POZNAN UNIVERSITY OF TECHNOLOGY

FACULTY OF ARCHITECTURE

Constructing semi-automated buildings' energy loads model to retrofit built heritage by using a Data-driven model and computer vision

Author: mgr inż. arch. Hassan Bazazzadeh

Supervisor: dr hab. inż. arch. Adam Nadolny, prof PP

A dissertation submitted to attain the degree of

DOCTOR OF PHILOSOPHY In

ARCHITECTURE AND URBAN PLANNING

Poznań 2023

Abstract

In the face of escalating concerns regarding climate change and environmental sustainability, the imperative to monitor and regulate energy consumption has gained paramount importance. Buildings, being significant contributors to overall energy usage, warrant focused attention, especially when confronting the complexities of heritage protection and assessing energy performance through non-destructive ways in existing structures. This research endeavors to tackle these challenges through the introduction of an innovative and user-friendly approach for estimating energy performance in heritage buildings, utilizing readily accessible tools: smartphones and data-driven algorithms. By harnessing the computational power of smartphones and deploying machine learning algorithms to capture essential geometrical attributes, this novel workflow showcases its capacity to accurately estimate a building's energy performance. This practical solution eliminates the necessity for costly computational resources and specialized expertise, making energy performance assessments more accessible and feasible while protecting heritage values. The research findings underscore the efficacy of this method as a part of retrofit measures for heritage buildings, even when using limited and fundamental geometric data, thereby enhancing the accessibility and utility of energy performance evaluations. This workflow holds significant promise for a diverse range of stakeholders, including researchers, architects, property owners, and government agencies. It empowers them with real-time, precise insights into the energy performance of existing structures. Consequently, this research constitutes a pivotal stride toward bolstering sustainable energy management practices and furnishes a tangible avenue for ameliorating the adverse environmental impact associated with buildings.
Keywords: Building energy consumption, heritage buildings, data-driven methods, Energy Audit, Poland

Abstrakt

W obliczu narastających obaw związanych ze zmianami klimatycznymi i zrównoważonym rozwojem środowiskowym, konieczność monitorowania i regulowania zużycia energii nabrała kluczowego znaczenia. Budynki, będące znaczącym źródłem ogólnego zużycia energii, wymagają szczególnej uwagi, zwłaszcza przy konfrontacji z złożonościami ochrony dziedzictwa i oceniania wydajności energetycznej w istniejących strukturach w sposób nieniszczący. Niniejsze badania mają na celu rozwiązanie tych wyzwań poprzez wprowadzenie innowacyjnego i przyjaznego dla użytkownika podejścia do szacowania wydajności energetycznej w budynkach zabytkowych, wykorzystując łatwo dostępne narzędzia: smartfony i algorytmy oparte na danych. Poprzez wykorzystanie mocy obliczeniowej smartfonów i zastosowanie algorytmów uczenia maszynowego do przechwytywania kluczowych atrybutów geometrycznych, nowatorski przepływ pracy demonstruje swoją zdolność do dokładnego oszacowania wydajności energetycznej budynku. To praktyczne rozwiązanie eliminuje potrzebę kosztownych zasobów obliczeniowych i specjalistycznej wiedzy, ułatwiając dostępność i wykonalność ocen wydajności energetycznej, jednocześnie chroniąc wartości dziedzictwa. Wyniki badań podkreślają skuteczność tej metody jako części działań modernizacyjnych w budynkach zabytkowych, nawet przy wykorzystaniu ograniczonych i podstawowych danych geometrycznych, zwiększając tym samym dostępność i użyteczność ocen wydajności energetycznej. Ten przepływ pracy ma znaczący potencjał dla szerokiego spektrum zainteresowanych stron, w tym badaczy, architektów, właścicieli nieruchomości i agencji rządowych. Umożliwia im on uzyskanie w czasie rzeczywistym dokładnych informacji na temat wydajności energetycznej istniejących struktur. W konsekwencji niniejsze badania stanowią kluczowy krok w kierunku wzmocnienia praktyk zarządzania zrównoważoną energią i dostarczają namacalnej drogi do zmniejszenia negatywnego wpływu budynków na środowisko.

Słowa kluczowe: Zużycie energii w budynkach, budynki zabytkowe, metody oparte na danych, Audyt energetyczny, Polska

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbr	Description	Abbr	Description
EAD	Energy Application domain	BES	Building energy Simulation
TMY	Typical Meteorological Year	BEM	Building Energy Modeling
TRY	Typical Reference Year	DSY	Design Summer Year
IDF	Input Data File	CSV	Comma-separated values
GCM	Global Circulation Models	UBEM	Urban Building Energy Modeling
SERG	Sustainable Energy Research Group	HadCM3	Hadley Centre Coupled Model, version 3
AR3	Third Assessment Report	IPCC	Intergovernmental Panel on Climate Change
AR5	Fifth Assessment Report	FS	Finkelstein-Schafer
EDA	Exploratory Data Analysis	PNNL	Pacific Northwest National Laboratory
EU	European Commission	ECSO	European Construction Sector Observatory
DHW	Domestic Hot Water	NECP	National Energy and Climate Plan
WM	Water Management	KAPE	National Energy Conservation Agency
PEB	Plus Energy Buildings	NFEP	National Fund for Environmental Protection
PH	Passive House	EPBD	Energy Performance of Buildings Directive
PV	Photovoltaic	RES	Renewable Energy Systems
EPC	Energy Performance Certificates	ETICS	External Thermal Insulation Construction Systems
EP	Energy Performance	HVAC	Heating, Ventilation, and Air Conditioning
NEP	Natural Environment Policy	EPB	Energy Performance of Buildings
BPS	building performance simulations	TFC	Total Final Consumption
GHG	Greenhouse Gas	BPIE	Buildings Performance Institute Europe
HRM3	Hadley Regional Model 3	SRES	Special Report Emissions Scenarios
EPI	Energy Performance Indicator	EUI	Energy Use Intensity
RF	Random Forest	ML	Machine Learning
ANN	Artificial Neural Networks	LSTM	Long Short-Term Memory networks
MLP	Multilayer Perceptrons	BPNN	Back Propagation Neural Networks
SVR	Support Vector Regression	MLR	Multiple Linear Regression
SVM	Support Vector Machines	XGB	Extreme Gradient Boosting
RF	Random Forests	RNN	Recurrent Neural Networks
FFN	Feed Forward Networks	CNN	Convolutional Neural Networks
RBFN	Radial Basis Function Networks	Bi-LSTM	Bidirectional LSTM
DSF	Double-Skin Facade	PSO	Particle Swarm Optimization
GA	Genetic Algorithms	BP	Back Propagation
RMSE	Root Mean Square Error	ESN	Echo State Networks
WT	Wavelet Transform	BECP	Building Energy Consumption Prediction
PLS	Partial Least Squares	WIO	wolf-Inspired Optimization
WD	Wavelet Decomposition	MNR	Multiple Nonlinear Regression
PB	Pattern-Based	Std	Standard Deviation
RC	Relative Compactness	CL	Characteristic Length
V	Volume	R	Rotation
GA	Glazing Area	EUI	Energy Use Intensity
WWR	Window To Wall Ratio	L	Length
H	Height	RA	Roof Area
MSE	Mean Squared Error	MAE	Mean Absolute Error

Author Contribution Statement

This thesis is an original work by the author and has not been previously published in its entirety. Small portions of this work may have appeared in the following publications, for which the author holds the copyright:

- Bazazzadeh, H., Pilechiha, P., Nadolny, A., Mahdavinejad, M., & Hashemi safaei, S. S. (2021). The impact assessment of climate change on building energy consumption in Poland. Energies, 14(14), 4084.
- Bazazzadeh, H., Nadolny, A., & Safaei, S. S. H. (2021). Climate change and building energy consumption: A review of the impact of weather parameters influenced by climate change on household heating and cooling demands of buildings. European Journal of Sustainable Development, 10(2), 1-1.

Except for the aforementioned publications, no part of this thesis has been published or submitted for publication elsewhere. The research, analysis, and writing presented herein were conducted solely by the author as part of doctoral studies at Faculty of Architecture and Urban Planning at Poznan University of Technology.

Acknowledgements

I would like to extend my deepest appreciation to my dear wife, Sara, whose support throughout this journey has been invaluable. Her encouragement and belief in me were a constant source of strength and inspiration.
I must also express profound gratitude to my supervisor, whose guidance and expertise were instrumental in shaping this work. The endless hours spent discussing the nuances of this research have been invaluable, and I couldn't have asked for a more supportive and insightful mentor.
My heartfelt thanks also go to my parents, who have always been my biggest cheerleaders. Their unwavering faith in me and sacrifices on my behalf have made all the difference in my academic journey. I am forever indebted to them for their love and support.
Lastly, I want to thank my friends, who have been my emotional backbone, offering moral support and enriching my life in countless ways. Their camaraderie made even the toughest of times seem manageable.

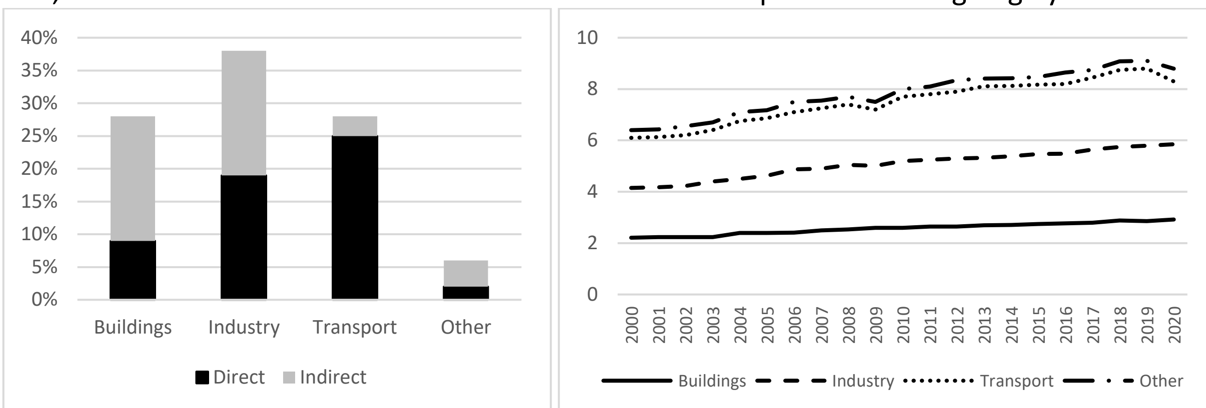
Chapter 1: Research outline

1.1. Abstract

This abstract provides an overview of the first chapter of the thesis, outlining the topics covered and the objectives of the research. In this chapter, the introduction sets the stage by discussing the global challenges of climate change, urban sustainability, and the need for energy efficiency in buildings. The Sustainable Development Goals and the European Union's long-term goals for reducing greenhouse gas emissions are introduced as driving forces for action. The problem statement emphasizes the importance of energy audits in existing buildings and highlights the challenges associated with data availability, invasive inspections, financial constraints, complex building systems, and the absence of standardized procedures. The goals of the research are then presented, focusing on facilitating the energy audit process, exploring improvements within traditional methodologies, and proposing a surrogate data-driven method. The research questions and hypotheses are outlined to guide the investigation and address the identified challenges. Finally, the methodology is briefly described, encompassing literature review, case studies, the use of LIDAR technology for 3D modeling, data generation, climate change considerations, data-driven model development, and the conclusion of the chapter. Through this research, the aim is to contribute to enhanced energy efficiency, streamlined audit procedures, and sustainable building retrofitting practices for a more sustainable future.

1.2. Introduction

Addressing today's global issues including climate change, resource depletion, and growing urbanization needs a strong insight about urban sustainability [1]. It describes a city's capacity to meet the requirements of present needs and also considering future generations and preserving the harmony of economic, social, and environmental concerns [2]. With the increasing average earth's surface temperate as a result of climate change, countries all around the world are trying to minimize CO_2 emissions and energy use [3]. Thus, the EU has set long-term goals to cut GHG emissions by 80-95% compared to 1990 levels by the year 2050 [4]. The urgent need for urban infrastructures including energy systems as well as housing has significantly multiplied due to the growth of the urban population as the result of economic and industrial development. Together with improving life quality, these developments have increased greenhouse gas emissions [5, 6]. A strong foundation for global collaboration is offered by the Sustainable Development Goals (SDGs), which were initially endorsed by the UNGA in 2015 to ensure a sustainable future for the planet. The 17 SDGs and the 169 targets that they contain, which make up the core of "Agenda 2030," provide a course for eradicating extreme poverty, combating injustice and inequality, and preserving the environment on Earth [7, 8]. It has been discovered that the future holds promising pathways towards achieving better energy access, improved air quality, and enhanced energy security without endangering our climate. Through exploring various resource, technology, and policy combinations, we have the ability to reach these objectives in innovative ways [9]. We must emphasize the development of renewable energy sources and stop relying on fossil fuels in order to reduce irreversible environmental damage. This entails a change in the way energy is produced away from fossil fuels, as well as attempts to increase energy efficiency and lower demand generally [10-12]. However, coal continues to play a substantial role in Poland's energy industry. The Polish government has agreed to speed up the process of phasing out coal and gradually close all coal mines by 2049, thanks to pressure from the European Union [13]. As International Energy Agency has stated the world's consumption can be broadly categorized into various sectors (See figure 1.1) including buildings, transport, industry and other areas such as agriculture, forestry, and fishing [14]. The total energy consumption across these sectors amounted to 9.1 Gtoe in 2019, with their relative contribution to the overall consumption remaining largely stable.



a b

Figure 1. 1. a. High global energy consumption sector, b. CO2 emissions by sectors
Adopted from [15-17]

The surge in population growth, urban development, towering infrastructure, and improved building amenities and conveniences, coupled with the growing preference for indoor lifestyles, have collectively contributed to the annual rise in building energy consumption by 1.2% since the turn of the century. This upward trajectory has remained persistent, even during challenging times like the global economic downturn of 2008 and the ongoing COVID-19 pandemic [18]. Forecasts suggest that without robust measures to regulate this trend, energy consumption in buildings will continue to escalate, especially in developing nations where it's gaining prominence [19].

The towering structures that dominate our skylines are notorious culprits when it comes to global energy consumption and CO2 emissions, accounting for a staggering one-third and one-quarter respectively. Shockingly, in some of the world's most energy-intensive nations, buildings consume an even larger percentage of energy, with the EU at 41%, the US at 34% , and Japan at 37% [14]. Recognizing their immense impact on the environment, buildings have been thrust into the spotlight of climate policies as potential drivers of energy efficiency and renewable energy. However, to successfully develop, evaluate, and monitor these policies, access to energy information is crucial, not just for the sector as a whole, but for specific building types and energy services as well [20]. Apart from their substantial contribution to primary energy consumption and greenhouse gas emissions, buildings are expected to play a vital role in promoting energy efficiency and renewable energy generation. Upon closer examination of Figure 1, it becomes apparent that the majority of greenhouse gas (GHG) emissions for building sector are indirect and the environmental impact of buildings varies depending on the emission factors of the energy production processes.

The building sector presents a vast opportunity for mitigating energy consumption and curbing CO2 emissions. Innovative measures such as building envelope enhancements, energy efficiency initiatives, and transitioning to renewable fuel sources for both residential and commercial buildings can contribute to this cause. It is no wonder that a multitude of national and international mandates and regulations have been established to boost the energy efficiency of buildings and facilitate the adoption of renewable energy sources.

Existing buildings account for a large proportion of this consumption. Thus, it is crucial to assess and improve the energy efficiency of existing buildings to reduce energy consumption and decrease the carbon footprint of the building sector. Existing buildings are often not as energy-efficient as newer buildings because they may lack modern insulation, efficient heating and cooling systems, and energy-efficient appliances. As a result, they consume more energy than necessary to maintain a comfortable indoor environment [21]. Retrofitting existing buildings, which involves making modifications to improve energy efficiency, is one solution to reduce the energy consumption of existing buildings. Retrofitting measures can include the installation of energy-efficient lighting, HVAC systems, and insulation, among others. studies have suggested that retrofitting existing buildings can lead to significant energy savings [22]. Additionally, it has been stated that retrofit measurements can be a cost-effective way to reduce energy consumption, with a payback period of less than five years for most measures [23]. The importance of retrofitting existing buildings is not limited to reducing energy consumption and carbon emissions, but it can also bring improved indoor air quality, reduced maintenance costs, increased the value of buildings, and enhanced occupant comfort [24]. The EU energy efficiency directive 2012/27/EU recognizes the existing building stock as the most significant potential sector for energy savings [25]. To achieve this, policymakers try to formulate strategies to encourage cost-effective deep renovations that significantly reduce a building's energy consumption.

To be more precise, fulfilling the ambitious energy and climate goals established by the European Union for 2030 and 2050 depends on the decarbonization of the Polish building sector. This is essential to achieving the aforementioned objectives as well as preserving the health and wellbeing of our communities. Significant economic prospects are also presented by this shift to low-carbon building practices, including the growth of the market for sustainable building materials in Poland and the creation of new jobs in the renewable energy sector. To draft such strategies, it is necessary to analyze the structure and energy consumption of the existing building stock. However, due to the lack of comprehensive data sets, it would be a challenging task. The identification of certain crucial parameters such as building age, function, or floor area is necessary to estimate a building's energy demand. Therefore, every possible strategy in this regard requires to estimate the energy performance of individual buildings accurately. Upon reviewing the relevant literature and recent studies, certain inadequacies have come to light. From the practical perspective the main obstacle is the data collection as in most cases the accurate and proper data availability is rare which prevents in-depth analysis of cases. This gap along with other problems in this process that may discourage owners, municipalities, energy departments or other involving parties to act effectively to address this issue. This research tries to focus on the required data availability challenge and propose a surrogate data-drive method which can be useful for all parties involved in this subject.

1.3. Problem Statement

The primary source of the global warming phenomenon is thought to be greater atmosphere emissions of greenhouse gases, such as carbon dioxide (CO2) [26]. It will take a significant shift in how countries' energy systems are now organized to achieve global warming gas (GHG) emission reduction goals. Together with decarbonizing the energy source, this also entails drastically lowering existing energy usage [27]. As mentioned before, due to the notable share of building stock in the total energy consumption globally, improving the energy efficiency of buildings has become crucial to cutting back on both fossil fuel usage and gas emissions. By increasing the energy performance of buildings in the European Union (EU) by 20%, it is predicted that 60 billion euros will be saved annually [28] and a huge step (Blue color in Figure 1.2) towards to fulfil 2020 and 2030 primary energy consumption targets (38% of the target)

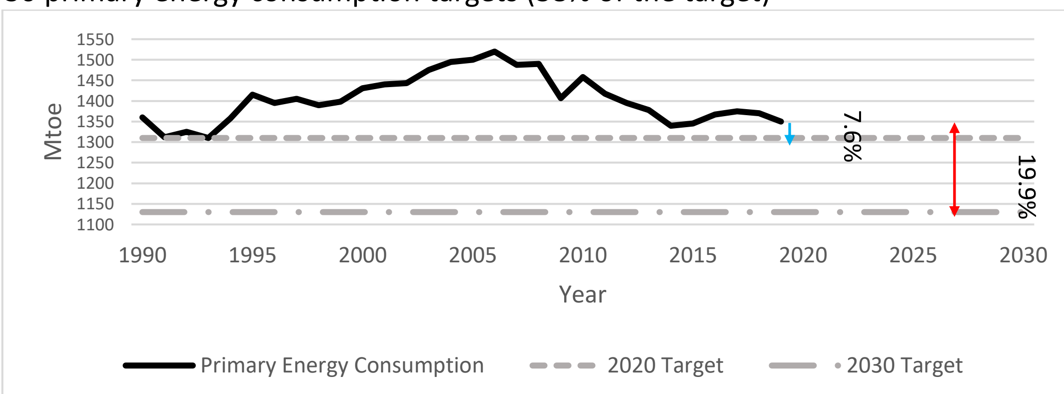


Figure 1. 2. Distance to 2020 and 2030 targets for primary energy consumption
Adopted from [29]

According to the Multi-annual Financial Framework (MFF) for the years 2014 to 2020, the Cohesion Policy budget allocated an astounding €80 billion to Poland over a seven-year period (See figure 1.3), making Poland the largest recipient of EU money. However, compared to the average of the EU, which is 3.9%, only €2.2 billion (or 2.8%), or 2.6%, has been allocated to improving building energy efficiency. The fact that massive amounts of money totaling €27 billion have been promised to Poland by international financial organizations (IFIs), such as the EBRD, EIB, and the World Bank, just 1.3% of which will go toward building restoration, further complicates the situation [30]. A serious problem that requires immediate action is the absence of funding for energy-efficient building modifications.

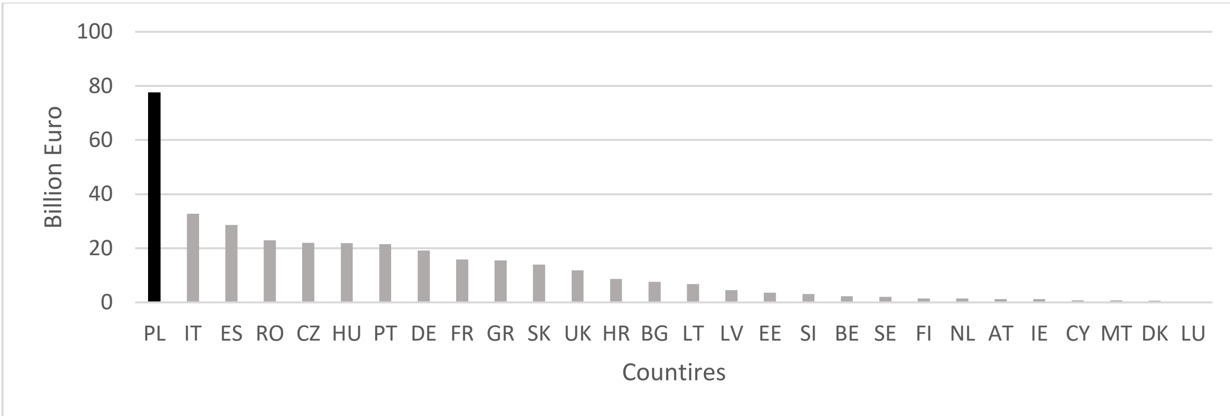


Figure 1. 3. Cohesion policy EU allocation funds 2014-2020

Adopted from [31]

A significant portion of this fund (more than 27 billion Euro) has been allocated in “Infrastructure and Environment” sector. With 33 of the top 50 most polluted cities in Europe (which is 36 recently, see figure 1.4), Poland has a serious and persistent air pollution problem.

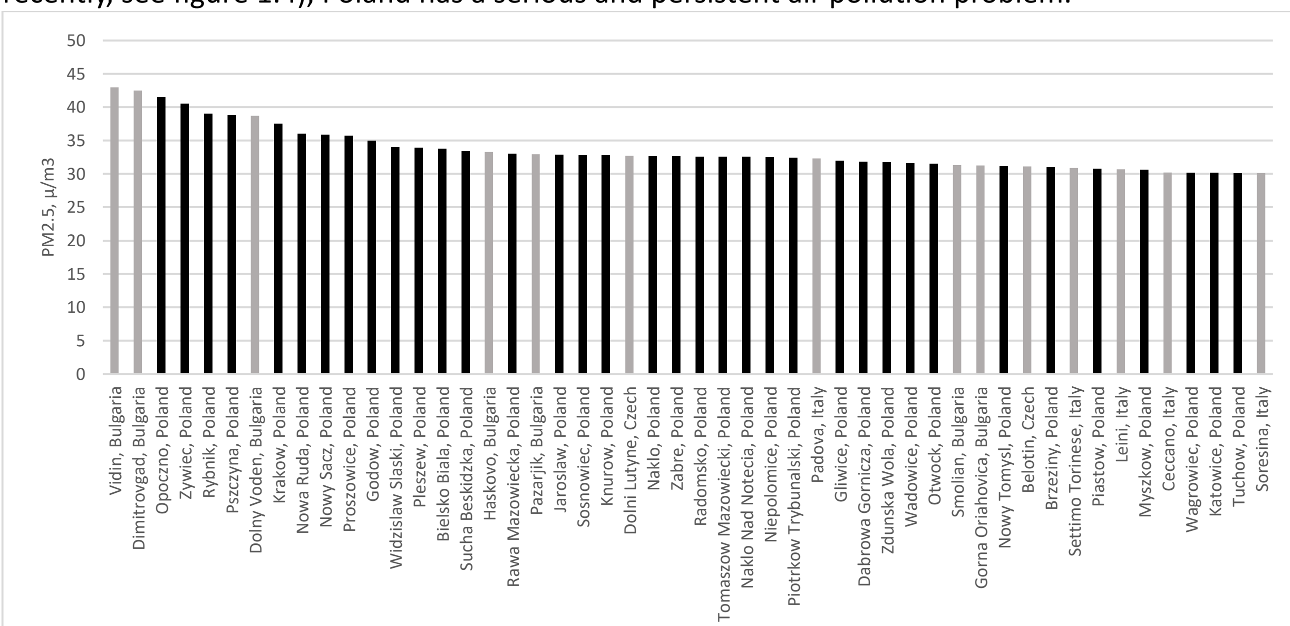


Figure 1. 4. Most polluted cities in Europe

adopted from [32], *black color shows cities in Poland

This disturbing reality is largely due to the inefficient energy use of existing structures, especially single-family homes, and their antiquated coal-fired boilers. Regrettably, Poland has not yet given this activity the proper priority, despite the enormous potential for building renovation to alleviate this problem, improve energy security, and foster citizen well-being. It's time for the nation to understand the value of sustainable building techniques and take concrete steps toward a cleaner, healthier future. An extraordinary yearly expenditure of €5.3 billion would be needed to renovate half of Poland's current building portfolio over the course of the next 20 years. This would dramatically increase the present restoration pace, which is currently less than 1% of floor

area every year, to a healthy 2.5% [33]. Redirecting financing from the European Union and other international financial institutions to improve building energy efficiency is necessary to achieve this ambitious aim. Furthermore, creative financing methods with higher leverage, such securing cash from building owners and other investors, might improve the usage of available funds.
When considering the limitations of improving the building stock, one aspect to take into account is the budget allocated for retrofit measurement. However, alongside financial constraints, there are additional challenges associated with retrofit actions. The process of retrofitting existing buildings is a complex endeavor that encompasses multiple areas of expertise, including built heritage (particularly relevant for historic buildings), energy efficiency, user functionality, material science, and more. Integrating the requirements of each of these domains into the retrofitting process adds to its inherent difficulty and complexity. One particular aspect of building retrofit that demands significant time, labor, and financial resources is the energy audit. Energy audits play a crucial role in driving energy retrofitting initiatives for already established structures, which often account for a substantial portion of energy consumption within urban areas. These audits serve various purposes beyond energy conservation, such as optimizing energy usage, managing costs, and addressing the environmental impacts associated with energy consumption. Therefore, the challenges inherent in retrofitting existing buildings extend beyond the financial limitations of measurement and encompass a wide array of considerations across different disciplines. The inclusion of energy audits as a labor-intensive and costly step underscores their importance in promoting energy retrofitting efforts and the overall improvement of urban building stocks [34].
According to Forbes, the cost associated with conducting a building energy audit at ASHRAE level 3 (refer to figure 1.5) for retrofit measurements is estimated to be approximately $600 [35]. This essential task not only requires a significant financial investment [36] but also demands the expertise of professionals who are capable of conducting precise measurements and thorough analyses as well as expensive computation resources. It is worth noting that completing such an audit can take several weeks to ensure its accuracy and comprehensiveness. From another perspective, the success of the energy audit process heavily relies on obtaining high-quality data encompassing various details, including material specifications, construction age, foundation type, and more [37]. However, for existing buildings, gathering this data can prove to be immensely challenging, and in some cases, nearly impossible, particularly when dealing with historic structures that may have been poorly documented over time. In such instances, conducting inspections becomes necessary, but this poses a critical dilemma, especially when dealing with buildings of historical significance.
Performing inspections to acquire information about the building materials and specific details can potentially cause damage or harm to the integrity of these cherished historic structures [38]. Consequently, in the case of buildings officially recognized as built heritage, this process is strictly prohibited to preserve their historical value and physical integrity. It is clear that conducting a building energy audit, specifically at ASHRAE level 3, entails substantial financial implications. Furthermore, the challenges associated with obtaining accurate data for existing buildings, particularly those with historical significance, highlight the delicate balance between the need for comprehensive information and the preservation of cultural heritage.

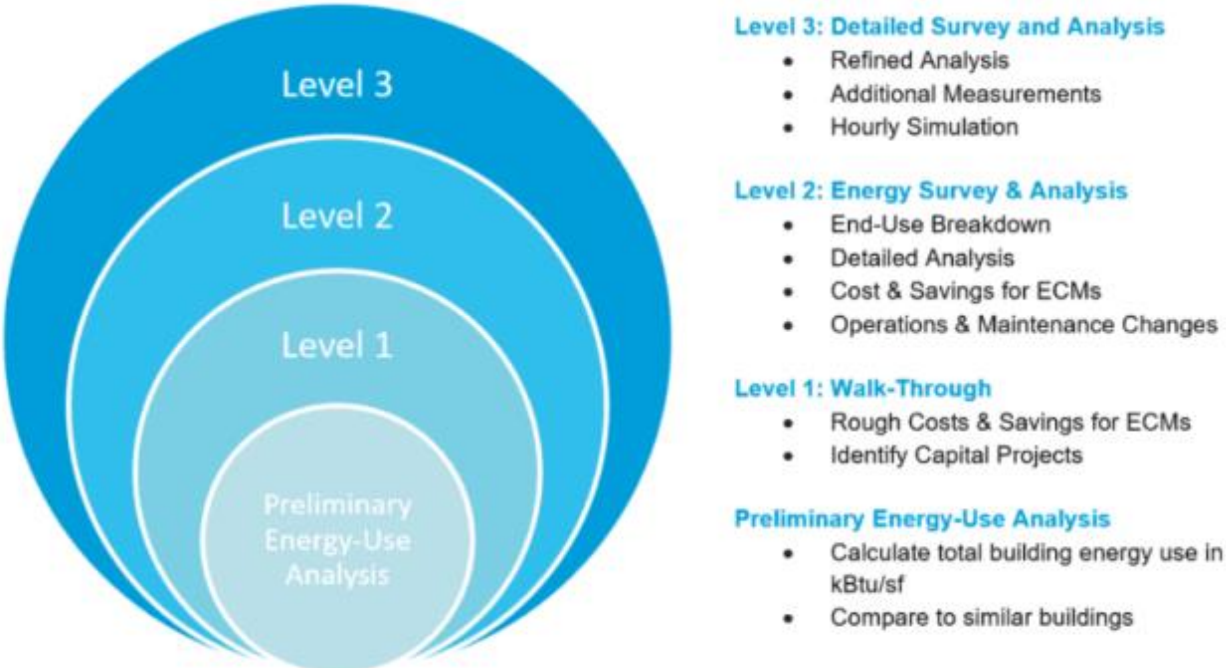


Figure 1. 5. Different levels of Energy audit

1.4. Goals

Therefore, in order to comprehensively address the multifaceted challenges associated with the energy audit process for existing buildings and pave the way for effective retrofitting initiatives, this research endeavors to achieve the following interconnected goals:

1. Facilitate the process of energy audit for existing buildings: The challenge here is to streamline and simplify the energy audit process, making it more accessible and efficient for existing buildings. This involves developing methodologies, guidelines, and tools that can guide auditors through the audit process, from data collection to analysis and reporting. The goal is to create a standardized and user-friendly approach that reduces the complexity and time required for energy audits, ultimately promoting their wider adoption.
2. Check the applicability and accuracy of data-driven methods: With the advancements in data analytics and machine learning, data-driven methods are gaining popularity for energy audits. However, their applicability and accuracy for existing buildings need to be thoroughly assessed. This challenge involves evaluating the performance of data-driven models and algorithms in accurately predicting energy consumption and identifying energy-saving opportunities in existing buildings. The goal is to assess the strengths and limitations of these methods and provide insights into their practical implementation in real-world scenarios.
3. Propose a surrogate data-driven method for energy audit of existing buildings: This challenge focuses on developing a surrogate data-driven method that can overcome the limitations and data requirements associated with traditional energy audits. The goal is to

explore alternative data sources and develop models that can estimate energy performance and identify potential energy-saving measures without the need for extensive data collection or invasive inspections. This surrogate method should leverage readily available data, such as building characteristics, occupancy patterns, weather data, and utility bills, to provide reliable and actionable insights for energy audit purposes.

1.5. Questions and Hypotheses

To achieve the research goals outlined, it is essential to address the following questions, as they serve as critical steppingstones toward attaining the desired outcomes (In figure 1.6). By answering these questions, the research aims to overcome challenges in energy audit processes, explore improvements within traditional methodologies, harness the potential of data-driven methods, and leverage technological advancements to offer surrogate approaches for energy audits in existing buildings. The insights gained from addressing these questions will pave the way for enhanced energy efficiency, streamlined audit procedures, and the promotion of sustainable practices in the realm of building retrofitting.

1. What are the significant challenges and hurdles encountered in the process of conducting comprehensive energy audits for existing buildings, considering factors such as limited data availability, invasive inspections, financial constraints, complex building systems, and the absence of standardized procedures?
2. In the realm of energy audits for existing buildings, what specific domains within the traditional audit process hold potential for improvement and optimization, encompassing aspects such as data collection methodologies, analysis techniques, reporting standards, and the incorporation of real-time monitoring systems?
3. To enhance the efficacy and efficiency of energy audits for existing buildings, how can data-driven methodologies, leveraging advanced analytics, machine learning algorithms, and modeling techniques, contribute to the refinement and streamlining of the audit process, enabling accurate predictions of energy consumption patterns, identification of energy-saving opportunities, and the customization of recommendations based on diverse building characteristics?
4. By synergistically integrating data-driven techniques with cutting-edge technological advancements, such as remote sensing, Internet of Things (IoT) devices, smart metering, and building energy modeling, how can a surrogate method for energy audits in existing buildings be developed, circumventing the need for invasive inspections and extensive data collection, while still ensuring robust energy performance evaluation, anomaly detection, and the simulation of retrofit measures?

In pursuit of addressing the aforementioned questions, this research puts forth several hypotheses that will be rigorously tested throughout the investigative process. These hypotheses serve as guiding principles, aiming to shed light on specific aspects related to energy audits for existing buildings. By subjecting these hypotheses to empirical examination, the research endeavors to contribute to the body of knowledge in the field and uncover valuable insights. The

outcomes of these hypothesis tests will provide a solid foundation for evidence-based conclusions and recommendations, ultimately driving advancements in energy auditing methodologies, data-driven approaches, and the development of surrogate methods for efficient energy audits in existing buildings.
For the first cluster of questions:

- Limited data availability: Gathering accurate and comprehensive data about existing buildings, including historical energy consumption, building characteristics, and occupancy patterns, can be difficult and time-consuming.
- Invasive inspections: Inspecting existing buildings to gather detailed information about materials, construction, and equipment can be invasive and potentially damaging, especially for buildings with historical significance.
- Cost and resource-intensive: Traditional energy audits often require significant financial resources, time, and expertise to conduct comprehensive measurements, analysis, and reporting.
- Complex building systems: Existing buildings may have complex and interconnected energy systems, making it challenging to identify energy-saving opportunities and assess the performance of individual components.
- Lack of standardized processes: The absence of standardized guidelines and methodologies for energy audits can lead to inconsistency in audit procedures and results.

For the second cluster:

- Streamlining data collection: Leveraging digital technologies, automation, and remote sensing techniques can simplify and expedite data collection processes, reducing the time and effort required to gather essential building and energy consumption data.
- Enhancing data analysis capabilities: Utilizing advanced data analytics, machine learning algorithms, and modeling techniques can enable more accurate and insightful analysis of energy performance, identifying specific areas for improvement and providing customized recommendations.
- Incorporating real-time monitoring: Integrating real-time energy monitoring systems can provide continuous data on energy usage, allowing for better understanding of patterns and enabling proactive energy management strategies.
- Standardizing audit procedures: Developing standardized guidelines and protocols for energy audits can ensure consistency, reliability, and comparability of results across different buildings and auditors.
- Incorporating impacts of climate change: on building performance by considering future climate

Finally, for the last cluster:

- Hypothesis: Using smartphones with technologies such as image processing, computer vision, and LIDAR scanners to gather information from building spaces. It is hypothesized that utilizing smartphones equipped with advanced technologies like image processing, computer vision, and LIDAR scanners can effectively capture spatial information from building spaces.

- Coupling the information gathered from smartphones with simulation data. It is hypothesized that by combining the data obtained from smartphones with simulation data, a comprehensive model can be developed that integrates real-life observations with simulated scenarios.
- Calibrating the simulation data with real-life cases. It is hypothesized that by calibrating the simulation data with real-life cases, the accuracy and relevance of the model can be further enhanced, ensuring its reliability and applicability in practical energy audit applications.
- Mixing all the components together to create a new model that offers accuracy and ease of use for non-technical users. It is hypothesized that by integrating smartphones, simulation data, and calibration techniques, a novel and user-friendly model can be developed. This model is expected to provide improved accuracy in energy audits and be easily accessible for non-technical users, simplifying the energy audit process for existing buildings.

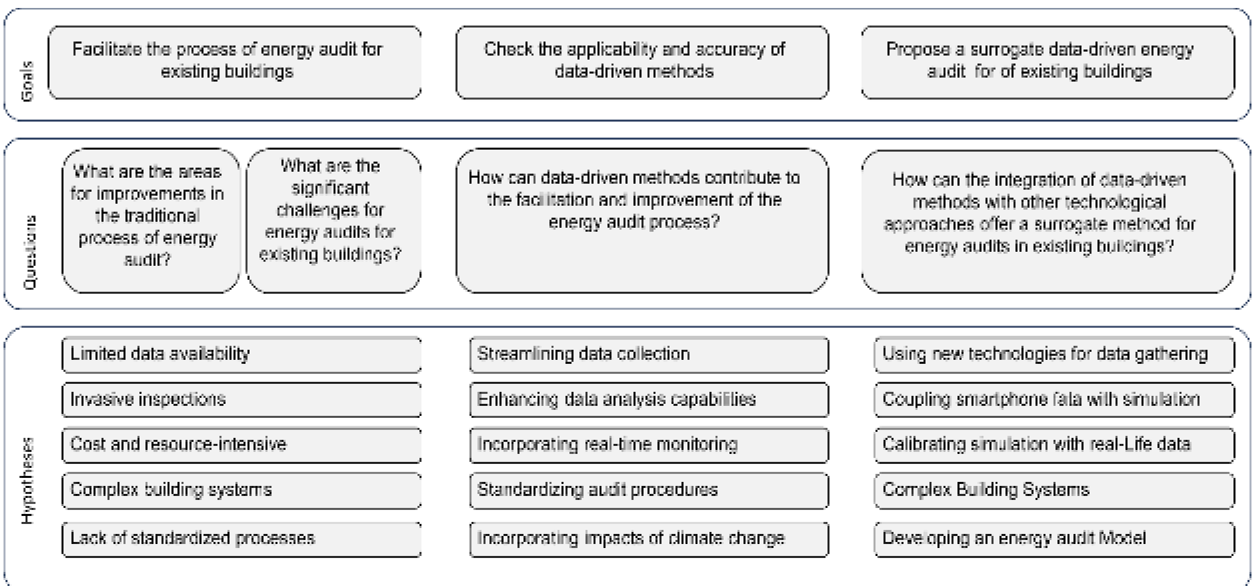


Figure 1. 6. goals, questions, and hypothesis of the research

Chapter 2: Methodology

2.1. Abstract

The methodology section of this research provides a comprehensive overview and framework of the current work in 7 main parts. Initially, the study begins with a foundational background that introduces existing technologies and paradigms, setting the stage for the methods used in this research. The primary focus is on the detailed process involved in 3D Model Generation, employing techniques such as image calibration, 3D point cloud generation, and alternative methods for model optimization. Subsequently, the paper delves into the systematic approach for Bigdata generation within the realm of architecture and building science. The workflow outlined ranges from initial geometry creation to intricate construction details. Special attention is given to EAD additions and energy simulation metrics. The section also introduces automation techniques that significantly reduce manual input, thereby improving efficiency and reliability. To address the pressing issue of climate change, this methodology incorporates it as a critical consideration. Data-driven methods are used to provide a scientific basis for design decisions and environmental impact assessments. Finally, an integrated workflow is discussed, which aims to combine all the aforementioned methodologies into a single, cohesive system for streamlined project execution. This approach provides a robust framework for professionals, researchers, and policy-makers aiming to create sustainable, efficient, and technologically advanced built environments.

2.2. Introduction

In order to reach the goals of this research, it is crucial to acknowledge that they lie at the intersection of various disciplines and fields, each playing a significant role in addressing the challenges and complexities of energy audits for existing buildings. This study aims to contribute to the advancement of energy-efficient buildings, sustainable development, artificial intelligence, automation in the construction and energy industry, and climate change mitigation.

By exploring the realm of energy-efficient buildings, this research seeks to uncover innovative strategies and technologies that can optimize energy consumption, reduce greenhouse gas emissions, and improve the overall sustainability of the built environment. Sustainable development principles will guide the investigation, ensuring that the proposed solutions align with the economic, social, and environmental dimensions of sustainability. The integration of artificial intelligence (AI) and advanced data analytics is another key aspect of this research. By harnessing the power of AI algorithms, machine learning techniques, and modeling methodologies, the study aims to develop data-driven models that can accurately predict energy consumption patterns, identify energy-saving opportunities, and provide customized recommendations tailored to the specific characteristics of each building. This intersection of AI and energy auditing holds great potential for streamlining processes, enhancing accuracy, and enabling more effective decision-making in building retrofit projects. Automation in the construction and energy industry is yet another critical area of focus. By leveraging technological advancements, such as digital automation, robotics, and smart systems, this research aims to streamline energy audit procedures, data collection processes, and analysis methodologies. Automating repetitive tasks and leveraging digital tools can improve efficiency, reduce human error, and accelerate the overall energy auditing process. Lastly, climate change considerations are woven throughout the research methodology. Given the urgent need to address the impacts of climate change, this study will incorporate future climate scenarios and adaptation strategies into the analysis. By assessing the resilience and sustainability of proposed retrofit measures under different climate conditions, the research aims to contribute to building resilience, adaptability, and long-term energy efficiency.

To achieve the ambitious goals set forth, a sophisticated methodology has been designed. The research will progress through a series of interconnected steps, including background research and case study analysis, advanced data collection techniques, climate change considerations, data-driven modeling, and comprehensive analysis of results. This systematic approach ensures a comprehensive investigation, informed decision-making, and the development of practical recommendations that can drive positive change in the building sector and contribute to global efforts towards a greener and more resilient future.

By integrating knowledge, methodologies, and expertise from various fields, this research strives to push the boundaries of energy auditing practices, promote sustainable retrofitting strategies, and foster collaboration among stakeholders in pursuit of a more energy-efficient and environmentally conscious built environment.

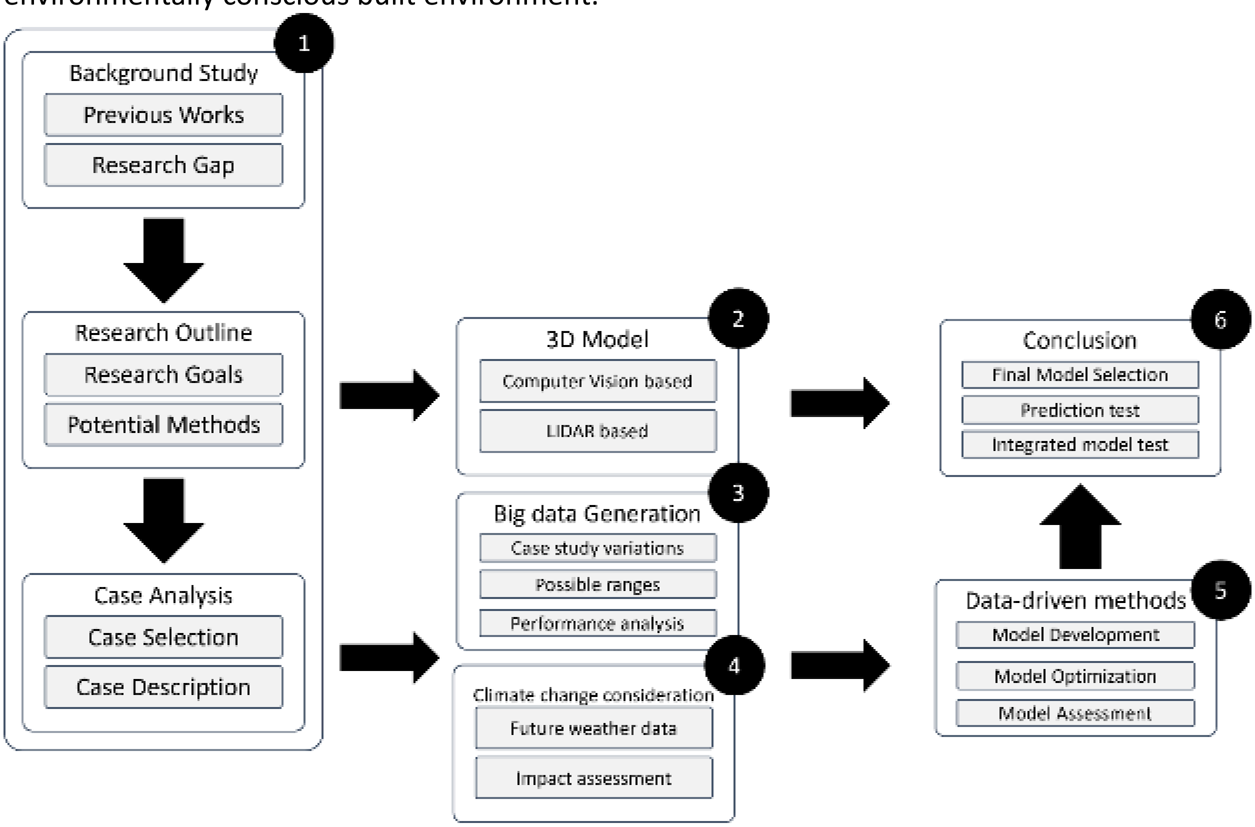


Figure 2. 1. Research Steps

The methodology of this research consists of six main steps, each contributing to the comprehensive investigation of energy audits for existing buildings (see Figure 2.1.).

- The first step involves conducting a background study and comprehensive case analysis to establish a solid theoretical foundation and gain insights into current practices, challenges, and advancements in the field of energy auditing. This study will provide a comprehensive understanding of existing methodologies and highlight areas for improvement.
- In the second step, LIDAR technology will be employed to create a simplified 3D model of the buildings under study. This technology enables precise measurements and captures detailed spatial information, allowing for a thorough analysis of building characteristics such as geometry, surface area, and volume.
- The third step focuses on generating a large and diverse dataset, encompassing various sources of information such as historical energy consumption data, weather data, occupant behavior, and building systems performance. This extensive dataset will serve as the basis for analysis and modeling, enabling a comprehensive evaluation of energy performance and the identification of potential energy-saving opportunities.
- The fourth step involves considering the impacts of climate change on building performance. Future climate scenarios and potential adaptation strategies will be integrated into the analysis, ensuring that the proposed energy retrofit measures are resilient and sustainable in the face of changing environmental conditions.
- In the fifth step, a data-driven model will be developed using advanced analytics, machine learning algorithms, and modeling techniques. This model will leverage the generated dataset to accurately predict energy consumption patterns, evaluate the effectiveness of different retrofit strategies, and provide customized recommendations tailored to the specific characteristics of each building.
- Finally, the research concludes with a comprehensive analysis of the results obtained from the data-driven model. The findings will be evaluated, interpreted, and compared with existing methodologies and industry standards. The research will culminate in a conclusive summary, highlighting the key insights, implications, and recommendations derived from the study.

By following this structured methodology, the research aims to contribute to the advancement of energy auditing practices for existing buildings, foster sustainable retrofitting strategies, and pave the way for more energy-efficient and environmentally conscious building practices. In the following chapter each of these steps will be discussed in detail.

2.3. Background Study

Within the initial phase of this research project, an extensive examination was undertaken in the realm of building energy analysis. This diligent exploration was pursued with the intention of gaining a profound understanding of the capabilities and limitations within this field, thereby establishing a comprehensive foundation for further investigation. By undertaking this study, the researcher aimed to acquire a holistic comprehension of the interconnected subjects within this discipline, thus facilitating the identification of existing gaps and paving the way for the advancement of knowledge within the chosen field of study [39]. This preliminary segment encompassed a comprehensive review of the general field of building energy analysis, delving into its theoretical underpinnings, methodologies, and practical applications. The primary objective was to assess the potential and shortcomings of this area, enabling the researcher to form a comprehensive overview of the topic. By meticulously examining the existing literature, seminal studies, and innovative approaches within the field, the researcher developed a solid foundation upon which the subsequent investigations could be built.

Conducting such an in-depth study proved instrumental in establishing a strong knowledge base and fostering a deeper understanding of the intricacies associated with building energy analysis. By meticulously scrutinizing various aspects, including energy consumption patterns, building performance evaluations, energy modeling techniques, and sustainability considerations, the researcher gained invaluable insights into the subject matter. This comprehensive comprehension laid the groundwork for identifying areas of research that required further exploration and investigation.

Moreover, this preliminary examination enabled the researcher to recognize and comprehend the existing gaps within the field of building energy analysis. By identifying areas where knowledge and understanding were lacking, the researcher could discern the key research questions and formulate hypotheses that would contribute to advancing the field. This initial exploration not only facilitated the development of a robust research framework but also ensured that the subsequent investigations would address critical gaps and provide meaningful contributions to the discipline [40]. By embarking on this preliminary journey, the researcher positioned themselves strategically to embark on the subsequent stages of the research project. Armed with a comprehensive overview of the topic and a keen awareness of the existing knowledge gaps, the researcher was well-equipped to design and execute rigorous empirical studies and innovative theoretical inquiries. This thorough grounding in the field of building energy analysis laid the groundwork for uncovering novel insights, advancing knowledge, and making significant contributions to both academia and practical applications within the field [41].

A literature review was conducted to gain a comprehensive understanding of the field of building energy analysis. The Web of Science was selected as a reliable data source for this review. Through a systematic approach, relevant scholarly materials were identified and analyzed. The review focused on key aspects of building energy analysis, aiming to identify gaps and contribute to the existing knowledge in the field. The selected literature underwent a critical evaluation, ensuring the reliability and quality of the findings. The review process helped establish a strong foundation for subsequent investigations within the chosen field.

Using the Reporting Standards for Systematic Evidence Syntheses (ROSES) process, a systematic literature review (SLR) was carried out [42]. The ROSES framework was employed to conduct a systematic literature review in the field of disaster management. Despite its original focus on environment management, the methodology proved suitable due to its consideration of complexities and variations in different scenarios and studies. The review process involved three steps: identification, screening, and eligibility, ensuring a comprehensive and systematic approach to document searching. A quality assessment was also conducted using adapted criteria to evaluate the reliability and credibility of the included studies. Overall, the ROSES framework facilitated a transparent and rigorous review process, providing a solid foundation for further research in disaster management. Its flexibility in accommodating diverse contexts and the inclusion of a quality assessment process enhanced the credibility and validity of the review findings. By implementing the ROSES framework, this review contributes to the advancement of knowledge in disaster management [43].

The process within the current project involved several detailed steps (Figure 2.2). It began with formulating the main query, which was derived from the project's goals and the researcher's initial background study. This initial step yielded a considerable number of works, with over 800 identified publications within the specified timeframe of 2020 to 2023. The subsequent filtering step was performed based on the field of study and the relevance to the specific topic of interest. All 832 papers were carefully analyzed during this filtering stage, resulting in a substantial reduction in the number of papers to be further examined. Ultimately, 40 papers were selected for full reading, representing a more refined and focused set of literature.

In the final step of filtering, the significance of the topic was taken into consideration. This step aimed to ensure that only the most relevant and impactful papers were included in the Data Extraction Table (DET) for further synthesis and analysis. Consequently, only 23 papers met the criteria and were selected for inclusion in the DET, representing the final set of papers that formed the basis for synthesizing the research findings. This multi-step filtering process allowed for a systematic and rigorous approach to document selection, ensuring that the most relevant and significant papers were included in the analysis. By carefully scrutinizing each stage of filtering, the researchers were able to narrow down the initial pool of publications to a final set of 23 papers that would contribute to the synthesis and generation of the final research outcomes.

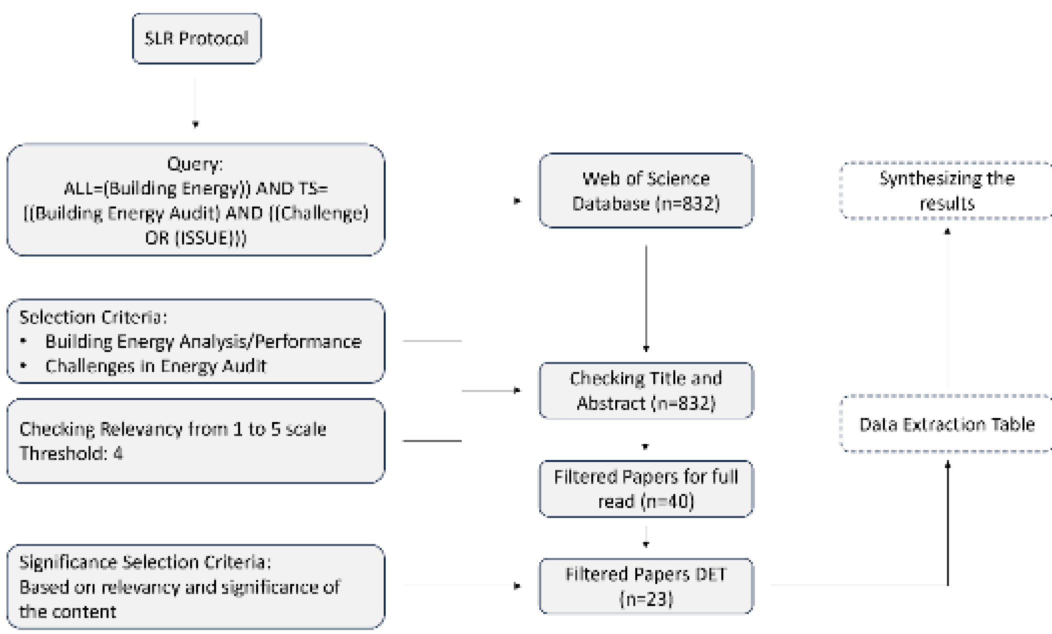


Figure 2. 2. Literature review workflow

2.4. 3D Model Generation of Existing Buildings

In the Architecture, Engineering, Construction, and Facility Management (AEC/FM) industry, there is a growing emphasis on existing buildings due to their significant potential for performance improvement and positive environmental impact [44]. Various AEC/FM applications targeting existing buildings necessitate the use of accurate 3D models that depict the as-built conditions of these structures [45]. These models play a crucial role in supporting a wide range of activities, including safety and health assessments, space planning, procurement, cost estimation, life cycle assessment, sustainability evaluation, performance monitoring, operations and maintenance, scheduling, as well as retrofit, refurbishment, and renovation planning.
By employing precise 3D models, stakeholders in the AEC/FM industry can effectively evaluate the current state of existing buildings and make informed decisions about various aspects related to their management and optimization. Safety and health assessments can be conducted to identify potential risks and ensure compliance with regulations. Space planning activities benefit from accurate representations of the building, allowing for efficient utilization and allocation of spaces. Procurement decisions can be made more effectively by considering the precise conditions of the building. Cost estimation becomes more reliable when based on detailed 3D models that capture the actual characteristics of the structure.
Furthermore, life cycle assessment and sustainability evaluations are enhanced through the utilization of 3D models that enable comprehensive analysis of the building's environmental impact and performance. Monitoring key performance indicators becomes more efficient by integrating the as-is conditions into the assessment process. Operations and maintenance activities can be optimized through access to precise 3D models, facilitating effective planning and execution. Scheduling tasks and activities are streamlined by leveraging accurate representations of the building. Lastly, retrofit, refurbishment, and renovation planning benefit from detailed 3D models that allow for better visualization and assessment of potential improvements [46].
The first method to discuss is using computer vision algorithm to have a Indoor 3D reconstruction: The first thing to know is that, this technique can be applied for different input data based on the goal of the work. Input source for data can be in the form of image (normal RGB image from normal cameras) or point cloud (from 3d scanners). Using image input for create a 3D reconstruction of the building may have various steps and methods, the one that we discuss in this project which is a standard one that can be used to create the indoor environment of a building is as follow (Figure 2.3.):

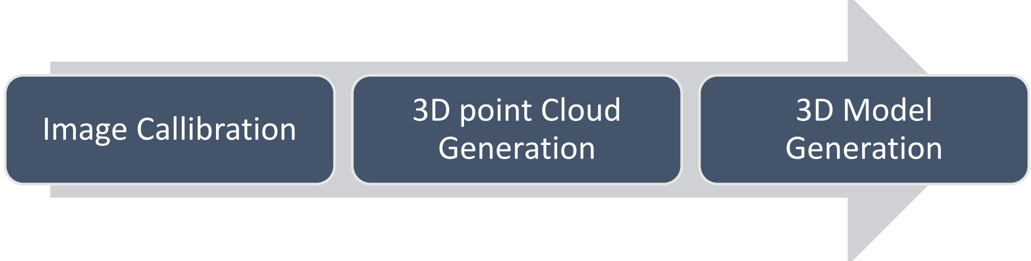


Figure 2. 3. Image-based 3D reconstruction workflow

- The first step in our pipeline is the calibration of the imaging system. This crucial process involves estimating various parameters of the cameras, such as focal length, scaling factor, and distortion. Additionally, the rotation and translation between the two cameras are determined. Accurate calibration is essential for ensuring precise and reliable 3D reconstructions.
- Next, the pipeline moves on to the generation of a 3D point cloud. Multiple views of a room are utilized to estimate a set of 3D points representing the spatial layout. This step employs a combination of structure-from-motion techniques and multi-view stereo methods to reconstruct the 3D environment. By leveraging the information from different camera perspectives, a comprehensive and detailed point cloud is created.
- Building upon the sparse set of 3D points obtained in the previous step, the subsequent stage focuses on generating a 3D model of the room. Specifically, this step involves the calculation of the room's walls in 3D. By analyzing the available data and leveraging geometric principles, the pipeline determines the precise positioning and shape of the walls within the reconstructed 3D environment.

2.4.1. Image Calibration

The camera calibration process in computer vision involves defining the physical properties of a camera through its intrinsic and extrinsic parameters [47]. Intrinsic parameters primarily describe the focal length and optical center of the camera, while extrinsic parameters specify the camera's physical location, including rotation and translation, with respect to a reference coordinate system. To perform camera calibration, one can utilize software tools such as the camera calibration toolbox in Matlab or Python. A common method involves using a checkerboard calibration object made from a piece of cardboard. Images of the checkerboard captured by the visible band camera are used for calibration.
By taking a series of images of the checkerboard from different perspectives (such as in Figure 2.4.), the intrinsic and extrinsic parameters of the visible band camera can be computed. It is recommended to capture at least 20 images to ensure accuracy in the calibration process. Through the analysis of these images, the software tools can estimate the intrinsic parameters, such as the focal length and optical center, as well as the extrinsic parameters, including the camera's rotation and translation. This camera calibration process provides a valuable means of determining the camera's intrinsic and extrinsic properties, enabling accurate measurements and precise positioning in subsequent image-based tasks. By utilizing the appropriate software tools and following the calibration procedure using the checkerboard calibration object, researchers and practitioners can obtain reliable intrinsic and extrinsic camera parameters for their computer vision applications.

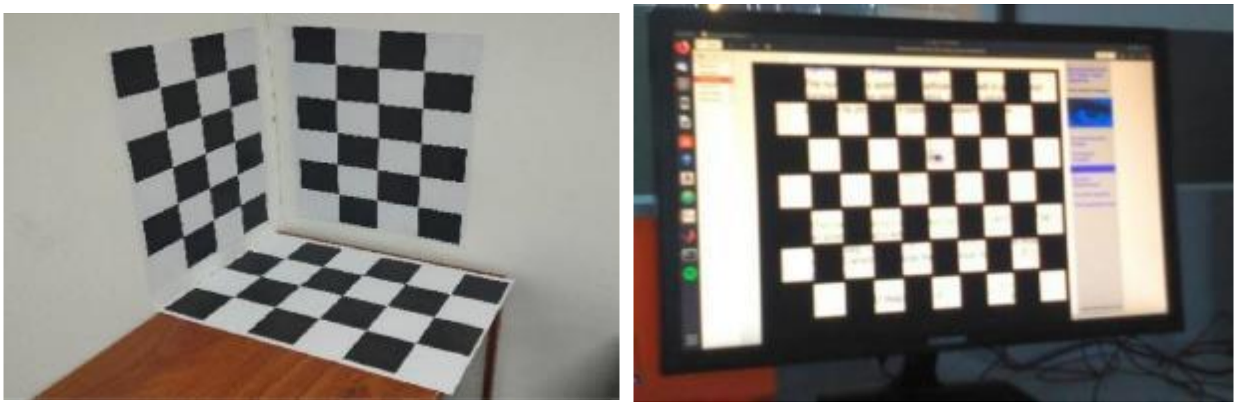


Figure 2. 4. Checkboard calibration method

2.4.2. 3D Point Cloud Generation

To generate a high-quality 3D point cloud of the room, relying on the widely-used technique of Structure from Motion (SfM) in computer vision is a standard practice [48]. SfM is a powerful method that utilizes a set of images captured from different positions, possibly with different cameras, to extract 3D information about the environment. The key principle behind SfM is to identify overlapping views in the images, where the visual information corresponds to the same 3D entities in the environment [49]. By analyzing these overlapping views, the positions of the cameras and the 3D coordinates of the pixels can be estimated. This allows us to reconstruct the spatial layout of the room in 3D. SfM takes advantage of the inherent redundancy in multiple images capturing the same scene from different angles. By leveraging the parallax effect and the correspondence of visual features across images, SfM can accurately estimate the camera poses and triangulate the 3D positions of the observed points.

By incorporating information from multiple images and exploiting the geometric relationships between them, SfM enables the generation of a dense and precise 3D point cloud. This point cloud represents the spatial structure of the room, capturing the detailed geometry and spatial relationships of objects within the environment. SfM has proven to be a valuable technique in various applications, such as 3D reconstruction, augmented reality, and scene understanding. Its ability to extract 3D information from a set of images makes it a versatile and powerful tool for generating accurate and detailed 3D representations of real-world scenes.

Keypoint Description: In the context of Structure from Motion (SfM), identifying meaningful pixels (keypoints) in images is crucial for accurate reconstruction. The Scale-Invariant Feature Transform (SIFT) method is employed to locate keypoints by analyzing intensity changes across different scales. Pixels with consistent changes are considered useful and are described using 128-dimensional vectors, summarizing intensity changes around the keypoints.[50].

Visual Feature Matching: After keypoints are identified and represented as feature vectors in each image, the next step involves matching keypoints that correspond to the same 3D points. This is achieved by comparing feature vectors across different views and identifying the closest matches. A cascaded method is utilized for matching, producing a set of potential matches between features from various views. Subsequently, a post-processing step involving AC-RANSAC is applied to eliminate geometrically incorrect matches [51].

Structure from Motion (SfM): SfM addresses both reconstructing matching 3D points and estimating relative 3D distances and poses between images. The method proposed by Moulon et al. is chosen for its robustness and adaptability. It begins with an initial 3D model based on the best matching images and progressively incorporates additional images for reconstruction.

Densification using Multi-View Stereo (MVS): To achieve a dense 3D point cloud necessary for accurate reconstruction, the sparse point cloud obtained from SfM is densified using an existing multi-view stereo algorithm. This algorithm employs interpolation to densify the given 3D point cloud [52].

Enhancement through Interaction: Further refinement involves manual interaction. The generated 3D point cloud requires adjustments to scale and the selection of planar regions on walls, doors, and windows through user input.

Scale and Pose Adjustment: The 3D point cloud obtained earlier needs correction for scale and pose. Three points are selected from captured images, and their actual coordinates are measured relative to a known origin point in the room. This information is then used to calculate a similarity transformation, correcting scale, orientation, and translation of the 3D point cloud. This process involves a non-linear least square optimization method. However, this procedure might introduce discrepancies due to challenges in accurately matching pixels to known 3D points, particularly with low-resolution images. The imperfections of real rooms, like rounded corners and tilted walls, further contribute to the noise in the process. Using this available information, we can determine the necessary transformation between the existing model and the target model, ensuring accurate scale and orientation alignment. To achieve this, a similarity transformation is calculated using the following formula:

$$\boldsymbol{x'} = \boldsymbol{Ax} + \boldsymbol{t}, \quad (1)$$

Here, $\boldsymbol{x} \in \boldsymbol{R3}$ represents a point in the original 3D model, which is devoid of scale or specific orientation. A denotes $\boldsymbol{R}^{3\times3}$ orthogonal matrix encompassing rotation and scaling elements. Meanwhile, $\boldsymbol{t} \in \boldsymbol{R}^3$, signifies translation. The outcome, x', corresponds to the 3D point after it has been adjusted for both scale and orientation [53]. This process is executed through a non-linear least square optimization technique, ensuring the solution minimizes the mean squares error ($\Sigma i(xi' - xi)^2/n$).

2.4.3. 3D Model Generation

The dense 3D model comprises a collection of points within a three-dimensional space. To accurately generate the 3D model, it is necessary to identify and estimate the surfaces as 3D planes. Equations for the planes corresponding to different boundary surfaces (walls, floor, and ceiling) are computed using points situated within a manually marked rectangular region on the images. It's important to ensure that only obstruction-free surface portions are marked to prevent misleading plane fitting caused by objects. This allows the algorithm to function effectively even in cluttered or obstructed environments. Two methods have been devised to robustly estimate the planes corresponding to the 3D point cloud. The first method, "Baseline Plane Estimation

from 3D Point Cloud" (BPE), relies on RANSAC and SVD to independently estimate a plane for each wall, ceiling, or floor. The second method, "Robust 3D structure estimation with geometric constraints" (RSEC), improves upon the first by enhancing the precision of the final model. RSEC builds upon BPE and leverages the assumption of the room's rectangular shape.
The baseline method closely resembles the approach used by Ham and Golparvar-Fard [53] for constructing a 3D model from visible band images, while the second method, RSEC, , introduced by Gursel Dinoa et al [54]

Baseline Plane Estimation Method: The BPE method employs least-squares plane fitting with RANSAC. As previously mentioned, the user selects a four-corner polygon on each wall where the enclosed points are coplanar. These bounded 3D points are then located within the 3D point cloud. The algorithm proceeds to fit a plane to each wall for constructing the room's 3D geometric model.
A set of points is randomly chosen from each surface to estimate a plane. The success of the plane fit is assessed by calculating the inlier ratio of the model estimation, which measures the proportion of points adhering to the estimated model. Inlier determination involves a distance threshold method. If the distance between a point and its corresponding estimated plane is below a specific threshold (hyperparameter), the point is classified as an inlier. This process is iterated, and the geometric model with the highest inlier ratio is retained as the best model. A least-squares error plane fitting algorithm is adapted for model estimation within the proposed pipeline. Given N 3D points, x1, …, xN, with xi ∈ ℝ3 and sampled through RANSAC, they are stacked in an N × 3 matrix denoted as X = [x1,x2,…,xN]T, where T represents transpose. As a common preprocessing step, the center point of the set is calculated and subtracted from all points to shift their center to the origin. Therefore, the new points and also the new 3xN point matrix would be as follows:

$$\boldsymbol{x} = \sum_{i=1}^{N} \boldsymbol{x}_{i,} \tag{2}$$

$$x'_i = x_i - \overline{x}, \tag{3}$$

$$X' = [x'_1, x'_2, \ldots, x'N]^T. \tag{4}$$

The general aim of this process of plane fitting is to find an appropriate normal vector $n \in \mathbb{R}3$ that minimizes the MSE (Mean Squared Error) of 3D points which are expected to be within the wall:

$$n^* \leftarrow \arg\min_n \sum_{i=1}^{N} |n^T X_i|^2 = \arg\min_n n^T X'^T X' n, \text{ s. t. } ||n||^2_2 = 1. \tag{5}$$

Lagrange multipliers method has been chosen for minimizing the cost function which is as folllows:

$$J\,(X'; n, \lambda) = n^T X'^T X' n + \lambda(1 - n^T n). \tag{6}$$

In order to minimize the cost function (J), the derivative of J should be equal to 0:

$$\frac{\partial J}{\partial n}\; 2X'^T X' n - 2\lambda n = 0, \tag{7}$$

$$X'^T X' n = \lambda n. \tag{8}$$

Where the vector of $\boldsymbol{n} \in \mathbb{R}3$ works with eq. (8) and the cost of $\boldsymbol{n^T X'^T X' n}$ should be found to eigenvalue:

$$n^T X'^T X' n = \lambda n^T n = \lambda\,. \qquad (9)$$

As the result, in order to minimize the cost, selection of eigenvector corresponds to the minimum eigenvalue:

$$n^* \leftarrow \underset{n}{\arg\min}\ \lambda\,. \qquad (10)$$

The plane equation is then calculated as:

$$n^* (x - x_0) = ax + by + cz + d\,. \qquad (11)$$

Where n* represents normal vector of the plane, x_0 as a fixed point on the plane and x is any arbitrary point on the plane. In this way, a,b, and c would be as follows Where n_x, n_y and n_z are z,y, and z component of n:

$$a = n^*_x, \quad b = n^*_y, \ c = n^*_z, \ d = -n^{*T} x_0, \qquad (12)$$

Robust 3D structure estimation with geometric constraints (RSEC): as proposed by Dino et al [54] an improvement to the baseline method, in order to exploit the 3D structure of the room, it has been assumed The meeting point of room surfaces forms a precise right angle of 90 degrees, and the estimation of planes incorporates this specific constraint. Consequently, within RSEC, the calculation of three surface normal vectors, each mutually perpendicular, becomes essential. To enforce this constraint, two distinct cost functions are needed for the entire room. It's important to note that in contrast to BPE, where each wall is treated separately, these functions account for both plane fitting and maintaining the orthogonality of the surfaces. The initial cost function is articulated as follows:

$$J_{fit} (X_1, X_1, \dots X_6, n_1,n_2,n_3) = n^T_1 X^T_1 n_1 + n^T_2 X^T_2 X_2 n_2 + n^T_1 X^T_3 X_3 n_1 + n^T_2 X^T_4 X_4 n_2 + n^T_3 X^T_5 X_5 n_3 + n^T_3 X^T_6 X_6 n_3 \qquad (12)$$

And in the simpler version it can be seen as follows:

$$J_{fit} (X_1, X_1 ,\dots X_6 ; n_1 , n_2 , n_3) = n^T_1 A n_1 + n^T_2 B n_2 + n^T_3 C n_3\,, \qquad (13)$$

And in this situation A, B, and C are positive metrics which are symmetric. The second cost function defines the orthogonality as follows:

$$J_{ort} (N) = \|N^T N - I\|^T_2 = tr\,((N^T N - I)\,(N^T N - I))\,, \qquad (14)$$

Where finally two cost function will be merged

$$J (X_1, X_1, \dots X_6, ; N, \lambda) = Jfit\,(X1, X1 ,\dots X6 ;N) + \lambda\, Jort\,(N) \qquad (15)$$

This final cost function will be minimized using “Nelder-Mead” method [55]. In this process different values for λ were tested for 200 images from an iPhone 14 pro.

In order to designate the positions of windows and doors in a three-dimensional space, users pick out the corners of these openings within the visible images. In doing so, we operate under the assumption that the windows and doors are situated on the same plane as the walls they're a part of. Following this selection process, we proceed to determine the precise 3D coordinates for each of the chosen corner points. This involves finding the point where a ray, projected from the selected 2D point, intersects with the corresponding plane of the wall. In Figure 2.5 the summary of the whole process is shown.

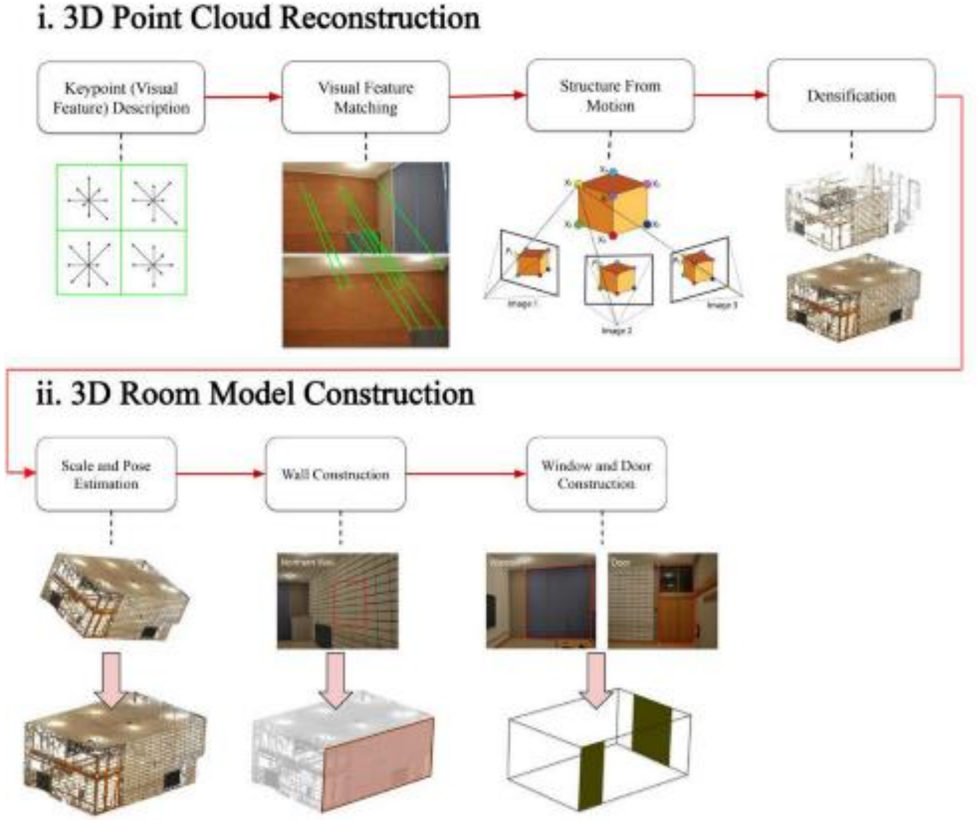


Figure 2. 5. Integrated method of point cloud generation and 3D model construction

2.4.4. Alternative Method

Within the realm of geosciences, the applications of terrestrial laser-scanning and airborne laser-scanning techniques (TLS & ALS) extend across diverse scales for conducting topographic land surveys [56, 57]. LiDAR stands as a prevalent method, employing laser pulses to measure distances based on pulse return times between a transmitting unit and a receiving unit [58]. The swift progression of digital processing methods, coupled with the emergence of novel technologies in remote sensing, is driving a revolution in the realm of digital twin of our surrounding [59]. The advancement of technology in multi-sensor portable devices has integrated numerous digital tools into tablets and smartphones. These devices have become essential components of an engineer's toolkit, signifying standard equipment. Building upon the concept of primary digital mapping, the utilization of smartphones in fieldwork has expanded to become a direct alternative to traditional approaches for collecting field data. Upon the introduction of novel features by Apple, global attention is invariably captured. In March 2020, the unveiling of the new iPad Pro included the incorporation of a LiDAR scanner, a development that piqued curiosity regarding the potential utility of this sensor (See Figure 2.6).

Figure 2. 6. Progress of evolution of digitally assisted fieldwork of smartphones, based on [60]

Apple announced the integration of a LiDAR scanner into the camera of the iPhone 12 Pro, prompting inquiry into the rationale behind embracing a technology with an established history spanning decades within the domain of mobile devices and now it continues and developed up to the latest version of Apple iPhone 14 pro as the state-of-the-art version of this technology with high accuracy. The succinct answer to this query resides in the enhancement of augmented reality (AR) experiences. For individuals involved in construction and residential refurbishments, the implications are more compelling: the technology facilitates precise, prompt, and realistic 3D modeling of indoor spaces. This advancement not only simplifies tasks but also augments customer services. A comprehensive exploration of LiDAR room scanning, encompassing its applications and advantages, is warranted. The LiDAR sensor, tailored for room scanning, exhibits the capability to measure distances to proximate objects within a 5-meter radius. However, its intrinsic potential remains dormant unless harnessed in conjunction with dedicated software applications like magicplan [61]. The amalgamation of augmented reality (AR) with artificial intelligence (AI) within this software engenders automated object detection and categorization. Beyond the rudimentary recognition of elements like floors and walls, the software enables faithful reproduction of entire spatial geometries. Within this LiDAR framework, a room scan entails the capture of data, subsequently identified, and processed by AI. This archival of data bears enduring significance, as the information remains perpetually accessible and exploitable.
Therefore, as an alternative, rapid, and user-friendly approach to creating simplified 3D models of indoor spaces leverages the power of lidar sensors in smartphones which has been very popular among researchers [60]and there are several studies where this method has been implemented for different purposes including but not limited to urban modelling [62], building 3D modelling [63], drop of liquid [64], wildlife studies [65] and a lot of several areas such as engineering (look at Figure 2.6) by search WoS database. As it is depicted in Figure 2.6 it has been widely used in various fields especially in remote sensing, geology, computer science and engineering. The last field (engineering) has by far highest number of applications for smartphone lidars and by taking a deeper look at the following figure, building technology has also been mentioned which can be integrated to the engineering sector.

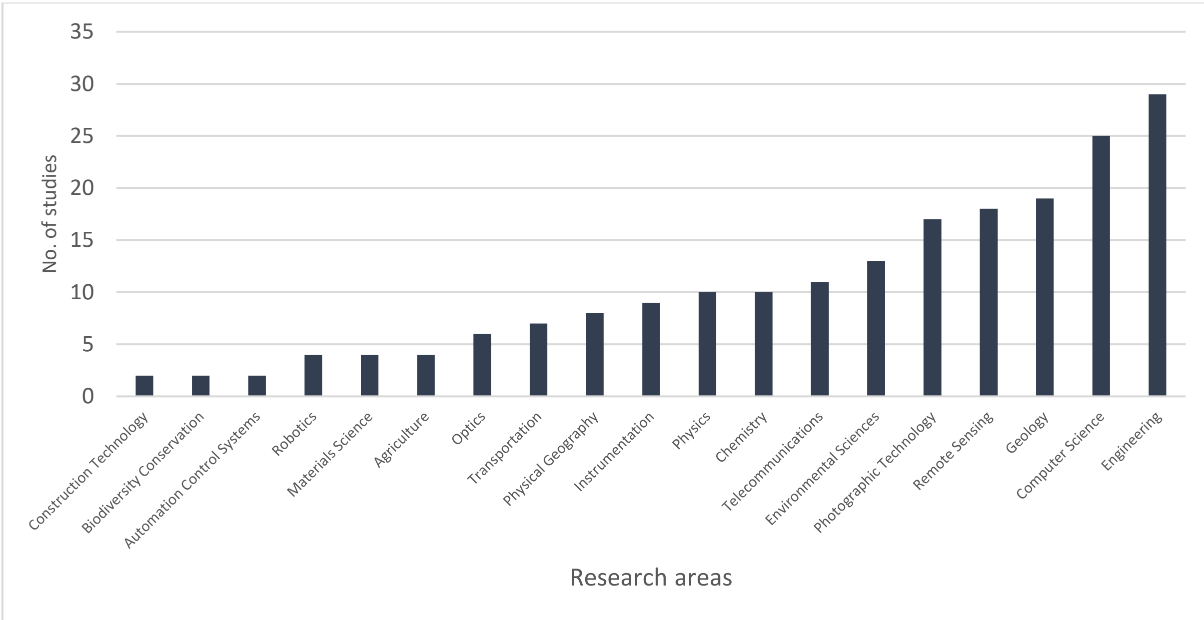


Figure 2. 6. Implementation of smartphone LIDAR sensors in different research areas, based on WoS

Another important aspect is that the implementation of this method is growing since these days more and more cell phone companies are adding LIDAR sensors to their products.

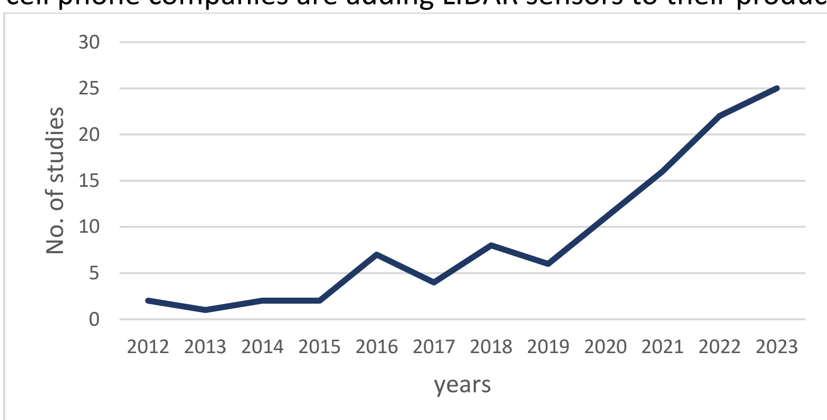


Figure 2. 7. Studies using smartphone LIDAR scanners in recent years, based on WoS

This change is obvious by considering the number of research studies that use this technology in recent years which shows a meaningful increase. This innovative method offers a convenient way [66] to generate accurate room representations without the need for intricate coding or technical expertise. Applications such as Magicplan [61] exemplify the seamless integration of lidar technology into everyday devices, transforming the way individuals capture and visualize their surroundings. Unlike the traditional computer vision-based methods, which involve intricate steps and complex algorithms, the lidar-based technique capitalizes on the inherent simplicity of lidar sensors. These sensors emit laser beams that bounce off surfaces and return to the device, allowing for the accurate measurement of distances and the creation of precise 3D point clouds. This process is quick, efficient, and requires minimal user input, making it an ideal solution for users seeking an intuitive way to model their indoor spaces. One standout feature of this approach is its computational efficiency. Lidar sensors provide real-time depth information, which significantly reduces the computational load compared to image-based methods. This means that the process of generating 3D models becomes almost instantaneous. Users can simply walk around the room with their smartphones, and the lidar sensor captures the spatial data necessary to create a virtual representation. The generated point cloud offers a foundation for further analysis and visualization.

The workflow is highly smooth and user friendly (Figure 2.8.) which starts with defining the project, location, function, and other required details by which the level of details can be increased. In the next step the type of field data acquisition must be selected so that in this case LIDAR sensor would be the optimum choice. Then, the application will ask to rotate the phone and point a stick towards the first corner and other corners in order. After detecting a close shape, the application will ask to rotate the phone towards ceiling for height detection. The preliminary plan is constructed in this stage which needs more details including opening shape, location, and type. One again by pointing camera towards openings they will be automatically detected but the type of opening must be confirmed (if it is a door or operable window or a non-operable window). It should be noted that in any stage of this 3D model creation including this stage there is a possibility for editing, removing, or adding elements. The next two steps are optional for a building not only a single space which means users can add each space to a whole to create the whole unit. Additionally, furniture can be added to the unit too. Finally, the CAD model can be exported to be used for different purposes.

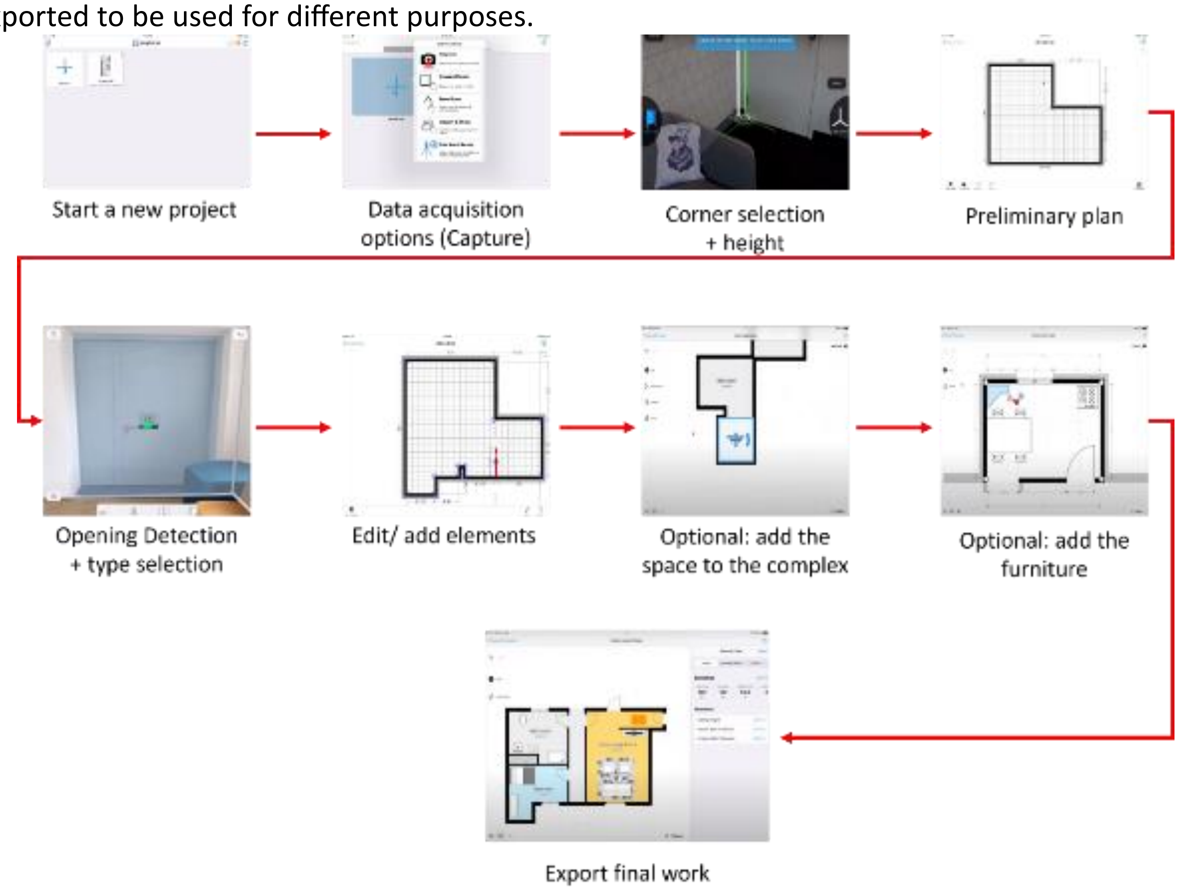


Figure 2. 8. 3D model creation workflows

Magicplan and similar applications build upon this lidar-based data by incorporating intelligent algorithms that identify edges, corners, and openings in the room. By processing the lidar data, these apps can detect the boundaries of the room, calculate its height, and recognize openings such as windows and doors. This information is then used to construct a simplified 3D model that captures the essential features of space. The beauty of this method lies in its accessibility – anyone with a compatible smartphone can use it, regardless of their technical background. Moreover, the lidar-based approach eliminates the need for labor-intensive manual measurements or complex 3D modeling software. Users can create accurate floor plans and room layouts with just a few taps on their screens. The user-friendly interface guides individuals through the scanning process, ensuring that the captured data aligns with the physical space. The benefits of this method extend beyond its ease of use. The generated 3D models serve various purposes, from aiding in interior design and furniture placement to assisting in renovation planning. The accuracy of the lidar-based approach ensures that users can make informed decisions based on reliable representations of their environments. Additionally, these models can be easily shared with architects, contractors, and other professionals, streamlining communication and collaboration.
Therefore, the benefits of this method compared to the previous one can be categorized as below:

- User-friendly interface
- Low-cost method
- Fast process
- Low-computational resource requirement
- High accuracy

2.5. Bigdata Generation

Conventionally, studies focused on optimizing building performance often utilize unchanging building shapes, with the parameters for optimization being the physical characteristics of materials or the configurations of building systems [67]. While this method is completely valid, it may not be applicable when the actual building is not considered. In better words, when talking about an archetype of a specific style of architecture or building rather than an actual case study the geometry and shape of the buildings within the studied cluster can be slightly or even considerably different. In this situation considering similarities is important but differences are more critical. In order to consider almost all possible cases within the studied cluster two sets of information were needed.
Firstly, the basic characteristics of the archetype of specific style of buildings that is the focus of this project. This archetype is the most frequent style of building in the cluster with the normal building components that have been used in similar buildings. Characteristics of the archetype can be categorized into the following three items:

- Spatial configuration:
  Number of bedrooms and other spaces

General shape and geometry of the house
Number of floors
Shape of each space

- Construction details:
  Normal indoor and outdoor walls
  Number of windows and the type of glazing and frame of windows
  Foundation details
- Operational details:
  Function of the building
  Schedule of occupancy
  Energy systems
  Operational settings

Secondly, the range of variations for each characteristic is also another important feature that helps the inclusion of majority of cases in the cluster. In this way through a deep field study the possible range of each feature of the cluster will be identified.

By having these two sets of information it would be possible to generate the big data for parametric analysis. In the process of constructing a model to predict energy consumption, a plethora of input parameters are considered. These parameters exhibit intricate interrelations, and when combined with variations in building types, lead to considerable disparities in energy consumption outcomes. Current methodologies and associated software tools utilized in energy simulations often fall short in terms of facilitating seamless data exchange and interoperability across various modeling and energy simulation platforms. An up-and-coming technique in the realm of energy analysis is parametric analysis [68]. Notwithstanding its potential, it yields voluminous numerical datasets, which can be daunting for many design professionals to decipher. This is particularly true for architects who are more inclined towards perceiving analysis results in more intuitive visual representations, rather than delving into extensive numerical datasets. It is worth noting that during the preliminary phases of design conceptualization, architects often endeavor to formulate a design that optimally balances heating, cooling, and lighting loads.

Parametric design can be understood as an approach that grapples with design variables related to the geometric properties, particularly from an architectural perspective. It essentially serves as a medium to articulate an idea, converting it into tangible actions using certain parameters, which, in this context, represent the significant attributes of the design. Such a parametric model, when visualized, presents a specific shape derived from a pre-determined set of parameters currently in use. These parameters are interconnected through an array of associations, enabling designers to revisit and adapt logical definitions whenever there's an alteration in preceding parameters at any juncture during the design phase. On a related note, Building Energy Modeling (BEM) is employed to project a structure's expected energy consumption and the consequent energy savings when juxtaposed with a standard benchmark. This projection aids in affirming a project's adherence to local, regional, or national energy standards. The accuracy of BEM hinges on Typical Meteorological Year (TMY) data and certain assumptions concerning the operation of the building [69]. It's pivotal to recognize that the accuracy of these predictions aligns closely with the underlying assumptions. In this regard, the process of energy process would be as Figure 2.9 in which the input data will be given to the energy simulation engine to have the building performance results.

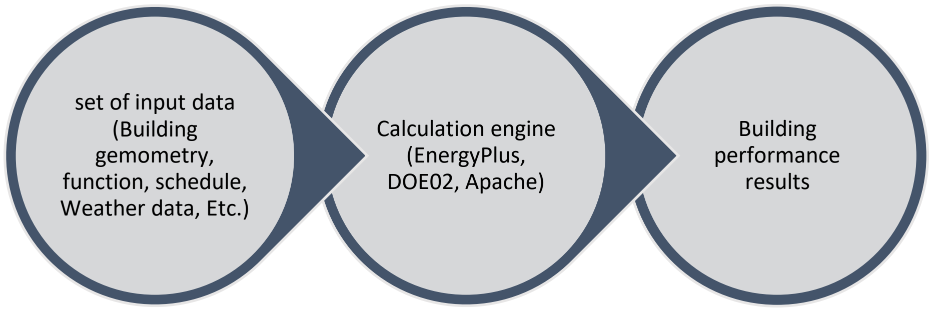


Figure 2. 9. 3D Energy simulation workflows

According to the field study and the two sets of data it would be possible to generate big data. Nevertheless, it is important to mention here that in this case because of two main reasons the focus scale of study is a single space rather than a complete house. The main reason is that most house or apartment owners in Poland are enthusiastic about having a deep insight about each space since it is quite normal there that each space is rented separately. It means that the general insight into the house in this case while being useful might not be the exact information that building owners are looking for. Another reason is that is that in almost most cases in Poland as the case study at least there is one space that has experienced undocumented renovation that can compromise the whole Building energy simulation (BES) result of the building. Therefore, the focus of this study would be on single spaces of the studied building cluster to be more consistent in terms of building characteristics as well as the results.

2.5.1. General Workflow

In this section, we delve into the intricate processes underlying big data generation, elucidating its general workflow. This exploration commences by leveraging two distinct sets of information derived from a previously mentioned field study - the Archetype and the variation range. The nuances and intricacies of these data sets will be unraveled in greater depth as we venture deeper into the discussion of this phase. From these foundational data layers emerges what we term the "geometry range." This encompasses nearly every conceivable shape and configuration of the space under study, all confined within the defined cluster boundaries. For readers seeking a more granulated understanding of the cluster's characteristics and significance, a thorough examination is reserved for the case study section.

Once the geometry range has been meticulously crafted, it paves the way for the incorporation of supplementary Energy Application domain (EDA) details. These details span a broad spectrum, capturing facets like energy systems, schedules of operation, openings, construction intricacies, and so forth. The culmination of this extensive data collation and processing is represented by the Input Data File (IDF). This file format is particularly pivotal for its compatibility with energy simulation engines, most notably, EnergyPlus. When this IDF file is fed into the energy simulation engine, it undergoes a rigorous analysis, subsequently yielding insights into the energy performance of the buildings under study. For a visual representation of this process, readers are directed to Figure 2.10, which provides a comprehensive overview.

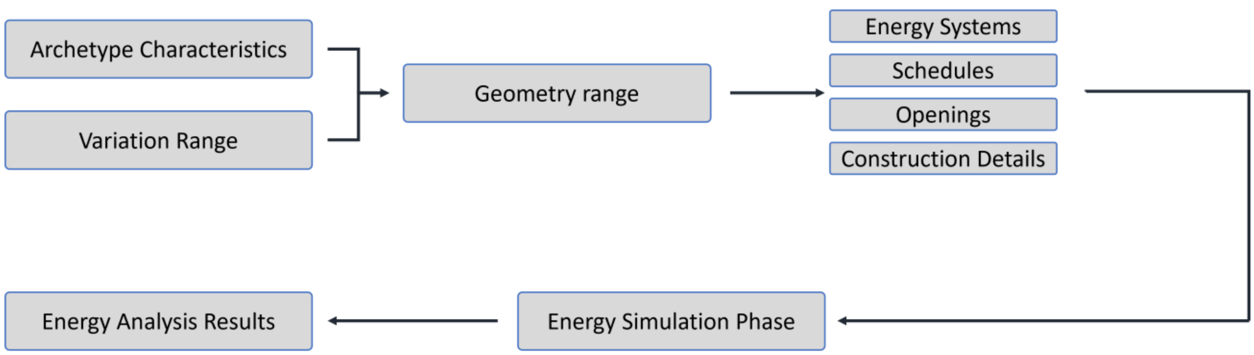


Figure 2. 10. General workflow of big data generation

In terms of tools that have been used in this workflow the process of automated bigdata generation using geometrical and environmental input begins in Rhino environment using Grasshopper plugin. The output in the form of geometry was then used as an input to Ladybug tools component to define the space as a thermal zone and then as a living space with defining function, schedule and the energy systems in Honeybee plugin. It led to the creation of the IDF file which was the input for EnergyPlus energy simulation engine to run the calculation and the results were extracted in a CSV file for further analysis (Figure 2.11).

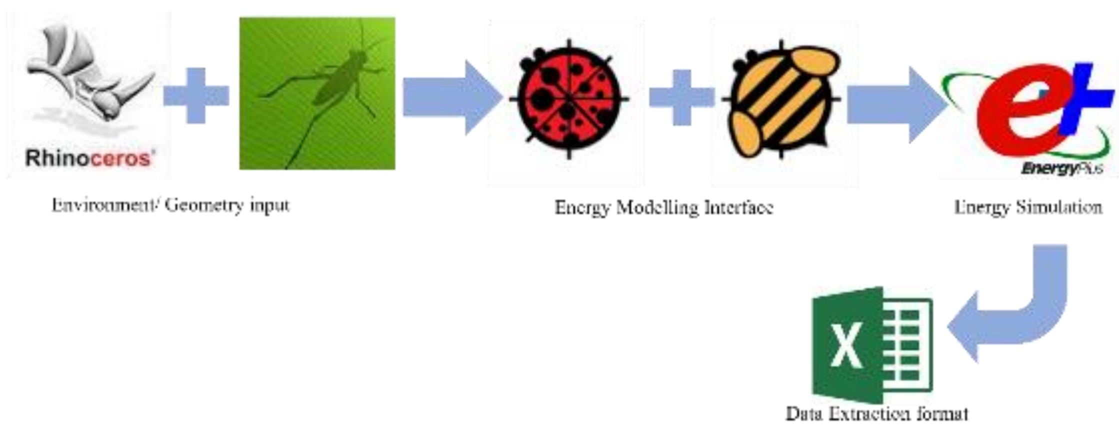


Figure 2. 11. Application used in different stages of general workflow

2.5.2. Geometry Creation

In this initial part of big data generation, it was very critical to set the archetype and variation in a form to include the highest possible rate of buildings in the cluster within the range so the results would be representing the behavior of the whole cluster. In this way the most basic geometrical feature of a single space. Hence, using features in Figure 2.12 was the beginning of the geometry creating in the first step of the workflow.

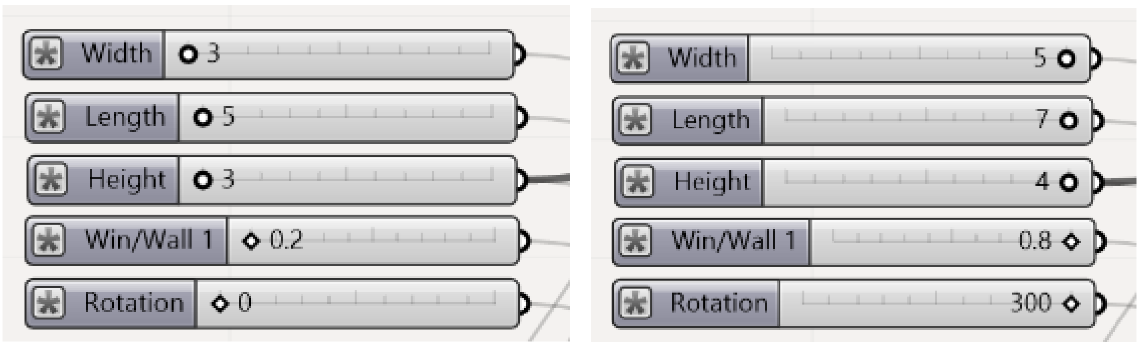


Figure 2.12. upper and lower range of input geometrical range

The basic geometrical features including height, width, Length, window to wall ratio and rotation variables with specific ranges and distance between each step created the whole dataset. Indeed, based on the field study for each single room of the cluster house the following range includes most cases the following range and steps have been used in this study (Table 2.1.).

Table 2.1. Variable range and steps

Variable	Min	Max	Step	No. of Variations
Height	3	4	0.25	5
Width	3	5	0.5	5
Length	5	7	0.5	5
Window to Wall Ration	0.2	0.8	0.2	4
Rotation	0	300	60	6
Total Variations				3000

Using these ranges as input created the geometry of the space using shoe-box modelling attitude based on Reinhart suggestion [70] which has been proven to be a useful method for comparing variations of a building cluster as it can speed up the process and avoid unnecessary calculations. Also, there are numerous studies that have validated the applicability of using shoebox modelling for BEM and UBEM [70, 71].

It is important to mention here that window to wall ratio in this study defined as the percentage of the total area of the exterior wall and the shape of window is the offset of the wall itself.

2.5.3. Construction Details

In this stage based on most frequent construction details and materials used in the building cluster, required information and measurements were translated in Honeybee material algorithm. In order to do this each layer of material should be defined accurately based on the following order in Table 2.2.

Table 2.2. Variable range and steps

Variable	Description
Name	Name of the material (for future reference and reporting)
Thickness	Thickness of the material layer [m]
Conductivity	Measure of material's ability to conduct heat [W/mK]
Density	Mass of the material per unit volume [kg/m³]
Specific Heat	Amount of heat energy required to raise the temperature of the material by 1°C [J/kg°C]
Roughness	Surface irregularities and texture that affect friction and interaction
Thermal Absorption	Capacity of a material to soak in and retain thermal energy
Solar Absorption	Ability of a material to capture solar radiation
Visual Absorption	Material's propensity to capture or reflect light in the visible spectrum

After defining each material based on these order and using Honeybee Opaque Material Component. Cobining these layers, will define the construction details of a specific surface such as exterior wall for instance. According to Figure 2.13.

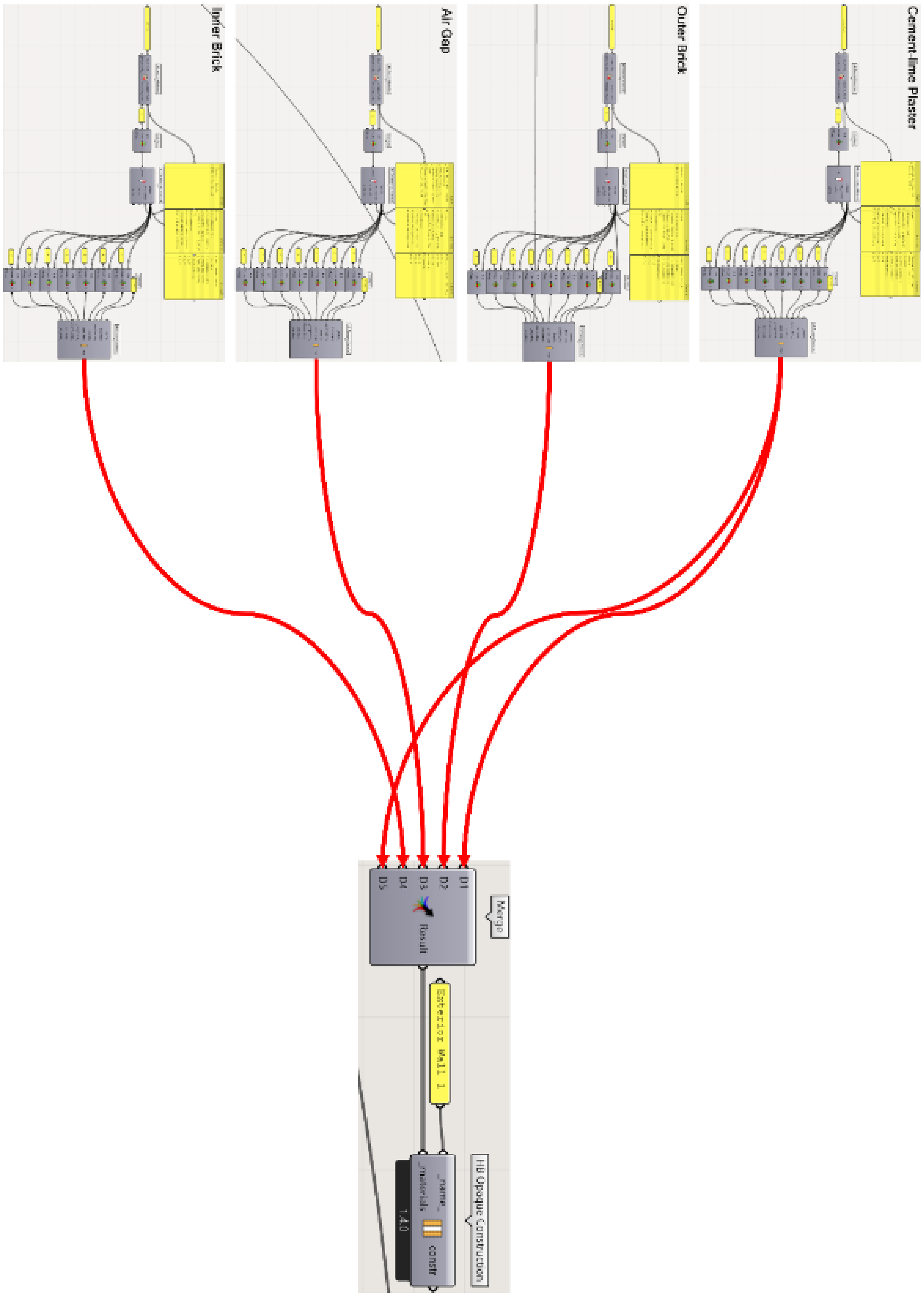


Figure 2.13. creation of different construction details using Honeybee components

2.5.4. EAD Additions

Then in order to apply EAD additions, there are two important layers off data that must be added to the basic model. Firstly, the energy systems and secondly occupants profile. In order to do it, it is vital to have deep enough knowledge about the original and current systems of the studied building cluster which are as follows:

- Heating: Originally most of buildings in the cluster would have been primarily heated with coal-fired boilers, with radiators distributing the heat throughout the house. Over the years, coal-fired boilers have been largely replaced. Now, it's more common to find gas-fired boilers or pellet stoves. Underfloor heating systems might also be present in some homes.
- Cooling: Typically buildings lacked air conditioning systems, relying on the natural cooling provided by window ventilation and thick walls. This condition continues which seems a bit stanger as the need for cooling is under a sharp increase in Poland and the studied region.
- Ventilation: These buildings predominantly depended on natural ventilation methods, such as opening windows and doors. While many houses continue to use natural ventilation, there's an increasing number that have incorporated mechanical ventilation systems, particularly in homes that have been upgraded to meet modern insulation standards.
- Domestic Hot Water (DHW): In original Systems coal-fired boilers were commonly used not only for heating but also for supplying hot water. Some homes also had stand-alone electric water heaters. Today, the trend leans towards using gas-fired boilers or even solar water heaters for DHW, both of which offer improved energy efficiency and a more consistent hot water supply.

More details about this type of buildings and the overview of the whole Polish building stock will be given in the case study section. These information coupled with the occupants profile in the format of user schedule have been defined to have a complete set of EAD additions (Table 2.3.).

Table 2.3. User profile variables

Variable in Honeybee component	**Description**
_room_or_program	Identifies the space type or function (e.g., office or bedroom)
occupancy_sch_	Indicates when the space is occupied
activity_sch_	Describes intensity of activities in the space in Watts
Lighting _sch_	Dictates when lights are on or off
Electric_equip_sch_	Shows when electric equipment is in use
Gas_equip_sch_	Represents operation times for gas-powered equipment
infiltration_sch_	Details air leak schedules into/out of the space
ventilation_sch_	Indicates mechanical ventilation operation times
Heating_setpy_sch_	Provides heating temperature setpoints
Cooling_stp_sch_	Gives cooling temperature setpoints

In energy simulations, it's crucial to recognize that while technologies, materials, and design play a significant role, it's ultimately user behavior that often determines the real-world outcomes. This means that even with state-of-the-art energy-efficient technologies, user behavior can either

augment or undermine the building's intended energy performance [72]. Two identically constructed spaces can present starkly different energy profiles solely because of variations in user behavior [73]. This could range from how often lights are left on, the temperature preferences set on thermostats, or even the frequency of window openings [74]. Such daily actions, while seemingly minor, cumulatively dictate a building's energy footprint .
Hence, the user profile in Honeybee is not just a feature; it's an acknowledgment of the paramount importance of user behavior. Without thoroughly accounting for it, our energy models and predictions remain notably abstracted from reality. To truly optimize for energy efficiency, it's imperative that architects, engineers, and designers deeply integrate an understanding of user behavior into their strategies, ensuring that buildings not only resonate with technological efficiency but also align seamlessly with the behavioral patterns of their occupants [75].

2.5.5. Energy Simulation

In the current phase of our big data generation process, the outcomes derived from the preceding steps are consolidated and structured in the format of a Honeybee room. This formatted data is then provided as an input to the Honeybee Annual Load Component (Figure 2.14.). Crucially, this component leverages the capabilities of the EnergyPlus engine, a sophisticated simulation software, to determine the heating, cooling, and lighting loads of the modelled space across an entire calendar year. It is worth noting that while the Honeybee Annual Load Component possesses the capability to compute loads from various equipment and can also account for domestic hot water demand, such computations are not the primary focus of this research. Thus, these particular calculations have been excluded from the purview of our study.

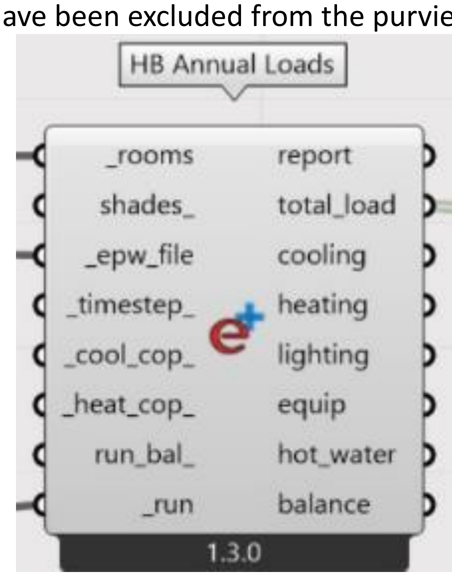


Figure 2.14. Honeybee annual load components

2.5.6. Automation

In the context of this research, a meticulous and comprehensive analysis of 300 potential options is imperative. Undertaking such an extensive analysis manually would undoubtedly be a daunting task, given its sheer scale and complexity. To circumvent this manual labor and introduce a degree of efficiency and automation into the process, we adopted the use of an advanced Grasshopper plugin named "Colibri" (Figure 2.15.). Colibri boasts a carefully crafted workflow that aids in seamless automation. At the outset, input variables, which are represented by a number slider, provide a concise overview of the variable range, highlighting its minimum, maximum, and specific increments or steps. These inputs are then channeled into the Colibri Iterator component. This Iterator is designed to calculate the total number of iterations, keeping in consideration the total items within each range and the cumulative count of impactful variables.

It's crucial to note that for every iteration, a holistic assessment is conducted. Beyond merely analyzing output variables like cooling and heating loads, there's a concerted effort to scrutinize both direct and indirect input values. These encompass parameters such as height, width, length, rotation, and the window to wall ratio. Furthermore, attributes like relative compactness, surface area, volume, and roof area are meticulously documented. Such an expansive approach ensures that all pertinent details of one iteration are captured, providing a rich data set for further evaluation. To effectively capture and store these outputs, the Colibri Parameters component is employed. This dedicated component archives the desired output after each simulation. Bringing the process to a close, the Colibri Aggregator consolidates the inputs and outputs, presenting them in an organized CSV format. This resultant file, with its depth of information, becomes a cornerstone for subsequent analyses, laying the foundation for insightful deductions and evaluations.

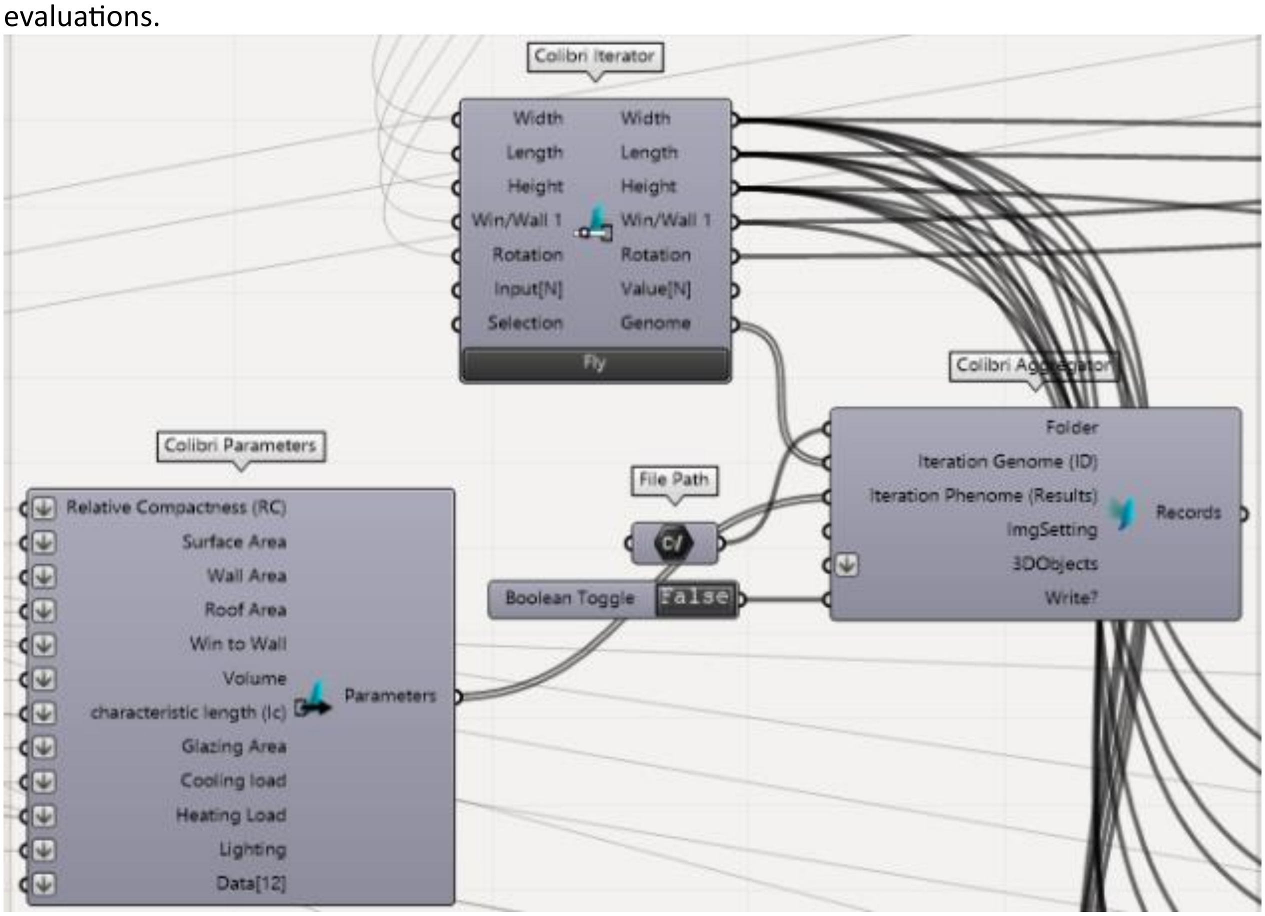


Figure 2.15. components of Colibri for automation

2.6. Climate Change Consideration

Within the ambit of this research, a salient objective has been to assess the ramifications of climate change on the performance metrics of buildings. While a comprehensive discourse on this subject will be delved into in Chapter 4, the purpose of this section is to proffer a concise overview of the workflow that underscores the need and validates the relevance of considering climate change effects on architectural edifices. Initially, a sweeping literature review was embarked upon to discern the interplay between meteorological variables and building energy performance. The nuances and multifaceted dimensions of this relationship, gleaned from an extensive body of existing research, are presented in depth in Chapter 4. Such an examination was deemed necessary to establish a robust foundational understanding upon which the subsequent analyses could be built. Post the literature review, the study ventured into an exhaustive simulation-based examination. This was orchestrated with the intent to gauge, with empirical precision, the impacts of shifting climatic patterns on the energy consumption metrics of buildings. Specifically, the geographic locus for this segment of the study was Poland, with a micro-focus on Poznań as a representative case study for this research. The choice of Poznań was predicated on its emblematic climatic conditions, making it an apt ground for such an inquiry.
What adds a layer of depth to this investigation is that it wasn't confined to a singular building archetype. On the contrary, the simulation encompassed an array of 16 diverse building typologies. By casting such a broad net, the intention was to derive a panoramic view of the potential impacts of climate change on various structural categories. This was instrumental in establishing a preliminary validation framework. Having acquired insights at this macro level, the research then converged its focus onto a specific building cluster, which remains the cynosure of this research. Through this phased and tiered approach, the study aims to present a comprehensive and nuanced understanding of the subject matter.
In this regard the methodology of this part can be presented as Figure 2.16. where all three steps of this part have been described.

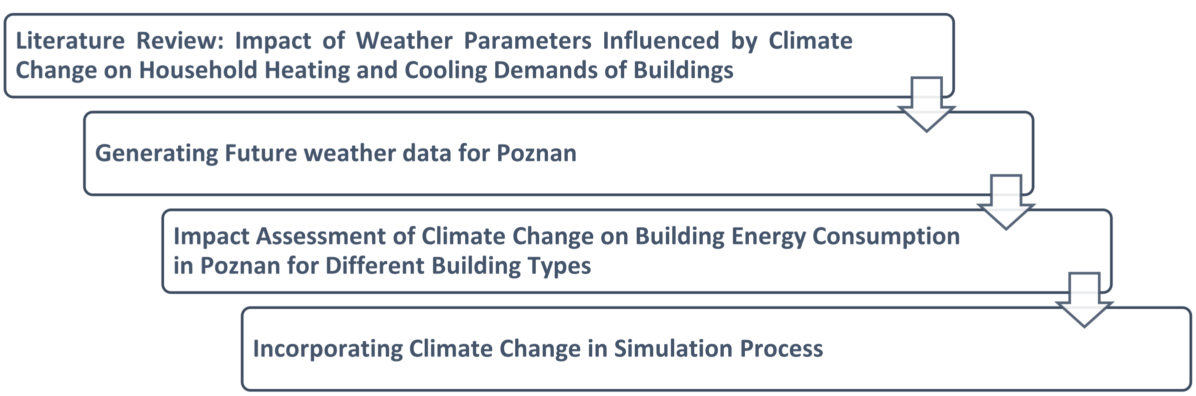


Figure 2.16. stages of considering climate change in the study

While the first stage of this process is straightforward in terms of methodology other parts need clarification. For this part research, a distinct set of historical data, encapsulated in the form of a Typical Meteorological Year (TMY file), was utilized. This particular file was selected due to its recency, especially when compared to other TMY files available for various locations throughout Poland. For the foundational phase of the study, a weather file sourced from the Ławica Airport weather station with the normal data from 2008 to 2018 (Table 2.4.) in Poznan was adopted as the primary benchmark for subsequent projections.

Table 2.4. used data for each month of the weather file.

Month	Representative year	Month	Representative year
January	2012	July	2012
February	2015	August	2011
March	2018	September	2012
April	2010	October	2008
May	2007	November	2012
June	2016	December	2008

Data for this step was derived from Global Climate Models (GCMs) and subsequently integrated into the Climate Change World Weather Generator (CCWorldWeatherGen) tool, a notable innovation by the Sustainable Energy Research Group (SERG) at the University of Southampton, England. Originally developed by Jentsch [76, 77], this Microsoft® Excel-based tool, known formally as the 'Climate Change World Weather Generator' and often referred to as CCWorldWeatherGen, harnesses the output data of HadCM3 (a contribution from the Met Office Hadley Centre for Climate Science and Services; 2010). The IPCC A2 emission scenario is employed to formulate prospective weather files through the application of the morphing method. A comprehensive assessment was conducted by Jentsch et al., where 23 GCMs under AR4 and six GCMs under AR3 were examined globally. It was determined that the HadCM3, in conjunction with the A2 emission scenario, was well-suited for the morphing technique. The A2 emission scenario, as outlined by the IPCC AR3, paints a trajectory characterized by an ongoing surge in the global population and a regionally focused economic growth. Consequently, the utilization of HadCM3, aligned with the A2 emission scenario, for the CCWorldWeatherGen tool that incorporates the morphing method, was deemed appropriate.

The CCWorldWeatherGen tool, due to its inherent capabilities, provides researchers with an efficient projection regarding future climatic trajectories [78]. Upon integration of the TMY file into the CCWorldWeatherGen, future climatic data for both 2050 and 2080 were generated, adhering to the principles of the A2 scenario. The data extracted from GCMs was then introduced into the CCWorldWeatherGen, facilitating a statistical downscaling of the baseline and the generation of weather data for subsequent scenarios. For the duration of this study, a presumption was made that building life cycles endure a minimum of 60 years. This suggests that no significant alterations would be observable in building components' performance, especially concerning thermal loads, during the entirety of the analysis period.

The study's progression led to building energy modeling being executed using EnergyPlus 9.0.1. The objective here was to simulate the potential consequences of climate change on the thermal loads of buildings, with the city of Poznan as the focal point. To ensure an encompassing perspective, 16 building prototypes as defined by the ASHRAE standard 90.1 were incorporated into the study. These prototypes, derived from the DOE's Commercial Reference Building Models, underwent modifications inspired by the Advanced Energy Design Guide series and the ASHRAE 90.1 committee. Detailed nuances and modeling methodologies related to these prototypes can be found in the reports released by the Pacific Northwest National Laboratory (PNNL) [79, 80] as in Figure 2.17.

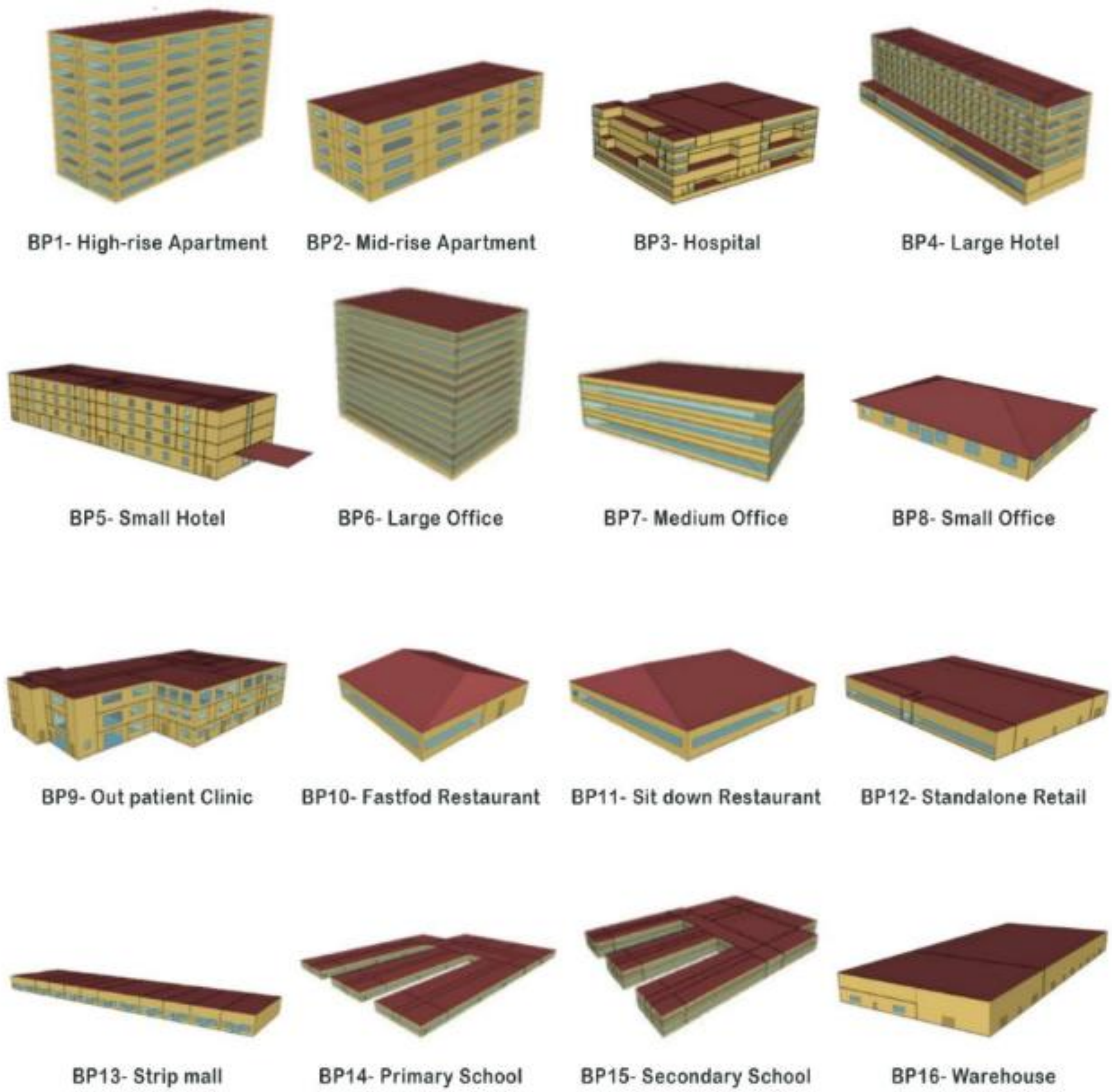


Figure 2.17. Building prototypes from the ASHRAE 90.1 standard [81]

Endowed with authentic building characteristics, these 16 prototypes were simulated using both current and anticipated weather data to gauge the relative effects of climate change on their energy performance metrics. A meticulous detailing of the prototypes' envelope components has been documented and is accessible in Table 2.5.

Table 2.5. Technical description of the prototype envelope

Building Type		U-value (W/(m2 K))			
		Roof	External Wall	Glazing	
				Window	Skylight
Apartment	High-rise (BP1)	0.18	0.31	2.65	-
	Mid-rise (BP2)	0.18	0.31	2.65	-
Hotel	Large (BP4)	0.18	0.45, 0.51	2.65	-
	Small (BP5)	0.18	0.31	2.65	-
Office	Large (BP6)	0.18	0.51	2.65	-
	Medium (BP7)	0.18	0.31	2.65	-
	Small (BP8)	0.15	0.29	2.65	-
Medical	Hospital (BP3)	0.18	0.45, 0.51	2.65	-
	Outpatient (BP9)	0.18	0.31	2.65	-
Restaurant	Fastfood (BP10)	0.15	0.29	2.65	-
	Sit-down (BP11)	0.15	0.31	2.65	-
Retail	Stand-alone (BP12)	0.18	0.51	2.65	2.96
	Strip-mall (BP13)	0.18	0.31	2.65	-
Educational	Primary School (BP14)	0.18	0.31	2.65	-
	Secondary School (BP15)	0.18	0.31	2.65	2.96
Warehouse	(BP16)	0.21	0.28, 0.47	2.65	2.96

Upon the meticulous execution of energy simulations for each individual prototype, utilizing both contemporary and forecasted climate weather files, a detailed comparative analysis was conducted. This analysis not only can reveal the influence of climate change on the thermal load within architectural structures but also underscored the intricate relationships between building characteristics and changing climatic conditions. Shedding light on these nuances, particularly the thermal responses of buildings to evolving climatic scenarios, was a pivotal and central objective of this comprehensive research document. Through this exploration, deeper insights were garnered, contributing significantly to the broader discourse on building resilience and climate adaptation. Therefore, the whole process of this step can be presented as in Figure 2.18.

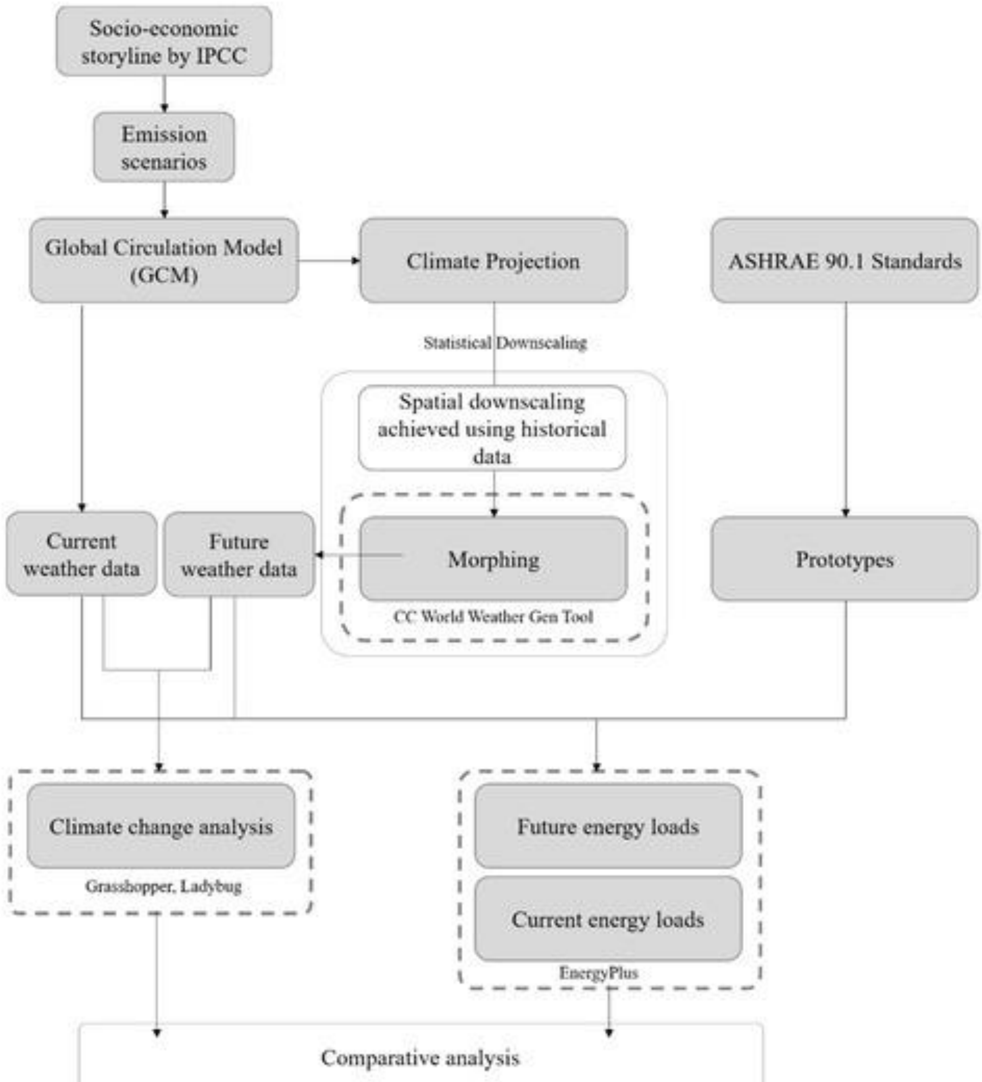


Figure 2.18. Workflow of incorporating climate change consideration

Finally, after careful consideration of the potential impacts of climate change on the behavior of buildings in terms of energy consumption, the generated weather file was used for the simulation and generating the big data.

2.7. Data-driven Methods

The adoption of surrogate data-driven models has been advocated as a solution to the challenge posed by very long energy simulation procedures, as noted in the work of Zhao & Magoules [82]. In essence, these models encapsulate mathematical relationships that capture the dynamics between specific inputs and desired outputs in the system under investigation. These relationships are often formulated based on empirical or simulated data that represent the physical phenomena being studied. For instance, the thermo-physical attributes of construction materials coupled with meteorological variables can serve as predictors for indoor environmental conditions, an approach that is pursued in the current research endeavor. Models that demonstrate a high level of precision act as expedient and accurate substitutes to traditional building performance simulation tools, particularly in scenarios that demand substantial computational resources [83]. The employment of such surrogate models necessitates a rigorous evaluation of both the data's reliability and the legitimacy of the inferred relationships. In the context of the present research, an analytical lens is placed on a specific facet of this methodology, namely the process of selecting and fine-tuning regression models tailored to a given dataset. This entails the deliberate choice of model types, structural configurations, and critical parameters that align optimally with the problem scope. While extant literature often contrasts linear models with their nonlinear counterparts or evaluates different genres of nonlinear models in the domain of building simulation [84], such studies usually concentrate on a restricted set of model parameters. The performance of a model in handling a specific dataset is fundamentally contingent on the selection of these parameters, a performance metric that inherently fluctuates across diverse datasets. Consequently, earlier research has not rendered an exhaustive critique of various nonlinear models, nor has it furnished ample guidelines regarding model selection criteria. In the present work, it is posited that the act of model selection ought to consider a myriad of factors, encompassing predictive accuracy, model intricacy, user-friendliness, and the stability of the model's predictions. In the current section, a very brief description of the workflow of the deployment of the data-driven methods in analyzing the generated data in previous steps is presented (Figure 2.19.).

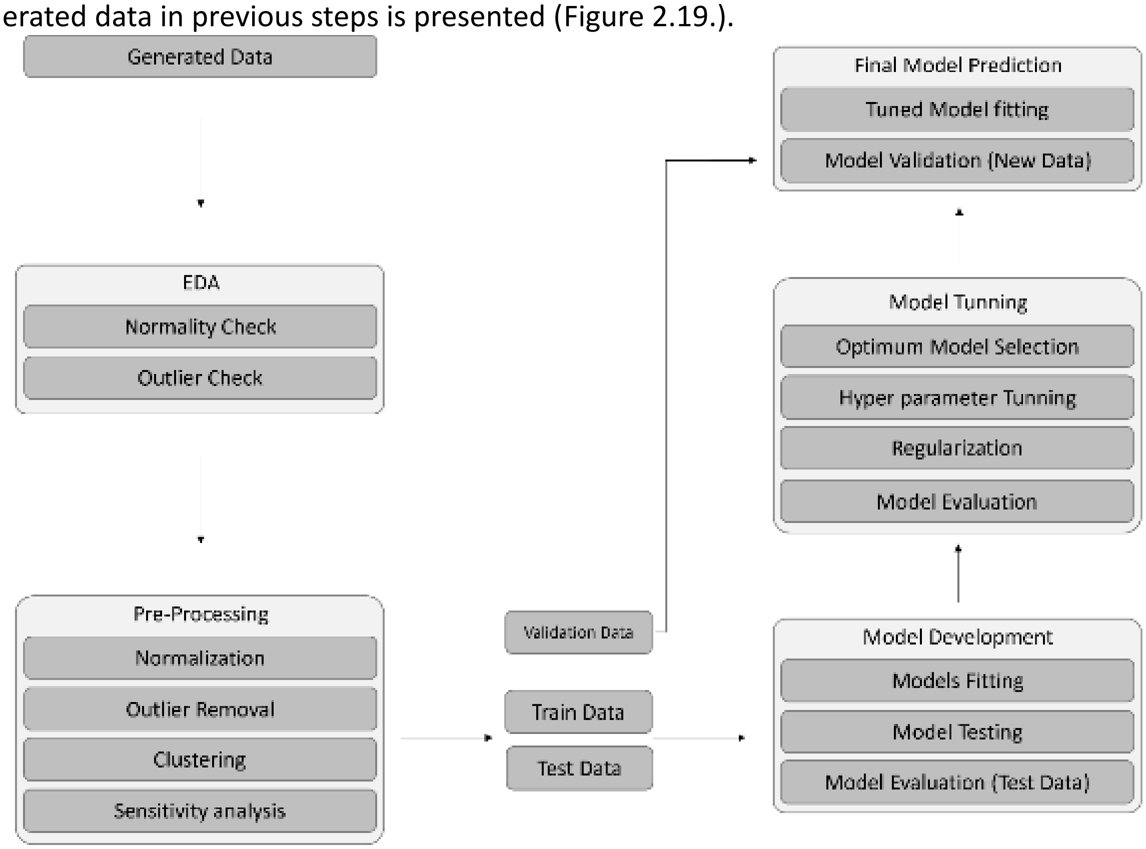


Figure 2.19. Data-driven method deployment workflow

The research approach was organized into several key stages to ensure the quality and rigor of the study. Initially, an expansive dataset generated in previous phases of the research served as the foundational input for data-driven methods, primarily employing Python and associated packages. The initial stage was dedicated to Exploratory Data Analysis (EDA), where the normality of the data was scrutinized, and outliers were assessed to determine if corrective measures were needed. Following the EDA, data pre-processing commenced. During this stage, data normalization procedures were executed, and outliers were systematically removed. Additionally, clustering techniques were employed, not only to improve the understanding of the data but also to introduce new categorical variables that could offer a more nuanced description of energy consumption behaviors. Furthermore, a sensitivity analysis was undertaken at this stage to reveal the most influential variables and to quantify the magnitude of their impact on energy consumption rates.

After the pre-processing phase, the data was segregated into training, testing, and validation sets. These subsets were earmarked for use in various machine learning (ML) models. The next phase involved model development; wherein multiple models were calibrated using the training data. An initial evaluation was then conducted using the testing data, providing preliminary feedback on the performance of each model. This feedback became instrumental in the next stage of model tuning. Hyperparameter tuning was executed, followed by the implementation of regularization techniques to avoid overfitting. Another round of model evaluation was conducted through cross-validation, ensuring the robustness of the selected models.

Lastly, the model boasting the best-tuned hyperparameters was applied to the training dataset. Its performance was subsequently assessed using the test dataset. Ultimately, the final predictive model was validated using the separate validation dataset. This rigorous process culminated in definitive conclusions regarding the applicability and reliability of the data-driven method in the context of building energy consumption.

2.8. Integrated Workflow

Upon completion of the aforementioned stages, it became critically important to establish seamless connections between each phase of the research to ensure the integrity and effectiveness of the overall process. To achieve this, the simplified 3D geometry that was extracted from the initial data was imported into the Grasshopper software environment (Figure 2.20.). Within Grasshopper, a series of computational operations were performed to isolate specific architectural and design features from the 3D geometry. These features encompassed a range of parameters, including but not limited to, the building's height, length, and width, as well as nuanced variables such as the window-to-wall ratio, characteristic length, and relative compactness, among others.

Once these features were accurately extracted, they served as the foundational input variables for the machine learning model that had been rigorously trained to predict energy consumption rates. The significance of these selected features cannot be overstated, as they represent key determinants in the model's ability to generate reliable energy consumption forecasts.

It is worth noting that during this integration stage, special attention was given to validation procedures. The validation data, which had not been a part of the original training or testing datasets, was also subjected to energy consumption simulations. These simulations were executed with the explicit aim of comparing the results against those produced by the data-driven model. This comparative analysis was undertaken as a crucial step in validating the performance of the machine learning model, thereby bolstering the reliability and generalizability of the research outcomes. This dual-pronged approach of both simulation and data-driven modeling provided a comprehensive and robust validation framework, reinforcing the credibility of the model's predictive capabilities.

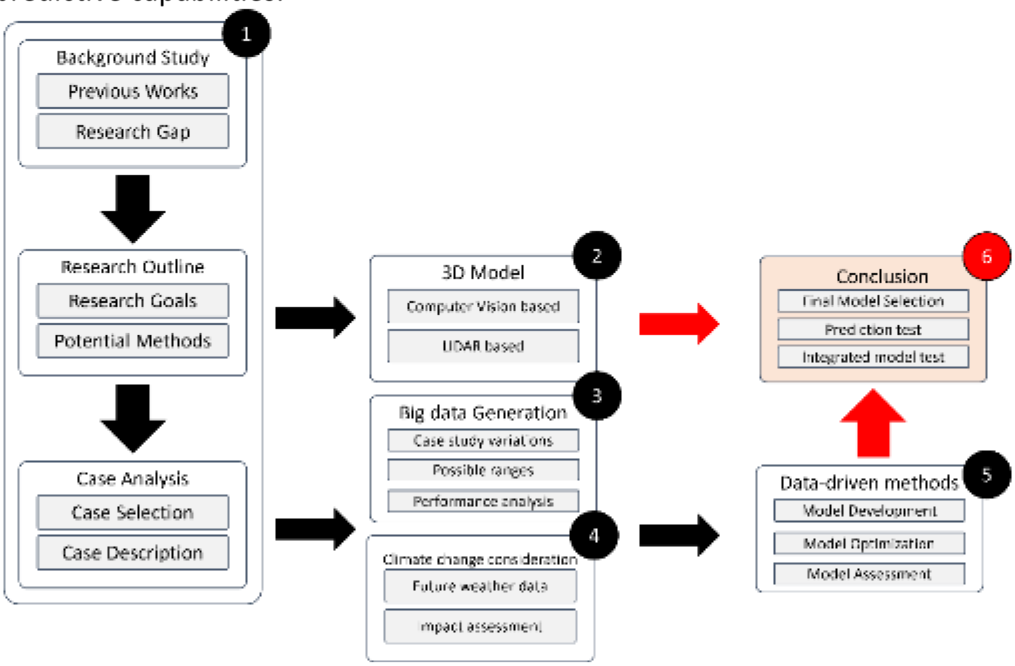


Figure 2.20. integration part of the workflow

Chapter 3: Research Context

3.1. Abstract

This chapter provides a detailed analysis of the energy efficiency of residential buildings in Poland, homing in a specific cluster of buildings constructed between 1945-1970 in the Wielkopolski region, especially in Poznan. The chapter categorizes buildings into various clusters based on their construction years and assesses their energy consumption patterns, revealing the 1945-1970 cluster as particularly energy-inefficient and therefore a priority for immediate intervention. The chapter evaluates various components of building energy consumption, such as building envelopes, HVAC systems, Domestic Hot Water (DHW), and potential renewable energy solutions. It also exposes Poland's deviation from EU energy consumption norms, particularly the country's significant reliance on coal. This situation underscores the urgency for energy-efficient upgrades, especially in rural areas with less stringent building codes. Within the specific focus on Poznan's 1945-1970 building cluster, this chapter discusses their unique architectural and construction characteristics, ranging from high-quality bricks to outdated heating solutions. It recommends data-driven methodologies like energy audits and machine learning for identifying the most impactful and cost-effective retrofit options. The overarching aim of this chapter is to offer targeted, actionable recommendations for improving the energy efficiency of this key building cluster. These proposed changes aim not only to benefit the individual owners of these buildings but also to contribute to broader environmental and economic stability objectives within Poland.

3.2. Introduction

Poland finds itself at a pivotal juncture, facing significant challenges and opportunities in its energy market, which is currently heavily reliant on coal (Figure 3.1.) even for residential heating (Figure 3.2.). Amidst increasing pressures from European Union policies to transition towards more sustainable energy sources, the country has set an ambitious target to phase out coal and shut down all its coal mines by 2049. This decision is particularly notable given that Poland is one of the fastest-growing economies in the Euro Zone, with its construction sector expected to expand rapidly. According to the European Construction Sector Observatory (ECSO), this sector is predicted to grow by 5.4% between 2020 and 2021 [85]. This upward trajectory in the construction sector is further substantiated by the data on commissioned apartments, which have seen a significant increase from 2014 to 2018. The Polish government has been proactive in implementing policies aimed at decarbonizing its energy sector, thereby positioning the country on a sustainable growth path. The role of decarbonizing the building sector in Poland is critically important in meeting the EU's climate and energy objectives for 2030 and 2050, as buildings in Poland account for a substantial 38% of total energy consumption [86] and contribute to 33% of energy-related greenhouse gas emissions [87].

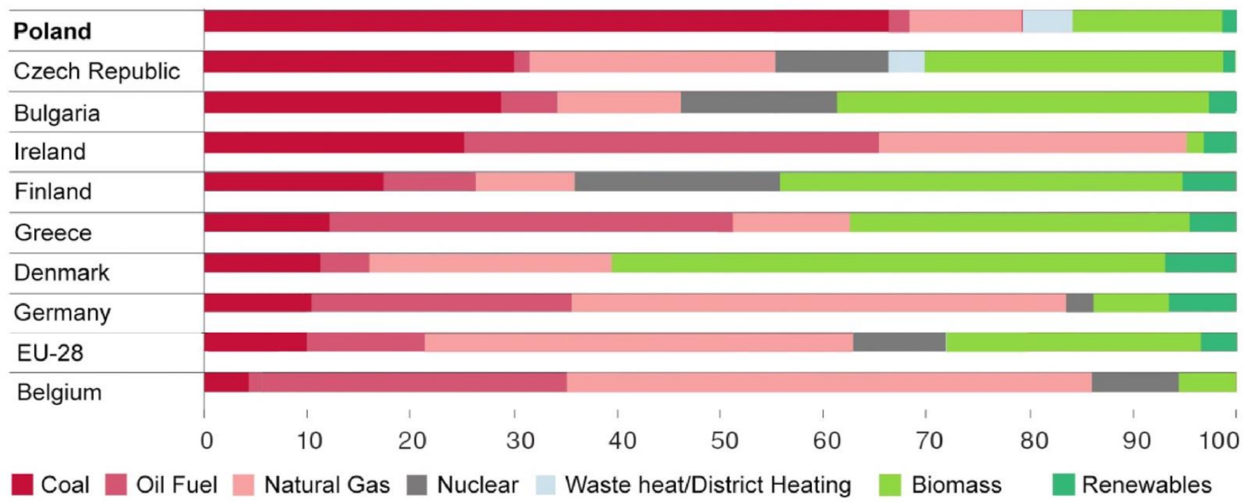


Figure 3.1. Fuel share of residential heating in Europe, based on [88]

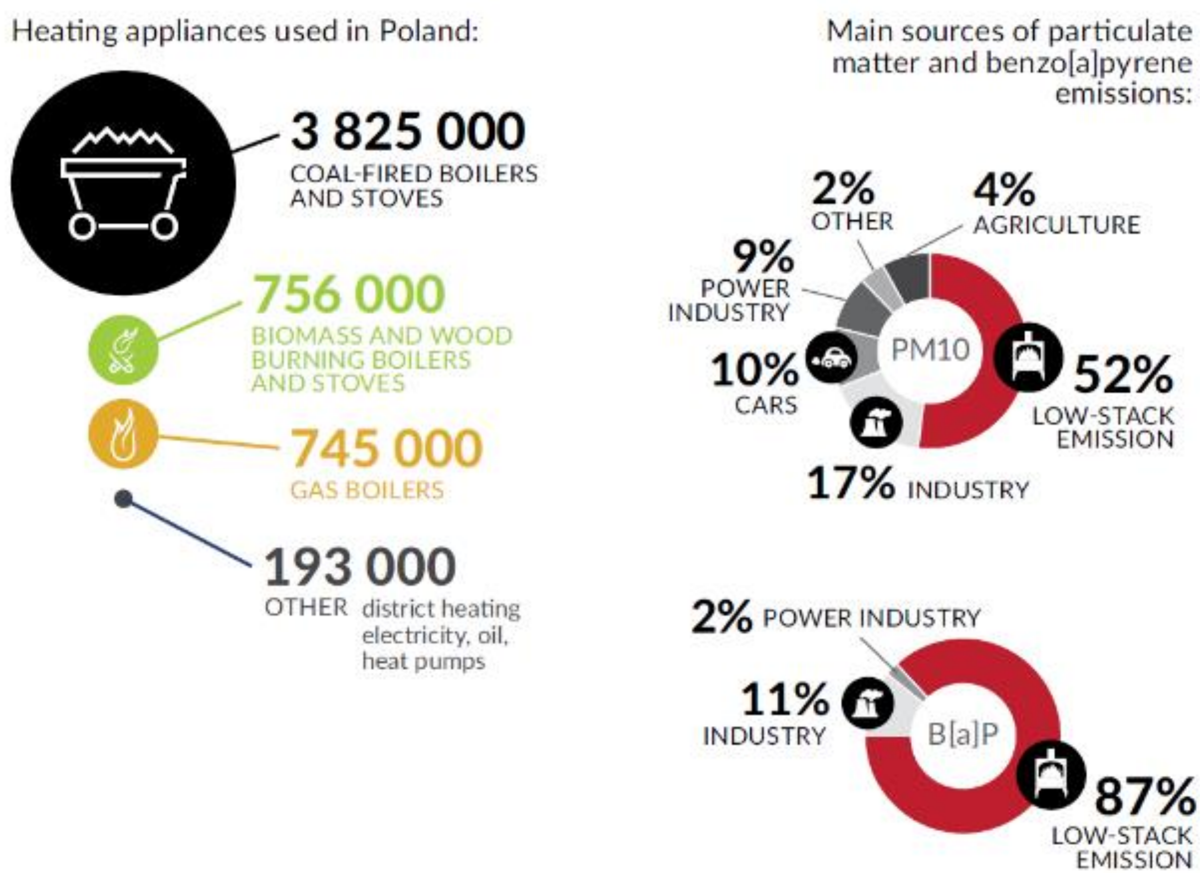


Figure 3.2. Heating appliances and main sources of pollution in Poland ,based on [86]

The residential sector is particularly worrisome; it depends more on solid fuels than any other EU country, contributing to deteriorating ambient air quality and raising concerns over energy security [89]. To comply with the European Union's net 55% emission reduction target by 2030, Poland must aim for a 20% reduction in energy-related greenhouse gas emissions from buildings compared to 2015 levels, a goal far more ambitious than the 7% initially outlined in the National

Energy and Climate Plan (NECP) [90]. Between 2008 and 2018, Poland ranked 21st among EU nations in terms of energy efficiency of households with an annual rate of energy efficiency improvements at 1.13%, it fell below the EU's average, and even more dramatically behind the improvement rates observed in the industrial sectors of Poland [88, 91]. However, there are initiatives in place for improving energy efficiency in multifamily buildings, and Poland does have an energy efficiency obligation scheme [92]. The residential sector remains the largest cluster of residential buildings (Table 3.1) and the highest energy consumer, mainly fueled by single-family buildings.

Table 3.1. Technical description of the prototype envelope, based on [93]

			Including						
				of which					
					of which				
		Total	**Inhabited**	**Residential**	**Single-family**	**Multi-family**	**Collective accommodation**	**Non-residential**	**Uninhabited**
Total	**Thousands**	6,047.1	5,567.6	5,542.6	5,007.5	535.1	3.3	21.0	479.5
Urban Areas		2,285.6	2,189.2	2,176.4	1,738.2	438.2	1.8	10.8	96.4
Rural Areas		3,761.5	3,378.4	3,366.2	3,269.3	96.9	1.4	10.3	383.1

Data from the National Energy Conservation Agency (KAPE), part of the TABULA project, reveals that nearly half of Polish households (approximately 5.5 million households) reside in single-family homes [94]. These homes predominantly rely on coal, biomass, and waste for heating and domestic hot water (DHW). The TABULA project offers a typological classification of residential buildings in Poland, categorized into single-family houses (SFH), terraced houses (TH), multifamily houses (MFH), and apartment blocks. Each category has been further divided into seven types based on building traditions and insulation levels [94]. In light of these complexities, this chapter aims to provide a detailed snapshot of the current state of energy efficiency in Poland's residential building sector as well as the studied cluster, which stands as one of the significant sources of pollution and energy consumption. The objective is to assess the energy efficiency landscape in this crucial sector, offering insights that could inform both policy and practice.

3.3. Financial Supports

The subject of financial factors influencing buildings' energy efficiency in Poland has been relatively under-researched, even though it remains a critical determinant for policy formulation and investment decisions. While there are a few studies that have delved into this area, their scope, focus, and findings are varied, leading to gaps in comprehensive understanding. In 2013, the Building Performance Institute Europe (BPIE) led one of the earliest initiatives to examine Poland's approach to cost-optimality in building energy efficiency [95]. The study revealed a considerable discrepancy between Poland's U-values and those suggested by the European Union's cost-optimal methodology [96]. The incongruence indicates that Poland's energy standards, particularly for gas and coal, are less stringent than what the EU recommends. This discovery raises questions about the effectiveness and adequacy of Poland's current regulatory framework concerning building energy efficiency.

In another study conducted by Basinska et al. in 2015 [97], the authors explored the optimal energy requirements for residential buildings in Poland over a span of 30 years. The research was rooted in Polish energy standards [98] and applied a comprehensive global cost calculation method. The study incorporated 28 technical variations that spanned envelope, HVAC systems, and economic facets of energy sources and pricing. Despite its thoroughness, the study stopped short of ranking energy-efficient measures by their payback time or overall cost-effectiveness. Additionally, it did not align itself with the EU's cost-optimality framework [98], but relied instead on Poland's 2008 energy efficiency regulations. A further contribution to this area was the 2016 BPIE status report [93], which provided a robust account of existing financing mechanisms for building energy performance improvements in Poland. This study reviewed operational funding schemes like the Thermo-modernization Fund, RY'S Fund, the Clean Air Fund (2018–2029), and others including KAWKA and SMEs-focused programs. It also examined the role of the National Fund for Environmental Protection and Water Management (NFEP&WM) in this context. The study assessed the effectiveness of current renovation technologies in Poland and evaluated them against three proposed renovation scenarios.

In 2019, Firląg [99] embarked on an in-depth study examining the cost-effectiveness of energy performance in single-family residences. The central objective of this research was to develop guidelines for Plus Energy Buildings (PEB) specifically tailored to the Polish housing landscape. Utilizing Poland's 2015 energy efficiency standards as a foundation [98], the study explored multiple avenues for boosting energy efficiency and incorporating renewable energy solutions in various building subsystems. These subsystems included thermal insulation components, ventilation infrastructure, and heating systems. The research offered pivotal insights and suggestions for new regulations slated to be introduced in 2021. These forthcoming rules specify that the annual demand for non-renewable primary energy (which encompasses heating, ventilation, cooling, domestic hot water provisioning, and lighting) should not exceed 70 kWh/m^2 per annum. Nonetheless, the study fell short in defining the cost-optimal benchmarks explicitly considered in Polish energy efficiency laws.

Fast-forwarding to 2020, Adamczyk et al. [100] investigated the economic implications of medium-level Thermo-Modernization initiatives for single-family houses. Their research unambiguously confirmed the absence of substantial financial gains for homeowners who opt for medium-level thermal upgrades. Especially when contrasted with conventional coal, biomass, and waste boilers, the elevated initial costs of efficient alternatives, such as natural gas and fuel oil boilers, inhibit any meaningful economic benefits for substantial renovations. Similarly, Golabeska [101] conducted an economic assessment to determine the expenses associated with constructing and maintaining a Passive House (PH)-certified dwelling. The results showed that despite higher initial costs relative to traditional construction, the economic viability of PH-certified homes becomes apparent only when operational costs are considered over a long-term horizon exceeding 30 years.

Supplementing these insights, Ksiezopolski et al. [91] carried out a study in 2020 focused on the potential for both energy and financial savings achievable through the renovation of rural single-family homes in Poland. Their research emphasized the implementation of various energy conservation strategies, such as enhancing the building's thermal envelope and transitioning from coal-based boilers to cleaner alternatives like gas, electric, and heat pump/Photovoltaic (PV) systems. The study concluded that, under the Polish Prosumer support mechanism, heat pumps powered by PV installations emerged as the most economically efficient heating solution for single-family homes. However, without the availability of green subsidies or other financial incentives, these systems remained economically unfeasible. Intriguingly, none of these two more recent studies [91, 101] adhered to the European Union's prescribed cost-optimality methodology, thereby marking a deviation from a standardized European approach to such analyses.

The existing research has successfully integrated the assumptions related to green subsidies and financial support funding into the evaluations of cost-effectiveness in the context of Poland's energy landscape. However, these studies exhibit limitations concerning the representation of national architectural archetypes. Ideally, the selection should have been founded on a thorough characterization of Poland's building typology [94]. Additionally, the studies fell short in aligning their future renovation or new construction scenarios with upcoming energy efficiency regulations and the requirements set forth by the Energy Performance of Buildings Directive (EPBD) concerning nearly and net-zero energy buildings.

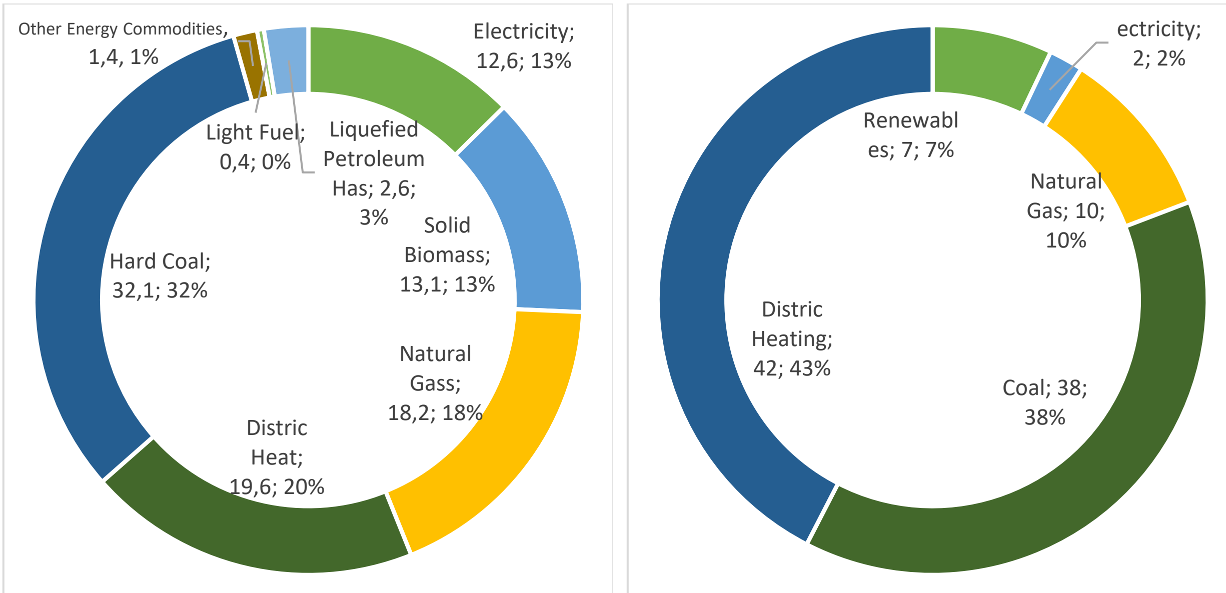
Consequently, Poland has yet to fully adopt the European Union's cost-optimal methodology, which aims to standardize the determination of energy-efficient measures in building construction and renovation [102]. Only a handful of studies have scrutinized the role that subsidies play in influencing the adoption of energy-efficient measures in both newly constructed and renovated residential settings [91].

3.4. Energy Sources

In Poland, non-renewable energy sources such as hard coal and natural gas continue to be the prevailing options for energy, particularly in the heating sector [103]. Over 40% of Polish households rely on district heating systems [86]. A detailed breakdown of energy consumption patterns in these households is provided in Table 3.2 [103]. As illustrated by the data, hard coal remains the principal fuel source for district heating, even though its usage dropped marginally from 86.7% in 2011 to 81.6% in 2017. When it comes to meeting the energy needs for space and water heating, coal still constitutes a significant 39% of the total energy share, as shown in Fig. 3.3.

Table 3.2. Energy Use by Source and Purpose in Polish Households for 2018, based on [103]

Energy community	Unit of Measure	Total	Space Heating	Water Heating	Cooking	Lighting + Electrical Appliances
Electricity	GWh	29284	1305	2118	3168	22693
	TJ	105422	4698	7625	11405	81695
Heat	TJ	157000	107250	49750	X	X
Natural gas	TJ (GCV)	165679	88532	43892	33255	X
Solid fuels	Thousands tons	10430	9365	935	130	X
	$TJ10^3$ t	267440	240132	23975	3333	X
Petroleum products	Thousands tons	580	90	34	456	X
	$TJ10^3$ t	26440	3930	1534	20976	X
Of Which:						
LPG	Thousands tons	500	20	24	24	X
	$TJ10^3$ t	23000	920	1104	20976	X
Heating Oil	Thousands tons	80	70	10	X	X
	$TJ10^3$ t	3440	3010	430	X	X
Energy from renewable sources Of Which:	TJ	112675	98676	11784	2215	X
Solar energy	TJ	2129	106	2023	X	X
Solid biofuels excluding charcoal	TJ	108015	96800	9000	2215	X
Geothermal energy and ambient heat	TJ	2531	1770	761	X	X
Energy sources in total	TJ	834656	543218	138560	71184	81695



a. b.

Figure 3.2. a. Structure of households energy consumption by various energy commodities, b. hare of energy sources used for heat demand in the residential sector

A study examining the current state of Poland's heating market revealed that the existing energy infrastructure is antiquated [104]. This outdated system is plagued by inefficiencies in energy production and elevated levels of pollutant emissions. Accordingly, there is an immediate necessity for modernization efforts, which should encompass a thorough evaluation of transmission losses. Furthermore, an improvement in the efficiency of district heating systems could be achieved by better matching the heat capacity demanded by users [104]. On a more positive note, there has been a discernible shift toward renewable energy adoption in Poland's district heating systems in recent years. Investments in the maintenance and development of local heat sources and distribution networks tailored for residential structures are currently underway [105]. At the individual building level, a study by Krawczyk in 2016 [29] examined natural gas usage in a renovated Polish household for heating purposes. The research offered recommendations for energy conservation that align with the Polish energy efficiency regulations of 2015, which set the primary energy consumption factor at 120 kWh/m^2 per year.

Even more intriguingly, a study by Ksiezopolski et al. in 2020 investigated the potential of transitioning from coal to cleaner energy sources for heating [91]. Specifically, the study probed the impacts of replacing traditional coal boilers with alternatives powered by gas, electricity, and heat pump/Photovoltaic (PV) systems. The research also explored the feasibility of incorporating zero-emission heat sources, such as PV/heat pump installations, especially in rural areas of Poland. According to the data presented in Table 3.3, the conversion factors for calculating primary energy from solid fuels in Poland are notably low. These factors are dictated by the Regulation of the Minister of Infrastructure and Development as of March 18, 2015, which outlines the methodology for determining a building's energy performance. The study highlights that, in the absence of financial incentives, transitioning away from coal in residential settings would not be economically feasible. Existing research, although limited, has begun to tackle the issue of reducing carbon emissions through heat generation for buildings and substituting conventional energy sources with renewable and other non-conventional alternatives.

Table 3.3. Non-Renewable Primary Energy Input Factors for Energy Carrier, based on [106]

No.	Method of supplying a building or part of a building in energy	Energy carrier type or energy type	coefficient of expenditure
1	Local energy production in the building	Fuel oil	1.10
2		Natural gas	
3		Hard coal	
4	District heating from cogeneration	Hard coal or natural gas	0.80
5		Biomass, biogas	0.15
6	District heating from a local heating	Hard coal	1.30
7	plant	Natural gas or fuel oil	1.20
8	Electrical grid system	Electricity	3.00
9	Local renewable energy sources	Solar energy	0.00
10		Wind energy	
11		Geothermal energy	
12		Biomass	0.20
13		Biogas	0.50

Poland's domestic energy landscape presents a rather paradoxical situation. While the use of local coal in electricity production and domestic coal mining are on the decline, the demand for imported coal remains elevated. A study by Zieleniec et al. [107] revealed that Poland imported 10 million tons of hard coal last year from countries like Russia, Chile, Colombia, the USA, and Kazakhstan. Concurrently, the role of gas in Poland's energy portfolio is growing; its share touched nearly 9% in 2018, mostly met through imports. Electricity generation from coal is dwindling, electricity imports are on the rise, and renewable energy is playing an increasingly significant role. Despite these shifts, the country's greenhouse gas emissions remain stagnant, registering over 412 million tons of CO2 equivalents in 2018, including close to 150 million tons solely from the power sector. European Union policies have the potential to steer Poland's governmental strategies toward enhanced energy efficiency and reduced CO2 emissions. A study that analyzed the European Union Emissions Trading Systems' impact on Poland’s conventional energy sector from 2008 to 2020, and projected up to 2050 [108], concluded that gas-fueled combined heat and power units might be less adversely impacted by EU regulations compared to hard coal-fired plants, which could become unprofitable post-2020. Nevertheless, the study emphasizes that Poland can't sidestep substantial decarbonization efforts in its power sector to align with post-Paris climate goals.

The Polish government has responded with its own strategic roadmap, known as the 2030 Natural Environment Policy (PEP2030) [32]. This plan aims to enhance air quality through various measures such as the modernization of district heating networks and the replacement of outdated stoves and boilers. Public awareness and sentiment regarding energy efficiency and conservation are also evolving, as highlighted in another study [33]. Motivations for energy savings are shifting from purely financial considerations to broader concerns like environmental degradation, air quality, and climate change. Although the public tends to focus more on electricity conservation rather than heating, the PEP2030 aims to encourage the use of electricity as a source of heat.

3.5. Policies and Regulations

The history of energy efficiency regulations in Poland has a long trajectory, starting from the post-World War II era. Initial formal guidelines focusing on permissible heat losses were put in place as early as 1957, denoted by the heat transfer coefficient [109]. Figure 3.3 illustrates the significant evolution in Poland's heat transfer requirements, which were ahead of the curve in comparison to many European countries. Notably, Poland updated its heat transfer requirements for walls in 1964, pre-dating the 1972 oil crisis. Subsequent updates in 1974 and 1982 [110] were influenced by the energy challenges arising from the oil crisis [111]. According to research by Wojcik in 2018, the energy usage in standard Polish buildings ranged from 240-380 kWh/m^2 up to the year 1985. A series of progressive reductions were then mandated: 160–200 kWh/m^2 between 1991–1992, further reduced to 120–160 kWh/m2 from 1993–1997, and since 1998 the energy usage has been capped at 90–120 kWh/m^2 [112].

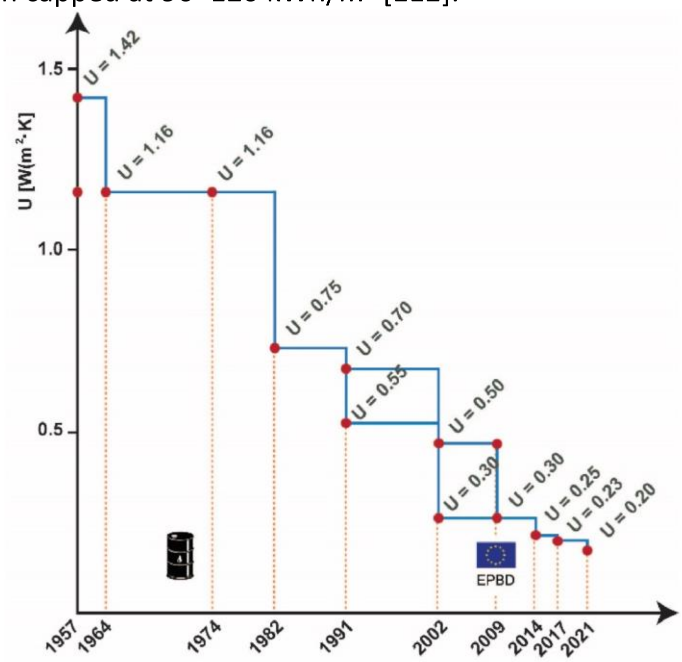


Figure 3.3. Requirements for the heat transfer coefficient U for walls in force in Poland, based on [113]

In an institutional development, the Polish Energy Conservation Agency (KAPE) was established in 1994, as a specialized body under the Ministry of Building Industry, Industry and Trade, and Environmental Protection [114]. KAPE played a pivotal role in formalizing Poland's climatic zoning, dividing the nation into five distinct climate zones as depicted in Figure 3.4. A ministerial committee came into existence in 1996, tasked with pinpointing the challenges and opportunities in building energy efficiency and the incorporation of Renewable Energy Systems (RES).

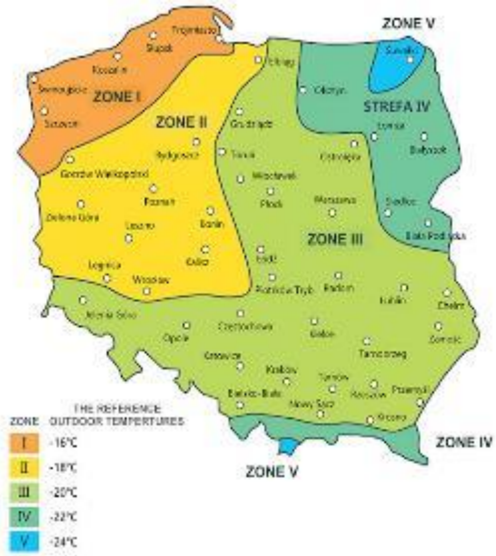


Figure 3.4. Climate zones in Poland, based on [106]

The Construction Law, initially enacted on July 7, 1994, underwent amendments to align with energy efficiency goals. With Poland becoming a member of the European Union in 2004, the European Parliament's Energy Performance of Buildings Directive (EPBD) was adopted by Poland in 2009. Since then, Energy Performance Certificates (EPCs) have become mandatory for new constructions and for buildings that are being sold or rented. These certificates are calculated based on the EN 13790 standard and express a building's energy efficiency through the primary energy use intensity indicator [98].

3.6. Building Envelope and Technologies

Poland's building energy conservation policies have seen a considerable transformation over the past two decades, both in terms of material science and methodology. A 2015 study by Zyczynska et al. tracked the progression of the U-value in building envelope components from 1974 to 2021, highlighting the increasingly stringent regulations [111]. Another study in 2017 by Wojcik et al. specifically dealt with the challenges and solutions for renovating buildings with historical facades [115]. They explored the effectiveness of Autoclaved Aerated Concrete (AAC) as an internal insulation material, particularly for its proficiency in managing heat and moisture.

Grygierek et al. in 2018 offered a detailed methodology to optimally select envelope parameters specifically for single-family houses featuring natural ventilation [116]. The study honed in on cost-effective solutions for building envelopes, including various types of insulation and airtightness techniques. However, it did not delve into nearly Zero-Energy Building (nZEB) targets or advanced envelope technologies. A subsequent 2018 study by Weglarz et al. compared the energy and carbon footprint of three distinct construction technologies: conventional construction, wooden frame building, and straw-bale construction, across a 40-year life cycle [117]. Interestingly, despite having higher energy consumption during its operational phase, the straw-bale technology emerged as the most ecologically sound option. The 2019 study by Kisilewicz et al. offered an intriguing take on active insulation systems as potential substitutes for

traditional passive insulation methods [118]. Using a concrete layer in the external wall embedded with a refrigerant-circulating pipe system, the study found that heat losses could be reduced by an average of 63%. While effective for both new and existing constructions, the system needs to be deactivated during the summer to permit natural cooling of the external walls.
Despite these advancements in energy efficiency research and thermo-modernization projects, the actual quality of most newly constructed and retrofitted energy-efficient building envelopes in Poland leaves much to be desired [119, 120]. Insights from experts and the observations show systemic issues like thermal bridges that lead to heat loss due to inconsistencies in the thermal envelope. These often arise from elements like windows, doors, structural beams, wall ties, pipes, cables, and cantilevers that interrupt the insulation layer [121]. Notably, the most prevalent building envelope construction systems are the External Thermal Insulation Construction Systems (ETICS) in walls and steel sheet roofs. While ETICS have certain advantages, issues related to their placement and insulation thickness have led to what is known as 'pseudo insulation' [122]. Additionally, the practice of ensuring envelope airtightness through blower door tests remains disappointingly rare.

Conversely, the growth rate of energy-efficient HVAC (Heating, Ventilation, and Air Conditioning) system sales in Poland has been sluggish, although discernible. Boilers remain a staple in the nation's residential architecture, as indicated in Figure 3.5. Over the last decade, gas boilers have gained prominence over their solid fuel counterparts, and there has been a modest uptick in the sales of electric boilers as well. Figure 3.5 offers a snapshot of the heating system market, which is primarily dominated by district heating in Poland. As of 2019, the country boasted 412 district heating systems with a combined capacity of 54.912 MWth [123]. An article delved into the anticipated future of Poland's district heating network in the context of evolving European policies [124]. It suggested that, although the long-term outlook points towards a decline in energy demand from these systems due to more stringent European regulations, the short-term prospects indicate an expansion and modernization of the existing infrastructure. Such modernization efforts will likely include a shift in fuel types, and technologies like biomass combustion and cogeneration are being considered, although they present their own sets of challenges.

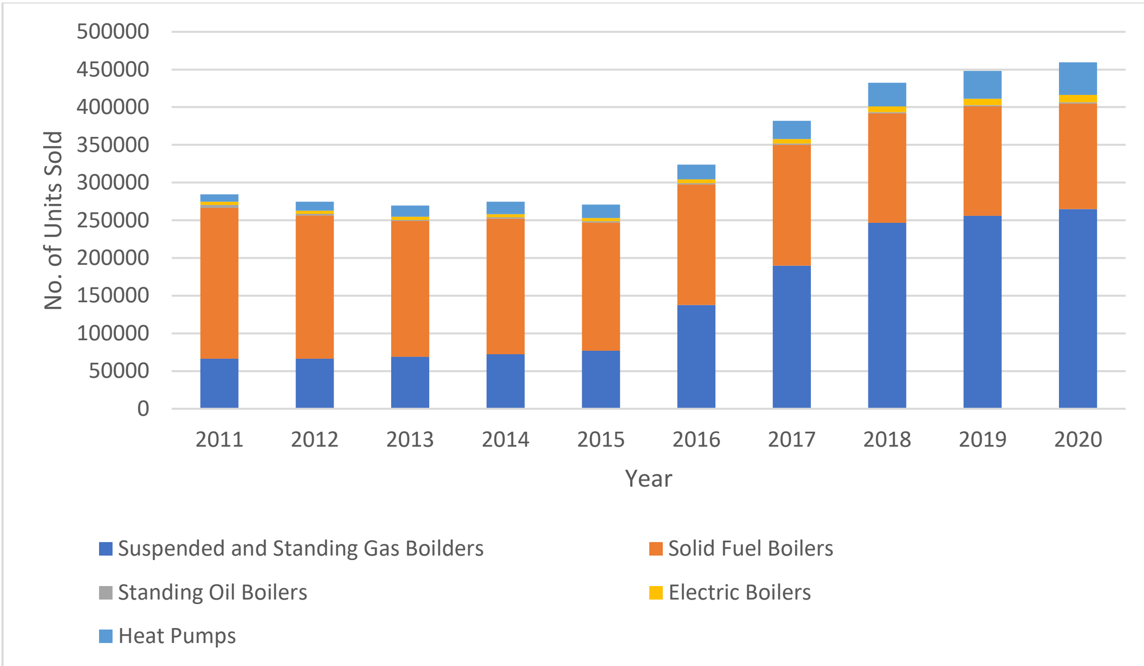


Figure 3.5. Sales of boilers and heat pumps without AC, based on [125]

When it comes to Domestic Hot Water (DHW) consumption, the field remains relatively under-researched in Poland [126]. According to a study by Bertelsen et al., the average household in Poland consumes approximately 2600 kWh/year for DHW [88]. Chmielewska conducted a comprehensive study spanning two years that examined 626 apartments to quantify DHW energy usage in multi-apartment complexes [127]. Cholewa et al. took a longer view, studying nine multifamily buildings over a 20-year period to identify ways to minimize energy consumption in older DHW systems [128]. The research discovered that a significant portion of energy wastage in existing structures, ranging between 56.7% and 70.5%, could be attributed to heat losses in the DHW systems. This highlights an urgent need for renovation strategies focused on improving the efficiency of hot water delivery systems in older buildings. Heat pumps are experiencing a surge in popularity in Poland, particularly for Domestic Hot Water (DHW) applications and less so for space heating [129]. Sales of heat pumps have nearly quintupled from 2011 to 2020, as indicated in Figure 3.4 Due to more stringent regulations regarding heat transfer coefficients and Energy Performance (EP) values in the Polish Energy Performance of Buildings (EPB) standards, there is an increasing demand for smaller, more efficient heating devices, thus fueling the growth of the heat pump market. A study conducted in 2014 by Flaga-Maryanczyk et al. empirically tested a ground-source heat exchanger in Passive House (PH)-certified single-family homes [78]. The study confirmed that such heat exchangers, when coupled with mechanical ventilation systems, effectively mitigated outdoor temperature variations, a critical factor during harsh winter months. The integration of Photovoltaic (PV) panels with heat pump systems has been explored as a potential avenue for enhanced energy efficiency. Romanska-Zapala et al. (2017) investigated this combination specifically for low-energy single-family homes in Poland during the winter season [130]. They found that the energy gains from utilizing a PV system in the winter were minimal due to low levels of solar irradiation. Their study indicated that even when equipped with batteries, the energy saved through a PV system was only 175 kWh of primary energy compared to a system

without batteries. However, the study also emphasized that a more accurate picture of energy savings could only be obtained by analyzing data across an entire year. Furthermore, a pipe ground-air heat exchanger (GAHE) has been investigated for its potential to work alongside an air handling unit to provide summer cooling. The study found that a GAHE could pre-condition the incoming air, either pre-heating or pre-cooling it, thereby reducing overall energy consumption. Optimizing such a system would require the ability to selectively source fresh air based on varying conditions and room-specific requirements.
In summary, recent trends in HVAC and renewable energy system sales demonstrate a marked increase in the adoption of heat pumps and solar collectors. Information provided by a representative of the Polish Ministry responsible for construction on October 6, 2021, at the 20th Thermomodernization Forum indicated that the residual energy requirements to be met through renewable energy in residential construction are approximately 50 kWh/m2 per year [131]. This trend may be further boosted by the introduction of dynamic electricity pricing schemes in Poland [132]. The development of grid flexibility and dynamic pricing models for prosumers is an area that requires further attention [133]. Figure 3.6 presents the percentage changes in the adoption of different heating devices and solar panels in Poland between the years 2011 and 2020.

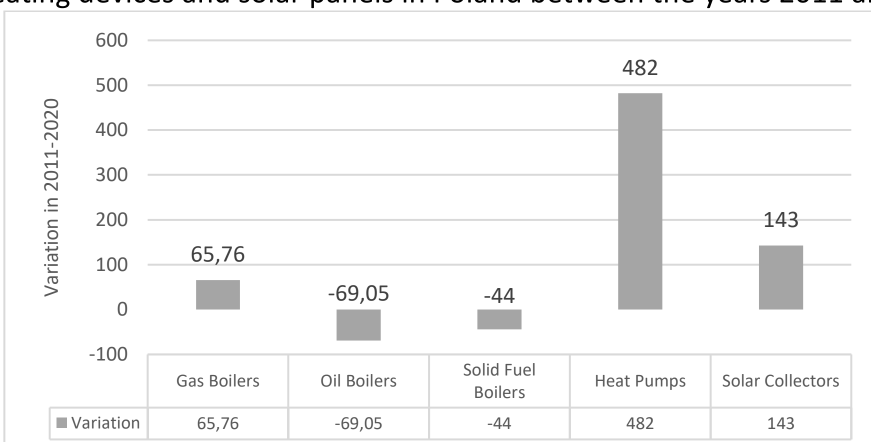


Figure 3.6. Variation of heating devices and solar panels sales in Poland

3.7. Building Clusters

To enhance the reliability and generalizability of the research findings, this study narrows its focus to a single cluster of buildings located in a specific region within Poland. By adopting such a targeted approach, the research aims to minimize the number of variables that could introduce high levels of uncertainty into the results. This strategy is predicated on the belief that a more nuanced and accurate understanding of a specific building cluster will yield more reliable data that can then be more confidently generalized to similar contexts. Understanding the existing landscape of buildings in Poland is a critical initial step in this endeavor. Before selecting the building cluster to focus on, a comprehensive survey of various clusters is necessary to gain a detailed understanding of each. This involves considering various factors, such as the number of buildings within each cluster, their architectural designs, age, usage patterns, and, most importantly, their current energy consumption levels or, put more precisely, their potential for energy savings. Identifying the clusters with the highest number of buildings and the greatest potential for energy savings is of particular importance. The goal is to prioritize clusters that not only represent a substantial proportion of the built environment but also have significant room for improvement in energy efficiency. By meticulously studying one such high-priority cluster, the research aims to derive insights that are both deep and widely applicable, thus increasing the value and impact of the study. In the context of this analysis, our initial focus will be on Poland's Building sector, then residential sector, examining the unique characteristics that distinguish each sub-group within this domain.

Table 3.4. Energy demand of Polish housing stock up to 2010, based on [93]

Year of Construction	**Buildings**		**Dwellings**		**Primary energy**	**Final (delivered) energy**
	Thousands	**%**	**Mln.**	**%**	**kWh/(m²a)**	**kWh/(m²a)**
Before 1918	413.30	7.71	1.21	9.01	>350	>300
1918-1944	828.20	15.55	1.54	11.46	300-350	260-300
1945-1970	1,367.50	25.50	3.71	27.62	250-300	220-260
1971-1978	676.50	12.61	2.16	16.08	210-250	190-220
1979-1988	763.50	14.24	2.20	16.38	160-210	140-190
1989-2002	698.40	13.02	1.52	11.31	140-180	125-160
2003-2010	616.02	11.48	1.09	8.14	100-150	90-120
Total	5,363.42	100.0	13.43	100.0		

As it is depicted in Table 3.4 in terms of the number of buildings 1945-1970 cluster is by far dominant in building stock in Poland with more than 1.3 million cases and has roughly 1/3 of Polish population. Considering the number of buildings this cluster is followed by 1918-1944 with more than 800 thousand cases, but the number of dwellers is almost 10% which shows that they are mainly unoccupied or occupied by a small portion of people in Poland. On the contrary, 1971-1978 cluster and 1979-1988 cluster they have around 1.3 million cases combined with around 32 % of dwellings in Poland. To be more precise, if we multiply the median of Primary energy for each cluster by the number of buildings that each cluster has, we can define a hyper-parameter that can show the total energy use intensity like variable that shows the impact of each cluster on the primary energy use in Poland energy sector if we consider a fix area for each house to rule out the area in this case (Figure 3.7).

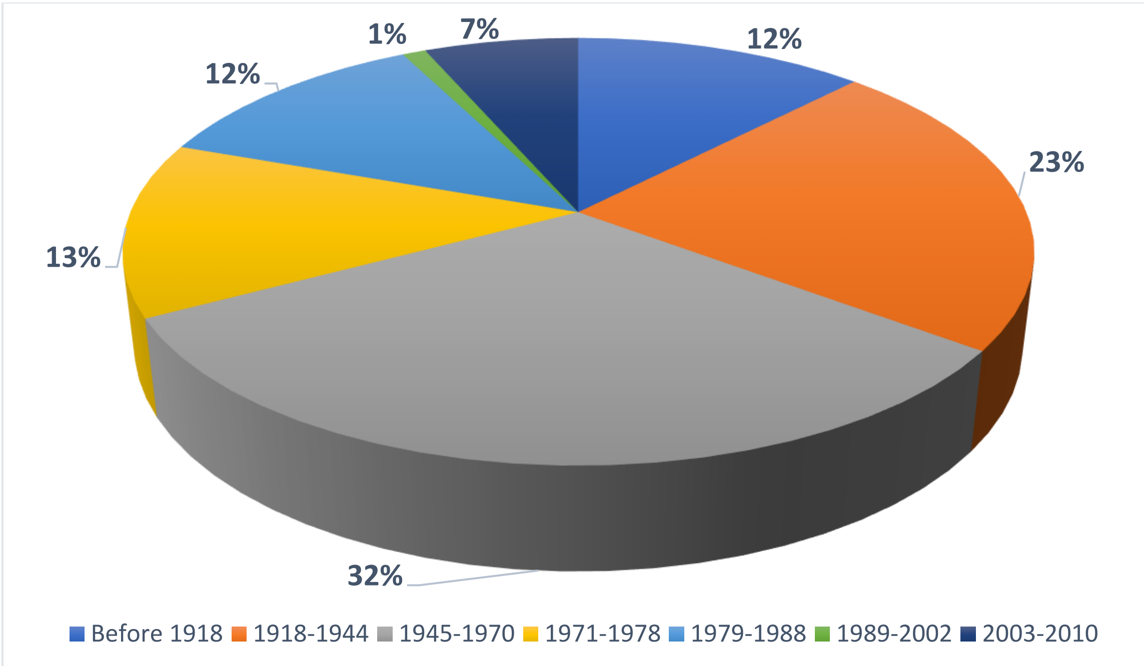


Figure 3.7. the impact of each cluster on the primary energy use in Poland energy sector

Based on figure 3.6 the dominancy of impacts of 1945-1970 building cluster on the total energy stock in Poland is quite obvious and by improving this cluster probably a significant impact on the total energy stock of Poland can be expected. Subsequently, we will zero in on a specific cluster that warrants our attention. Based on a comprehensive study from 2013 concerning energy efficiency in the country, it was found that an alarming 72% of single-family homes in Poland are categorized as having either low or very low energy standards [134]. Concurrently, coal is the primary energy source for 70% of these single-family residences (Table 3.5). This translates to around 3.5 million coal-fired boilers in active use, collectively responsible for the consumption of over 9 million tons of coal every year. Adding another layer of concern, approximately 28.8% of these boilers have been operational for more than a decade. Furthermore, a considerable number of these installations—around 3 million—are based on manually-fed boilers. This is a significantly outdated technology, notorious for contributing to high levels of air pollution.

When it comes to household energy consumption patterns in Poland, the data reveals a striking divergence from the rest of the European Union member states. Specifically, Poland holds the dubious distinction of having the highest per capita coal consumption across the EU. To put this in perspective, the amount of coal consumed per resident in Poland is a staggering ten times the average recorded across the other 27 EU countries (Table 3.6). For a more granular understanding, Table 3.5 lays out the types of energy carriers used in Polish households. As corroborated by earlier discussions, coal is overwhelmingly the dominant energy source, highlighting a critical area in need of transformation to align Poland more closely with broader European energy sustainability objectives.

Table 3.5. Energy consumption in households in Poland, based on [93]

Energy carrier	**Units of Masure**		**Household consumption**		**National share of energy carrier usage in households**
		Original units	**PJ**		**%**
Total			821.3	100.0	
Electricity	TWh	28	101.9	12.41	19.1
District heat	PJ		180.0	21.92	53.9
Natural gas	PJ		141.4	17.22	22.7
LPG*	thousand t	500	23.7	2.88	20.5
Heating oil	thousand t	87	3.8	0.46	9.6
Hard coal	thousand t	9,200	243.8	29.69	12.2
Lignite	thousand t	450	3.6	0.44	0.7
Coke	thousand t	450	5.3	0.65	6.5
Fuel wood	PJ		116.9	14.23	58.0
Solar energy	PJ		0.4	0.05	69.3
Geothermal energy**	PJ		0.5	0.06	77.1

* Consumption for household purposes only (excluding fuels consumed by cars)

**Households use geothermal energy obtained indirectly from a heating company network

Table 3.6. Energy consumption in households in Europe, based on [93]

Country	**Energy Consumption in Households**		**Share of Households in National Energy Consumption**
	TJ	**GJ/inhabitant**	**%**
Latvia	55,166	27	31
Denmark	182,957	33	23
Hungary	231,140	23	22
Lithuania	63,950	21	22
Romania	329,067	16	22
Ireland	114,360	25	20
Greece	228,082	21	20
Austria	269,813	32	20
Poland	795,745	21	19
UK	1,500,500	24	18
Italy	1,311,299	22	18
Germany	2,216,246	27	17
Slovenia	49,106	24	16
Estonia	39,203	29	15
Finland	211,224	**39**	14
Sweden	291,259	31	14
France	**1,546,935**	24	14
Czech Republic	246,700	23	14
Spain	679,154	15	13
Belgium	310,040	28	12
Netherlands	408,220	24	12
Slovakia	88,814	16	12
Cyprus	12,877	15	12

Bulgaria	99,649	14	12
Portugal	116,659	11	12
Luxembourg	17,867	34	9
Malta	3,051	**7**	6

It is very important to know the purpose of energy consumption in households to see which area has the potential for energy saving (Table 3.7). Also by focusing on changes during recent years on different areas it would be possible to understand which area used to be low but for some reasons (maybe behavior changes) the figure may experience changes (Figure 3.8).

Table 3.7. Final energy consumption in households in Poland in 2022, based on [93]

Purpose of use	GWh	%
Total	341788	100
HVAC	228900	66.97134
DHW	51688	15.12282
Cooking	23000	6.729318
Lighting	4700	1.375121
Electrical appliances	33500	9.801397

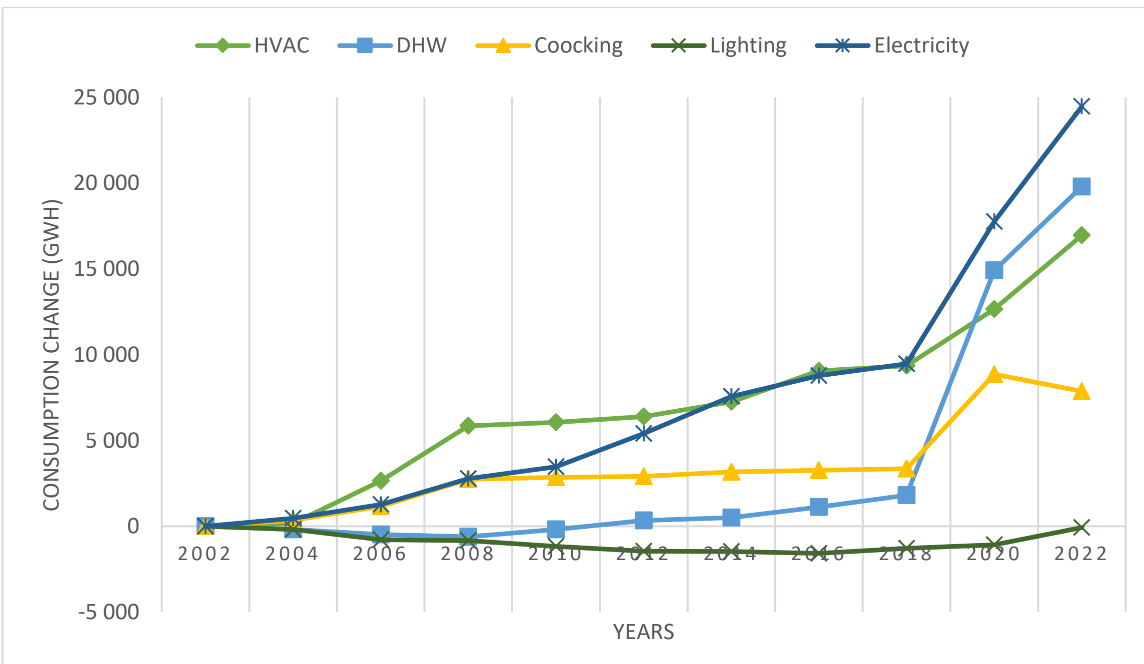


Figure 3.8. changes in household resource consumption in Poland in recent years, based on [93]

The data previously mentioned clearly indicates a pressing need for energy-efficient upgrades across Poland's residential buildings. This is largely attributable to historically lax energy performance standards, which in many instances still persist today. The situation is particularly dire in the realm of single-family residences situated in rural areas. As illustrated by Table 3.8, a significant proportion of both single-family and multi-family dwellings either lack insulation altogether or are only partially insulated, exacerbating the demand for energy.

Table 3.8. Standard of buildings based on the criterion of thermal insulation, based on [93]

Efficiency level of the building	Number of buildings		Building characteristics
	Thousands	% of total	
Very high standard	45	1.2	• Modernised/modern installation • Wall insulation minimum 15 cm • Roof insulation • Energy efficient, triple glazed windows
High standard	335	6.7	• Modernised/modern installation • Wall insulation minimum 11 cm • Roof insulation • Double glazed windows
Average standard	1000	20.1	• Modernised/modern installation • Wall insulation 8-10 cm • Roof insulation • Double glazed windows
Low standard	1700	34	• Buildings with wall insulation layer thinner than 8 cm
Very low standard	1900	38	• Uninsulated buildings

In the context of existing residential structures, the ideal metrics for demand concerning non-renewable primary energy and final energy for functions like heating, ventilation, and Domestic Hot Water (DHW) heating are as follows:

- Existing buildings, built after 1970:
  30-45 kWh/(m^2year);
- Older buildings, built before 1970:
  45-70 kWh/(m^2year);
  Taking into account the share of RES, EP could be reduced to 50-75 kWh/(m^2year).

Addressing these inefficiencies is of the utmost importance not just for environmental sustainability but also for economic stability. Inefficient energy use translates into higher utility bills for residents and a greater overall consumption of non-renewable energy sources. The widespread lack of insulation and other energy-saving measures presents both a challenge and an opportunity. It's a challenge because retrofitting existing buildings is often more complex and costly than incorporating energy-efficient designs from the outset. Yet, it's an opportunity for creating jobs, stimulating the construction industry, and adopting newer, cleaner technologies. Thus, there is an imperative to revise building codes, incentivize energy-efficient renovations, and foster public awareness about the financial and ecological benefits of energy-efficient buildings. This would not only reduce the country's carbon footprint but also alleviate some of the energy burdens on Polish households, particularly those in rural areas where energy inefficiencies are most pronounced. Finally, considering the number of cases and retrofitted portion of each luster (Table 3.9) the focus cluster (1945-1970) has more than 1 million cases that need energy retrofit.

Table 3.9. Thermo-modernization statistics of different building clusters in Poland

Year of Construction	Buildings Thousands	Percent of stock that has been thermo-modernized	No. of building that has not been thermo-modernized Thousands
Before 1918	413.30	7%	384.36
1918-1944	828.20	7%	770.22
1945-1970	**1,367.50**	**11%**	**1,217.07**
1971-1978	676.50	16%	568.26
1979-1988	763.50	15%	648.975
1989-2002	698.40	10%	628.56
After 2002	616.02	New buildings constructed under prevailing obligatory	
Total	5,363.42	after 2008 energy performance standards	

That is why it is better to focus in this group to have more information about it especially in the studied region. The building cluster in Wielkopolska region, particularly in Poznan, were commonly built between the early 1950s and late 1960s. Typically constructed in place of older farm buildings, these structures often integrate features or foundations from previous constructions. Building usually occurred in two phases: initial construction involved the basement and ground floor, temporarily capped with a makeshift roof. The second phase included adding an upper floor and finalizing the structure with a hipped roof featuring a short ridge. A small dormer often appeared on one side of the roof, which would later be subject to expansion and renovations. Quality has been a hallmark of these houses; Local brickyards supplied solid, high-quality bricks for both external and internal walls. A unique feature of these buildings was the use of an inverted three-layer wall system for the outer walls, consisting of cement-lime plaster, a masonry wall, an air void, and a pressure brick wall, finished with internal cement-lime plaster. Ground floors often featured concrete flooring that would be retrofitted later with modern insulation solutions. The ceiling structure usually relied on flat brick ceilings laid on steel I-beams, without thermal insulation.

Figure 3.9. Street and panoramic view of a case at Mickiewicza Street, Poznan

Distinctive features of this architectural style are readily apparent in various aspects of the house's design. One of the most notable is the proportions of the house's volume. These proportions create a specific aesthetic balance, often influenced by regional architectural trends and local landscapes. Another defining characteristic is the window composition. The layout, size, and type of windows are thoughtfully integrated into the overall design. They serve not just as openings for natural light but also as essential elements that complete the architectural narrative of the house. Furthermore, the roof's geometry is an important feature that adds to the uniqueness of this style. Typically, two different types of roofs are associated with this architectural form. Each type has its own set of characteristics that contribute to the visual appeal and functionality of the house. These key elements—ranging from the proportions of the structure to the design of the windows and the geometry of the roof—come together to define this specific type of house. These choices often take into account local climate conditions, available building materials, and cultural preferences, resulting in a design that is both practical and aesthetically pleasing, as well as representative of its locality (Figure 3.9,10, 11, 12).

Figure 3.10. Flat roof case (the most frequent case)

Stairs within these homes were generally constructed from reinforced concrete leading to the basement and the first floor, with wooden structures often employed for stairs between upper levels. Thermal insulation, typically in the form of mineral wool, was common in the roof structure. Roofs were commonly covered with ceramic plain tiles arranged in a lace pattern. Dormers were often a mix of steel and wooden structures, featuring zinc-steel sheet flashings.

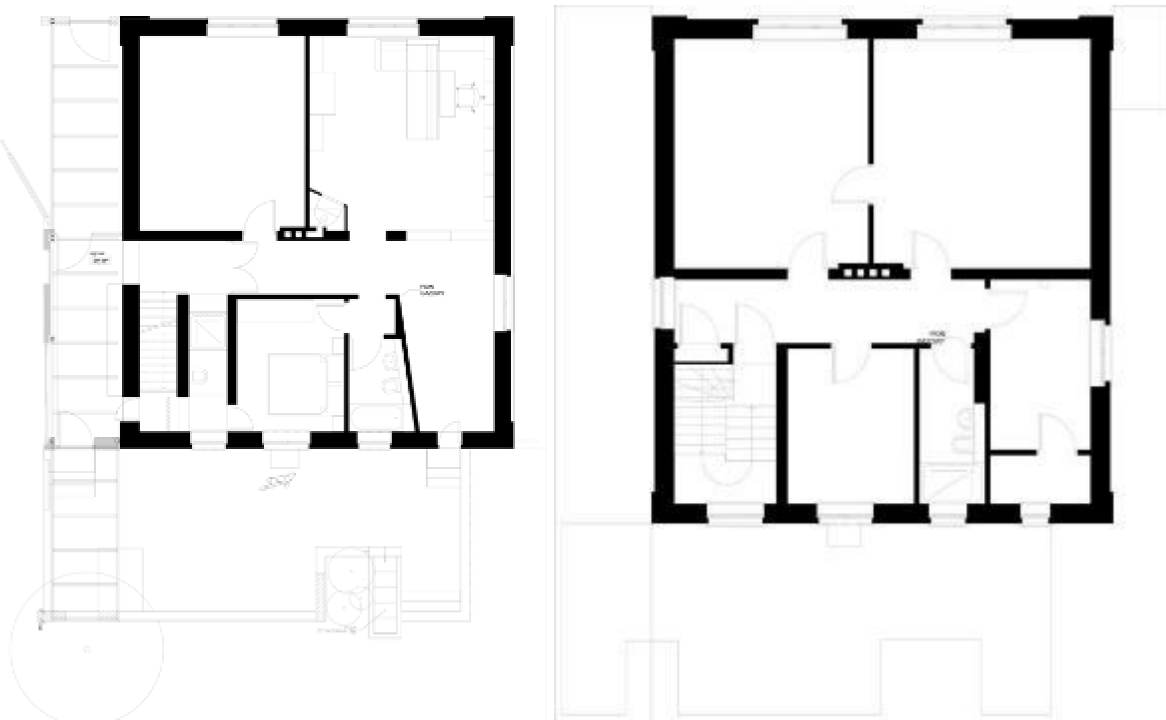


Figure 3.11. Archetype plan for the cluster

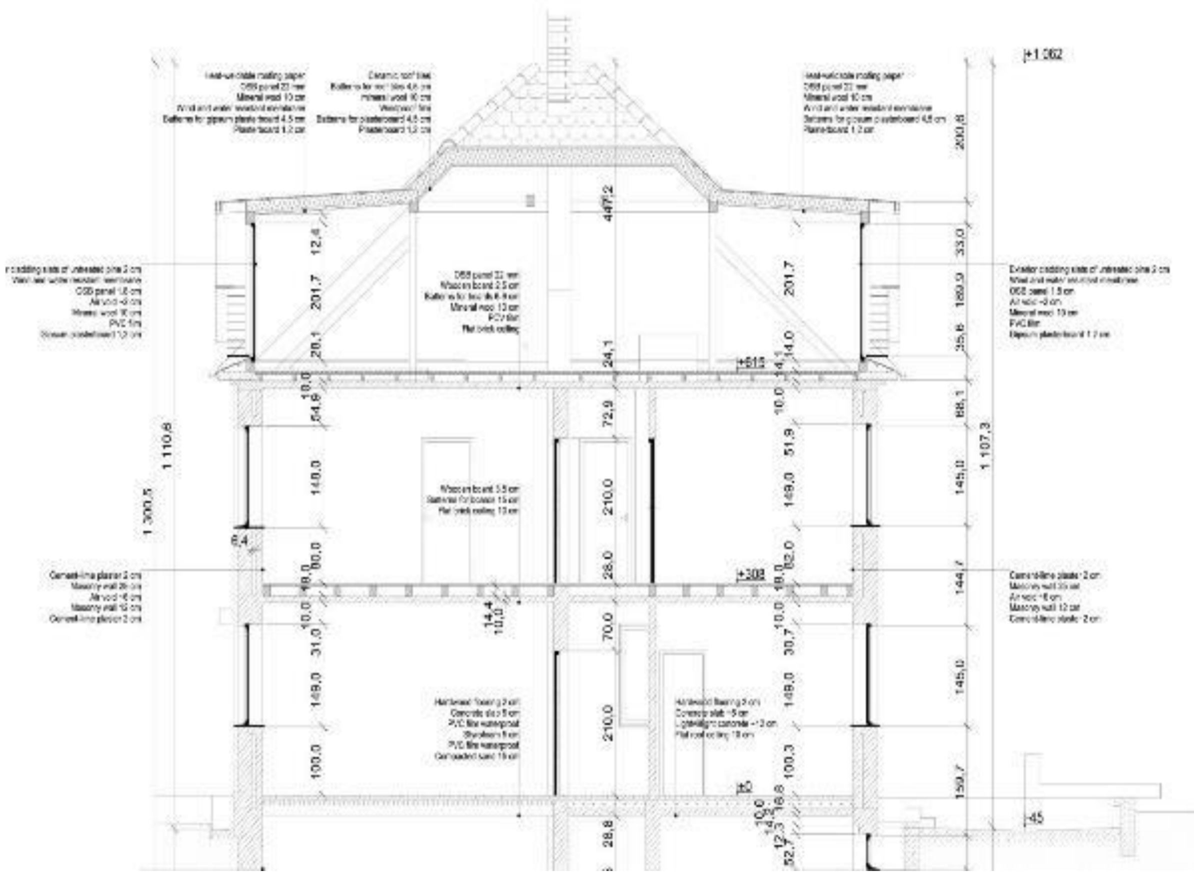


Figure 3.12. Archetype section and construction details for the cluster

Vestibules were often later additions to these homes, built with a wooden frame and insulated with thick mineral wool. The roofing material for these vestibules usually matched the primary roof, but was flat and based on wooden rafters. Joinery of the era often employed wooden doors and casement windows. Over time, it is common for these to be replaced with modern equivalents, such as PVC glazed doors and tilt-and-turn windows. Originally, heating solutions included tiled stoves that were later converted to gas or coal stoves placed in the basement. By the 21st century, many of these homes have been updated to feature more contemporary heating solutions like central fireplaces.

3.8. Conclusion

Following an extensive analysis of Poland's building inventory, particularly within the residential sector, our focus narrowed to a particular group of buildings in Wielkopolski, Poznan. Our collected data indicates that this specific cluster is the most commonly encountered example requiring immediate attention for energy retrofits. Undertaking energy audits for these buildings represents a significant stride in enhancing their energy efficiency. The overarching objective of this project is to streamline this retrofitting process through the utilization of data-driven methodologies. Key features that make this cluster stand out for immediate action include specific architectural traits and energy usage patterns that deviate from modern efficiency norms. The concentration of these buildings in a particular geographical locale adds an additional layer of urgency, as it magnifies the environmental impact of their inefficiencies. Therefore, energy audits, which would serve as comprehensive assessments of energy use and potential savings, are of paramount importance for these buildings. Not only would audits identify areas in need of improvement, but they could also pave the way for cost-effective solutions that could be universally applied to similar structures in the region. The ultimate goal of this initiative is to facilitate the transformation of these energy-inefficient buildings into models of sustainability. By applying data-driven methods, such as machine learning algorithms to predict energy use, or statistical analysis to identify the most impactful retrofit options, we aim to expedite and simplify the retrofit process. This will not only benefit the individual building owners through decreased energy costs, but it will also contribute to broader environmental sustainability goals.

Chapter 4: Climate Change Consideration

4.1. Abstract

This chapter investigates the localized impact of climate change on building energy consumption, specifically focusing on heating and cooling energy demands in Poznan, Poland. As climate change poses a significant threat to global energy sustainability, it becomes imperative to scrutinize how it might alter energy consumption patterns. Notably, the impact of climate change on building energy consumption varies depending on geographical and climatic conditions. Utilizing statistical downscaling methods, this chapter generates future weather data for the years 2050 and 2080 in Poznan, based on the HadCM3 and A2 greenhouse gas (GHG) scenario. This data serves as the foundation for simulating energy demands in 16 building prototypes, according to the ASHRAE 90.1 standard.

Our analysis reveals an average increase in cooling load by 135% and a decrease in heating load by 40% by the year 2080. While the total thermal load of buildings is currently decreasing due to a higher share of heating load, the chapter posits that if no mitigative steps are taken, the increasing cooling demand could lead to a surge in both thermal loads and associated GHG emissions. These findings provide valuable insights for urban planners, policymakers, and building engineers who are engaged in designing sustainable, climate-resilient buildings. It suggests the critical need for preemptive strategies to control increasing cooling demands, thereby contributing to broader environmental sustainability goals.

4.2. Introduction

The surge in urban population, fueled by economic and industrial advancements, has placed unprecedented demands on urban infrastructures, including energy systems and housing [135]. These infrastructural developments, while elevating quality of life, have also led to elevated levels of greenhouse gas emissions. The consequential impacts of these emissions, such as shifting weather patterns, extreme climatic conditions, and global warming, are well-established [136]. According to the Fifth Assessment Report by the IPCC (AR5), these climatic disruptions are anticipated to elevate global mean surface temperatures by approximately 2.5 – 4.5 °C by the century's end [137]. Importantly, these external climatic factors are instrumental in determining the energy consumption patterns of buildings [138]. Buildings account for a staggering 67% of global energy demand and contribute significantly to total GHG emissions [139]. Specifically, in the E.U. and the U.S., buildings accounted for up to 40% of energy demand in 2019 [140, 141].
Furthermore, a concerning trend has emerged: total energy-related CO_2 emissions from the building sector, which had plateaued between 2013 and 2016, spiked to an all-time high of 10 $GtCO_2$ in 2019 [142]. This uptick has prompted a multidisciplinary response from scholars across various fields aiming to address the energy and emissions crisis in the building sector [143]. However, climate change has far-reaching impacts that extend beyond the scope of energy and emissions. Its ripple effects touch upon various facets of urban living, such as public health, water resource management, economic stability, and political governance [144]. Consequently,

adapting to the changing climate stands as one of the most pressing challenges of our times, prompting numerous international initiatives aimed at exploring sustainable solutions [145].
Recent scholarship has categorized the impacts of climate change on the building sector into three core aspects: HVAC systems, heating and cooling demand, and peak power demand [146]. To scrutinize these categories, building performance simulations (BPS) have proven invaluable. These simulations provide a holistic view of a building's energy consumption over its lifecycle and have revealed the significance of thermal loads, particularly in heating and cooling systems [147]. The thermal behavior of a building is primarily governed by three elements: the physics of the building itself, the microclimate of the surrounding environment, and the required internal thermal comfort levels [148]. Given the pronounced influence of weather variables on energy consumption, strategic management of heating and cooling demand alone can significantly optimize a building's energy efficiency [149]. Within the context of the European Union (Figure 4.1), a decline in energy consumption, particularly in thermal load, has been observed since 2008 [93]. However, this decline is not uniform across countries, as evidenced by the average annual energy consumption per square meter, which ranged from 55 kWh/m^2 in Malta to 300 kWh/m^2 in Romania as of 2013 [150]. Even among countries with similar climates, there are noteworthy differences—for example, Sweden's average consumption is 18% lower than that of Finland [151]. Thus, it is critical to conduct country-specific analyses that consider both current and projected climatic conditions to comprehensively understand and address the impacts of climate change on building energy consumption [152].

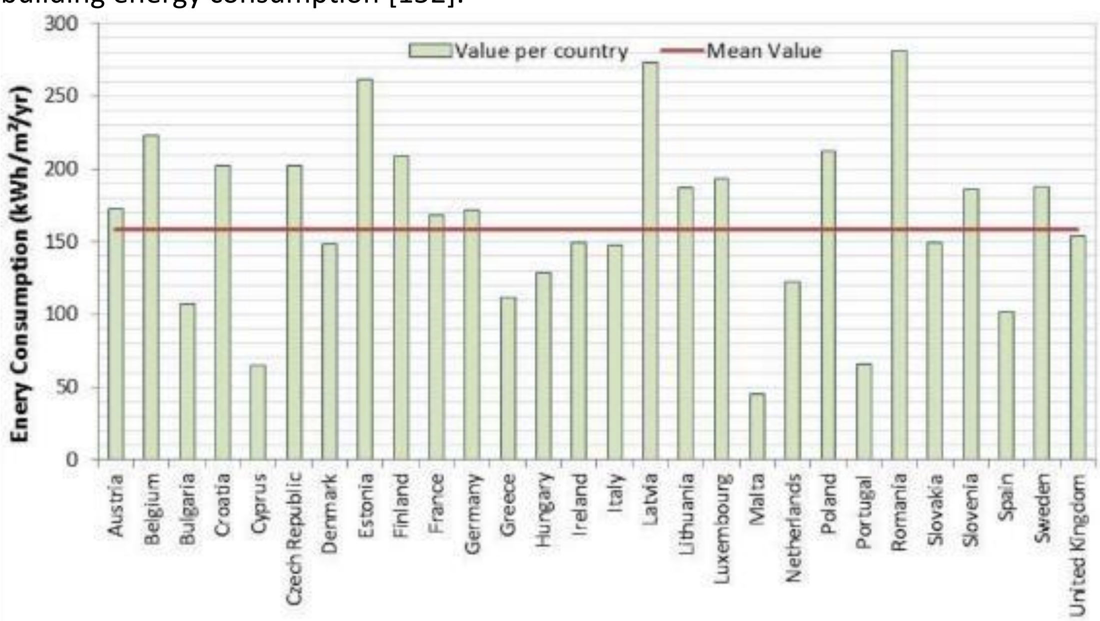


Figure 4.1. Annual energy consumption of buildings per m^2 on average, based on [147]

Poland stands as a notable case study in the European Union, ranking among the top five countries in terms of energy consumption per square meter in buildings, exceeding 200 kWh/m^2/yr [148]. Despite its high consumption, the country has shown a remarkable improvement in energy efficiency over recent years. Specifically, Poland has managed to cut its energy intensity by at least 50%, owing in large part to initiatives like the Thermo-modernization and Rehabilitation Fund program. Nevertheless, coal remains the dominant energy source in

Poland, constituting a significant obstacle in reducing greenhouse gas (GHG) emissions and furthering the energy efficiency of the building sector. In response, the draft energy policy of Poland has strategically prioritized the reduction of GHG emissions while enhancing building sector efficiency (Figure 4.2). The "Polish National Strategy for Adaptation to Climate Change (SPA 2020)" was developed to set a clear agenda for sectors most susceptible to climate-induced changes, including the building industry. In 2019, the building sector emerged as the highest energy-consuming sector in Poland, accounting for nearly one-third of the nation's total energy consumption [149]. To break it down further, residential buildings were responsible for 29% of Poland's Total Final Consumption (TFC) in 2014, while commercial buildings contributed roughly 17%. Alarmingly, these sectors also accounted for approximately 18% of CO2 emissions since 1990, indicating a severe need for sector-specific mitigation strategies. What's more, according to a 2016 survey by The Buildings Performance Institute Europe (BPIE), nearly three-quarters of Poland's buildings were evaluated as having either low or deficient energy efficiency standards [93]. This prevalence of subpar efficiency levels places Poland in a particularly precarious position vis-à-vis climate change.

Given the high energy consumption rates and existing inefficiencies, the impact of climate change on Poland's building sector becomes an issue of paramount importance. Climate-induced changes can exacerbate already strained energy demands and further contribute to GHG emissions. For these reasons, an impact assessment focusing on climate change's role in shaping Poland's building sector is not only vital but also timely.

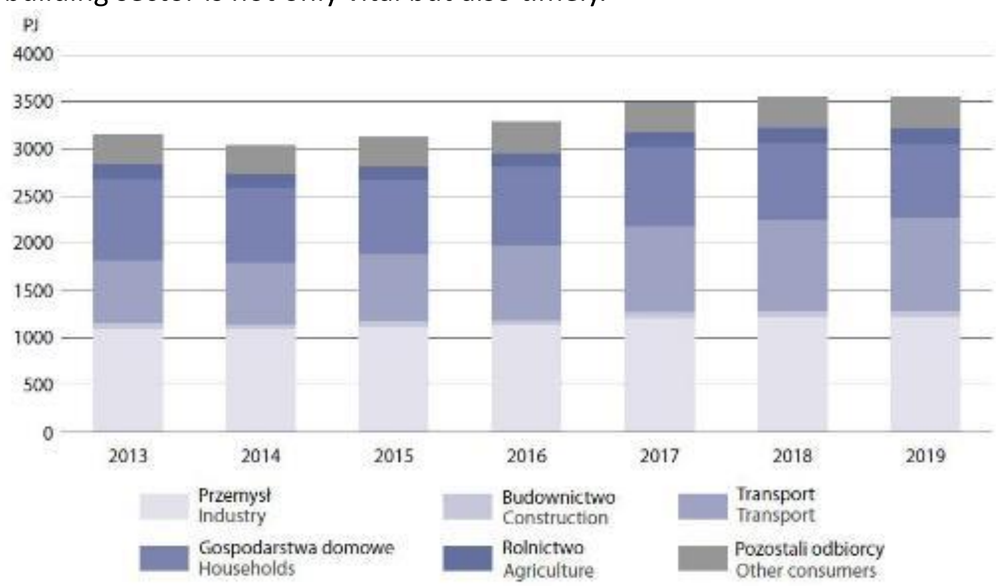


Figure 4.2. Poland's energy consumption by sectors, based on [150]

Poland has been progressively leaning towards enhancing energy efficiency across various sectors, and the building industry stands as a pivotal arena for such improvements. Despite this positive trajectory, the sector is hampered by limitations such as insufficient and unreliable data related to energy use, which could potentially allow loopholes and circumventions within the system [93]. Recognizing the urgency of this challenge, a robust national coalition spearheaded by the 'Build Desk' initiative has been established. The goal of this coalition is to bridge the data gaps by undertaking comprehensive regional studies that scrutinize energy consumption patterns in the building sector. Among the cities spotlighted for such studies is Poznan, a metropolis that holds a significant place in Poland's energy landscape. Not only is Poznan one of Poland's oldest cities, but it also ranks as the fifth-largest city in the country with a population exceeding 534,813 as of the 2019 census [151]. Its prominence is further underscored by its rapid urban growth, which naturally leads to escalating demands on its energy infrastructure (Figure 4.3).

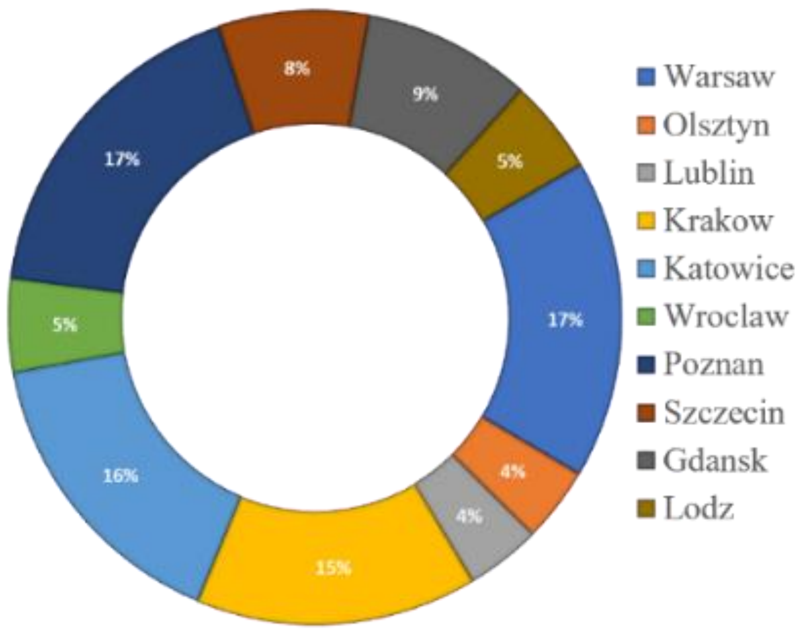


Figure 4.3. Share of Polish region in the building sector, based on [150]

Poznan serves as a microcosm that reflects broader national trends and challenges. The city's accelerating urban development and its importance in Poland's energy matrix make it a critical candidate for detailed energy consumption analyses. These comprehensive studies are anticipated to yield valuable insights that can inform policy adjustments and inspire energy-efficient solutions, not just for Poznan but for Poland as a whole. Moreover, the data generated could serve as a template for similar endeavors in other cities or even countries grappling with the complexities of energy efficiency in the building sector. In alignment with the E.U.'s ambitious Clean Energy Package objectives, Poznan has been actively participating in an EU-Horizon 2020 initiative titled "Energy Island Communities for Energy Transition." This groundbreaking project, with a generous funding pool of €6,694,000, aims to seamlessly integrate and demonstrate scalable solutions that can substantially elevate energy efficiency across sectors. One of the focal points of this initiative is the "Warta Campus in Poznan," earmarked as an 'energy island' for exhaustive study [153].

However, despite these ongoing efforts, there remains an acute need for a detailed examination of the energy performance within various sectors, most notably the building industry. This analysis becomes particularly relevant when taking into account both current climatic conditions and future variables, which notably include the pressing issue of climate change (see Table 4.1). According to the Köppen climate classification, Poznan falls under the 'Cfb' climate zone, characterized by temperate oceanic climate conditions. In this zone, temperatures remain moderate year-round, with the coldest month averaging above 0°C or −3°C, and at least four months registering average temperatures above 10°C [153].

Table 4.1. A short description of the research case study.

Country	City	Latitude	Longitude	Time zone	KGC
Poland	Poznan	52.42N	16.83E	1.00	Cfb
Population	Elevation	Weather Data	Heating DB 99.6%	Cooling DB 0.4%	Cooling 0.4% MCWB
533,830	302 (m)	2004–2018 (TMY)	− 14.27 °C	29.77 °C	19.22 °C

Given the significant role that buildings play in Poland's overall energy consumption—and in Poznan, in particular—this study intends to serve as a timely response to this analytical gap. It aspires to critically evaluate the impact of climate change on the thermal load within Poznan's building sector, a subject of heightened relevance in the context of the E.U.'s overarching energy-efficient research program. By doing so, this chapter aims to offer valuable insights that can guide policymakers, researchers, and practitioners in crafting resilient, adaptive strategies that can mitigate the negative implications of climate change on building energy consumption patterns.

4.3. Climate and Building Performance

It has been proven that the environment type and climatic conditions affect several aspects of energy consumption of buildings. This impact can be categorized into the following three groups:

- HVAC system (Heating, ventilation, and air conditioning)
- Heating and cooling demand
- Power peak demand [11]

4.3.1. HVAC system

Today, the prevailing architectural philosophy in the European Union leans towards leveraging natural ventilation systems over electrical cooling devices like air conditioners in building designs. This approach aims to minimize environmental impact. However, as summers continue to heat up, posing an increased likelihood of extreme temperature spikes, a gradual but noticeable surge in HVAC-related energy consumption is becoming evident [154]. One key indicator of this shift comes from predictions around the future energy demands in Athens. Projections indicate that by the year 2080, the energy required for air conditioning in the months of July and August could escalate by a staggering 30% [155]. Such data underpins the limitations of natural ventilation in an era of climate unpredictability. The United Kingdom serves as another poignant example. As of now, nearly 40% of large buildings are equipped with air conditioning systems, a significant increase compared to just 10% in 1994 [156]. Notably, the majority of buildings constructed before 1990 were designed with natural ventilation in mind, a clear signal of how thermal management strategies are evolving [157].

As average and extreme temperatures rise, buildings are becoming both uncomfortable and costly to maintain in terms of thermal regulation. This situation creates an imperative need for retrofitting existing mechanical systems, such as ventilation. Enhanced HVAC technologies are becoming indispensable for mitigating increased humidity and the heat gain of outdoor air during hot summers. Complicating matters is the finding that the efficacy of HVAC systems exhibits a strong negative correlation with rising wet and dry bulb temperatures (WBT and DBT). This implies that as WBT/DBT values climb, the operational efficiency of these systems degrades [158]. As a consequence, buildings—particularly in warmer regions experiencing rising WBT/DBT—will likely witness a significant uptick in energy consumption dedicated to HVAC systems.

4.3.2. Heating and cooling demand

One of the most immediate and palpable impacts of climate change on building energy consumption manifests in the rising rates of both cooling and heating energy. This surge is principally driven by an increasingly felt need for thermal comfort within buildings, particularly during sweltering summers and freezing winters [159, 160]. Research indicates that global temperature increases over recent years have led to more uncomfortable summers but milder winters [161]. The changes in energy requirements for heating and cooling are generally attributed to fluctuations in 'degree-days,' a meteorological indicator used to evaluate energy demand for thermal comfort [162]. This variable is profoundly influenced by the geographical location under study, creating diverse impact scenarios across regions. Studies have discussed the possible consequences of climate change on cooling and heating loads—critical determinants of a building's energy consumption profile—often leveraging the degree-day approach for their analyses [163-165].

A consensus emerging from computational studies suggests a notable decline in heating energy requirements paired with a significant uptick in cooling energy demands [166, 167]. More precisely, these investigations have found that the rate of decline in heating energy requirements is likely to outstrip the falling rates of heating degree-days. On the flip side, cooling energy requirements are projected to see a substantial increase. To quantify, estimates suggest a staggering 2100% increase in cooling energy demands from 1975 to 2085, according to cooling degree-day metrics. This body of research essentially corroborates earlier studies, collectively emphasizing a future with reduced heating and increased cooling needs in buildings. Geographic and climatic nuances mean that hot countries like Mexico are expected to experience higher energy demands predominantly in the summer and spring seasons. Conversely, colder nations like Canada and Norway are likely to see diminished energy demands during winter. For countries with moderate climates, such as Italy, the increased cooling energy demands in summers are expected to be offset by lower heating energy needs in winters and springs [163].

4.3.3. Power peak demand

Research addressing the implications of climate change on power peak demand within buildings offers a mixed yet eye-opening landscape. The prevailing trend, as underscored by numerous studies, is a notable decrease in the heating energy requirements juxtaposed with a discernible uptick in the cooling energy demands [168]. This is not an inconsequential development; it speaks to broader shifts in energy consumption patterns that may well reshape our understanding of sustainable building design and management. There exists a divergence of academic opinion on whether the burgeoning demands for cooling will effectively neutralize the reductions in heating requirements. Some researchers posit that the two trends will reach an equilibrium, thus leaving overall power consumption relatively unchanged [169]. Others, however, are less sanguine, speculating that cooling demands could completely overshadow heating demands, potentially leading to an energy crisis in the power sector.

It's crucial to appreciate the complexity of these trends. Climate change's influence on building energy consumption isn't just a simple equation of rising temperatures and cooling demands. Various external variables such as different climate change scenarios, building designs, and construction materials also contribute to the equation. For instance, research conducted by Huang's team suggests that colder areas may actually experience a drastic reduction in energy demands due to milder winters [170]. Moreover, although conventional wisdom might suggest that lighting and plug loads are the primary energy guzzlers in buildings, the evidence points otherwise. Cooling loads are emerging as a significant—if not the most significant—contributor to power consumption in buildings. This trend is not just a passing phase; projections indicate that the demand for cooling could surge by over half in the next century [144].

4.4. Building Energy Consumption Projection

The prediction of energy demand, particularly focusing on thermal load, has become a cornerstone of research endeavors aiming for sustainability through the optimization of energy consumption [171-178]. This critical knowledge aids in the integrated management of building systems and services [179]. One of the most intricate areas within this research landscape is the evaluation and projection of climate change’s impacts on building energy consumption, specifically thermal load. A range of studies have addressed this topic, revealing often conflicting findings and insights. For instance, Rosenthal et al. were among the pioneering scholars who posited that global warming might not universally lead to increased energy usage, particularly in colder regions [180]. They estimated that a 1.0°C global temperature rise by the end of 2010 could result in the U.S. saving over $5.5 billion (1991 USD) in reduced energy costs. This hypothesis diverges from earlier studies that projected a net increase in U.S. energy consumption due to climate change [181].

In the European context, Nik in 2006 forecasted a significant decline in heating load but a substantial rise in cooling load in Stockholm by the end of the 21st century, based on various uncertainty factors [182]. Christenson and colleagues arrived at similar conclusions for Switzerland up until 2085 [167]. Likewise, Hosseini et al. predicted for Montreal a downward trend in heating load and an upward trajectory in cooling load between 2020 and 2080, focusing their study on a one-story commercial building [183]. Taking a broader view, Cellura analyzed various scenarios across 15 cities in southern Europe and concluded that thermal load in buildings would generally increase, driven by a relative decline in heating demand and a surge in cooling demand [184]. Employing an array of general circulation models (GCMs) through the morphing method, Cellura projected an increase in thermal load ranging between 50–119% for the study areas. Crawley suggests that in regions where heating is the dominant thermal load, the overall load is likely to decrease in a warming climate [185]. This observation is further bolstered by Triana, Lamberts, and Sassi, who, focusing on the performance of social housing in Sao Paulo and Salvador, argue that an uptick in cooling consumption will consequently drive up the thermal load [186]. Contrarily, a host of other scholars maintain that a variety of factors could lead to a net increase in thermal load in numerous scenarios [187, 188]. Zooming in on specific geographical case studies, Shen examined the impact of climate change on annual energy use for both residential and office buildings across four U.S. cities for the years 2040–2069. According to the IPCC SRES A2 scenario for greenhouse gas emissions, Shen found a somewhat paradoxical outcome: an overall increase in energy consumption for residential buildings, but a decrease for office buildings [189].

In more recent research, Moazami, Nik, and their team have made critical advancements by generating future weather data for Geneva. Their findings underscore the importance of accounting for extreme weather conditions to enhance the reliability of future climate projections. They found a notable increase—around 20%—in cooling load compared to under standard conditions [135]. Similarly, Berardi and Jafarpur [137] utilized statistical downscaling methods to project future energy consumption in Toronto by 2070, based on the Köppen-Geiger climate classification for the Dfb zone [190]. Their projections suggest that while heating needs might slightly decline, a sharp increase in cooling demands could be anticipated. Adding to this body of work, Velashjerdi Farahani et al. aimed to quantify the effects of employing different passive measures to mitigate overheating risks in both old and new apartments in southern Finland. Using generated future weather data under two separate scenarios targeting the year 2050, they offered valuable insights into adaptive measures for thermal load management [191].

4.5. Weather Data

Typical Meteorological Year (TMY) files serve as one of the most widely utilized types of weather data for gauging a building's energy efficiency and corresponding emissions. These files are comprehensive data sets, featuring 8,760 hourly values that detail an array of climatic parameters, providing a representative snapshot of weather for a given location over a year [191]. These parameters are crucial in the simulation models that predict energy usage in buildings and thus have broad applications in research and real-world scenarios. The methodology for generating TMY files can vary, but the Finkelstein-Schafer (FS) statistics method is often the choice for many researchers and practitioners. Developed by Hall et al., the FS method has gained wide acclaim for its robustness and applicability, becoming a standard tool in the field [192]. This method requires the amalgamation of weather data over a considerable time frame to ensure that the generated TMY files offer a 'typical' weather profile for the location in question. As an

example, Petrakis et al. formulated TMY files for Nicosia, Cyprus, by synthesizing seven years of weather data [193].
It's worth mentioning that TMY is not the only type of meteorological file available. There are other types, such as Typical Reference Year (TRY) and Design Summer Year (DSY). These variants are also generated, analyzed, and adapted through a multitude of approaches to create specialized weather datasets for specific applications [194, 195]. However, the critical backbone for generating any of these meteorological files is the availability of a comprehensive historical dataset. This archive acts as the source material from which each month is meticulously chosen to represent that month's typical weather conditions over multiple years. The richness and completeness of this historical data are vital to the reliability and credibility of the resulting TMY files, or any other types of meteorological files for that matter.

4.6. Climate Models and Projection

When it comes to predicting the future climate conditions that will impact building performance, General Circulation Models (GCMs) have been a mainstay in the realm of research. However, they have a significant limitation for this specific application; they typically provide daily or monthly data, which falls short of the hourly data requirements for building performance simulations (BPSs) [196]. Recognizing the intricacies and uncertainties in climate modeling, some researchers have opted for a multifaceted approach. In a notable example, Berardi and Jafarpur employed a combination of models in their recent study. They used Hadley Regional Model 3 (HRM3) alongside Hadley Climate Model 3 (HadCM3) for dynamical downscaling of future climate data. This was an expansion of their earlier work where they had used HadCM3 for statistical downscaling, signifying a comprehensive methodology for climate projections [197]. Country-specific models and collaborations further enrich climate projection studies. For instance, the Polish-Norwegian CHASE-PL project offers a specialized glimpse into Poland's future climatic conditions. The project's approach involved downscaling GCM projections using an ensemble of nine Regional Climate Models (RCMs) derived from the EURO-CORDEX initiative. The findings were alarming but vital: under moderate scenarios, mean annual temperatures in Poland could rise by 1-2°C within this century. Under more extreme scenarios, this figure could soar to 4°C [198-202].
Researchers have a variety of tools at their disposal to integrate the impacts of climate change into weather data. These can broadly be classified into two categories [203]. The first leans heavily on historical weather data, encompassing techniques like the imposed offset method, extrapolating statistical method, and the stochastic weather model. The second category pivots towards the use of numeric climate models, such as downscaling GCMs to produce localized future weather data using RCMs. Downscaling itself can be executed via two avenues: statistical and dynamic. Dynamic downscaling is computationally demanding, employing regional-scale forces in conjunction with lateral boundary conditions to formulate more localized RCMs.
Statistical downscaling has gained prominence in the field of climate modeling for building design due to its comparatively low computational intensity. This approach involves two key steps: first, establishing a statistical relationship between large-scale and local climate variables; second, utilizing this pre-established relationship for simulating local climate conditions [204]. Stephen

Belcher, in 2005, further advanced the domain by introducing a specialized form of statistical downscaling known as "morphing" [205]. The "morphing" method was proposed as a more practical alternative to dynamic downscaling for building design projects, which often cannot afford the computational expenses associated with the latter. Dynamic downscaling provides fine-scale detail but at a high computational cost, making it less feasible for practical applications. Conversely, another method called stochastic weather generation, while computationally inexpensive, has its shortcomings too. It relies heavily on large datasets for training, and even then, the generated weather series may lack meteorological consistency [206, 207].

Statistical downscaling, and more recently, morphing, have risen in prominence for their efficiency and speed of calculation. Unlike dynamic downscaling, which often necessitates specialized requirements for data, such as precise representation of topography and mesoscale processes, morphing has fewer data constraints [76, 208-212]. This makes it particularly appealing for those who are operating within the limitations of available computational resources or tight project timelines. Scholars like Nik have continued to employ other methodologies, such as Regional Climate Models (RCMs), for dynamic downscaling, especially when the need for a more detailed topographical and mesoscale representation is essential [182]. However, the general trend suggests a shift towards statistical downscaling techniques like morphing, as it offers a balance between accuracy and computational efficiency.

4.7. Impact Assessment

The present section initiates with an exposition of the results obtained from weather projections, followed by a comparison of these projections with existing conditions in Poznan. This provides a foundational understanding of the likely future climate trends affecting the region. Subsequently, these weather datasets are used to construct a prototype for assessing the impact of climate change on thermal load in buildings. The comparative analysis between current and projected weather data, spanning the next 30 and 60 years, reveals specific trends in great detail. This study incorporates 13 pivotal variables that include a spectrum of meteorological and radiative factors: dry bulb temperature, dew point temperature, relative humidity, direct normal radiation, global horizontal radiation, diffuse horizontal radiation, horizontal infrared radiation, direct normal illuminance, global horizontal illuminance, diffuse horizontal illuminance, ground temperature, total sky cover, and atmospheric pressure. The forecasted future weather conditions point toward a significantly warmer climate in Poznan, coupled with increased levels of direct illuminance and radiation, as well as decreased humidity and cloud cover (as summarized in Table 4.2). To put the numerical data into perspective, temperature-related variables are anticipated to rise by an average of 4°C over the specified time frame. In parallel, variables related to radiation and illuminance are projected to increase by around 14.3 Wh/m^2 and 463 lux, respectively.

However, the scenario is not uniform across all variables. For instance, while there's a considerable reduction in total sky cover, with a decline by nearly 15%, the change in atmospheric pressure is negligible. To further elucidate, the most substantial increment was observed in direct normal illuminance, surging by about one-third. On the flip side, atmospheric pressure registered the most negligible change. In contrast, total sky cover exhibited the most significant reduction, dropping by 13%, while diffuse horizontal illuminance saw the least decrease, less than 1%.

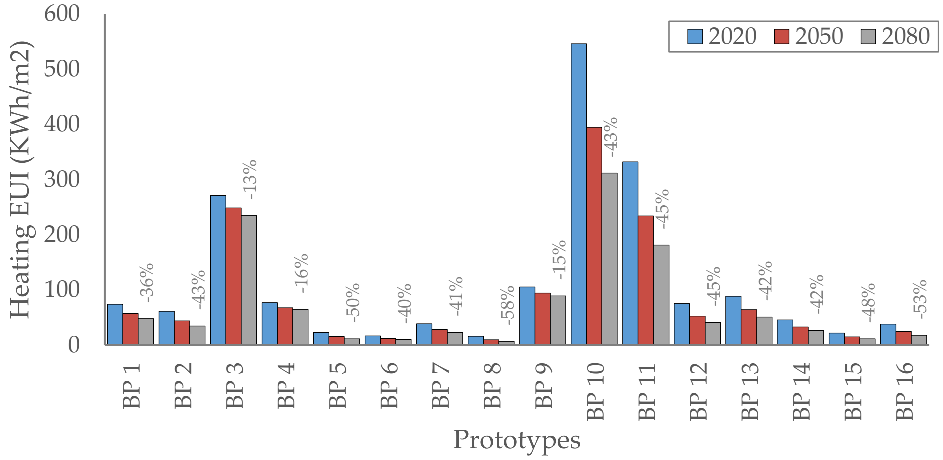
Table 4.2. Relative changes of weather parameters for 2050 and 2080 compared to 2020.

Weather Parameters	Absolute Value	Relative Changes compare to 2020	
	2020	2050	2080
Dew point temperature (°C)	4.2	6.0	7.2
Average ground temperature (°C)	8.0	10.8	12.6
Dry bulb temperature (°C)	8.2	10.9	12.6
Direct normal radiation (Wh/m^2)	59.1	81.3	86.4
Diffuse horizontal radiation (Wh/m^2)	75.9	72.5	71.4
Global horizontal radiation (Wh/m^2)	109.7	113.0	115.7
Horizontal infrared radiation (Wh/m^2)	303.0	322.1	331.5
Direct normal illuminance (lux)	7015	7385	7875
Diffuse horizontal illuminance (lux)	8578.5	8609.7	8527.7
Global horizontal illuminance (lux)	12,157	12,434	12,739
Atmospheric pressure (Pa)	101,325	101,272	102,249
Relative humidity (%)	78.1	74.6	72.7
Total sky cover (0–10)	4.5	4.2	4.3

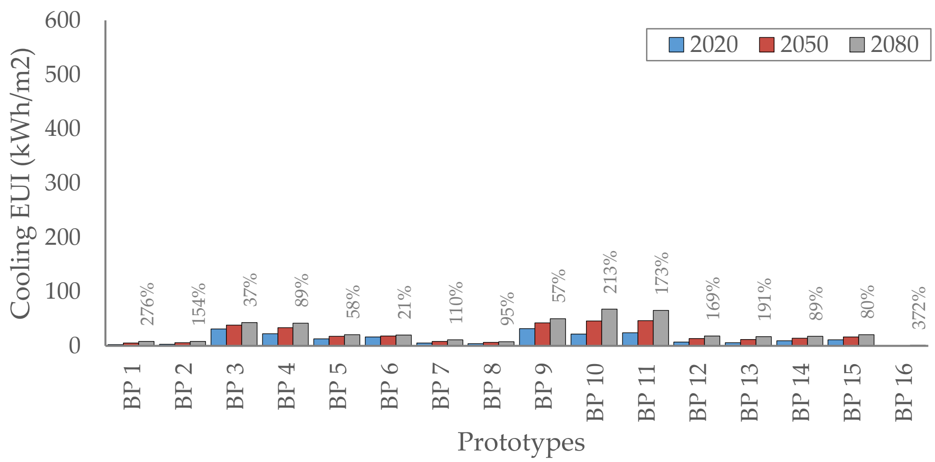
The subsequent phase in evaluating the influence of climate change on buildings focuses on scrutinizing the energy consumption, specifically the heating and cooling energy use intensity (EUI), as well as the thermal load for an array of 16 building prototypes. Post-simulation, the data for heating and cooling EUI were visually represented through scatterplot graphs. Trendlines were then plotted for each prototype, providing an interpretive lens into the changes in heating and cooling demands over the study period.

For heating EUI, the data presented a pronounced decline across the different building prototypes, boasting an average reduction of more than 41 kWh/m^2 (as depicted in Figure 4.4 a). To quantify, the baseline average heating load in 2020 stood at approximately 114 kWh/m^2. Projections for the years 2050 and 2080 indicated reductions to 87 kWh/m^2 and 72 kWh/m^2, respectively, translating into decreases of about 25% and 40%. Delving deeper into specific prototypes, the heating load for the fast-food restaurant building (BP10) exhibited a substantial reduction, plummeting by 324 kWh/m^2. Conversely, the large office building (BP06) registered only a modest decline, amounting to around 6 kWh/m2. On a broader scale, while the average change in heating load for the building prototypes hovered around a 40% reduction, the steepest decline was observed in the small office building (BP08) at approximately 57%, whereas the hospital building (BP03) had the least reduction at about 13%.

Turning our attention to cooling EUI, an unequivocal upward trajectory was observed, evidenced by an average increase of 13 kWh/m^2 (refer to Figure 4.4 b). To put it in numbers, the average cooling load in the baseline year of 2020 was roughly 13 kWh/m^2. This escalated to 1.5 and 2 times the original value for the years 2050 and 2080, respectively. Amongst all building performance simulations (BPS), the fast-food restaurant building (BP10) recorded the most dramatic surge in cooling load, rising by 46 kWh/m^2. In contrast, the cooling load in the warehouse building (BP16) increased marginally by just 0.7 kWh/m^2. On an average scale, the prototypes exhibited a 135% escalation in cooling load. The warehouse building (BP16) outpaced others with a massive 371% uptick, while the large office building (BP06) registered the lowest growth rate at 20%.



(**a**)



(**b**)

Figure 4.4. Heating EUI (a) and cooling EUI (b) in building prototypes from 2020 to 2080; the numbers demonstrate relative changes of values for 2080 compared to 2020.

The comparison of heating and cooling loads between the years 2020 and 2080 presents noteworthy differences. While a decline in heating load was observed across all cases, an increase in cooling load was inevitable. One key observation is the rate of change: although the absolute average change for cooling load was lesser at 13 kWh/m^2 compared to over 41 kWh/m^2 for the heating load, the percentages of these changes reveal another aspect of the story. The cooling load increased by approximately 135%, contrasting with a 40% decrease in heating load. This data highlights a more accelerated rate of increase in cooling needs compared to the relatively modest decrease in heating requirements for the region of Poznan over the study period. Delving into specific building prototypes, the most significant relative change in heating load was observed in the small office building (BP08) with a decrease of around 60% or approximately 10 kWh/m^2. Conversely, the highest relative change in cooling load was found in the warehouse building (BP16), with an impressive 372% increase, albeit the absolute figure of this change was the lowest at just 0.7 kWh/m^2.

On the flip side, the hospital building (BP03) had the lowest relative decrease in heating load at 13%, and the large office building (BP06) displayed the lowest relative increase in cooling load at 58%. These findings imply that, due to the ongoing trend of global warming, cooling demands are anticipated to rise at a faster pace than the declining needs for heating. Importantly, while heating load has historically been the dominant factor in overall thermal load, the increasing rate of cooling load suggests a future shift in this paradigm. In essence, the share of cooling load in overall thermal load, though currently marginal, is poised to become significantly more relevant as the years progress.

4.8. Discussion

The need for reliable data is paramount for accurate scenario projections, especially in the context of the built environment where precise energy metrics are crucial. However, the limitations in data availability can impede the robustness of such studies. For instance, in the case of Poznan, only the HadCM3 climate model and IPCC A2 emission scenario were available for climate projections. This underscores the significance of having multiple datasets and climate models to improve the accuracy of findings. When comparing the relative rates of change in heating and cooling demand, it is crucial to consider their initial magnitudes. Although the relative rate of cooling demand change was about three times that of the heating load, its share in the total thermal load of 2020 was negligible. Thus, this significant rate of change does not translate into a drastic alteration in the overall thermal load pattern. The study found that most building prototypes (BPs) saw a slight decline in thermal loads, with an average decrease of approximately 28 kWh/m^2.

Breaking down the rates, the average rate of thermal load change was a decrease of around 20%. The minimum change rate was a 1% increase in the outpatient healthcare building (BP09), whereas the most significant decrease was observed in the warehouse building (BP16) at more than 50%. This highlights that despite the rise in cooling load, its impact on the overall thermal load remains marginal in most cases. For example, in buildings like the large hotel (BP04) and the outpatient healthcare building (BP09), the decrease in heating load was less sharp compared to other cases where thermal load increased (Figure 4.5).

Looking forward to the year 2080, the study predicts an average thermal load of around 98 kWh/m^2, marking a decrease of more than 22% from the 2020 average of approximately 127 kWh/m^2. The smallest decrease was recorded for the secondary school building (BP15) at 1.3 kWh/m^2, while the largest decrease was in the fast-food restaurant building (BP10) at 188 kWh/m^2.

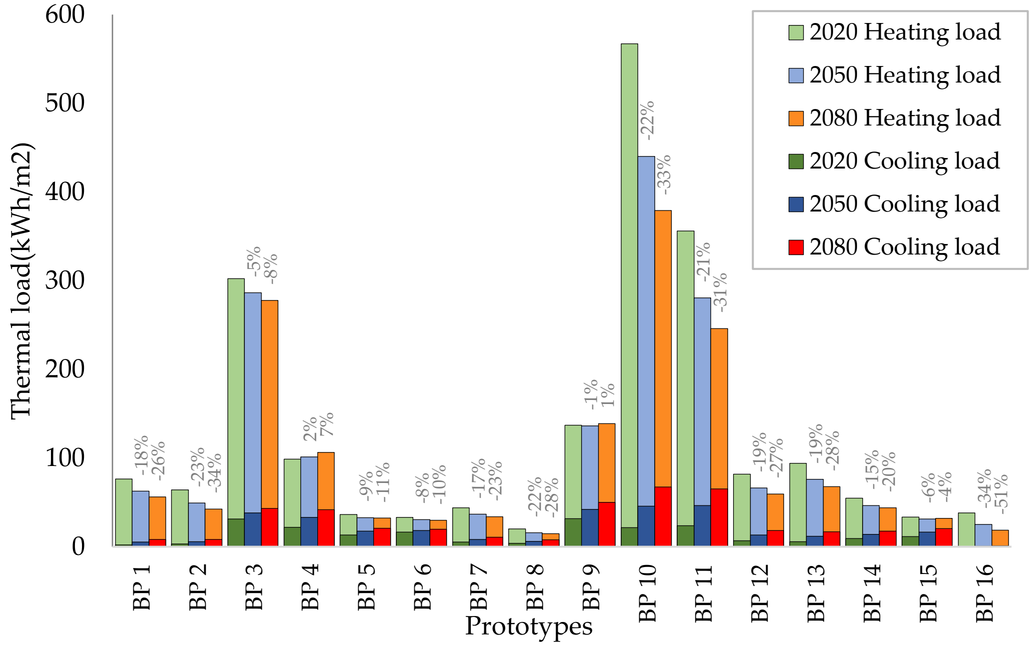


Figure 4.5. Change in the thermal load of building prototypes in 2020–2080

Even though cooling loads are on the rise, their impact on the overall thermal load remains less significant compared to heating loads across most building prototypes (BPs). A closer examination of the data (as indicated in Figure 4.6) reveals that in 25% of the studied prototypes (BP05, BP06, BP08, and BP15), the contributions of cooling and heating loads by the year 2080 are expected to reverse when compared to another group of the same size (BP01, BP02, BP03, and BP16).

This finding points to a nuanced yet meaningful trend: the more balanced the initial contributions of heating and cooling loads, the higher the likelihood for these contributions to switch predominance by 2080. In other words, for building prototypes where heating and cooling loads are closely matched, a reversal in their respective contributions to the overall thermal load is more likely to occur in the future.

Therefore, while heating loads currently dominate the thermal balance, the trend suggests a gradual and selective shift towards cooling loads becoming more impactful in certain building prototypes. This shift emphasizes the need for more adaptable building designs and energy management strategies that can account for these evolving thermal load dynamics.

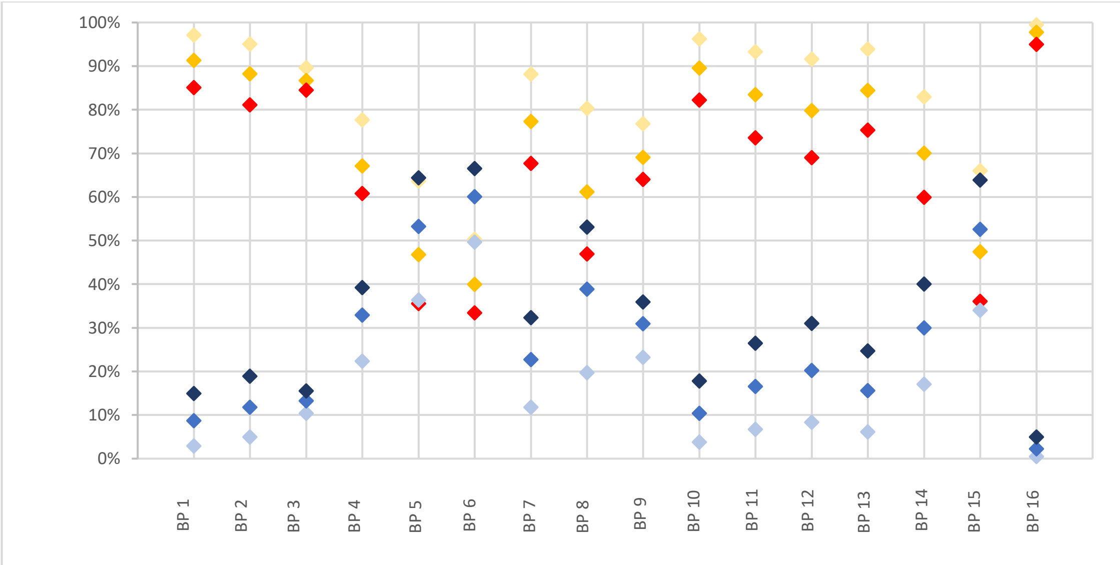


Figure 4.6. The share of thermal load of building prototypes in 2020, 2050, and 2080.

Therefore, considering the earlier discussions, it can be concluded that, in general, climate change heavily changed the energy performance of all building prototypes through increasing cooling and decreasing heating load. In this sense, as the contribution of cooling was more petite than heating in most cases, the total thermal load of most cases was reduced due to the decreased heating load. Furthermore, in one quarter of cases, the cooling increment altered heating and cooling contributions in thermal load. Finally, whenever increased heating load was reported, the total thermal load showed an upward trend in the study period.

4.9. Conclusion

As we navigate a time when the building sector accounts for a critical portion of global energy consumption, the need to understand the impacts of climate change on building energy performance is more crucial than ever. This chapter focused on these impacts for the city of Poznan in Poland. Utilizing future weather data, generated through the HadCM3 climate model and factoring in the IPCC SRES A2 greenhouse gas scenario, a picture of the city's energy future emerged. According to the projections, a significant uptick in average temperature in Poznan over the next 60 years is likely, potentially shifting the city's climate zone. Current dominant heating loads are expected to see a decline by about 40% by 2080, while the initially negligible cooling loads will sharply rise by around 135%.

In 25% of the building prototypes studied, this shift resulted in a change in the primary contributor to total thermal load from heating to cooling. In contrast, the majority of cases will still see heating as the major contributor, albeit with a reduced total thermal load. These insights suggest that future building projects in Poznan should prioritize energy-efficient cooling systems to counterbalance the increased greenhouse gas emissions that would otherwise occur.

One limitation of the study is its reliance on a single climate model, HadCM3, under the A2 emission scenario. Future research should diversify the climate models and scenarios employed to ensure broader applicability of the findings. As climate change continues to alter environmental variables, the resiliency of our buildings must be front and center in planning and policy decisions. For cities like Poznan, or others with similar climatic conditions, preparations must be made for a future where cooling requirements will assume a more significant role. Consequently, methods to reduce cooling loads could serve as a critical strategy in reducing a building's carbon footprint, thereby helping to curb the vicious cycle of climate change.

In summary, the insights from this chapter offer a valuable framework for architects, engineers, policymakers, and other stakeholders tasked with shaping our urban environments. They also underscore the need for future research that includes more comprehensive studies covering major cities across Poland, thereby providing a richer database for informed decision-making at a national level.

Chapter 5: Data-driven Methods Application

5.1. Abstract

This chapter provides a comprehensive review of machine learning algorithms commonly employed in the domain of building energy consumption prediction. A selection of methods ranging from traditional statistical models like Multiple Linear Regression (MLR) to more advanced machine learning techniques such as Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), Random Forests (RF), and Extreme Gradient Boosting (XGB) are discussed. Each algorithm is explored in detail, offering insights into its underlying mechanics, advantages, and limitations. Case studies and real-world applications are presented for each method to highlight their empirical performance and suitability for various scenarios, including short-term and long-term forecasting, as well as high-dimensional and noisy data environments. Although the selection of the optimal algorithm is recognized as a complex task—dependent on myriad factors like data availability, feature dimensionality, and specific project requirements—the chapter aims to provide an initial layer of filtration for decision-making. By comprehending the foundational principles and capabilities of these algorithms, readers are guided in aligning algorithmic choices with specific challenges and requirements. The chapter serves as an essential resource for researchers, practitioners, and policy-makers in the field, providing the foundational knowledge required for informed selection among various machine learning options for predicting building energy consumption.

5.2. Introduction

Greenhouse gas emissions, predominantly carbon dioxide (CO_2), are identified as the main instigators of global warming. Various studies highlight the building sector as a significant contributor to these emissions, accounting for 46% in the UK [213], 40% in the USA, and 27% in Australia [214]. Consequently, enhancing the energy efficiency of buildings has become a priority, not only to curb emissions but also to reduce fossil fuel consumption. For context, a 20% improvement in the energy performance of European Union (EU) buildings could lead to an estimated annual saving of 60 billion Euros [215]. To effectively mitigate the impacts of greenhouse gas emissions, substantial changes are required in various areas. These include shifts in human behavior related to energy consumption and the production of environmentally friendly products [216]. Within this multifaceted landscape, two strategies stand out as particularly impactful: the development of new, energy-efficient buildings and the optimization of energy usage in existing structures.

The cornerstone of enhancing building energy consumption lies in accurate and reliable measurement. Building energy assessment methods serve this purpose well, offering decision-makers a comparative Energy Performance Indicator (EPI) or Energy Use Intensity (EUI) [217, 218]. These indices generally represent the energy consumed by a building over a specified period, normalized by its floor area, and expressed as kWh/m^2/period.

The evaluation of energy use in buildings can be categorized into four primary techniques: engineering calculations, simulation-based benchmarking, statistical modeling, and Machine Learning (ML). Engineering approaches utilize the principles of physics to estimate energy

consumption either for entire buildings or specific subsystems. These techniques often employ advanced mathematical models and building dynamics to provide highly accurate estimations of energy use, taking into account various internal and external factors such as climate conditions, construction materials, and HVAC systems. On the other hand, simulation-based methods use specialized software and computer models to replicate building performance under set conditions. These simulations are versatile and can be applied to diverse aspects like lighting and HVAC design. With the availability of historical building energy data, top-down methods using statistical modeling have gained traction. These techniques often employ regression analyses to create what are known as data-driven surrogate models, leveraging existing data instead of intricate system details. ML, a subset of artificial intelligence, offers a unique approach to learning from data through algorithmic processes. Given its foundational ties to computational statistics, ML can also be viewed as a specialized form of statistical modeling.
Simulation-based approach enables architects and engineers to evaluate how a building's form, materials, and systems will influence its thermal efficiency before it's even built. The traditional approach to finding the best design through simulation involves a manual, iterative process that can be labor-intensive, thereby limiting the range of design alternatives.

This constraint is addressed by using optimization techniques, capable of evaluating thousands of possibilities [219]. While this reduces the need for specialized human labor, it can be computationally taxing and time-consuming. To circumvent this, surrogate or data-driven models have been suggested as a solution [220]. These models establish mathematical connections between input variables and desired outputs, such as thermal properties of materials and weather conditions to predict indoor climates. When adequately precise, these models can offer quick and accurate alternatives to traditional simulation tools during a resource-intensive design process [221].
The use of surrogate models comes with its own set of considerations, primarily concerning data accuracy and the validity of the derived relationships. This paper delves into an overlooked aspect: the selection and tuning of appropriate regression models for specific datasets. Existing literature on the use of surrogate models in building simulations commonly focuses on comparing linear models to non-linear ones, usually optimizing only a limited number of model parameters [26]. We argue that model selection should not solely be based on predictive accuracy but also take into account factors like model complexity, user-friendliness, and prediction consistency.
This study concentrates on regression algorithms aimed at linking specific building features with corresponding performance indicators, such as geometric-based features (height, width, window to wall ration, Etc.) with the energy needed for heating. Generally, regression methods are categorized into either supervised or unsupervised learning. In supervised learning, where the target outcome is predefined, the model aims to map input features (X) to one or more output variables (Y), mirroring the behavior of the real system under study. Unsupervised learning, in contrast, doesn't have a specified 'output' to predict and often employs techniques like clustering to categorize data based on inherent similarities. In the context of this study, a dataset that originates from a physics-based simulator has been used, making this endeavor a supervised learning task. This simulator serves as the benchmark for evaluating the performance of various predictive models.

Attention of this work primarily is directed towards nonlinear regression models. In such models, inputs can't be linearized even after applying transformations. For instance, although a polynomial model can be transformed into a linear model using squared or cubed inputs, the same is not possible for a Random Forest (RF) regression model. The utility of machine learning (ML) models in the realm of building analysis was initially demonstrated by Kalogirou, Neocleous, and Schizas in 1997 [222], who estimated heating loads based on building envelope characteristics and the set temperature. Subsequent research has consistently shown that nonlinear models outperform linear ones in predicting both simulated and real-world data. To give a comprehensive understanding, we provide a literature review categorized by the specific type of model employed in each study.

5.3. ML algorithms

Machine learning algorithms that are commonly used for predictive modeling encompass a wide variety of approaches, such as Long Short-Term Memory networks (LSTM), Artificial Neural Networks (ANN), and Back Propagation Neural Networks (BPNN). Additionally, Multilayer Perceptrons (MLP) are often used for tasks that require complex function approximation. On the statistical front, methods like Support Vector Regression (SVR) and Multiple Linear Regression (MLR) offer robust ways to model relationships in data. For classification tasks, Support Vector Machines (SVM) are a popular choice due to their effectiveness in high-dimensional spaces. Ensemble methods like Random Forests (RF) and Extreme Gradient Boosting (XGB) are also frequently deployed for both classification and regression problems, as they combine multiple weak learners to create a more robust model.

5.3.1. LSTM

Long Short-Term Memory networks (LSTM) represent a specialized type of Recurrent Neural Networks (RNN) that are particularly effective in addressing issues like vanishing and exploding gradients, especially during the training of long data sequences. The architecture of Long Short-Term Memory (LSTM) is composed of a singular yet critical component known as the memory unit or LSTM unit (Figure 5.1). This unit itself is an ensemble of four distinct feedforward neural networks, each having an input layer and an output layer. In every one of these neural networks, each input neuron is directly connected to all the output neurons, resulting in four layers that are fully connected within the LSTM unit. Out of these four neural networks, three serve the specialized function of information selection and management. These are known as the forget gate, input gate, and output gate. These gates handle three essential memory operations: purging information from memory (via the forget gate), adding new data into memory (through the input gate), and utilizing the stored information for computations (by means of the output gate).The remaining fourth neural network within the LSTM unit is referred to as the candidate memory. This network is tasked with generating new candidate data that could potentially be added to the memory unit.

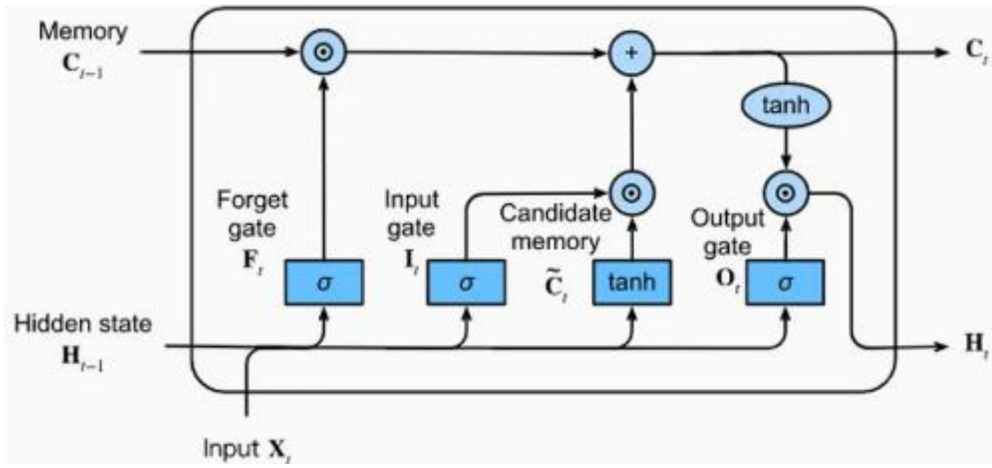


Figure 5.1. Architecture of LSTM unit, based on [223]

LSTMs have the edge over traditional RNNs when handling long sequences. For instance, Sendra-Arranz and Gutierrez [224] employed LSTM networks to forecast the daily energy usage of HVAC systems. Similarly, Pittí et al. [225] designed an LSTM-based model for predicting the daily consumption of a heat pump located at the Teatro Real in Spain. In a more intricate application, Rosemary et al. [226] employed an encoder-decoder LSTM model to carry out hourly and day-ahead predictions for residential high-voltage alternating current use and solar power generation, integrating load history and meteorological variables. Wang et al. [227] used LSTMs for predicting various parameters like electricity consumption, lighting loads, occupancy rates, and internal heat gains in double-office buildings across the United States. Jogunola et al. [228] utilized Convolutional Neural Networks (CNN) for feature extraction and then employed LSTMs to forecast energy consumption in diverse building types. Das et al. [229] took it a step further by developing a bidirectional LSTM (Bi-LSTM) model to offer one-day and one-week electricity consumption forecasts. Ullah et al. [230] experimented with multiple variants of LSTM, including Bi-LSTM and multilayer LSTM (M-LSTM), to predict home energy usage. Li et al. [231] combined K-means clustering with LSTM to predict energy loads at the granularity of individual building floors. In a hybrid approach, Kim and Cho [232] merged CNN with LSTM to capitalize on both spatial and temporal features for accurate residential energy consumption prediction. He and Tsang [233] introduced a novel hybrid model, coupling improved fully integrated empirical modal decomposition with adaptive noise (iCEEMDAN) and LSTM, for precise short-term load forecasting in educational institutions. Ijaz et al. [234] employed convolutional LSTM for spatial feature extraction and Bi-LSTM for sequence learning, thereby minimizing energy consumption prediction errors.

Chalapathy et al. [235] demonstrated an LSTM-based Recurrent Neural Network with multiple-input multiple-output (RNN-MIMO) architecture that excelled in both one-hour and one-day multi-step prediction scenarios across various settings like office buildings, hospitals, and shopping malls. Compared to existing shallow machine learning models like SVR and Extreme Gradient Boosting (XGB), this RNN-MIMO model showed a 33% increase in average accuracy. However, Salah et al. [236] noted that poor performance in LSTM prediction could result from improperly set hyperparameters or anomalies in energy consumption data. To address this, they used two evolutionary metaheuristics, namely the genetic algorithm (GA) and particle swarm optimization (PSO), to fine-tune the LSTM model's performance for power load prediction, outperforming traditional models like SVR, Random Forest (RF), ANN, and manually-tuned LSTM configurations.

5.3.2. Artificial Neural Network

Neural networks are widely recognized as leading machine learning techniques in the domain of building energy prediction. They are particularly adept at modeling complex, non-linear systems. With the application of specialized methods, Artificial Neural Networks (ANNs) can achieve a level of immunity to noise and faults, making them effective at learning essential patterns in building systems [237]. The conceptual foundation of ANNs is derived from neurobiology. A variety of ANN architectures have been developed for diverse applications, such as Feed Forward Networks (FFN), Radial Basis Function Networks (RBFN), and Recurrent Neural Networks (RNN). These networks typically comprise multiple layers (at least two) of neurons connected by activation functions. Common activation functions utilized in these networks include linear, sigmoid, and hard-limit functions [238]. A neural network is typically structured with three main layers. The first of these is the input layer, which comprises input neurons responsible for forwarding data to the next layer, known as the hidden layer. Within the hidden layer, calculations are performed on the received input data, and the results are then sent to the output layer. Components of the hidden layer include weights, activation functions, and cost functions (Figure 5.2).

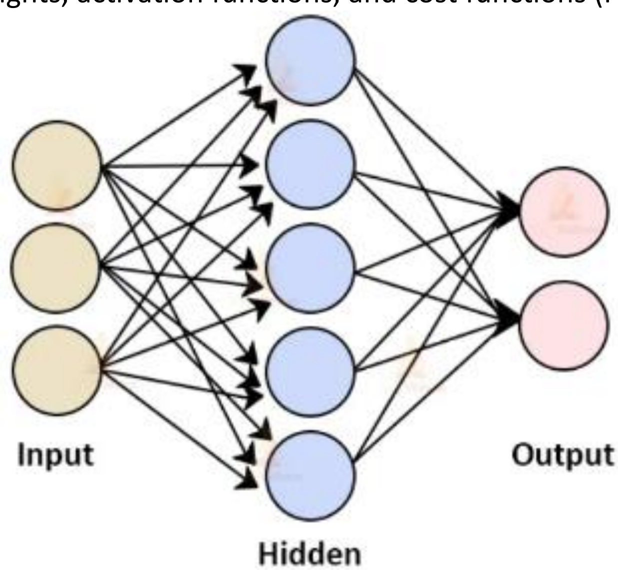


Figure 5.2. Common architecture of ANN, based on [239]

The term 'weight' refers to the numerical values that define the connections between neurons. These weights play a crucial role in determining the network's learning capabilities. As the artificial neural network undergoes the learning process, these weights between neurons are dynamically adjusted. Artificial Neural Networks (ANNs) have shown promising results in a variety of intricate tasks that require high temporal resolution, particularly in forecasting short-term heating loads within buildings [31]. For instance, Yaser et al. [240] leveraged ANNs for forecasting the daily energy consumption associated with a laboratory fan coil system. In a different setting, Byeongmo et al. [241] introduced an ANN-based control strategy for managing the climate in a

Double-Skin Facade (DSF) office building situated in a humid, hot environment, achieving a cost savings of 4.5%. Muralitharan et al. [242] optimized an ANN using Genetic Algorithms (GA) and Particle Swarm Optimization (PSO) to automatically adjust its hyperparameters.
Beyond standard ANNs, specialized variants like Multilayer Perceptrons (MLP), Back Propagation (BP) neural networks, Elman neural networks, and Echo State Networks (ESN) have also been explored. Andrew et al. [243] employed an MLP network to optimize energy usage in HVAC systems while maintaining a comfortable thermal environment, even with fluctuating occupancy levels. Mitali et al. [244] utilized a BP neural network to make predictions about residential HVAC energy consumption. Ruiz et al. [245] developed a methodology based on the Elman neural network to enhance energy efficiency in university buildings without sacrificing comfort or health. All these cases and applications show the strength of ANN in general for the specific case that this research is looking for.

5.3.3. Support Vector Machine

Support Vector Machines (SVMs) have gained considerable recognition for their effectiveness in solving classification problems, where the aim is to categorize data into distinct classes. While they are primarily known for this role, their application extends into regression analysis as well, albeit less commonly. In the context of regression, these models go by the name of Support Vector Regression (SVR). SVR models aim to predict continuous outcomes as opposed to discrete classes, and they share many of the same foundational principles with their classification counterparts. Despite being less documented, SVRs are gaining traction for their ability to handle complex, high-dimensional data in predictive modeling.
SVR seeks to minimize prediction error by identifying the optimal hyperplane and narrowing the gap between predicted and observed values. In the equation provided, reducing the value of 'w' is equivalent to maximizing the margin, as illustrated in Fig. 5.3.

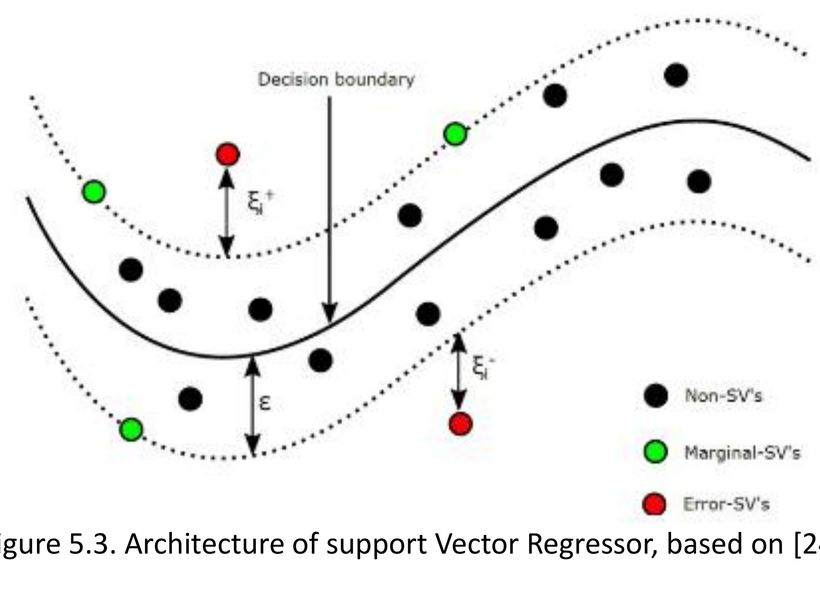


Figure 5.3. Architecture of support Vector Regressor, based on [245]

In the realm of Building Energy Consumption Prediction (BECP), both Support Vector Machine (SVM) and Support Vector Regression (SVR) rely on nonlinear mapping techniques. These map

input data into a high-dimensional space to execute linear regression, ultimately achieving a nonlinear regression effect on the original input data. For example, Zhong et al. employed SVR to forecast the cooling load in a large office structure located in Tianjin [246]. Similarly, Li et al. [247] created a moment-to-moment building cooling load predictive model, which was based on SVM and Back Propagation (BP) algorithms. They applied this model to an office building in Guangzhou, finding that SVM outperformed BP in terms of prediction accuracy. Ding et al. [248] fashioned both GA-SVR and GA-WD-SVR models to predict the cooling load in office buildings at varying time scales. They found that GA-SVR was superior for next-day cooling load predictions, while GA-WD-SVR excelled in one-hour cooling load forecasting. Paudel et al. [249] used SVM to estimate the thermal loads in a low-energy building, demonstrating that the model was more accurate when relying on relevant data (RMSE = 3.4) compared to using all available data (RMSE = 7.1).
Seyedzadeh et al. [84] conducted a comparative analysis of multiple algorithms including SVM, Random Forest (RF), Recurrent Neural Networks (RNN), and Extreme Gradient Boosting (XGB) for predicting the cooling and heating loads in both commercial and residential structures. They concluded that SVM was the most effective choice for relatively straightforward datasets. To further fine-tune prediction accuracy, Zhao and Liu [250] first applied wavelet transform (WT) for noise reduction on historical energy data. They then utilized low-correlation features for Partial Least Squares (PLS) predictions and high-correlation features for SVM-based predictions, which significantly improved predictive accuracy. Finally, Ngo et al. [251] introduced an innovative time series wolf-inspired optimization SVR model (WIO-SVR) designed for predicting energy consumption across multiple buildings.

5.3.4. Random Forest

A Random Forest is essentially a collection of decision trees, constructed in a specific manner that introduces randomness into the process. Each tree within the ensemble is generated from a unique subset of rows and, for each node within the tree, a distinct set of features is chosen for making the split. Every tree in the ensemble contributes its own individual prediction. These multiple predictions are subsequently averaged to arrive at a single, consolidated output (Figure 5.4).The practice of averaging the predictions from multiple trees enhances the performance of a Random Forest over that of a single Decision Tree. This approach not only improves the model's overall accuracy but also minimizes the likelihood of overfitting the data. In the context of Random Forest Regressor, the final prediction is essentially an aggregate, calculated as the average of the individual predictions generated by each of the trees within the forest.

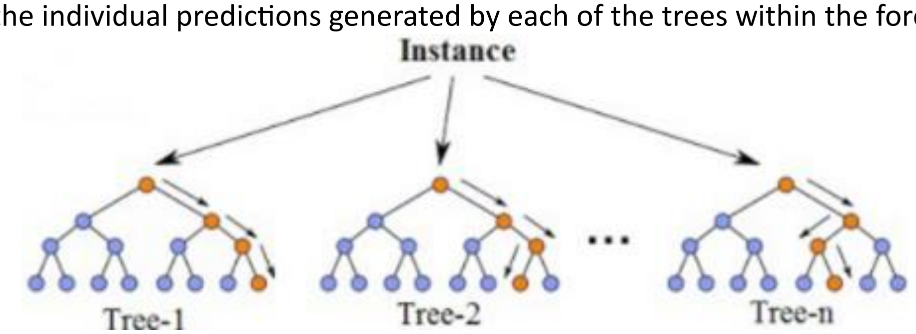


Figure 5.4. Conceptual framework of random forest

One of the key advantages of RF is its ability to automatically conduct feature selection, identifying interactions between various variables without requiring manual feature selection. Additionally, it can maintain high levels of predictive accuracy even when some features are missing. In applied research, Wang et al. [252] utilized the RF model to forecast hourly electricity consumption in two educational buildings in Florida. Seyedzadeha et al. [253] employed RF to predict the cooling and heating loads in a range of residential and commercial structures. Further, Rana et al. [254] used a specialized version of RF, referred to as a divided number regression forest, to predict the cooling load for a large retail shopping center and an office building in Australia over a one-month period. Ahmad et al. [255] took a multi-model approach, incorporating a binomial decision tree, tight regression Gaussian process model, stepwise Gaussian process regression, and generalized linear regression models to forecast electricity consumption on a monthly, quarterly, and annual basis.

5.3.5. Extreme Gradient Boosting

Extreme Gradient Boosting (XGB) is an advanced implementation of gradient-boosted decision trees specifically engineered for speed and high performance. One of its major strengths is its ability to effectively handle nonlinear relationships in data without requiring substantial fine-tuning [256]. For example, João et al. [257] introduced a hyperparametric adaptive model based on the Jaya algorithm and utilized XGB to accurately predict energy consumption patterns in residential structures. Similarly, Lu et al. [258] leveraged the capabilities of XGB to forecast the energy requirements of a water tower. They chose XGB particularly for its prowess in refining predictions by smoothing out raw data that showed significant fluctuations. Furthermore, Feng et al. [259] employed XGB to estimate cooling loads in three different homes located in the United States, each subjected to varying climatic conditions—hot, humid, cold, and dry. Therefore, it is fair to say that XGB has proven to be a versatile and powerful tool for diverse applications, particularly in scenarios demanding rapid and accurate predictions.

5.3.6. LGBM Regressor

The LightGBM (LGBM) algorithm serves as a sophisticated learning framework that capitalizes on tree-based learning algorithms. Designed with an emphasis on efficiency and distributed computing, LightGBM brings several advantages to the table. Among these are rapid training speeds, heightened efficiency, minimal memory consumption, and impressive prediction accuracy. Moreover, it comes with built-in support for parallel computations, making it particularly well-suited for handling large-scale data management tasks [260]. One of the distinguishing features that sets LightGBM apart from its counterpart, XGBoost (XGB), is its utilization of a histogram-based learning algorithm. This specific approach not only accelerates the model training process but also significantly reduces the memory footprint required. The histogram-based technique allows LightGBM to construct compressed histograms, which in turn, speeds up both training and prediction, while concurrently reducing memory consumption [261]. In recent times, the LightGBM algorithm has found applicability in various studies across different domains. It has been adopted for prediction tasks in diverse fields, showcasing its versatility and effectiveness [262, 263]. Whether employed in predictive analytics, classification tasks, or other machine learning applications, LightGBM consistently proves its mettle as a robust and scalable solution.

5.3.7. Multiple Linear Regression

Multiple Linear Regression (MLR) serves as a mathematical framework for quantifying the linear relationships between input variables and energy consumption in buildings [264]. This approach has been employed in a variety of settings to predict energy usage effectively. For instance, Fumo and Biswas applied linear regression models to forecast the energy consumption of HVAC systems in residential buildings at different time scales. Their study revealed that a quadratic regression model produced superior outcomes for short-term (1-hour) predictions, but did not maintain this advantage for longer periods (1-day) [265]. Similarly, Fan and Ding [266] used a specialized form of MLR known as Multiple Nonlinear Regression (MNR) to estimate the hourly cooling load of a large library. This approach adapted the basic principles of MLR to a more complex nonlinear context, thereby enhancing its predictive capabilities. Furthermore, Chen et al. [267] tackled the issue of weak generalizability in models trained on limited data samples by introducing a PB-MLR (Pattern-based Multiple Linear Regression) model. Their innovative model was designed to predict the cooling load of office buildings with greater accuracy over a time-wise scale. Overall, MLR and its various adaptations provide a robust and versatile toolset for predicting energy consumption, with the flexibility to address both linear and nonlinear scenarios.

5.4. Conclusion

In this chapter, a comprehensive range of machine learning algorithms has been explored, each with its prominence in the domain of building energy consumption prediction. The spectrum spans from traditional approaches like Multiple Linear Regression (MLR) to more advanced, computationally intensive algorithms such as Artificial Neural Networks (ANNs) and Extreme Gradient Boosting (XGB). With each algorithm offering its unique set of advantages, complexities, and limitations, the vast array of options is indicative of the differential suitability these algorithms present for various datasets and prediction challenges. The intricacies inherent in building energy systems—affected by a multitude of variables, from occupancy patterns to external climatic conditions—suggest that a one-size-fits-all solution is elusive. Consequently, the task of selecting the most appropriate algorithm from among this wide array is far from trivial. The optimal choice can vary based on myriad factors, including data availability, feature dimensionality, the need for model interpretability, and the desired prediction accuracy. Furthermore, the unique characteristics and patterns embedded in each dataset can often make some algorithms more attuned to capturing them than others. Amidst this complexity, the insights derived from this chapter become instrumental. While it is recognized that the optimal algorithmic choice will likely differ across datasets, an understanding of the foundational principles, capabilities, and limitations of these algorithms can provide an initial layer of filtration in the decision-making process. Knowledge of the core mechanics of each algorithm aids in aligning algorithmic choices with the specific requirements of a project. For instance, if high temporal resolution is sought, it might be inferred that ANNs or XGB could be preferable. Conversely, for scenarios where interpretability holds precedence, simpler algorithms like MLR could be deemed more fitting.

Furthermore, the studies and applications presented for each algorithm in this chapter offer insights into their empirical performance. These real-world applications can be viewed as benchmarks, shedding light on potential performance metrics of each algorithm under analogous conditions or similar application domains.

In conclusion, while further steps, such as cross-validation or hyperparameter tuning, might be required in the quest to select the best-suited algorithm, the foundational knowledge provided in this chapter serves as an essential backdrop. This foundational understanding offers invaluable guidance in navigating the plethora of machine learning algorithms available for building energy consumption prediction. Such knowledge underscores the importance of informed decisions, thereby facilitating the effective and nuanced utilization of these potent computational tools.

Chapter 6: Results

6.1. Abstract

This chapter delves into the comprehensive process of deploying and fine-tuning machine learning models for a specific application, emphasizing the optimization of Artificial Neural Network (ANN) as the optimum model. Initially, a thorough exploratory data analysis is conducted, encapsulating aspects like statistical inference, inter-relationships among variables, distribution normality, input/output relationships, and the impact of categorical variables. Subsequent sections address the deployment of various machine learning models, including LightGBM Regressor, Random Forest, and Support Vector Machine, culminating in the deployment and fine-tuning of ANN. Methodical model tuning is performed to enhance performance metrics, with a particular focus on the mitigation of overfitting through input normalization and model regularization. The chapter concludes with an evaluation of the final models, highlighting the efficacy of the final ANN model through a real-life case study involving 3D reconstruction and energy usage intensity (EUI) simulation. Comparative analysis validates that the final model shows a significant performance improvement over the base model, substantiating its potential for delivering more reliable results in similar applications.

6.2. Exploratory Data Analysis

In this section, the findings from the Exploratory Data Analysis (EDA) of the generated dataset are presented. For a comprehensive understanding of the dataset, it is advised to examine both the structure or "shape" of the data as well as the mathematical characteristics of each individual variable. These mathematical aspects serve essentially as statistical inferences, providing a robust framework that can accurately model and describe the behavior and distribution of the dataset. When the dimensions of the dataset are considered, it is notable that it consists of 3,000 rows of data. Each row is comprised of 11 distinct columns, representing various features or variables. These columns are labeled as follows: W for Width; L for Length; H for Height; WWR, which stands for Window to Wall Ratio; R representing Rotation; RC indicating Relative Compactness; RA for Roof Area; V for Volume; CL representing Characteristic Length; GA for Glazing Area; and EUI, which indicates Energy Use Intensity (Figure 6.1).

	W	L	H	WWR	R	RC	RA	V	CL	GA	EUI
0	3.00	5.00	3.00	0.20	0	0.95	15.00	45.00	0.58	1.80	68.39
1	3.50	5.00	3.00	0.20	0	0.95	17.50	52.50	0.61	2.10	68.04
2	4.00	5.00	3.00	0.20	0	0.95	20.00	60.00	0.64	2.40	67.78
3	4.50	5.00	3.00	0.20	0	0.95	22.50	67.50	0.66	2.70	67.56
4	5.00	5.00	3.00	0.20	0	0.94	25.00	75.00	0.68	3.00	67.41

Figure 6.1. view from the generated dataset

6.2.1. Statistical Inference

By considering both the number of variables and their respective value ranges, a refined understanding of the dataset can be developed. This, in turn, sets the stage for more targeted analysis and the potential construction of predictive models. Initially, as per standard data science workflows, the presence of null values in the dataset should be scrutinized. However, given that this dataset is the output of simulation-based data generation, no null values are present. To ensure the data quality, this absence of null values has been confirmed through checks. Subsequently, the focus shifts to the statistical inferences associated with each variable. For the purposes of this research, the statistical metrics examined include the number of observations, the mean, standard deviation, minimum value, first quartile, second quartile (or median), third quartile, and maximum value, as detailed in Table 6.1.

Table 6.1. Description of the dataset

	W	L	H	WWR	R	RC	RA	V	CL	GA	EUI
Count	3000	3000	3000	3000	3000	3000	3000	3000	3000	3000	3000
Mean	4.00	6.00	3.50	0.50	150.00	0.94	24.00	84.00	0.70	7.00	86.54
Std	0.71	0.71	0.35	0.22	102.49	0.01	5.12	19.92	0.06	3.50	14.06
Min	3.00	5.00	3.00	0.20	0.00	0.91	15.00	45.00	0.58	1.80	60.13
25%	3.50	5.5	3.25	0.35	60.00	0.93	20.00	68.25	0.66	3.86	75.19
50%	4.00	6.00	3.50	0.50	150.00	0.94	24.00	82.50	0.70	6.75	85.24
75%	4.50	6.50	3.75	0.65	240.00	0.95	27.50	97.50	0.74	9.60	96.02
Max	5.00	7.00	4.00	0.80	300.00	0.97	35.00	140.00	0.84	16.00	131.87

In the subsequent phase, a recommendation has been made to decrease the data's dimensionality by eliminating extraneous variables. While certain attributes in the analysis may initially appear crucial, their behavior might closely mirror that of other parameters. Consequently, such attributes could be deemed redundant and safely disregarded in the context of machine learning predictions.

6.2.2. Inter-relationship

To comprehend the interrelationships among the variables, the utilization of a correlation matrix is highly advisable. This tool is valuable not only for assessing statistical correlations but also for providing a visual means to easily grasp the behavior of each attribute. This visual insight allows for effortless comparisons between different variables, enhancing the overall understanding of the dataset's characteristics (Fig 6.1).

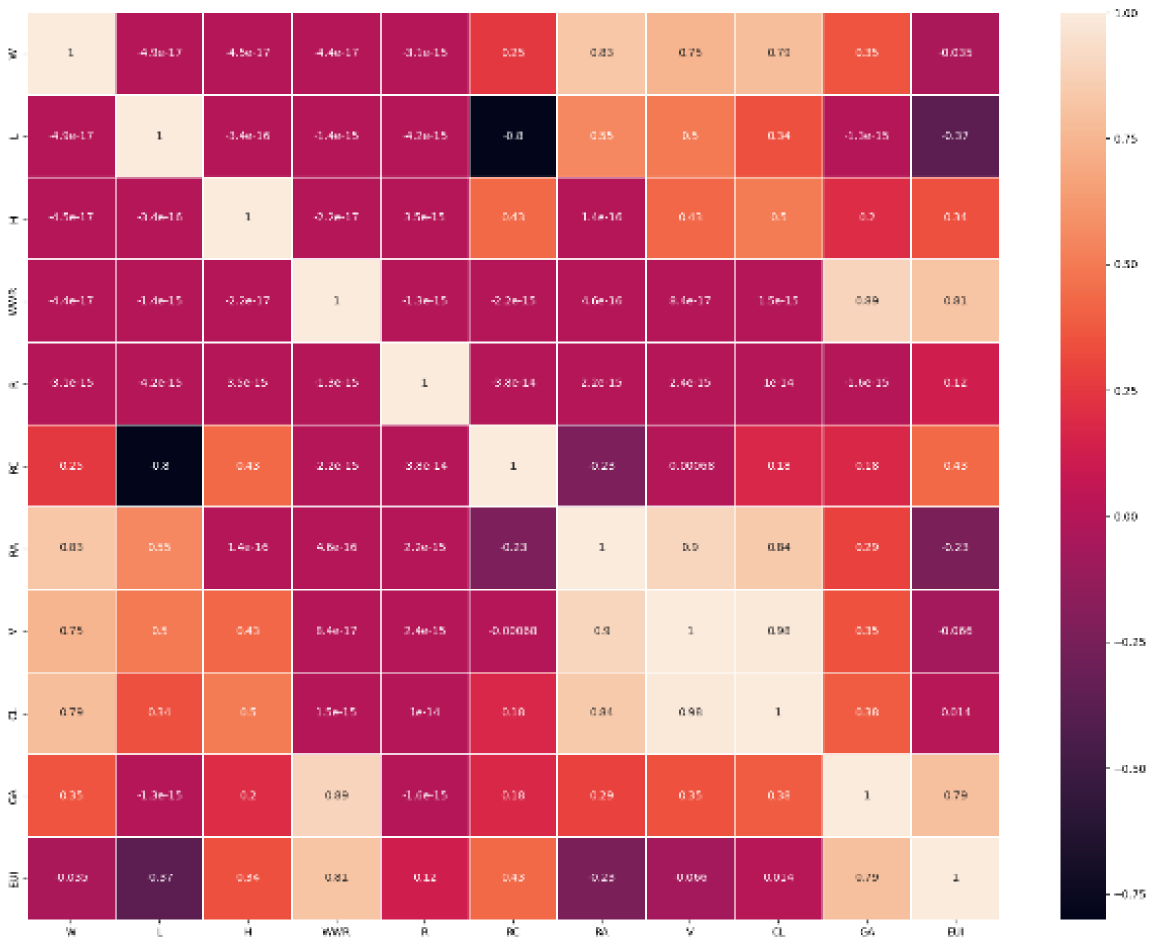


Figure 6.1. correlation matrix of the variables

In the study, several key insights are derived that merit consideration for future analyses. Specifically, a notable degree of similarity is observed between the variables Glazing Area (GA) and Window-to-Wall Ratio (WWR), which exhibit an 89% resemblance. Given this substantial overlap, the removal of one of these variables from future analyses is recommended to minimize redundancy. An even higher similarity rate of 98% is observed between Characteristic Length (CL) and Volume (V), thereby making an even stronger case for the elimination of one of these variables to enhance the model's efficiency.

From another perspective, the importance of closely examining both input features and output labels in the context of machine learning algorithms is emphasized. Understanding the effect of each input feature on the output label is considered critical for the construction of effective predictive models. Features and labels are not mere data points; rather, they are identified as playing pivotal roles in determining the algorithm's predictive accuracy and reliability. The focus of the research is on assessing the impact of various geometric-based features on Energy Use Intensity (EUI), which serves as the output label for this study. These features are categorized into two levels: The first level includes basic geometrical dimensions such as Length (L), Width (W), Height (H), Ratio (R), and Window-to-Wall Ratio (WWR). The second level comprises more complex variables like Relative Compactness (RC), Roof Area (RA), Volume (V), Characteristic Length (CL), and Glazing Area (GA), which are derived from the primary geometric attributes. This can also be seen with more details, scatter plot will show not only the similarity in the behavior for each pair of variables, but also the type of relationship (Figure 6.2).

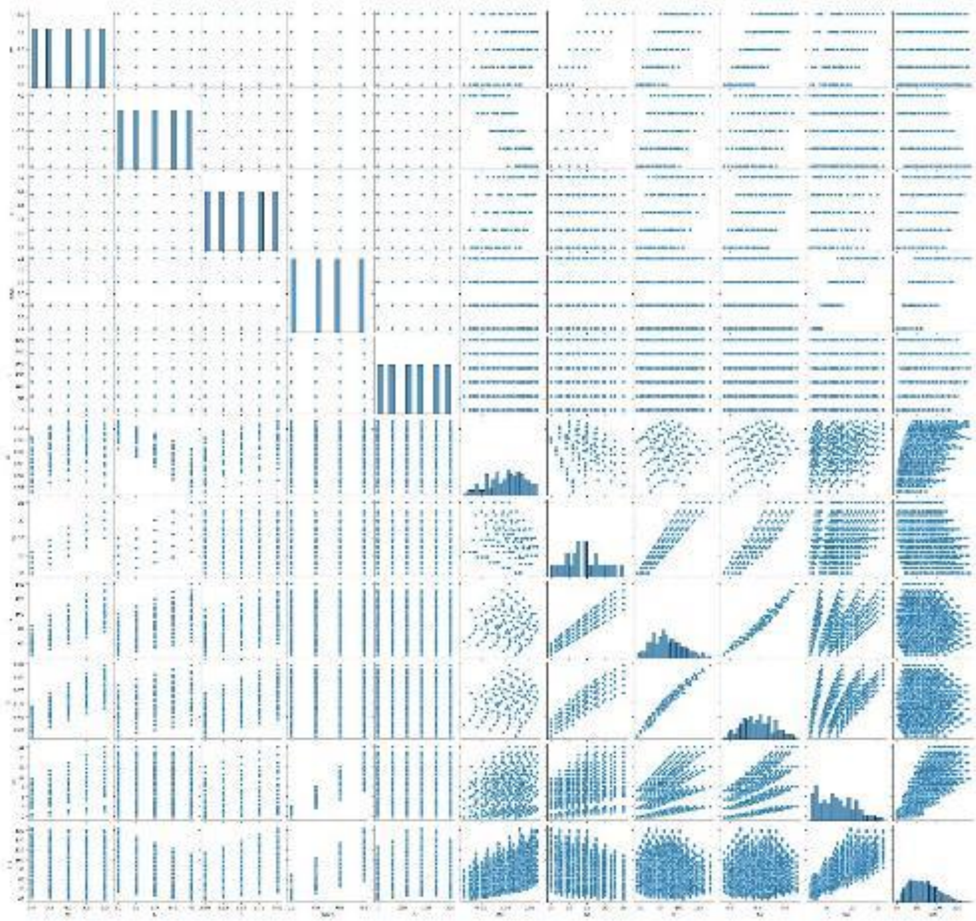
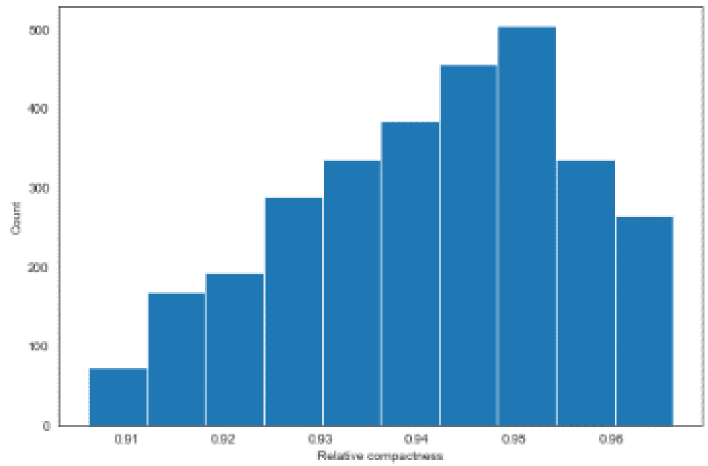


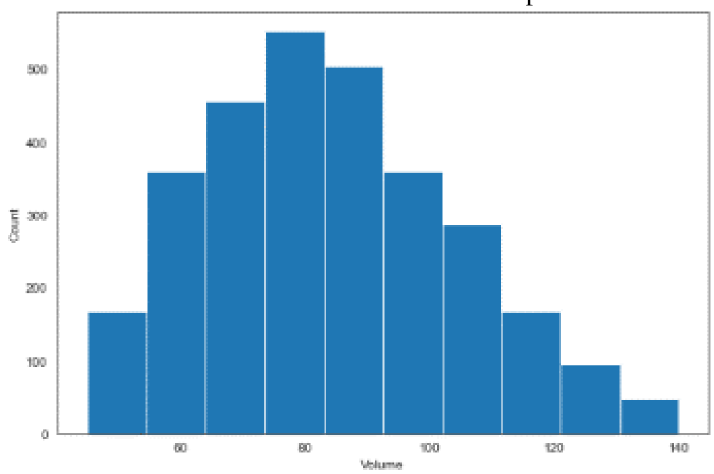
Figure 6.2. scatter plot show of pair variables

6.3.3. Distribution, normality

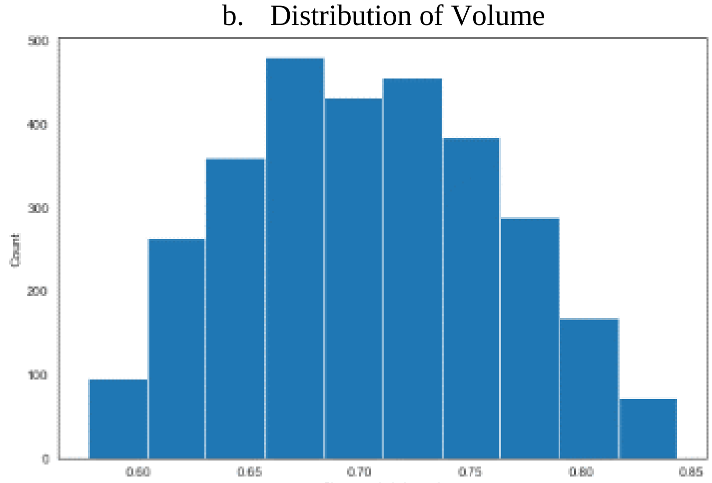
Nonetheless, for a more nuanced understanding of the data, examining the distribution of each variable is advisable. Such scrutiny is particularly crucial for conducting normality checks, as data that does not conform to a normal distribution can pose challenges in the machine learning process. Distribution of the level 2 features as well as output can provide more information about the skewness of the data.



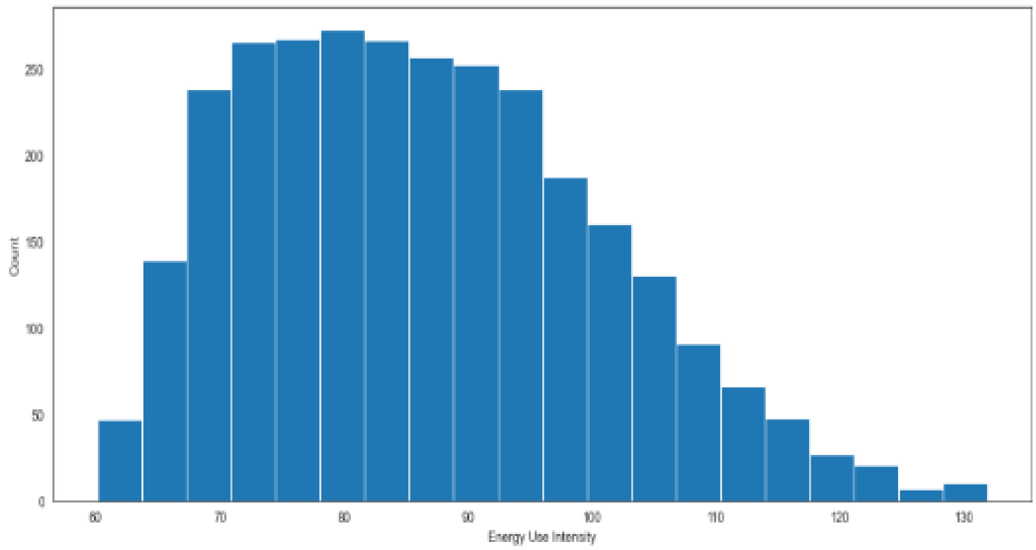
a. Distribution of Relative compactness



b. Distribution of Volume



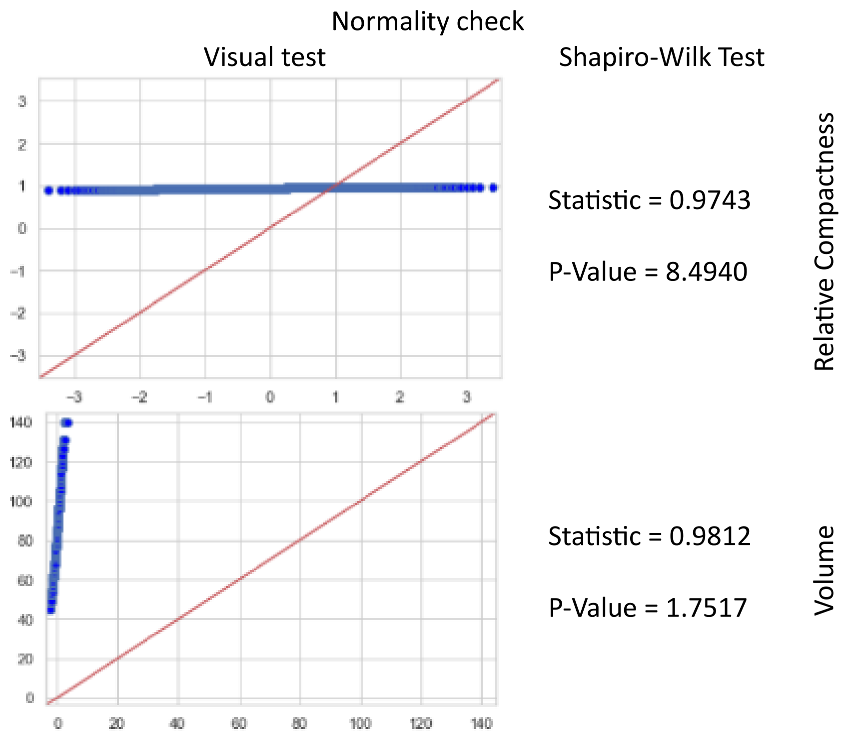
c. Distribution of Characteristic length



d. Distribution of EUI

Figure 6.3. Distribution of level 2 inputs and output

Upon closely examining Figure 6.3, it becomes evident that nearly all the variables investigated in prior stages exhibit skewness, with the exception of Characteristic Length (Figure 6.4). This observation is not trivial; it holds significant implications for the implementation of machine learning algorithms aimed at making predictions. Skewed data can introduce elements of uncertainty and inaccuracy, thereby potentially compromising the robustness of predictive models. Therefore, it is recommended to account for this aspect of data distribution when designing and deploying machine learning algorithms.



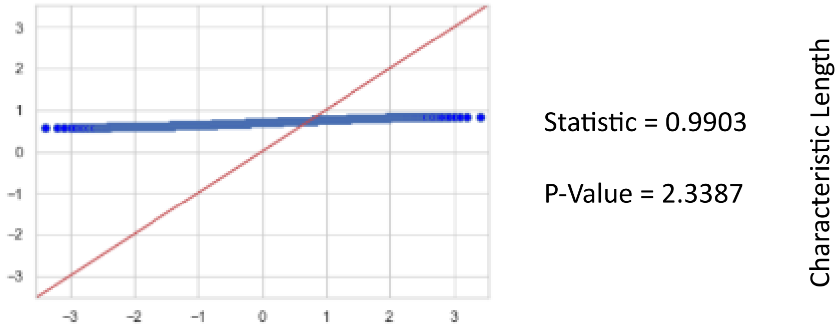


Figure 6.4. Input normality check

6.3.4. Input/Output relationship

In the subsequent phase, to better understand the influence exerted by each input variable on the output variable, which in this case is Energy Use Intensity (EUI), a preliminary visual analysis is undertaken. Utilizing scatter-plot graphs paired with fitted linear regression lines, a foundational understanding of the relationships between each input variable and the EUI output can be gleaned. This graphical approach serves as an effective way to visually capture and assess the interactions between the variables, providing initial insights that may guide further in-depth analysis (Figure 6.5).

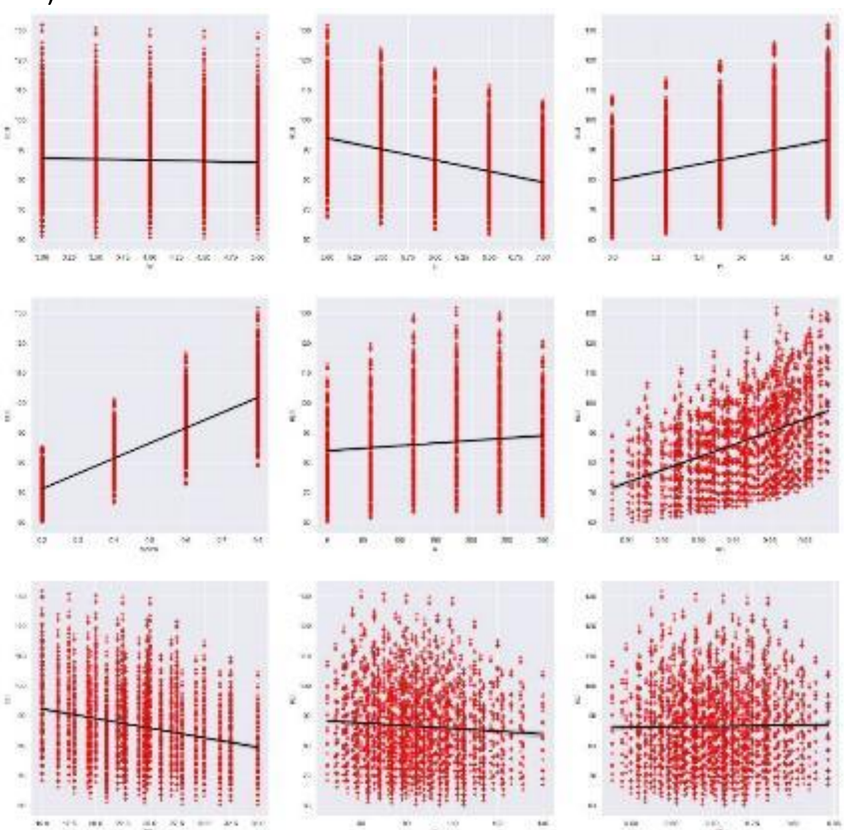


Figure 6.5. Input/output relationship

Beyond simple visual analysis, the Pearson correlation coefficient serves as a powerful statistical tool to quantify the relationships between the features and the targeted output, in this case, Energy Use Intensity (EUI). According to Figures 6.6 and 6.7, it becomes abundantly clear that the Window to Wall Ratio (WWR) has the most pronounced and positive correlation with EUI. This implies that an increase in the window area is associated with higher energy consumption, a conclusion that intuitively makes sense. Glazing Area (GA) as expected exhibits a behavior largely similar to that of WWR, as observed in Figure 6.2. This similarity in behavior suggests that incorporating both WWR and GA into machine learning models might be redundant. Therefore, only one of these variables may be necessary for effective model deployment. Following WWR and GA, Relative Compactness (RC) emerges as having the next strongest positive correlation with EUI. This further illuminates the variables that are pivotal in shaping energy usage patterns. Additionally, when examining the interrelationships among Characteristic Length (CL), Volume (V), and Roof Area (RA), it becomes evident from Figure 6.7 that these variables exhibit highly similar behaviors. Consequently, choosing just one of these variables for inclusion in the machine learning model would likely suffice, as incorporating all would not provide additional unique information.

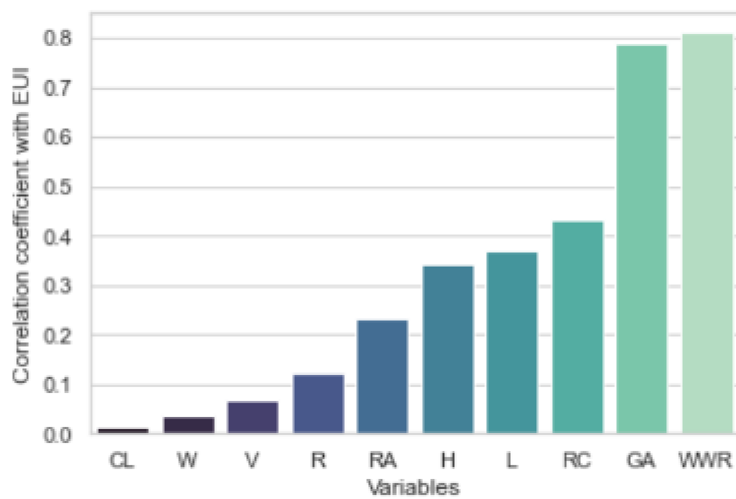


Figure 6.6. Input/output correlation coefficient

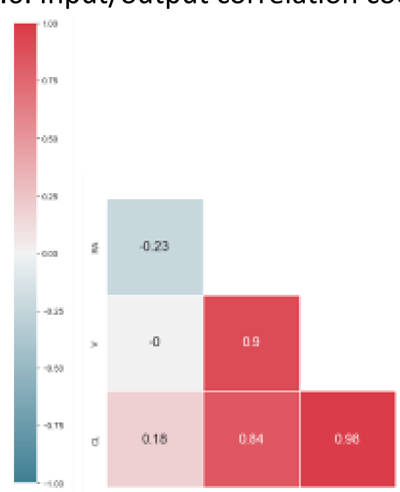


Figure 6.7. Inter-relationship of a subset of features

6.3.5. Categorical additions

To assess whether the generated dataset can serve as an accurate representation of the broader population in the context of building clusters, an additional categorical variable has been introduced. This variable serves to categorize the Energy Use Intensity (EUI) into five distinct levels: 'Very Low,' 'Low,' 'Medium,' 'High,' and 'Very High' (Table 6.2). By doing so, it becomes possible to more effectively evaluate the data's coverage and its ability to encapsulate the range of EUI rates that may be encountered in the overall population. This categorization allows for a nuanced interpretation of the data, enabling stakeholders to better understand the spectrum of energy use across different building configurations. It can also provide valuable insights into how representative the dataset is of various energy usage scenarios, which is crucial for ensuring that any subsequent analyses or predictive models developed are both robust and generalizable.

Table 6.2. categorical values description

EUI Rate	Count	Mean	Std	Min	25%	50%	75%	Max
Very Low	691.00	69.37	3.36	60.13	67.18	69.78	72.15	74.46
Low	1065.00	81.54	4.10	74.48	77.97	81.52	85.06	88.80
Medium	842.00	95.23	4.05	88.83	91.63	94.89	98.65	103.16
High	337.00	108.85	3.91	103.18	105.48	108.17	111.74	117.27
Very High	65.00	122.82	4.02	117.73	119.62	121.98	124.94	131.87

Upon closer examination of this newly introduced categorical data, along with the distribution of cases across each category, it is noteworthy that more than half of the instances fall within the 'Low' and 'Very Low' categories. This distribution could be seen as favorable when considering the intra-group conditions, as illustrated in Figure 6.8. However, relying solely on these categories may not provide a fully accurate understanding of the energy consumption levels within this specific building cluster. To obtain a more precise insight into the energy consumption patterns, alternative approaches or supplementary metrics may be warranted. The goal is to ensure that the understanding of energy usage within this cluster is both comprehensive and nuanced, allowing for more effective future analyses or interventions. Thus, while the initial categorization serves as a helpful starting point, additional methods should be explored for a more accurate assessment of the cluster's energy consumption levels.

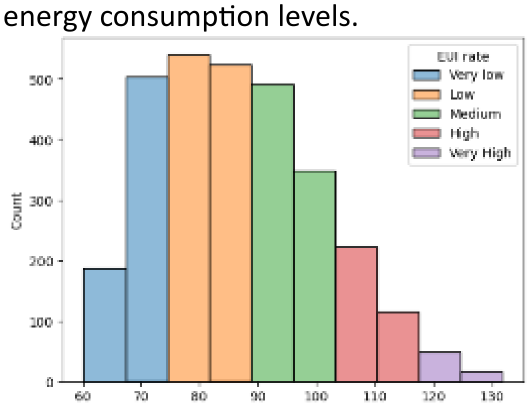


Figure 6.8. EUI categorical values count

To further refine the analysis, the dataset has been evaluated against the legally acceptable Energy Use Intensity (EUI) standards as determined by the Polish government, which were introduced in Chapter 3. According to these standards, an EUI rate higher than 70 kWh/m^{2}y is deemed unacceptable. Consequently, an additional categorical variable has been generated to assess the compliance of each case within the dataset in accordance with Polish regulations. The findings reveal that only 11% of the cases in the dataset meet the acceptable EUI rate as per Polish governmental standards. Intriguingly, this percentage aligns precisely with the proportion of acceptable cases within the broader population of building clusters in Poland, as shown in Figure 6.9. This alignment suggests that the simulation-based dataset serves as a reliable representative of the building cluster under study, especially in the context of compliance with Polish regulations. Therefore, the dataset not only helps in understanding the nuances of EUI rates but also validates its own representativeness by closely mirroring the compliance rates observed in the wider population. This adds a layer of credibility to the dataset, affirming its utility as a tool for further research and analysis in line with Polish energy consumption regulations.

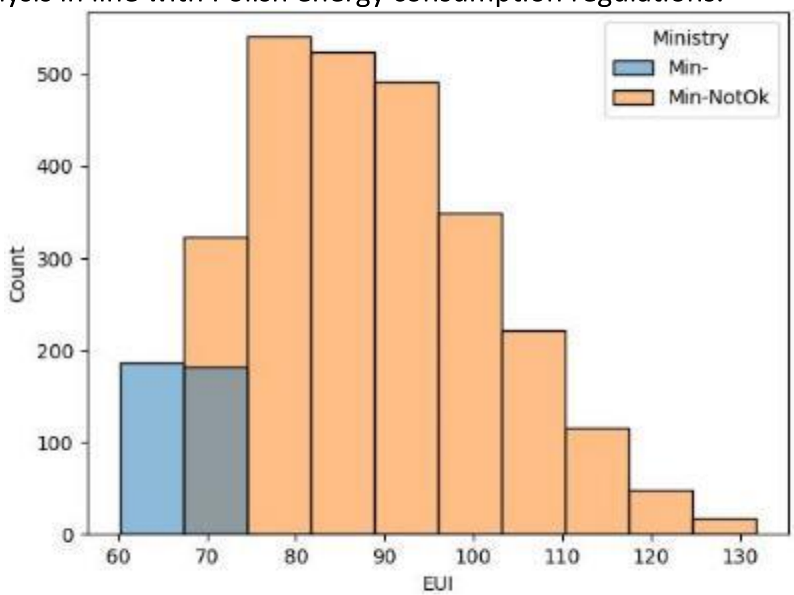
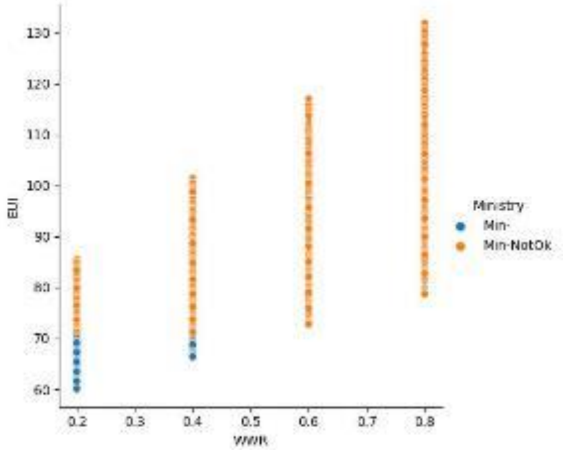
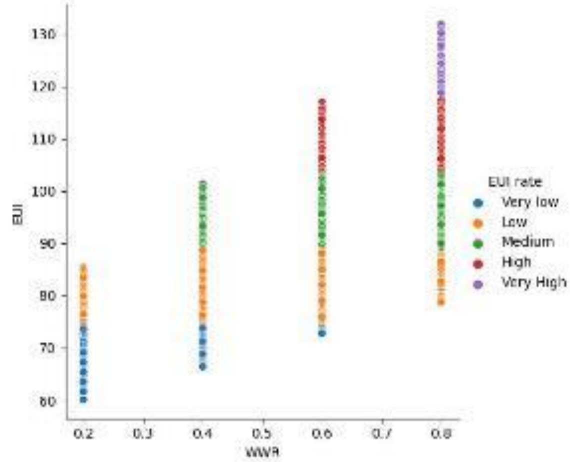


Figure 6.9. Status of cases categorized by being acceptable by Polish Ministry of Energy

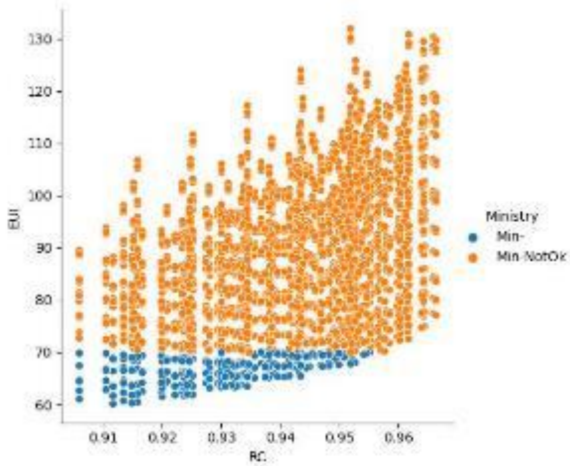
In a parallel vein, an examination of the impacts exerted by individual variables on the Energy Use Intensity (EUI) rate and its compliance with Polish government standards offers further insights, as illustrated in Figure 6.10. When it comes to the Window to Wall Ratio (WWR), it is observed that only values of 0.2 and 0.4 meet the acceptable EUI criteria set forth by Polish regulations. Conversely, a WWR of 0.8 is associated with a "Very High" EUI rating, making it unacceptable according to these standards. For the variable of Relative Compactness (RC), which is a continuous variable, the resulting graphs take on different shapes but convey a similar thematic message. Specifically, higher values of RC are linked to increased EUI rates, further complicating the quest for compliance with Polish regulations. Height (H) exhibits behavior similar to that of RC, in that greater height correlates with elevated EUI levels. In contrast, variables such as Characteristic Length (CL), Rotation (R), and Volume (V) appear to have negligible impact on EUI rates according to the analyzed data.



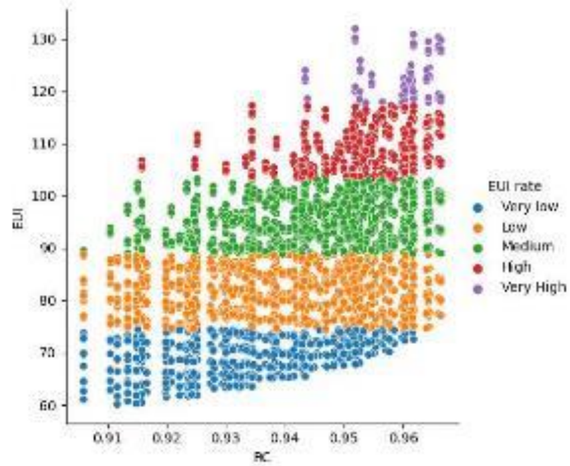
a. WWR and government category



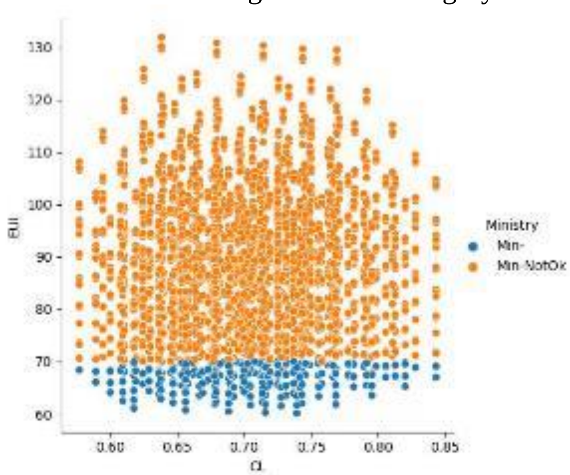
b. WWR and EUI rate



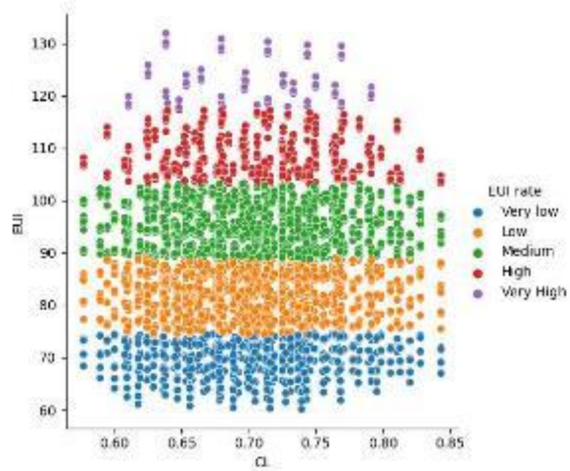
c. RC and government category



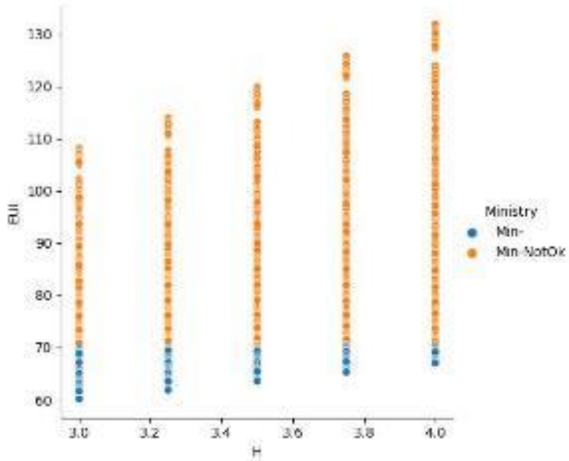
d. RC and EUI rate



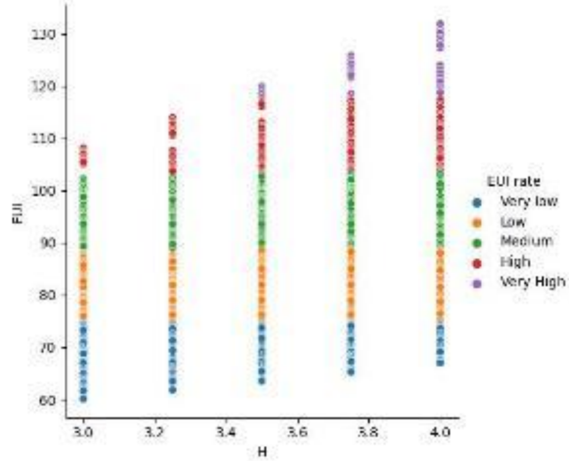
e. CL and government category



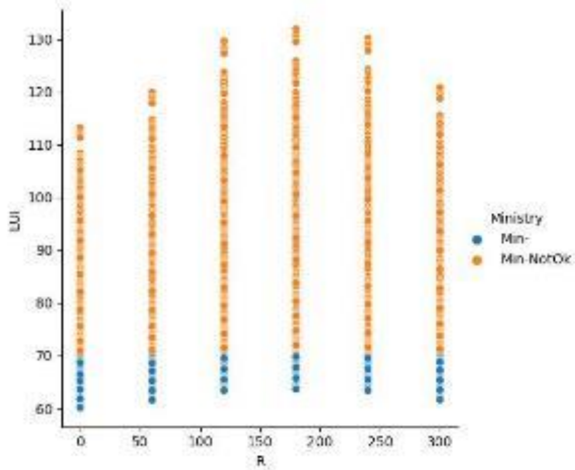
f. CL and EUI rate



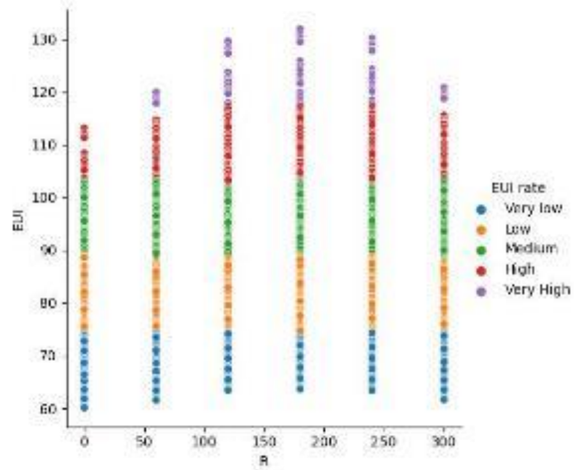
g. H and government category



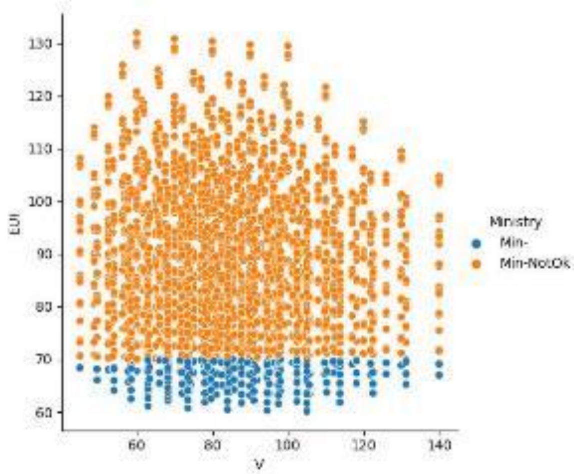
h. H and EUI rate



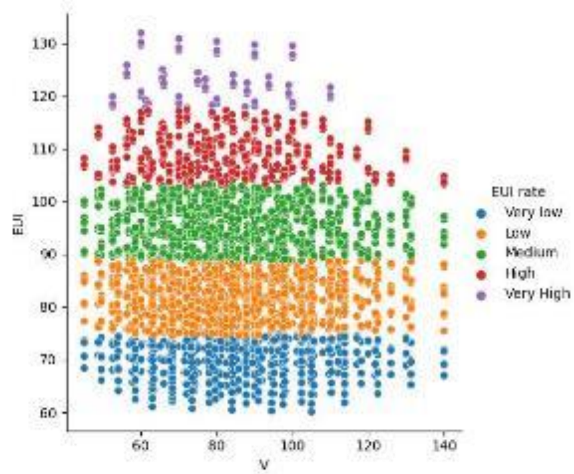
i. R and government category



j. R and EUI rate



k. V and government category



l. V and EUI rate

Figure 6.10. Features and governmental categories/EUI rate

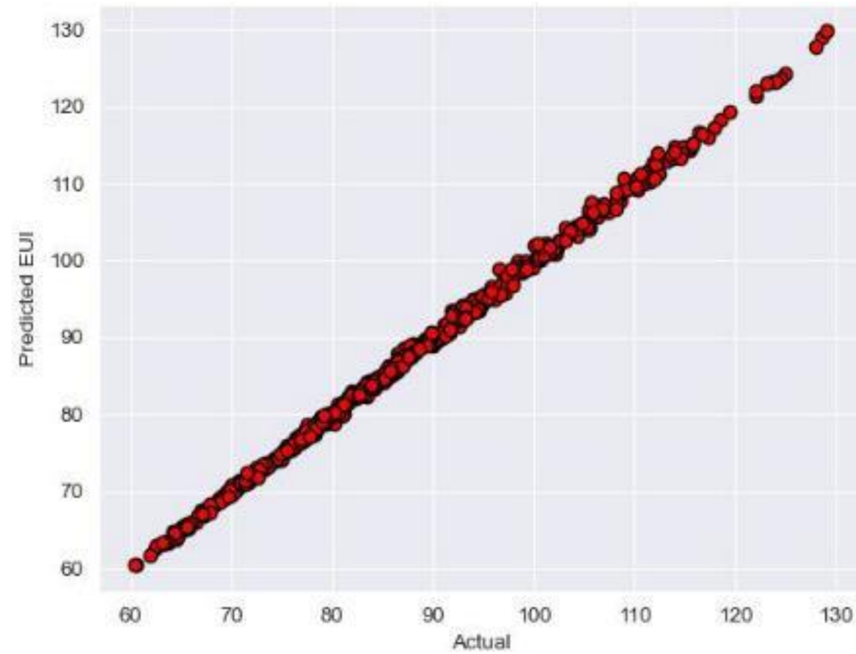
6.4. ML Deployment

Upon completing the data exploration phase and gaining insights into how various input variables affect Energy Use Intensity (EUI) as the output variable, the next logical step involves deploying a range of basic machine learning models. The goal is to compare their performance metrics, thereby gaining preliminary insights into which model holds the most promise as a potential final choice for more rigorous analysis. In the initial modeling phase, all 10 features were included in the various models tested. While it may appear redundant to include all 10, especially given that some have been observed to exert a negligible impact on EUI, this approach serves a specific purpose. By incorporating all features, the complexity of the question being posed to the model is fully maintained. In this way, the initial analysis can serve as a baseline for further refinement and feature selection. For this stage, a comprehensive testing strategy was employed using the lazy regressor, which facilitated the evaluation of almost all available regression algorithms. According to the results presented in Table 6.3, several models exhibited an R2-score exceeding 0.90. This suggests that these high-performing models are strong candidates for more detailed analyses in subsequent stages of the research process.

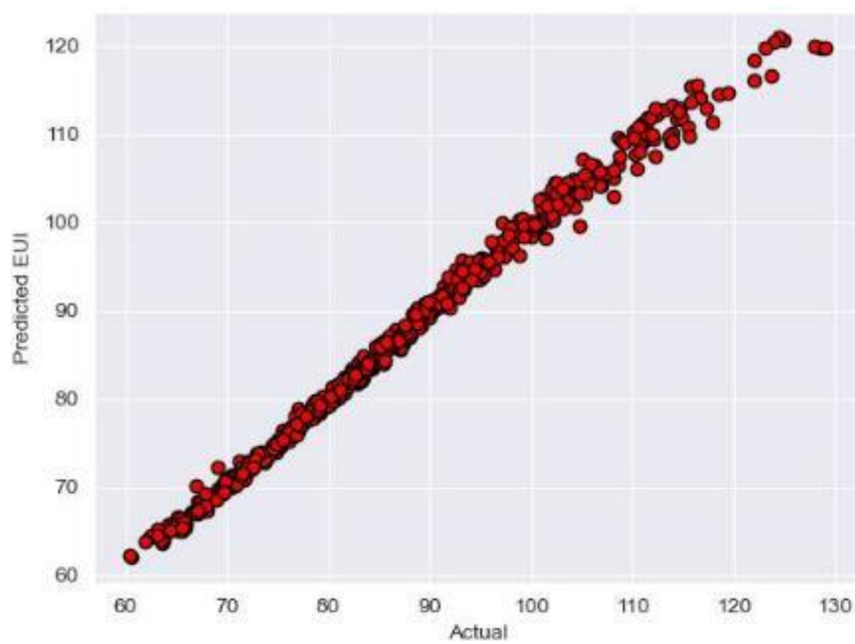
Table 6.3. Top 10 basic regression model comparison

Model	R-Suqared	RMSE	MSE	MAE	Time Taken
LGBM Regressor	0.99	0.53	0.28	0.40	0.18
SVR	0.99	1.40	1.97	0.85	0.43
ANN	0.97	2.29	5.28	1.34	0.85
Random Forest Regressor	0.92	3.89	15.20	2.98	0.64
Elastic Net	0.86	5.33	18.85	3.10	0.01
Gamma Regressor	0.82	6.04	19.75	3.35	0.01
Tweedie Regressor	0.81	6.09	19.90	4.10	0.01
LarsCV	0.81	6.10	21.06	4.49	0.02
MLP Regressor	0.78	6.70	22.00	5.03	1.90
Orthogonal Matching Pursuit	0.63	8.62	23.08	6.80	0.01

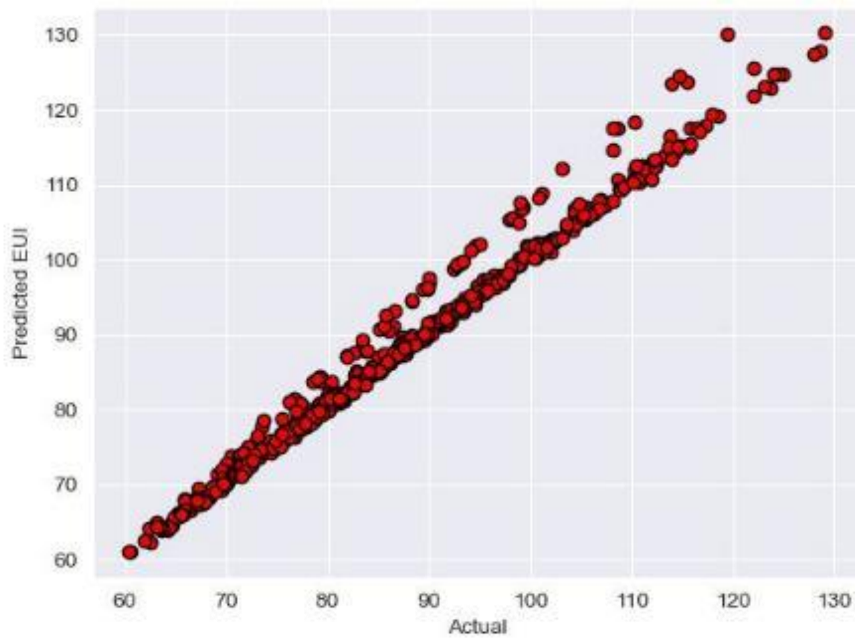
To attain a nuanced and thorough evaluation of how well the selected basic machine learning algorithms are performing on the dataset, a carefully designed investigative approach was implemented. Specifically, the top four algorithms (Table 6.3), which initially demonstrated the most promising performance metrics, were singled out for a more exhaustive analysis. One of the key elements of this analysis involved the creation of residual graphs, a valuable tool for visually scrutinizing the effectiveness of each model. In these residual graphs, the X-axis represents the actual output values from the dataset, while the Y-axis displays the corresponding predicted values produced by the model. By plotting these axes against each other, the residual differences between the actual and predicted values can be visualized. This graphical representation allows for a more granular examination of each model's performance, providing insights that go beyond what traditional numerical metrics can offer. In an ideal scenario, all the data points in the residual graph would closely align with a 45-degree diagonal line running through the graph. Such a pattern would indicate that the model is unbiased and has accurately captured the underlying data patterns. This would also suggest that the model is neither underfitting nor overfitting the data, thus representing a well-balanced predictive tool. Importantly, these residual graphs have been generated before the implementation of any optimization techniques or hyperparameter tuning on any of the models. These 'pre-tuned' visualizations provide an initial, unfiltered view of each model's strengths and weaknesses. This crucial, baseline performance data is illustrated comprehensively in Figure 6.11, serving as a cornerstone for subsequent evaluations and optimizations.



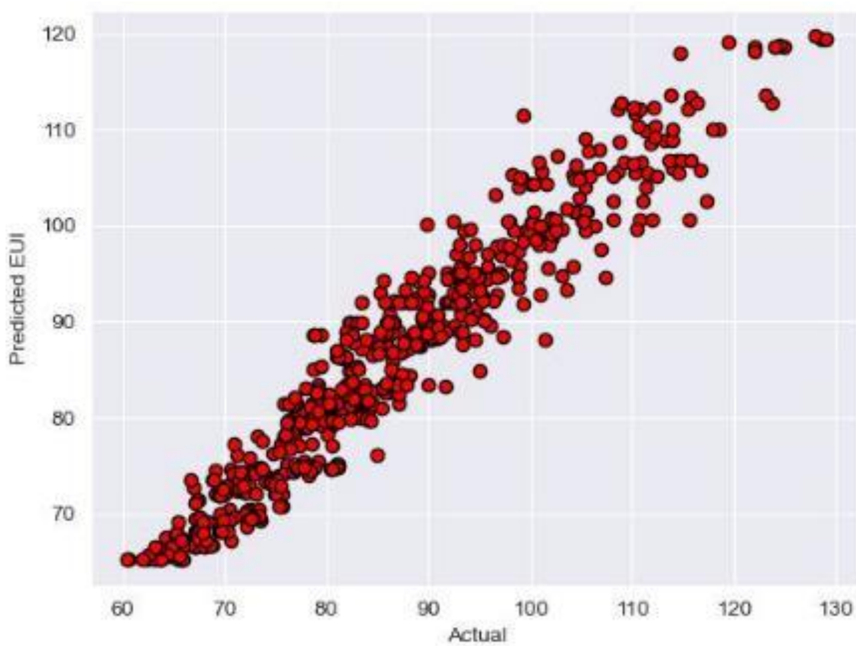
a. LGBM Regressor



b. Support Vector Machine



c. ANN



d. Random Forest

Figure 6.11. residual analysis of selected model with all input features

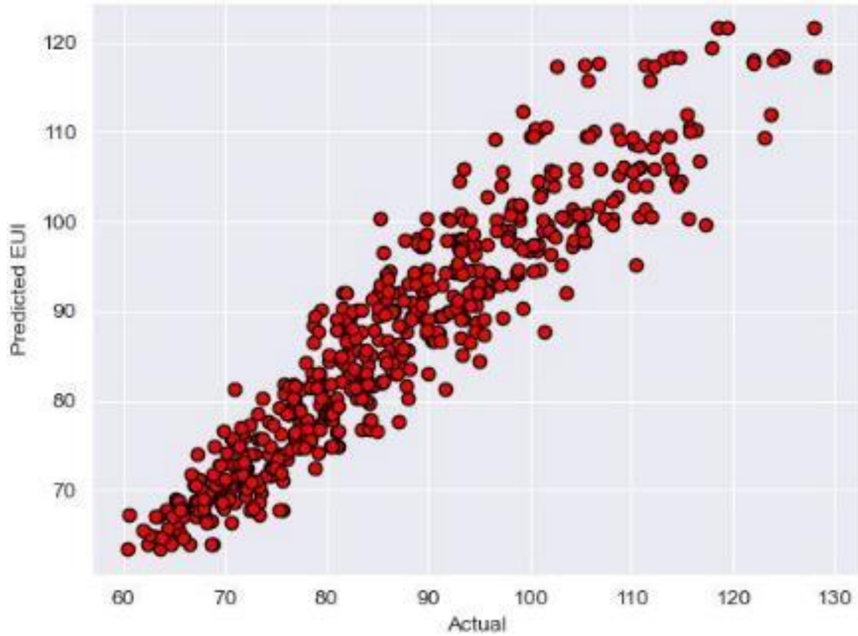
It's worth noting an additional, highly practical dimension to this approach. In real-world scenarios, acquiring comprehensive data sets about existing building clusters is often fraught with challenges. Whether due to bureaucratic hurdles, proprietary restrictions, or logistical constraints, complete data may not be readily available. This makes the scalability and efficiency of machine learning models especially crucial from a practical standpoint. By rigorously testing these models using only a narrow set of variables, namely Relative Compactness (RC) and Window to Wall Ratio (WWR), the research aims to address this very issue. Should the models

prove effective at predicting Energy Use Intensity (EUI) using just these two inputs, the implications would be significant and far-reaching. It would mean that reliable predictions could be made even in scenarios where only minimal information is accessible. In other words, this would open the door to more agile, efficient, and pragmatic energy assessments of buildings, bypassing the need for extensive and sometimes impractical data collection efforts.
This focused evaluation, therefore, not only serves to scrutinize the models' theoretical robustness and scalability but also has the potential to significantly impact their real-world applicability. If successful, this approach would provide a powerful tool that could be deployed in a variety of settings, contributing to both academic research and practical solutions for energy management. Such an advancement would be particularly advantageous for policy makers, urban planners, and sustainability experts who often operate under constraints of limited data. It could potentially revolutionize how Energy Use Intensity is predicted and managed, thereby making a substantive contribution to the field.
The findings from the exercise of employing merely RC (Relative Compactness) and WWR (Window to Wall Ratio) as the lone input variables to estimate EUI (Energy Use Intensity) are detailed in Table 6.4. A notable change in the performance metrics is evident when comparing these outcomes to previous models that utilized a broader set of features. Specifically, the maximum R2-score experienced a reduction, dropping from an impressive 0.99 to a lower value of 0.87. Concurrently, the Root Mean Square Error (RMSE) demonstrated an increase, escalating from a previous low of 0.53 to a more elevated figure of 4.86. Such alterations in the performance metrics were, in fact, anticipated, given the reduction in the number of utilized features from 10 to a mere 2. This made the task of predicting EUI significantly more challenging than in the initial models, where a comprehensive feature set was used. The increased difficulty is reflected in the diminished R2-score and the heightened RMSE, underscoring the complexities of accurately predicting EUI with a restricted set of variables.

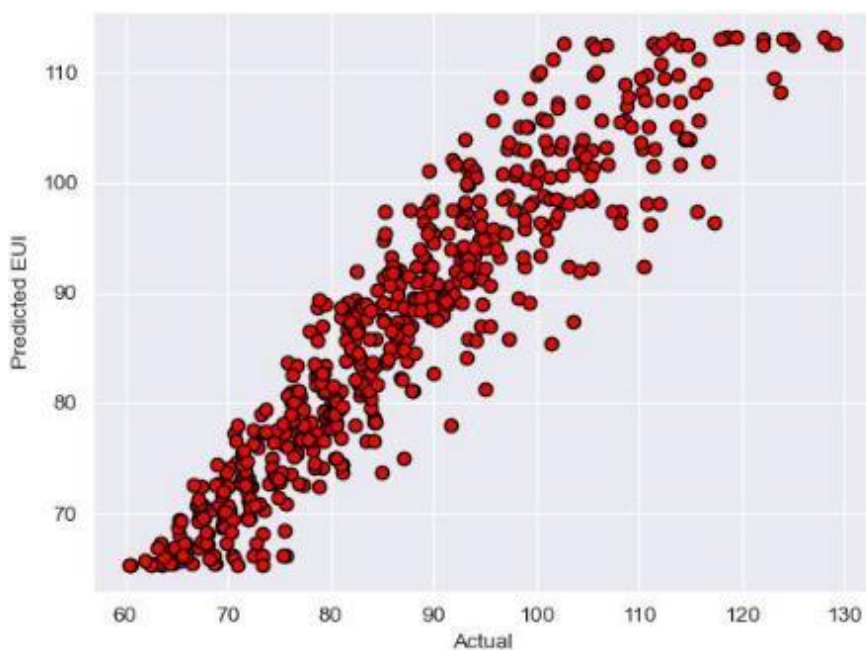
Table 6.4. Top 10 basic regression model comparison with RC and WWR

Model	R-Suqared	RMSE	MSE	MAE	Time Taken
LGBM Regressor	0.87	4.86	24.02	3.88	0.13
Random Forest Regressor	0.87	5.09	25.91	3.94	0.16
SVR	0.86	5.24	27.47	3.92	0.44
ANN	0.84	5.61	31.52	4.24	1.20
Gaussian Process Regressor	0.83	5.70	28.36	4.02	0.46
K Neighbors Regressor	0.83	5.75	28.57	4.10	0.02
Poisson Regressor	0.82	5.82	29.46	4.16	0.01
LarsCV	0.82	5.85	31.53	4.23	0.01
Bayesian Ridge	0.82	5.90	31.56	4.27	0.01
SGD Regressor	0.82	5.93	31.60	4.30	0.01

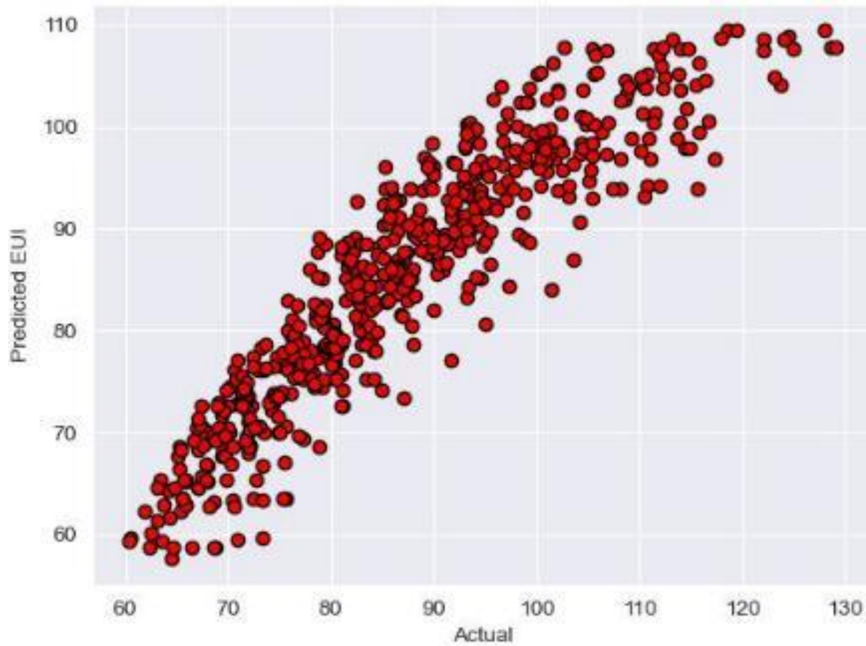
In a manner analogous to the previous scenario involving all features, residual analysis was once more employed to shed light on the performance of the model pre-tuning, this time focusing solely on RC (Relative Compactness) and WWR (Window to Wall Ratio) as the input features (Fig. 6.12). This subsequent analysis corroborated the observed decline in the model's overall accuracy and performance. However, it is worth highlighting that the magnitude of this decline was not excessively steep, especially when one considers the substantial reduction in the number of contributing features. This somewhat restrained downturn in performance metrics implies that the model's scalability remains relatively robust, even when the feature set is drastically curtailed. Furthermore, it offers an encouraging indicator that, post-tuning, there exists a high likelihood of achieving results that are more or less commensurate with those obtained using a more extensive set of features. Therefore, this suggests that the models are promisingly adaptable and could yield reliable predictions even when operating on a limited set of variables.



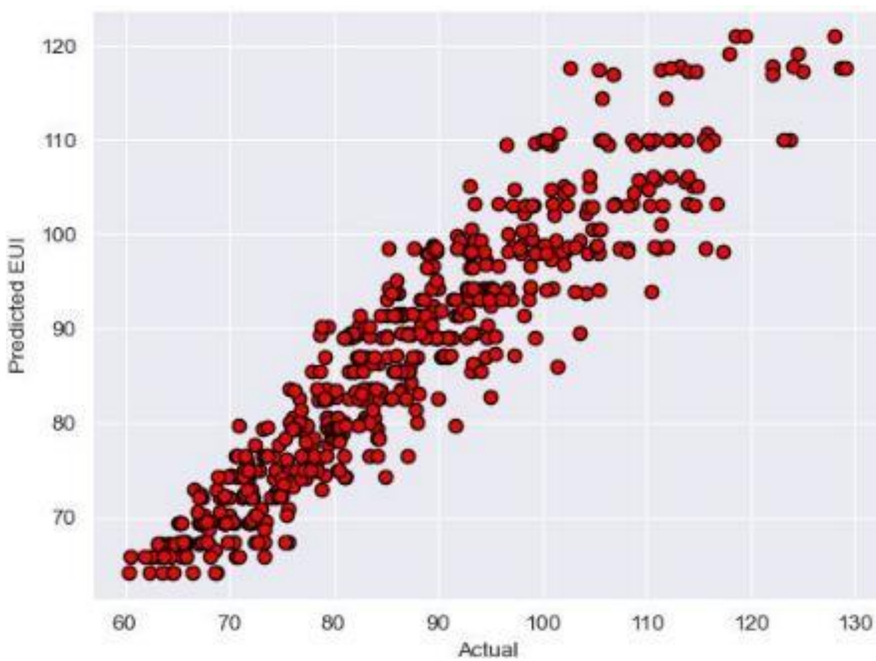
a. LGBM Regressor



b. Support Vector Machine



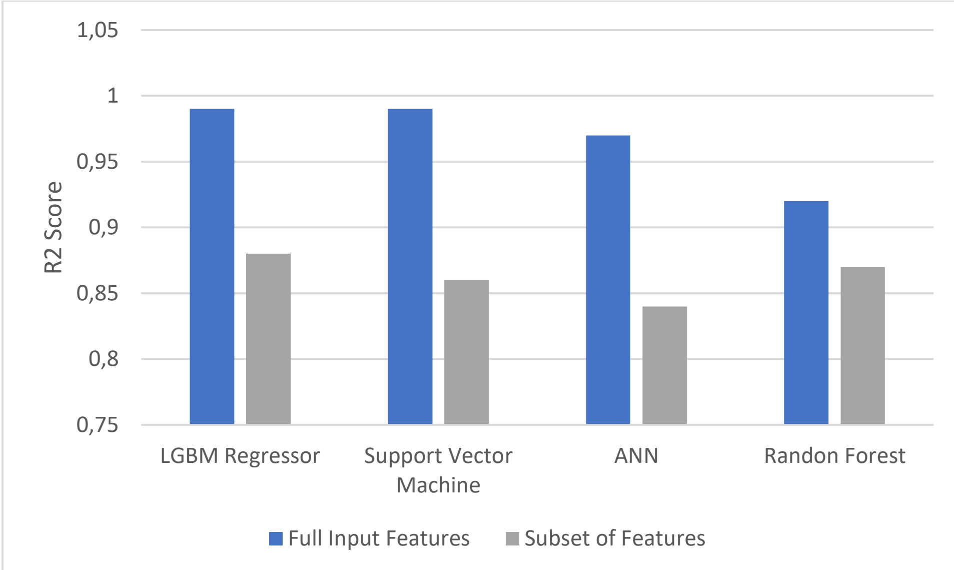
c. ANN



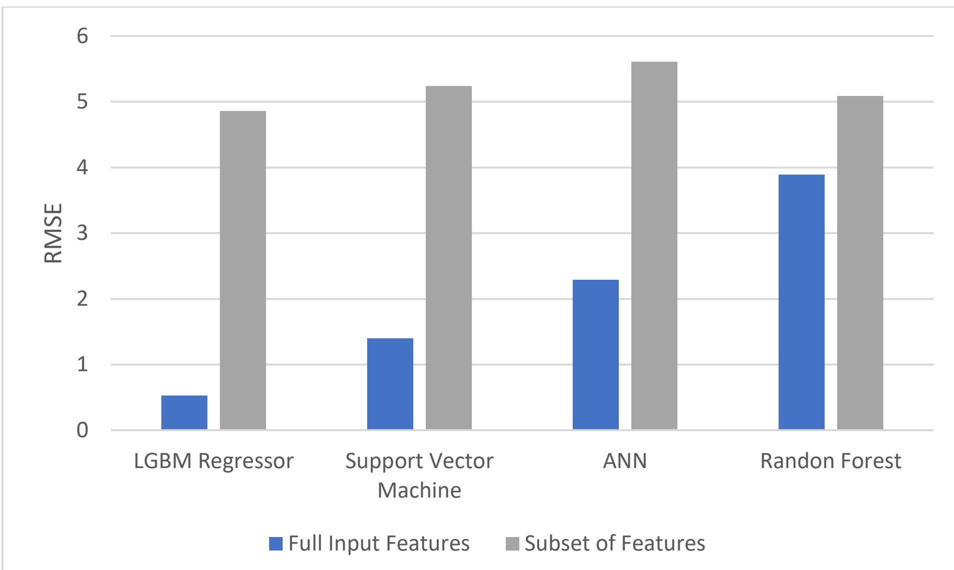
d. Random Forest

Figure 6.12. residual analysis of selected model with all input features

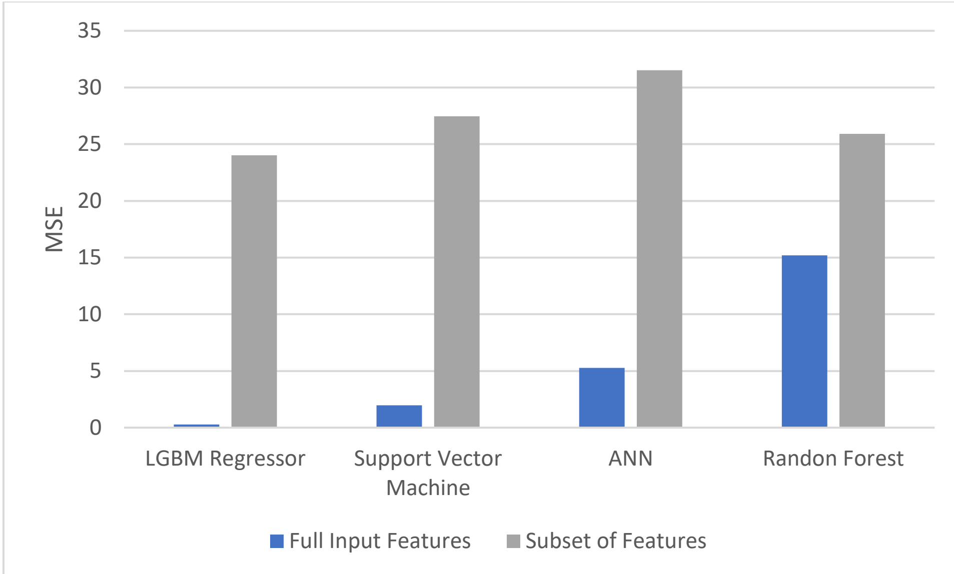
Before embarking on the model-tuning phase, it's advantageous to gain a more nuanced understanding of how the dimensionality reduction from 10 to just 2 input features affects the model's performance. To facilitate this, a visual comparison of the residuals for each model configuration—with 10 features as opposed to just 2—is recommended (Fig 6.13). By juxtaposing the residuals in this manner, one can easily discern the extent to which the model's predictive ability is impacted by the reduced number of variables. This step serves not only as a preliminary evaluation but also as a foundation for subsequent tuning processes, by highlighting the areas where improvement is needed and establishing a baseline against which the effectiveness of later adjustments can be measured.



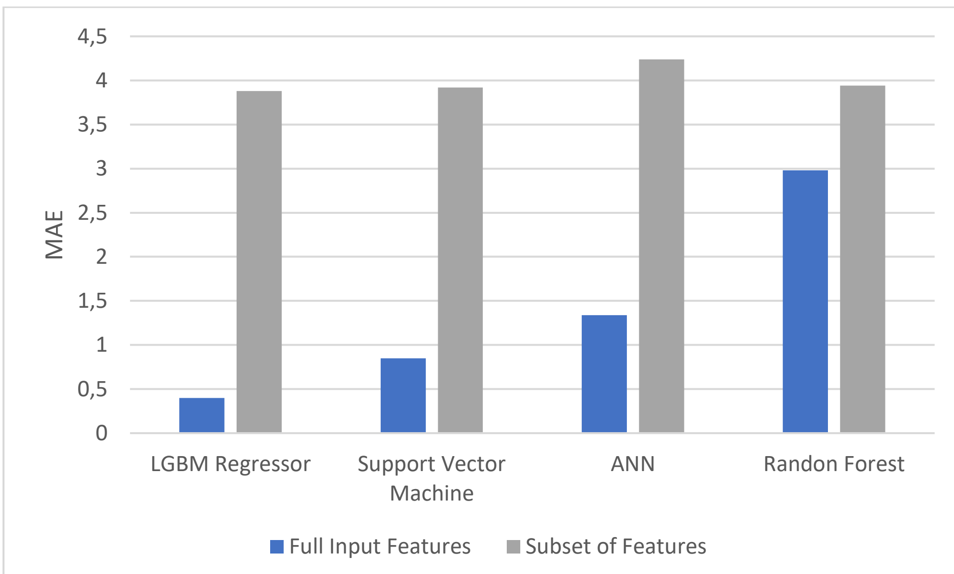
a. R2-Score



b. RMSE



c. MSE



d. MAE

Figure 6.13. Comparison of ML models with full and subset of input features

As can be observed from the comparative analysis, the performance degradation was least severe for the Random Forest algorithm, while Support Vector Regression (SVR) and Light Gradient Boosting Machine Regression (LGBMR) experienced the most significant declines in efficacy. Interestingly, Artificial Neural Networks (ANN) exhibited moderate variations in performance metrics such as R-squared score and error rates. In light of this, ANN emerges as a balanced choice, offering a relatively stable level of accuracy and error minimization irrespective of whether all features are considered or only a subset of two. This makes ANN a potentially optimal choice for modeling, given its demonstrated resilience to the loss of feature dimensions, and could serve as a benchmark for future comparative studies.

6.5. Model Tuning

In the first round of model deployment, which utilized all available features, an R-squared score of 0.99 was achieved, indicating a near-perfect fit of the model to the data. Given this high level of accuracy, there's minimal room for improvement when it comes to reducing errors or further boosting the R-squared score. As a result, the primary focus pivots towards the more challenging but practical scenario of deploying the model using only two key input features. Opting to proceed with just two features serves a dual purpose. First, it tests the model's capability to still produce reasonably accurate predictions with far fewer variables, which can be critical for real-world applications where obtaining comprehensive data may be difficult or costly. Second, focusing on a model with fewer features naturally requires fewer computational resources, making it more scalable and easier to implement in various settings.

The overarching goal becomes to approximate, as closely as possible, the high level of performance achieved in the initial deployment with all features. In this way, the effort concentrates on maximizing the efficacy of a more streamlined and computationally efficient model. If this goal can be accomplished, it will affirm the model's robustness and versatility, thereby making it an especially valuable tool for real-world applications where data limitations often exist. In this part each of selected model were subjected to model tunning one by one as follows:

6.5.3. LightGBM Regressor

Fine-tuning a machine learning model like the LightGBM Regressor involves adjusting a series of hyperparameters to optimize the model's performance. In the case of the LightGBM Regressor, key hyperparameters include 'n_estimators', 'learning_rate', and 'max_depth'. These parameters can have a substantial impact on the predictive power, speed, and generalizability of the model. Below is a breakdown of each hyperparameter and the ranges that will be considered in the tuning process:

- n_estimators: [50, 100, 200, 300, 400]
  This hyperparameter controls the number of boosting rounds or trees to be built during the LightGBM training process. The idea is to find a trade-off between model performance and computational efficiency. Too many trees can lead to overfitting, whereas too few might result in underperformance.

- learning_rate: [0.1, 0.01, 0.001, 0.0001, 0.00001]
  The learning rate or "shrinkage" is a factor by which to scale the contribution of each tree as it is added to the model. A high learning rate could allow the model to learn quickly but may risk overshooting the optimal solution. On the other hand, a low learning rate will make the model learn slowly, potentially requiring many trees to achieve good performance, but it's often more precise.

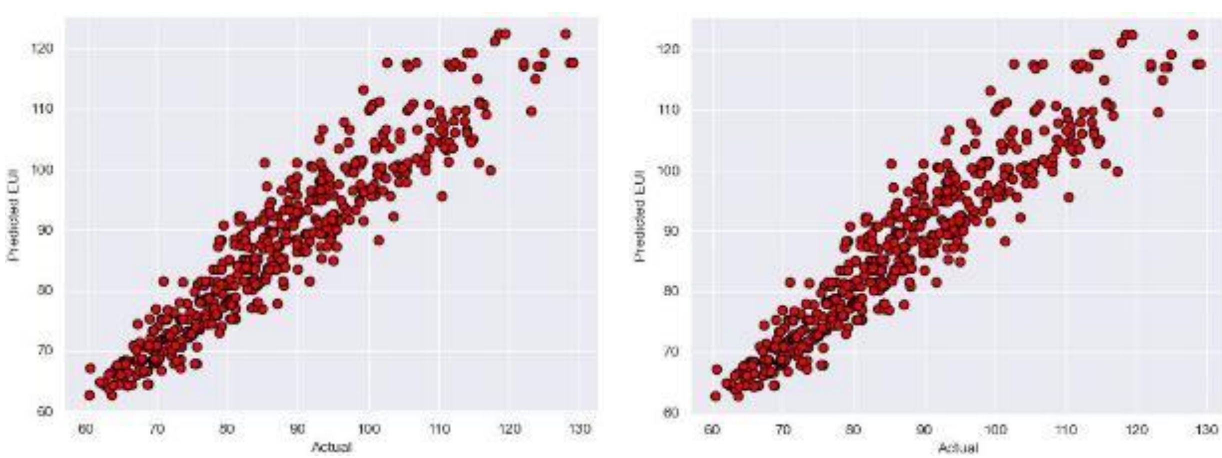
- max_depth: [5, 10, 20, 30, 40]

This parameter sets the maximum depth of each decision tree used in the boosting process. A deeper tree captures more complexity but is also more likely to overfit, especially when the depth is too high. A shallower tree might not capture the underlying patterns in the data well but is computationally less expensive.

The tuning process involves running the model through a combination of these hyperparameter values, typically using techniques like grid search or random search to systematically explore the hyperparameter space. By doing so, we aim to identify the combination that yields the best performance based on a chosen evaluation metric, such as R-squared or Root Mean Square Error (RMSE). After the optimal set of hyperparameters has been identified, the model can be retrained using these settings, ideally improving both its predictive accuracy and computational efficiency. This fine-tuned model is then more likely to perform well on unseen data, making it a valuable tool for various practical applications. Final selection of hyper-parameters are as follows: 'learning_rate': 0.1, 'max_depth': 5, 'n_estimators': 300. With the aforementioned fine-tuned hyper-parameters, the performance of the model improved (Table 6.5, Fig. 6.14).

Table 6.5. Performance of fine-tuned LightGBM Regressor

Model	R2-Score	RSME	MSE	MAE
Basic LightGBM Regressor	0.87	4.86	24.02	3.88
Fine-tuned LightGBM Regressor	0.88	4.82	23.27	3.87



a. Fine-tuned LightGBM Regressor b. Basic LightGBM Regressor

Figure 6.14. Fine-tuned and basic LightGBM Regressor performance

The performance of the model has evidently improved following the fine-tuning process, which is an encouraging development. However, it's important to note that the magnitude of these improvements is not uniformly significant across all evaluation metrics. This underscores the importance of considering multiple performance indicators when assessing the true efficacy of model optimization. While fine-tuning may have moved the needle in a favorable direction, the incremental gains in some metrics might not be dramatic enough to suggest a transformational change in the model's predictive power. The subtle improvements post-tuning could result from the model already being relatively well-calibrated before the optimization process. Alternatively, it could indicate that we are approaching the limits of what can be achieved with this particular algorithm and dataset. Therefore, while any improvement is generally a positive outcome, it's crucial to weigh these gains against the computational cost and complexity involved in the fine-tuning process.

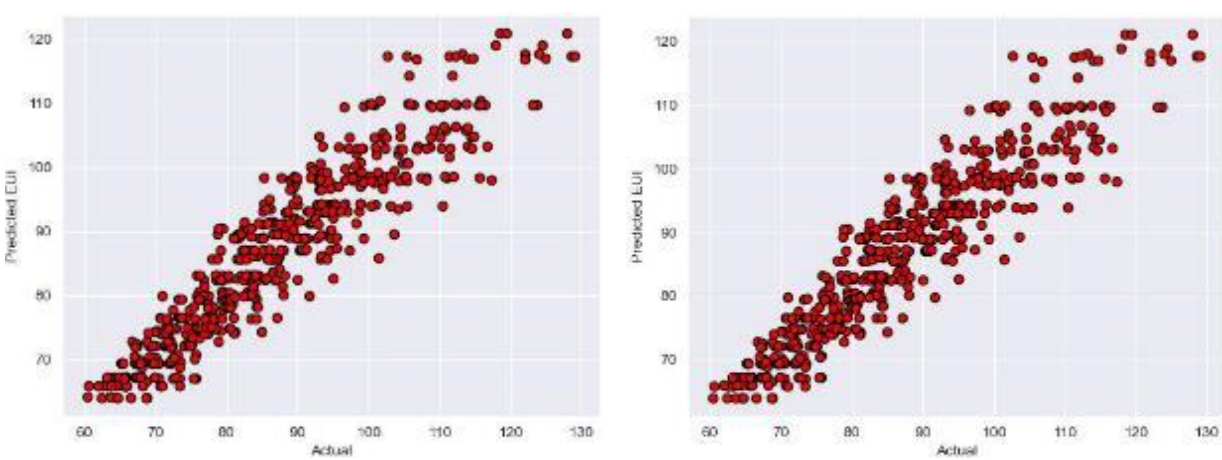
6.5.4. Random Forest

In this section, we delve into the results obtained from fine-tuning the Random Forest model. For the purpose of optimization, the hyperparameters considered are 'n_estimators' with values [50, 100, 500, 1000, 2000] and 'max_depth' with values [2, 3, 5, 7, 9]. Upon completion of the fine-tuning process, the optimal hyperparameters emerged as 'max_depth': 5 and 'n_estimators': 100 (Table 6.6, Fig. 6.15). However, it's worth noting that, akin to the experience with fine-tuning the LightGBM Regressor, the degree of improvement in the Random Forest model's performance did not meet our initial expectations for significance. The results, while improved, have not undergone a dramatic transformation, which might lead one to question the real-world impact of these incremental gains.

The lack of significant improvement after fine-tuning might suggest that the original model was already reasonably well-calibrated or that we have reached the limit of what this algorithm can deliver with the available data. As with the LightGBM Regressor, these marginal gains need to be carefully assessed in the context of computational expense and the complexity introduced by the tuning process.

Table 6.6. Performance of fine-tuned random forest

Model	R2-Score	RSME	MSE	MAE
Basic random forest	0.87	5.09	25.91	3.94
Fine-tuned random forest	0.87	5.08	25.85	3.93



a. Fine-tuned random forest b. Basic random forest

Figure 6.15. Fine-tuned and basic random forest performance

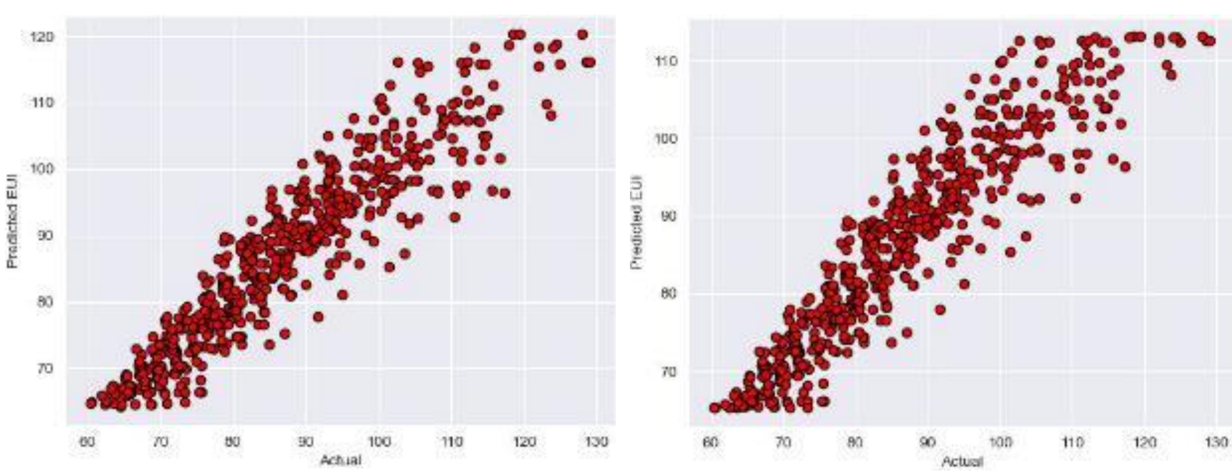
6.5.5. Support Vector Machine

In this part, we delve into the results of the fine-tuning process applied to the Support Vector Machine (SVM) regression model. The hyperparameters targeted for optimization were 'C', with potential values in the set [0.1, 1, 10, 100]; 'gamma', with options among [0.1, 1, 10, 100]; and 'epsilon', with candidates [0.01, 0.1, 1]. After conducting the hyperparameter search, the optimal configuration for the model was identified as C=100, epsilon=1, and gamma=1. Each of these hyperparameters plays a pivotal role in the SVM model's performance. The 'C' parameter acts as a regularization term and controls the trade-off between a low error rate on the training set and a simplified model to prevent overfitting. A higher 'C' value tends to generate a more intricate model at the risk of overfitting. 'Gamma' serves to influence the reach of individual training instances, affecting the shape of the hyperplane. A high 'gamma' value can make the model more complex, but it also increases the risk of overfitting. Lastly, 'epsilon' determines the acceptable range within which deviations between predicted and actual values are allowed without incurring any penalty.

Upon fine-tuning, the performance metrics displayed marginal changes. The R2 score remained constant, showing that the model's explanatory power did not improve. The RMSE (Root Mean Square Error) showed a modest reduction from 5.24 to 5.17. Similarly, the MSE (Mean Square Error) experienced a minor decrease from 27.47 to 26.82. The MAE (Mean Absolute Error) also fell from 3.92 to 3.90 (Table 6.7 and Fig 6.16). While these changes indicate some improvement, the level of enhancement was not particularly significant. This outcome aligns with the experience from fine-tuning other algorithms like Random Forest and LightGBM Regressor. Although the model was fine-tuned, the improvements were marginal and did not substantially elevate the model's overall performance.

Table 6.7. Performance of fine-tuned SVM

Model	R2-Score	RSME	MSE	MAE
Basic random forest	0.86	5.24	27.47	3.92
Fine-tuned random forest	0.86	5.17	26.82	3.90

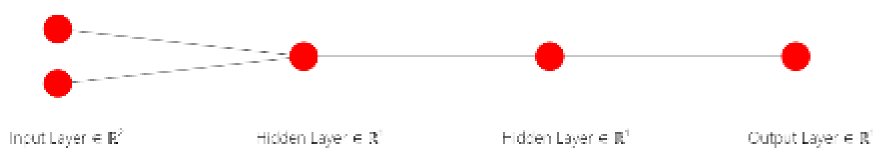


a. Fine-tuned SVM b. Basic SVM

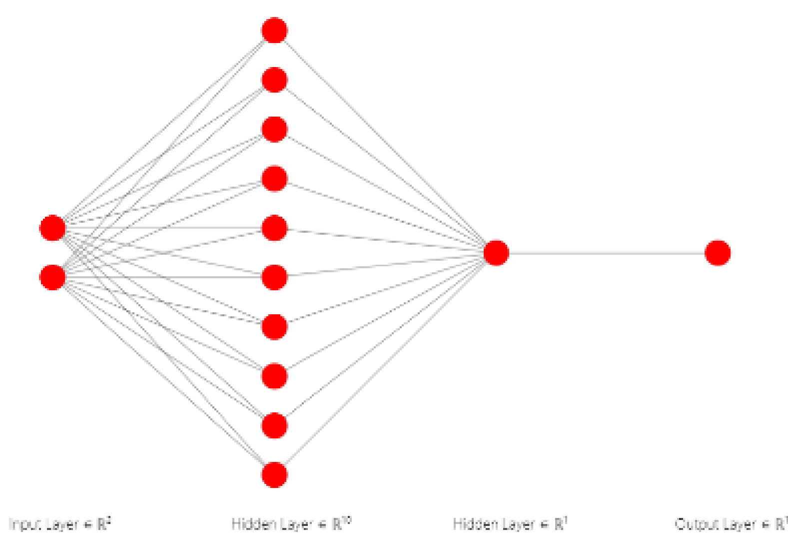
Figure 6.16. Fine-tuned and basic SVM performance

6.5.6. ANN

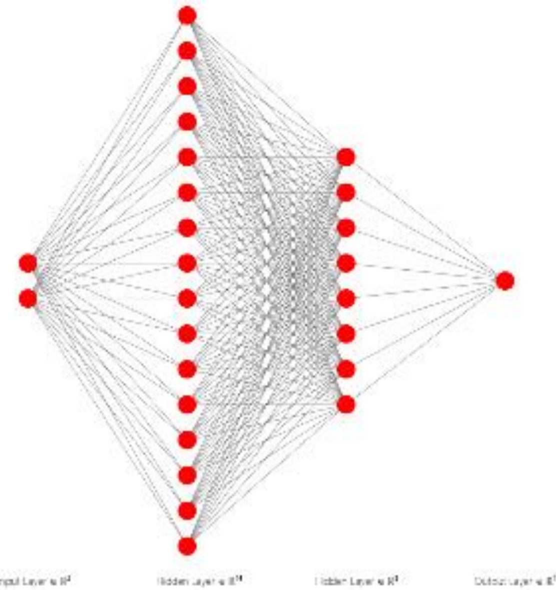
In this section, we explore the intricate procedure of fine-tuning the Artificial Neural Network (ANN) model. It's crucial to note that the tuning process of an ANN model can be broadly divided into two major components: architectural refinement and hyperparameter optimization. Both of these aspects play pivotal roles in enhancing the model's performance and should be systematically addressed to achieve optimal results. The first stage in the tuning process involves improving the architecture of the neural network. Architectural refinement essentially entails modifying the network's structure, including aspects like the number of layers, the number of neurons in each layer, and the types of activation functions used. These changes can significantly influence the model's capacity to capture complex relationships in the data and thus can drastically affect the overall performance. Once an effective architecture is established, the second stage involves hyperparameter optimization. This could involve tuning parameters such as the learning rate, batch size, and the choice of optimization algorithm, among others. The aim here is to fine-tune these variables to enable faster and more stable convergence during training, thereby enhancing the model's predictive accuracy. In ANN model in order to improve the architecture and then hyper-parameters one variables should be used as the metric for optimization and according to the other models' results in this context MAE sounds an appropriate metric since the range of MAE variation for tuned model is very specific between 3.87 and 3.93. Therefore, from now on only this meter and its graph will be used for evaluating the model. Through 4 stages of architecture improvement, the number of layers and neurons in each layer process of model optimization was conducted. Considering Fig. 6.17.



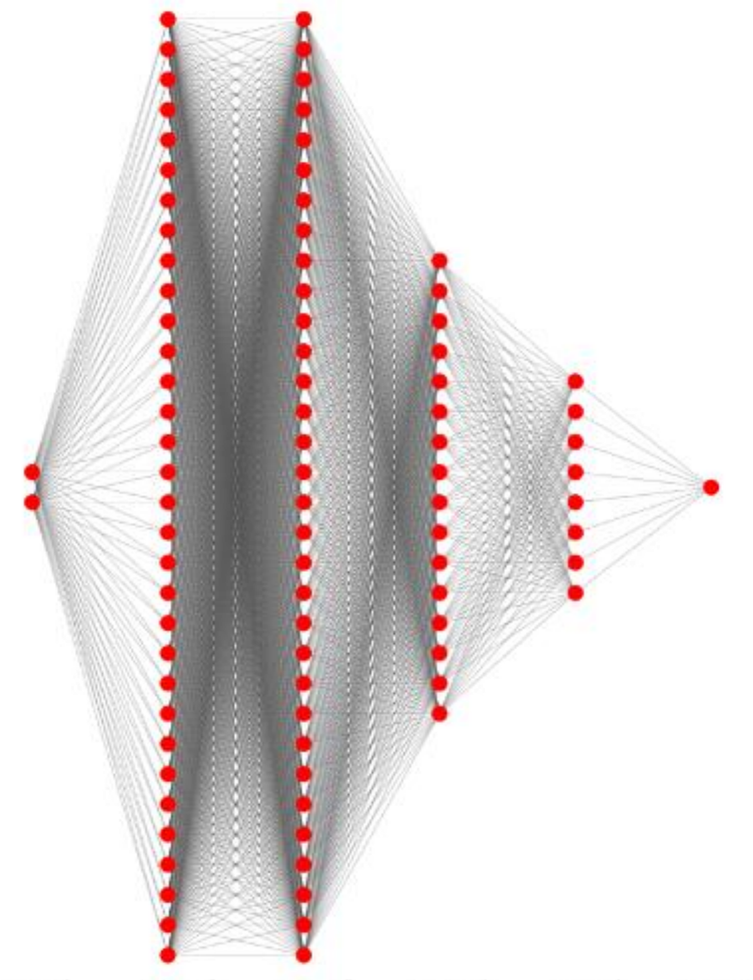
a. Base architecture with 2 hidden layer and one neuron in each layer



b. 2^{nd} architecture with 2 hidden layers with 10 and 1 neurons



c. 3^{rd} architecture with 2 hidden layers with 16 and 8 neurons



d. 4^{th} architecture with 4 hidden layers 32,32,16,8 neurons

Figure 6.17. Architecture improvement of deployed ANN algorithm

In the process of refining our model, the primary focus was on improving the Mean Absolute Error (MAE) metric. As previously discussed, we were able to reduce the MAE from an initial value of 4.24 down to 3.90, marking a notable enhancement in the model's performance (Table 6.8). While MAE served as the cornerstone for our optimization efforts, other performance metrics also exhibited slight improvements, though their significance was relatively negligible in comparison. By achieving a lower MAE, the model has become more accurate in its predictions, thereby increasing its utility and reliability. It's worth mentioning that even though the other metrics showed only modest gains, they still contribute to a more robust and trustworthy model. Therefore, while the MAE was the focal point of our improvement strategy, the subtle improvements in other metrics should not be entirely discounted.

Table 6.8. ANN architecture improvement

Model	R2-Score	RMSE	MSE	MAE
Based ANN Model	0.84	5.61	31.52	4.24
2nd architecture	0.84	5.32	30.65	4.20
3rd architecture	0.85	5.26	29.53	4.01
4th architecture	0.85	5.17	27.37	3.90

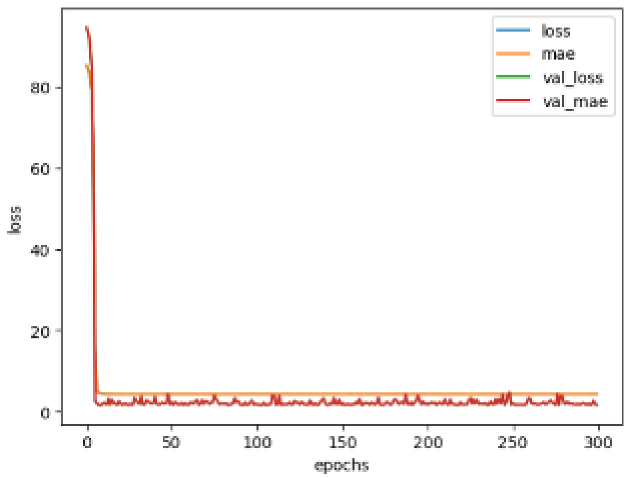
The initial architecture of the model was improved before attention was turned to hyperparameter tuning to enhance performance further. Various optimizer algorithms and learning rates were tested rigorously to identify the most effective combination. Upon determining the optimal settings, focus was shifted to addressing the issue of overfitting, a common concern in machine learning models. To mitigate this issue, two key strategies were employed: input normalization and model regularization. A learning rate of 0.004 was selected for the final model, and the Adam optimizer was utilized. Adam is known for dynamically adjusting the learning rate during training, offering faster convergence and reducing the sensitivity to the initial learning rate setting.

To counteract overfitting, two dropout layers with an intensity of 0.03 were incorporated, positioned strategically at the first and third hidden layers of the neural network. Dropout layers function by randomly setting a fraction of input units to zero during each training iteration, thus enhancing the model's ability to generalize to unseen data. This technique led to the random deactivation of about 3% of the neurons in these layers during each training iteration. In addition, the Standard Scaler was utilized to normalize the input data. This step is crucial as it ensures that all input features are on the same scale, thereby improving the learning efficacy of the model. By employing these techniques—the Adam optimizer for efficient and adaptive learning, dropout layers for robust generalization, and Standard Scaler for feature normalization—a more accurate, robust, and well-adapted model was achieved. Overall, this comprehensive approach to hyperparameter tuning and regularization led to the construction of a significantly more robust and accurate model (Table 6.9).

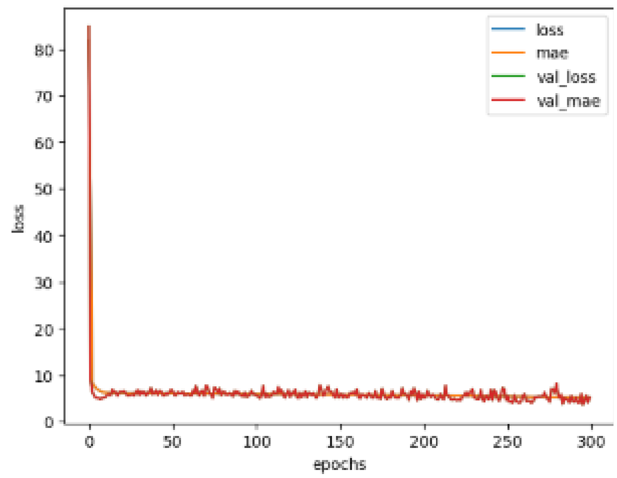
Table 6.9. ANN model improvement

Model	R2-Score	RMSE	MSE	MAE
Final ANN Architecture	0.85	5.17	27.37	3.90
Final fine-tuned model	0.88	4.20	24.47	3.70

In the final assessment of the model, the test and train losses were compared to evaluate the presence of any overfitting prior to the introduction of regularization techniques. Upon examination of Figure 6.18. a, a slight disparity between the test and train loss was observed. Although this difference was marginal, it indicated an opportunity for further refinement. Subsequently, the incorporation of dropout layers led to a convergence of test and train loss values, bringing them closer together. This improved state is visibly captured in Figure 6.16. b, confirming the efficacy of the regularization measures in mitigating overfitting.



a. Before dropout regularization



b. After dropout regularization

Figure 6.18. Impact of regularization on the performance of the final model

6.6. Final Model Testing

To evaluate the efficacy of the final model and the overall workflow, a real-life case study was conducted. The studied case was a room in a house in at Mickiewicza Street in Poznan. A 3D reconstruction of a room was generated through the proposed workflow, employing lidar-based software known as Magic Plan on an iPhone 14 Pro. The resulting 3D representation of the room is depicted in Figure 6.19, serving as a practical test of the model's performance and the workflow's applicability.

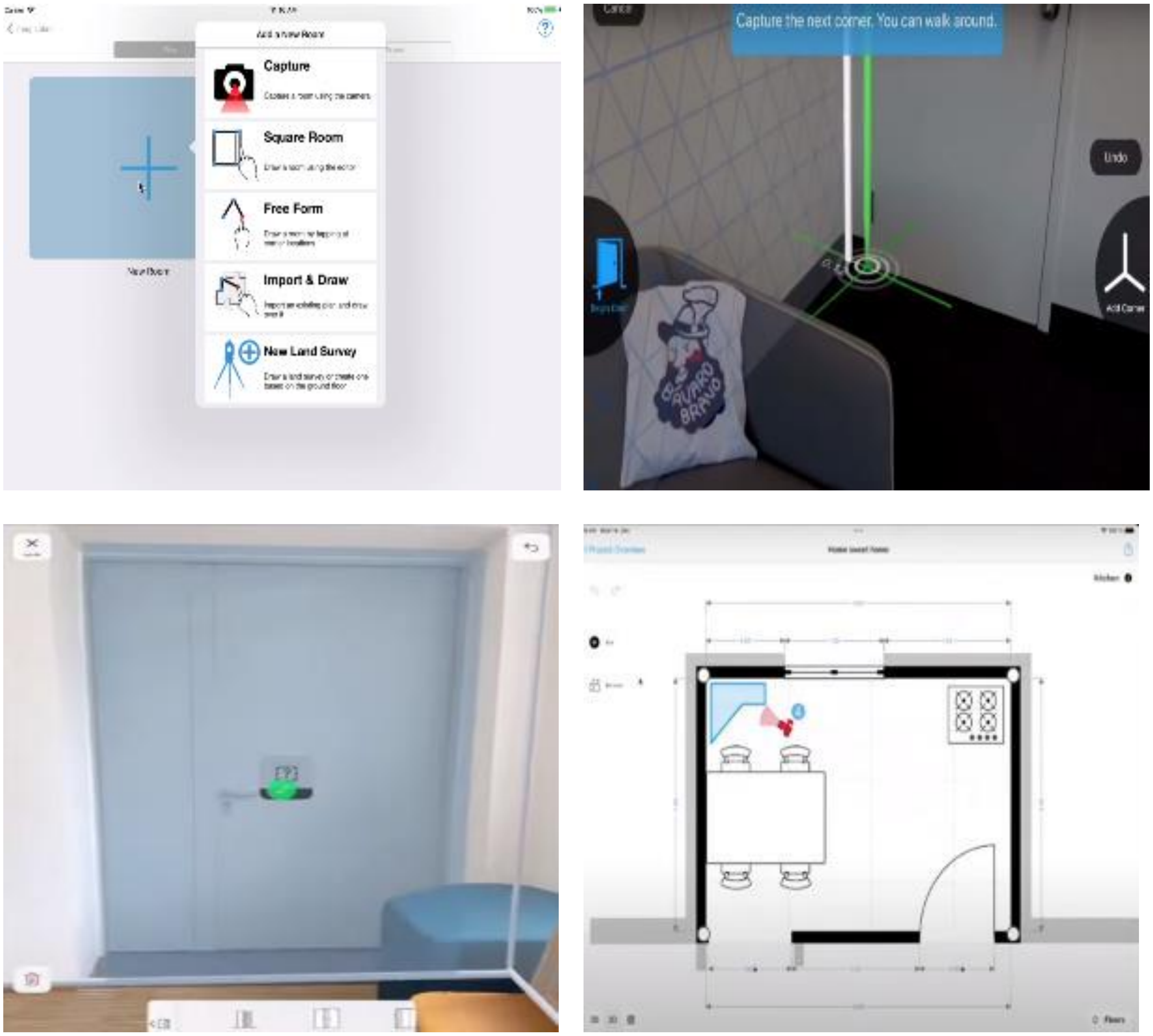


Figure 6.19. steps of 3D reconstruction of a room

Subsequently, the generated 3D model was augmented with scheduling and material information, along with details of the HVAC system, to produce an IDF file. This IDF file was then used as input for the EnergyPlus software, paired with the appropriate weather file, to simulate the Energy Use Intensity (EUI) of the room. The simulation yielded an EUI value of 81.34 KWh/m^2 y. This simulated EUI served as the baseline or "actual" value for further evaluations. When the Window-to-Wall Ratio (WWR) and the Roof Construction (RC) variables were input into the selected, fine-tuned machine learning algorithm, a predicted EUI value was generated for comparative analysis.
To gain deeper insights into the performance of various Artificial Neural Network (ANN) models, all were subjected to the same validation stage, the results of which are presented in Table 6.10. It is evident from the data that the final model's performance surpasses that of the base model substantially. Moreover, the absence of overfitting in the final model suggests that it could potentially deliver superior and more reliable performance on similar validation datasets.

Table 6.10. comparative analysis of actual value and predicted one with different models.

Model	Actual Value	Predicted value
Base ANN model	81.34	4.25
1st Architecture	81.34	17.69
2nd Architecture	81.34	58.65
3rd Architecture	81.34	41.53
4th Architecture	81.34	92.89
Tuned model	81.34	87.06
Final model	81.34	84.97

Chapter 7: Conclusion

7.1. Motivation & Significance

The growing urgency to combat climate change has rendered optimizing energy consumption in buildings a priority for global sustainable development. Buildings account for a significant portion of global energy use and are therefore pivotal in the ongoing battle against environmental degradation. Furthermore, the optimization of energy consumption doesn't only have environmental benefits; it also offers an economic advantage and can make living conditions more resilient against extreme weather events, which are becoming more frequent due to climate change. Energy optimization in buildings serves as a dual-purpose mechanism, not just mitigating the negative impacts of climate change but also preparing structures to withstand the very outcomes of global warming. For instance, buildings optimized for energy consumption can maintain temperature more efficiently, reducing the need for excessive heating or cooling. This is increasingly crucial as extreme weather patterns become more unpredictable, necessitating indoor environments that can maintain comfort and safety with minimum energy expenditure. Therefore, having an optimized energy consumption figure can pave the way for a future that is not just sustainable, but also resilient in the face of environmental challenges.

Achieving a net-zero future—a scenario where human activities emit no more carbon than the Earth can naturally absorb—is intrinsically tied to understanding and enhancing the energy performance of buildings. The traditional method to gain such insights has been through simulation-based energy audits. These audits are exhaustive studies that measure a building's energy usage and recommend changes to make it more efficient. However, while effective, this method has its drawbacks. It is a time-consuming process, often requiring specialized equipment and expertise. This makes it expensive and labor-intensive, restricting its scalability and widespread adoption. Given the urgent need to accelerate progress, these limitations underscore the necessity for alternative solutions that can deliver accurate insights into a building's energy performance without the burdens of a traditional audit. The drawbacks of simulation-based energy audits signal the importance of developing other methods, which can facilitate a faster, more cost-effective evaluation of building energy performance. This is where machine learning models, particularly optimized Artificial Neural Networks (ANNs), offer promising alternatives. These models can rapidly analyze multiple parameters that influence a building's energy consumption, delivering results that are just as accurate as traditional methods but in a fraction of the time and at a fraction of the cost.

Therefore, the application of machine learning in understanding building energy consumption isn't merely an academic exercise but a practical necessity, one that holds the key to unlocking a more sustainable, resilient future. By making the process of energy audit more efficient, economical, and accessible, machine learning models serve as invaluable tools in the global push toward a net-zero and climate-resilient future.

7.2. Robustness and innovation

To confront the urgent challenges related to building energy optimization and its impact on climate change, this thesis introduces a groundbreaking workflow as an alternative to traditional simulation-based energy audits. Utilizing contemporary technology, the methodology employs a

LiDAR-based application on smartphones to create 3D models of rooms or entire buildings. These 3D models serve as a foundation for extracting specific geometrical features that, in turn, become inputs for data-driven algorithms designed to estimate a building's energy performance. The core innovation of this work lies in its simplicity and efficiency. With only two geometrical-based features extracted from the 3D model, the proposed algorithm can offer a surprisingly accurate estimate of a building's future energy performance, even under the complex variables introduced by climate change. This offers a rapid, cost-effective, and scalable means to predict how a building will fare in terms of energy consumption in both current and future climatic conditions. Importantly, the level of uncertainty in these estimates decreases when considering present-day conditions, further solidifying the method's reliability and potential for widespread application.

The significance of this approach is multi-faceted. For one, it circumvents the laborious, time-consuming, and costly process associated with traditional energy audits. By doing so, it democratizes access to vital information that could expedite efforts to make buildings more energy-efficient and climate-resilient. The use of smartphone technology for data collection is particularly noteworthy, as it makes the workflow highly accessible. With smartphones being ubiquitous, the data acquisition stage becomes not just cheaper but also much quicker. Another remarkable aspect is the algorithm's ability to make accurate predictions with minimal input, minimizing the complexity traditionally associated with such audits. The geometrical-based features act as a streamlined yet effective set of variables that can predict a complex outcome—energy performance—thereby simplifying the usually intricate process of data collection and analysis. This makes the methodology ideal for quick assessments, perhaps even in settings where detailed, traditional audits are not feasible due to resource constraints.

Furthermore, the workflow is designed to be adaptable to future climate scenarios, a feature that is increasingly important as the impacts of climate change become more pronounced. The ability to assess a building's future energy performance under different climate conditions not only makes the model forward-looking but also invaluable for long-term planning. It serves as a tool for architects, urban planners, and policy-makers to make data-driven decisions that are not just responsive to current energy needs but are also sustainable in the long run.

7.3. Results

The results of this research provide compelling evidence of both the challenges and opportunities in using machine learning to predict building energy performance. Initial analysis of the generated dataset revealed that only 11% of cases fell within the acceptable range of Energy Use Intensity (EUI) as defined by the Polish Ministry of Energy. This serves to highlight the immense scope for improvement in energy consumption patterns in buildings, underscoring the critical need for advanced predictive models. Upon deploying various machine learning algorithms on the complete feature set, an array of algorithms displayed an R-squared score of 0.99. While this may seem promising, it is important to note that these results were achieved under the assumption that construction details remain constant across the dataset. Given that this is seldom the case in real-world scenarios, the high R-squared score, in this case, may not be fully generalizable.

In contrast, when algorithms were tested with just two input features, the R-squared score dipped to approximately 0.84. However, this apparent limitation was viewed as an opportunity: the lower score indicated a higher room for improvement and optimization, a potential that was thoroughly explored in subsequent stages of the research. The selection of Artificial Neural Networks (ANN) as the algorithm of choice emerged from a careful analysis of both training and test loss metrics. While ANNs initially yielded an R-squared score of 0.84, similar to other algorithms when restricted to just two features, further architectural improvements and hyperparameter tuning significantly improved this figure. The final architecture, fortified with an optimal combination of dropout layers and learning rates, achieved an impressive R-squared score of 0.88. In the context of only two feature inputs and the complexity of energy performance modeling, this result can be deemed as highly satisfactory.

This accomplishment is particularly noteworthy given that the algorithm was able to make fairly accurate energy performance predictions based solely on two geometric features extracted from 3D models generated using LiDAR technology. The efficiency of this approach not only offers a faster and cost-effective alternative to traditional simulation-based energy audits but also opens the door to scalable applications. This means that the method could feasibly be applied to a broad range of buildings with varying structural complexities and energy use patterns, offering a universally applicable tool for enhancing energy efficiency on a wider scale.

The strength and generalizability of the research were further validated through a rigorous real-life case study. In this practical scenario, a building's Energy Use Intensity (EUI) was first calculated using traditional simulation-based analysis as a baseline for comparative purposes. The data-driven model, designed and refined in this study, was then applied to forecast the building's EUI. Impressively, the model displayed a high degree of accuracy, closely mirroring the actual EUI value generated from simulation-based methods. This validation serves multiple purposes. Not only does it substantiate the model's reliability, but it also underscores its practical applicability and scalability. The fact that such a close match could be achieved using only two geometric features from the model bolsters its robustness. It essentially proves that the model doesn't require an exhaustive list of variables to make accurate predictions, emphasizing its efficiency and scalability. The real-world validation not only endorses the reliability of the model but also adds credence to its practical applicability across different scenarios. Given that the model could reach such a high level of approximation in a complex, real-world environment, it suggests that it could be effectively applied on a much larger scale. The empirical success in this case study adds a layer of confidence, showing that this data-driven approach can serve as a rapid and accurate alternative to traditional, more laborious simulation-based methods. The successful validation ultimately accentuates the transformative potential of the research. In a world where rapid and accurate evaluation of building energy performance is crucial for addressing the adverse impacts of climate change, the methodological innovations presented in this study represent a meaningful contribution. By demonstrating its robustness and scalability through real-world validation, the model has shown that it can be a cornerstone for future building energy assessments. This efficient, scalable, and empirically validated tool is not just an academic endeavor; it's a real-world solution to a pressing global challenge.

7.4. Implementations

The implementation of this research offers various practical applications that can significantly impact the field of building energy performance assessment. One of the most immediate uses of

this work is for building owners considering renovations or energy audits. Traditionally, gaining insights into a building's energy performance requires a comprehensive, and often costly, energy audit. However, the data-driven model developed in this research provides a quicker, more accessible initial assessment. This is particularly valuable for building owners who may need to understand how their building's energy usage compares to local or even national regulations before embarking on more intensive audits or renovations. In the academic sphere, the workflow presents a robust framework that can be reproduced by other researchers in different geographical regions. While the study itself was limited to a cluster of residential buildings in Poland, the methodology is designed in such a way that it can be adapted to different building types and climates. This scalability enhances the research's value, opening avenues for further validation and refinement of the model across diverse settings.
Moreover, the developed model could be integrated into governmental plans for monitoring the energy performance of existing buildings or even for the energy certification of new constructions. Traditional methods for such monitoring or certification often require substantial financial and time resources. The model offers a cost-effective, yet reliable, alternative that can streamline these processes, thereby accelerating the move towards more sustainable building practices. Another promising avenue for the implementation of this research is its potential incorporation into the emerging field of digital twins for cities. Digital twins are virtual replicas of physical systems, and they are increasingly being used for various forms of urban planning and management. The model developed in this research can add a semantic layer to these digital twins, providing real-time or predictive insights into the energy performance of buildings. This added layer of information can significantly enrich the capabilities of digital twins, making them more comprehensive tools for urban sustainability planning.
The results of this work have shown its potential, with a relatively high degree of accuracy in predicting energy performance based on limited input features. During validation with a real-life case study, the model demonstrated its robustness and further confirmed its potential for scalability. These outcomes not only underscore the research's immediate practical applications but also establish it as a strong foundation for future work in the field.

7.5. Limitations

One of the primary limitations of this research stems from computational constraints, which restricted the scope of testing to a single cluster of residential buildings in Poland. While the model demonstrated promising results within this localized setting, the question remains as to how effectively it can be generalized to different types of buildings, or even different geographical locations with varying climate conditions. The computational limitations essentially constrain the model's capacity for wider validation, and by extension, its immediate applicability on a broader scale.
Another significant limitation is the model's simplification of building construction details and energy systems. The study was conducted based on a single archetype that represents the cluster of residential buildings. This approach inherently glosses over the potential variations in construction materials, insulation levels, HVAC systems, and other factors that could significantly influence a building's energy performance. While this simplification made the problem more

tractable given the available resources, it raises questions about the model's capability to accurately predict energy performance in settings where these variables differ.
The third major limitation is the consideration of only one climate change scenario. Climate change is a complex phenomenon with multiple possible trajectories, each influenced by a variety of factors such as greenhouse gas emissions, technological advancements, and global policies. By focusing on just one climate change scenario, the study narrows its applicability. Real-world conditions could differ substantially, leading to divergent energy performance outcomes that the current model might not accurately predict.
Each of these limitations, whether they pertain to the scope of testing, the simplification of variables, or the narrow focus on a single climate change scenario, presents challenges for the generalizability and scalability of the model. While the research provides valuable insights and a promising framework for building energy performance assessment, these limitations suggest avenues for future work. Extending the model to incorporate more variables and scenarios, and testing it across broader contexts, would be essential steps for enhancing its robustness and reliability.

7.6. Future works

Future work could focus on several key areas to further refine and expand the utility of the developed model. One of the most critical aspects is increasing its scalability by incorporating more variables such as different building types, materials, and varying weather conditions. The current model was developed and tested on a specific cluster of residential buildings in Poland. While it showed robustness and predictive accuracy within that context, its generalizability to other building types and climatic conditions remains an open question. By broadening the dataset to include these variables, the model could become a more universally applicable tool for predicting building energy performance. Another promising avenue for future research is the development of parallel models that consider various renovation measures and their potential impacts on energy performance. These could be designed to provide building owners or policymakers with more actionable insights. For example, a parallel model could predict how different types of insulation or HVAC systems might affect a building's energy usage. These predictions could be made not only for current conditions but also for future climates based on different climate change scenarios. This would offer a multi-dimensional view that could significantly aid in long-term planning and decision-making processes for both building renovations and new constructions.
Furthermore, the integration of the model into digital twin systems for cities could be examined more thoroughly. While the current study suggests this as a potential application, detailed research would be needed to understand how best to implement this and what the potential benefits and drawbacks might be.
Also, it should be noted that the execution of these suggested future works would require significant computational resources. Complex models that take into account diverse building types, materials, and climatic conditions would necessitate more powerful computing capabilities for data processing, model training, and validation. Furthermore, to improve the generalizability

and reliability of the model, accurate and comprehensive data about the building stock in various regions would be crucial. Collecting such data might involve large-scale surveys, satellite imagery analysis, or partnerships with governmental and private institutions that have access to relevant databases.

The quality of data fed into the model is as crucial as the computational power, given that machine learning models are heavily dependent on the data they are trained on. Poor or inaccurate data can severely limit the effectiveness and reliability of the model, making the need for high-quality data gathering paramount. Therefore, while expanding the scale and scope of the model holds considerable promise for making it a universally applicable tool, it also implies a significant commitment in terms of computational resources and data quality. Overall, the successful realization of these future work suggestions would necessitate a multidisciplinary approach, bringing together expertise in computing, data science, architecture, urban planning, and climate science.

8. Reference

1. Watson, V., *'The planned city sweeps the poor away…': Urban planning and 21st century urbanisation.* Progress in Planning, 2009. **72**(3): p. 151-193.
2. Mori, K. and A. Christodoulou, *Review of sustainability indices and indicators: Towards a new City Sustainability Index (CSI).* Environmental Impact Assessment Review, 2012. **32**(1): p. 94-106.
3. IEA, *Key World Energy Statistics 2019*. 2019, Paris, France: International Energy Agency.
4. EU, *binding annual greenhouse gas emission reductions by Member States from 2021 to 2030 for a resilient Energy Union*. 2016, European Comission: Brussels, Belgium.
5. Omer, A.M., *Energy, environment and sustainable development.* Renewable and Sustainable Energy Reviews, 2008. **12**(9): p. 2265-2300.
6. Bazazzadeh, H., et al. *The Impact Assessment of Climate Change on Building Energy Consumption in Poland*. Energies, 2021. **14**, DOI: 10.3390/en14144084.
7. UN DESA, *Sustainable Development Goal 7: Ensure Access to Affordable, Reliable, Sustainable and Modern Energy for All*. 2017, New Yok, USA: United Nations Department of Economic and Social Affairs (UN DESA).
8. Gielen, D., et al., *The role of renewable energy in the global energy transformation.* Energy Strategy Reviews, 2019. **24**: p. 38-50.
9. Riahi, K., et al., *Energy Pathways for Sustainable Development*, in *Global Energy Assessment: Toward a Sustainable Future*, T. Global Energy Assessment Writing, Editor. 2012, Cambridge University Press: Cambridge. p. 1205-1306.
10. Davidson, D.J., *Exnovating for a renewable energy transition.* Nature Energy, 2019. **4**(4): p. 254-256.
11. York, R. and S.E. Bell, *Energy transitions or additions?: Why a transition from fossil fuels requires more than the growth of renewable energy.* Energy Research & Social Science, 2019. **51**: p. 40-43.
12. Quitzow, R., et al., *The COVID-19 crisis deepens the gulf between leaders and laggards in the global energy transition.* Energy Research & Social Science, 2021. **74**: p. 101981.
13. Malec, M., *The prospects for decarbonisation in the context of reported resources and energy policy goals: The case of Poland.* Energy Policy, 2022. **161**: p. 112763.
14. IEA, *World Energy Balances*. 2021, Paris, France: International Energy Agency.
15. IEA, *Net Zero by 2050*. 2021, Paris, France: International Energy Agency.
16. IEA, *Global Energy Review 2021*. 2021, Paris, France: International Energy Agency.
17. IEA, *Greenhouse Gas Emissions from Energy*. 2021, Paris, France: International Energy Agency.
18. Levesque, A., et al., *How much energy will buildings consume in 2100? A global perspective within a scenario framework.* Energy, 2018. **148**: p. 514-527.
19. González-Torres, M., et al., *A review on buildings energy information: Trends, end-uses, fuels and drivers.* Energy Reports, 2022. **8**: p. 626-637.
20. Mavromatidis, G., et al., *A strategy for reducing CO2 emissions from buildings with the Kaya identity – A Swiss energy system analysis and a case study.* Energy Policy, 2016. **88**: p. 343-354.
21. Aksoezen, M., et al., *Building age as an indicator for energy consumption.* Energy and Buildings, 2015. **87**: p. 74-86.
22. Radwan, A.F., et al., *Retrofitting of existing buildings to achieve better energy-efficiency in commercial building case study: Hospital in Egypt.* Alexandria Engineering Journal, 2016. **55**(4): p. 3061-3071.
23. Asdrubali, F., et al., *Energy and environmental payback times for an NZEB retrofit.* Building and Environment, 2019. **147**: p. 461-472.

24. Hall, M.R., et al., *Analysis of UK domestic building retrofit scenarios based on the E.ON Retrofit Research House using energetic hygrothermics simulation – Energy efficiency, indoor air quality, occupant comfort, and mould growth potential.* Building and Environment, 2013. **70**: p. 48-59.
25. EU Parliment, *DIRECTIVE 2012/27/EU on Energy Efficiency, Amending Directives 2009/125/EC and 2010/30/EU and Repealing Directives 2004/8/EC and 2006/32/EC*, E. Parliment, Editor. 2012, EU Parliment.
26. Seyedzadeh, S., et al., *Machine learning for estimation of building energy consumption and performance: a review.* Visualization in Engineering, 2018. **6**(1): p. 5.
27. Wang, D., et al., *CESAR: A bottom-up building stock modelling tool for Switzerland to address sustainable energy transformation strategies.* Energy and Buildings, 2018. **169**: p. 9-26.
28. Li, X., C.P. Bowers, and T. Schnier, *Classification of Energy Consumption in Buildings With Outlier Detection.* IEEE Transactions on Industrial Electronics, 2010. **57**(11): p. 3639-3644.
29. EU *Primary and final energy consumption slowly decreasing*. 2021.
30. Breznitz, D. and D. Ornston, R.a.D. section, Editor. 2017, European Bank: Frankfurt, Germany.
31. EPRS, *Total EU allocations of Cohesion Policy 2014 – 2020 (€ Billion, Current prices)*. 2014, European Parlimentary Research Service: EPRS website.
32. The Economist, *Why 33 of the 50 most-polluted towns in Europe are in Poland*. 2018, The Economic, Europe Section.
33. BPIE, *Financing Renovation of Buildings in Poland*. 2018, Brussel, Belgium: Buildings Performance Institute Europe.
34. Dall'O', G., A. Speccher, and E. Bruni, *The Green Energy Audit, a new procedure for the sustainable auditing of existing buildings integrated with the LEED Protocols.* Sustainable Cities and Society, 2012. **3**: p. 54-65.
35. Deane, B., *Ask The Contractor: Things to Know About Home Energy Audits*, in *Forbes HOME*. 2022: Forbes webpage.
36. Jagarajan, R., et al., *Green retrofitting – A review of current status, implementations and challenges.* Renewable and Sustainable Energy Reviews, 2017. **67**: p. 1360-1368.
37. Rodriguez, A., S.T. Smith, and B. Potter, *Sensitivity analysis for building energy audit calculation methods: Handling the uncertainties in small power load estimation.* Energy, 2022. **238**: p. 121511.
38. Jiao, J., Q. Xia, and F. Shi, *Nondestructive inspection of a brick–timber structure in a modern architectural heritage building: Lecture hall of the Anyuan Miners' Club, China.* Frontiers of Architectural Research, 2019. **8**(3): p. 348-358.
39. Neuman, S.B., T. Kaefer, and A. Pinkham, *Building Background Knowledge.* The Reading Teacher, 2014. **68**(2): p. 145-148.
40. Müller-Bloch, C. and J. Kranz. *A Framework for Rigorously Identifying Research Gaps in Qualitative Literature Reviews*. in *Proceedings/International Conference on Information Systems (ICIS)*. 2015.
41. Paul, J. and A.R. Criado, *The art of writing literature review: What do we know and what do we need to know?* International Business Review, 2020. **29**(4): p. 101717.
42. Haddaway, N.R., et al., *ROSES RepOrting standards for Systematic Evidence Syntheses: pro forma, flow-diagram and descriptive summary of the plan and conduct of environmental systematic reviews and systematic maps.* Environmental Evidence, 2018. **7**(1): p. 7.
43. Hong, Q.N., A. Gonzalez-Reyes, and P. Pluye, *Improving the usefulness of a tool for appraising the quality of qualitative, quantitative and mixed methods studies, the Mixed Methods Appraisal Tool (MMAT).* Journal of Evaluation in Clinical Practice, 2018. **24**(3): p. 459-467.
44. Volk, R., J. Stengel, and F. Schultmann, *Building Information Modeling (BIM) for existing buildings — Literature review and future needs.* Automation in Construction, 2014. **38**: p. 109-127.
45. Jung, J., et al., *Automated 3D volumetric reconstruction of multiple-room building interiors for as-built BIM.* Advanced Engineering Informatics, 2018. **38**: p. 811-825.

46. Li, Y., et al., *3D LiDAR and multi-technology collaboration for preservation of built heritage in China: A review.* International Journal of Applied Earth Observation and Geoinformation, 2023. **116**: p. 103156.
47. Wang, Q., et al., *3D-CALI: Automatic calibration for camera and LiDAR using 3D checkerboard.* Measurement, 2022. **203**: p. 111971.
48. Yang, Y., et al., *Evaluation of structure from motion (SfM) photogrammetry on the measurement of rill and interrill erosion in a typical loess.* Geomorphology, 2021. **385**: p. 107734.
49. Liu, L. and I. Stamos, *A systematic approach for 2D-image to 3D-range registration in urban environments.* Computer Vision and Image Understanding, 2012. **116**(1): p. 25-37.
50. Al-khafaji, S.L., et al., *Spectral-Spatial Scale Invariant Feature Transform for Hyperspectral Images.* IEEE Transactions on Image Processing, 2018. **27**(2): p. 837-850.
51. Moulon, P., P. Monasse, and R. Marlet. *Adaptive Structure from Motion with a Contrario Model Estimation*. in *Computer Vision – ACCV 2012*. 2013. Berlin, Heidelberg: Springer Berlin Heidelberg.
52. Furukawa, Y. and J. Ponce, *Accurate, Dense, and Robust Multiview Stereopsis.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010. **32**(8): p. 1362-1376.
53. Umeyama, S., *Least-squares estimation of transformation parameters between two point patterns.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991. **13**(4): p. 376-380.
54. Dino, I.G., et al., *Image-based construction of building energy models using computer vision.* Automation in Construction, 2020. **116**: p. 103231.
55. Nelder, J.A. and R. Mead, *A Simplex Method for Function Minimization.* The Computer Journal, 1965. **7**(4): p. 308-313.
56. Jaboyedoff, M., et al., *Use of LIDAR in landslide investigations: a review.* Natural Hazards, 2012. **61**(1): p. 5-28.
57. Telling, J., et al., *Review of Earth science research using terrestrial laser scanning.* Earth-Science Reviews, 2017. **169**: p. 35-68.
58. Lohani, B. and S. Ghosh, *Airborne LiDAR Technology: A Review of Data Collection and Processing Systems.* Proceedings of the National Academy of Sciences, India Section A: Physical Sciences, 2017. **87**(4): p. 567-579.
59. Luetzenburg, G., A. Kroon, and A.A. Bjørk, *Evaluation of the Apple iPhone 12 Pro LiDAR for an Application in Geosciences.* Scientific Reports, 2021. **11**(1): p. 22221.
60. Tavani, S., et al., *Smartphone assisted fieldwork: Towards the digital transition of geoscience fieldwork using LiDAR-equipped iPhones.* Earth-Science Reviews, 2022. **227**: p. 103969.
61. Sankar, A. and S. Seitz, *Capturing indoor scenes with smartphones*, in *Proceedings of the 25th annual ACM symposium on User interface software and technology*. 2012, Association for Computing Machinery: Cambridge, Massachusetts, USA. p. 403–412.
62. Costantino, D., et al. *Smartphone LiDAR Technologies for Surveying and Reality Modelling in Urban Scenarios: Evaluation Methods, Performance and Challenges*. Applied System Innovation, 2022. **5**, DOI: 10.3390/asi5040063.
63. Xiong, X., et al., *Automatic creation of semantically rich 3D building models from laser scanner data.* Automation in Construction, 2013. **31**: p. 325-337.
64. Chan, J., et al., *Testing a Drop of Liquid Using Smartphone LiDAR.* Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., 2022. **6**(1): p. Article 3.
65. Stitt, J.M., et al., *Smartphone LIDAR can measure tree cavity dimensions for wildlife studies.* Wildlife Society Bulletin, 2019. **43**(1): p. 159-166.
66. Lehtola, V.V., et al. *Comparison of the Selected State-Of-The-Art 3D Indoor Scanning and Point Cloud Generation Methods*. Remote Sensing, 2017. **9**, DOI: 10.3390/rs9080796.
67. Azari, R., et al., *Multi-objective optimization of building envelope design for life cycle environmental performance.* Energy and Buildings, 2016. **126**: p. 524-534.

68. Elbeltagi, E., et al., *Visualized strategy for predicting buildings energy consumption during early design stage using parametric analysis.* Journal of Building Engineering, 2017. **13**: p. 127-136.
69. AIA, *An Architect's Guide to Integrating Energy Modeling in the Design Process*. American Institute of Architects. 2012, California, USA: American Institute of Architects.
70. Dogan, T. and C. Reinhart, *Shoeboxer: An algorithm for abstracted rapid multi-zone urban building energy model generation and simulation.* Energy and Buildings, 2017. **140**: p. 140-153.
71. Battini, F., G. Pernigotto, and A. Gasparella, *District-level validation of a shoeboxing simplification algorithm to speed-up Urban Building Energy Modeling simulations.* Applied Energy, 2023. **349**: p. 121570.
72. Laskari, M., et al., *On the impact of user behaviour on heating energy consumption and indoor temperature in residential buildings.* Energy and Buildings, 2022. **255**: p. 111657.
73. Hax, D.R., et al., *Influence of user behavior on energy consumption in a university building versus automation costs.* Energy and Buildings, 2022. **256**: p. 111730.
74. Nikdel, L., et al., *Datasets for occupancy profiles in apartment-style student housing for occupant behavior studies and application in building energy simulation.* Data in Brief, 2021. **37**: p. 107205.
75. Causone, F., et al., *A data-driven procedure to model occupancy and occupant-related electric load profiles in residential buildings for energy simulation.* Energy and Buildings, 2019. **202**: p. 109342.
76. Fiorillo, D., et al., *Assessing the Impact of Climate Change on Future Water Demand using Weather Data.* Water Resources Management, 2021. **35**(5): p. 1449-1462.
77. Jentsch, M., *Climate Change Weather File Generators*. 2010, niversity of Southampton Press: Southampton, UK.
78. Jentsch, M.F., et al., *Transforming existing weather data for worldwide locations to enable energy and building performance simulation under future climates.* Renewable Energy, 2013. **55**: p. 514-524.
79. Deru, M., et al., *US Department of Energy commercial reference building models of the national building stock*. 2011, Washington, DC, USA: US Department of Energy Press.
80. Zhang, J., et al., *Energy and Energy Cost Savings Analysis of the IECC for Commercial Buildings*. 2013, Pacific Northwest National Lab.(PNNL), Richland, WA (United States): Washington, USA.
81. Bazazzadeh, H., et al., *The Impact Assessment of Climate Change on Building Energy Consumption in Poland.* Energies, 2021. **14**(14): p. 4084.
82. Zhao, H.-X. and F. Magoulès, *Feature Selection for Predicting Building Energy Consumption Based on Statistical Learning Method.* Journal of Algorithms & Computational Technology, 2012. **6**(1): p. 59-77.
83. Parag, R., E.K. Mohammad, and A. Marilyne. *Gaussian-Process-Based Emulators for Building Performance Simulation*. in *Building simulation* 2017. San Francisco, USA: IBPSA.
84. Seyedzadeh, S., et al., *Tuning machine learning models for prediction of building energy loads.* Sustainable Cities and Society, 2019. **47**: p. 101484.
85. (ECSO), E.C.S.O., *Country Profile: Poland*. 2019, European Commission.
86. Dworakowska, A., et al., *Energy Efficiency in Poland. 2015 Review*. 2016, Institute of Environmental Economics: Krakow.
87. EU, *A Renovation Wave for Europe - greening our buildings, creating jobs, improving lives.* 2020: Brussel, Belgium.
88. Bertelsen, N. and B. Vad Mathiesen, *EU-28 Residential Heat Supply and Consumption: Historical Development and Status.* Energies, 2020. **13**(8): p. 1894.
89. World Bank, *Poland Energy Efficiency in Single Family Buildings Program (P170131)*. 2020, The World Bank: Washington, D.C., USA.
90. European Comission, *Assessment of the Final National Energy and Climate Plan of Poland*. 2020, European Commission: Brussels, Belgium.

91. Księżopolski, K., et al., *The Economic Effects of New Patterns of Energy Efficiency and Heat Sources in Rural Single-Family Houses in Poland.* Energies, 2020. **13**(23): p. 6358.
92. Rosenow, J., et al., *Evaluating the Polish White Certificate scheme.* Energy Policy, 2020. **144**: p. 111689.
93. BPIE, *Financing Building Energy Performance Improvement in Poland*. 2016, Buildings Performance Institute Europe: Brussels, Belgium.
94. NAPE, *Polish Building Typology*. 2012, Narodowa Agencja Poszanowania Energii: Warsaw, Poland.
95. H. Atanasiu , et al., *Implementing the cost-optimal methodology in EU countries: Lessons learned from three case studies*. 2013, Buildings Performance Institute Europe: Brussels, Belgium.
96. Hamdy, M., K. Sirén, and S. Attia, *Impact of financial assumptions on the cost optimality towards nearly zero energy buildings – A case study.* Energy and Buildings, 2017. **153**: p. 421-438.
97. Basinska, M., H. Koczyk, and E. Szczechowiak, *Sensitivity analysis in determining the optimum energy for residential buildings in Polish conditions.* Energy and Buildings, 2015. **107**: p. 307-318.
98. ISO, E., *Energy Performance of Buildings—Calculation of Energy Use for Space Heating and Cooling*. 2008, International Organization for Standardization: Geneva, Switzerland.
99. Firląg, S., *Cost-Optimal Plus Energy Building in a Cold Climate.* Energies, 2019. **12**(20): p. 3841.
100. Adamczyk, J. and R. Dylewski, *Ecological and Economic Benefits of the "Medium" Level of the Building Thermo-Modernization: A Case Study in Poland.* Energies, 2020. **13**(17): p. 4509.
101. Golabeska, E., *The profitalbility of investing in passive housing in Poland.* Economics and Environment, 2020. **74**(3).
102. Thomas Boermans, et al., *Cost optimal building performance requirements*. 2011, European Council for an Energy Efficient Economy: Stockholm, Sweden.
103. Statistics Poland, *Energy consumption in households in 2015*. 2017, Department of Statistical Publishing: Warsaw, Poland.
104. Regulski, B., et al., *Rynek ciepłowniczy w Polsce.* Rynek Energii, 2014. **113**(4): p. 9-16.
105. G. Berent-Kowalska, S. Peryt, and R. Wnuk, *Energy Efficiency Policies and Measures in Poland*. 2012, Polish National Energy Conservation Agency: Warsaw, Poland.
106. Attia, S., et al., *Energy efficiency in the polish residential building stock: A literature review.* Journal of Building Engineering, 2022. **45**: p. 103461.
107. A. Zieleniec and K. Koszniec, *Forum Energii*. 2020, Transformacja Energetyczna W Polsce: Warsaw, Poland.
108. Skoczkowski, T., et al., *Impact assessment of climate policy on Poland's power sector.* Mitigation and Adaptation Strategies for Global Change, 2018. **23**(8): p. 1303-1349.
109. R. W´ojcik, *Docieplanie Budynk´ow Od Wewnątrz*. 2017, Docieplanie Budynk´ow Od Wewnątrz, Grupa Medium: Warsaw, Poland.
110. PMI, *Polish Law: Dz.U. 1974, PN-74/B-03404 Wsp´ołczynnik Przenikania Ciepła Dla Przegr´od Budowlanych*, M.o.I.a.D. Poland, Editor. 1974: Warsaw, Poland.
111. Życzyńska, A. and T. Cholewa, *The modifications to the requirements on energy savings and thermal insulation of buildings in Poland in the years 1974-2021.* Budownictwo i Architektura, 2015. **14**(1): p. 145--154.
112. Pogorzelski, J., *Normalizacja ochrony cieplnej budynków w krajach UE iw Polsce. Przegląd bieżącej sytuacji.* Prace Instytutu Techniki Budowlanej, 1999. **28**(4): p. 39-51.
113. Pawłowski, K., *ROJEKTOWANIE przegr´od zewnetrznych W swietle NOWYCH wymagan CIEPLNYCH*, B.U.o.S.a. Technology, Editor. 2014: Bydgoszcz, Poland.
114. Chwieduk, D., *The role of Polish Energy Conservation Agency in promoting energy efficiency and renewable energy in Poland.* Renewable Energy, 1997. **10**(2): p. 323-326.
115. Robert, W. and K. Piotr, *On rehabilitation of buildings with historical façades.* Energy Procedia, 2017. **132**: p. 927-932.

116. Piasecki, M., et al., *Implementation of the Indoor Environmental Quality (IEQ) Model for the Assessment of a Retrofitted Historical Masonry Building.* Energies, 2020. **13**(22): p. 6051.
117. Węglarz, A. and M. Pierzchalski. *Comparing construction technologies of single family housing with regard of minimizing embodied energy and embodied carbon*. in *E3S Web of Conferences*. 2018. EDP Sciences.
118. Kisilewicz, T., M. Fedorczak-Cisak, and T. Barkanyi, *Active thermal insulation as an element limiting heat loss through external walls.* Energy and Buildings, 2019. **205**: p. 109541.
119. Lepkova, N., et al., *Comparative Analysis of Newly-Built Housing Quality in Poland and Lithuania.* Real Estate Management and Valuation, 3917. **25**(4): p. 85-98.
120. Blaszczok, M. and A. Baranowski, *Thermal Improvement in Residential Buildings in View of the Indoor Air Quality – Case Study for Polish Dwelling.* Architecture, Civil Engineering, Environment, 3918. **11**(3): p. 121-130.
121. Fedorczak-Cisak, M., et al., *Thermal and Vibration Comfort Analysis of a Nearly Zero-Energy Building in Poland.* Sustainability, 2018. **10**(10): p. 3774.
122. Attia, S., *Net Zero Energy Buildings (NZEB): Concepts, frameworks and roadmap for project analysis and implementation*. 2018, Quebec, Canada: Butterworth-Heinemann.
123. Euroheat & Power, *Plan to decarbonise district heating at PGE Group*. 2023, Euroheat & Power: Euroheat & Power Website.
124. Wojdyga, K. and M. Chorzelski, *Chances for Polish district heating systems.* Energy Procedia, 2017. **116**: p. 106-118.
125. Starościk, J., *Heating device market in Poland in 2018*. 2019, Warsaw, Poland: Association of Heating Device Manufacturers and Importer.
126. Cholewa, T. and A. Siuta-Olcha, *Long term experimental evaluation of the influence of heat cost allocators on energy consumption in a multifamily building.* Energy and Buildings, 2015. **104**: p. 122-130.
127. Chmielewska, A., M. Szulgowska-Zgrzywa, and J. Danielewicz. *Domestic hot water consumption in multi-apartment buildings*. in *E3S Web of Conferences*. 2017. EDP Sciences.
128. Cholewa, T., A. Siuta-Olcha, and R. Anasiewicz, *On the possibilities to increase energy efficiency of domestic hot water preparation systems in existing buildings – Long term field research.* Journal of Cleaner Production, 2019. **217**: p. 194-203.
129. Zimny, J., P. Michalak, and K. Szczotka, *Polish heat pump market between 2000 and 2013: European background, current state and development prospects.* Renewable and Sustainable Energy Reviews, 2015. **48**: p. 791-812.
130. Romańska-Zapała, A., M. Furtak, and M. Dechnik, *Cooperation of Horizontal Ground Heat Exchanger with the Ventilation Unit During Summer – Case Study.* IOP Conference Series: Materials Science and Engineering, 2017. **245**(5): p. 052027.
131. Wąsowska, B., *Long-Term Building Renovation Strategy*. 2021, Department of Energy Efficiency and Transformation: Warsaw, Poland.
132. Karlsen, S.S., M. Hamdy, and S. Attia, *Methodology to assess business models of dynamic pricing tariffs in all-electric houses.* Energy and Buildings, 2020. **207**: p. 109586.
133. Jarnut, M., S. Wermiński, and B. Waśkowicz, *Comparative analysis of selected energy storage technologies for prosumer-owned microgrids.* Renewable and Sustainable Energy Reviews, 2017. **74**: p. 925-937.
134. Zaborowski, M. and A. Dworakowska, *Energy Efficiency in Poland 2013 Review*. European Climate Foundation: Warsaw, Poland.
135. Moazami, A., et al., *Impacts of future weather data typology on building energy performance – Investigating long-term patterns of climate change and extreme weather conditions.* Applied Energy, 2019. **238**: p. 696-720.

136. IPCC, *Synthesis Report Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. 2014, IPCC: Geneva, Switzerland.
137. Berardi, U. and P. Jafarpur, *Assessing the impact of climate change on building heating and cooling energy demand in Canada.* Renewable and Sustainable Energy Reviews, 2020. **121**: p. 109681.
138. Madlener, R. and Y. Sunak, *Impacts of urbanization on urban structures and energy demand: What can we learn for urban energy planning and urbanization management?* Sustainable Cities and Society, 2011. **1**(1): p. 45-53.
139. Cao, X., X. Dai, and J. Liu, *Building energy-consumption status worldwide and the state-of-the-art technologies for zero-energy buildings during the past decade.* Energy and Buildings, 2016. **128**: p. 198-213.
140. Bazazzadeh, H., A. Nadolny, and S.S.H. Safaei, *Climate Change and Building Energy Consumption: A Review of the Impact of Weather Parameters Influenced by Climate Change on Household Heating and Cooling Demands of Buildings.* European Journal of Sustainable Development, 2021. **10**(2): p. 1.
141. IEA, *Tracking Buildings*. 2020, IEA: Paris, France.
142. Tabrizikahou, A. and P. Nowotarski, *Mitigating the Energy Consumption and the Carbon Emission in the Building Structures by Optimization of the Construction Processes.* Energies, 2021. **14**(11): p. 3287.
143. Mancini, F. and G. Lo Basso, *How Climate Change Affects the Building Energy Consumptions Due to Cooling, Heating, and Electricity Demands of Italian Residential Sector.* Energies, 2020. **13**(2): p. 410.
144. Yau, Y.H. and S. Hasbi, *A review of climate change impacts on commercial buildings and their technical services in the tropics.* Renewable and Sustainable Energy Reviews, 2013. **18**: p. 430-441.
145. Kheiri, F., *A review on optimization methods applied in energy-efficient building geometry and envelope design.* Renewable and Sustainable Energy Reviews, 2018. **92**: p. 897-920.
146. Gero, J.S., N. D'Cruz, and A.D. Radford, *Energy in context: A multicriteria model for building design.* Building and Environment, 1983. **18**(3): p. 99-107.
147. European Commission, *Energy Consumption of Buildings per m²*. 2013, European Commission: Brussels, Belgium.
148. Stuggins, G., A. Sharabaroff, and Y. Semikolenova, *Energy Efficiency; Lessons Learned from Success Stories*. Europe and Central Asia Reports. 2013: The World Bank. 68.
149. IEA, *Energy Policies of IEA Countries: Poland 2016 Review*. 2017, IEA: Paris, France.
150. Build Desk Polska, *Energy Condition of Buildings in Poland*. 2011, Build Desk Polska: Cigacice, Poland.
151. Parysek, J.J. and L. Mierzejewska, *Poznań.* Cities, 2006. **23**(4): p. 291-305.
152. Build Desk Polska, *Analytics Report; Energy Condition of Buildings in Poland*. 2019, Build Desk Polska: Cigacice, Poland.
153. EU, *Energy Island Communities for Energy Transition*. 2021, Energy Island Communities for Energy Transition: Brussels, Belgium.
154. Pérez-Lombard, L., J. Ortiz, and C. Pout, *A review on buildings energy consumption information.* Energy and Buildings, 2008. **40**(3): p. 394-398.
155. Giannakopoulos, C. and B.E. Psiloglou, *Trends in energy load demand for Athens, Greece*

weather and non-weather related factors. Climate Research, 2006. **31**(1): p. 97-108.
156. Carbon Trust, *airconditioning: maximisingcomfort, minimising energy consumption*. 2007, Carbon Trust: London, UK.
157. ODPM, *age of commercial and industrial stock: local authority: level 2004 England and Wales*. 2005, Office of Deputy Prime Minister: London, UK.

158. Delfani, S., M. Karami, and H. Pasdarshahri, *The effects of climate change on energy consumption of cooling systems in Tehran.* Energy and Buildings, 2010. **42**(10): p. 1952-1957.
159. Lam, J.C., et al., *Building energy efficiency in different climates.* Energy Conversion and Management, 2008. **49**(8): p. 2354-2366.
160. Lam, J.C., et al., *Weather data analysis and design implications for different climatic zones in China.* Building and Environment, 2005. **40**(2): p. 277-296.
161. Humphreys, M., *Outdoor temperatures and comfort indoors.* Batiment International, Building Research and Practice, 1978. **6**(2): p. 92-92.
162. Frank, T., *Climate change impacts on building heating and cooling energy demand in Switzerland.* Energy and Buildings, 2005. **37**(11): p. 1175-1185.
163. Belzer, D.B., M.J. Scott, and R.D. Sands, *Climate Change Impacts on U.S. Commercial Building Energy Consumption: An Analysis Using Sample Survey Data.* Energy Sources, 1996. **18**(2): p. 177-201.
164. Considine, T.J., *The impacts of weather variations on energy demand and carbon emissions.* Resource and Energy Economics, 2000. **22**(4): p. 295-314.
165. Camilleri, M., R. Jaques, and N. Isaacs, *Impacts of climate change on building performance in New Zealand.* Building Research & Information, 2001. **29**(6): p. 440-450.
166. Aguiar, R., M. Oliveira, and H. Gonccedilalves, *Climate change impacts on the thermal performance of Portuguese buildings. Results of the SIAM study.* Building Services Engineering Research and Technology, 2002. **23**(4): p. 223-231.
167. Christenson, M., H. Manz, and D. Gyalistras, *Climate warming impact on degree-days and building energy demand in Switzerland.* Energy Conversion and Management, 2006. **47**(6): p. 671-686.
168. Kyle, P., et al., *Climate Policy and the Long-Term Evolution of the U.S. Buildings Sector.* The Energy Journal, 2010. **31**(2): p. 145-172.
169. Hadley, S.W., et al., *Responses of energy use to climate change: A climate modeling study.* Geophysical Research Letters, 2006. **33**(17).
170. Huang, W.Z., M. Zaheeruddin, and S.H. Cho, *Dynamic simulation of energy management control functions for HVAC systems in buildings.* Energy Conversion and Management, 2006. **47**(7): p. 926-943.
171. Seyedzadeh, S., et al., *Machine learning modelling for predicting non-domestic buildings energy performance: A model to support deep energy retrofit decision-making.* Applied Energy, 2020. **279**: p. 115908.
172. Eslamirad, N., et al., *Thermal comfort prediction by applying supervised machine learning in green sidewalks of Tehran.* Smart and Sustainable Built Environment, 2020. **9**(4): p. 361-374.
173. Ghaffour, W., M.N. Ouissi, and M.A. Velay Dabat, *Analysis of urban thermal environments based on the perception and simulation of the microclimate in the historic city of Tlemcen.* Smart and Sustainable Built Environment, 2021. **10**(2): p. 141-168.
174. Hp, T., R. C, and D. Md, *Developing a building performance score model for assessing the sustainability of buildings.* Smart and Sustainable Built Environment, 2022. **11**(1): p. 143-161.
175. Rahmouni, S. and R. Smail, *A design approach towards sustainable buildings in Algeria.* Smart and Sustainable Built Environment, 2020. **9**(3): p. 229-245.
176. Dell'Anna, F., et al., *Designing a decision support system to evaluate the environmental and extra-economic performances of a nearly zero-energy building.* Smart and Sustainable Built Environment, 2020. **9**(4): p. 413-442.
177. Seyedzadeh, S. and F. Pour Rahimian, *Building Energy Data-Driven Model Improved by Multi-objective Optimisation*, in *Data-Driven Modelling of Non-Domestic Buildings Energy Performance: Supporting Building Retrofit Planning*, S. Seyedzadeh and F. Pour Rahimian, Editors. 2021, Springer International Publishing: Cham. p. 99-109.

178. Weiss, T., *Energy flexibility and shiftable heating power of building components and technologies.* Smart and Sustainable Built Environment, 2021. **10**(4): p. 740-758.
179. Sheikhkhoshkar, M., et al., *Automated planning of concrete joint layouts with 4D-BIM.* Automation in Construction, 2019. **107**: p. 102943.
180. Rosenthal, D.H., H.K. Gruenspecht, and E.A. Moran, *Effects of Global Warming on Energy Use for Space Heating and Cooling in the United States.* The Energy Journal, 1995. **16**(2): p. 77-96.
181. Nordhaus, W.D., *To Slow or Not to Slow: The Economics of The Greenhouse Effect.* The Economic Journal, 1991. **101**(407): p. 920-937.
182. Nik, V.M. and A. Sasic Kalagasidis, *Impact study of the climate change on the energy performance of the building stock in Stockholm considering four climate uncertainties.* Building and Environment, 2013. **60**: p. 291-304.
183. Hosseini, M., F. Tardy, and B. Lee, *Cooling and heating energy performance of a building with a variety of roof designs; the effects of future weather data in a cold climate.* Journal of Building Engineering, 2018. **17**: p. 107-114.
184. Cellura, M., et al., *Climate change and the building sector: Modelling and energy implications to an office building in southern Europe.* Energy for Sustainable Development, 2018. **45**: p. 46-65.
185. Crawley, D.B., *Estimating the impacts of climate change and urbanization on building performance.* Journal of Building Performance Simulation, 2008. **1**(2): p. 91-115.
186. Triana, M.A., R. Lamberts, and P. Sassi, *Should we consider climate change for Brazilian social housing? Assessment of energy efficiency adaptation measures.* Energy and Buildings, 2018. **158**: p. 1379-1392.
187. Wang, H. and Q. Chen, *Impact of climate change heating and cooling energy use in buildings in the United States.* Energy and Buildings, 2014. **82**: p. 428-436.
188. Cox, R.A., et al., *Simple future weather files for estimating heating and cooling demand.* Building and Environment, 2015. **83**: p. 104-114.
189. Shen, P., *Impacts of climate change on U.S. building energy use by using downscaled hourly future weather data.* Energy and Buildings, 2017. **134**: p. 61-70.
190. Kottek, M., et al., *World map of the Köppen-Geiger climate classification updated.* Meteorologische Zeitschrift, 2006. **15**(3).
191. Velashjerdi Farahani, A., et al. *Overheating Risk and Energy Demand of Nordic Old and New Apartment Buildings during Average and Extreme Weather Conditions under a Changing Climate*. Applied Sciences, 2021. **11**, DOI: 10.3390/app11093972.
192. Herrera, M., et al., *A review of current and future weather data for building simulation.* Building Services Engineering Research and Technology, 2017. **38**(5): p. 602-627.
193. Lam, J.C., S.C.M. Hui, and A.L.S. Chan, *A Statistical Approach to the Development of a Typical Meteorological Year for Hong Kong.* Architectural Science Review, 1996. **39**(4): p. 201-209.
194. Petrakis, M., et al., *Generation of a "typical meteorological year" for Nicosia, Cyprus.* Renewable Energy, 1998. **13**(3): p. 381-388.
195. Bilbao, J., et al., *Test Reference Year Generation and Evaluation Methods in the Continental Mediterranean Area.* Journal of Applied Meteorology, 2004. **43**(2): p. 390-400.
196. Chan, A.L.S., et al., *Generation of a typical meteorological year for Hong Kong.* Energy Conversion and Management, 2006. **47**(1): p. 87-96.
197. Pouriya, J. and B. Umberto, *Building energy demand within a climate change perspective: The need for future weather file.* IOP Conference Series: Materials Science and Engineering, 2019. **609**(7): p. 072037.
198. Kundzewicz, Z.W., et al., *Assessment of climate change and associated impact on selected sectors in Poland.* Acta Geophysica, 2018. **66**(6): p. 1509-1523.

199. Mezghani, A., et al., *CHASE-PL Climate Projection dataset over Poland – bias adjustment of EURO-CORDEX simulations.* Earth Syst. Sci. Data, 2017. **9**(2): p. 905-925.
200. Thomson, A.M., et al., *RCP4.5: a pathway for stabilization of radiative forcing by 2100.* Climatic Change, 2011. **109**(1): p. 77.
201. Riahi, K., et al., *RCP 8.5—A scenario of comparatively high greenhouse gas emissions.* Climatic Change, 2011. **109**(1): p. 33.
202. Horton, B.P., et al., *Estimating global mean sea-level rise and its uncertainties by 2100 and 2300 from an expert survey.* npj Climate and Atmospheric Science, 2020. **3**(1): p. 18.
203. Guan, L., *Preparation of future weather data to study the impact of climate change on buildings.* Building and Environment, 2009. **44**(4): p. 793-800.
204. Laflamme, E.M., E. Linder, and Y. Pan, *Statistical downscaling of regional climate model output to achieve projections of precipitation extremes.* Weather and Climate Extremes, 2016. **12**: p. 15-23.
205. Belcher, S.E., J.N. Hacker, and D.S. Powell, *Constructing design weather data for future climates.* Building Services Engineering Research and Technology, 2005. **26**(1): p. 49-61.
206. Wilby, R.L. and T.M.L. Wigley, *Downscaling general circulation model output: a review of methods and limitations.* Progress in Physical Geography: Earth and Environment, 1997. **21**(4): p. 530-548.
207. van Paassen, A.H. and Q.X. Luo, *Weather data generator to study climate change on buildings.* Building Services Engineering Research and Technology, 2002. **23**(4): p. 251-258.
208. Zhang, L., et al., *Comparison of Statistical and Dynamic Downscaling Techniques in Generating High-Resolution Temperatures in China from CMIP5 GCMs.* Journal of Applied Meteorology and Climatology, 2020. **59**(2): p. 207-235.
209. P.Tootkaboni, M., et al. *A Comparative Analysis of Different Future Weather Data for Building Energy Performance Simulation*. Climate, 2021. **9**, DOI: 10.3390/cli9020037.
210. Liu, S., et al., *Development and application of future design weather data for evaluating the building thermal-energy performance in subtropical Hong Kong.* Energy and Buildings, 2020. **209**: p. 109696.
211. Hosseini, M., A. Bigtashi, and B. Lee, *Generating future weather files under climate change scenarios to support building energy simulation – A machine learning approach.* Energy and Buildings, 2021. **230**: p. 110543.
212. Yassaghi, H., P.L. Gurian, and S. Hoque, *Propagating downscaled future weather file uncertainties into building energy use.* Applied Energy, 2020. **278**: p. 115655.
213. Kelly, S., D. Crawford-Brown, and M.G. Pollitt, *Building performance evaluation and certification in the UK: Is SAP fit for purpose?* Renewable and Sustainable Energy Reviews, 2012. **16**(9): p. 6861-6878.
214. Filippín, C., *Benchmarking the energy efficiency and greenhouse gases emissions of school buildings in central Argentina.* Building and Environment, 2000. **35**(5): p. 407-414.
215. Xuemei, L., et al. *A Novel Hybrid Approach of KPCA and SVM for Building Cooling Load Prediction*. in *2010 Third International Conference on Knowledge Discovery and Data Mining*. 2010.
216. Abrahamse, W., et al., *The effect of tailored information, goal setting, and tailored feedback on household energy use, energy-related behaviors, and behavioral antecedents.* Journal of Environmental Psychology, 2007. **27**(4): p. 265-276.
217. Hong, T., et al., *A review on sustainable construction management strategies for monitoring, diagnosing, and retrofitting the building's dynamic energy performance: Focused on the operation and maintenance phase.* Applied Energy, 2015. **155**: p. 671-707.
218. Nikolaou, T., et al., *Review and state of the art on methodologies of buildings' energy-efficiency classification.* Managing indoor environments and energy in buildings with integrated intelligent systems, 2015: p. 13-31.

219. Si, J., *Green retrofit of existing non-domestic buildings as a multi criteria decision making process*. 2017, UCL (University College London).
220. Zhao, H.-x. and F. Magoulès, *A review on the prediction of building energy consumption.* Renewable and Sustainable Energy Reviews, 2012. **16**(6): p. 3586-3592.
221. Rastogi, P., *On the sensitivity of buildings to climate: the interaction of weather and building envelopes in determining future building energy consumption*. 2016, École Politechnique Federale de Lausanne (EPFL),: Lausanne, Switzerland.
222. Kalogirou, S., C. Neocleous, and C. Schizas. *Heating load estimation using artificial neural networks*. in *Brussels, Belgium: Proceedings of CLIMA 2000 Conference on CD-ROM*. 1997.
223. Zhang, A., et al., *Dive into deep learning.* arXiv preprint arXiv:2106.11342, 2021.
224. Huang, Y. and C. Li, *Accurate heating, ventilation and air conditioning system load prediction for residential buildings using improved ant colony optimization and wavelet neural network.* Journal of Building Engineering, 2021. **35**: p. 101972.
225. Mendoza-Pittí, L., et al. *Developing a Long Short-Term Memory-Based Model for Forecasting the Daily Energy Consumption of Heating, Ventilation, and Air Conditioning Systems in Buildings*. Applied Sciences, 2021. **11**, DOI: 10.3390/app11156722.
226. Alden, R.E., et al., *Artificial Intelligence Method for the Forecast and Separation of Total and HVAC Loads With Application to Energy Management of Smart and NZE Homes.* IEEE Access, 2021. **9**: p. 160497-160509.
227. Wang, Z., T. Hong, and M.A. Piette, *Data fusion in predicting internal heat gains for office buildings through a deep learning approach.* Applied Energy, 2019. **240**: p. 386-398.
228. Jogunola, O., et al. *CBLSTM-AE: A Hybrid Deep Learning Framework for Predicting Energy Consumption*. Energies, 2022. **15**, DOI: 10.3390/en15030810.
229. Das, A., et al., *Occupant-centric miscellaneous electric loads prediction in buildings using state-of-the-art deep learning methods.* Applied Energy, 2020. **269**: p. 115135.
230. Ullah, F.U., et al. *Diving Deep into Short-Term Electricity Load Forecasting: Comparative Analysis and a Novel Framework*. Mathematics, 2021. **9**, DOI: 10.3390/math9060611.
231. Li, W., et al., *A clustering-based approach for "cross-scale" load prediction on building level in HVAC systems.* Applied Energy, 2021. **282**: p. 116223.
232. Kim, T.-Y. and S.-B. Cho, *Predicting residential energy consumption using CNN-LSTM neural networks.* Energy, 2019. **182**: p. 72-81.
233. He, Y. and K.F. Tsang, *Universities power energy management: A novel hybrid model based on iCEEMDAN and Bayesian optimized LSTM.* Energy Reports, 2021. **7**: p. 6473-6488.
234. Haq, I.U., et al. *Sequential Learning-Based Energy Consumption Prediction Model for Residential and Commercial Sectors*. Mathematics, 2021. **9**, DOI: 10.3390/math9060605.
235. Chalapathy, R., N.L.D. Khoa, and S. Sethuvenkatraman, *Comparing multi-step ahead building cooling load prediction using shallow machine learning and deep learning models.* Sustainable Energy, Grids and Networks, 2021. **28**: p. 100543.
236. Bouktif, S., et al. *Multi-Sequence LSTM-RNN Deep Learning and Metaheuristics for Electric Load Forecasting*. Energies, 2020. **13**, DOI: 10.3390/en13020391.
237. Tso, G.K.F. and K.K.W. Yau, *Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks.* Energy, 2007. **32**(9): p. 1761-1768.
238. Park, Y.S. and S. Lek, *Chapter 7 - Artificial Neural Networks: Multilayer Perceptron for Ecological Modeling*, in *Developments in Environmental Modelling*, S.E. Jørgensen, Editor. 2016, Elsevier. p. 123-140.
239. Gupta, T.K. and K. Raza, *Chapter 7 - Optimization of ANN Architecture: A Review on Nature-Inspired Techniques*, in *Machine Learning in Bio-Signal Analysis and Diagnostic Imaging*, N. Dey, et al., Editors. 2019, Academic Press. p. 159-182.

240. Alamin, Y.I., et al., *An Artificial Neural Network (ANN) model to predict the electric load profile for an HVAC system⁎⁎This work has been funded by the grant from the Spanish Ministry of Economy and Competitiveness (ENERPRO DPI 2014-56364-C2-1-R). Yaser I. Alamin is a fellow of the MARHABA, an Erasmus Mundus Lot 3 project. José Domingo Álvarez is a fellow of the Spanish 'Ramón y Cajal' contract program, co-financed by the European Social Fund. Antonio Ruano acknowledges the support of FCT through IDMEC, under LAETA grant UID/EMS/50022/2013.* IFAC-PapersOnLine, 2018. **51**(10): p. 26-31.
241. Seo, B., et al. *Application of Artificial Neural Network for the Optimum Control of HVAC Systems in Double-Skinned Office Buildings*. Energies, 2019. **12**, DOI: 10.3390/en12244754.
242. Muralitharan, K., R. Sakthivel, and R. Vishnuvarthan, *Neural network based optimization approach for energy demand prediction in smart grid.* Neurocomputing, 2018. **273**: p. 199-208.
243. Kusiak, A., G. Xu, and Z. Zhang, *Minimization of energy consumption in HVAC systems with data-driven models and an interior-point method.* Energy Conversion and Management, 2014. **85**: p. 146-153.
244. Ray, M., P. Samal, and C.K. Panigrahi, *Implementation of a Hybrid Technique for the Predictive Control of the Residential Heating Ventilation and Air Conditioning Systems.* Engineering, Technology & Applied Science Research, 2022. **12**(3): p. 8772-8776.
245. Ruiz, L.G.B., et al., *Energy consumption forecasting based on Elman neural networks with evolutive optimization.* Expert Systems with Applications, 2018. **92**: p. 380-389.
246. Zhong, H., et al., *Vector field-based support vector regression for building energy consumption prediction.* Applied Energy, 2019. **242**: p. 403-414.
247. Li, Q., et al., *Applying support vector machine to predict hourly cooling load in the building.* Applied Energy, 2009. **86**(10): p. 2249-2256.
248. Ding, Y., Q. Zhang, and T. Yuan, *Research on short-term and ultra-short-term cooling load prediction models for office buildings.* Energy and Buildings, 2017. **154**: p. 254-267.
249. Paudel, S., et al., *A relevant data selection method for energy consumption prediction of low energy building based on support vector machine.* Energy and Buildings, 2017. **138**: p. 240-256.
250. Zhao, J. and X. Liu, *A hybrid method of dynamic cooling and heating load forecasting for office buildings based on artificial intelligence and regression analysis.* Energy and Buildings, 2018. **174**: p. 293-308.
251. Ngo, N.-T., et al., *Proposing a hybrid metaheuristic optimization algorithm and machine learning model for energy use forecast in non-residential buildings.* Scientific Reports, 2022. **12**(1): p. 1065.
252. Wang, Z., et al., *Random Forest based hourly building energy prediction.* Energy and Buildings, 2018. **171**: p. 11-25.
253. Seyedzadeh, S., et al., *Data driven model improved by multi-objective optimisation for prediction of building energy loads.* Automation in Construction, 2020. **116**: p. 103188.
254. Rana, M., S. Sethuvenkatraman, and M. Goldsworthy, *A data-driven approach based on quantile regression forest to forecast cooling load for commercial buildings.* Sustainable Cities and Society, 2022. **76**: p. 103511.
255. Ahmad, T., et al., *Supervised based machine learning models for short, medium and long-term energy prediction in distinct building environment.* Energy, 2018. **158**: p. 17-32.
256. Kamel, E., S. Sheikh, and X. Huang, *Data-driven predictive models for residential building energy use based on the segregation of heating and cooling days.* Energy, 2020. **206**: p. 118045.
257. Sauer, J., et al., *Extreme gradient boosting model based on improved Jaya optimizer applied to forecasting energy consumption in residential buildings.* Evolving Systems, 2022. **13**(4): p. 577-588.
258. Lu, H., et al., *Short-term prediction of building energy consumption employing an improved extreme gradient boosting model: A case study of an intake tower.* Energy, 2020. **203**: p. 117756.

259. Feng, Y., et al., *Space cooling energy usage prediction based on utility data for residential buildings using machine learning methods.* Applied Energy, 2021. **291**: p. 116814.
260. Ke, G., et al., *Lightgbm: A highly efficient gradient boosting decision tree.* Advances in neural information processing systems, 2017. **30**.
261. Fan, J., et al., *Comparison of Support Vector Machine and Extreme Gradient Boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in China.* Energy Conversion and Management, 2018. **164**: p. 102-111.
262. Sun, X., M. Liu, and Z. Sima, *A novel cryptocurrency price trend forecasting model based on LightGBM.* Finance Research Letters, 2020. **32**: p. 101084.
263. Ju, Y., et al., *A Model Combining Convolutional Neural Network and LightGBM Algorithm for Ultra-Short-Term Wind Power Forecasting.* IEEE Access, 2019. **7**: p. 28309-28318.
264. Qiang, G., et al., *An improved office building cooling load prediction model based on multivariable linear regression.* Energy and Buildings, 2015. **107**: p. 445-455.
265. Fumo, N. and M.A. Rafe Biswas, *Regression analysis for prediction of residential energy consumption.* Renewable and Sustainable Energy Reviews, 2015. **47**: p. 332-343.
266. Fan, C. and Y. Ding, *Cooling load prediction and optimal operation of HVAC systems using a multiple nonlinear regression model.* Energy and Buildings, 2019. **197**: p. 7-17.
267. Chen, S., et al., *An online physical-based multiple linear regression model for building's hourly cooling load prediction.* Energy and Buildings, 2022. **254**: p. 111574.