of 64 units with ReLU activation f(x)=max(0,x) **Output Layer:** Dense layer with 1 unit to produce the final prediction: $y_{pred} = \hat{y}$ **Step 2.** Loss Function and Optimizer **Use means squared error (MSE):** $MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y_i})^2$ y_i is the actual value of the sample i; $\hat{y_i}$ is the model's prediction for the sample i; n is the total number of sample. **Using the Adam algorithm for optimization:** $m_t = \beta_1 m_{t-1} + (1 - \beta_1)\nabla_\theta J(\theta)$ $v_t = \beta_2 v_{t-1} + (1 - \beta_2)\nabla_\theta^2 J(\theta)$

$$\widehat{m_t} = \frac{m_t}{1-\beta_1^t}, \widehat{v_t} = \frac{v_t}{1-\beta_2^t}$$

$$\theta_t = \theta_{t-1} - \alpha \frac{\widehat{m_t}}{\sqrt{\widehat{v_t} + \varepsilon}}$$

m_t is the average of the gradients;

v_t is the variance of the gradients;

β_1 *and* β_2 are the decay coefficients;

α is the learning rate;

ε is a small value to avoid division by zero.

Step 3. Training with Learning Rate Reduction Callback

Callback ReduceLROnPlateau: monitors validation loss during training:

$$\eta_t = \eta_{t-1} * factor$$

η_t is the learning rate at the time t;

η_{t-1} is the learning rate of the previous epoch;

factor is the reduction factor.

Step 4. Model Training

Using the Keras fit method

Feed the neural network and parameters according to the loss function and optimizer:

$$\widehat{y_t} = f(X_t; \theta)$$

X_t are the input data at the time t;

$\widehat{y_t}$ is the model's prediction for X_t;

θ are model weights and biases.

Step 5. Completion of the Process

Predict values for new input data

Perform time series predictions based on the training obtained

End code.

5.2.2.2 Extreme Gradient Boosting (XGBoost)

The XGBoost algorithm is used for target detection and prediction due to its incremental tuning capability. At each iteration, it minimizes the residual of the previous model by constructing a new tree that fits in the direction of the negative gradient, allowing the training effect of a tree to directly influence the input samples for building the next model. To improve computational efficiency, XGBoost applies second-order Taylor expansion to the objective function, effectively approximating the generalization error and simplifying calculations. Furthermore, XGBoost mitigates prediction volatility and reduces the risk of overfitting by including a regularization term in the objective function, which improves the robustness and generalization of the model (Li *et al.*, 2023).

The XGBoost gradient algorithm can model complex, nonlinear interactions in data. It is suitable for identifying patterns in solar irradiance. This model uses a set of decision trees that operate cooperatively to refine predictions, with each tree evaluating different aspects of the data. In addition, XGBoost generates variable importance ratings to identify the main determinants of the time series (Saigustia; Pijarski, 2023). The XGBoost algorithm, after establishing each node, reduces losses as shown in Equation (5).

$$L^j = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(j-1)} + f_j(x_i)\right) + \Omega(f_j)$$
$$L_{split} = \frac{1}{2}\left[\frac{(\sum_{i \in lL} g_i)^2}{\sum_{i \in lL} h_i + \lambda} + \frac{(\sum_{i \in lR} g_i)^2}{\sum_{i \in lR} h_i + \lambda} + \frac{(\sum_{i \in l} g_i)^2}{\sum_{i \in l} h_i + \lambda}\right] - \gamma \quad (5)$$

The objective function (L^j) in XGBoost is designed to optimize the model at each boosting iteration by minimizing the loss function while adding a regularization term. The summation term represents the loss function $l(y_i, \hat{y}_i^{(j-1)}+f_j(x_i))$, which measures the difference between the actual and predicted values. The function $f_j(x_i)$ represents the newly added tree at iteration j, and $\Omega(f_j)$ is a regularization term that controls the complexity of the model to prevent overfitting. The split gain function (L_{split}) quantifies the loss reduction achieved by splitting a node in a decision tree. It is computed using the first-order gradient statistics (g_i) and second-order statistics (h_i), where L and R denote the left and right child nodes after the split. The denominator terms include λ, a regularization parameter that smooths the gain calculation. The term γ represents the penalty for introducing a new leaf node, which prevents unnecessary splits. A higher split gain indicates a more effective partition of the data, contributing to improved model accuracy.

In Equation (6), l is the subset of observations available at the current node, lL and lR are the subsets of observations available at the left and right nodes after splitting. The XGBoost algorithm defines the functions as in Equation (6) (Li *et al.*, 2023).

$$g_i = \partial_{\hat{y_i}}^{(j-1)} l(y_i, \hat{y_i}^{(j-1)})$$
$$h_i = \partial_{\hat{y_i}^{(j-1)}}^{2} l(y_i, \hat{y_i}^{(j-1)}) \quad (6)$$

In the XGBoost algorithm, g_i and h_i represent the first-order and second-order gradients, respectively, of the loss function $l(y_i, \hat{y}_i^{(j-1)})$ with respect to the predicted value from the previous boosting iteration. The first-order gradient (g_i) measures the direction and magnitude of the loss function's change, indicating how adjustments to the prediction affect the model's error. The second-order gradient (h_i) captures the curvature of the loss function, allowing for an adaptive learning rate based on confidence in the gradient estimation. These gradient statistics are used in the tree-splitting process and the computation of the split gain to determine the optimal node partitioning, ultimately improving model performance.

Like LSTM, the algorithm partitions the dataset into three subsets: training, testing, and validation. Initially, 80% of the data is allocated to training, 10% to testing, and the remaining 10% to validation, which allows evaluation of the model's performance on samples not seen during tuning.

Then, the XGBoost model is configured with the mean squared error (MSE) function as the objective, which is suitable for regression tasks. The parameter n_estimators=100 defines the number of trees, while learning_rate=0.1. Finally, prediction is performed on the validation set, and the mean squared error is calculated as an evaluation metric. This calculation of the error on the validation set provides an objective measure of model performance, which can be used for tuning and comparisons with other models.

5.2.2.3 Feedforward Neural Network (FNN)

An Artificial Neural Network is a nonlinear model that aims to simulate how the human nervous system performs specific tasks. With an appropriate and limited set of processing units, neural networks could recognize patterns and adapt to empirical data. Several studies have demonstrated their effectiveness in classification tasks, pattern recognition and, especially, in making accurate predictions. Among the types of neural networks, feedforward networks stand out as one of the most prominent structures. Recent advances in literature have established the theoretical basis for the ability of feedforward networks to approximate functions in a universal way. It has been shown that, with enough hidden units and properly adjusted parameters, feedforward networks can approximate any arbitrary function. In this architecture, neurons are organized in layers, allowing the transmission of information in a hierarchical and sequential manner. The general formulation of the FNN is shown in Equation 7 (Saâdaoui, 2017).

$$y_t = N\left[\beta_0 + \sum_{i=1}^{q} \beta_1 M(\alpha_i^T x_t)\right] + \in_t = f(x_t, \Theta) + \in_t \tag{7}$$

The term β_0 represents the bias, acting as an offset in the prediction. The summation $[\sum_{i=1}^{q} \beta_1 M(\alpha_i^T x_t)]$ models the hidden layer processing, where $\alpha_i^T x_t$ denotes the weighted sum of inputs, and $M(\alpha_i^T x_t)$ is the activation function that introduces nonlinearity to the model. The coefficient β_1 determines the contribution of each transformed input to the final output. Additionally, the function $f(x_t,\Theta)$ represents a general nonlinear mapping parameterized by Θ, which may correspond to the network weights and biases learned during training. The term ϵ_t captures residual errors or noise in the predictions. This formulation allows the FNN to approximate complex functions and relationships within the input data, making it suitable for applications in forecasting, classification, and regression tasks.

Considering the network structure described in Equation (8), a central numerical issue is estimating the set of unknown parameters Θ for a given data sample. For this purpose, backpropagation, together with its variants, is widely used as one of the most effective learning algorithms. Backpropagation is an iterative estimation method that allows the calculation of Θ recursively. In this process, a starting point is randomly selected and then updated according to the following procedure of Equation (8) (Saâdaoui, 2017).

$$\hat{\Theta}^{(i+1)} = \hat{\Theta}^{(i)} + \gamma \nabla f(x_t, \hat{\Theta}^{(i)})[y_t - f(x_t, \hat{\Theta}^{(i)})] \tag{8}$$

The term $\hat{\Theta}^{(i+1)}$ represents the updated parameter set at iteration i+1, while $\hat{\Theta}^{(i)}$ denotes the parameters from the previous iteration. The update is influenced by the learning rate γ, which controls the step size in the optimization process. The gradient $\nabla f(x_t, \hat{\Theta}^{(i)})$ represents the partial derivative of the model function $f(x_t, \hat{\Theta}^{(i)})$ with respect to the parameters, indicating the direction in which the model should adjust its weights. The term $[y_t - f(x_t, \hat{\Theta}^{(i)})]$ measures the prediction error at time t, guiding the update to minimize the difference between actual and predicted values.

The FNN uses the TensorFlow Keras library. Initially, the model is configured to predict a specific time variable with a 1-day observation window (look_back=1), meaning that the current value is predicted based on the immediately previous value. The create_dataset function organizes the training, validation, and test data, generating the X and y variables suitable for time series modeling.

The model itself is a multilayer perceptron (MLP) network with two hidden layers of 10 neurons each, activated by the ReLU function, which facilitates the modeling of non-linear patterns. The output layer is a dense layer with a single neuron, suitable for predicting continuous values. The Adam optimizer is used for efficient adjustment of the weights, while the MSE loss function minimizes the mean squared error, a common metric for regression.

To monitor the accuracy during training, a custom Callback (MAE_Percentage_Callback) is added. This calculates and displays the MAE and MAPE for both the training and validation sets at the end of each epoch. Finally, the MSE, the MAE, and the MAPE are calculated to quantify the accuracy of the predictions and verify the generalization ability of the model on unseen data. The architecture of the algorithm is shown in Algorithm 3.

Algorithm 3: Defining and Training the FNN Model
Step 1. For each input sequence $\{x_t\}$ with time window L $X_t = \{x_{t-L+1}, \dots, x_{t-1}, x_t\}$ and $y_{t+1} = x_{t+1}$ X is the feature matrix y is the vector of target values **Step 2. Input** layer with n_1 neurons, where n_1=10 $h^{(1)}$=f($W^{(1)} * X + b^{(1)}$) $W^{(1)}$ and $b^{(1)}$ are the weights and biases of the input layer f is the ReLU activation function $h^{(1)}$ is the activation of the first hidden layer $h^{(2)}$ is the activation of the second hidden layer **Hidden** layer with n_2=10 neurons $h^{(2)}$=f($W^{(2)} * h^{(1)} + b^{(2)}$) **Step 3. Output** layer with a single neuron, which provides the prediction $\hat{y}$ $\hat{y}$=$W^{(3)} * h^{(2)} + b^{(3)}$ End code

5.2.3 Evaluation Metrics

To evaluate the proposed model, this paper applies the MAPE and MAE, as presented in Equation 9. The MAE measures the average absolute difference between actual and predicted values. It indicates how much, on average, predictions deviate from the real values without considering direction. The MAPE expresses this error as a percentage of the actual values.

$$MAE = \frac{\sum_{i=1}^{N} |y(i) - \hat{y}(i)|}{N}$$
$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_{(i)} \widehat{y_{(i)}}}{y_{(i)}} \right| \times 100 \quad (9)$$

In this context, $y_{(i)}$ represents the actual solar energy value, while $\hat{y}_{(i)}$ is the predicted solar energy value. N denotes the number of samples. The defined statistical evaluation tests were applied to measure the forecasting system's effectiveness. Variations in the evaluation metrics provide insights into the stability of the proposed model.

5.2.4 BIM Model Generation

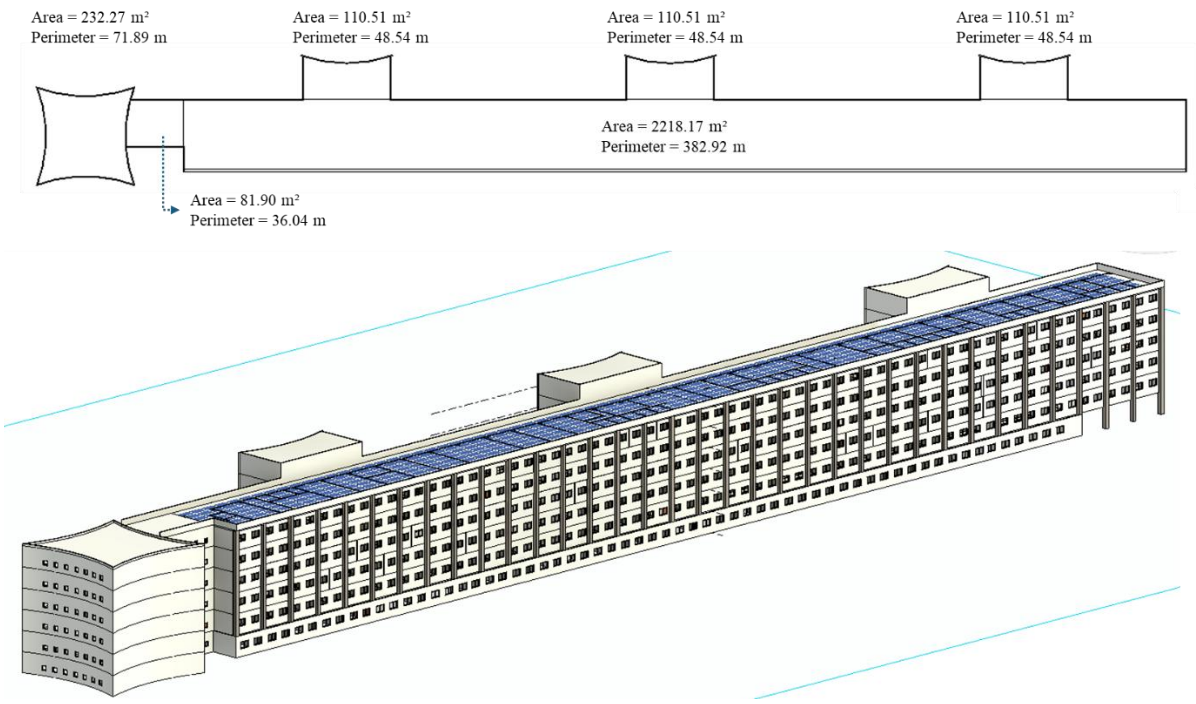
School of Engineering, a building of UFPE, was chosen to model and simulate PV production. It is worth noting that this building has solar modules installed on its roof, which allows the comparison of the simulation model proposed in this study with the real production case.

However, all the plans for this building are still in 2D format in Autocad. Therefore, to be able to develop this research the 2D was modeled in Revit software (2025). The BIM mandate in Brazil for public institutions, established by Decree No. 10,306/2020, aims to promote the adoption of BIM in engineering works and services in the public sector. The implementation of BIM in public contracts in Brazil is divided into phases, with progressive deadlines that require the use of this technology from conception to management of the project life cycle. The main objectives of this requirement are to improve efficiency, transparency, and quality control in public contracts, in addition to reducing costs and waste throughout the construction execution and maintenance stages. The Decree establishes BIM as an essential instrument for the modernization of the construction sector, encouraging data interoperability, collaboration between the agents involved, and the adoption of innovative practices.

Furthermore, it is worth noting that UFPE has developed strategic actions to restructure the energy matrix of its Campi to align with environmental guidelines and sustainability practices in the generation of renewable energy. The institution's focus is to reduce dependence on energy suppliers by using PV energy due to its low maintenance and operating costs.

The chosen CTG UFPE building has 6 floors and several classrooms, professors' officers, laboratories, and auditoriums for various engineering courses. The floor plan of the roof and the 3D model generated in Revit are shown in Figure 15.

Figure 15 -Representation of the roof and 3D model generated in Revit



A Dynamo routine was implemented to extract the usable area of the building's roof, as illustrated in Figure 4. Subsequently, the dimensions and efficiency of the currently installed solar modules were input into the routine to determine the number of modules. This data was then utilized to estimate the predicted photovoltaic energy production using SolarisBIM.AI. Dynamo offers an intuitive, node-based visual programming environment that facilitates automated data import. Dynamo is a plugin of Revit and integrates this software but can be used on other platforms. It is useful for automating and enhancing design and modeling tasks. Its features streamline the process of collecting, analyzing, and visualizing data within the BIM context. The node-based structure allows for the creation of customized workflows by interconnecting different nodes to meet the specific needs of each project (Cho *et al.*, 2024).

5.2.5 Quantification of PV energy production

The DL algorithms provide the solar radiation in kWh/m² for the desired day(s). The manufacturer's manual provides the average conversion efficiency of the modules, typically between 15% and 22% (or 0.15 to 0.22). The area required for the calculation is extracted from the 3D BIM model. The calculation is achieved by using Equation 10. The number of solar modules (N_{SP}) is obtained by the division between total area of the building's roof (A_{roof}) by area of a solar module (A_{SP}). The daily PV production (D_{PV}) is calculated by multiplying the number of solar modules by the average daily solar radiation per square meter ($D_{average}$) and the efficiency of the modules (E). The annual PV energy production (PV_{anual}) is then obtained by multiplying the daily energy output by 366 days, accounting for leap years. This estimation provides a straightforward method to assess photovoltaic energy generation based on module specifications and local solar conditions. It is worth noting that this model disregards the losses of the conversion stage of the PV system (inverter).

$$N_{SP} = \frac{A_{roof}}{A_{SP}} \qquad (9.1)$$

$$D_{PV}\ (kWh) = N_{SP} \times (D_{average}) \times E \qquad (9.2) \qquad (10)$$

$$PV_{anual} = D_{PV} * 366 \qquad (9.3)$$

The PV energy production in kWh was further compared to the CO_2 emission factor to find the amount of CO_2 avoided. The formula is shown in Equation 11. In Brazil, the Ministry of Science, Technology and Innovation (MCTI) reports that the annual Average Emission Factor (tCO2/kWh) is 0.0000467 tCO_2/kWh (MCTI, 2023).

$$CO_{2avoided} = PV_{anual}\ (kWh) \times CO_{2\ factor}\ (tCO_2/kWh) \qquad (11)$$

5.2.6 Proposed model comparison

The model developed in this research is evaluated by comparing the existing energy production in the UFPE school building, which has a PV power generation plant with a production capacity of 273.24 kWp. The total investment was $333,450 dollars, with resources from the Ministry of Education (MEC), including installing two other poles at the Center for

Applied Social Sciences (CCSA) and the administration building. The UFPE dchool building generation pole began operating in June 2021 (UFPE, 2021).

The proposed methodology was compared with the Solarius PV software simulations to validate its effectiveness in predicting solar irradiance. Solarius PV is widely used in industry to simulate the performance of solar systems with high precision, integrating climate, sunlight, and solar module efficiency data (Di Giovanni *et al.*, 2024). In this context, the BIM model generated in Revit was exported in IFC format and imported into Solarius PV.

5.3 Results

5.3.1 Experiment

LSTM and FNN demonstrated good effectiveness in predicting variations in the data, with a good overlap between predicted and observed values in both the validation and test sets. This indicates that the model could capture the underlying temporal trends in the data, generating predictions that were close to the actual observations. However, FNN demonstrated the best performance among the three classifiers tested. The overlap in the FNN network of predicted and observed values reflects the model's ability to generalize, as shown in Figure 16 and Figure 17.

In contrast, the XGBoost model did not perform as well in capturing the temporal variations in the data, as indicated by a lower overlap between predicted and observed values. Additionally, while LSTM demonstrated reasonable effectiveness, its performance was still inferior to that of FNN, which exhibited better performance.

Figure 16 -Predicted and observed values in the test set in FNN algorithm

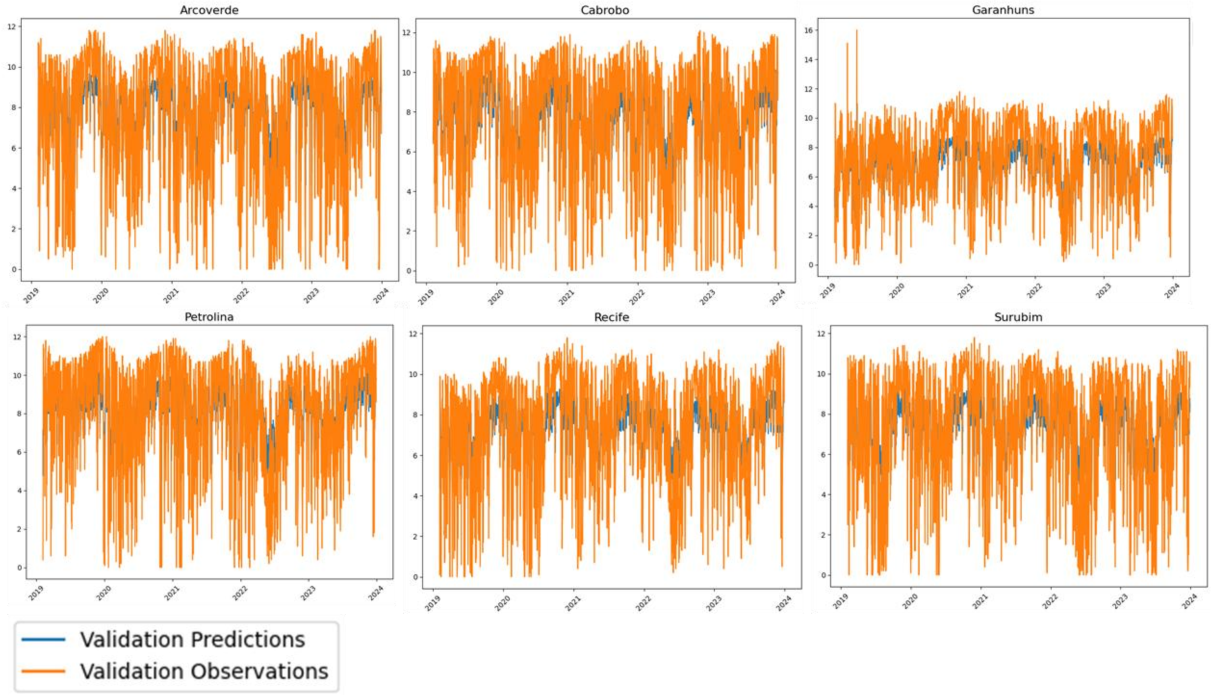
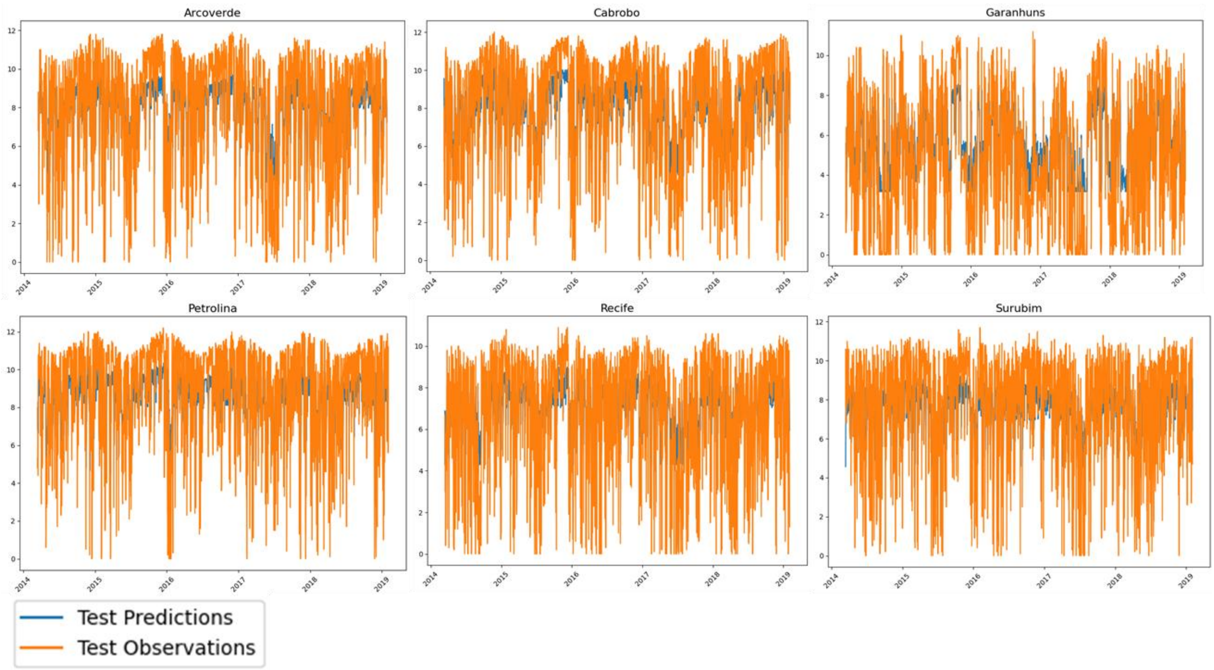


Figure 17 - Predicted and observed values in the validation set in FNN algorithm



In the LSTM algorithm, the cities Arcoverde, Petrolina, and Cabrobó have MAPE percentage errors for the validation set ranging from 31.7328% (Petrolina) to 39.1658% (Arcoverde). Sertão (São Francisco Pernambucano and Sertão Pernambucano), with drier and more stable climate characteristics throughout the year, provides data with less seasonal variation for the model to capture patterns. In Agreste Pernambucano, which includes Garanhuns and Surubim, the model presents greater variability in percentage errors, especially in Garanhuns, where the MAPE validation reaches 54.8246%. This mesoregion has a humid climate and can be influenced by local phenomena. The city of Surubim, despite presenting a high percentage error in testing (42.9990%), has a slightly more balanced performance in validation (37.4277%). The climate variation of the Agreste, which has a transition between the dry climate of the Sertão and the more humid climate of the Zona da Mata, can make it difficult to capture patterns. In the Metropolitan Region of Recife, with its urban characteristics and high climate variation due to the influence of the ocean, it presents a percentage error of 46.9155% in testing and 45.7837% in validation. The model results are summarized in Table 8.

For XGBoost, Recife presented a MAPE of 43.76% in training, while Arcoverde had a higher MAPE, around 67.2%, which indicates that the model had difficulty learning complex patterns in the locations. In validation, errors increase for all cities, with Recife standing out, where the MAPE reaches 76.66%. The best results are in the FNN algorithm. Cities such as Arcoverde and Petrolina present MAEs around 2.07 and 2.11, respectively, with MAPE varying between 26% and 26.6%, which indicates good accuracy in these regions. Recife has a MAPE of 33.28% and MAE of 2.31.

Table 8 - Model Evaluation Metrics

Collection points	**Mean Absolute error (MAE)**	**Mean Absolute Percentage Error (MAPE)**	**Validation (MAE)**	**Validation (MAPE)**
		LSTM		
Arcoverde	1.8647	39.9136	1.8453	39.1658
Cabrobo	1.9144	40.8204	1.8904	37.1235
Garanhuns	1.8220	45.3659	1.7889	54.8246
Petrolina	1.9691	41.0090	1.8716	31.7328
Surubim	1.8820	42.9990	1.8410	37.4277
Recife	1.9406	46.9155	1.8413	45.7837
		XGBoost		
Arcoverde	1.7610	67.2044	2.0811	61.7300
Cabrobo	1.8107	67.4638	2.1141	75.1386
Garanhuns	1.0574	50.8401	2.1414	53.2101
Petrolina	1.8400	67.6806	2.0256	59.1410
Surubim	1.8286	65.4670	2.0343	62.3113
Recife	1.0907	43.7614	2.3363	76.6583
		FNN		
Arcoverde	2.0671	26.6168	2.0904	27.5519
Cabrobo	2.0903	26.8951	2.1251	27.8433
Garanhuns	2.2850	35.9728	1.8467	25.0528
Petrolina	2.1141	26.2415	2.0454	25.4250
Surubim	2.1417	28.8155	2.0402	27.5487
Recife	2.3154	33.2809	1.9939	32.6219

Note: Epoch 100/100

FNN had the lowest MAPE value across all cities. This is because FNN, being a sequential memoryless network, makes predictions considering only static patterns, which can work well when there is less complexity or temporal variability. In this way, FNN can smooth out the values and have a lower percentage error, since it is not trying to capture more detailed sequences.

LSTM had an intermediate performance with higher error percentages than FNN, but lower than XGBoost. This is because LSTM has memory cells that can capture temporal relationships and long-term dependencies, which is advantageous for sequential data, but can introduce some level of complexity and increase the percentage error, especially if variability is high. LSTM tends to be more accurate on data where seasonal or sequential patterns are present, but the adjustment to each city may not be optimal in all cases.

XGBoost had the highest MAPE values across almost all collection points. This is because XGBoost, a decision tree-based model, is excellent for tabular data but can struggle with temporal dependencies, especially in data with temporal variability. It overfits static or low-frequency data but loses accuracy on complex time series, which explains the higher percentage error.

5.3.2 Comparison of the proposed methodology

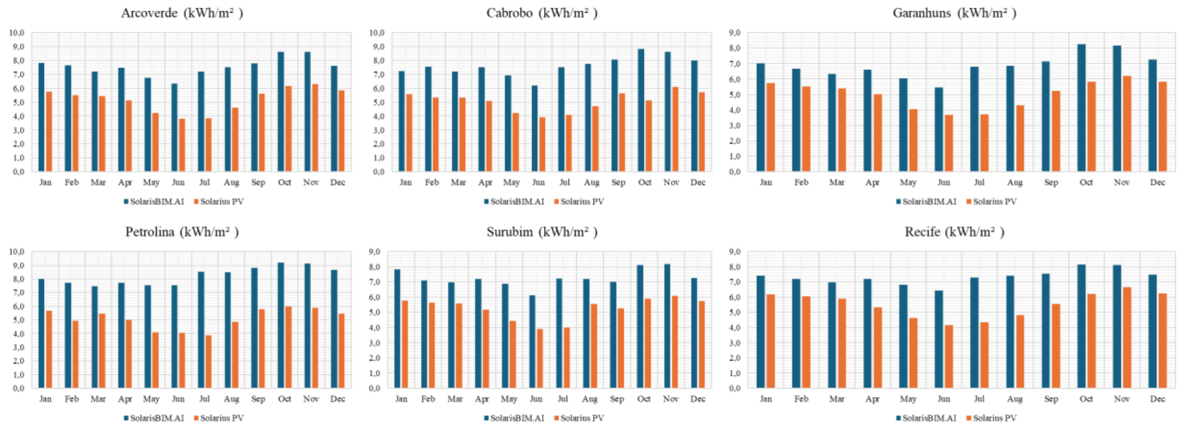
Considering this study's objectives, which prioritize minimizing relative errors in energy production forecasting across different locations, the FNN model was selected as the most suitable approach. The FNN consistently achieved the lowest MAPE values across most collection points in both training and validation phases, indicating greater reliability in proportional accuracy. However, it is worth noting that LSTM also presented good results when considering the MAE metric.

Noting that the FNN demonstrated better performance across the time series, this paper used its solar radiation predictions as input to the solar radiation data in the process of BIM quantifying solar energy production. This quantification process is called SolarisBIM.AI.

Figure 18 compares the solar irradiance predicted by the algorithm to the solar irradiance provided by the Solarius PV software. Solarius PV, a BIM software, enables the modeling and analysis of PV energy production using solar irradiance values from climate databases incorporated into the software. However, Solarius PV does not disclose detailed information about the algorithms used in sky modeling or about the matrix calculation methodology it uses (Di Giovanni *et al.*, 2024).

There is an average percentage variation of 30% between the daily average solar radiation values provided by Solarius PV and those predicted by SolarisBIM.AI. In some cities, such as Petrolina and Cabrobó, the variation reaches 38% and 33%. There is a smaller discrepancy in other cities, such as Recife (25%) and Garanhuns (27%). Despite the differences, the values predicted by the proposed model follow the variations of high and low solar radiation values, as shown in Figure 18.

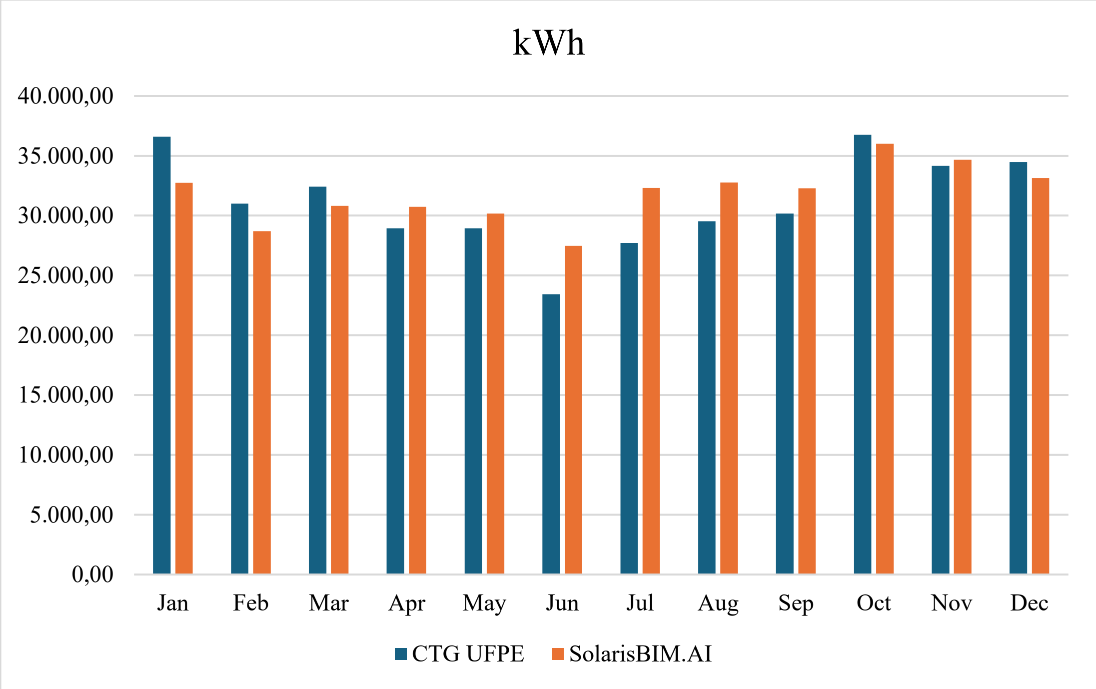
Figure 18 - Comparison between predicted values for the year 2023 and Solarius PV



The results of Di Giovanni *et al.* (2024) also showed differences in terms of comparison of solar radiation provided by Solarius PV. The results of Di Giovanni *et al.* (2024) show that the Solarius PV software overestimated the energy production of the PV plant and underestimated the global radiation received by the modules. In addition, the Solarius PV software, with the climate data provided by UNI 10349:2016, presents a relative error of 21.13%. In contrast, the simulation with the radiation data provided by PVGIS shows a relative error of 22.11% (Di Giovanni *et al.*, 2024).

The daily production data for the year 2023 from CTG UFPE were collected to compare with the predicted production data for the year 2023 from SolarisBIM.AI. The solar module model of CTG UFPE is the Astronergy CHSM6612P/HV-345, whose module efficiency is 18%. In total, CTG UFPE has 792 modules distributed on the roof of the building. Therefore, as shown in Figure 19, the predicted values from SolarisBIM.AI are very close to the actual values of CTG UFPE. For the year 2023, UFPE's actual solar energy production was 374,204.50 kWh, while the value predicted by SolarisBIM.AI was 381,802.31 kWh, establishing an annual difference of 2% more from the real case to the prediction. Although CTG UFPE has a total roof area of 2,863.87 m², only 1,536.80 m² are covered with solar modules.

Figure 19 - Actual solar energy production at CTG UFPE vs. Production predicted by SolarisBIM.AI

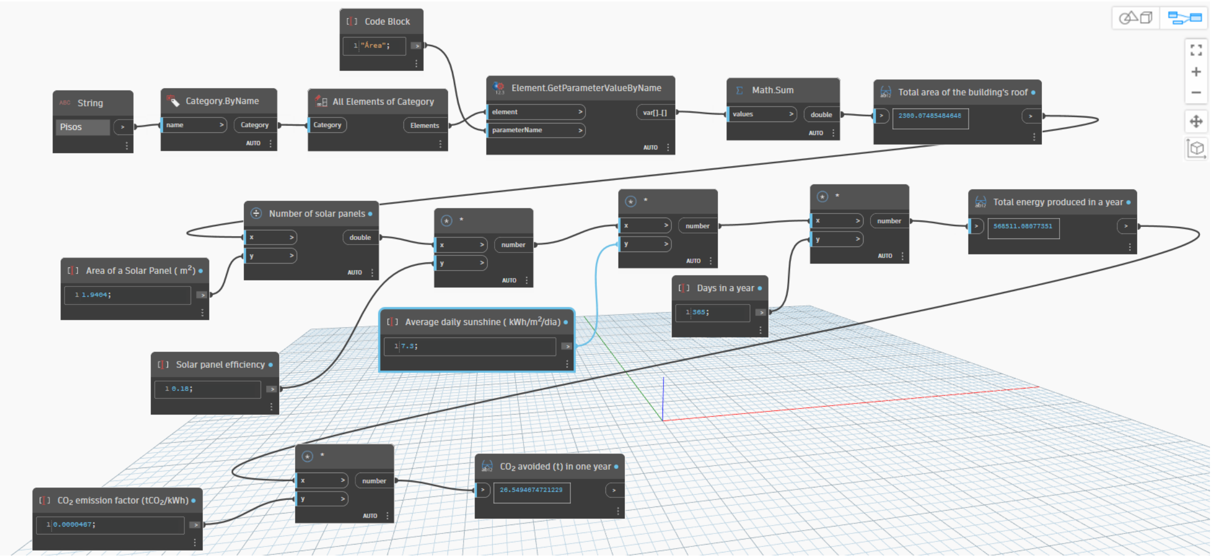


Previous research, such as Kazanasmaz *et al.* (2009), developed an FNN model with six building parameters, two weather parameters, and five weather parameters as inputs to predict daylight illuminance, and also obtained an average error of 2.2% (Kazanasmaz; Günaydin; Binol, 2009). These findings support assumptions investigated in research, such as that of (Liu et al., 2023b).

5.3.3 BIM routine for quantifying solar energy produced in a year and CO_2 avoid

Considering that the CTG UFPE has a useful area of 2,300.07 m² (removing access to stairs, cleaning, and water tanks), a routine was created in Dynamo focused on automating the quantification of roof areas of the UFPE CTG building in Revit, calculating the energy efficiency of a solar module system and, additionally, estimating the avoided CO_2 emissions based on the energy generated. The routine extracts the total roof area and uses the specific parameters of the solar module model, such as the efficiency of the solar modules and the average radiation, to calculate the annual energy production. It then multiplies this value by an adjusted emission factor to obtain avoided emissions, as shown in Figure 20 and described in Algorithm 4.

Figure 20 - Dynamo routine for building roof area extraction



Algorithm 4: PV Energy Calculation using Dynamo
Step 1. Extract the roof area Select **floor category** (roof) in Revit Retrieve **all elements** in this category Extract the **area** parameter for each element Compute the **total roof area** by summing all retrieved values **Step 2.** Determine the number of solar modules Define the **area of a single solar module** (m^2). Compute the **maximum number of solar modules** that fit in the total available area: $N_{SP} =$ **Step 3.** Calculate daily and annual energy production Define the **solar module efficiency** Define the **average daily solar radiation** ($kWh/m^2/day$). Compute **daily energy production**: D_{PV} (kWh) = N_{SP} ($D_{average}$) E Multiply by **366 days** to get the **annual energy production**: $PV_{anual} = D_{PV} * 366$ **Step 4.** Calculate CO_2 emissions avoided Define the **CO_2 emission factor** (tCO_2/kWh). Compute the **total CO_2 emissions avoided per year**: $CO_{2avoided} = PV_{anual}$ (kWh) $CO_{2\ factor}$ (tCO_2/kWh)

Step 5. Output Results Total usable roof area (m^2) Number of solar modules Total energy produced annually (kWh) CO_2 emissions avoided per year (tons) End code

This Dynamo routine highlights the importance of modeling in the design phase for BIM projects focused on PV installations. The application on roofs minimizes energy losses, optimizes installation costs, and maximizes solar radiation capture. Thus, the location, characteristics, and dimensions of roofs become fundamental when analyzing installations' viability, production, and profitability (Martín-Jiménez *et al.*, 2020). This research contributes to the AECO sector by expanding the use of two-dimensional data to integrate the BIM into an automated design process, aiming to achieve optimal solar energy production values.

To quantify solar energy production and CO_2, parameters such as solar module model, roof area, module efficiency, and CO_2 emission factor (0.0000467 tCO_2/kWh) were applied to the Dynamo routine. The Astronergy CHSM6612P/HV-345 model was randomly selected for this simulation, with a total roof area of 2,300.07 m^2. The number of solar modules required to cover this area was calculated based on the area of one module, which is 1.9404 m^2, resulting in approximately 1,185 solar modules. The module efficiency provided by Astronergy is 18% (meaning that this system converts 18% of the incident solar energy into useful electricity). In addition, to calculate the CO_2 reduction, the emission factor of 0.0000467 tCO_2/kWh was used, which indicates the amount of carbon dioxide avoided for each kilowatt-hour of energy generated.

The potential daily energy generation in the cities ranged from 1,468,79 kWh (Garanhuns) to 1,757.23 kWh (Petrolina), with the highest daily production occurring in the Sertão region. The avoided CO_2 results show that Petrolina, which has the largest energy production, also avoided the largest amount of CO_2, with an annual reduction of 29.95 tons. Arcoverde also stood out in CO_2 avoidance (27.46 tons per year). Garanhuns and Surubim, with smaller energy productions, also presented a lower CO_2 avoidance (25.04 and 26.41 tons, respectively). These results are summarized in Table 9.

Table 9 - Simulation of annual solar energy production and avoided CO_2

Results	Arcoverde	Cabrobó	Garanhuns	Petrolina	Surubim	Recife
Daily energy (kWh)	1,611.19	1,627.08	1,468,79	1,757.23	1,549.19	1,565.26
Annual energy (kWh)	588,082.77	593,884.27	536,109.17	641,389.23	565,455.11	571,319.09
CO_2 avoided (t)	27,46	27,73	25,04	29,95	26,41	26,68

The Sertão region, characterized by more intense solar radiation, generates greater energy production and, consequently, greater CO_2 avoidance, which supports the strategic installation of PV systems in these areas to maximize economic and environmental benefits. The analysis of solar energy production and CO_2 reduction avoided in the different cities of Pernambuco shows the importance of efficient planning and design in implementing PV systems.

Using tools such as Dynamo and Revit, which integrate with BIM, can assist in analyzing local solar conditions and the precise dimensioning of PV systems, considering the specific characteristics of each region. Thus, in cities such as Petrolina and Cabrobó, where solar radiation is more intense, the BIM model can optimize the layout of the systems, ensuring greater efficiency in energy production. On the other hand, in areas such as Garanhuns, where solar radiation is more moderate, the use of BIM can help to plan alternative solutions, such as the use of more efficient modules or the inclusion of complementary systems, such as storage batteries, to maximize the use of the energy generated.

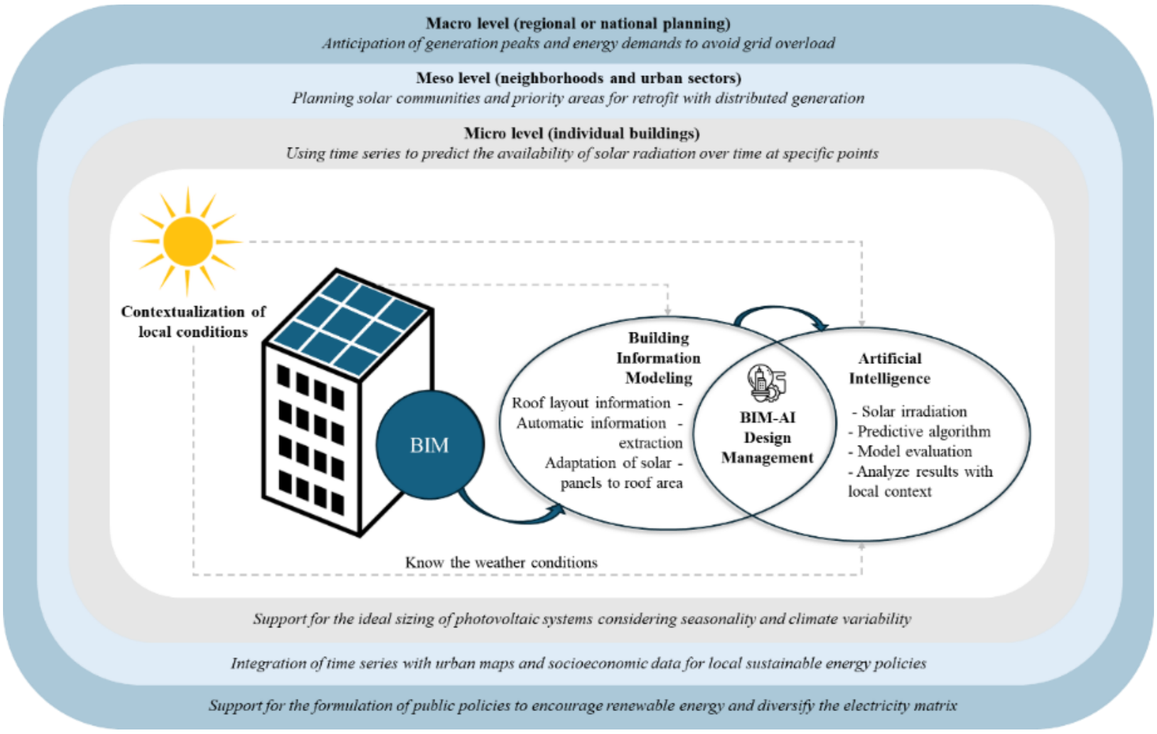
5.4 Discussion

This chapter establishes an automated process for predicting and quantifying solar energy production, highlighting the role of BIM and project planning in the Design phase. Figure 21 highlights that the design phase is essential for the proper operational functioning of a rooftop PV system. Using BIM from the early stages of project construction contributes to optimized information management, facilitating the design process and multidisciplinary collaboration between the different stages of the project (Di Giovanni *et al.*, 2024). Applying solar radiation data and BIM information at the design stage helps designers optimize PV plants adapted to local climatic and regional conditions.

Some factors can influence the adaptation of solar modules to building roofs that can be optimized or avoided in the design phase. The available area on roofs must be analyzed in advance since space limitations can restrict the installation of PV systems in buildings. Several structural components, such as chimneys, machine rooms, exhaust fans, and plumbing outlets, also impact the physical layout for installing solar modules. In addition, self-shading between modules in close rows can reduce the amount of solar radiation received, compromising efficiency. Another important point is the variety of sizes of PV modules available on the market, which allows for certain flexibility in adapting these modules to the roofs of different buildings (Barbón *et al.*, 2022).

Furthermore, on a macro scale, energy planning needs to anticipate and mitigate risks associated with grid overload, especially in light of the rise in distributed generation and the growing electrification of cities. The use of forecasts based on solar radiation time series, combined with deep learning models and data integrated via BIM, allows policymakers to identify consumption and generation patterns with greater precision. With this information, it is possible to define load balancing strategies, such as encouraging the adoption of energy storage systems, implementing dynamic tariffs and the intelligent expansion of transmission infrastructure.

Figure 21 - SolarisBIM.AI - Smart sustainable building planning with solar radiation forecasting for photovoltaic energy production with a BIM-driven deep learning model



From a theoretical perspective, this study contributes to sustainability research in the built environment by articulating its implications across micro, meso, and macro levels of application. At the micro scale, the proposed SolarisBIM.AI framework addresses individual building projects by enabling the automated estimation of photovoltaic energy production and avoided CO_2 emissions directly from BIM models. This allows project designers to optimize energy strategies during the early design phase, enhancing building-level sustainability and energy efficiency (Gan, 2022; Wijayarathne; Gunawan; Schultmann, 2024).

At the meso scale, the model supports professional practices within architectural and engineering firms by integrating artificial intelligence into routine design workflows. The automated process developed in Dynamo addresses the standardization of sustainable design practices, improves decision-making across teams, and fosters digital transformation in the AECO sector. At the macro scale, this research supports broader urban and policy agendas by contributing to the operationalization of net-zero carbon targets and energy transition goals. The ability to quantify the environmental impact of design alternatives at scale can inform municipal planning strategies, energy resilience programs, and sustainability certifications. Thus, the integration of BIM and AI presented in this study has the potential to align with national and global policies for decarbonization, particularly in the context of rapidly urbanizing regions and climate-sensitive zones (Chen; Gou, 2024; D'adamo *et al.*, 2024).

In the sustainability domain, beyond the economic and environmental benefits, PV energy promotes the democratization of access to electricity, as it is a decentralized source that reinforces the security of energy supply. Unlike centralized generation, PV systems can be installed in different locations and do not necessarily require a connection to the electricity grid. As a result, owners of these systems are protected against increases in energy costs and supply interruptions caused by extreme events (Kim; Kim; Kim, 2017; Qiu; Yang, 2024). For example, government measures have been widely adopted to encourage the implementation of PV projects, such as the feed-in tariff (FIT). The FIT offers project operators a fixed payment per kilowatt-hour (kWh) of electricity generated from PV systems. This policy has played a key role in stimulating the development of solar energy in several countries, both developed countries such as the USA, New Zealand, Japan, and Italy, and developing countries such as Chile, Vietnam, and China, which have also implemented a series of policies to encourage the expansion of PV projects (Hu; Song; Zhao, 2024).

For BIM modelers, these incentives increase the demand for infrastructure projects incorporating renewable energy sources, requiring the integration of PV systems with architectural and structural design. AI complements this need by enabling optimized analysis of performance data, climate simulations, and energy generation forecasting, increasing efficiency in the planning and maintenance of solar systems (Ling *et al.*, 2023; Shao; Meng; Che, 2025).

6 FROM COST TO EFFICIENCY: A BIM-DRIVEN APPROACH FOR PHOTOVOLTAIC MODULES ALLOCATION TO MAXIMIZING ENERGY PRODUCTION WHILE MINIMIZING COSTS

6.1 Chapter Introduction

Addressing the urgent challenge of energy sustainability requires expanding renewable sources and driving the rapid adoption of innovative technologies in the Architecture, Engineering, Construction, and Operations (AECO) sector (Winkler *et al.*, 2024). This includes the production of photovoltaic (PV) energy through project planning from the initial phases of buildings and extending throughout their entire life cycle (Serat *et al.*, 2025; Szalay *et al.*, 2022). For this planning to be efficient, optimized, and data-driven, design processes must provide the best configurations for allocating solar modules on building roofs (Di Giovanni *et al.*, 2024). To this, simulations of PV plant projects and integration between the different disciplines that make up a building are possible through digital representations made by BIM models. In this context, this chapter integrates two research topics to provide an automated design process for simulating alternative layouts of PV systems on a building´s roof.

First, this chapter argues that BIM models are the basis for simulating PV array layout alternatives. BIM models store and integrate data throughout the building life cycle through geometric and non-geometric information about building objects. When applied in the early design phases, the models provide the information needed to document and coordinate solutions that justify design alternatives (Palha *et al.*, 2024). Furthermore, as sustainable demands are placed on the AECO industry, new technologies and automated design processes are needed for the long-term growth of the sector (Araújo *et al.*, 2020b; Araujo; Alves, 2025). The association of BIM models and PV systems establishes a data-driven workflow for the AECO industry. Thus, during the planning and development phase of the project, BIM modelers analyze the compliance of the detailed design of the PV system, structural load of the PV system, power generation, and electrical flow with the other disciplines and structures of the building. After the project is completed, the BIM model becomes a data platform and a management system for the PV system by analyzing the system's performance during operation (Lin *et al.*, 2021; Wang *et al.*, 2024).

Second, isolated planning of PV systems can perpetuate interoperability problems between software and accentuate conflicts between the different disciplines involved (Palha *et al.*, 2024). Li *et al.* (2021) highlight that the conventional workflow using BIM for developing

PV systems usually involves using multiple platforms. In this process, the architectural model with its geometric information is transferred to simulation software, either through the Industry Foundation Classes (IFC) format or by rebuilding the model in the simulation tool. After analysis, the data generated is integrated back into the architectural model for the insertion of the photovoltaic system. In contrast, the workflow proposed by researchers such as Li *et al.* (2021) and in this chapter simplifies this process by using a single BIM platform, Revit, associated with the Dynamo plugin.

The ideal workflow for designing and constructing PV systems should start at the planning and design phase of the building. For example, Wang *et al.* (2024) highlight that in the design phase, simulation technologies can be employed to optimize the PV layout to maximize solar power generation efficiency. This includes analyzing the efficiency of different PV arrays to enable refinement of the asset design at an early stage. From the information on the construction site, it is possible to analyze factors such as solar radiation and specific conditions of roof structural elements (Shao *et al.*, 2024). The design phase has three main parts: system design, electrical design, and structural design. This chapter specifically discusses the system design in buildings. The system design includes the layout of the PV array, system size, module location, and how to integrate the system efficiently (Lin *et al.*, 2021).

Nevertheless, determining the efficiency of PV systems considers aspects such as solar radiation (to estimate total energy production), available area on the building roof, PV panel brands and models, and the financial analysis for each PV panel model and brand. Several solar design and simulation tools on the market are specific to PV, such as RET Screen, Homer Pro, SAM, PVsyst, PVwatts, Polysun, and PVSol (Jing Yang *et al.*, 2024; Lin *et al.*, 2021). However, the analyses performed with these tools are based on the geometry of the building's roof, which must be created manually (which generates rework).

This chapter argues that the AECO literature can benefit from developing strategies geared toward the advanced planning of PV plants. The focus is on guiding the planning of the building roof design to the PV project and establishing better layout alternatives in the case of retrofit projects. For this, Alves *et al.* (2025) argue that BIM can be considered a source of information for developing smart projects based on data management. Furthermore, Palha *et al.* (2024) argue that geometric and non-geometric representations of construction elements are developed using BIM to mitigate recurring failures in the traditional method in 2D Computer-Aided Design and overcome limitations in the exchange of information in collaborative environments. Palha *et al.* (2024) also argue that the application of BIM is more consolidated in the initial stages of projects, especially in project documentation and coordination. Lin *et al.*

(2021) highlight that the lack of support tools in the initial design phase is one of the main problems hindering the use of photovoltaic tools in designing PV layouts.

Previous research demonstrates advances in the planning process for integrating BIM with photovoltaic systems. Li *et al.* (2021) developed a Revit API built-in class and user control classes for PV module planning, including filters, user selection, bounding box, sun and shadow setting, reference intersection, family manager, and spatial field manager. Di Giovanni *et al.* (2024) use solar radiation data on building surfaces, accessed through Insight 360 (Revit plug-in), and apply different machine learning techniques to predict long-term photovoltaic energy generation. (Changsaar *et al.*, 2022) developed a photovoltaic system in an Eco-Home model and demonstrated that it has the potential to generate annual energy savings of 26,552 kWh, with an estimated maximum payback period of 18 years — a value considered adequate within the standard useful life of photovoltaic systems. (Szalay *et al.*, 2022) developed BIM models to analyze life cycle carbon neutrality and prove that this can be achieved through additional photovoltaic panels installed on the roof, establishing a model for nearly zero energy buildings (nZEB). (Lu *et al.*, 2022) developed a method based on the IFC standard in BIM to estimate the potential area and location of photovoltaic modules on the surface of buildings, considering windows (47,589.793 kWh/year), roofs (141,126.304 kWh/year), and facades (284,060.393 kWh/year). However, these approaches still have limitations when it comes to automating the PV module allocation process directly in Revit models. The body of knowledge can advance the proposition of a model to analyze the cost-energy production relationship considering different sizes and brands of PV modules.

This research aims to develop an automated process optimization model for allocating solar modules, seeking to maximize photovoltaic energy production (kWh/day) while minimizing implementation costs. This research integrates visual programming in Dynamo with programming in Python to analyze different combinations of PV modules, considering the dimensions of 21 PV modules from 4 brands for allocation on the roof of a building. The algorithm identifies the most efficient configuration of photovoltaic cost-production. It uses Dynamo to extract information on the families of PV modules and the available roof area from a BIM model in Revit. Finally, the model automatically allocates the best arrangement of PV modules directly in the Revit model.

This research advances the body of knowledge on BIM and sustainable projects by proposing an automated process for selecting and allocating PV modules. First, the method starts by selecting solar models available on the market. Then, based on research related to costs associated with solar energy implementation, the final cost of the modules for customers is estimated. Third, the production of the modules is estimated. By automatically extracting information from the building's roof via Dynamo, the algorithm selects the best allocation solution, based on maximizing solar energy production and minimizing cost. Finally, Dynamo allocates the best configuration directly in the BIM model.

6.2 Methodological Approach

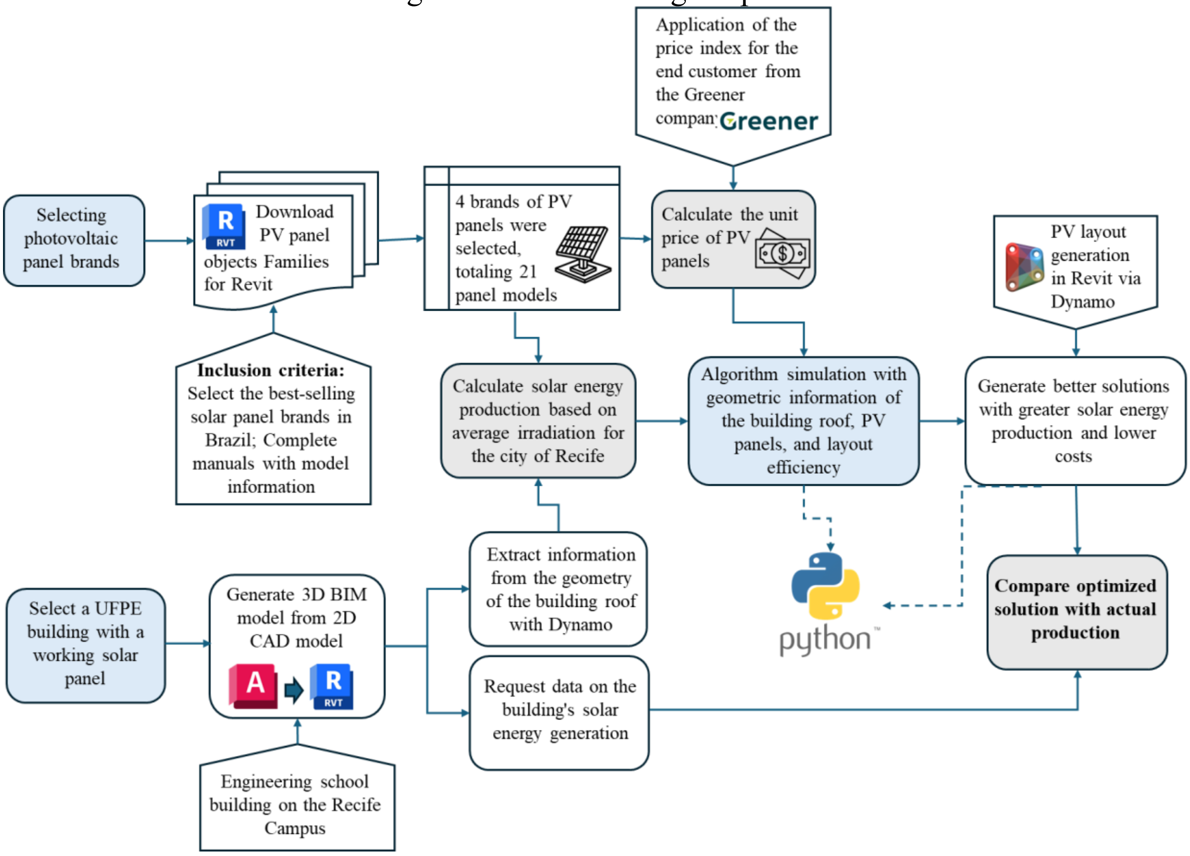
This chapter implements an automated process model to determine the best PV module allocation within an available area. It seeks a balance between PV energy production (kWh/day) and cost (USD $). The best PV module configuration is given by maximizing energy production (kWh/day) and minimizing total cost (USD $).

The methodological process of this research is structured into two simultaneous stages. First, four PV module brands were selected, and their manuals were reviewed to gather technical specifications of the series and models available. Next, corresponding BIM object families representing these panels were identified and imported into Revit. With the geometric and technical data of the PV modules integrated into the BIM environment, it was then possible to calculate their unit price, which served as a input for assessing the efficiency of the photovoltaic layout.

The second stage involves selecting the case to be analyzed and compared with the simulation results proposed in this article. The UFPE engineering school building was chosen to compose the case. This building already has a PV system, and as it is a public entity, data on solar energy generation can be made available upon request. Thus, the building's floor plan was converted from 2D CAD (made in AutoCAD) to a 3D BIM model in Revit. The extraction of information on the geometry of the building's roof was done with Dynamo, a Revit plugin that allows the development of automated projects through visual programming. Dynamo also has the functionality of allowing the insertion of code in Python. Based on information on the geometry of the building's roof and the average daily solar radiation in the city of Recife-PE, it was possible to estimate the daily production of solar energy for the year 2023.

With the information on cost and energy production, an efficiency indicator was created to analyze the different PV layout combinations generated by the algorithm. Finally, with the information on the actual solar energy production of the selected building, it was possible to compare the PV layout generated by the algorithm with the solution currently implemented by UFPE. This methodological process is shown in Figure 22 and described in detail in the following subsections.

Figure 22 - Methodological process



It is important to note that PV cells can be classified into different types according to the semiconductor material and the manufacturing process used, directly influencing solar systems' efficiency, cost, and application. The main types of PV cells include monocrystalline (Mono-Si), polycrystalline (Poly-Si), and thin-film (Thin-Film), in addition to emerging technologies, such as perovskite cells. Monocrystalline cells, highlighted in this study as the technology used in simulations, are known for their high conversion efficiency and superior performance in low irradiation and high-temperature conditions. Produced from a single high-purity silicon crystal, these cells have a uniform and dark appearance and higher power density, making them ideal for projects where the available space for installation is limited (Pupin *et al.*, 2023). In the context of this article, the simulations were performed with modules based on mono-crystalline cells, considering their high energy efficiency and wide application in BIPV systems integrated into the architecture of buildings.

6.2.1 Case selection

The algorithm for PV module allocation was implemented in an existing building at the UFPE. The building already has a photovoltaic plant installed and operating since the second half of 2022. To this end, the algorithm follows the BIM-oriented approach by extracting information about the geometry of the building's roof.

Brazil has reached an installed capacity of 55 GW in photovoltaic (PV) energy, consolidating its position as the sixth largest country in the world in terms of installed capacity and positioning this source as the second largest national energy matrix, far surpassing wind energy. Distributed generation (DG), which includes systems installed in buildings, represents the largest part of this capacity. In February 2025, solar PV energy was responsible for 12.10% of all electricity supply in the country, contributing to avoiding the emission of more than 66.6 million tons of CO_2 and boosting the economy by generating around 1.6 million new jobs. At the state level, Pernambuco ranks 7th among Brazilian states in solar energy generation (ABSOLAR, 2025). Despite the growth of the national market, most of the PV modules and panels used are still imported from China due to Brazil's limited manufacturing capacity of photovoltaic components (Pupin *et al.*, 2023).

UFPE, a public federal higher education institution, is one of Brazil's leading higher education institutions, recognized for its academic excellence and innovation. Founded in 1946 and headquartered in Recife, it also has Caruaru and Vitória de Santo Antão campuses. With a wide range of undergraduate and graduate courses in various areas of knowledge, UFPE stands out, especially in technology, exact sciences, humanities, and health. The university is committed to innovation, internationalization, and social impact (UFPE, 2025).

UFPE is committed to sustainability and energy efficiency. As part of its policy to reduce environmental impact, it is installing PV modules on its buildings. The initiative is part of a program to expand the use of renewable energy on campus, reduce dependence on the conventional electricity grid, and generate long-term savings (UFPE, 2021). However, the logic of project selection and prioritization in higher education institutions in Brazil, as exemplified in the case of UFPE, is strongly influenced by a cost-effectiveness assessment approach, where the costs of similar alternatives are compared to support decision-making. In the institutional context, project demands arrive through two distinct channels: through the university presidency or through organic requests entered the corporate system (SIPAC). Projects forwarded by the presidency have clear priority, as they are part of the annual strategic planning. They generally have guaranteed resources, such as parliamentary amendments, which result in larger-scale projects with financial impact. In contrast, projects coming from SIPAC are more varied in scope and value and may be requested continuously throughout the year, which creates challenges for planning and resource allocation. Currently, the predominant criterion for developing these projects follows the FIFO logic (first in, first out) (Nascimento *et al.*, 2023).

Energy consumption in public academic buildings is four times higher than in other types of buildings. Academic buildings generally require large amounts of energy to keep classrooms cool and laboratory equipment running. This demand is even greater in full-time educational institutions (Chen *et al.*, 2025).

Furthermore, the Brazilian Federal Government, through Decree Nº. 10,306, of April 2, 2020 (Brazil, 2020), establishes the mandatory use of BIM in engineering works and services performed by federal public administration agencies and entities. Implementation occurs in phases, starting with building designs and modeling and progressively expanding to the execution, maintenance, and management of the life cycle of buildings. In this context, this article contributes to UFPE by implementing the BIM design process by converting a 2D CAD project to BIM in Revit and structuring an automated method for analyzing the efficiency of PV module allocation.

The selected UFPE building called the Engineering School Building, located on the Recife - PE Campus, has complete daily records of solar energy production with no missing values for 2022 and 2023. For the comparisons explored in the results of this article, the records from 2023 are used. The data on daily PV energy production was made available by the High Voltage and Public Lighting Management (GATIP) linked to the Infrastructure Superintendence (Sinfra) of UFPE. Sinfra provides building and urban maintenance and conservation services. It aims to preserve the integrity of the infrastructure and guarantee the quality of access to the University's facilities, ensuring environmental Sustainability (UFPE, 2025).

Finally, UFPE is part of the Federal Institutions of Higher Education (FIHL) of the Federal Government. FIHLs have limited financial and personnel resources to develop projects, despite having a high demand for requests. In addition, decision methods for selecting project portfolios from public institutions must consider the wide range of benefits associated with the costs involved (Nascimento *et al.*, 2023). Thus, this research also advances by providing an automated strategy to select different layout alternatives for PV projects in buildings based on costs, PV energy production, and efficiency of the adopted solution.

3.2 Calculating the production and cost of PV modules

The family was downloaded and imported into the Revit model for each indicated PV module brand. The simulation considered all four brands' models for the series of modules available in the manual. Table 10 shows information on the brands and models of the PV modules.

The calculation of the unit price of the PV modules was based on market research carried out by Greener, a renowned Brazilian company specializing in market intelligence, consulting, and studies for the photovoltaic solar energy sector. Greener is recognized for providing detailed reports on the photovoltaic market, financial analyses, feasibility studies, and strategic data, assisting companies and investors in the renewable energy sector (GREENER, 2024).

According to Greener's research, the end consumer pays, on average, USD $ 1.09 per watt of maximum nominal power (Pmax) of each PV module. Therefore, to estimate the unit cost of each panel model, its maximum nominal power (Pmax) was multiplied by the reference value of USD $ 1.09/W. Thus, the cost was calculated according to Equation 12.

$$Cost = Pmax\ (W) \times \frac{\$\ 1.09}{W} \tag{12}$$

According to Chapter 5, the average daily solar radiation in the year 2023 for the city of Recife, Pernambuco, Brazil, is 7.6 kWh/m² (kilowatt-hours per square meter). Thus, the daily production of PV energy is the multiplication of the value of solar radiation (SR), the efficiency of the module (η), and the area of the panel (A). Thus, the PV production was calculated according to Equation 13.

$$PV\ production\ \left(\frac{kWh}{day}\right) = SR\ \left(\frac{kWh}{m^2\ /\ day}\right) \times \eta \times A\ (m^2) \tag{13}$$

Equation 12 and Equation 13 calculated the production and costs of each PV module. The results are summarized in Table 10.

Table 10 - General information on solar modules, cost, and unit production

Panel brand	ID	Nominal Max. Power (Pmax)	Efficiency (η)	Unit Cost	PV Production (kWh/day)
A Dimension: 1,16 m x 2,07 m Area: 2,401 m²	1	475	0,198	$515.83	3,613
	2	480	0,200	$521.26	3,649
	3	485	0,202	$526.69	3,686
	4	490	0,204	$532.12	3,722
	5	495	0,207	$537.55	3,777
	6	500	0,209	$542.98	3,814
B Dimension: 0,99 m x 1,62 m Area: 1,634 m²	7	300	0,184	$325.79	2,284
	8	310	0,190	$336,65	2,358
	9	320	0,196	$347.51	2,433
C Dimension: 1,30 m x 2,38 m Area: 3,106 m²	10	640	0,206	$695.02	4,863
	11	645	0,208	$700.45	4,910
	12	650	0,209	$705.88	4,934
	13	655	0,211	$711.31	4,981

	14	660	0,212	$716.74	5,004
	15	665	0,214	$722.17	5,052
	16	670	0,216	$727.60	5,099
	17	210	0,1923	$228.05	1,595
D	18	215	0,1969	$233.48	1,634
Dimension: 0,7 m x 1,56 m	19	220	0,2015	$238.91	1,672
Area: 1,092 m²	20	225	0,206	$244.34	1,709
	21	230	0,2106	$249.77	1,747

6.2.2 Algorithm implementation

6.2.2.1 Algorithm objectives

The process implemented in this paper seeks the best arrangement of PV modules that produce the most solar energy at the lowest possible cost through the efficiency parameter. The algorithm's first objective is to maximize the amount of solar energy generated, as defined in Equation 14.

$$P_t = \sum_{i=1}^{n} MaxQ_i \times P\ (kWh/day)_i \tag{14}$$

Where:

- P_t is the PV module configuration's total energy production (kWh/day).
- n is the number of different types of PV modules in the combination.
- $MaxQ_i$ is the maximum amount of panel i that can be installed in the available area.
- Production $(kWh/day)_i$ is the daily energy production per panel i.

The second objective is to minimize the cost of PV modules, which is defined as shown in Equation 14.

$$C_t = \sum_{i=1}^{n} MaxQ_i \times P_{price(R\$)i} \tag{15}$$

Where:

- C_t is the total cost (USD $) of the PV module configuration.
- $MaxQ_i$ is the price of each panel i (Unit Cost).

The algorithm uses a function derived from the two objectives to represent efficiency, which is given by the ratio between total production and total cost, as shown in Equation 15. The algorithm then seeks to maximize efficiency, balancing energy production and price.

$$\eta_{(kWh/USD\ \$)} = \frac{P_t}{C_T} \tag{16}$$

6.2.2.2 Logical structure of the algorithm in Revit via Dynamo

First, the available area on the building roof is extracted in Dynamo via ‘Select Model Element’ (Figure 23). Then, for each combination of PV modules, the algorithm selects the corresponding panels from the database (families imported into the 3D Revit model – Figure 24a) and calculates the maximum quantity of each panel installed in the available area (Figure 24b). The maximum number of units of each panel is calculated by dividing the available area by the area of each panel, and the total production and cost are obtained by multiplying the quantities by their respective daily production and prices. This results in total energy production value and cost per configuration. The efficiency of each combination is defined as the ratio between the total production (in kWh/day) and the total cost (in USD $).

After defining the best allocation, the one with the highest efficiency, the algorithm identifies the families (brands of PV panels) and models and allocates them according to the best configuration. The algorithm via Dynamo finds the families imported into Revit by the ID of the panel model.

Figure 23 - Dynamo structure for extracting information from the building´s roof and module allocation

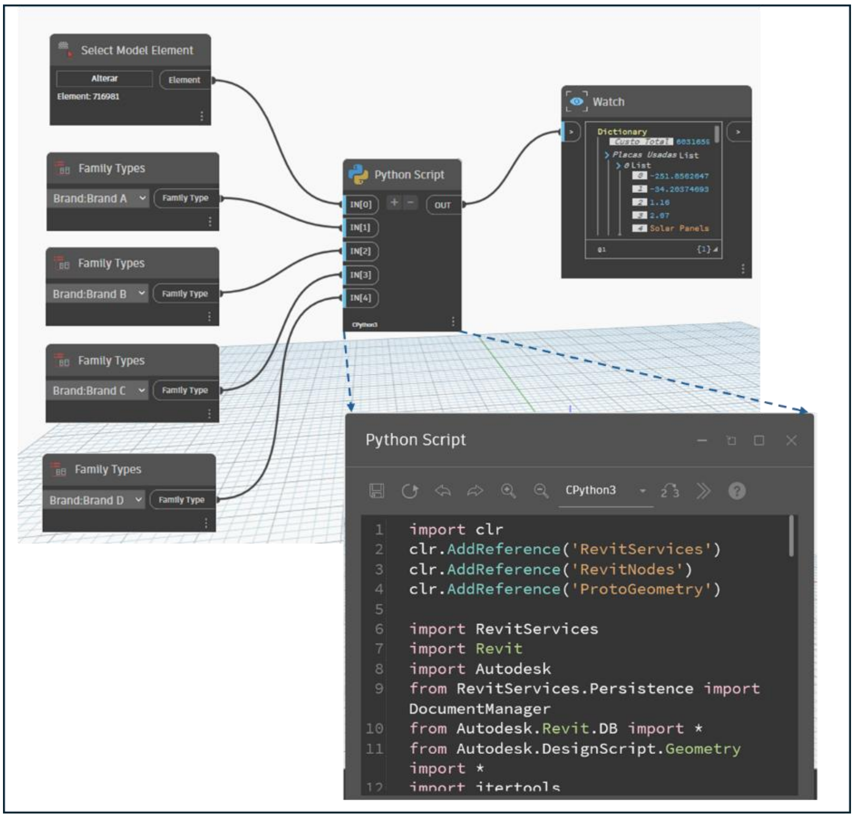
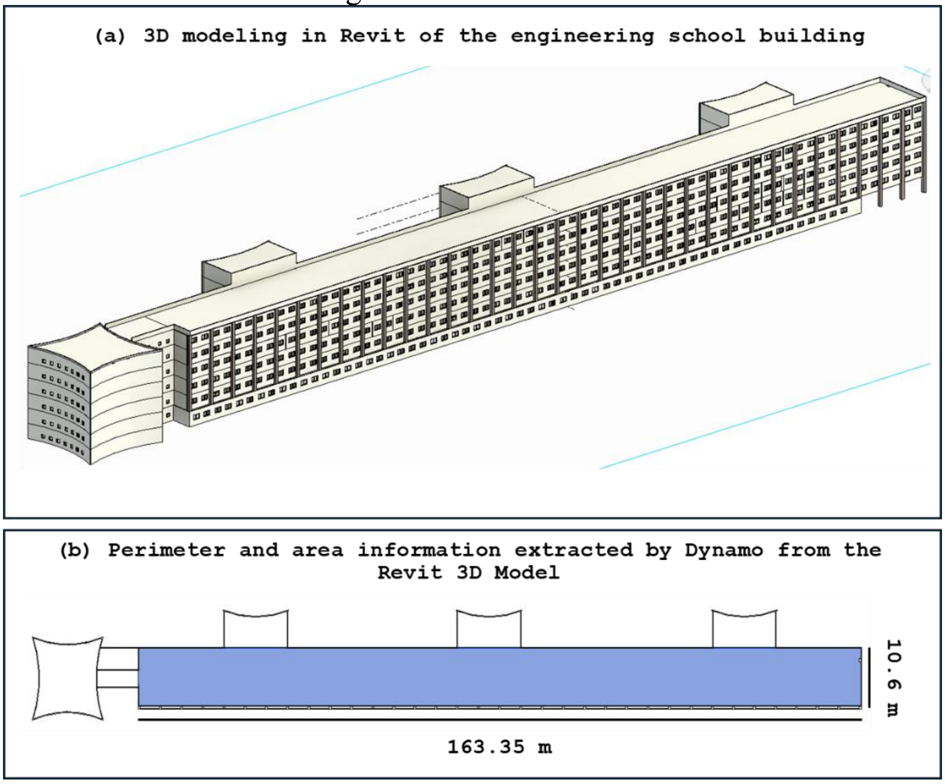


Figure 24 - Revit model



Algorithm 5 uses Dynamo to optimize and automatically allocate PV modules in a Revit model. The algorithm first extracts information from the model's coverage using Dynamo, selecting the model elements (Step 1). It then generates all possible combinations of PV modules and prepares a list to store the results (Step 2). During the simulation, the algorithm calculates the maximum number of each panel allocated in the available area, creating a temporary data frame with this information (Step 3). It then calculates the total production and cost of the panel configuration, as well as the efficiency of the configuration, which is the ratio between output and cost (Step 4). The results are stored and converted to a final 'DataFrame' (Step 5). The algorithm then identifies the best configuration, with the highest efficiency, to be used for the allocation (Step 6). Finally, the transaction in Dynamo/Revit is started to create instances of PV modules in the model, calculating the positions of each panel and creating instances in Revit, which are added to the list of panels, and the transaction is completed (Step 7). The algorithm returns to the list of created PV module instances as output.

Algorithm 5: Automatic PV module optimization and allocation

Step 1. Extract information from the BIM model
    Select the roof element in the Revit model
    available_area ← ***Extract*** available roof area from selected element
 	
Step 2. Generate combinations of solar modules

possible_combinations ← Generate ***all*** combinations of available PV modules
results_sampled ← Empty list ***to*** store simulation results

Step 3. Run simulations for each combination
For each combination ***in*** possible_combinations:
selected_panels ← ***Select*** modules from current combination
df_comb ← ***Create*** DataFrame with module data
For each module **i** in df_comb:
$MaxQ_i \leftarrow \text{roof}\left(\frac{available\_area}{Area_i}\right)$
$Cost_i \leftarrow Pmax_i \times 1.09$ # Equation (12)
$PV\_production_i \leftarrow SR \times \eta \times Area_i$ # Equation (13)
$TotalProduction_i \leftarrow MaxQ_i \times PVproduction_i$
$TotalCost_i \leftarrow MaxQ_i \times Cost_i$

$P_t \leftarrow$ Sum ***of all*** $TotalProduction_i$ # Equation (14)
$C_t \leftarrow$ Sum ***of all*** $TotalCost_i$ # Equation (15)
$\eta \leftarrow \frac{P_t}{C_t}$ # Equation (16)
Store [df_comb, P_t, C_t, η] ***in*** results_sampled

Step 4. Select the best configuration
df_results_sampled ← ***Convert*** results_sampled into a DataFrame
best_index ← Index ***of*** the highest efficiency (η) ***in*** df_results_sampled
best_combination ← ***Corresponding*** combination ***with*** highest efficiency

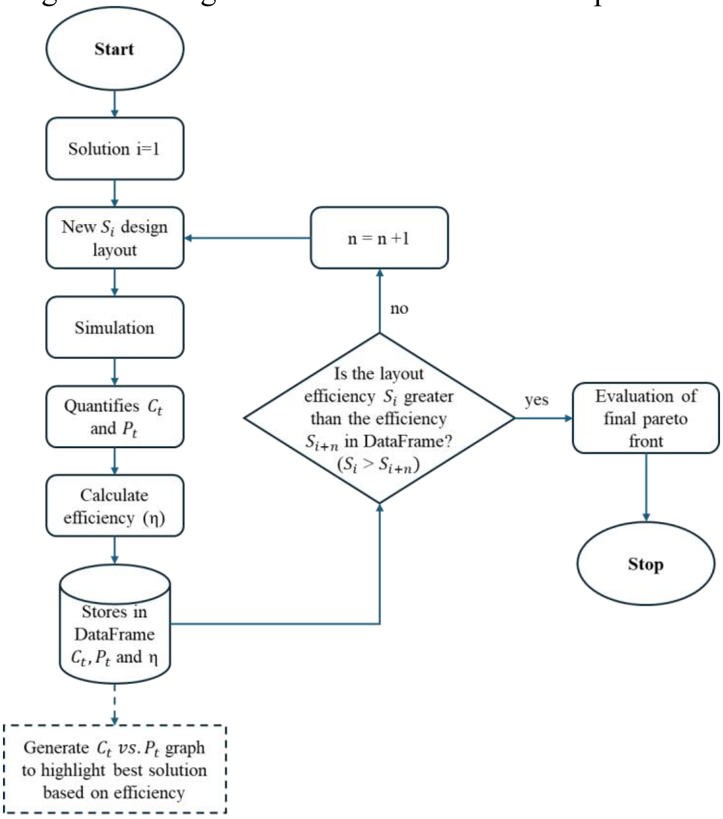
Step 5. Automatically place modules in the Revit model
Start transaction ***in*** Dynamo/Revit
For each module type **i** in best_combination:
For j ***from*** 0 to $MaxQ_i$ - 1, do:
x ← minPt.X + i × (panel_width + spacing) + (panel_width / 2)
y ← minPt.Y + j × (panel_length + spacing) + (panel_length / 2)
z ← minPt.Z
position ← XYZ(x, y, z)
PV_module ← **Search** family instance at position in Revit
Add PV_module ***to*** the module list
OUT ← modules

End transaction

6.3 Results

The results were developed through two experiments that follow the logical structure shown in Figure 25. The algorithm stores all simulated solutions in a DataFrame, highlighting the best solution through graphs. Only the best solution is allocated to the building's roof through Dynamo.

Figure 25 - Logical structure of the results experiment



6.3.1 Total costs and Production in PV module Allocation

Table 11 shows the different brands of PV modules (A, B, C, and D) and presents production and cost data associated with each set of panels. In all brands, the increase in production is related to the rise in cost. For example, brand A and model ID 1 require 721 panels allocated in a total area of 1731.51 m². In addition, the daily production in ID 1 is 2,605.21 kWh and rises to 2,749.94 kWh in ID 6—the total cost increases from USD \$ 371,915.83 to USD \$ 391,490.35. In addition, despite having the smallest number of panels (557), brand C has the highest production, with 2,840.36 kWh in ID 16, but requires greater investments.

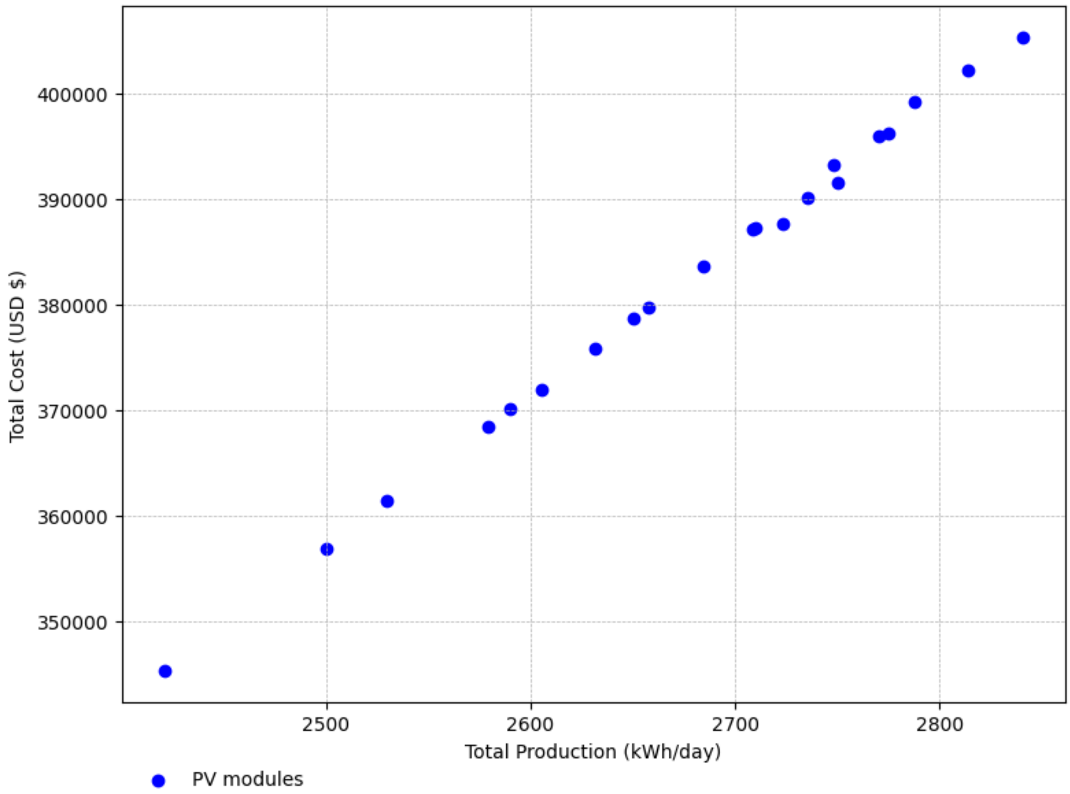
Table 11 - Single PV modules allocation

Brand	**ID**	**Total of panels**	**Total Production (kWh/day)**	**Total Cost (USD \$)**
A	1	721	2605.207	\$371,915.83
	2	721	2631.523	\$375,830.74
	3	721	2657.838	\$379,745.64
	4	721	2684.153	\$383,660.54
	5	721	2723.626	\$387,575.45
	6	721	2749.941	\$391,490.35
B	7	1060	2421.343	\$345,336.84

	8	1060	2500.300	$356,848.07
	9	1060	2579.257	$368,359.30
	10	557	2708.860	$387,124.77
	11	557	2735.160	$390,149.18
	12	557	2748.310	$393,173.60
C	13	557	2774.609	$396,198.01
	14	557	2787.759	$399,222.42
	15	557	2814.059	$402,246.83
	16	557	2840.358	$405,271.25
	17	1586	2531.154	$361,691.47
	18	1586	2591.702	$370,303.18
D	19	1586	2652.250	$378,914.88
	20	1586	2711.481	$387,526.58
	21	1586	2772.029	$396,138.28

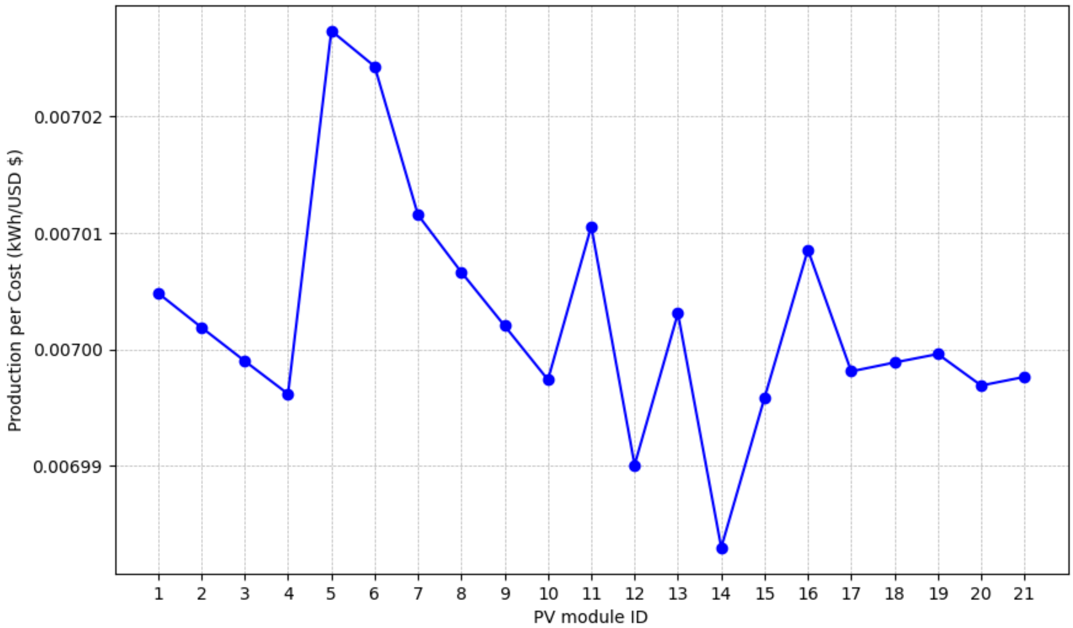
As shown in Figure 26, increased production is associated with increased costs for all brands of PV modules. To increase production capacity, greater financial investments are required. Therefore, it is necessary to find a balance between production and investment.

Figure 26 - Correlation between total cost and total production of solar modules



An index is applied to divide the total energy production (kWh) by the total cost (USD $) to assess the efficiency of PV modules. Thus, Figure 27 shows the efficiency of each model considered in this article. This index reflects the amount of energy generated per monetary unit invested. Among the models presented, Brand A ID 5 has the highest efficiency among the others analyzed.

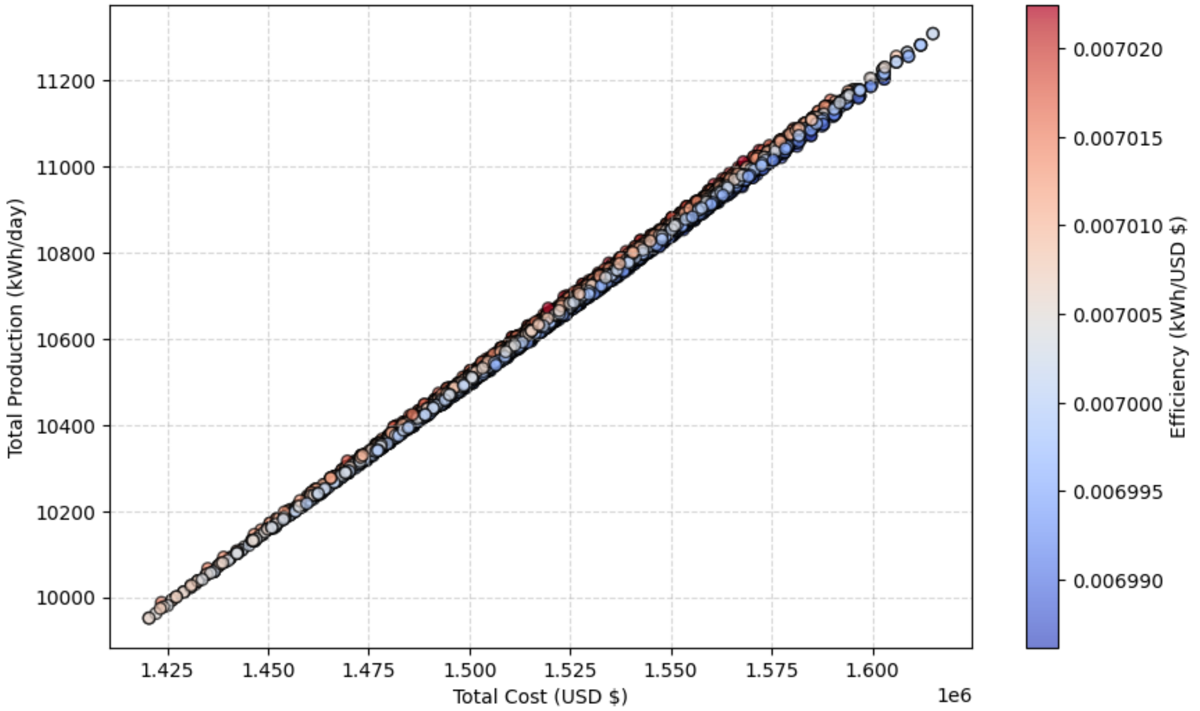
Figure 27 - PV module efficiency



6.3.2 Experiment 1: PV module allocation optimization algorithm

Experiment 1 considers all photovoltaic modules for simulation. The algorithm combines the different models within the restriction of the total roof area. When the maximum number of panels is allocated, the algorithm records the total cost (in USD $) and the total energy production (in kWh/day) for different combinations of PV modules. Thus, Figure 28 shows a configuration of PV modules at each point in the graph, with energy production on the Y axis and the total cost on the X axis. The efficiency of each configuration, which is the ratio between energy production and cost, is represented by a color scale on the right side of the graph, ranging from blue (least efficient) to red (most efficient). The higher the efficiency value, the greater the cost-benefit of the solution analyzed.

Figure 28 - Experiment 1 for selecting the best arrangement of PV modules considering the 21 PV modules



In experiment 1, using only a single PV module model maximizes production and reduces costs. In this case, allocating seven PV modules of Brand A, ID 5 is the highest efficiency. This configuration's total production is 2723.63 kWh/day, and its total cost is USD $387,575.45. The efficiency of this configuration is 0,007027345. The arrangement of the panels in the BIM model is shown in Figure 29.

Figure 29 - Arrangement of solar modules in Experiment 1

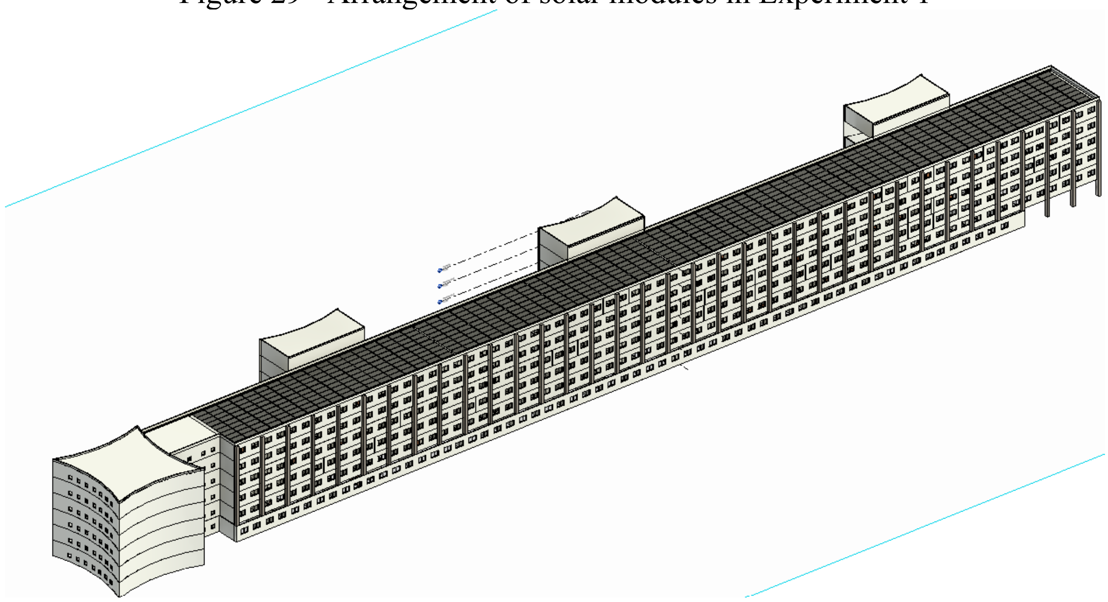
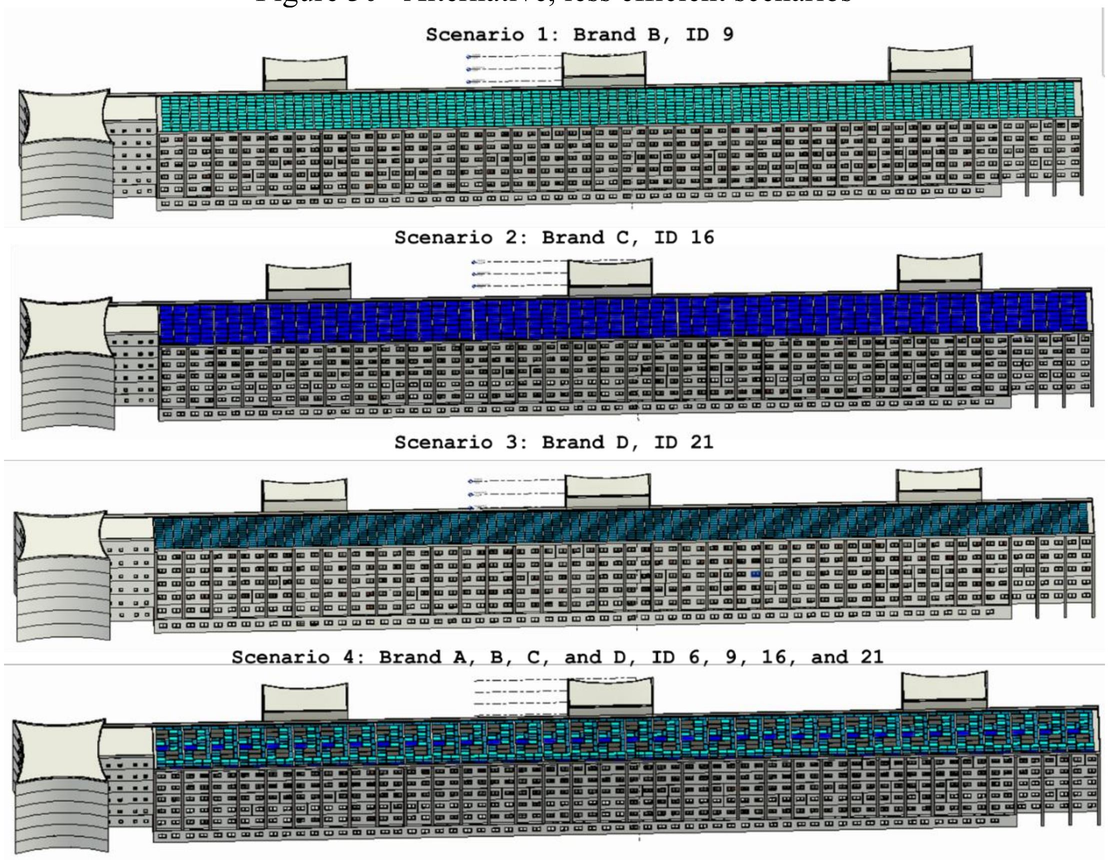


Figure 30 shows other alternative, less efficient scenarios. When different brands are applied in the same plant, as in scenario 4, many empty spaces arise, reducing PV energy production. In addition, this alternative would be unviable in the long term due to high maintenance costs.

Figure 30 - Alternative, less efficient scenarios



6.3.3 Experiment 2: Best allocation considering all four brands

The algorithm implemented in Experiment 1 showed that the best configuration would be one that considers only a single panel across the entire roof area of the building. However, in Experiment 2, an algorithm was implemented that divides the total roof area into four equal parts (i.e., four rectangles measuring 432.88 m² each). Each area has at least one panel of each brand. This algorithm makes it possible to determine efficiency when different brands and models exist in a single project. Thus, Figure 31 shows the relationship between PV energy production, costs, and efficiency.

Figure 31 - Experiment 2 for selecting the best arrangement of solar modules considering the 21 PV modules and requiring at least one panel of each brand

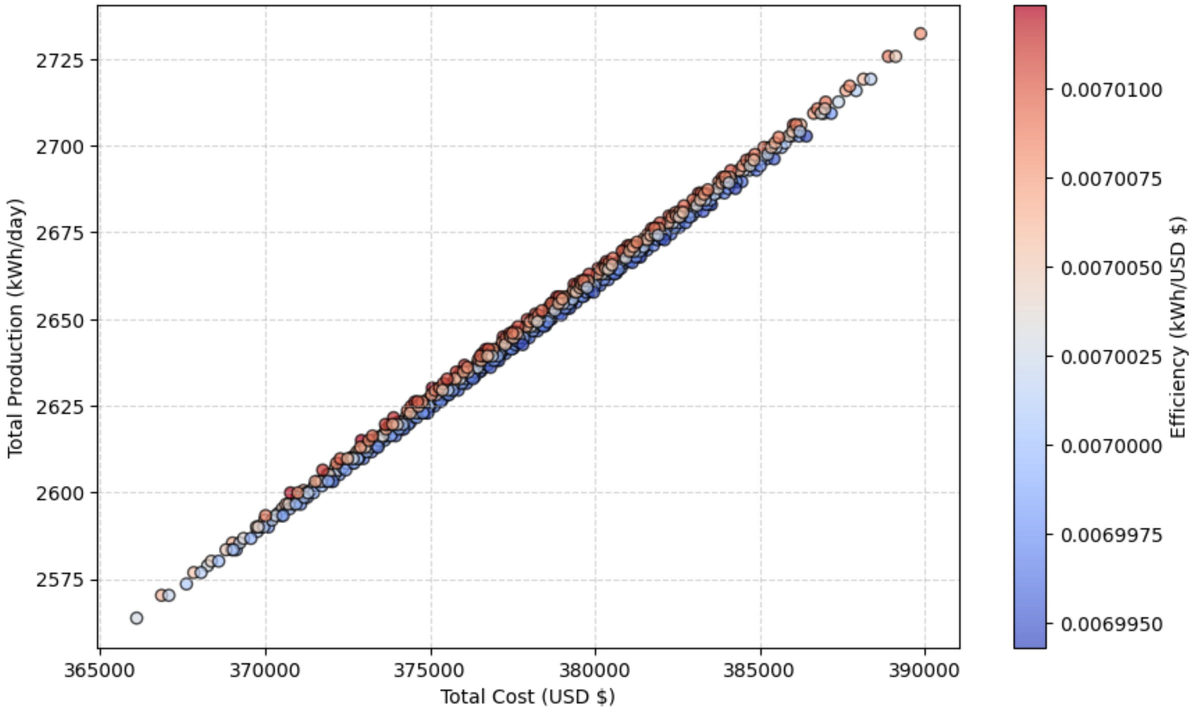
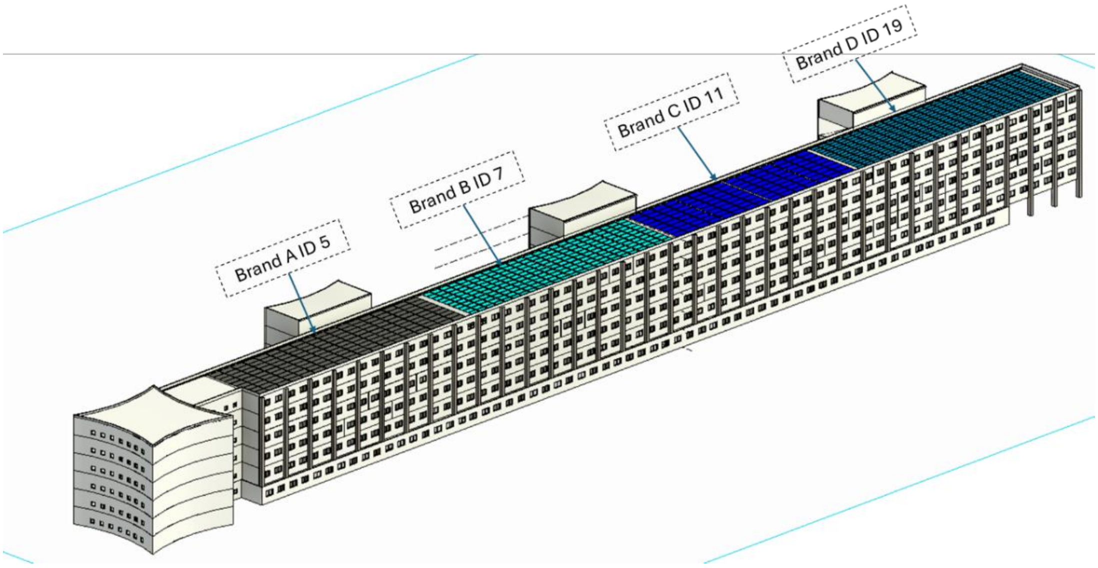


Table 12 shows the information on the best arrangement of the PV modules. In Experiment 2, it was also observed that the best configuration for the layout of the PV modules is when only one model is considered. In the four areas designated for simulation, only one single model per area generates the best cost-benefit ratio. The total number of PV modules is 980, with a total production of 2,630.09 kWh/day and a total cost of USD $ 370,504.08. The efficiency of 0,007012411 kWh/USD $ indicates the amount of energy generated per real spend and reflects the cost-benefit ratio of the configurations. The arrangement of the panels in Experiment 2 is shown in Figure 32.

Table 12 - Allocation of PV brand models

Brand	**ID model**	**Total of PV Modules**	**Total Production (kWh/day)**	**Total Cost (USD $)**
A	5	180	679.962	$ 96,759.47
B	7	265	605.335	$ 86,008.42
C	11	139	682.562	$ 94,560.39
D	19	396	662.226	$ 93,175.79
	Total:	969	2630.09	$ 370,504.08
		Efficiency: 0,007012411 kWh/USD $		

Figure 32 - Arrangement of solar modules in Experiment 2



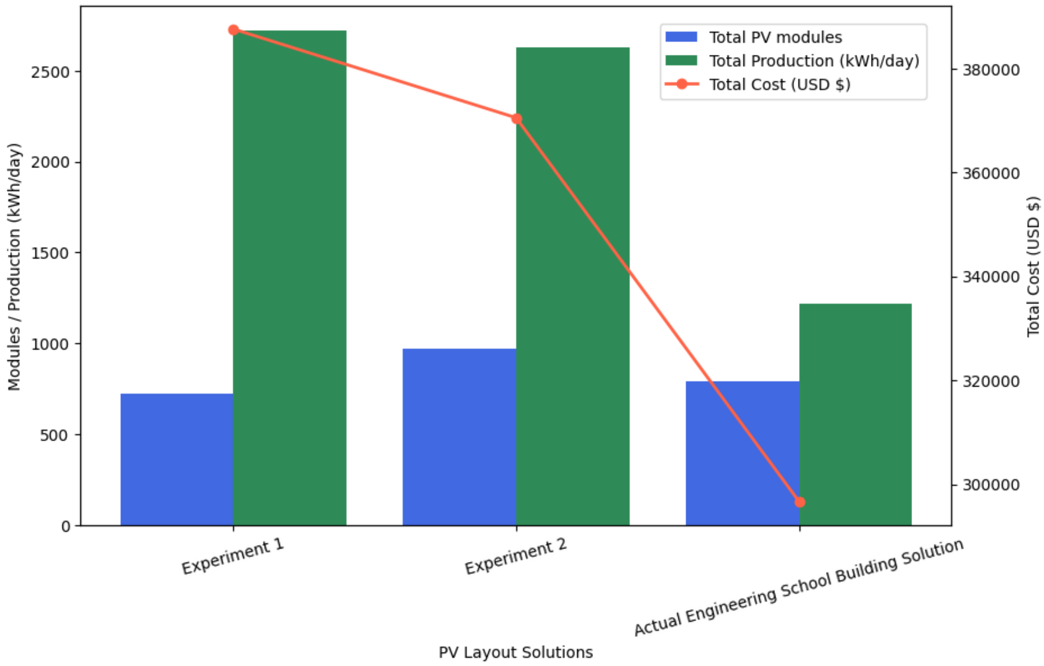
6.3.4 Case study for comparison of PV layout solutions

The results of Experiment 1 and Experiment 2 are compared with the solution currently adopted by UFPE. Figure 33 compares the three photovoltaic layout solutions regarding the total installed panels, total energy production (kWh/day), and total system cost (USD $).

Experiment 1 used 721 PV modules and obtained the highest daily energy production, with 2723.63 kWh/day. However, this solution also presented the highest total cost, reaching USD $ 387,575.45. Experiment 2 was the one that used the greatest number of panels, totaling 969 units. Energy production was lower than in Experiment 1, with 2630.09 kWh/day. On the other hand, this layout presented a lower cost than Experiment 1, totaling USD $ 370,504.08. The solution currently implemented in the School of Engineering building has 792 photovoltaic modules, but its energy production was the lowest among the three options, registering only 1,220.1 kWh/day. On the other hand, this solution presented the lowest total cost, being USD $ 296,729.05.

The solution currently adopted by the School of Engineering is to implement a panel whose production was discontinued in 2020. This model and brand were not considered in the simulations carried out in the study. This means that the Experiments 1 and 2 analyses were based on newer technologies available on the market, possibly more efficient and with better energy performance. In addition, the solutions proposed by the algorithm consider the entire helpful area available on the building's roof.

Figure 33 - Comparison between solutions



6.4 Discussion

This research is a BIM model-based approach to quantify the PV energy production and associated costs in the simulated PV layouts. PV panel manufacturers' manuals were consulted to provide technical information, including characteristics of the photovoltaic modules, such as dimensions and conversion efficiency. This data serves as the basis for calculating the energy production of the PV system. The evaluation process developed from this BIM-PV integration allowed for estimating the energy generation and calculating the costs associated with the PV modules. For end customers, this design process enables creating and evaluating PV layouts to justify the design alternatives and solutions proposed by AECO industry professionals.

With the rapid global expansion of renewable energy, approaches that integrate PV systems with existing architectural structures are needed, taking advantage of underutilized spaces such as rooftops to avoid the need for dedicated land for PV plants. Increasing the power generation capacity of rooftop PV systems requires a focus on optimizing the layout of PV modules (Cao *et al.*, 2024). In this chapter, it is argued that photovoltaic systems can be selected based on cost-efficiency criteria as a viable solution to mitigate energy shortages, reduce environmental impacts, and contribute to carbon neutrality (He *et al.*, 2024).

In the traditional workflow of PV modules in building projects, the project boundaries and parameters are defined, directly impacting their execution. The decisions made by the first actors involved influence the structuring of the project and determine the agreements that will be signed in the subsequent phases. These contracts act as connecting points, guiding the activities of various participants. In addition to the client and, occasionally, the PV specialist consultants, architects, and technical consultants, such as electrical engineers, structural designers, and fire safety experts, are involved in this initial phase (Kathiravel *et al.*, 2024). During the design phase, architects and technical consultants prepare technical specification documents to transfer essential information to the procurement phase. In the procurement phase, the client is responsible for selecting PV system suppliers, electrical contractors, and contractors based on their technical capabilities and the proposals submitted regarding cost and compliance with specifications. To strengthen integration between the agents involved and ensure the continuity of the project, contracts for civil works and installation are signed, and, when necessary, technical support from a PV specialist can be provided (Winkler, 2024; Winkler *et al.*, 2024).

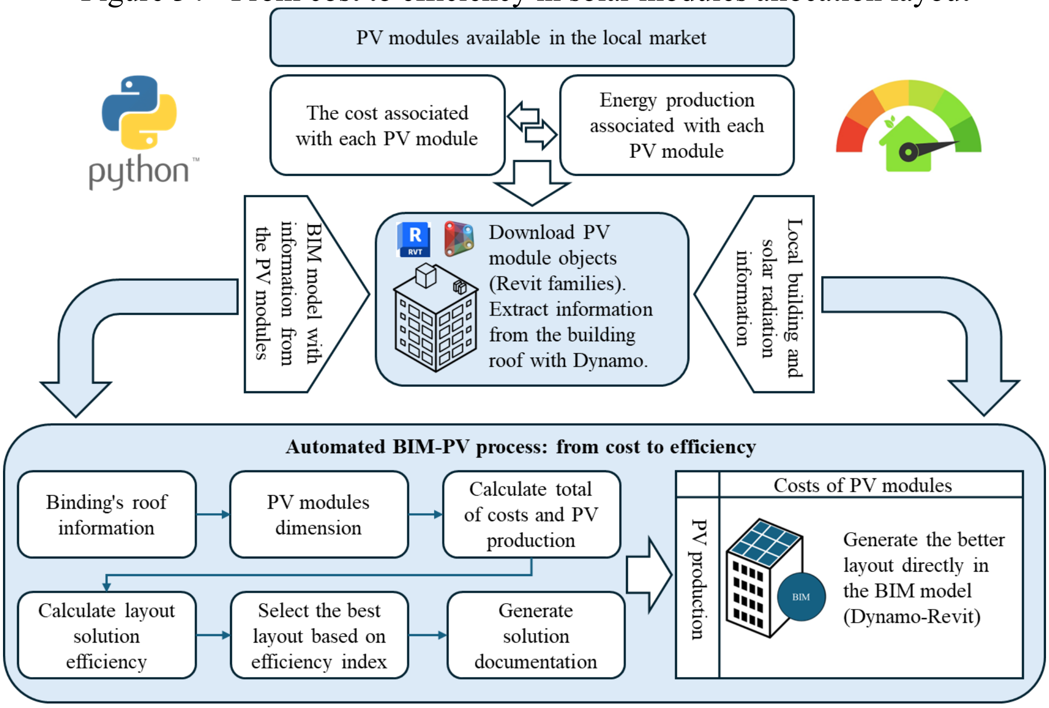
Strategies and approaches to energy efficiency in buildings have been widely debated globally. In general, current efforts focus on two main axes: the direct reduction of energy consumption in buildings and the increased incorporation of renewable energy sources on-site, promoting an indirect reduction in demand (Liu *et al.*, 2023).

However, the workflow proposed in this chapter positions BIM modelers as key players in generating optimized PV systems. BIM can integrate different stakeholders in project development, eliminating segmentation between specialists, clients, and suppliers to reduce conflicts and project nonconformities. For PV system simulations, BIM modelers can look for PV module manuals that contain the necessary information that influences the decision and choice of brands. In this article, the decision process is based on the efficiency of the solution, which is the relationship between the area and the cost of the solution based on the building's roof area (Abouelaziz; Jouane, 2023; Chen *et al.*, 2022; Ning *et al.*, 2018).

Due to the emerging demand for renewable energy sources, evaluating the potential for PV energy generation on urban-scale rooftops is necessary. Through BIM, smart projects can identify the appropriate roof area for solar photovoltaic installation and perform an economic feasibility assessment. The methods should consider the installation size, the costs, and the expected benefits of solar PV modules (Dong; Zhong, 2025). In this context, this paper proposes an automated workflow for simulating layouts for PV modules. Figure 34 highlights the digital element BIM models for centralizing building information. The main information from the BIM models discussed in this article is the area of the PV modules, the area of the building's roof, information on solar radiation, costs of the PV modules, and the PV energy production of the different models to be simulated.

Integrating this information is feasible through the Dynamo plugin in Revit, which applies to the multi-objective algorithm to select the best layout based on the highest efficiency index in the tested solutions. In this context, we highlight the role of integrating stakeholders from AECO industry projects into BIM models for decision-making on implementing sustainable solutions, given the high initial investments associated with the proposed solutions. Specifically, we highlight the role of BIM modelers in controlling and centralizing information and then performing simulations to select PV modules. The goal is to maximize PV energy production while selecting the lowest-cost solution. These goals were achieved through the efficiency index calculated based on the integrated information from the BIM model.

Figure 34 - From cost to efficiency in solar modules allocation layout



Urban surfaces, such as roofs and facades, have a high potential for capturing solar energy. This potential can be used as a subsidy in the pre-design of solar generation systems to define guidelines, planning recommendations, and good practices that promote exploiting this energy source. The advantage of using BIM models in planning PV systems is that the accuracy depends on the spatial information available and generated (Manni *et al.*, 2022). The paper argues that photovoltaic energy is one of the main renewable sources and brings economic, environmental, and social benefits. PV systems installed on rooftops help to produce electricity at the point of consumption, reducing the need for energy from the electricity grid and, consequently, carbon emissions (Wei *et al.*, 2024; Zhang *et al.*, 2024). In addition to the positive environmental impact, photovoltaic technology has great potential to reduce long-term electricity costs and improve distributed generation's economic viability, making energy more affordable and sustainable (Zhang *et al.*, 2023).

From an environmental perspective, adopting PV systems in buildings directly reduces greenhouse gas emissions since the electricity generated is clean and renewable. Furthermore, by decentralizing energy production, PV systems reduce losses in electricity transport along the electrical grid, making energy use more efficient. Furthermore, in regions where access to electricity is limited or unstable, photovoltaic systems offer a solution to guarantee energy supply, even in remote communities (Chong *et al.*, 2024).

It is worth noting that the PV systems market is more present in Europe, the United States, and China, driven by collaboration between government, manufacturers, developers, and end users. To expand this technology, it is essential that governments adopt incentive policies, offer financial support, and promote Research and Development (R&D) projects aligned with local needs. From the developers' perspective, improving research and design methods and investing in innovation are essential for the sector's evolution. For manufacturers, the focus should be on increasing module efficiency and enhancing integration solutions. Finally, awareness and engagement of end users play an essential role in disseminating technology, enabling broader and more efficient adoption of photovoltaic systems (Liu *et al.*, 2021; Poshnath *et al.*, 2023).

Future applications in the BIM-PV topic could include simulations of PV modules to assess energy, thermal performance and comfort, with a focus on mitigating carbon emissions as well as climate change impacts. In the field of BIM research, future research could involve global and regional simulation using climate models. For this, geographic information systems and satellite imagery could be used in conjunction with climate models to analyze and visualize the spatial aspects of solar energy potential, resource availability and system performance. Software such as EnergyPlus and Rhinoceros (Grasshopper) could be applied for energy system modeling, efficiency as well as optimization at the building to city scale. These analyses could extend to carbon emissions, life cycle analysis, benchmarking, probabilistic analysis, statistical learning algorithms, machine learning optimization methods and zero energy design analysis (He *et al.*, 2024).

7 FINAL REMARKS FOR THE THESIS CONCLUSION

This thesis aimed to investigate how the integration of BIM and AI applications can promote a BIM-driven design process to support the planning and renovation of sustainable buildings, with a specific focus on photovoltaic systems in the AECO sector. To this, the results presented achieved three specific objectives.

First, the recent literature on artificial intelligence and BIM in the context of AECO projects was examined. BIM and AI application domains were mapped, and which capabilities needed for the development of smart projects. The combined capabilities of BIM and AI encompass the set of skills and functionalities that a BIM platform offers to enhance the lifecycle of projects in architecture, engineering, and construction. This includes the integrated digital representation provided by BIM, covering detailed geometry, materials, and component data. AI capabilities combine technical skills for manipulating data and implementing advanced AI techniques, and management skills for efficient strategies. For example, applying data mining capabilities drives innovative processes, while predictive algorithms allow you to anticipate results and trends, contributing to a proactive approach to continuous project management and optimization. These combined capabilities are fundamental to creating potential benefits for point cloud processing, design customization, simulations, cost modeling, and scheduling.

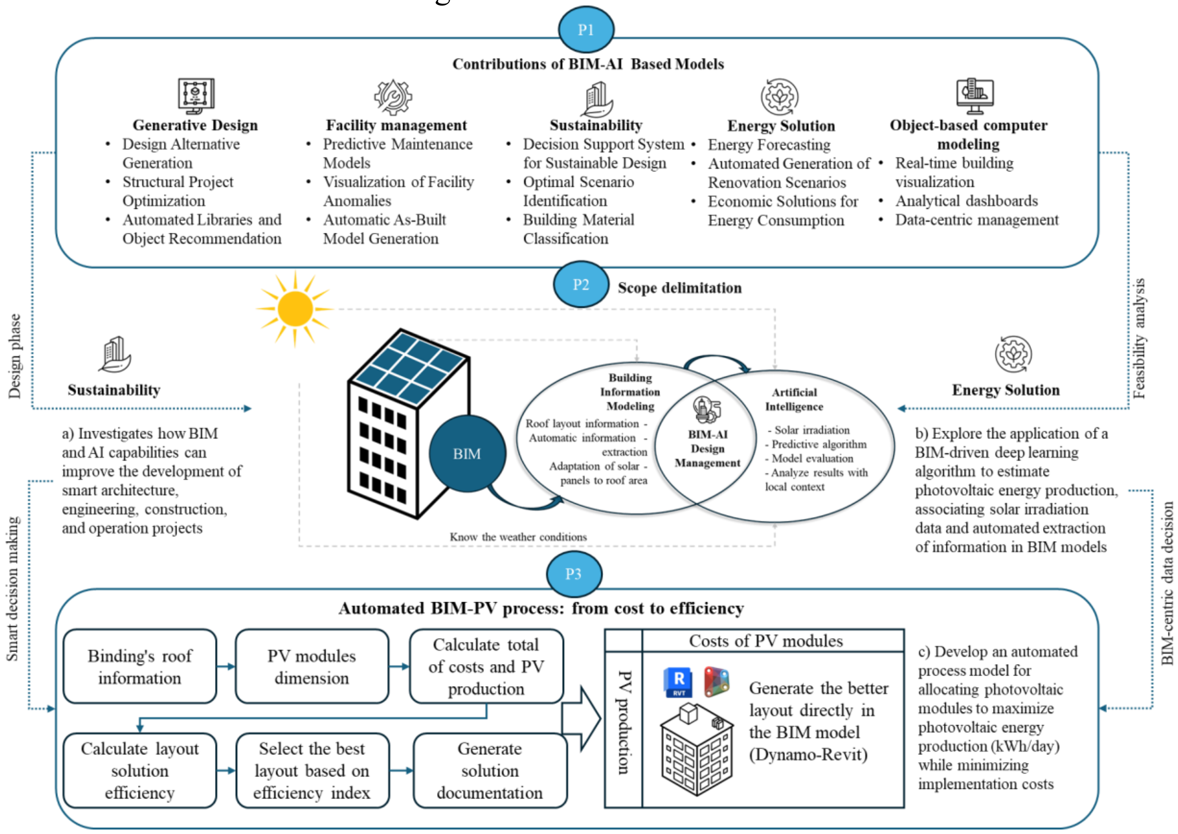
Second, with a specific focus on smart sustainable projects, the thesis develops a BIM-AI approach called SolarisBIM.AI, based on deep learning and BIM data extraction to estimate PV energy production by associating solar radiation data. The results were compared with commercial BIM software (for solar irradiance) and a confirmed case of solar energy production. It was argued that this integrated approach of technologies in the design phase can enhance solar module performance by considering each project's specific characteristics and location. Answering the research question introduced in this article (#RQ3), the results demonstrate the potential of integrating deep learning models with BIM automation tools to support early-stage PV system planning and energy performance assessment. The results indicate that integrating deep learning algorithms into BIM workflows can provide reliable energy production estimates before construction begins. Furthermore, the methodology was validated using a real case study in Pernambuco, Brazil, where the predicted energy generation closely matched the actual output of an operational PV system. The automated process also quantified the corresponding CO_2 emissions avoided, reinforcing the environmental benefits of early PV integration. By embedding energy simulation and forecasting into the BIM environment, SolarisBIM.AI empowers AECO professionals to optimize building design for solar efficiency, assess the feasibility of PV systems during the conceptual stage, and contribute to net-zero energy goals through data-driven, replicable practices.

Third, based on solar energy production prediction data, the design process for planning photovoltaic energy production is improved. The focus is on the design and placement of solar panels in Revit models via Dynamo. The research proposes an efficiency index that is the relationship between energy production and the cost of PV modules. In this context, BIM is the source of data and information about the building, allowing the best configuration of the PV layout given the elements on the building's roof. In this research thesis stage, it is argued that the total costs and efficiency per panel need to be carefully analyzed to ensure the economic and operational viability of the project.

In this way, Figure 35 shows that the results of this thesis highlight the role of BIM in the management and digital representation of building data. It is argued that the AECO industry needs to capture market opportunities and customer requirements as technology develops. However, this future orientation of AECO organizations must be preceded by data management and processing capabilities. Such capabilities are formed through the interdisciplinary skills of data science, engineering, and architecture. The needs of the industry in the face of digital transformation open opportunities for the proactive exploration of data as a strategic mechanism to meet customer needs.

This theorization is explored in Chapter 4, which guides the definition of the scope of the Thesis, as shown in Figure 35. Different research fronts of the AECO industry recognize the promising path of BIM-AI applications, but there is still a barrier to technical skills for developing smart projects. Associated with customers' needs (public or private), the new trends in both the academic community and the practical community are to develop sustainable solutions for buildings. To contribute to the development of sustainable projects, Chapter 5 uses time series to predict solar energy production. In this chapter, an automated process is established to predict solar irradiation in different mesoregions of Pernambuco, and the amount of energy produced and the amount of CO_2 avoided are subsequently calculated. The perspective explored in Chapter 5 is the position of the proactive use of AI for generating predictions, still in the design phase. This highlights the role of the BIM modeler in thinking of sustainable alternatives that are efficiently accommodated to the structure and architecture of the building. Specifically, in the generation of solar energy, the proposed methodology can support feasibility studies, providing documentation that proves the potential of the building to generate renewable energy. These results can also be integrated into return on investment and life cycle analyses. Chapter 6 establishes an initial metric for selecting PV layouts through an automated BIM-driven process. The strategy allows decision-makers to determine which manufacturers and brands best suit their needs. Using the Revit model, the best solution for the building's needs can be planned at the lowest possible cost and with the highest energy production.

Figure 35 - Thesis Overview



These contributions position the thesis in the field of automation of sustainable design processes by integrating BIM modeling with artificial intelligence and computational automation in the context of energy planning for buildings. The novelty lies in the creation of a methodological flow that combines forecasting of solar generation by time series (Chapter 5) with the optimized allocation of photovoltaic modules guided by BIM (Chapter 6). Unlike traditional approaches that treat sustainability as a later stage in the project life cycle, this research anticipates decisions even in the design phase. It is articulated that the BIM modeler assumes a role in the formulation of environmentally responsible solutions. Energy forecasts and feasibility metrics are operationalized within the modeling environment to simulate strategies for how sustainable projects can be designed.

The Sustainable Development Goal (SDG) Agenda for 2030 includes the provision of clean and sustainable energy and the realization of sustainable communities. An advantage of PV energy in buildings is that it produces renewable energy on-site without requiring additional land area. One of BIM-PV's promising strategies is combining energy efficiency with sustainable design, contributing to reducing the carbon footprint and promoting energy self-sufficiency.

The BIM-PV framework can effectively guide the AECO industry toward sustainable outcomes. This integration can help establish indicators to enhance the attainment of sustainable building certifications, such as the Building Research Establishment Environmental Assessment Method (BREEAM), Leadership in Energy and Environmental Design (LEED), GREEN STAR (GS), and the Comprehensive Assessment System for Built Environment Efficiency. BIM-PV framework also enables the planning of renewable energy generation directly in the building’s construction elements.

In the context of projects aimed at energy self-sufficiency, this approach increases the areas available for solar capture, optimizing energy generation and directly contributing to achieving energy neutrality goals, as in the case of net-zero energy buildings (NZEBs). Therefore, the integration between BIM and photovoltaic systems (BIM-PV) guides the energy planning of buildings by adopting sustainable solutions. By incorporating technical, spatial, and economic variables at the design stage, this approach expands the role of BIM as a decision-making support tool. Thus, the research reinforces the potential of BIM-PV as a technological solution for sustainable urban development, with a direct impact on reducing the carbon footprint.

7.1 Hypotheses and research questions

This thesis also answered 4 research questions that support four hypotheses. The research questions (#RQ1) and (#RQ2) were answered by identifying 4 capabilities associated with BIM and 16 capabilities related to AI that are fundamental to promoting innovation, automation, and efficiency in AECO projects. Based on this conceptual basis, three theoretical propositions were proposed that support understanding how BIM and AI can generate value and transform traditional practices in the sector. The evidence indicated that combining these technologies allows for greater analytical capacity, prediction results, and solutions customization. By acting as a digital database, including geometry, materials, schedules, and technical information, BIM provides the ideal environment for applying AI algorithms aimed at pattern extraction, simulations, predictive analysis, and project optimization. This thesis argued that BIM capabilities are associated with the integrated digital representation and management of information throughout the project lifecycle, and AI capabilities involve both technical skills for data manipulation and strategic skills for the application of advanced algorithms and machine learning techniques.

Research question 3 (#RQ3) was answered by developing a SolarisBIM.IA process, which combined time series with deep learning algorithms and parametric data extraction from the BIM model to predict solar energy production and avoided environmental impact. The results were compared with commercial software and real production cases, validating the accuracy of the integrated approach. The main contribution lies in the automation of the photovoltaic system design process, allowing designers to simultaneously consider the spatial constraints of the building and the local energy potential. In addition, by anticipating the quantification of renewable energy and avoiding emissions even in the design phase, the methodology offers technical support for feasibility studies, environmental certifications, and policies to encourage sustainable construction.

To answer research question 4 (#RQ4), a method was developed that begins with selecting solar modules available on the market, incorporating performance and cost data. Based on a database of solar system implementation prices, the final cost to the client was estimated. The energy production of each model was estimated. By automatically extracting information from the building's roof using scripts developed in Dynamo, the algorithm selects the most efficient allocation solution, considering the maximization of energy generation and the minimization of total costs. Finally, the ideal configuration is automatically positioned in the BIM model, respecting the geometric and technical limitations of the roof. The main contribution of this stage of the research lies in the formulation of an efficiency index that relates the cost of the system to the amount of energy generated. Finally, this approach expands the potential for replication of the methodology in different types of buildings and stages of the life cycle, including new and retrofit projects.

7.2 Contributions to literature

The identification and mapping of BIM and AI capabilities guide the functionalities available to optimize the project lifecycle. Furthermore, the analysis of the mapped capabilities and their relationship with the development of smart projects indicates the potential benefits generated by project automation in different phases. Based on the core codes identified and the relationships established in the proposed framework, the thesis suggests that practitioners identify organizational capabilities and invest in building and strengthening them. This involves massive investments in technologies and knowledge focused on data science and engineering, aiming for a multidisciplinary approach oriented towards innovation. Through capability mapping, the three suggested propositions integrate BIM and AI for application on AECO topics that sustain a theoretical framework. The framework suggests an approach for analyzing and processing data from BIM models and their respective contributions to disciplines in the AECO sector, which generate potential benefits.

In the context of 6D BIM, which covers projects' sustainability and environmental impact, the study offers a strategy to estimate renewable energy generation and avoid CO_2. This strategy gives designers, engineers, and managers another way to analyze buildings' energy efficiency and sustainability. This paper offers a design process that can serve as a strategy to predictively quantify the sustainable actions of the project in the design phase. This opens the way for incentive programs, such as specific credit lines for sustainable buildings and tax exemptions for projects that adopt energy forecasting technology, which can encourage the construction industry to adopt more sustainable and technologically advanced practices.

7.3 Management contributions

The design process results established in this thesis highlight the importance of automation in the planning process of PV systems. In addition, the use of solar radiation data and integration with BIM models offers a promising path to improve the sustainability and economic viability of PV projects. Applying time series with Deep Learning models expands the system's ability to adapt to different environmental conditions based on real production analyses. This research also offers theoretical and managerial contributions to the AECO industry. The research establishes a BIM and AI-based design process to support automated data extraction and predictive quantification of solar energy generation, introducing the implementation of deep learning techniques applied to solar radiation time series. The practical application of the model to estimate energy produced and CO_2 avoided highlights the role BIM and deep learning can play in sustainability through renewable energy. This study provides an example of the literature on how digital technologies can be applied to promote sustainable projects in the AECO industry.

Also, it offers a BIM-oriented approach to PV system projects applied to buildings by presenting a comparative analysis between different layout solutions and the automated allocation of PV modules. The results reveal that, although solutions with a larger number of panels can generate more energy. In the professional field, this research directly contributes to improving the practices of engineers, designers, and consultants who develop and implement photovoltaic systems. The method developed in this article can be applied to projects still in the design phase and building retrofit projects to generate documents that can serve as a technical reference for dimensioning and selecting equipment in new projects. In addition, the methodology applied in comparative analysis can be replicated or adapted in other contexts, building typologies and expanding the practical applicability of the study.

7.4 Limitations and future directions for research development

Has some limitations that may provide opportunities for future research. The sample articles and databases used in this literature review were selected based on specific criteria and search strings. The article was developed based on a thorough structuring of the theoretical background; it is possible that some studies were not included in the sample. Furthermore, the sample was restricted until 2023, which may limit the timeliness of the information presented in the future. The search strategies adopted, such as selected databases, search filters, and exclusion criteria, can restrict the sample of articles. Subjective qualitative analysis is another limitation that may affect the results. These limitations highlight the importance of cautiously interpreting results and indicate areas requiring further investigation. Therefore, future research can apply the findings of this research through other methodological approaches. Case studies can drive an in-depth understanding of how BIM and AI capabilities are used in the organizational context to generate innovation. Research can advance the knowledge of which contextual phenomena, such as stakeholders, market, and environment, influence the development of BIM and AI capabilities. Furthermore, quantitative research can apply the variables mapped in research to quantify their correlation and test the proposed framework, mathematically validating the suggested propositions.

Although the solar radiation data used to forecast PV energy production were based on time series, the analyses of meteorological data are conditioned by the availability of meteorological monitoring stations. Despite the advantages of deep learning for forecasting PV energy production, the deep learning model may be sensitive to the training data. The scarcity of data may result in a model with suboptimal performance, limiting the generalization of predictions for different design conditions. Future research could address the automation of PV plant planning on existing building roofs using 3D point cloud techniques from aerial LiDAR technology or aerial orthoimagery. There is a fertile path in analyzing AI algorithms for point cloud segmentation and application in BIM, specifically for the optimized generation of PV plants. New research could also propose an analysis for optimizing PV plant layouts, addressing installation costs, module positioning, available area, maintenance, and return on investment of PV systems.

The analysis does not consider the entire life cycle of PV modules and the costs associated with maintaining the solutions. In addition, the study used estimated cost information based on current market values without considering possible regional variations, maintenance costs throughout the life cycle, or tax incentives that could alter the economic viability of the proposed solutions. In this context, new studies can advance the analysis of the life cycle of components, aspects related to maintenance, and the evaluation of suppliers and manufacturers that guarantee technical support throughout the system's useful life. To this end, longitudinal databases will be necessary to evaluate these parameters based on solutions already implemented and compare them with those planned in the design phase of the projects.

Future research could explore integrating LiDAR (Light Detection and Ranging) data to map surfaces and structures where PV modules will be installed. This technology allows for detailed analysis of the building's geometry, considering factors such as inclination, shading, and obstructions, which can impact the energy efficiency of PV systems. Future studies could investigate how point cloud information generated by LiDAR sensors can be directly integrated into BIM models and energy simulation software, optimizing the positioning and orientation of solar modules.

Future research could explore increasing the complexity of the algorithms used for photovoltaic module allocation on building rooftops. One promising direction is to integrate information about the natural slope of the building roof (e.g., rainwater drainage). Such data could be used to assess whether modules should be placed following the original slope or whether alternative inclinations, supported by stainless steel structures, could provide higher efficiency. In buildings without natural slopes (i.e., flat roofs), the direction and inclination of the modules could be optimized to maximize energy generation for the selected module type.

Another avenue for future research is to evaluate the influence of elevated objects on rooftop PV performance, such as water tanks, lightning rods, and antennas. By calculating the shadow profile cast by these objects over the course of the year, it would be possible to test different strategies to mitigate shading effects, for instance, repositioning the object, adjusting the orientation of PV modules, or redesigning their layout.

Another area for future research is the application of aerial photogrammetry with drones to capture spatial data of buildings. Photogrammetry can generate high-resolution 3D models and orthomosaics for topographic analysis and planning of large-scale solar projects. In addition, future research could deepen the use of these technologies in retrofit strategies, combining them with real-time energy performance analysis and smart monitoring systems.

Finally, future research may incorporate life cycle assessments (LCA) as a strategic tool for risk management in implementing PV module systems. The application of LCA allows the evaluation of environmental, economic, and social impacts throughout all stages of the photovoltaic system, from the extraction of raw materials, manufacturing of modules, transportation, installation, and operation to the decommissioning and recycling of components at the end of their useful life. For example, by considering LCA, it is possible to identify risks related to the carbon footprint of the materials used, maintenance costs over time, or even the challenges associated with the disposal and reuse of components such as crystalline silicon and heavy metals in some modules.

Future research can also apply the capabilities identified in this thesis to establish indicators for organizations' transition to digital transformation. Research methods such as Structural Equation Modeling can be applied to quantify the correlation between these capabilities and how they are interrelated in a theoretical model. Case studies can also be developed to understand how companies have overcome the interdisciplinary barrier of the topic since BIM-AI integration is reflected in areas such as building construction and computer science.

The findings of this thesis mainly focus on project planning in BIM models. Future research can focus on understanding how to deal with the operational barriers of sustainable projects in terms of maintenance and operation. In the field of photovoltaics, this includes investigating the actual performance of systems over time, smart monitoring strategies, predictive maintenance, and integration with BIM-based building management systems (BMS). Furthermore, there is room to explore how real-time data analysis collected by IoT sensors can be incorporated into BIM models to feed back into design decisions and optimize energy performance throughout the building's life cycle.

A promising development of this research is the incorporation of LCOE (Levelized Cost of Energy) as an evaluation index for PV projects integrated into the BIM environment. LCOE is widely used in the energy sector because it represents the total cost of energy generation over the system's useful life, considering both initial installation costs (CAPEX) and operational and maintenance costs (OPEX). Adopting this metric could improve the economic feasibility analysis of simulated layouts.

In addition, a factor that has not yet been explored but is of great practical relevance is the need for spacing between photovoltaic modules to allow access for cleaning and maintenance. This variable directly impacts the coverage density and the usable area available for installation. In this scenario, using higher-power modules may be more advantageous, as it reduces the total number of panels and, consequently, the area sacrificed for the circulation of the technical team.

7.5 Research Development

The doctoral program at UFPE for developing this thesis was oriented toward the themes of innovation, technology, and sustainability. This was reflected in the subjects studied, such as BIM, Python, Machine Learning, Sustainability, Project Management, and Applied Statistics. The highest grade was achieved in all these subjects, as well as in the qualification exam. A one-year teaching internship was also carried out on the graduate subject of Tecnologia da Construção Civil I.

During de research activities, it participated in scientific events with the presentation of articles, such as Simpósio Brasileiro de Gestão e Economia da Construção (SIBRAGEC) and the Simpósio Brasileiro de Tecnologia da Informação e Comunicação na Construção (SBTIC) in 2023 and Encontro Nacional de Tecnologia do Ambiente Construído (ENTAC) in 2024.

The chapters of this thesis generated three articles. The first was published in Automation in Construction, impact factor 9.6, qualis A1 (see Alves *et al., 2025*). The second was submitted to the Journal of Construction Engineering and Management, impact factor 5.1, qualis A1. The third was submitted to Building and Environment, impact factor 7.1, qualis A1. A fourth article published in a journal with qualis A2, reflecting theoretical research on the integration of PV and BIM systems (see “BIM-Based Framework for Photovoltaic Systems: Advancing Technologies, Overcoming Challenges, and Enhancing Sustainable Building Performance”).

In addition to these articles, partnerships were established with the research group on Decision Support in BIM–ADBIM. This resulted in two other articles submitted to international journals.

REFERENCES

ABDIRAD, Hamid; MATHUR, Pegah. Artificial intelligence for BIM content management and delivery: Case study of association rule mining for construction detailing. **Advanced Engineering Informatics**, *[S. l.]*, v. 50, n. August, p. 101414, 2021. DOI: 10.1016/j.aei.2021.101414. Disponível em: https://doi.org/10.1016/j.aei.2021.101414.

ABDULFATTAH, Basem S.; ABDELSALAM, Hassan A.; ABDELSALAM, Mai; BOLPAGNI, Marzia; THURAIRAJAH, Niraj; PEREZ, Laura Florez; BUTT, Talib E. Predicting implications of design changes in BIM-based construction projects through machine learning. **Automation in Construction**, *[S. l.]*, v. 155, n. February, p. 105057, 2023. DOI: 10.1016/j.autcon.2023.105057.

ABOUELAZIZ, Ilyass; JOUANE, Youssef. Photogrammetry and deep learning for energy production prediction and building-integrated photovoltaics decarbonization. **Building Simulation**, *[S. l.]*, p. 189–205, 2023. DOI: 10.1007/s12273-023-1089-y.

ABSOLAR, Associação Brasileira de Energia Solar Fotovoltaica. **Panorama da solar fotovoltaica no Brasil e no mundo**. 2025. Disponível em: https://www.absolar.org.br/

AKSOY TIRMIKÇI, Ceyda; YAVUZ, Cenk; ÖZKURT, Cem; ÇARKLI YAVUZ, Burcu. Machine learning-assisted evaluation of PVSOL software using a real-time rooftop PV system: a case study in Kocaeli, Turkey, with a focus on diffuse solar radiation. **International Journal of Low-Carbon Technologies**, *[S. l.]*, v. 20, p. 223–233, 2025. DOI: 10.1093/ijlct/ctae292.

ALAWI, Omer A.; KAMAR, Haslinda Mohamed; YASEEN, Zaher Mundher. Optimizing building energy performance predictions: A comparative study of artificial intelligence models. **Journal of Building Engineering**, *[S. l.]*, v. 88, p. 109247, 2024. DOI: 10.1016/j.jobe.2024.109247.

ALMEIDA-FILHO, Adiel Teixeira De; DE LIMA SILVA, Diogo Ferreira; FERREIRA, Luciano. Financial modelling with multiple criteria decision making: A systematic literature review. **Journal of the Operational Research Society**, *[S. l.]*, v. 72, n. 10, p. 2161–2179, 2021. DOI: 10.1080/01605682.2020.1772021.

ALVES, Gabriela Tenório de Morais; NASCIMENTO, Claudia R. S. de M. S.; SANTOS, Eduardo B. Dos; SOUZA, Kleyton M. N. De; FERNANDES, Bruna S.; PALHA, Rachel Perez. Integration potential between REVIT and LEED: a review. **Architectural Engineering and Design Management**, *[S. l.]*, v. 20, n. 3, p. 510–525, 2024. DOI: 10.1080/17452007.2023.2259387.

ALVES, Josivan Leite; DE CARVALHO, Marly Monteiro. Bridging Knowledge Management and Capabilities in Innovative Projects: An Integrative Framework. **Project Management Journal**, *[S. l.]*, 2023. DOI: 10.1177/87569728231217493.

ALVES, Josivan Leite; PALHA, Rachel Perez; ALMEIDA FILHO, Adiel Teixeira De. Towards an integrative framework for BIM and artificial intelligence capabilities in smart architecture, engineering, construction, and operations projects. **Automation in**

Construction, *[S. l.]*, v. 174, n. 106168, p. 1–22, 2025. DOI: 10.1016/j.autcon.2025.106168. Disponível em: https://linkinghub.elsevier.com/retrieve/pii/S0926580525002080.

ALZARA, Majed; ATTIA, Yehia Abdelhamid; MAHFOUZ, Sameh Youssef; YOSRI, Ahmed M.; EHAB, Ahmed. Building a Genetic Algorithm-Based and BIM-Based 5D Time and Cost Optimization Model. **IEEE Access**, *[S. l.]*, v. 11, n. September, p. 122502–122515, 2023. a. DOI: 10.1109/ACCESS.2023.3317137.

ARAÚJO, Adolpho Guido; PEREIRA CARNEIRO, Arnaldo Manoel; PALHA, Rachel Perez. Sustainable construction management: A systematic review of the literature with meta-analysis. **Journal of Cleaner Production**, *[S. l.]*, v. 256, p. 120350, 2020. b. DOI: 10.1016/j.jclepro.2020.120350. Disponível em: https://linkinghub.elsevier.com/retrieve/pii/S0959652620303978.

ARAUJO, Cleber Marchi Bernardo De; ALVES, Josivan Leite. How to unlock BIM capabilities in the design phase to project success for long-term organization development? **Engineering, Construction and Architectural Management**, *[S. l.]*, 2025. DOI: 10.1108/ECAM-08-2024-1038. Disponível em: https://www.emerald.com/insight/content/doi/10.1108/ECAM-08-2024-1038/full/html.

ARIA, Massimo; CUCCURULLO, Corrado. bibliometrix : An R-tool for comprehensive science mapping analysis. **Journal of Informetrics**, *[S. l.]*, v. 11, n. 4, p. 959–975, 2017. DOI: 10.1016/j.joi.2017.08.007.

ARSIWALA, Arva; ELGHAISH, Faris; ZOHER, Mohammed. Digital twin with Machine learning for predictive monitoring of CO2 equivalent from existing buildings. **Energy and Buildings**, *[S. l.]*, v. 284, p. 112851, 2023. DOI: 10.1016/j.enbuild.2023.112851. Disponível em: https://doi.org/10.1016/j.enbuild.2023.112851.

ASIF, Muhammad; NAEEM, Ghinwa; KHALID, Muhammad. Digitalization for sustainable buildings: Technologies, applications, potential, and challenges. **Journal of Cleaner Production**, *[S. l.]*, v. 450, p. 141814, 2024. DOI: 10.1016/j.jclepro.2024.141814.

AZHAR, Salman. Building Information Modeling (BIM): Trends, Benefits, Risks, and Challenges for the AEC Industry. **Leadership and Management in Engineering**, *[S. l.]*, v. 11, n. 3, p. 241–252, 2011. DOI: 10.1061/(ASCE)LM.1943-5630.0000127.

BADUGE, Shanaka Kristombu; THILAKARATHNA, Sadeep; PERERA, Jude Shalitha; ARASHPOUR, Mehrdad; SHARAFI, Pejman; TEODOSIO, Bertrand; SHRINGI, Amkit; MENDIS, Priyan. Artificial intelligence and smart vision for building and construction 4.0: Machine and deep learning methods and applications. **Automation in Construction**, *[S. l.]*, v. 141, n. June, p. 104440, 2022. DOI: 10.1016/j.autcon.2022.104440. Disponível em: https://doi.org/10.1016/j.autcon.2022.104440.

BARBÓN, A.; GHODBANE, M.; BAYÓN, L.; SAID, Z. A general algorithm for the optimization of photovoltaic modules layout on irregular rooftop shapes. **Journal of Cleaner Production**, *[S. l.]*, v. 365, p. 132774, 2022. DOI: 10.1016/j.jclepro.2022.132774.

BARBOSA, Eduardo Jose; CAVALCANTI, Marcelo Cabral; AZEVEDO, Gustavo Medeiros Souza; NETO, Rafael Cavalcanti; BARBOSA, Eduardo Augusto Oliveira; BRADASCHIA, Fabricio. Hybrid GMPPT Technique for Photovoltaic Series Based on Fractional Characteristic Curve. **IEEE Journal of Photovoltaics**, *[S. l.]*, v. 14, n. 1, p. 170–177, 2024. DOI: 10.1109/JPHOTOV.2023.3323774.

BLOCH, Tanya; SACKS, Rafael. Clustering Information Types for Semantic Enrichment of Building Information Models to Support Automated Code Compliance Checking. **Journal of Computing in Civil Engineering**, *[S. l.]*, v. 34, n. 6, p. 1–11, 2020. DOI: 10.1061/(asce)cp.1943-5487.0000922.

BOJE, Calin; GUERRIERO, Annie; KUBICKI, Sylvain; REZGUI, Yacine. Towards a semantic Construction Digital Twin: Directions for future research. **Automation in Construction**, *[S. l.]*, v. 114, n. January, p. 103179, 2020. DOI: 10.1016/j.autcon.2020.103179. Disponível em: https://doi.org/10.1016/j.autcon.2020.103179.

BRAHMA, Banalaxmi; WADHVANI, Rajesh. Solar Irradiance Forecasting Based on Deep Learning Methodologies and Multi-Site Data. **Symmetry**, *[S. l.]*, v. 12, n. 11, p. 1830, 2020. DOI: 10.3390/sym12111830.

BRAZIL. Building Information Modelling na execução direta ou indireta de obras e serviços de engenharia realizada pelos órgãos e pelas entidades da administração pública federal. Estabelece a utilização do Building Information Modelling na execução direta ou indireta de obras e serviços de engenharia realizada pelos órgãos e pelas entidades da administração pública federal, no âmbito da Estratégia Nacional de Disseminação do Building Information Modelling - Estratégia BIM BR, instituída pelo Decreto nº 9.983, de 22 de agosto de 2019. **https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/decreto/D10306.htm**, Brazil, 2 abr. 2020. p. 1–2.

CAO, Haiyun; HUANG, Minghao. Building Information Modeling Technology Capabilities: Operationalizing the Multidimensional Construct. **Sustainability**, *[S. l.]*, v. 15, n. 20, p. 14755, 2023. DOI: 10.3390/su152014755.

CAO, Zhixiang; LIU, Yangshaohua; BAI, Yuqing; WANG, Yi; YE, Shengjun; CAO, Haibin. Study on the optimal layout of roof vents and rooftop photovoltaic of the industrial workshop. **Building and Environment**, *[S. l.]*, v. 260, p. 111624, 2024. DOI: 10.1016/j.buildenv.2024.111624.

CARREIRA, Paulo; CASTELO, Tiago; GOMES, Cristina Caramelo; FERREIRA, Alfredo; RIBEIRO, Cláudia; COSTA, Antonio Aguiar. Virtual reality as integration environments for facilities management: Application and users perception. **Engineering, Construction and Architectural Management**, *[S. l.]*, v. 25, n. 1, p. 90–112, 2018. DOI: 10.1108/ECAM-09-2016-0198.

CASSANDRO, Jacopo; MIRARCHI, Claudio; GHOLAMZADEHMIR, Maryam; PAVAN, Alberto. Advancements and prospects in building information modeling (BIM) for construction: a review. **Engineering, Construction and Architectural Management**, *[S. l.]*, 2024. DOI: 10.1108/ECAM-04-2024-0435.

CATERINO, Nicola; NUZZO, Iolanda; IANNIELLO, Antonio; VARCHETTA, Giorgio; COSENZA, Edoardo. A BIM-based decision-making framework for optimal seismic retrofit of existing buildings. **Engineering Structures**, *[S. l.]*, v. 242, n. May, p. 112544, 2021. DOI: 10.1016/j.engstruct.2021.112544.

ÇETIN, Sultan; GRUIS, Vincent; STRAUB, Ad. Digitalization for a circular economy in the building industry: Multiple-case study of Dutch social housing organizations. **Resources, Conservation and Recycling Advances**, *[S. l.]*, v. 15, n. August, p. 200110, 2022. DOI: 10.1016/j.rcradv.2022.200110.

CHANGSAAR, Chai; ABIDIN, Nur IzieAdiana; KHOSO, Ali Raza; LUENHUI, Ling; YAOLI, Xiong; HUNCHUEN, Gui. Optimising energy performance of an Eco-Home using Building Information Modelling (BIM). **Innovative Infrastructure Solutions**, *[S. l.]*, v. 7, n. 2, p. 140, 2022. DOI: 10.1007/s41062-022-00747-6.

CHEN, Siwei; GOU, Zhonghua. City-roof coupling: Unveiling the spatial configuration and correlations of green roofs and solar roofs in 26 global cities. **Cities**, *[S. l.]*, v. 147, p. 104780, 2024. DOI: 10.1016/j.cities.2023.104780.

CHEN, Tianyi; SUN, Huixuan; TAI, Kong Fai; HENG, Chye Kiang. Analysis of the barriers to implementing building integrated photovoltaics in Singapore using an interpretive structural modelling approach. **Journal of Cleaner Production**, *[S. l.]*, v. 365, p. 132652, 2022. DOI: 10.1016/j.jclepro.2022.132652.

CHEN, Xichen; CHANG-RICHARDS, Alice; LING, Florence Yean Yng; YIU, Tak Wing; PELOSI, Antony; YANG, Nan. Digital technologies in the AEC sector: a comparative study of digital competence among industry practitioners. **International Journal of Construction Management**, *[S. l.]*, v. 0, n. 0, p. 1–14, 2023. a. DOI: 10.1080/15623599.2024.2304453.

CHEN, Zhengshu; CUI, Yanqiu; SONG, Dexuan; ZHENG, Haichao; DING, Xin; YANG, Haoran. Data-driven Approach of Academic Building-integrated Photovoltaic System Based on Carbon Emission, Energy Payback Time and Comfort: Considering Climate Change. **Building and Environment**, *[S. l.]*, v. 270, p. 112489, 2025. DOI: 10.1016/j.buildenv.2024.112489.

CHEN, Zhengyi; LAI, Zhijie; SONG, Changhao; ZHANG, Xiao; CHENG, Jack C. P. Smart camera placement for building surveillance using OpenBIM and an efficient Bi-level optimization approach. **Journal of Building Engineering**, *[S. l.]*, v. 77, n. June, p. 107257, 2023. b. DOI: 10.1016/j.jobe.2023.107257.

CHENG, Jack C. P.; CHEN, Weiwei; CHEN, Keyu; WANG, Qian. Data-driven predictive maintenance planning framework for MEP components based on BIM and IoT using machine learning algorithms. **Automation in Construction**, *[S. l.]*, v. 112, n. August 2018, p. 103087, 2020. DOI: 10.1016/j.autcon.2020.103087. Disponível em: https://doi.org/10.1016/j.autcon.2020.103087.

CHO, Junhwi; KIM, Chaehyeon; SONG, Yooseob; KANG, Julian; YEON, Jaeheum. Lumped record management method using BIM and dynamo for spalling maintenance. **Automation in Construction**, *[S. l.]*, v. 160, p. 105324, 2024. DOI: 10.1016/j.autcon.2024.105324.

CHONG, Shijia; YOU, Jialin; WU, Jing; CHANG, I. Shin. Booming solar energy drives land value enhancement: Evidence from 648 photovoltaic projects in China. **Journal of Cleaner Production**, *[S. l.]*, v. 484, p. 144270, 2024. DOI: 10.1016/j.jclepro.2024.144270.

CROCE, Valeria; CAROTI, Gabriella; PIEMONTE, Andrea; DE LUCA, Livio; VÉRON, Philippe. H-BIM and Artificial Intelligence: Classification of Architectural Heritage for Semi-Automatic Scan-to-BIM Reconstruction. **Sensors**, *[S. l.]*, v. 23, n. 5, p. 2497, 2023. DOI: 10.3390/s23052497. Disponível em: https://www.mdpi.com/1424-8220/23/5/2497.

D'ADAMO, Idiano; GASTALDI, Massimo; KOH, S. C. Lenny; VIGIANO, Alessandro. Lighting the future of sustainable cities with energy communities: An economic analysis for incentive policy. **Cities**, *[S. l.]*, v. 147, p. 104828, 2024. DOI: 10.1016/j.cities.2024.104828.

DARKO, Amos; CHAN, Albert P. C.; ADABRE, Michael A.; EDWARDS, David J.; HOSSEINI, M. Reza; AMEYAW, Ernest E. Artificial intelligence in the AEC industry: Scientometric analysis and visualization of research activities. **Automation in Construction**, *[S. l.]*, v. 112, p. 103081, 2020. DOI: 10.1016/j.autcon.2020.103081.

DI GIOVANNI, Gianni; ROTILIO, Marianna; GIUSTI, Letizia; EHTSHAM, Muhammad. Exploiting building information modeling and machine learning for optimizing rooftop photovoltaic systems. **Energy and Buildings**, *[S. l.]*, v. 313, p. 114250, 2024. DOI: 10.1016/j.enbuild.2024.114250.

DONG, Cunzhuang; ZHONG, Qing. Evaluating solar photovoltaic potential of buildings based on the installation parameters of photovoltaic modules. **Solar Energy**, *[S. l.]*, v. 288, p. 113304, 2025. DOI: 10.1016/j.solener.2025.113304.

ERIŞEN, Serdar. A Systematic Approach to Optimizing Energy-Efficient Automated Systems with Learning Models for Thermal Comfort Control in Indoor Spaces. **Buildings**, *[S. l.]*, v. 13, n. 7, 2023. DOI: 10.3390/buildings13071824.

FENZ, Stefan; GIANNAKIS, Giorgos; BERGMAYR, Julia; IOUSEF, Samy. RenoDSS – A BIM-based building renovation decision support system. **Energy and Buildings**, *[S. l.]*, v. 288, p. 112999, 2023. DOI: 10.1016/j.enbuild.2023.112999.

FRÍAS, Ernesto; PINTO, José; SOUSA, Ricardo; LORENZO, Henrique; DÍAZ-VILARIÑO, Lucía. Exploiting BIM Objects for Synthetic Data Generation toward Indoor Point Cloud Classification Using Deep Learning. **Journal of Computing in Civil Engineering**, *[S. l.]*, v. 36, n. 6, 2022. DOI: 10.1061/(ASCE)CP.1943-5487.0001039. Disponível em: https://ascelibrary.org/doi/10.1061/%28ASCE%29CP.1943-5487.0001039.

GAN, Vincent J. L. BIM-Based Building Geometric Modeling and Automatic Generative Design for Sustainable Offsite Construction. **Journal of Construction Engineering and Management**, *[S. l.]*, v. 148, n. 10, 2022. DOI: 10.1061/(ASCE)CO.1943-7862.0002369.

GAO, Hao; KOCH, Christian; WU, Yupeng. Building information modelling based building energy modelling: A review. **Applied Energy**, *[S. l.]*, v. 238, p. 320–343, 2019. DOI: 10.1016/j.apenergy.2019.01.032.

GARCIA-GAGO, Jesús; SÁNCHEZ-APARICIO, Luis Javier; SOILÁN, Mario; GONZÁLEZ-AGUILERA, Diego. HBIM for supporting the diagnosis of historical buildings: case study of the Master Gate of San Francisco in Portugal. **Automation in Construction**, *[S. l.]*, v. 141, n. April, 2022. DOI: 10.1016/j.autcon.2022.104453.

GREENER. **Strategic Study: Distributed Generation | 2023 Data | Release Feb. 2024**. 2024. https://dg-report-2025.greener.com.br/

HE, Rui; LI, Mingkai; GAN, Vincent J. L.; MA, Jun. BIM-enabled computerized design and digital fabrication of industrialized buildings: A case study. **Journal of Cleaner Production**, *[S. l.]*, v. 278, p. 123505, 2021. DOI: 10.1016/j.jclepro.2020.123505.

HE, Yueer; HII, Daniel Jun Chung; WONG, Nyuk Hien. Solar photovoltaics deployment impact on urban temperature: Review and assessment recommendations. **Building and Environment**, *[S. l.]*, v. 264, p. 111920, 2024. DOI: 10.1016/j.buildenv.2024.111920.

HEIDARI, Aliakbar; PEYVASTEHGAR, Yaghowb; AMANZADEGAN, Mohammad. A systematic review of the BIM in construction: from smart building management to interoperability of BIM & AI. **Architectural Science Review**, *[S. l.]*, 2023. DOI: 10.1080/00038628.2023.2243247.

HONG, Ying; XIE, Haiyan; BHUMBRA, Gary; BRILAKIS, Ioannis. Comparing Natural Language Processing Methods to Cluster Construction Schedules. **Journal of Construction Engineering and Management**, *[S. l.]*, v. 147, n. 10, p. 1–11, 2021. DOI: 10.1061/(asce)co.1943-7862.0002165.

HOU, Fangli; MA, Jun; KWOK, Helen H. L.; CHENG, Jack C. P. Prediction and optimization of thermal comfort, IAQ and energy consumption of typical air-conditioned rooms based on a hybrid prediction model. **Building and Environment**, *[S. l.]*, v. 225, n. August, p. 109576, 2022. DOI: 10.1016/j.buildenv.2022.109576. Disponível em: https://doi.org/10.1016/j.buildenv.2022.109576.

HSU, Hsieh Chih; CHANG, Shen; CHEN, Chien Chih; WU, I. Chen. Knowledge-based system for resolving design clashes in building information models. **Automation in Construction**, *[S. l.]*, v. 110, n. March 2019, p. 103001, 2020. DOI: 10.1016/j.autcon.2019.103001.

HU, Yibo; SONG, Jinbo; ZHAO, Tingting. Evolutionary game analysis of the intelligent upgrading of smart solar photovoltaic projects. **Engineering, Construction and Architectural Management**, *[S. l.]*, v. 31, n. 5, p. 1835–1856, 2024. DOI: 10.1108/ECAM-07-2021-0631.

HUANG, Chien Hsun; HSIEH, Shang Hsien. Predicting BIM labor cost with random forest and simple linear regression. **Automation in Construction**, *[S. l.]*, v. 118, n. May, p. 103280, 2020. DOI: 10.1016/j.autcon.2020.103280. Disponível em: https://doi.org/10.1016/j.autcon.2020.103280.

HUANG, M. Q.; NINIĆ, J.; ZHANG, Q. B. BIM, machine learning and computer vision techniques in underground construction: Current status and future perspectives. **Tunnelling**

and Underground Space Technology, *[S. l.]*, v. 108, p. 103677, 2021. DOI: 10.1016/j.tust.2020.103677.

IBGE. **IBGE Cidades**. 2024. Disponível em: https://bdmep.inmet.gov.br/#. Acesso em: 5 nov. 2024.

INMET. **Banco de Dados Meteorológicos do INMET**. 2023. Disponível em: https://bdmep.inmet.gov.br/#. Acesso em: 5 nov. 2024.

JANG, Suhyung; LEE, Ghang; OH, Jiseok; LEE, Junghun; KOO, Bonsang. Automated detailing of exterior walls using NADIA: Natural-language-based architectural detailing through interaction with AI. **Advanced Engineering Informatics**, *[S. l.]*, v. 61, p. 102532, 2024. DOI: 10.1016/j.aei.2024.102532.

JIANG, Yuhan; HAN, Sisi; BAI, Yong. Scan4Façade: Automated As-Is Façade Modeling of Historic High-Rise Buildings Using Drones and AI. **Journal of Architectural Engineering**, *[S. l.]*, v. 28, n. 4, p. 1–22, 2022. DOI: 10.1061/(asce)ae.1943-5568.0000564.

JING YANG, Rebecca *et al.* Digitalising BIPV energy simulation: A cross tool investigation. **Energy and Buildings**, *[S. l.]*, v. 318, p. 114484, 2024. DOI: 10.1016/j.enbuild.2024.114484.

KATHIRAVEL, Rojini; ZHU, Shiyao; FENG, Haibo. LCA of net-zero energy residential buildings with different HVAC systems across Canadian climates: A BIM-based fuzzy approach. **Energy and Buildings**, *[S. l.]*, v. 306, p. 113905, 2024. DOI: 10.1016/j.enbuild.2024.113905.

KAYHANI, Navid; MCCABE, Brenda; SANKARAN, Bharath. Semantic-aware quality assessment of building elements using graph neural networks. **Automation in Construction**, *[S. l.]*, v. 155, n. July, p. 105054, 2023. DOI: 10.1016/j.autcon.2023.105054.

KAZANASMAZ, Tuğçe; GÜNAYDIN, Murat; BINOL, Selcen. Artificial neural networks to predict daylight illuminance in office buildings. **Building and Environment**, *[S. l.]*, v. 44, n. 8, p. 1751–1757, 2009. DOI: 10.1016/j.buildenv.2008.11.012.

KIM, Byungil; KIM, Kyeongseok; KIM, Changyoon. Determining the optimal installation timing of building integrated photovoltaic systems. **Journal of Cleaner Production**, *[S. l.]*, v. 140, p. 1322–1329, 2017. DOI: 10.1016/j.jclepro.2016.10.020.

KIM, Hyunsoo; KIM, Changwan. 3D as-built modeling from incomplete point clouds using connectivity relations. **Automation in Construction**, *[S. l.]*, v. 130, n. July, p. 103855, 2021. DOI: 10.1016/j.autcon.2021.103855.

KUMARI, Pratima; TOSHNIWAL, Durga. Deep learning models for solar irradiance forecasting: A comprehensive review. **Journal of Cleaner Production**, *[S. l.]*, v. 318, p. 128566, 2021. DOI: 10.1016/j.jclepro.2021.128566.

LEON-GARZA, Hugo; HAGRAS, Hani; PEÑA-RIOS, Anasol; CONWAY, Anthony; OWUSU, Gilbert. A type-2 fuzzy system-based approach for image data fusion to create building information models. **Information Fusion**, *[S. l.]*, v. 88, n. July, p. 115–125, 2022. DOI: 10.1016/j.inffus.2022.07.007.

LI, Nan; HAN, Yiming; GAO, Feng. Theory and practice of BIM skills of construction management professional based on conceive–design–implement–operate engineering teaching mode. **Computer Applications in Engineering Education**, *[S. l.]*, 2024. DOI: 10.1002/cae.22719.

LI, Sihao; WANG, Jiali; XU, Zhao. Automated compliance checking for BIM models based on Chinese-NLP and knowledge graph: an integrative conceptual framework. **Engineering, Construction and Architectural Management**, *[S. l.]*, 2024. DOI: 10.1108/ECAM-10-2023-1037.

LI, Zhen; LU, Tieding; HE, Xiaoxing; MONTILLET, Jean-Philippe; TAO, Rui. An improved cyclic multi model-eXtreme gradient boosting (CMM-XGBoost) forecasting algorithm on the GNSS vertical time series. **Advances in Space Research**, *[S. l.]*, v. 71, n. 1, p. 912–935, 2023. DOI: 10.1016/j.asr.2022.08.038.

LIN, Penghui; WU, Maozhi; ZHANG, Limao. Probabilistic safety risk assessment in large-diameter tunnel construction using an interactive and explainable tree-based pipeline optimization method. **Applied Soft Computing**, *[S. l.]*, v. 143, p. 110376, 2023. DOI: 10.1016/j.asoc.2023.110376. Disponível em: https://doi.org/10.1016/j.asoc.2023.110376.

LIN, Qihang; KENSEK, K.; SCHILER, M.; CHOI, J. Streamlining sustainable design in building information modeling BIM-based PV design and analysis tools. **Architectural Science Review**, *[S. l.]*, v. 64, n. 6, p. 467–477, 2021. DOI: 10.1080/00038628.2021.1884525.

LIN, Tzu Hsuan; HUANG, Yu Hua; PUTRANTO, Alan. Intelligent question and answer system for building information modeling and artificial intelligence of things based on the bidirectional encoder representations from transformers model. **Automation in Construction**, *[S. l.]*, v. 142, n. July, p. 104483, 2022. DOI: 10.1016/j.autcon.2022.104483.

LING, Florence Y. Y.; HENG, Gerald Tze Hon; CHANG-RICHARDS, Alice; CHEN, Xichen; YIU, Tak Wing. Impact of Digital Technology Adoption on the Comparative Advantage of Architectural, Engineering, and Construction Firms in Singapore. **Journal of Construction Engineering and Management**, *[S. l.]*, v. 149, n. 12, 2023. DOI: 10.1061/JCEMD4.COENG-13743.

LINS, Eduardo José Melo; PALHA, Rachel Perez; SOBRAL, Maria do Carmo Martins; ARAÚJO, Adolpho Guido De; MARQUES, Érika Alves Tavares. Application of Building Information Modelling in Construction and Demolition Waste Management: Systematic Review and Future Trends Supported by a Conceptual Framework. **Sustainability**, *[S. l.]*, v. 16, n. 21, p. 9425, 2024. DOI: 10.3390/su16219425.

LIU, Ke; XU, Xiaodong; HUANG, Wenxin; ZHANG, Ran; KONG, Lingyu; WANG, Xi. A multi-objective optimization framework for designing urban block forms considering daylight, energy consumption, and photovoltaic energy potential. **Building and Environment**, *[S. l.]*, v. 242, p. 110585, 2023. a. DOI: 10.1016/j.buildenv.2023.110585.

LIU, Qiuping; CHEN, Yaodong; LIU, Yang; LEI, Yuanfang; WANG, Yibo; HU, Pantin. A review and guide on selecting and optimizing machine learning algorithms for daylight prediction. **Building and Environment**, *[S. l.]*, v. 244, p. 110822, 2023. b. DOI: 10.1016/j.buildenv.2023.110822.

LIU, Zhijian; ZHANG, Yulong; YUAN, Xitao; LIU, Yuanwei; XU, Jinliang; ZHANG, Shicong; HE, Bao-jie. A comprehensive study of feasibility and applicability of building integrated photovoltaic (BIPV) systems in regions with high solar irradiance. **Journal of Cleaner Production**, *[S. l.]*, v. 307, p. 127240, 2021. DOI: 10.1016/j.jclepro.2021.127240.

LONG, Xu; MAO, Ming-hui; SU, Tian-xiong; SU, Yu-tai; TIAN, Meng-ke. Machine learning method to predict dynamic compressive response of concrete-like material at high strain rates. **Defence Technology**, *[S. l.]*, v. 23, p. 100–111, 2023. DOI: 10.1016/j.dt.2022.02.003.

LÓPEZ, Irene Del Hierro; OLIVIERI, Lorenzo. Comprehensive review of building-integrated photovoltaics in the renovation of heritage buildings. **Journal of Building Engineering**, *[S. l.]*, v. 108, p. 112883, 2025. DOI: 10.1016/j.jobe.2025.112883.

LU, Xiu; LI, Guannan; ZHOU, Liangchen; HAO, Lisha; LV, Guonian; LIN, Bingxian. Photovoltaic potential estimation for various surface components of urban residential buildings based on Industry Foundation Classes data. **Energy Science & Engineering**, *[S. l.]*, v. 10, n. 10, p. 3741–3765, 2022. DOI: 10.1002/ese3.1253.

MA, Jong Won; LEITE, Fernanda. Performance boosting of conventional deep learning-based semantic segmentation leveraging unsupervised clustering. **Automation in Construction**, *[S. l.]*, v. 136, n. February, p. 104167, 2022. DOI: 10.1016/j.autcon.2022.104167.

MANDIČÁK, Tomáš; SPIŠÁKOVÁ, Marcela; MÉSÁROŠ, Peter. Sustainable Design and Building Information Modeling of Construction Project Management towards a Circular Economy. **Sustainability**, *[S. l.]*, v. 16, n. 11, p. 4376, 2024. DOI: 10.3390/su16114376.

MANNI, Mattia; NOCENTE, Alessandro; KONG, Gefei; SKEIE, Kristian; FAN, Hongchao; LOBACCARO, Gabriele. Solar energy digitalization at high latitudes: A model chain combining solar irradiation models, a LiDAR scanner, and high-detail 3D building model. **Frontiers in Energy Research**, *[S. l.]*, v. 10, 2022. DOI: 10.3389/fenrg.2022.1082092.

MARROQUIN, Roberto; DUBOIS, Julien; NICOLLE, Christophe. Ontology for a Panoptes building: Exploiting contextual information and a smart camera network. **Semantic Web**, *[S. l.]*, v. 9, n. 6, p. 803–828, 2018. DOI: 10.3233/SW-180298.

MARTÍN-JIMÉNEZ, J.; DEL POZO, S.; SÁNCHEZ-APARICIO, M.; LAGÜELA, S. Multi-scale roof characterization from LiDAR data and aerial orthoimagery: Automatic computation of building photovoltaic capacity. **Automation in Construction**, *[S. l.]*, v. 109, p. 102965, 2020. DOI: 10.1016/j.autcon.2019.102965.

MARZOUK, Mohamed; ZAHER, Mohamed. Artificial intelligence exploitation in facility management using deep learning. **Construction Innovation**, *[S. l.]*, v. 20, n. 4, p. 609–624, 2020. a. DOI: 10.1108/CI-12-2019-0138. Disponível em: https://www.emerald.com/insight/content/doi/10.1108/CI-12-2019-0138/full/html.

MATRONE, Francesca; MARTINI, Massimo. Transfer Learning and Performance Enhancement Techniques for Deep Semantic Segmentation of Built Heritage Point Clouds.

Virtual Archaeology Review, *[S. l.]*, v. 12, n. 25, p. 73–84, 2021. DOI: 10.4995/var.2021.15318.

MCTI. **Fatores de emissão de CO2 pela geração de energia elétrica no Sistema Interligado Nacional do Brasil - Ano Base 2023**. 2023. https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/sirene/dados-e-ferramentas/fatores-de-emissao

MEHRABAN, Mohammad H.; ALNASER, Aljawharah A.; SEPASGOZAR, Samad M. E. Building Information Modeling and AI Algorithms for Optimizing Energy Performance in Hot Climates: A Comparative Study of Riyadh and Dubai. **Buildings**, *[S. l.]*, v. 14, n. 9, p. 2748, 2024. DOI: 10.3390/buildings14092748.

MESCHINI, Silvia; PELLEGRINI, Laura; LOCATELLI, Mirko; ACCARDO, Daniele; TAGLIABUE, Lavinia Chiara; DI GIUDA, Giuseppe Martino; AVENA, Marco. Toward cognitive digital twins using a BIM-GIS asset management system for a diffused university. **Frontiers in Built Environment**, *[S. l.]*, v. 8, n. December, p. 1–28, 2022. DOI: 10.3389/fbuil.2022.959475.

MIKALEF, Patrick; GUPTA, Manjul. Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. **Information & Management**, *[S. l.]*, v. 58, n. 3, p. 103434, 2021. DOI: 10.1016/j.im.2021.103434.

MOHAMMADI, Masoud; RASHIDI, Maria; YU, Yang; SAMALI, Bijan. Integration of TLS-derived Bridge Information Modeling (BrIM) with a Decision Support System (DSS) for digital twinning and asset management of bridge infrastructures. **Computers in Industry**, *[S. l.]*, v. 147, n. January, p. 103881, 2023. DOI: 10.1016/j.compind.2023.103881.

MOUSAVI, Milad; TOHIDIFAR, Ali; ALVANCHI, Amin. BIM and machine learning in seismic damage prediction for non-structural exterior infill walls. **Automation in Construction**, *[S. l.]*, v. 139, n. May, p. 104288, 2022. DOI: 10.1016/j.autcon.2022.104288.

MUHAMMAD, Ilyas; YING, Khaw; NITHISH, Muthuchamy; XIN, Jin; XINGE, Zhao; CHEAH, Chien Chern. Robot-Assisted Object Detection for Construction Automation: Data and Information-Driven Approach. **IEEE/ASME Transactions on Mechatronics**, *[S. l.]*, v. 26, n. 6, p. 2845–2856, 2021. DOI: 10.1109/TMECH.2021.3100306.

MULERO-PALENCIA, Sofía; ÁLVAREZ-DÍAZ, Sonia; ANDRÉS-CHICOTE, Manuel. Machine Learning for the Improvement of Deep Renovation Building Projects Using As-Built BIM Models. **Sustainability**, *[S. l.]*, v. 13, n. 12, p. 6576, 2021. DOI: 10.3390/su13126576.

MUNIANDAY, Praveena; RADZI, Afiqah R.; ESA, Muneera; RAHMAN, Rahimi A. Optimal Strategies for Improving Organizational BIM Capabilities: PLS-SEM Approach. **Journal of Management in Engineering**, *[S. l.]*, v. 38, n. 3, 2022. DOI: 10.1061/(ASCE)ME.1943-5479.0001038.

MUTAVHATSINDI, Tendani; SIGAUKE, Caston; MBUVHA, Rendani. Forecasting Hourly Global Horizontal Solar Irradiance in South Africa Using Machine Learning Models. **IEEE Access**, *[S. l.]*, v. 8, p. 198872–198885, 2020. DOI: 10.1109/ACCESS.2020.3034690.

MYINT, Nwe Ni; SHAFIQUE, Muhammad; ZHOU, Xiangming; ZHENG, Zhuang. Net zero carbon buildings: A review on recent advances, knowledge gaps and research directions. **Case Studies in Construction Materials**, *[S. l.]*, v. 22, p. e04200, 2025. DOI: 10.1016/j.cscm.2024.e04200.

NADERI, Mojdeh; NAZARI, Ahad; SHAFAAT, Ali; ABRISHAMI, Sepehr. Enhancing accuracy in construction overhead cost estimation: a novel integration of activity-based costing and building information modelling. **Smart and Sustainable Built Environment**, *[S. l.]*, 2024. DOI: 10.1108/SASBE-07-2023-0180.

NASCIMENTO, C. R. S. M. S.; DE ALMEIDA-FILHO, A. T.; PALHA, R. P. A TOPSIS-Based Decision Model to Establish Priorities for Sequencing the Design of Construction Projects in the Public Sector. **Mathematical Problems in Engineering**, *[S. l.]*, v. 2023, n. 1, 2023. DOI: 10.1155/2023/1414294.

NASCIMENTO, Cláudia Rafaela Saraiva de Melo Simões; ALMEIDA-FILHO, Adiel Teixeira De; PALHA, Rachel Perez. A TOPSIS-based framework for construction projects' portfolio selection in the public sector. **Engineering, Construction and Architectural Management**, *[S. l.]*, 2023. DOI: 10.1108/ECAM-05-2023-0534.

NING, Gui; KAN, He; ZHIFENG, Qiu; WEIHUA, Gui; GEERT, Deconinck. e-BIM: a BIM-centric design and analysis software for Building Integrated Photovoltaics. **Automation in Construction**, *[S. l.]*, v. 87, p. 127–137, 2018. DOI: 10.1016/j.autcon.2017.10.020.

NOVEMBRI, Gabriele; ROSSINI, Francesco Livio. Swarm modelling framework to improve design support systems capabilities. **Journal of Information Technology in Construction**, *[S. l.]*, v. 25, n. April, p. 398–415, 2020. DOI: 10.36680/j.itcon.2020.023.

OLU-AJAYI, Razak; ALAKA, Hafiz; SULAIMON, Ismail; SUNMOLA, Funlade; AJAYI, Saheed. Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques. **Journal of Building Engineering**, *[S. l.]*, v. 45, p. 103406, 2022. DOI: 10.1016/j.jobe.2021.103406.

PADALA, S. P. Sreenivas; SKANDA, M. Prabhanjan. BIM-based multi-objective optimization framework for volumetric analysis of building projects. **Journal of Engineering, Design and Technology**, *[S. l.]*, 2024. DOI: 10.1108/JEDT-07-2023-0309.

PALHA, Rachel Perez; HÜTTL, Ricardo Maciel Castro; COSTA E SILVA, Angelo Just Da. BIM interoperability for small residential construction integrating warranty and maintenance management. **Automation in Construction**, *[S. l.]*, v. 166, p. 105639, 2024. DOI: 10.1016/j.autcon.2024.105639.

PAN, Yue; ZHANG, Limao. Roles of artificial intelligence in construction engineering and management: A critical review and future trends. **Automation in Construction**, *[S. l.]*, v. 122, p. 103517, 2021. a. DOI: 10.1016/j.autcon.2020.103517. Disponível em: https://linkinghub.elsevier.com/retrieve/pii/S0926580520310979.

PEIMAN, Farshad; KHALILZADEH, Mohammad; SHAHSAVARI-POUR, Nasser; RAVANSHADNIA, Mehdi. Estimation of building project completion duration using a natural gradient boosting ensemble model and legal and institutional variables. **Engineering,**

Construction and Architectural Management, *[S. l.]*, 2023. a. DOI: 10.1108/ECAM-12-2022-1170.

PIERDICCA, Roberto; PAOLANTI, Marina; MATRONE, Francesca; MARTINI, Massimo; MORBIDONI, Christian; MALINVERNI, Eva Savina; FRONTONI, Emanuele; LINGUA, Andrea Maria. Point cloud semantic segmentation using a deep learning framework for cultural heritage. **Remote Sensing**, *[S. l.]*, v. 12, n. 6, p. 1–23, 2020. DOI: 10.3390/rs12061005.

POSHNATH, Aravind; RISMANCHI, Behzad; RAJABIFARD, Abbas. Adoption of Renewable Energy Systems in common properties of multi-owned buildings: Introduction of 'Energy Entitlement'. **Energy Policy**, *[S. l.]*, v. 174, p. 113465, 2023. DOI: 10.1016/j.enpol.2023.113465.

POUX, F.; MATTES, C.; SELMAN, Z.; KOBBELT, L. Automatic region-growing system for the segmentation of large point clouds. **Automation in Construction**, *[S. l.]*, v. 138, n. July 2021, p. 104250, 2022. DOI: 10.1016/j.autcon.2022.104250.

PUPIN, Priscila Carvalho; PERAZZINI, Maisa Tonon Bitti; GRILLO RENÓ, Maria Luiza; PERAZZINI, Hugo; HADDAD, Jamil; YAMACHITA, Roberto Akira. Life cycle assessment for producing monocrystalline photovoltaic panels: a case study of Brazil. **Energy Sources, Part A: Recovery, Utilization, and Environmental Effects**, *[S. l.]*, v. 45, n. 4, p. 12924–12937, 2023. DOI: 10.1080/15567036.2023.2278724.

QI, Charles R.; SU, Hao; MO, Kaichun; GUIBAS, Leonidas J. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. **Computer Vision and Pattern Recognition**, *[S. l.]*, v. 2, 2017. DOI: https://doi.org/10.48550/arXiv.1612.00593.

QIU, Andong; YANG, Zhouwang. Variable-depth large neighborhood search algorithm for cable routing in distributed photovoltaic systems. **Automation in Construction**, *[S. l.]*, v. 168, p. 105839, 2024. DOI: 10.1016/j.autcon.2024.105839.

RAFSANJANI, Hamed Nabizadeh; NABIZADEH, Amir Hossein. Towards digital architecture, engineering, and construction (AEC) industry through virtual design and construction (VDC) and digital twin. **Energy and Built Environment**, *[S. l.]*, v. 4, n. 2, p. 169–178, 2023. DOI: 10.1016/j.enbenv.2021.10.004.

RANGASAMY, Veerakumar; YANG, Jyh-Bin. The convergence of BIM, AI and IoT: Reshaping the future of prefabricated construction. **Journal of Building Engineering**, *[S. l.]*, v. 84, p. 108606, 2024. DOI: 10.1016/j.jobe.2024.108606.

RATAJCZAK, Julia; SIEGELE, Dietmar; NIEDERWIESER, Elias. Maximizing Energy Efficiency and Daylight Performance in Office Buildings in BIM through RBFOpt Model-Based Optimization: The GENIUS Project. **Buildings**, *[S. l.]*, v. 13, n. 7, 2023. DOI: 10.3390/buildings13071790.

ROGAGE, Kay; DOUKARI, Omar. 3D object recognition using deep learning for automatically generating semantic BIM data. **Automation in Construction**, *[S. l.]*, v. 162, p. 105366, 2024. DOI: 10.1016/j.autcon.2024.105366.

SAÂDAOUI, Foued. A seasonal feedforward neural network to forecast electricity prices. **Neural Computing and Applications**, *[S. l.]*, v. 28, n. 4, p. 835–847, 2017. DOI: 10.1007/s00521-016-2356-y.

SAIGUSTIA, Candra; PIJARSKI, Paweł. Time Series Analysis and Forecasting of Solar Generation in Spain Using eXtreme Gradient Boosting: A Machine Learning Approach. **Energies**, *[S. l.]*, v. 16, n. 22, p. 7618, 2023. DOI: 10.3390/en16227618.

SAMMAR, Muhammad Jibreel; SAEED, Muhammad Anwaar; MOHSIN, Syed Muhammad; AKBER, Syed Muhammad Abrar; BUKHSH, Rasool; ABAZEED, Mohammed; ALI, Mohammed. Illuminating the Future: A Comprehensive Review of AI-Based Solar Irradiance Prediction Models. **IEEE Access**, *[S. l.]*, v. 12, p. 114394–114415, 2024. DOI: 10.1109/ACCESS.2024.3402096.

SCHERZ, Marco; HOXHA, Endrit; KREINER, Helmuth; PASSER, Alexander; VAFADARNIKJOO, Amin. A hierarchical reference-based know-why model for design support of sustainable building envelopes. **Automation in Construction**, *[S. l.]*, v. 139, n. April, p. 104276, 2022. DOI: 10.1016/j.autcon.2022.104276. Disponível em: https://doi.org/10.1016/j.autcon.2022.104276.

SERAT, Zainullah; CHEN, Xin; ZUO, Hongyang; LI, Jun. Design strategies for building rooftop photovoltaic systems: Efficiency and grid integration. **Journal of Building Engineering**, *[S. l.]*, v. 100, p. 111693, 2025. DOI: 10.1016/j.jobe.2024.111693.

SHAO, Bilin; MENG, Jie; CHE, Wanbo. Long-Term Energy Usage Prediction in Public Buildings Using Aggregated Modal Decomposition and GRU. **Journal of Construction Engineering and Management**, *[S. l.]*, v. 151, n. 8, 2025. DOI: 10.1061/JCEMD4.COENG-15410.

SHAO, Chuanyong; MIGAN-DUBOIS, Anne; DIALLO, Demba. Performance of BIPV system under partial shading condition. **Solar Energy**, *[S. l.]*, v. 283, p. 112969, 2024. DOI: 10.1016/j.solener.2024.112969.

SHAO, Jie; ZHANG, Wuming; SHEN, Aojie; MELLADO, Nicolas; CAI, Shangshu; LUO, Lei; WANG, Nan; YAN, Guangjian; ZHOU, Guoqing. Seed point set-based building roof extraction from airborne LiDAR point clouds using a top-down strategy. **Automation in Construction**, *[S. l.]*, v. 126, p. 103660, 2021. DOI: 10.1016/j.autcon.2021.103660.

SHEN, Yuxuan; PAN, Yue. BIM-supported automatic energy performance analysis for green building design using explainable machine learning and multi-objective optimization. **Applied Energy**, *[S. l.]*, v. 333, n. August 2022, p. 120575, 2023. DOI: 10.1016/j.apenergy.2022.120575.

SHU, Jiangpeng; LI, Wenhao; ZHANG, Congguang; GAO, Yifan; XIANG, Yiqiang; MA, Ling. Point cloud-based dimensional quality assessment of precast concrete components using deep learning. **Journal of Building Engineering**, *[S. l.]*, v. 70, n. December 2022, p. 106391, 2023. DOI: 10.1016/j.jobe.2023.106391.

SONG, Xia; XU, Bo; ZHAO, Zhenzhen. Can people experience romantic love for artificial intelligence? An empirical study of intelligent assistants. **Information & Management**, *[S. l.]*, v. 59, n. 2, p. 103595, 2022. DOI: 10.1016/j.im.2022.103595.

SZALAY, Zsuzsa; SZAGRI, Dóra; BIHARI, Ádám; NAGY, Balázs; KISS, Benedek; HORVÁTH, Miklós; MEDGYASSZAY, Péter. Development of a life cycle net zero carbon compact house concept. **Energy Reports**, *[S. l.]*, v. 8, p. 12987–13013, 2022. DOI: 10.1016/j.egyr.2022.09.197.

TAO, Hai; ALAWI, Omer A.; HOMOD, Raad Z.; MOHAMMED, Mustafa KA.; GOLIATT, Leonardo; TOGUN, Hussein; SHAFIK, Shafik S.; HEDDAM, Salim; YASEEN, Zaher Mundher. Data driven insights for parabolic trough solar collectors: Artificial intelligence-based energy and exergy performance analysis. **Journal of Cleaner Production**, *[S. l.]*, v. 443, p. 141069, 2024. DOI: 10.1016/j.jclepro.2024.141069.

TAVOLARE, Riccardo; CABRERA, Elena; VERDOSCIA, Cesare; BULDO, Michele. A point cloud classification method for the Scan-to-BIM process in Architectural Heritage. **Disegnarecon**, *[S. l.]*, v. 16, n. 30, p. 201–208, 2023. DOI: 10.20365/disegnarecon.30.2023.20.

TIAN, Jia; OOKA, Ryozo; LEE, Doyun. Multi-scale solar radiation and photovoltaic power forecasting with machine learning algorithms in urban environment: A state-of-the-art review. **Journal of Cleaner Production**, *[S. l.]*, v. 426, p. 139040, 2023. DOI: 10.1016/j.jclepro.2023.139040.

TIXIER, Antoine J. P.; HALLOWELL, Matthew R.; RAJAGOPALAN, Balaji; BOWMAN, Dean. Application of machine learning to construction injury prediction. **Automation in Construction**, *[S. l.]*, v. 69, p. 102–114, 2016. DOI: 10.1016/j.autcon.2016.05.016.

TRANFIELD, David; DENYER, David; SMART, Palminder. Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. **British Journal of Management**, *[S. l.]*, v. 14, n. 3, p. 207–222, 2003. DOI: 10.1111/1467-8551.00375.

UFPE. **UFPE inaugura cinco usinas fotovoltaicas dentro das comemorações dos 75 anos da instituição**. 2021. https://www.ufpe.br/inicio/-/asset_publisher/55e3vpMwmIA2/content/ufpe-celebra-75-anos-de-fundacao-com-uma-serie-de-eventos-este-mes/40615#:~:text=UFPE%20inaugura%20cinco%20usinas%20fotovoltaicas,Honoris%20Causa%20no%20dia%2013

UFPE. **Conheça a UFPE**. 2025. https://www.ufpe.br/institucional/a-instituicao

URBIETA, Martin; URBIETA, Matias; LABORDE, Tomas; VILLARREAL, Guillermo; ROSSI, Gustavo. Generating BIM model from structural and architectural plans using Artificial Intelligence. **Journal of Building Engineering**, *[S. l.]*, v. 78, n. September, p. 107672, 2023. DOI: 10.1016/j.jobe.2023.107672. Disponível em: https://doi.org/10.1016/j.jobe.2023.107672.

UTKUCU, Duygu; YING, Huaquan; WANG, Zijian; SACKS, Rafael. Classification of architectural and MEP BIM objects for building performance evaluation. **Advanced Engineering Informatics**, *[S. l.]*, v. 61, p. 102503, 2024. DOI: 10.1016/j.aei.2024.102503.

VAN DER ZWAAG, Marco; WANG, Tong; BAKKER, Hans; VAN NEDERVEEN, Sander; SCHUURMAN, A. C. B.; BOSMA, Douwe. Evaluating building circularity in the early design phase. **Automation in Construction**, *[S. l.]*, v. 152, n. May, p. 104941, 2023. DOI: 10.1016/j.autcon.2023.104941.

VILLA, Valentina; BRUNO, Giulia; ALIEV, Khurshid; PIANTANIDA, Paolo; CORNELI, Alessandra; ANTONELLI, Dario. Machine Learning Framework for the Sustainable Maintenance of Building Facilities. **Sustainability (Switzerland)**, *[S. l.]*, v. 14, n. 2, p. 1–17, 2022. a. DOI: 10.3390/su14020681.

VOLK, Rebekka; STENGEL, Julian; SCHULTMANN, Frank. Building Information Modeling (BIM) for existing buildings — Literature review and future needs. **Automation in Construction**, *[S. l.]*, v. 38, p. 109–127, 2014. DOI: 10.1016/j.autcon.2013.10.023.

WANG, Chaofeng; YU, Qian; LAW, Kincho H.; MCKENNA, Frank; YU, Stella X.; TACIROGLU, Ertugrul; ZSARNÓCZAY, Adam; ELHADDAD, Wael; CETINER, Barbaros. Machine learning-based regional scale intelligent modeling of building information for natural hazard risk management. **Automation in Construction**, *[S. l.]*, v. 122, n. October 2020, 2021. DOI: 10.1016/j.autcon.2020.103474.

WANG, Kaiyang. Mapping the global knowledge landscape of digital transformation in the AEC industry: a scientometric analysis. **Engineering, Construction and Architectural Management**, *[S. l.]*, 2024. DOI: 10.1108/ECAM-11-2023-1174.

WANG, Lufeng; LIU, Jiepeng; ZENG, Yan; CHENG, Guozhong; HU, Huifeng; HU, Jiahao; HUANG, Xuesi. Automated building layout generation using deep learning and graph algorithms. **Automation in Construction**, *[S. l.]*, v. 154, n. July, p. 105036, 2023. DOI: 10.1016/j.autcon.2023.105036. Disponível em: https://doi.org/10.1016/j.autcon.2023.105036.

WANG, Tao; GAN, Vincent J. L. Automated joint 3D reconstruction and visual inspection for buildings using computer vision and transfer learning. **Automation in Construction**, *[S. l.]*, v. 149, n. May 2022, p. 104810, 2023. b. DOI: 10.1016/j.autcon.2023.104810. Disponível em: https://doi.org/10.1016/j.autcon.2023.104810.

WANG, Wanting; XU, Kaiyan; SONG, Shenghui; BAO, Yuxin; XIANG, Changying. From BIM to digital twin in BIPV: A review of current knowledge. **Sustainable Energy Technologies and Assessments**, *[S. l.]*, v. 67, p. 103855, 2024. DOI: 10.1016/j.seta.2024.103855.

WANG, Yongqi; ZHANG, Limao; YU, Hongbo; TIONG, Robert L. K. Detecting logical relationships in mechanical, electrical, and plumbing (MEP) systems with BIM using graph matching. **Advanced Engineering Informatics**, *[S. l.]*, v. 54, n. October, p. 101770, 2022. DOI: 10.1016/j.aei.2022.101770. Disponível em: https://doi.org/10.1016/j.aei.2022.101770.

WANG, Yue; WANG, Shouxiang; ZHAO, Qianyu. Experimental investigation on fire propagation characteristics and influencing factors in rooftop photovoltaic modules. **Journal of Building Engineering**, *[S. l.]*, v. 110, p. 113063, 2025. DOI: 10.1016/j.jobe.2025.113063.

WEI, Lim Jun; ISLAM, M. M.; HASANUZZAMAN, M.; CUCE, Erdem. Energy consumption, power generation and performance analysis of solar photovoltaic module based building roof. **Journal of Building Engineering**, *[S. l.]*, v. 90, p. 109361, 2024. DOI: 10.1016/j.jobe.2024.109361.

WEI, Y.; AKINCI, B. A vision and learning-based indoor localization and semantic mapping framework for facility operations and management. **Automation in Construction**, *[S. l.]*, v. 107, n. August, p. 102915, 2019. DOI: 10.1016/j.autcon.2019.102915.

WIJAYARATHNE, Navodi; GUNAWAN, Indra; SCHULTMANN, Frank. Dynamic Capabilities in Digital Transformation: A Systematic Review of Their Role in the Construction Industry. **Journal of Construction Engineering and Management**, *[S. l.]*, v. 150, n. 11, 2024. DOI: 10.1061/JCEMD4.COENG-15055.

WINKLER, Charlotta. Implementing solar photovoltaic systems in buildings: a case of systemic innovation in the construction sector. **Construction Innovation**, *[S. l.]*, v. 24, n. 7, p. 102–123, 2024. DOI: 10.1108/CI-10-2022-0264.

WINKLER, Charlotta; PEREZ VICO, Eugenia; WIDÉN, Kristian. Challenges to business ecosystem alignment when implementing solar photovoltaic systems in the Swedish built environment. **Building Research & Information**, *[S. l.]*, v. 52, n. 5, p. 497–514, 2024. DOI: 10.1080/09613218.2023.2256435.

WU, Bin; MAALEK, Reza. Renovation or Redevelopment: The Case of Smart Decision-Support in Aging Buildings. **Smart Cities**, *[S. l.]*, v. 6, n. 4, p. 1922–1936, 2023. DOI: 10.3390/smartcities6040089.

WU, Haitao; ZHONG, Botao; LI, Heng; LOVE, Peter; PAN, Xing; ZHAO, Neng. Combining computer vision with semantic reasoning for on-site safety management in construction. **Journal of Building Engineering**, *[S. l.]*, v. 42, n. October 2020, p. 103036, 2021. DOI: 10.1016/j.jobe.2021.103036. Disponível em: https://doi.org/10.1016/j.jobe.2021.103036.

WU, Ping; YANG, Linxi; LI, Wangxin; HUANG, Jiamin; XU, Yidong. Construction Safety Risk Assessment and Early Warning of Nearshore Tunnel Based on BIM Technology. **Journal of Marine Science and Engineering**, *[S. l.]*, v. 11, n. 10, 2023. DOI: 10.3390/jmse11101996.

XIA, Jiahao; GONG, Jie. Precise indoor localization with 3D facility scan data. **Computer-Aided Civil and Infrastructure Engineering**, *[S. l.]*, v. 37, n. 10, p. 1243–1259, 2022. DOI: 10.1111/mice.12795.

YAN, Ke; SHEN, Hengle; WANG, Lei; ZHOU, Huiming; XU, Meiling; MO, Yuchang. Short-Term Solar Irradiance Forecasting Based on a Hybrid Deep Learning Methodology. **Information**, *[S. l.]*, v. 11, n. 1, p. 32, 2020. DOI: 10.3390/info11010032.

YILMAZ, Gokcen; AKCAMETE, Asli; DEMIRORS, Onur. BIM-CAREM: Assessing the BIM capabilities of design, construction and facilities management processes in the construction industry. **Computers in Industry**, *[S. l.]*, v. 147, p. 103861, 2023. DOI: 10.1016/j.compind.2023.103861.

YIN, Chao; WANG, Boyu; GAN, Vincent J. L.; WANG, Mingzhu; CHENG, Jack C. P. Automated semantic segmentation of industrial point clouds using ResPointNet++. **Automation in Construction**, *[S. l.]*, v. 130, n. July, p. 103874, 2021. DOI: 10.1016/j.autcon.2021.103874. Disponível em: https://doi.org/10.1016/j.autcon.2021.103874.

YING, Huaquan; SACKS, Rafael; DEGANI, Amir. Synthetic image data generation using BIM and computer graphics for building scene understanding. **Automation in Construction**, *[S. l.]*, v. 154, n. July, p. 105016, 2023. DOI: 10.1016/j.autcon.2023.105016.

ZABIN, Asem; GONZÁLEZ, Vicente A.; ZOU, Yang; AMOR, Robert. Applications of machine learning to BIM: A systematic literature review. **Advanced Engineering Informatics**, *[S. l.]*, v. 51, n. April 2021, 2022. DOI: 10.1016/j.aei.2021.101474.

ZAWADA, Karol; RYBAK-NIEDZIÓŁKA, Kinga; DONDEREWICZ, Mikołaj; STARZYK, Agnieszka. Digitization of AEC Industries Based on BIM and 4.0 Technologies. **Buildings**, *[S. l.]*, v. 14, n. 5, p. 1350, 2024. DOI: 10.3390/buildings14051350.

ZHANG, Anshan; WANG, Feiliang; LI, Huanyu; PANG, Bo; YANG, Jian. Carbon emissions accounting and estimation of carbon reduction potential in the operation phase of residential areas based on digital twin. **Applied Energy**, *[S. l.]*, v. 376, p. 123155, 2024. DOI: 10.1016/j.apenergy.2024.123155.

ZHANG, Fan; CHAN, Albert P. C.; DARKO, Amos; CHEN, Zhengyi; LI, Dezhi. Integrated applications of building information modeling and artificial intelligence techniques in the AEC/FM industry. **Automation in Construction**, *[S. l.]*, v. 139, p. 104289, 2022. DOI: 10.1016/j.autcon.2022.104289.

ZHANG, Lifei; YU, Jingyu; SHI, Qingyu; KONG, Quan. An evaluation of the economic benefits of rooftop distributed photovoltaic projects in the whole county in China. **Journal of Cleaner Production**, *[S. l.]*, v. 432, p. 139744, 2023. DOI: 10.1016/j.jclepro.2023.139744.

ZHANG, Ruichuan; EL-GOHARY, Nora. Transformer-based approach for automated context-aware IFC-regulation semantic information alignment. **Automation in Construction**, *[S. l.]*, v. 145, n. August 2022, p. 104540, 2023. DOI: 10.1016/j.autcon.2022.104540.

ZHANG, Xiang; SAELENS, Dirk; ROELS, Staf. Quantifying dynamic solar gains in buildings: Measurement, simulation and data-driven modelling. **Renewable and Sustainable Energy Reviews**, *[S. l.]*, v. 212, p. 115221, 2025. DOI: 10.1016/j.rser.2024.115221.

ZHANG, Yihan; CAO, Yuxin; CHEN, Tianyi; LUCCHI, Elena. Optimized Community-Level Distributed Photovoltaic Generation (DPVG): Aesthetic, Technical, Economic, and Environmental Assessment of Building Integrated Photovoltaic (BIPV) Systems. **Journal of Building Engineering**, *[S. l.]*, p. 112085, 2025. DOI: 10.1016/j.jobe.2025.112085.

ZHENG, Junwen; FISCHER, Martin. Dynamic prompt-based virtual assistant framework for BIM information search. **Automation in Construction**, *[S. l.]*, v. 155, n. April, p. 105067, 2023. DOI: 10.1016/j.autcon.2023.105067.

ZHI, Yuan; SUN, Tao; YANG, Xudong. A physical model with meteorological forecasting for hourly rooftop photovoltaic power prediction. **Journal of Building Engineering**, *[S. l.]*, v. 75, p. 106997, 2023. DOI: 10.1016/j.jobe.2023.106997.

ZHOU, Xiaoping; ZHAO, Jichao; WANG, Jia; HUANG, Xiaoyuan; LI, Xiaofei; GUO, Maozu; XIE, Peng. Parallel computing-based online geometry triangulation for building information modeling utilizing big data. **Automation in Construction**, *[S. l.]*, v. 107, n. July, p. 102942, 2019. DOI: 10.1016/j.autcon.2019.102942.

ZHOU, Yu Cheng; ZHENG, Zhe; LIN, Jia Rui; LU, Xin Zheng. Integrating NLP and context-free grammar for complex rule interpretation towards automated compliance checking. **Computers in Industry**, *[S. l.]*, v. 142, p. 103746, 2022. DOI: 10.1016/j.compind.2022.103746.

ZHU, Tianqing; YE, Dayong; WANG, Wei; ZHOU, Wanlei; YU, Philip S. More Than Privacy: Applying Differential Privacy in Key Areas of Artificial Intelligence. **IEEE Transactions on Knowledge and Data Engineering**, *[S. l.]*, v. 34, n. 6, p. 2824–2843, 2022. DOI: 10.1109/TKDE.2020.3014246.

APPENDIX A – SLR SOURCE IDENTIFICATION OF FINAL SAMPLE

Id	Source	Id	Source	Id	Source
1	Ergan and Akinci (2012)	61	Çetin et al. (2021)	121	Wu et al. (2022)
2	Wooa and Menassa (2014)	62	Sresakoolchai and Kaewunruen (2021)	122	Opoku et al. (2022)
3	Golparvar-Fard et al. (2015)	63	Abdirad and Mathur (2021)	123	Arabameri et al. (2022)
4	Chen et al. (2015)	64	Bosch-Sijtsema et al. (2021)	124	Kim et al. (2022)
5	Jalaei et al. (2015)	65	Cheng et al. (2021)	125	Pan and Zhang (2023))
6	Kim et al. (2015)	66	Liu and Jiang (2021)	126	Omar and Mahdjoubi (2023)
7	Yang et al. (2016)	67	Du (2021)	127	Kor and Yitmen (2023)
8	Chen and Pan (2016)	68	Spallone and Palma (2021)	128	Morfidis et al. (2023)
9	Solihin et al. (2016)	69	Johansen et al. (2021)	129	Perez and Tah (2023)
10	Tixier et al. (2016)	70	Chen et al. (2021)	130	Tavolare et al. (2023)
11	Chardon et al. (2016)	71	Matrone and Martini (2021)	131	Heidari et al. (2023)
12	Juszczyk (2016)	72	Li et al. (2021)	132	Mohammadi et al. (2023)
13	Bloch and Sacks (2018)	73	He et al. (2021)	133	Ceccarelli et al. (2023)
14	Bruno et al. (2018)	74	Musella et al. (2021)	134	Abd (2023)
15	Kim et al. (2018)	75	González et al. (2021)	135	Croce et al. (2023)
16	Marroquin et al. (2018)	76	Hong et al. (2021)	136	Lien and Dolgorsuren (2023)
17	Li et al (2018)	77	Yitmen et al. (2021)	137	Wu and Maalek (2023)
18	Mahankali et al. (2018)	78	Jiang et al. (2022)	138	Ratajczak et al. (2023)
19	McArthur et al. (2018)	79	Turjo et al. (2022)	139	Choi and Lee (2023)
20	Carreira et al. (2018)	80	Çetin et al. (2022)	140	Almufarrej and Erfani (2023)
21	Kamari et al. (2018)	81	Wang et al. (2022)	141	Pal et al. (2023)
22	Barazzetti (2018)	82	Dobrucali et al. (2022)	142	Jia et al. (2023)
23	Hu et al. (2019)	83	Meschini et al. (2022)	143	Urbieta et al. (2023)
24	Juszczyk et al. (2019)	84	Scherz et al. (2022)	144	Dou et al. (2023)
25	Petrova et al. (2019)	85	Li et al. (2022)	145	Zheng and Fischer (2023)
26	Cheng and Chang (2019)	86	Sun and Liu (2022)	146	Shao et al. (2023)
27	Sha et al. (2019)	87	Cai (2022)	147	Yang and Xia (2023)
28	Livshits et al. (2019)	88	Mahmudnia et al. (2022)	148	Valery et al. (2023)
29	Bianconi et al. (2019)	89	Hajirasouli et al. (2022)	149	Jradi et al. (2023)
30	Bongiorno et al. (2019)	90	Yu et al. (2022)	150	Arsiwala et al. (2023)
31	Acharya et al. (2019)	91	Xu et al. (2022)	151	Rafsanjani and Nabizadeh (2023)
32	Karan and Asadi (2019)	92	Elghaish et al. (2022)	152	Erisen (2023)
33	Zhou et al. (2019)	93	Leon-Garza et al (2022)	153	Ureña-Pliego et al. (2023)
34	Dawood et al. (2019)	94	Doukari et al. (2022)	154	Billi et al. (2023)
35	Lu et al (2019)	95	Wang et al. (2022)	155	Peiman et al. (2023)
36	Boje et al. (2020)	96	Rodrigues et al. (2022)	156	Kim et al. (2023)
37	Doukari and Greenwood (2020)	97	Farghaly et al (2022)	157	Alzara et al. (2023)
38	Döllner (2020)	98	Kanyilmaz et al. (2022)	158	Luo et al. (2023)
39	Bienvenido-Huertas et al (2020)	99	Xia and Gong (2022)	159	Haznedar et al. (2023)
40	Soman and Whyte (2020)	100	Lee et al. (2022)	160	Fenz et al. (2023)
41	An et al. (2020)	101	Singh et al. (2022)	161	Galera-Zarco and Floros (2023)
42	Sacks et al. (2020)	102	Sampaio et al. (2022)	162	Trzeciak and Brilakis (2023)
43	Novembri and Rossini (2020)	103	Zabin et al. (2022)	163	Akomea-Frimpong et al. (2023)
44	Hetemi et al. (2020)	104	Garcia-Gago et al. (2022)	164	Chen et al. (2023)
45	You and Fang (2020)	105	Igwe et al. (2022)	165	Wang et al. (2023)
46	Sacks et al. (2020)	106	Shahzad et al. (2022)	166	Lin et al. (2023)
47	Bloch and Sacks (2020)	107	Sun and Kim (2022)	167	Zwaag et al. (2023)
48	Arashpour et al. (2020)	108	Xu et al. (2022)	168	Yang and Mao (2023)
49	Tak et al. (2020)	109	Lin et al. (2022)	169	Wu et al. (2023)
50	Eber (2020)	110	Baduge et al. (2022)	170	Abouelaziz and Jouane (2023)
51	Hsu et al. (2020)	111	Wang et al. (2022)	171	Bazzan et al. (2023)
52	Frías et al. (2020)	112	Zhang et al. (2022)	172	Hellenborn et al. (2023)
53	Darko et al. (2020)	113	Villa et al. (2022)	173	Hariri-Ardebili et al. (2023)
54	Marzouk and Zaher (2020)	114	Collins et al. (2022)	174	Na et al. (2023)
55	Caterino et al. (2021)	115	Doukari et al. (2022)	175	Rampini and Cecconi (2023)
56	Pan and Zhang (2021)	116	Frías et al. (2022)	176	Chen et al. (2023)
57	Mulero-Palencia et al. (2021)	117	Megahed and Hassan (2022)	177	Mohanta and Das (2023)
58	Ma et al. (2021)	118	Herrera-Martín et al. (2022)	178	Mohamed and Marzouk (2023)
59	Karan et al. (2021)	119	Onososen and Musonda (2022)	179	Joshi (2023)
60	Ilyas et al. (2021)	120	Lehtola et al. (2022)	180	Brilakis et al. (2010)

Id	Source	Id	Source	Id	Source
181	Golparvar-Fard et al. (2015)	241	Hosamo et al. (2022)	301	P Liu et al. (2023)
182	Zhang and El-Gohary (2016)	242	Kanna et al. (2022)	302	Lin et al. (2023)
183	Han and Golparvar-Fard (2017)	243	Tang et al. (2022)	303	Yu et al. (2023)
184	Tixier et al. (2017)	244	B Wang et al (2022)	304	Abdulfattah et al. (2023)
185	Vandecasteele et al. (2017)	245	Hou et al. (2022)	305	Cui et al. (2023)
186	Shi and O'Brien (2018)	246	X Wu et al. (2022)	306	Han et al. (2023)
187	Koo and Shin (2018)	247	Pan and Zhang (2022)	307	Kayhani et al. (2023)
188	Chen et al. (2018)	248	Y Zhou et al. (2022)	308	Artopoulos et al. (2023)
189	Lei et al. (2019)	249	Ma and Leite (2022)	309	Xie et al. (2023)
190	Baeka et al. (2019)	250	Feist et al. (2022)	310	Kiavarz et al. (2023)
191	Koo et al. (2019)	251	Z Wang et al. (2022)	311	Basu et al. (2023)
192	Chen et al. (2019)	252	Mousavi et al. (2022)	312	Hsieh and Ruan (2023)
193	Braun and Borrmann (2019)	253	Poux et al. (2022)	313	Zhang and El-Gohary (2023)
194	Jung and Lee (2019)	254	Emunds et al. (2022)	314	Shen and Pan (2023)
195	Lin and Huang (2019)	255	H Wang et al. (2022)	315	Xiang and Rashidi (2023)
196	Sakhakarmi et al. (2019)	256	H Hosamo et al. (2022)	316	Huang and Liang (2023)
197	Wei and Akinci (2019)	257	Mokhtari et al. (2022)	317	H Kiavarz et al. (2023)
198	Song et al. (2020)	258	J Cheng et al. (2022)	318	S Park et al. (2023)
199	Angah and Chen (2020)	259	Shon et al. (2022)	319	Su et al. (2023)
200	M.M. Singh et al. (2020)	260	M Singh et al. (2022)	320	Grandio et al. (2023)
201	Quinn et al. (2020)	261	Buruzs et al. (2022)	321	Wong et al. (2023)
202	Bouabdallaoui et al. (2020)	262	Fazeli et al. (2022)	322	Chen et al. (2023)
203	Huang and Hsieh (2020)	263	Abdelrahman et al. (2022)	323	Shu et al. (2023)
204	Braun et al. (2020)	264	Tan et al. (2022)	324	Wang and Gan (2023)
205	Czerniawski and Leite (2020)	265	Banihashemi et al. (2022)		
206	Pierdicca et al. (2020)	266	Wei and Akinci (2022)		
207	Ma et al. (2020)	267	Wusu et al. (2022)		
208	Y. Zhao et al. (2020)	268	W Wang et al. (2022)		
209	Liu et al. (2020)	269	Geyter et al. (2022)		
210	S. Zeng et al. (2020)	270	Soh et al. (2022)		
211	Pan and Zhang (2020)	271	Matthews et al. (2022)		
212	Cheng et al. (2020)	272	Zhai et al. (2022)		
213	Rahimian et al. (2020)	273	Kim et al. (2022)		
214	Boonstra et al. (2020)	274	Xiao et al. (2022)		
215	Ma et al. (2021)	275	Gao et al. (2022)		
216	Ma and Pan (2021)	276	B Wang et al. (2022)		
217	Su et al. (2021)	277	Watfa et al. (2022)		
218	Martínez-Rocamora et al. (2021)	278	Xu et al. (2022)		
219	Hou et al. (2021)	279	Y Cheng et al. (2022)		
220	Y. Zhao et al. (2021)	280	Shoar et al. (2022)		
221	Chow et al. (2021)	281	T Wang et al. (2022)		
222	Perez-Perez et al. (2021)	282	Zhang and Zou (2023)		
223	Yin et al. (2021)	283	Park and Yun (2023)		
224	Kim and Kim (2021)	284	Forth et al. (2023)		
225	Ryu et al. (2021)	285	Sobhkhiz and El-Diraby (2023)		
226	Chuang and Sung (2021)	286	Tang et al. (2023)		
227	Koo et al. (2021)	287	Wei et al. (2023)		
228	Cheng et al. (2021)	288	Barkokebas et al. (2023)		
229	B Koo et al. (2021)	289	Korus et al. (2023)		
230	Zhou and El-Gohary (2021)	290	Hassaan et al. (2023)		
231	Croce et al. (2021)	291	Liu et al. (2023)		
232	Hong et al. (2021)	292	Saini et al. (2023)		
233	Czerniawski et al. (2021)	293	Martens et al. (2023)		
234	Wu et al. (2021)	294	Ying et al. (2023)		
235	Y Zhou et al. (2021)	295	Park et al. (2023)		
236	C Yin et al. (2021)	296	L Wang et al. (2023)		
237	Wang et al. (2021)	297	Chen and Xue (2023)		
238	C Wu et al. (2021)	298	Wang and Chang (2023)		
239	Sanhudo et al. (2021)	299	Ghorbany et al. (2023)		
240	H Wu et al. (2021)	300	Kellner et al. (2023)		

Access to sample metadata:
https://drive.google.com/drive/folders/1NYtiw9NoWBekzZ75og90VoslPJoiiopn?usp=sharing

APPENDIX B – FINANCIAL SUPPORT

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.