

Development of an AI-based surrogate model to select optimal building envelope retrofit solutions considering heat waves in the Netherlands

Isabella Lucchese

Delft University of Technology
Master Architecture, Urbanism and Building Sciences
Track Building Technology
Climate Design & Computational Design Graduation Studio

Author:

Isabella Lucchese – 5258528

Mentors:

Azarakhsh Rafiee, *Geomatics, TU Delft*

Willem van der Spoel, *Building Technology, TU Delft*

University Supervisor:

Amin Jalilzadeh, *Geomatics, TU Delft*

Company Advisers:

Jonathan Ciurlanti *Digital Team, Arup Amsterdam*

Filique Nijenmanting *Building Physics Team, Arup Amsterdam*

Filippo Giunta *Digital Team, Arup Amsterdam*

Acknowledgements

During my master's studies in the Building Technology track at TU Delft, I received substantial support and encouragement from the AE+T faculty, whose commitment to innovation and sustainable development were crucial in advancing my academic development. Additionally, I found the Dutch educational approach very effective to intellectual growth; faculty members demonstrated a genuine openness to student perspectives and fostered collaborative discussions. This horizontal approach created an environment in which students and professors engaged as peers, which greatly enriched my analytical and critical thinking skills.

Before completing my studies, I had always aspired to explore the field of artificial intelligence, and I am profoundly grateful to the mentors who supported me in this endeavour. I want to extend my thanks to Azarakhsh Rafiee for her kindness and generosity in sharing her knowledge. I am also grateful to Willem van der Spoel, whose great advice elevated the quality of my project. Additionally, I thank Amin Jalilzadeh for his patience and readiness to assist me, always with a smile. My appreciation also goes to Jonathan Ciurlanti and Filique Nijenmanting from Arup, who, with their practical expertise, always encouraged me to create something with genuine value.

I extend my deepest gratitude to my family and friends, whose steadfast support has been foundational to my academic pursuits. Their encouragement provided the resilience necessary to deliver this project, making this journey both possible and profoundly meaningful.

Abstract

The building sector represents the highest share of operational energy consumption across all sectors, with a significant portion attributed to the inefficiency of the existing building stock. In this context, building retrofit plays a crucial role in enhancing energy efficiency and reducing environmental impact. However, conventional models for assessing retrofit scenarios are highly computationally expensive, thereby slowing down the retrofit process. This research addresses this challenge by developing an AI-based surrogate model using Multi-Task Learning (MTL). The proposed MTL model significantly reduces computational costs while simultaneously predicting energy consumption, costs, embodied carbon, and thermal comfort. Additionally, Multi-Objective Optimization (MOO) and Multi-Criteria Decision Making (MCDM) techniques are employed to select optimal retrofit solutions. Results demonstrate that the MTL model accelerates the retrofit simulation process from 90 minutes to just 2 seconds, highlighting its potential to streamline and enhance retrofit decision-making processes.

Keywords: building retrofit, energy efficiency, AI-based model, multi-task learning, surrogate models, computational efficiency, embodied carbon, energy consumption prediction, retrofit simulation

Table of Contents

Acknowledgements	3
Abstract	4
1. Research framework	7
1.1 Background	7
1.2 Problem statement	7
1.3 Research objectives.....	8
1.4 Research questions	9
1.5 Research methodology	9
1.6 Boundary conditions	11
2. Literature review	12
2.1 Retrofit policies	12
2.2 Retrofit process	15
2.3 Energy Simulation Models	15
2.3.1 Physics-Based Models	15
2.3.2 Data-based models	15
2.3.3 AI-based surrogate models in retrofitting	16
3. Base building and retrofit scenarios selection	17
3.1 Archetype selection	17
3.2 Retrofit scenarios selection.....	17
4. Weather file	22
4.1 Base file	22
4.2 Heat waves predictions.....	22
4.3 Future weather file creation.....	23
5. EnergyPlus simulation model	24
5.1 Model overview.....	24
5.2 Input data.....	24
5.2.1 Base IDF file	24
5.2.2 Retrofit parameters	28
5.2.3 Weather file.....	30
5.2.4 Output data.....	30
5.3 Intercomparison	31
6. Development of the AI-based surrogate model	32
6.1 Multi-Task Learning model overview.....	32
6.1.1 Overall workflow.....	32
6.1.2 Introduction to Multi Task Learning	34
6.1.3 Architecture of Multi-Task Learning.....	34
6.1.4 Benefits of Multi-Task Learning.....	35
6.1.5 Challenges in Multi-Task Learning	36
6.1.6 Dataset preprocessing.....	36

6.1.7	Input data	36
6.1.8	Output data.....	37
6.1.9	Dataset analysis and Interpretation	37
6.2	Architectures of the Model	39
6.2.1	Architecture Options Overview	39
6.2.2	Option 1: Task-specific architectures based on literature review	40
6.2.3	Option 2: Task-specific architectures based on data analysis.....	44
6.3	Training functions	48
6.3.1	Loss function	48
6.3.2	Early stopping mechanism	50
6.3.3	Weighted Sum Method.....	51
6.3.4	Multi-Gradient Descent Algorithm (MGDA)	51
6.4	Model evaluation	53
6.4.1	Model options evaluation	53
6.4.2	Evaluation of best-ranked model	57
6.4.3	Time efficiency	61
7.	<i>Selection of optimal retrofit solutions</i>	63
7.1	Multi-Objectives Optimization	63
7.1.1	Optimization based on dataset.....	63
7.1.2	Optimization based on constraints	65
7.2	Multi-Criteria Decision Making.....	66
7.2.1	Compromise programming	67
7.2.2	Pseudo-weights.....	67
8.	<i>Results</i>	69
8.1	Optimization based on dataset	69
8.2	Optimization based on constraints.....	70
9.	<i>Results discussion</i>	73
9.1	Optimization based on dataset	73
9.2	Optimization based on constraints.....	75
10.	<i>Conclusion</i>	77
10.1	Answering research questions.....	77
10.2	Research Limitations	79
11.	<i>Reflection</i>	81
12.	<i>References</i>	83
Appendix A	86
Appendix B	91

1. Research framework

1.1 Background

The large amount of greenhouse gases (GHG) we are releasing into the environment is driving significant climate change across the planet. As a consequence, we are witnessing a marked increase in the frequency and intensity of extreme weather events, a trend that is likely to worsen in the future. Events such as severe floods, strong heat waves, hurricanes, and persistent droughts are becoming increasingly common, presenting serious challenges to our society. These intensifying climatic disruptions will continue to strain our communities, economies, and ecosystems, requiring urgent efforts to mitigate their impacts and adapt to the changing climate.

1.2 Problem statement

According to the European Commission, the operational energy used in the building sector accounts for nearly 40% of total energy consumption and 36% of CO₂ emissions (Cuffe, 2020). Consequently, reducing the environmental footprint of the building sector is a crucial step toward achieving a more sustainable environment. Policymakers seem to have a strong understanding of the need to improve the energy efficiency of buildings, leading to the establishment of energy-saving targets at both national and continental levels. For instance, in Europe, these targets are set by the Energy Performance of Buildings Directive (EPBD), its subsequent amendments, and the most recent directive with a focus on building retrofits.

One of the key goals of the EPBD is for all new buildings to be nearly zero-energy buildings. While this is a necessary measure, it is not sufficient on its own to reduce the overall environmental impact of buildings. The implementation of advanced and efficient energy models to guide construction practices can ensure that new buildings operate efficiently and sustainably, potentially achieving nearly zero-energy status. However, constructing new energy-efficient buildings will only stabilize current levels of energy consumption and greenhouse gas (GHG) emissions, without contributing to their reduction.

The European building stock comprises a significant number of aging structures, with 35% of buildings being over 50 years old and a retrofit rate of less than 1.2% per year (Cuffe, 2020). In the Netherlands, approximately 46.8% of the buildings were constructed more than 50 years ago (ministerie van Binnenlandse Zaken en Koninkrijksrelaties, 2022). Therefore, to meaningfully reduce energy consumption and GHG emissions, special attention must be given to the existing building stock. Retrofitting existing buildings is one of the most promising strategies in this regard.

Building retrofits are often seen as optional decisions made by landlords, requiring substantial capital investment. Unfortunately, retrofitting is typically not a priority for low-income individuals. Without retrofitting, these residents not only continue to contribute to CO₂ emissions but also face numerous other consequences. For example, as heat waves become more frequent, indoor temperatures in their homes could reach dangerously high levels, increasing the risk of hospitalization and overburdening healthcare systems. The discomfort caused by high indoor temperatures can lead to serious health issues and poor living conditions, which in turn may have negative psychological effects and potentially increase crime rates among residents.

In the Netherlands, low-income families often reside in buildings managed by housing corporations. This means that these corporations have significant influence over a large portion

of the building stock and are responsible for the living conditions of many low-income families. Given that housing corporations can impact a considerable percentage of the building stock and have a vested interest in upgrading their properties to comply with new regulations, they emerge as key stakeholders in this project.

Building retrofit involves a wide range of interventions and it is considered a highly complex process, typically carried out by energy consultants. These consultants are equipped with specialized engineering tools to provide retrofit solutions to decision-makers. While these models offer great accuracy, they are also computationally expensive, which can slow down the entire process.

For this reason, it is crucial to develop computationally efficient tools that can enhance the process and simplify building retrofits. These models could significantly accelerate the identification and implementation of effective retrofit measures, reducing the time and resources required for analysis. By making retrofitting more accessible and scalable, particularly for housing corporations and smaller stakeholders, such tools could help target interventions with high environmental and social impact. This would support efforts to lower energy consumption and CO₂ emissions while potentially improving living conditions for vulnerable populations and contributing to long-term sustainability goals.

1.3 Research objectives

The primary goal of this project is to ensure good living conditions for low-income families while actively contributing to the reduction of the environmental crisis. This involves providing efficient, future-proof housing solutions that not only cater to the needs of low-income families but also reduce the CO₂ emissions associated with the built environment. To achieve this overarching goal, the project focuses on developing an AI-based model designed to make the building retrofit process faster and more accessible.

In pursuit of this goal, the following objectives must be accomplished:

A – Development of an AI-based surrogate model

This project aims to develop an AI-driven surrogate model to select optimal building envelope retrofit solutions for terraced houses, considering heat waves in the Netherlands. The choice of developing a surrogate model has been made since it offers a good trade-off between accuracy and computational efficiency. The model will be built using machine learning (ML) algorithms, specifically artificial neural networks (ANNs). The surrogate model will serve four primary purposes:

1. Environmental Impact: embodied carbon of materials used in envelope retrofit solutions;
2. Economic Impact: investment costs associated with the retrofit solution;
3. Thermal Comfort: indoor thermal comfort,
4. Energy consumption: annual energy consumption of the retrofitted building.

B – Development of a decision-making tool

Implement a tool that leverages Multi-Objective Optimization (MOO) and Multi-Criteria Decision-Making (MCDM) methodologies to identify optimal retrofit solutions. Within this framework, four objectives are minimized, with thermal comfort as the sole objective targeted for maximization. The "optimal" solution is determined by balancing the outcomes of the optimization process with the client's specific preferences and priorities.

C – Projections of heat waves

Project future heat wave patterns to better understand their potential impact on building performance and occupant comfort.

D – Identification of conventional envelope retrofit solutions

Identify and evaluate existing conventional envelope retrofit solutions to ensure they meet the project's environmental and economic goals.

1.4 Research questions

Given the problem statement mentioned above, the main research question of this project is:

How to develop an AI-based surrogate model to select optimal building envelope retrofit solutions for a terraced house in the Netherland considering the effect of heat waves in future weather?

To address the research question, the following sub-questions have been formulated:

- Which is the complete workflow of a model that selects optimal building envelope retrofit solutions considering heat waves in the Netherlands?
- Which type of AI is it better to implement for a surrogate model that identifies optimal building envelope retrofit solutions terraced houses?
- What are the heat waves projection data for the future?
- Which are commonly used building retrofit envelope solutions for a terraced house in the Netherlands?
- Which are optimal building envelope retrofit solutions considering heat waves in the Netherlands?

1.5 Research methodology

Literature research

To establish the overall framework of the project and identify a relevant research gap, a comprehensive state-of-the-art review is conducted across selected areas of research. Initially, an investigation into European retrofit policies, with a particular emphasis on the Dutch context, is undertaken. This is followed by an examination of the standard phases of the retrofit process. Subsequently, a focused study is carried out on the energy simulation models used in the retrofitting process. This includes an analysis of the advantages and limitations of the two main categories: physics-based models and data-based models. Lastly, specific attention is given to AI-based models employed in this context, exploring how they are typically integrated into the multi-criteria decision-making process.

Base building and retrofit scenarios selection

A typical Dutch building archetype is selected as the primary subject of analysis in this research, with potential expansion to other archetypes if time permits. Details regarding the geometry, materials, and systems of the base building are then defined, alongside reliable sources to guide the selection of common retrofit options for this archetype. Decisions are made regarding which costs associated with each retrofit option will be considered. Additionally, a choice is made about which aspects of environmental impact will be included in the analysis. Following this, reliable sources for both costs and environmental impacts associated with each retrofit scenario are identified. A final list of retrofit scenarios is created, specifying the thermal properties, technical details, cost and embodied carbon of each intervention.

Weather file generation

A location for the analysis building is selected, along with a base weather file corresponding to that location. Decisions are also made regarding the specific time horizons to be included in the

analysis. Heat wave projections for the chosen location and selected time horizons are researched. Finally, the base weather file is modified to incorporate heat wave data for each time scenario.

EnergyPlus simulations

To perform simulations with EnergyPlus, detailed information on the retrofit scenarios, the weather files for analysis, and the IDF file of the base building are required. Initially, a Grasshopper script is created using LadyBug Tools and Honeybee to generate the IDF file for the base model under analysis. Once this is completed, all necessary inputs for running the EnergyPlus simulations are prepared. For each retrofit scenario, the model generates specific IDF files, and simulations are conducted for each weather file. The desired outputs from these simulations are daily energy consumption and maximum indoor temperature.

AI-Based surrogate development

The most suitable type of AI for the research is selected, with a detailed overview of input and output features. Specific libraries are imported, and the input dataset is preprocessed. The architecture of the selected model is then implemented, and the model is trained, validated and tested. After that, the implementation and testing of the Multi-Objective Optimization to ensure it functions as intended.

Decision-making tool development

The most suitable Multi Objectives Optimization algorithm and Multi Criteria Decision Making methodology are selected and implemented in the workflow.

Results and conclusions

The model is run multiple times, applying various constraints within the MOO and MCDM. The resulting outputs are then analyzed and interpreted. The study concludes with a discussion and evaluation of the tool development process, an examination of its limitations, and recommendations for future improvements.

The overall research methodology is summarized and presented in Figure 1.

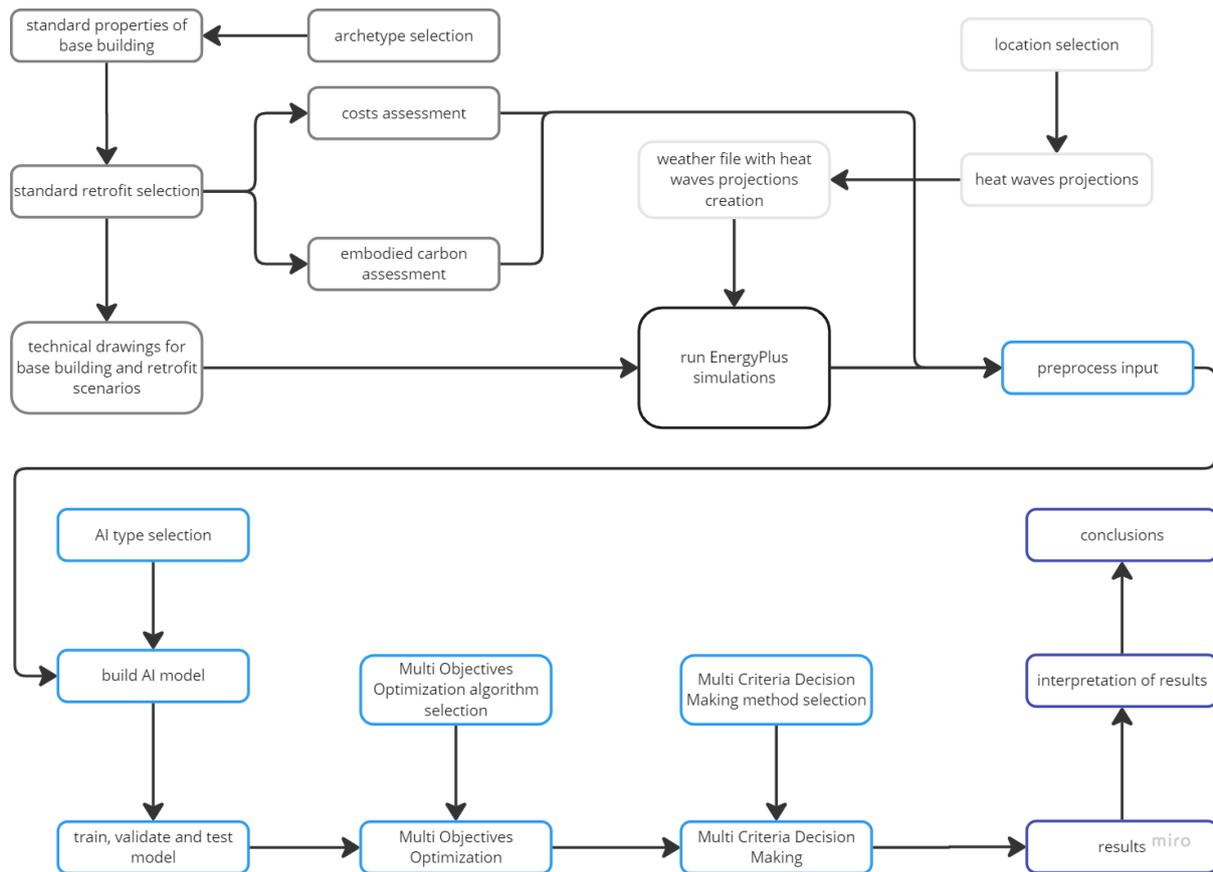


Figure 1 - Overall research methodology (Source: Own Source)

1.6 Boundary conditions

The retrofit scenarios examined in this study are confined to modifications of the building envelope, focusing specifically on the façade, ground floor, windows (both glass and frame), and roof. These interventions are primarily characterized by the addition of insulation to the existing structural components, with the exception of the window retrofits, which involve the replacement of the glass and the frame across all windows in the building.

These retrofit measures are designed for a mid-unit terraced house, representative of those constructed between 1946 and 1964. The final selection of the optimal retrofit solution utilizes Lelystad as the reference location. Furthermore, the analysis does not incorporate the influence of the surrounding building context on the performance outcomes.

2. Literature review

2.1 Retrofit policies

The European and Dutch building stock and related energy consumption

The energy inefficiency of many European buildings largely stems from the relatively late establishment of energy-related regulations and functional standards across various countries. The earliest building energy codes were introduced in the 1960s within Scandinavian nations (Furtado et al., 2023). Gradually, other European countries developed and updated these codes, particularly in response to European directives. From an energy perspective, the European Union released its initial Energy Performance of Buildings Directive in 2002 (Economidou et al., 2020). By that time, however, most European countries had already developed their national energy standards, with initial versions dating back to the early 1990s. Consequently, much of the masonry building stock across Europe was constructed before energy efficiency requirements were implemented.

Significant energy efficiency issues in European buildings are mainly due to inadequate insulation in building envelopes and poor thermal characteristics in façade and windows. Furthermore, inefficient heating, cooling, and water-heating systems in older buildings contribute to substantial energy consumption. Outdated lighting systems further exacerbate energy use and related CO₂ emissions.

(European Commission, 2020) highlights that energy use for space heating accounts for the largest share in the residential sector, with an average of 67.74% of total energy consumption across Europe. In colder climates such as those of the Netherlands, Denmark, and Finland, the need for heating is closely linked to the climate, whereas in moderate climates, like those in Italy and Croatia, inefficiencies in building design and insulation play a significant role.

Final household energy consumption for heating in 2020 (see Figure 2) underscores the need for some countries to improve the efficiency of their building stock. The graph shown shows that the Netherlands, together with Germany, Italy, Spain, France and Poland, rank above the European average consumption line of 7,618.

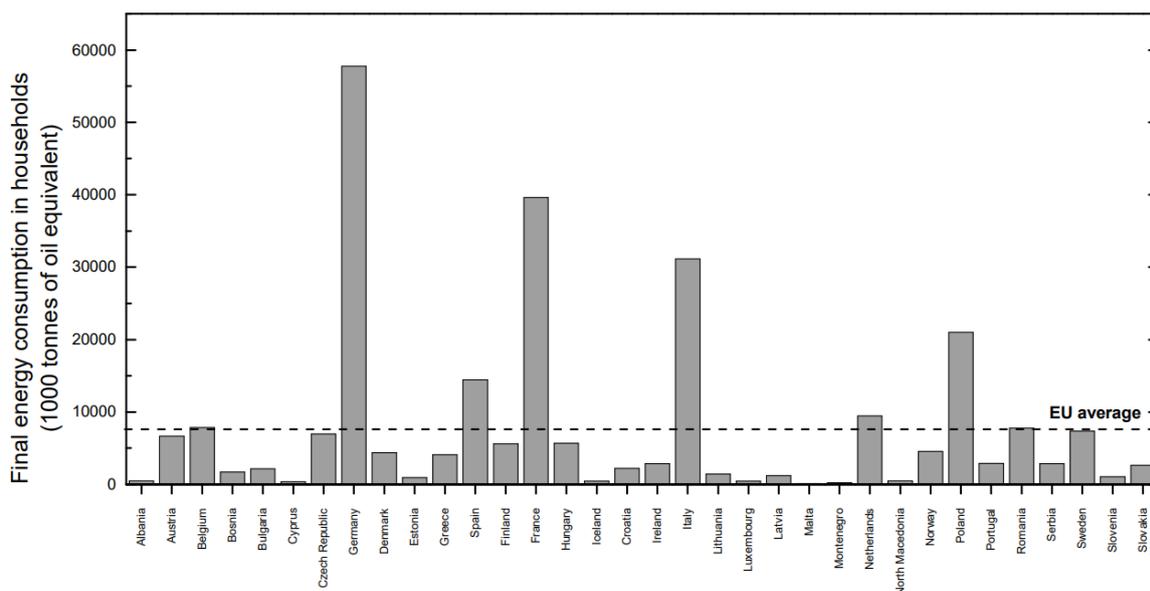


Figure 2 Final energy consumption in households in 2020 (Source: Furtado et al., 2023)

In the specific context of the Netherlands, useful data are provided by the European Union (n.d.). As indicated in Figure 3, the residential sector in the Netherlands had a final energy consumption of 423.45 PJ in 2021, which was significantly higher than that of the service sector. Furthermore, Figure 4 illustrates that the direct greenhouse gas (GHG) emissions from the residential sector amounted to 17.54 Mt CO₂eq in the same year.

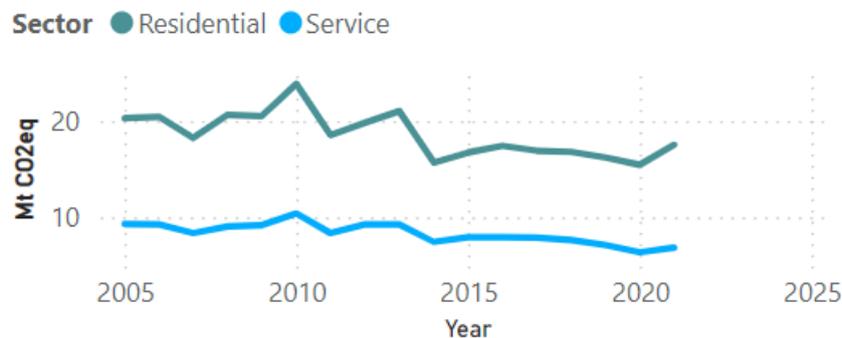


Figure 3 - Final energy consumption in the Netherlands from residential and service sectors (Source: Eu, n.d.)

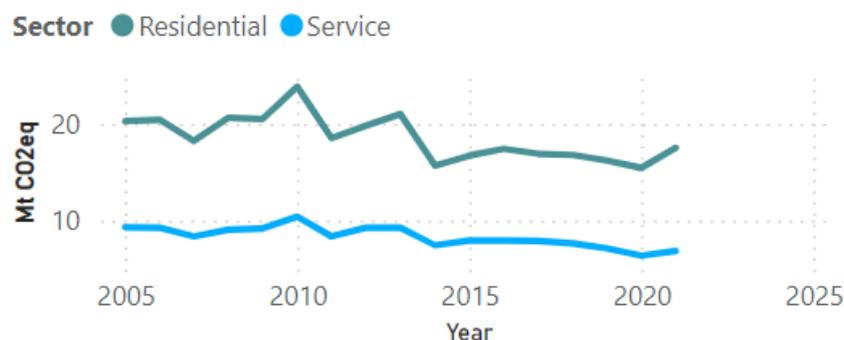


Figure 4 - Direct GHG emissions in residential and services sectors in the Netherlands (Source: Eu, n.d.)

Despite these high values, data from the European Union (n.d.) indicate that in 2022, 5.3% of the total Dutch population was unable to keep their homes adequately warm during the colder periods of the year. Additionally, approximately 14.8% of the population lived in dwellings suffering from structural issues such as leaking roofs, damp walls, or rotten window frames and floors.

These statistics, both at the Dutch and European levels, underscore the urgent need to address the energy inefficiencies of buildings. Energy retrofitting emerges as a promising solution to not only reduce energy consumption but also to lower related greenhouse gas emissions. In response, many European states are now implementing incentives for citizens to renovate their homes to enhance energy performance.

European policies to lower down GHG through incentivizing building retrofit

The European Union is actively committed to achieving a carbon-neutral building stock, acknowledging the significant role that the built environment plays in greenhouse gas (GHG) emissions. To this end, the EU has developed and enacted a series of regulatory frameworks and policies aimed at reducing emissions from buildings, with a special emphasis on promoting retrofitting measures.

A key element of these efforts is the European Green Deal, which aims for a substantial reduction in net GHG emissions, reaching climate neutrality by 2050. As a key component of

this initiative, the Commission proposed a renovation initiative in 2020 to bring together stakeholders from the construction and architecture sectors, local governments, and other relevant parties. This initiative includes innovative financing schemes under InvestEU, specifically targeting housing associations and energy service companies. By organizing renovation efforts into larger projects, the initiative aims to benefit from economies of scale, thereby reducing costs and improving access to financing. A particular focus is placed on social housing renovations, which are essential for supporting low-income households facing high energy costs.

In addition to this, the Fit for 55 legislative package reinforces the EU's 2030 climate target, seeking a 55% reduction in emissions (Council of the European Union, n.d.). It establishes a comprehensive framework for transforming the building sector by promoting energy efficiency, renewable energy integration, and sustainable heating systems. This package includes the updated Energy Efficiency Directive (EED) and Renewable Energy Directive (RED), both of which set binding targets for energy efficiency and renewable energy adoption in buildings, focusing on renovations to achieve significant emissions reductions. These targets are critical to supporting the transition to a low-carbon economy, where retrofits of existing structures play a central role in reducing energy consumption.

The Social Climate Fund, introduced under the Fit for 55 package, further supports these efforts. With a budget of approximately €86.7 billion from 2026 onwards, this fund aims to assist vulnerable communities and low-income households in meeting the costs of the green transition. Each EU member state is tasked with creating a national plan to allocate these funds, which could include financing for improved insulation, efficient heating systems like heat pumps, and other upgrades. By supporting those most affected by energy costs, the Social Climate Fund is intended to foster equity in the EU's climate agenda, ensuring that vulnerable people are not left behind in the transition.

A notable expansion within Fit for 55 is the adaptation of the EU Emissions Trading System (EU ETS) to include sectors such as buildings, road transport, and fuel. This new, self-standing ETS creates a pricing mechanism for carbon emissions in the building sector, which incentivizes property owners to invest in energy-efficient renovations. By embedding emissions from buildings in the ETS, the EU aims to create market signals that promote greener building practices, helping to reach both the 2030 and 2050 climate targets.

In the Netherlands, policies are in place to meet national goals that complement EU-wide strategies. The Dutch government has introduced the Energy Performance Subsidy Scheme (EPS), which provides grants to homeowners for improving insulation and installing energy-efficient systems. Additionally, the Netherlands has set a target to make all buildings energy-neutral by 2050, a goal that reflects the EU's broader climate ambitions. Through regional energy strategies and local partnerships, the Netherlands actively promotes building retrofits to meet national and EU targets, demonstrating its commitment to the shared vision of a sustainable, carbon-neutral future.

To date, numerous strategies have been implemented to encourage building retrofits, a trend that is expected to grow in the future. This expansion will be supported not only by increasing funding to carry out retrofit measures but also by lowering the accepted thresholds for building energy consumption.

2.2 Retrofit process

The building retrofit process typically encompasses a series of structured phases. As outlined by Deb et al. (2021), a generic retrofit process is divided into five key phases. The first phase involves a pre-retrofit survey, during which the scope and objectives are established in collaboration with the building owner. In the second phase, energy audits and performance assessments are conducted to identify areas of energy inefficiency and potential savings. The third phase is focused on exploring retrofit options, utilizing energy models alongside economic analysis and risk assessment tools. Implementation and commissioning follow as the fourth phase, and finally, the validation and verification phase ensures the projected energy savings align with actual performance.

Energy simulation tools are crucial in the third phase, where they enable stakeholders to evaluate different retrofit options before implementation. These tools simulate how various retrofit measures affect the building's energy use, facilitating informed decision-making by predicting performance outcomes and optimizing retrofit strategies.

2.3 Energy Simulation Models

Energy simulation models for building retrofitting fall into two primary categories: **physics-based models** and **data-driven models**. Each category with its own set of strengths and limitations.

2.3.1 *Physics-Based Models*

Physics-Based Models use specific thermodynamic principles to simulate the energy flow within a building. They rely on precise physical parameters like material thermal properties, environmental factors, and building geometry to calculate energy performance. These models provide robust frameworks for studying the impact of retrofit options on building envelope, mechanical, and electrical systems. They are especially useful in retrofit projects involving complex interactions between various building components and environmental conditions. Crawley et al. (2018) conducted a comprehensive comparison of 20 widely used physics-based energy modelling tools, from which three primary software solutions emerged as the most prevalent. EnergyPlus stands out as a versatile and extensively used engine, commonly integrated within platforms such as OpenStudio and BEopt, making it well-suited for detailed energy simulations in building retrofitting. Another widely adopted engine, DOE-2, is incorporated into tools like eQuest and GBS, and is valued for its ability to achieve a balance between simulation speed and accuracy. TRNSYS is particularly known for its adaptability in modelling complex systems, making it a preferred choice for dynamic energy simulations, especially in buildings with unique HVAC configurations. Additionally, specialized tools like IES-VE and TAS enable customized simulations and are frequently utilized in niche retrofit applications or projects with specific system requirements.

While physics-based models are invaluable for their precision, they face limitations. These models often require extensive and detailed building data, which can be challenging to obtain, especially for existing structures. Moreover, they are computationally demanding, which significantly delays the assessment of retrofit scenarios. There is also a gap in modelling capacity for innovative retrofit technologies, such as adaptive building materials or smart HVAC systems, which require dynamic, real-time performance adjustments. (Kamel et al., 2019).

2.3.2 *Data-based models*

Unlike physics-based models, data-driven approaches rely on patterns and correlations from historical or synthetic data rather than physical laws. By establishing mathematical

relationships between variables, these models handle large data sets efficiently, making them well-suited for rapid simulations across diverse climate zones and building types.

Data-driven models are particularly valuable in scenarios where detailed building data is unavailable or when accelerated simulations are needed, such as evaluating multiple retrofit options. Black-box models, for instance, can draw on pre-existing databases to generate reliable energy estimates without complex input parameters.

Although they may lack the granularity of physics-based models, data-driven approaches are adaptable and efficient, offering a practical alternative for energy estimation when time and data constraints are present. (Deb et al., 2021)

2.3.3 *AI-based surrogate models in retrofitting*

Machine Learning (ML) has become a fundamental part of modern life, embedded in applications that range from personalized recommendations and speech recognition to complex systems like autonomous vehicles and market forecasting. Its ability to process large volumes of data and identify patterns has made ML indispensable in diverse fields, enabling innovations that continue to reshape how we interact with technology and make decisions.

Another promising topic in machine learning is building retrofitting. However, it can be a challenging approach due to the limited availability of consistent and reliable building data. Many studies integrate ML within the traditional engineering-based retrofit models. For instance, Magnier and Haghghat (2009) pioneered a hybrid optimization approach for building design by combining Artificial Neural Networks with Genetic Algorithms. In their method, ANNs were trained to predict building energy demand, which was then used in GAs to assess design solutions based on total energy consumption and thermal comfort. Building on this, Asadi et al. (2014) adapted the method for building retrofits, using synthetic data for single and multi-objective optimization. This approach not only identified optimal envelope and system retrofits but also explored how various design variables impact outcomes, focusing on energy consumption and retrofit costs. Ascione et al. (2017) further refined this approach by optimizing both energy systems and building envelopes, targeting energy use and comfort hours. They included a sensitivity analysis to reduce the ANN model inputs, improving simulation efficiency. Since GA-based retrofits often require extensive simulations (Costa-Carrapiço et al., 2019), methods to reduce simulation demands have emerged. Prada et al. (2018) combined support vector machines with GAs to develop efficient retrofit strategies, while Safarzaghegan Gilan et al. (2015) employed an active learning Gaussian process alongside GAs to optimize building designs. Active learning in this context minimized the sample size needed to train models for energy demand prediction, resulting in a faster process than traditional simulation-optimization methods. However, this approach lacked a comparative reliability metric. Yuan et al. (2019) used Gaussian processes as a meta-model to predict energy demands across retrofit scenarios, which were subsequently ranked for cost-effectiveness.

A key objective in coupling ML with optimization models is to reduce simulation costs, highlighting the need for computationally efficient ML models in this research. Literature suggests that Multi-Task Learning (MTL) models may be a strong candidate due to their ability to predict multiple outcomes simultaneously by utilizing shared and task-specific layers (Zhang & Yang, 2017). This approach could be particularly valuable in the context of this research, where a single MTL model could predict energy consumption, costs, embodied carbon, and comfort levels concurrently, avoiding the need to create and run four separate models. Since no prior research has been found that integrates an MTL model with an optimization algorithm in this context, this study aims to implement and assess its effectiveness and potential.

3. Base building and retrofit scenarios selection

3.1 Archetype selection

The primary focus of this study is the terraced house, selected for its representation of a typical residential building in the Netherlands. Terraced houses typically consist of a row of 5 to 7 units, with the units at each end of the row differing slightly in terms of materials and energy efficiency from those in the middle. Given that most units are located in the central part of the row, this research will specifically concentrate on one of these central units.

The differences in the shape, materials, and energy efficiency of these houses depend largely on when they were built. Therefore, it was necessary to choose a specific period for the buildings being studied to get accurate baseline data for this research. The (ministerie van Binnenlandse Zaken en Koninkrijksrelaties, 2022) document was a key resource in making this choice, since it provides detailed examples of typical homes in the Netherlands. The section on terraced houses in this document is divided into eight parts, each part focusing on buildings from a specific time period, ranging from before 1945 up to 2018.

Given that housing corporations, which lease units to families in need, are a primary focus of this study, it was assumed that the typical buildings leased by these corporations were constructed between 1946 and 1964. This period is significant because approximately 45 percent of these units are leased, presumably mostly by housing corporations. Buildings from this era typically have lower energy efficiency compared to structures built in other periods, highlighting a pressing need for retrofit interventions (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, 2022). Therefore, this study uses a terraced house unit built between 1946 and 1964 as the standard example for analysis.

3.2 Retrofit scenarios selection

Subsequently, TABULA WebTool (n.d.) was taken into consideration. The TABULA Webtool (n.d.) is a platform co-funded by the EU, aimed at standardizing residential building typologies and their energy performance across Europe. It offers detailed data, standardized calculation procedures, and practical examples of energy-saving measures. This source was crucial for defining the details of the base building and its associated retrofit scenarios. Once the archetype of the terraced house built from 1946 to 1964 was selected, it was possible to access a table summarizing the type of construction, a technical detail image, and the U-values for the roof, walls, floors, and windows. These data are available for the building in three scenarios: existing state, usual refurbishment, and advanced refurbishment. Usual refurbishment is defined as adding insulation to meet current standards, while advanced refurbishment aims for insulation near Zero Energy Building (ZEB) standards. This source enabled the definition of the U-values for envelope parameters (ground floor, façade, roof, windows) for both the existing state and the retrofit scenarios, categorizing the retrofit options for each component between current standard and nZEB.

Since one of the objectives of the research deals with embodied carbon of retrofit measures, it was decided to include for each type of retrofit scenario both a commonly used insulation material and one with a low embodied carbon value. Consequently, for each of the four building envelope parameters - ground floor, façade, roof, windows - there are two types of retrofit -

current standards and nZEB standards - and for each type of retrofit, there are two insulation options - a commonly used and an environmentally low-impact insulation material.

Materials and build ups of the base building

Concerning the existing state, the stratigraphy and thermal properties of the four building envelope components were sourced from the TABULA WebTool (n.d.). Specifically, the technical detail images were determined based on the engineering interpretations provided by the author under the supervision of ARUP expert Filique Nijenmanting. Thus, the building is assumed to have a non-insulated tiled pitched roof with structural timber elements and an air cavity, a solid clay-brick façade, a non-insulated timber ground floor with an air cavity, and wood-framed double-glazed windows. The U-values for these components are respectively 2.08 W/m²K, 2.22 W/m²K, 2.44 W/m²K, and 2.9 W/m²K.

Selection of material and calculation of cost for commonly used scenario

Given the threshold U-values defined by the TABULA Webtool (n.d.) for retrofit interventions, the type of retrofit measure and its cost were determined using Regelhulpen voor bedrijven (n.d.), a platform developed by the Netherlands Enterprise Agency. By selecting the building type, the envelope component to be modified, and the desired thermal performance, the website suggests the type of intervention to be executed and provides a cost estimate per square meter. The cost indicators consist of four items: materials, equipment, labor, and storage.

The interventions considered in this research are those suggested by the platform in the case of a residential building, multi-family house, medium size of building, and single approach. Each retrofit scenario was selected based on the type of existing build-up, and the intervention with the U-value closest to the threshold suggested by TABULA Webtool (n.d.). It is important to note that for the retrofit interventions related to the façade, external insulation measures were selected. This choice is due to external insulation being more suitable for a residential building with a solid brick façade. External insulation slows down the transmission of external temperatures into the building, keeping interiors cooler during the day and warmer at night. This delay in thermal transfer makes it also particularly effective for heat wave resilience, as it prevents extreme outdoor temperatures from quickly impacting the internal environment, thus maintaining a more stable and comfortable indoor climate.

The index from Regelhulpen voor bedrijven (n.d.) of each specific scenario considered in this research, along with the related insulation material used and the cost, is shown Figure 7.

Selection of material and calculation of cost for low embodied carbon scenario

The article by Cosentino et al. (2023) includes the graph shown in Figure 5. From this graph, it is evident that straw bale insulation has the lowest emission values among the insulations considered, with hemp fiber insulation being the second lowest. After researching market availability in the Netherlands and consulting with experts at ARUP, who contributed to this project, hemp fiber insulation was found to be more commonly used and readily available. For this reason, this material was chosen for consideration in all low environmental impact retrofit scenarios. Specifically, the thermal performance and cost of this insulation were taken from (Thermohanf Thermo Hennep Combi Jute (n.d.)).

The total cost calculated for each retrofit option that employs hemp fiber insulation in this research is based on the same values for equipment, labor, and storage as those used in retrofit scenarios with typical environmental impacts. However, this cost estimate for the hemp fiber retrofit option does not include the price of commonly used insulation materials. Instead, it specifically accounts for the cost of the recommended hemp fiber insulation, Thermohanf Thermo Hennep combi Jute (n.d.), which is chosen for its thickness that satisfies the required thermal performance standards.

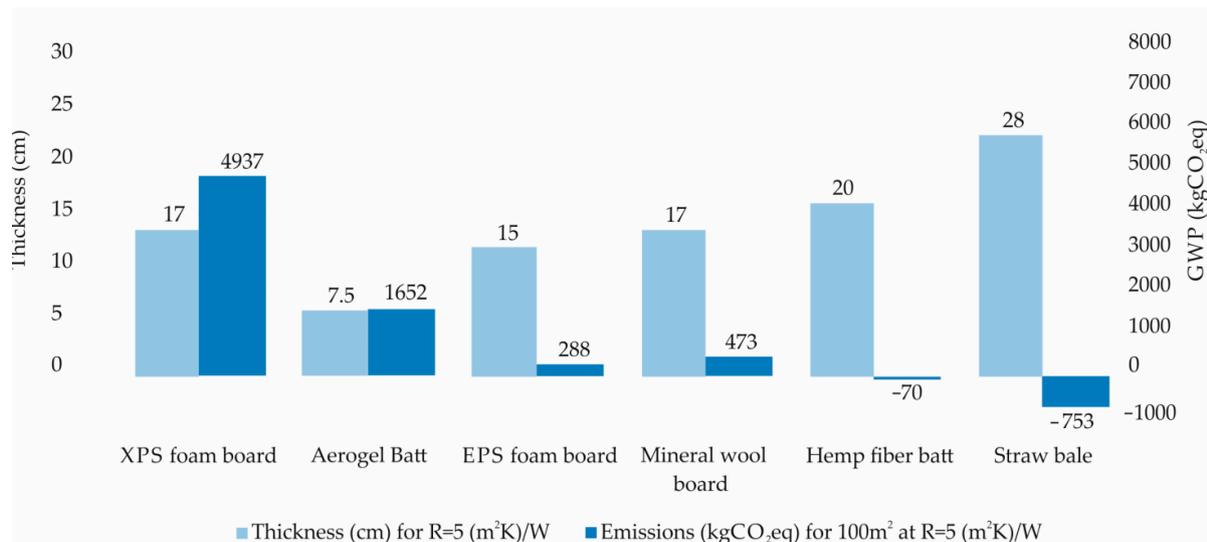


Figure 5 – Insulation materials carbon emissions X thickness for a thermal resistance of 5 (m²K)/W (Source: Cosentino et al.)

Calculation for Global Warming Potential of retrofit scenario

The environmental impact considered in the research for the retrofit scenarios relates to the embodied carbon of the insulation materials used. Specifically, it concerns the Global Warming Potential (GWP) expressed in kgCO_{2e}/m². The embodied carbon is calculated from the total CO_{2e} emissions produced during the Life Cycle Assessment phases A1-A3 (raw material supply, transport, manufacturing) of the material's life cycle (see Figure 6). This data for each retrofit option has been sourced from Bsr (n.d.), a database with many Environmental Product Declaration of common construction. A final table that summarizes all retrofit scenarios, U-values, insulation materials, their thickness, the cost of the retrofit, and the GWP is also displayed in Figure 7. From this figure, it is noteworthy that the embodied carbon consistently increases when retrofitting to near Zero Energy Building (ZEB) standards compared to current standards. This occurs because achieving higher energy efficiency requires the use of more insulating materials, which in turn elevates the embodied carbon. Additionally, it is observed that, for the same thermal performance threshold, scenarios utilizing low-impact materials are always more costly than those employing standard materials.

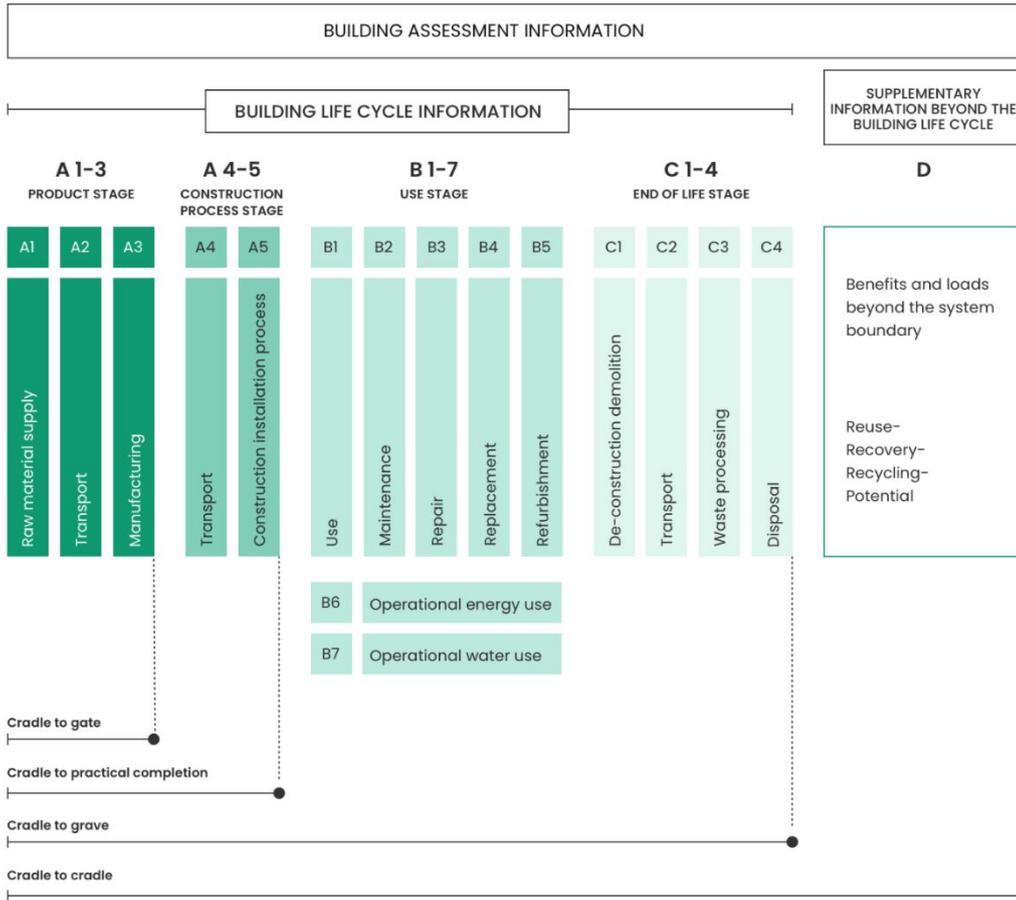


Figure 6 - Life-cycle stages (Source: Adapted from Life Cycle Stages by One Click LCA (n.d.), retrieved from <https://oneclicklca.zendesk.com/hc/en-us/articles/360015064999-Life-Cycle-Stages>)

	component	U-Factor (W/m2K)	level of retrofit	environmental impact	material	thickness of insulation (mm)	cost (euro/m2)	embodied carbon (kgCO2/m2)
existing state	floor	2,44	\	\	massive wooden floor, no insulation	\	0	\
	façade	2,22	\	\	solid walls	\	0	\
	roof	2,08	\	\	tiles - wook panel - air - wood panel	\	0	\
	window	2,9	\	\	wooden frame, double glazing	\	0	\
retrofit	floor	0,26	current standard	typical	PIR insulation - WB002f (U=0,21)	100	59,7	10
				low	hemp fiber insulation	174	77	5,92
		0,18	to nZEB	typical	resol insulation - WB002h (U=0,18)	100	87,93	11,18
				low	hemp fiber insulation	206	108	7,00
	façade	0,25	current standard	typical	EPS insulation - WB270 (U=0,25)	120	182,2	9,36
				low	hemp fiber insulation	142	179	4,83
		0,18	to nZEB	typical	EPS insulation - WB224 (U=0,15)	220	200	17,16
				low	hemp fiber insulation	249	222	8,50
	roof	0,25	current standard	typical	mineral wool - WB230 (U=0,22)	140	89,54	23,29
				low	hemp fiber insulation	181	105	4,76
		0,15	to nZEB	typical	PIR insulation - WB167 (U=0,12)	185	101,96	18,5
				low	hemp fiber insulation	314	139	10,68
	window	1,8	current standard	typical	WB221 - HR++ double glazing, plastic frame (U=1,2)	\	184	70
				low	WB252 - HR++ double glazing, wooden frame (U=1,2)	\	485	50
		1	to nZEB	typical	WB220 - triple glazing, plastic frame (U = 0,8)	\	295	150
				low	WB161b - triple glazing, wooden frame (U = 0,8)	\	622,5	120

Figure 7 – Summary of thermal technical details, cost and embodied carbon of retrofit and existing scenarios of four envelope parameters (Source: Own Work)

Technical details

Technical details related to the existing state and each retrofit option were developed based on images sourced from *TABULA WebTool* (n.d.). To ensure that these details do not have condensation issues, they were double-checked using *ubakus.de | Graphical Editor* (n.d.). These details are illustrated in Appendix A.

4. Weather file

4.1 Base file

Given the need to consider future heat wave conditions and the requirement for a EnergyPlus Weather File (EPW) for Energy Plus simulations, it was necessary to select a specific location for the building under study. As mentioned previously, the building archetype in this research is the terraced house. Therefore, we chose to investigate the municipality with the highest percentage of terraced houses in the Netherlands. According to a map shown by the Centraal Bureau voor de Statistiek (2023), which details the distribution of dwelling types by municipality in the Netherlands, Lelystad has the highest proportion of terraced houses, at 66.8%. After selecting this location, the corresponding weather file was downloaded from [Climate.onebuilding.org](https://climate.onebuilding.org) (n.d.), a provider of weather files worldwide.

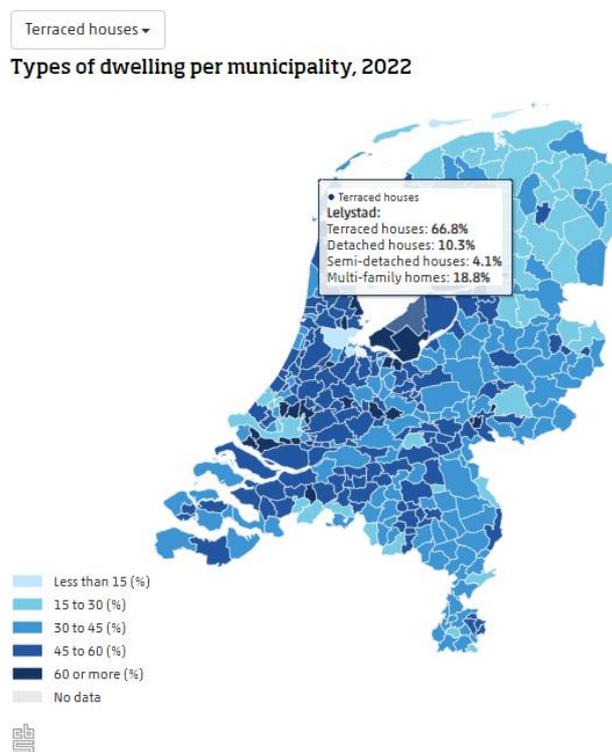


Figure 8 - Types of dwellings per municipality in the Netherlands (Source: Centraal Bureau voor de Statistiek, 2023)

4.2 Heat waves predictions

The Royal Netherlands Meteorological Institute (KNMI) is the Dutch national weather service that focuses on forecasting weather, monitoring climate, and seismic activity. KNMI has conducted significant research on climate change and future weather predictions, such as the "KNMI '23 Climate Scenarios". This document investigates the impacts of climate changes and their projections, offering detailed analyses of expected changes in temperature, precipitation, sea-level rise, and extreme weather events. These projections provide critical data that supports the development of adaptation and mitigation strategies. In addition to this, KNMI has developed the KNMI Klimaatscenario's (n.d.) tool that offers advanced features to explore future climate projections by transforming historical weather data into scenarios indicative of

future climates. Once a user selects a location, a time horizon, the historical meteorological data to use as a basis, and an emissions scenario, the tool provides future climate projections.

According to KNMI, a heat wave is defined as a period when the maximum temperature reaches at least 25°C for at least five consecutive days, which must include at least three days with temperatures exceeding 30°C. Given this definition, to project future heat waves, it was necessary to obtain predictions of daily future temperatures. The KNMI Klimaatscenario's (n.d.) tool was therefore used to access data on daily minimum and maximum temperatures.

Even though it is uncertain how much global efforts will reduce emissions, the projections used in this study considered the high CO₂ emissions scenario (H). This choice was made because it accounts for the worst-case scenario, ensuring that the effects of climate change are not underestimated. Specifically, the Climate Scenario Tool employed the Hd scenario, which corresponds to high CO₂, dessicant. Additionally, the transformation of historical weather data from 2020 has been used to generate the considered climate projections. This particular year was selected because it was the hottest year in the last hundred years, an important characteristic when using its weather data to estimate future heat waves. The time horizons considered in this research were the current scenario, and the 2050 and 2100 scenarios, aiming to analyse the present state, a near future scenario, and a distant future scenario.

4.3 Future weather file creation

Temperature projections for the two time horizons were analyzed, particularly by determining the number and duration of heat waves. These data are available for the three years considered in Appendix B. From those tables, it is observed that in Lelystad in 2020, one heat wave was recorded, lasting 12 days. The predictions from KNMI indicate that in 2050 there will be 2 heat waves, one lasting 5 days and the other 16 days, and in 2100 there will be 3 heat waves, lasting 16 days, 9 days, and 25 days respectively. KNMI projections clearly show a significant increase in both the number of heat waves and their duration over the considered time horizons.

The weather files in EPW format were generated by taking the base file downloaded from a [Climate.onebuilding.org](https://climate.onebuilding.org) (n.d.) and replacing the maximum and minimum hourly temperatures of the base weather file with the values from KNMI's predictions. This replacement was done for all the days identified as heat wave days by KNMI. By executing this procedure with the projections for both 2050 and 2100, weather files in EPW format were obtained that include heat wave projections for both future time horizons.

5. EnergyPlus simulation model

5.1 Model overview

A model taken from GitHub (Amin-Jalilzadeh-TU, n.d.) is utilized to conduct EnergyPlus simulations and generate the dataset for training the AI-based model. This model operates within the Visual Studio Code environment and employs the EnergyPlus engine alongside libraries such as Pandas and Eppy. The process begins with the input of the **base IDF file**, **EPW files** for the weather scenarios, and the **parameters that will vary for each retrofit scenario**, focusing particularly on the properties of the envelope components and infiltration levels. Primarily, the model functions by **creating IDF files for every retrofit scenario**, using the given IDF as base and changing the retrofitted parameters. Subsequently, **it executes all the IDF files using EnergyPlus** as the simulation engine, using first 2020 weather file, then 2050 and then 2100. After the simulations, it **merges the results into a single Comma-Separated Values(CSV) file for each weather scenario**, ultimately producing a CSV file that contains all simulations for each scenario. The outputs include daily simulations for gas consumption, which accounts for energy used by the heating system and domestic hot water system. They also include daily building electricity, representing the electricity consumed by lighting, equipment, and appliances. Daily facility electricity is also included, covering the electricity used by HVAC systems, such as fans, cooling systems, and heat pumps. Additionally, it includes data on the maximum daily indoor temperature. The maximum indoor temperature is a critical factor in this project as it aims to enhance resilience to heat waves, making it an essential metric for evaluating potential overheating and thermal discomfort during such events.

5.2 Input data

To run the simulations, the necessary inputs include the IDF file of the base building, the parameters that will change for each retrofit scenario, and the weather scenarios.

5.2.1 Base IDF file

As previously mentioned, the base building considered in this research is an in-between terraced house unit built between 1964 and 1974. The **IDF file** for this building was **created using** Rhinoceros, along with **Grasshopper**, Ladybug, and Honeybee Tools. All **data used for modelling the base building are detailed in Table 1**; unspecified values in the table default to LadyBug Tools parameters. Primary data sources include (ministerie van Binnenlandse Zaken en Koninkrijksrelaties, 2022), (TABULA WebTool, n.d.), (Stichting Koninklijk Nederlands Normalisatie Instituut, 2024), and (ASHRAE, 2021). Notably, the assumptions regarding the thickness of each envelope component are based on the technical details discussed in Chapter 3.

		unit		legend	
geometry	year	-	-	1946 - 1964	
	rooms	-	-	3	source
	floors	-	-	3	voorbeeldwoningen 2022
	total gfa	m ²	-	98	TABULA
	height	m ²	-	9	NTA 8800
	ground floor	m ²	-	44	assumption
	closed façade	m ²	-	40	2021 ASHRAE Handbook
	pitched roof	m ²	-	48	Stichting Bouwresearch. (1981).
	windows	m ²	-	18	
	orientation	*	-	0	
	location	-	-	-	Lelystad
materials	ground floor	thickness	m	0,12	
		thermal resistance	m ² K/W	0,41	
		density	kg/m ³	370	
		conductivity	W/MK	0,29	
		specific heat	J/kg*K	1699	
	roof	thickness	m	0,18	
		thermal resistance	m ² K/W	0,48	
		density	kg/m ³	437	
		conductivity	W/MK	0,37	
		specific heat	J/kg*K	1511	
	external wall	thickness	m	0,3	
		thermal resistance	m ² K/W	0,45	
		density	kg/m ³	1800	
		conductivity	W/MK	0,66	
		specific heat	J/kg*K	840	
	internal wall	thickness	m	0,2	
		thermal resistance	m ² K/W	0,25	
		density	kg/m ³	1800	
		conductivity	W/MK	0,66	
		specific heat	J/kg*K	840	
glass window	U Factor	W/m ² K	2,9		
	Solar Heat Gain Coefficient	-	0,7		
infiltration		m ³ /s m ²	0,0005		
HVAC	heating system	heating type	-	local heating with gas boiler incl. exhaust	
		heating device	-	radiators	
	lighting system			not taken into account	
	electric equipment			not taken into account	
	ventilation system	ventilation type	-	natural ventilation supply with mechanical exhaust	
		heat recovery system	-	no	
		flow per area	l/s m ²	0,042	
hot water system		-	boiler		
cooling system		-	no		
Occupant and control	heating setpoint		°C	20°C between 10 and 20 16°C between 20 and 10	
	schedule			Ladybug standard	

Table 1 - Details of base building used in the Grasshopper Script (Source: Own Work)

The Grasshopper script was created to generate a building energy model, enabling the creation of an IDF file usable in EnergyPlus simulations. While this script can also run simulations, it was not used as the main tool to produce the training dataset for the AI-based model because it requires modelling and running each retrofit scenario individually. In contrast, with the EnergyPlus simulation model used in this research, the script only needs to be executed once to obtain results from all retrofit scenarios for each weather scenario. The Grasshopper script follows a structured workflow that starts with defining key input parameters, continues through geometric modelling and material property assignments, and ends with the export of a fully defined model file.

At the beginning of the script, a square with a length of 4.7 meters and a depth of 7 meters is created to represent the base of the building. This square is extruded upward by 6 meters, forming the ground floor and the first floor. Next, the roof is constructed by adding a ridge line 4.7 meters above the last floor slab, creating the attic. Once the **external envelope of the building is complete**, three **floor surfaces are added**.

This building is then **duplicated and pasted twice side by side to form a row of three terraced houses**. These geometries are input into the “HB Intersection Solids” component to define which spaces are internal and which are external. For example, the shared walls between the three housing units are classified as internal walls, while the non-shared walls are classified as external.

Subsequently, the geometry obtained through “HB Intersection Solids” is input into the “HB Room from Solids” component to **create the building’s rooms**. For simplicity, this script applies the same attributes, such as schedules, energy loads, and setpoints, to all rooms. Additionally, this component receives inputs for a construction set, program, and roof angle - the latter set to 46 degrees.

To create the construction set used as input for the “HB Room from Solids” parameter, the “HB ConstructionSet” component is utilized. This component assigns materials to each part of the building, such as the ground floor, roof, façade, or windows. Using the “HB Opaque Material” components, materials for each opaque envelope parameter are created. Subsequently, **properties such as thickness, conductivity, density, specific heat, roughness, thermal absorptance, and solar and visual absorptance are assigned to each material (See Table 1). Specifically, roughness, solar absorptance, thermal absorptance, and visible absorptance were selected according to the model’s user manual**. These values are consistent across all envelope parameters and are presented in Table 2.

Roughness	Thermal absorptance	Solar absorptance	Visible absorptance
Smooth	0,9	0,8	0,7

Table 2 - Details of materials of base building (Source: Energy Plus Manual)

In this section, the **temperatures of the crawl space** are also provided as input and sourced by Castenmiller, Es, and Stichting Bouwresearch (1993). For a floor slab with a thermal resistance of 0.41 m²K/W, the monthly temperatures shown in the Table 3 are considered.

	Crawl space temperature (°C)
January	10,7
February	10,7
March	11,7

April	12,7
May	14,5
June	15,9
July	17,5
August	17,5
September	15,9
October	14,2
November	12,6
December	11,4

Table 3 - Crawl space temperature considered for modelling the base building (Source: Own Work)

The glass for the windows is assigned using the “HB Subface Subset” component, specifying the **U Factor and solar heat gain coefficient (See Table 1)**. It is important to note that **thermal bridges were not considered** in this calculation. Including them would reduce the building's insulation performance, thereby increasing the energy required to maintain it.

The output of this component is then input into “HB Room from Solids.”

To create the program input for the “HB Room from Solids” parameter, the “HB ProgramType” component is used. In this component, no inputs are provided for people, lighting, or electric equipment values, so Honeybee uses its default standard values. Instead, an **infiltration rate of 0.0005 m³/sm² and 0.042 l/sm² for natural ventilation** supply with mechanical exhaust are input. The ventilation is continuous during the day. Additionally, **no cooling set point is assigned because the building is not considered to have a cooling system**. However, a heating set point is provided based on recommendations for a residential building from NEN (2024). **The heating set points are set to 20 degrees Celsius from 10:00 to 20:00 and 16 degrees Celsius from 20:00 to 10:00.**

After this, adjacencies are resolved using the “HB Solve Adjacency” component. Since the building units are supposed to be the equal, the **shared walls are considered adiabatic**. The output of this component is passed through “HB Apertures by Ratio” to create the windows. Based on the Ministry of the Interior and Kingdom Relations (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, 2022), **32% glazing is assigned to the façade**, applied only to the north and south façades. The final modeling step adds a gas boiler to the building using the “HB HVAC Template” component.

All the modelled geometric and material properties, internal loads, and HVAC systems are assembled into a **Honeybee model** using the “HB Model” component, containing all necessary details for simulating energy flows within the building. This model, along with the weather file for the 2020 scenario, is input into the **“HB Model to OSM” component, which produces various outputs, including the IDF file**. The overall workflow of the Grasshopper script is summarized in Figure 10.

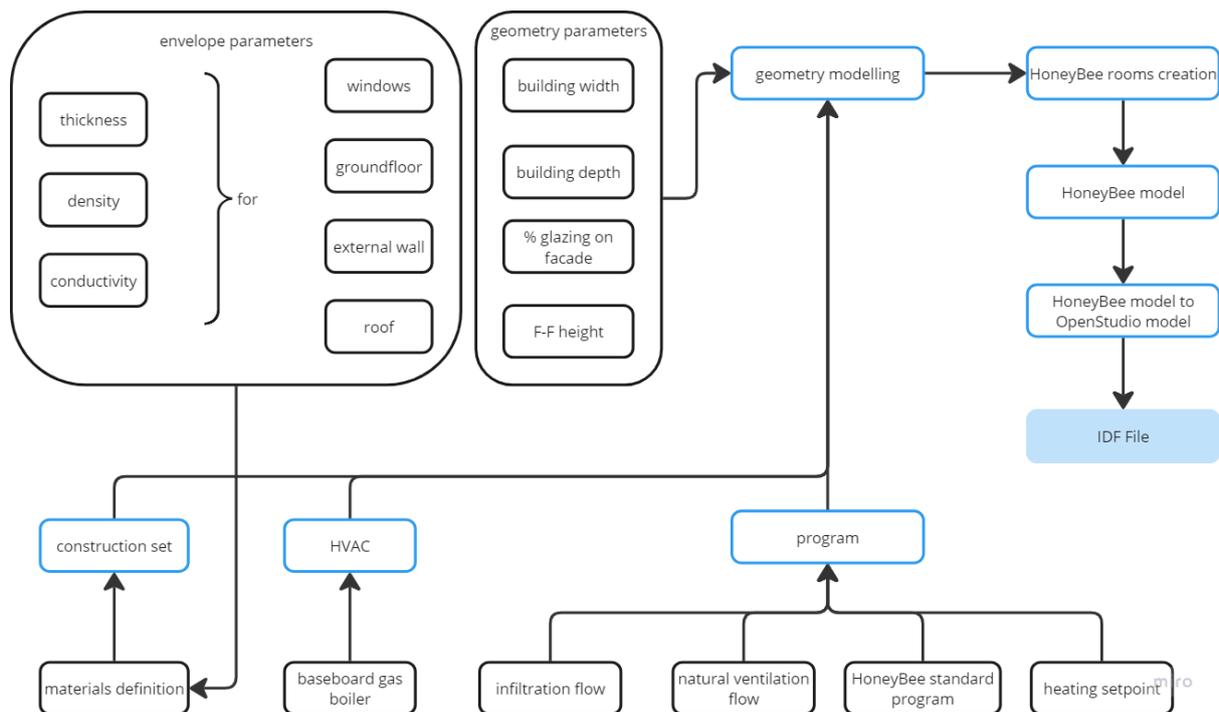


Figure 9 - Parametric building modelling workflow (Source: Own Work)

5.2.2 Retrofit parameters

The retrofit scenarios are integrated into the model by specifying the values for **thickness, conductivity, and U-Factor for each contemplated intervention**, corresponding to each parameter of the building envelope. These values are detailed in Table 4. Thickness is based on the drawn technical details, conductivity is sourced from the ASHRAE Handbook: Fundamentals (2021 edition, pp. 26.11–26.21), and U Factor is based on Tabula (n.d.).

component	scenario	specifications	U factor (W/m ² K)	thickness (m)	conductivity(W/MK)
windows	current	wooden frame, double glazing	2,9		
	retrofit 1	plastic frame, HR++ glazing	1,2		
	retrofit 2	wooden frame, HR++ glazing	1,21		
	retrofit 3	plastic frame, triple glazing	0,8		
	retrofit 4	wooden frame, triple glazing	0,81		
ground floor	current	wooden floor		0,12	0,29
	retrofit 1	PIR insulation in the cavity		0,12	0,03
	retrofit 2	hemp fiber insulation in the cavity and on top of the floor		0,2	0,04
	retrofit 3	resol insulation in the cavity		0,11	0,02
	retrofit 4	hemp fiber insulation in the cavity and on top of the floor		0,17	0,03
roof	current	wooden pitched roof		0,18	0,37
	retrofit 1	mineral wool in the cavity and below		0,18	0,04
	retrofit 2	hemp fiber in the cavity and below		0,19	0,04
	retrofit 3	PIR insulation in the cavity and below		0,17	0,02
	retrofit 4	hemp fiber insulation in the cavity and below		0,175	0,02
façade	current	massive clay bricks façade		0,3	0,66
	retrofit 1	external EPS insulation		0,42	0,1
	retrofit 2	external hemp fiber insulation		0,44	0,10
	retrofit 3	external EPS insulation		0,52	0,08
	retrofit 4	external hemp fiber insulation		0,54	0,08

Table 4 - Values of retrofit parameters (Source: Own Work)

It is important to note that for the scenarios that retrofit ground floor, the **temperature of the crawl space** shown in Table 5 are considered. They are sourced by Castenmiller, Es, and Stichting Bouwresearch (1993).

	Crawl space temperature (°C)
January	5,4
February	5,3
March	6,9

April	8,9
May	11,4
June	13,5
July	15,0
August	15,2
September	13,6
October	11,0
November	8,4
December	6,4

Table 5 - Crawl space temperature considered for scenarios retrofitting ground floor (Source: Own Work)

An other parameter that varies in the retrofitted options compared to the base building is the infiltration. In fact, when a building is better insulated is also has a lower infiltration value. For this reason, the **infiltration of retrofit options is set to 0.0002 m³/sm²**.

By inputting these details, the model is informed of all the retrofit specifics for the four parameters - ground floor, façade, roof, and windows. With these inputs provided, the model can generate all possible combinations of retrofit interventions. For instance, as referenced in the aforementioned Table, it might create a building option where the windows have Retrofit 1, the ground floor has Retrofit 4, the roof has Retrofit 2, and the façade remains unchanged, thereby staying in the current scenario.

Given that four parameters are considered, and each parameter has five scenarios including four retrofit options and one current scenario, the total number of building retrofit combinations is 5⁴, which equals 625. These scenarios will be assessed for three different weather files, resulting in a total of 1875 simulations, calculated as 625 combinations multiplied by 3 weather files.

5.2.3 *Weather file*

The weather files used as input for the simulations are those described in Chapter 4. This includes the EPW files for the current, 2050, and 2100 scenarios.

5.2.4 *Output data*

The final output of the model consists of three CSV files, one for each weather scenario. Each file contains the results of the simulations for all 625 retrofit combinations. For each building configuration, there are columns detailing the thermal properties of the envelope components as well as daily data for every day of the year. Specifically, these files include daily simulations for gas consumption, which accounts for energy used by the heating system and domestic hot water system. They also include building electricity, representing the electricity consumed by lighting, equipment, and appliances. Facility electricity is also included, covering the electricity used by HVAC systems, such as fans, cooling systems, and heat pumps. Additionally, the files provide data on the maximum indoor temperature. In these simulations the average maximum temperature between the ground and first floors was taken into account, as the second floor functions as an attic and is not typically used during the day. These three CSV files are crucial for the development of the AI-based model, as they provide the data used to train the model.

5.3 Intercomparison

To ensure the reliability of the simulation results, the energy consumption due to room heating obtained from the simulations was compared with that derived from the Warmteprofielgenerator (n.d.). The Warmteprofielgenerator proved to be extremely valuable in this context, as it is an online tool that allows users to create thermal profiles for various types of homes by customizing building details to obtain accurate energy consumption data. This provides concrete examples that are instrumental in conducting energy analyses and assessing thermal efficiency.

Specifically, this comparison analysed the simulation results of the base building against those of the building with the highest retrofit for all the considered envelope parameters. The comparison is presented in Tables 6 and 7.

	unit	Annual energy consumption for room heating
Warmteprofielgenerator	kWh/m ²	130
This research	kWh/m ²	136

Table 6 - Annual energy consumption due to room heating for base building in analysis (Source: Warmteprofielgenerator (n.d.))

	unit	Annual energy consumption for room heating
Warmteprofielgenerator	kWh/m ²	26
This research	kWh/m ²	28

Table 7- Annual energy consumption due to room heating for base building in analysis (Source: Warmteprofielgenerator (n.d.))

The intercomparison between calculation models demonstrates a strong correlation. It is important to recognize that actual energy consumption in residences can vary widely due to numerous factors, including user behaviour. Notably, the energy consumption of well-insulated homes is often significantly higher than what model calculations predict.

Additionally, while the EnergyPlus model utilized in this research provides data on gas consumption, facility electricity, and building electricity, this study concentrates exclusively on gas consumption. This focus is based on the retrofit scenarios examined, which do not affect the building's electricity usage. Consequently, including electricity consumption in the analysis is considered unnecessary. Therefore, the annual energy consumption data presented in the Tabula WebTool (n.d.) pertains solely to the energy required for heating.

6. Development of the AI-based surrogate model

6.1 Multi-Task Learning model overview

6.1.1 Overall workflow

In the pursuit of identifying optimal retrofit solutions, the integration of artificial intelligence (AI) surrogate models has emerged as a promising approach. Among the various AI methodologies, Multi-Task Learning (MTL) stands out due to its ability to simultaneously address multiple interrelated tasks. However, traditional MTL approaches often encounter significant challenges, particularly when **tasks exhibit conflicting objectives**. These conflicts make it difficult to **optimize all tasks simultaneously** using a single solution, as improvements in one task may lead to deteriorations in another.

To address these challenges, this research frames the **MTL problem as a Multi-Objective Optimization (MOO)** task, leveraging advanced methods to balance the trade-offs between multiple objectives effectively. A MTL model was developed using PyTorch, drawing on methodologies inspired by GumGum Tech (2020a, 2020b). The model is designed to simultaneously predict energy consumption, cost, carbon emissions, and comfort days across various building configurations. By interpreting MTL as an MOO problem, the model seeks to find Pareto-optimal solutions where no objective can be improved without worsening another, thus ensuring a balanced performance across all tasks.

Two distinct architectures - one based on literature review and one based on data analysis - and **two training approaches** - the weighted sum method and the Multi Gradient Descent Algorithm - were developed and subsequently compared. The **weighted sum method** represents a traditional approach where task-specific losses are combined using predefined weights, which often struggle with balancing conflicting objectives. In contrast, the **Multi Gradient Descent Algorithm** employs gradient-based techniques to dynamically adjust the learning process by considering the gradients of each task's loss function. This ensures that all tasks make progress without significantly deteriorating the performance of any individual task. The performance of these models was evaluated using root squared error, mean root squared error, and mean square error metrics to identify the most effective approach.

Upon selecting the top-performing model, two different MOO techniques were developed to select sets of optimal retrofit solutions: one utilizing Pareto optimization and the other combining NSGA II (Non-dominated Sorting Genetic Algorithm II) with Pareto optimization. Users have the flexibility to choose between these two optimization methods based on their specific needs and preferences.

Finally, the specific optimal retrofit solution was selected through Multi-Criteria Decision Making (MCDM), which takes into account the weights assigned by the client to each objective. This step ensures that the final decision prioritizes the tasks based on the client's preferences, effectively translating the Pareto-optimal solutions into a tailored retrofit strategy. The overall workflow of the model, encompassing MTL, MOO, and MCDM processes, is illustrated in Figure 11.

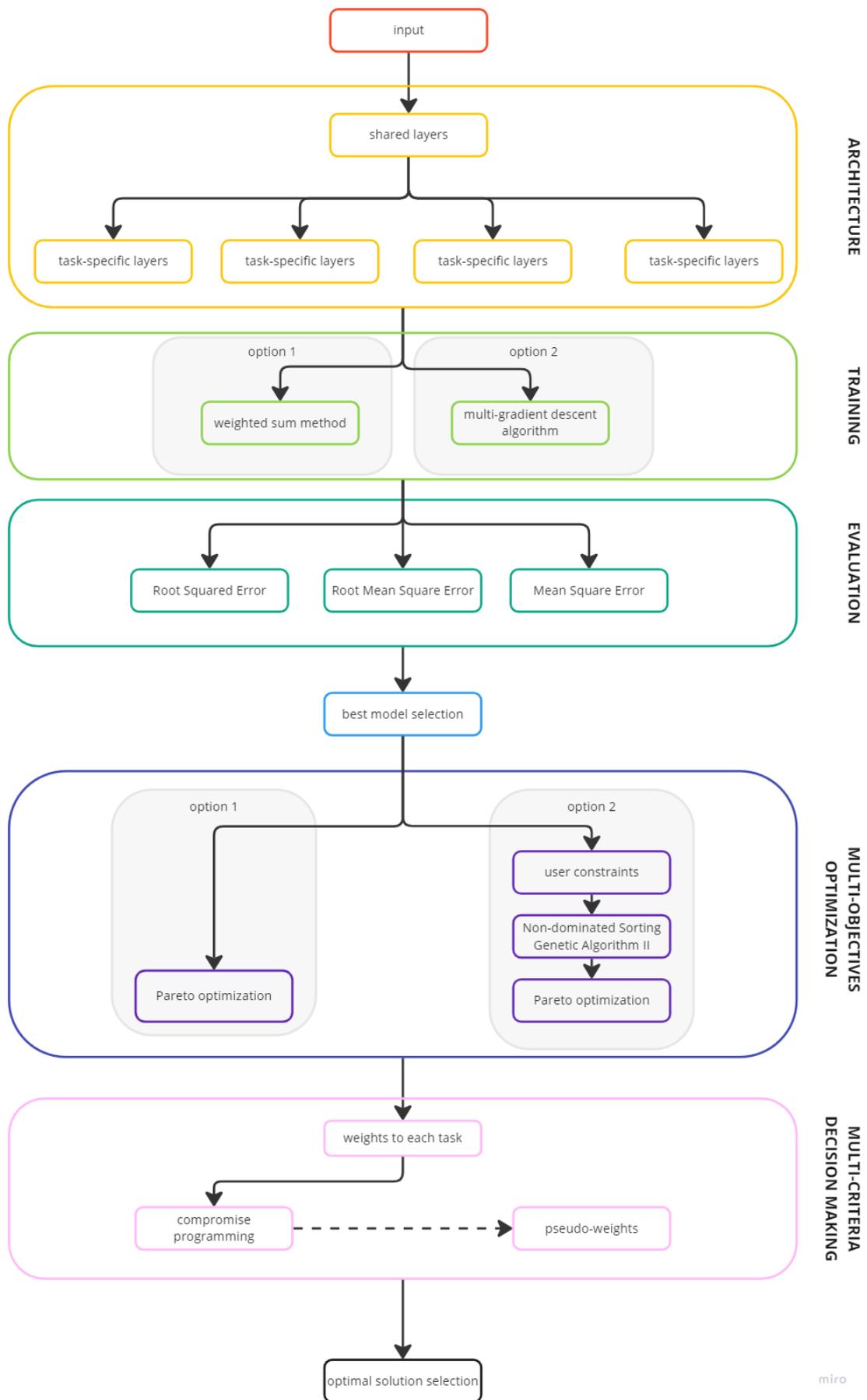


Figure 10 - Overall workflow to select optimal retrofit solutions (Source: Own Work)

6.1.2 Introduction to Multi Task Learning

MTL is a deep learning approach where a **single model is trained to perform multiple tasks simultaneously** by sharing portions of the model's layers. This shared learning leverages commonalities among tasks, enhancing efficiency. The foundational work by Caruana (1997) demonstrated that joint learning of related tasks can lead to superior performance compared to single-task learning, establishing the theoretical underpinnings of MTL.

The selection of retrofit solutions encompasses various interconnected objectives, including **cost, environmental impact, thermal comfort, and energy consumption**. These tasks are inherently **interrelated**; for instance, strategies to minimize costs may influence energy consumption and thermal comfort, while efforts to reduce environmental impact might affect overall expenses. Utilizing an MTL model allows for the simultaneous handling of these tasks, capturing the intricate interdependencies among them in an integrated manner. This approach offers several advantages, such as lower computational costs, higher accuracy and good scalability, further discussed in chapter 6.1.3.

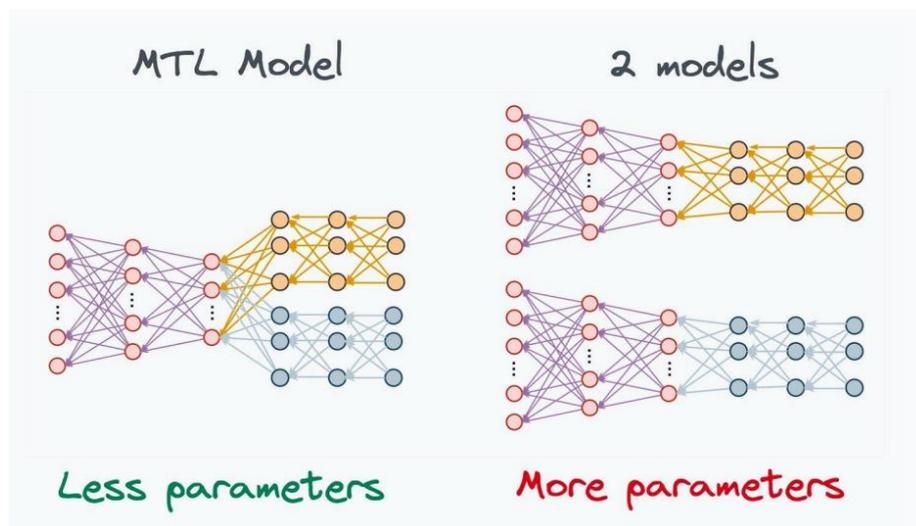


Figure 11 - Overall structure of MTL model compared to multiple models (Source: Chawla, 2024)

6.1.3 Architecture of Multi-Task Learning

The architecture of an MTL model is characterized by its division into **shared and task-specific layers**, facilitating the simultaneous learning of multiple tasks (See Figure 12). Shared parameters constitute the parts of the model that are common across all tasks. These components are responsible for extracting features from the input data that are relevant to multiple tasks, thereby promoting knowledge transfer and reducing redundancy. Typically, shared layers form the backbone of the model, capturing the underlying patterns and structures inherent in the data.

Complementing the shared layers are the task-specific layers, which are unique to each task. These components cater to the specific nuances and requirements of individual objectives, refining the shared representations to produce accurate predictions tailored to each task without interference from others. This structure enables the model to **learn both generalizable features** applicable to all tasks **and specialized features** unique to each task, fostering a comprehensive understanding of the data. By balancing shared and task-specific parameters,

MTL models can effectively leverage commonalities among tasks while accommodating their unique characteristics.

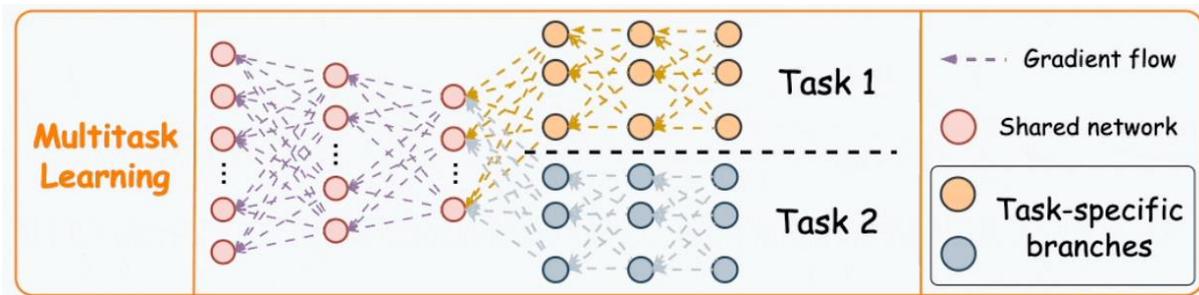


Figure 12 - Architecture of MTL (Source: Chawla, 2024)

Academic research supports this structural approach. For example, Ruder (2017) emphasizes the importance of designing architectures that balance shared and task-specific components to optimize performance across all tasks. Similarly, Liu et al. (2019) demonstrate how task-specific layers can enhance the model's ability to handle fine-grained distinctions within each task, further validating the structural benefits of MTL in complex applications.

6.1.4 Benefits of Multi-Task Learning

Adopting MTL in the context of retrofit solution selection offers several key advantages. As showed by Ruder (2017) and Liu et al. (2019), one of the primary benefits is **improved generalization**. Sharing information across related tasks acts as an inductive bias, enabling the model to learn more robust and generalizable features. This is particularly beneficial when tasks are related, as the model can leverage shared structures in the data to enhance its predictive capabilities. For example, predicting both energy consumption and thermal comfort allows the model to understand the underlying factors that influence both, leading to more accurate and reliable predictions.

Another significant advantage is the **reduction of overfitting** (Ruder, 2017). By sharing parameters among tasks, MTL serves as a regularization mechanism. This is especially useful when some tasks have smaller datasets, as the shared information helps prevent the model from overfitting to specific task data, thereby improving overall model performance. Additionally, MTL enhances **computational efficiency** (Kamali et al., 2019). A single MTL model can handle multiple tasks simultaneously, reducing the overall computational resources required compared to training and maintaining separate models for each task. This efficiency translates to faster training times and lower operational costs, making MTL a cost-effective solution for complex prediction tasks.

Scalability is another critical benefit of MTL. The inherent scalability of MTL allows for the easy addition of new tasks or objectives. This flexibility is crucial for adapting to changing research needs and incorporating additional predictive objectives without the need for extensive model redevelopment. For instance, integrating a new objective such as indoor air quality prediction into the existing MTL framework can be accomplished with minimal adjustments, ensuring that the model remains relevant and adaptable to evolving requirements. This scalability is supported by academic findings, where Ruder (2017) discusses how scalable MTL architectures can accommodate an expanding set of tasks without significant architectural overhauls.

6.1.5 *Challenges in Multi-Task Learning*

Despite its numerous benefits, Multi-Task Learning (MTL) presents several challenges that must be addressed to ensure optimal performance. A primary challenge is the **need for trade-offs**, which involves **balancing the performance across all tasks without disproportionately sacrificing one for the sake of others**.

To address these challenges, this research frames the MTL problem as a Multi-Objective Optimization (MOO) task. Unlike traditional MTL approaches that aim to minimize a single aggregated loss function, MOO seeks to minimize multiple loss functions simultaneously. This perspective acknowledges that it is rare for all objectives to be perfectly minimized at the same time, as improving one objective may worsen another. Instead, the goal is to identify Pareto optimal solutions—situations where no objective can be improved without degrading at least one other objective. By interpreting MTL through the lens of MOO, the model can effectively navigate the inherent trade-offs between conflicting objectives, ensuring a balanced performance across all tasks. This approach not only mitigates the difficulties associated with optimizing multiple, competing objectives but also enhances the model's ability to deliver comprehensive and reliable insights for retrofit solution selection.

6.1.6 *Dataset preprocessing*

The dataset discussed in Section 5.3, together with the data about costs and embodied carbon, was preprocessed by the Multi-Task Learning model, allowing it to more easily identify relationships among data. This preprocessing aimed to select only the data pertinent to the MTL model. The choices on how to preprocess the data were driven by the model's primary objectives: predicting energy consumption, thermal comfort, cost, and environmental impact in a retrofit scenario under varying weather conditions.

6.1.7 *Input data*

To develop a surrogate model, it is essential that its **inputs align with those used in EnergyPlus simulations**. This alignment ensures that the surrogate model accurately replicates the conditions and parameters of the original simulations, facilitating reliable predictions.

Specifically, consistent with EnergyPlus, the input data for the MTL model includes the time horizon and the retrofit scenario. This is represented by the thermal properties of the building envelope components, specifically the thermal resistance (R_c values) of the roof, floor, façade, and the U-Factor of the windows. Consequently, the input dataset for the MTL model is composed of five parameters: **floor, roof, external wall thermal resistance, windows U-Factor, and time horizon**. The U-Factor for windows is utilized instead of thermal resistance because it is more commonly employed in industry practice.

Only the mentioned thermal properties of the envelope parameters were considered because they are essential for determining the heat transfer between the indoor and outdoor spaces, thus assessing the building's energy consumption and thermal comfort. While other parameters could have been included alongside the R_c and U-Factor, they were omitted to avoid complicating the input data, making it more challenging to manage. This decision is a limitation of the project and a point that could be explored more thoroughly in future studies.

6.1.8 *Output data*

The output dataset of the MTL model corresponds to the four tasks that the model is designed to predict: **annual energy consumption, cost, embodied carbon, and thermal comfort**. The outputs from the simulations run with Energy Plus provided daily energy consumption data for each retrofit scenario considered. However, for the purposes of this research, it is feasible to consider annual energy consumption instead of daily. Indeed, the overarching goal of the project is to minimize the building's energy usage, which can pertain to either daily or annual consumption. Utilizing an annual value helps to avoid the need to manage time-series data, which would make the model significantly more complex and time-consuming. Thus, annual energy consumption was calculated by summing the 365 daily values of energy consumption obtained from the Energy Plus simulations.

The cost parameter in the MTL model refers to a value in euros per square meter of the building, calculated by summing the costs of the four envelope parameters. The investment costs of the retrofit interventions are the ones detailed in Chapter 3.2.

The environmental impact is assessed in terms of the embodied carbon of the retrofit materials used, referring to the values shown in Chapter 3.2. The total embodied carbon for each retrofit scenario is the sum of the embodied carbon from each intervention for each envelope parameter.

Thermal comfort was evaluated based on daily maximum indoor temperatures, derived from simulations run with EnergyPlus. According to Ioannou and Itard (2017), in the Netherlands, the indoor thermal comfort of a residential building ranges between 18 and 26 degrees Celsius. Based on this range, each day was classified as either in comfort (within the range) or out of comfort (outside the range). After this classification analysis, the comfortable days throughout the year were added together. Thus, for each retrofit scenario, the number of comfortable days was calculated. This value was subsequently utilized by the model to account for the building's thermal comfort.

6.1.9 *Dataset analysis and Interpretation*

After preprocessing the data, the correlation matrix between input and output variables is further examined to guide the architecture of the multi-task learning (MTL) model (Architecture Option 2: Task-specific architectures based on data analysis, see Chapter 6.3.3). This step is essential to understand the strength and direction of relationships, both positive and negative, between inputs and outputs. Understanding these correlations allows for the identification of tasks that share common underlying patterns, indicating that they can effectively leverage shared layers within the model. Conversely, inputs with weak correlations suggest more complex relationships, necessitating additional shared layers to capture these nuances accurately. By tailoring the model architecture based on these insights, the MTL framework can efficiently balance shared and task-specific parameters, enhancing both computational efficiency and predictive performance. This approach is driven by the hypothesis that tasks with strong correlations benefit from shared representations, thereby improving learning efficiency, while maintaining flexibility for tasks with less straightforward relationships. So, the correlation matrix serves as a critical tool in optimizing the model's structure, ensuring it is well-aligned with the data relationships.

The generated correlation matrix is illustrated in Figure 14. This matrix is a table displaying the correlation coefficients between variables, providing a summary of how each pair of variables in the dataset is related. The values in the matrix range from -1 to 1, where 1 indicates a perfect positive correlation, -1 signifies a perfect negative correlation, and 0 represents no correlation

at all. The diagonal entries of the matrix, all marked as 1.00, denote perfect self-correlation for each variable.

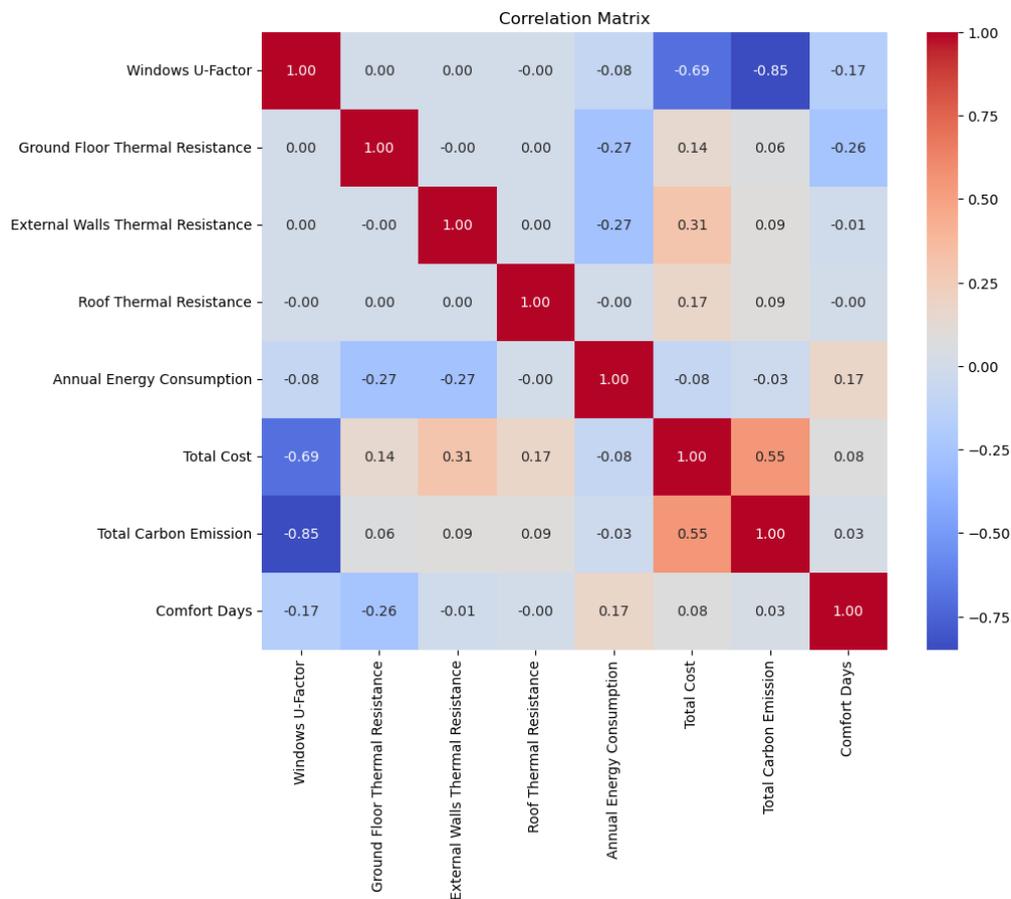


Figure 13 - Correlation Matrix between Input and Output parameter of the MTL Model (source: Own Work)

The matrix reveals that a lower Windows U-Factor is associated with higher total cost (-0.69) and total carbon emission (-0.85). This implies that improvements to the thermal performance of windows, while beneficial for reducing heat loss, are expensive not only in financial terms but also in terms of CO_{2e} emissions. Such findings suggest that enhancing window insulation is a costly endeavor, both economically and environmentally.

In terms of broader economic and environmental impacts, there is a moderate correlation between total cost and total carbon emission (0.55). This relationship indicates that more costly retrofit interventions are also more environmentally impactful. This correlation likely arises because achieving better energy performance often necessitates the use of more insulating material, thereby increasing both the cost and the environmental footprint. This finding is particularly intriguing in the context of low embodied carbon insulating materials, which, despite criticisms of their high cost, do not demonstrate this trend according to the matrix. This suggests a potential reconsideration of the economic and environmental trade-offs involved in selecting building materials.

Furthermore, among the envelope parameters considered, the external wall shows the highest correlation with total cost, indicating that retrofitting the facade is the most expensive envelope intervention. Simultaneously, the retrofit of both the facade and the ground floor has the most

significant effect on reducing annual energy consumption - both correlated at -0.27 with it. These two envelope parameters - facade and ground floor - emerge as the most influential on the building's energy performance among those studied.

Additionally, there is a weak positive correlation between annual energy consumption and comfort days (0.17), suggesting that higher energy consumption may be associated with increased indoor comfort. This relationship implies that buildings consuming more energy potentially maintain conditions that are more conducive to occupant comfort, highlighting a possible trade-off between energy efficiency and comfort levels within buildings.

In conclusion, the analysis of the correlation matrix offers an important insight into the relationships between key building performance metrics. It highlights the high costs and environmental impacts associated with improving window insulation and suggests that facade and ground floor retrofits are crucial for enhancing energy efficiency. The matrix also challenges the prevailing views on the cost-effectiveness of low embodied carbon materials, indicating a need for a nuanced understanding of material selection in sustainable building practices.

6.2 Architectures of the model

6.2.1 *Architecture options overview*

A key feature of the Multi-Task Learning model is its architecture, which incorporates both shared and task-specific layers that contribute to the final predictions for various tasks. Given this characteristic, this study aimed to explore the potential of this feature by experimenting with different architectures for the shared and task-specific layers. Specifically, two architectures were explored:

1. Assigning a distinct type of artificial neural network (ANN) to each task-specific layer, following the recommendations from four key studies—Fan et al. (2017), Yun et al. (2022), Altikat et al. (2021), and Escandón et al. (2019). These papers have developed ANNs tailored for prediction tasks similar to those addressed in this study, ensuring that each layer utilizes the most suitable network architecture for its specific purpose;
2. Assigning a distinct type of ANNs to each task-specific layer based on the best architecture for each task as determined by the data correlations analyzed in Section 6.2.3. This approach aims to tailor the architecture to the unique characteristics and correlations of the data pertaining to each task.

Both options are implemented and compared to determine the optimal architecture for the MTL model.

The development of the MTL model was primarily guided by the GitHub repository by Yaringal (n.d.) and the article by K (2022), both of which provide instructions for constructing an MTL model using PyTorch. NumPy was employed to handle arrays and execute essential mathematical operations, facilitating data processing within the model. Pandas was utilized to manage and manipulate structured data in DataFrame format, thereby streamlining data analysis. PyTorch served as the framework for constructing, training, and optimizing the neural network, leveraging its robustness in deep learning applications. To visualize data and results, Matplotlib was used, which aided in interpreting and understanding the model's performance. The Adam optimizer was selected with a learning rate of 0.001 to refine the model parameters effectively, and the training process was configured to run for 100 epochs, allowing sufficient iterations for the model to learn from the data. Additionally, the dataset was divided into 70% for training, 15% for validation, and 15% for testing. This division ensured a comprehensive

approach to training the model while enabling effective validation and generalization testing on unseen data.

6.2.2 *Option 1: Task-specific architectures based on literature review*

This architecture option is illustrated in Figure 14. As the image shows, it begins with two shared fully connected layers that extract general patterns relevant to each target variable. These shared layers transform the four input features into a high-dimensional representation.

The first shared layer is designed with 128 neurons, which provides a sufficient capacity to capture complex relationships in the input data while avoiding excessive dimensionality that could lead to overfitting. Empirical testing showed that starting with a higher dimensional representation improved the model's ability to generalize across tasks. The choice of 128 neurons balances computational efficiency with representational capacity, allowing the model to learn a broad range of patterns without excessive complexity. The second shared layer reduces the dimensionality to 64 neurons. This step-down in neurons is deliberate, as it focuses the model's representation before branching into task-specific pathways. By progressively reducing the dimensionality, the architecture avoids unnecessary complexity and encourages the shared layers to extract the most relevant features for each task.

In this approach, each task in the MTL model branches into separate pathways, customized to meet its particular requirements. The model utilizes distinct Artificial Neural Networks (ANNs) tailored for each of the four key tasks: cost, energy consumption, embodied carbon, and thermal comfort prediction.

For predicting energy consumption, the study by Fan, C., Xiao, F., & Zhao, Y. (2017) was used as reference. It explores the application of a Deep Neural Network (DNN) specifically for cooling energy prediction. The findings suggest that an effective DNN model for this task does not require a deeply layered architecture; instead, it operates optimally with just two hidden layers. The number of neurons in each layer is determined by an empirical rule commonly used in neural network design, which calculates the neuron count as half the sum of the number of inputs and outputs. For this study, this method yields 4.5, subsequently rounded down to 4 neurons per layer.

In cost prediction, Yun, S. (2022) presents an ANNs for construction costs prediction. It is configured with an input layer, an output layer, and two hidden layers of 100 and 64 nodes, respectively. This specific structure is tailored to enhance the accuracy of predicting both costs effectively.

Regarding environmental impacts, Altikat, S. (2021) investigates the use of two distinct ANN configurations to predict embodied carbon emissions. One model utilizes a single learning function with linear transfer functions and 8 neurons, achieving 95.56% accuracy. Another model employs a deep learning neural network (DLNN) approach with 14 neurons in the first hidden layer and 10 in the second, attaining a higher accuracy of 98.29%. This increase in precision indicates that the DLNN is particularly suitable for accurately forecasting CO₂ emissions, thus recommending its use in assessing environmental impacts.

Finally, for thermal comfort prediction, the methodology used in this research is drawn from Escandón, R., Ascione, F., Bianco, N., Mauro, G. M., Suárez, R., & Sendra, J. J. (2019), which employs a Multi-Layer Perceptron (MLP). This model integrates 18 input parameters that reflect various building characteristics, a single hidden layer with 6 neurons, and an output layer that predicts the annual percentage of discomfort hours. Such a setup is instrumental in providing insights into the expected comfort levels within buildings, facilitating the design of environments that enhance occupant comfort.

All mentioned papers used ReLU as activation function and Adam as optimizer. The MTL model with this architecture has been trained, validated and tested. To assess its performances training and validation losses graphs, predicted vs. actual values graphs and percentage error distribution graphs have been generated and can be seen in Figures 15, 16, and 17 respectively. These graphs are further discussed in Section 6.5.

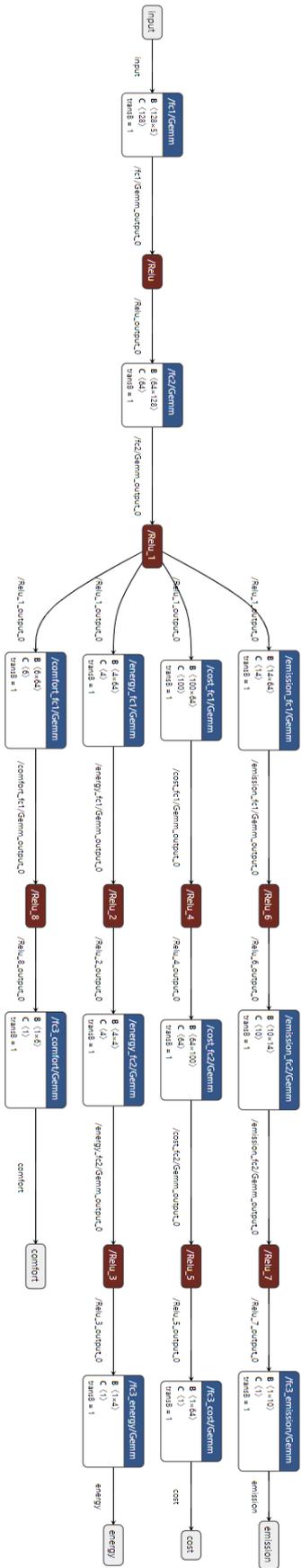


Figure 14 - Visualization of architecture option 1 (Source: Own Work)

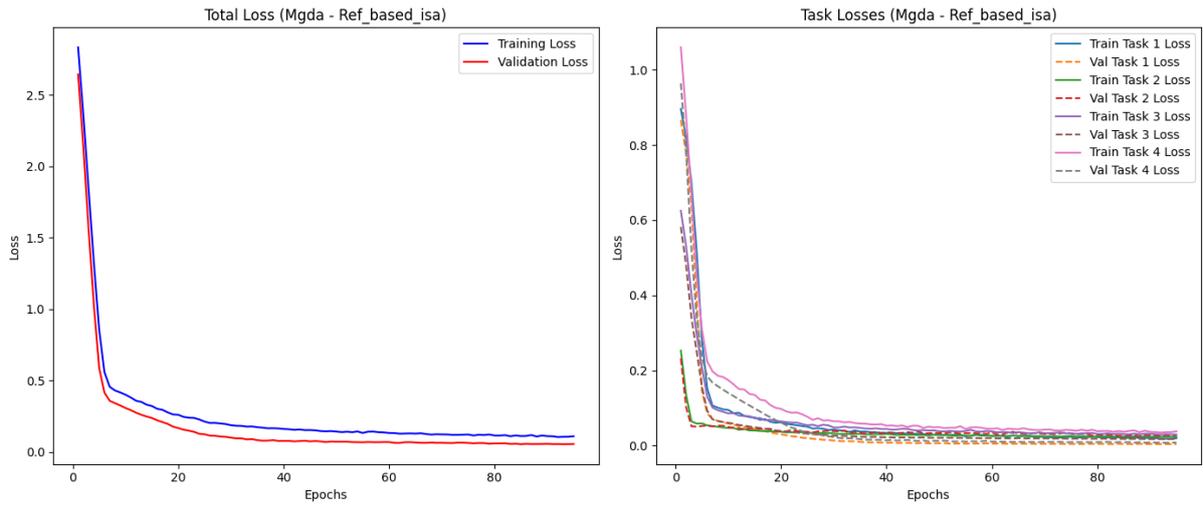


Figure 15 - Training vs. validation losses graphs for architecture option 1 (Source: Own Work)

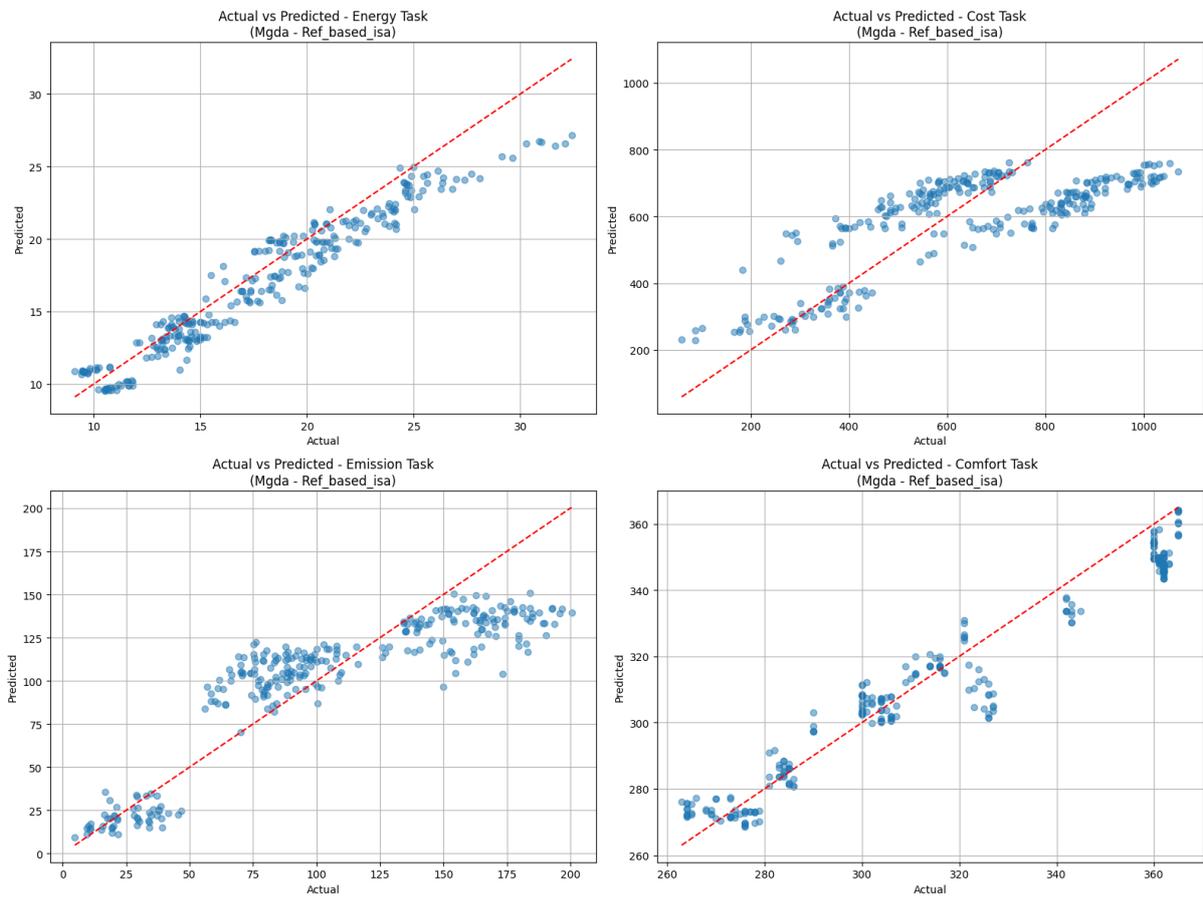


Figure 16 - Predicted vs. actual values graphs for architecture option 1 (Source: Own Work)

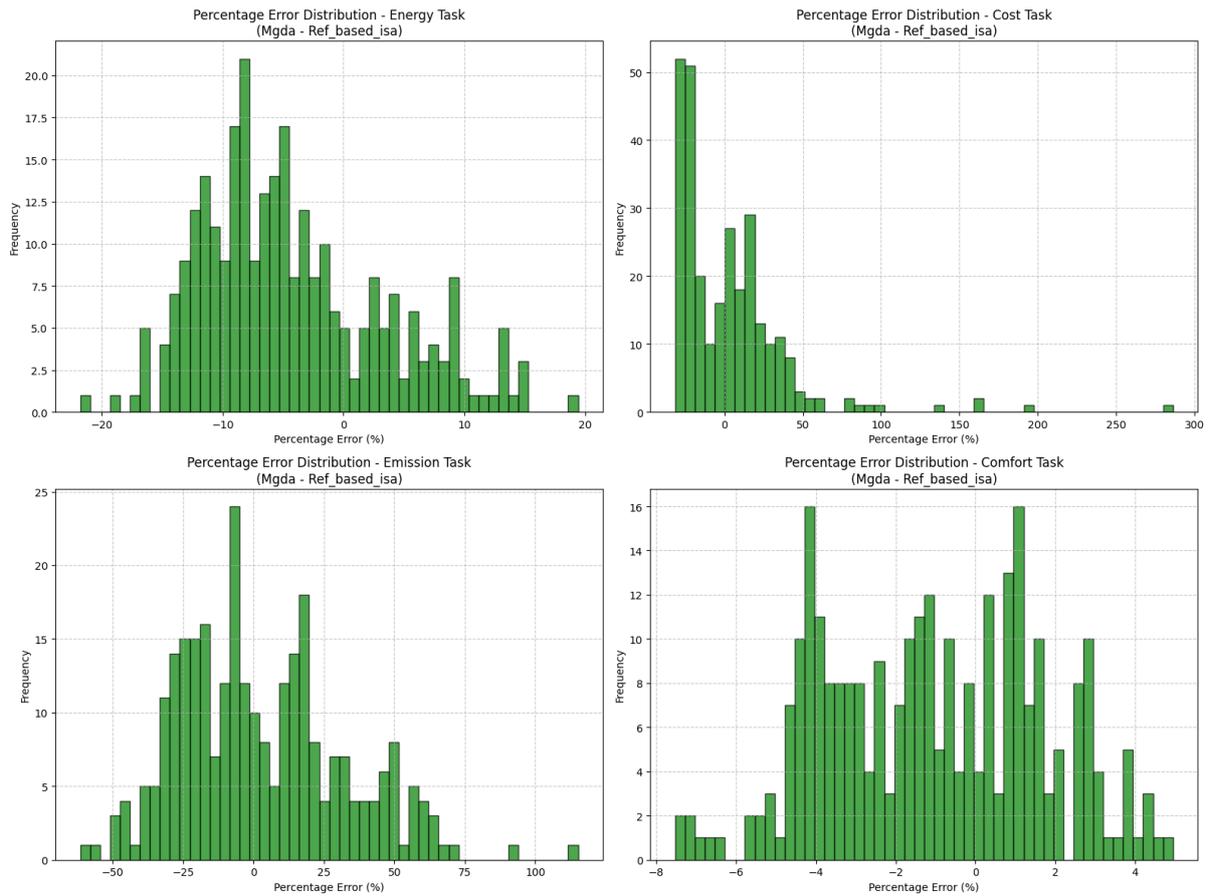


Figure 17 - Percentage error distribution graphs for architecture option 1 (Source: Own Work)

6.2.3 Option 2: Task-specific architectures based on data analysis

This architecture option is illustrated in Figure 18. Like the previously mentioned option, this architecture begins with two shared fully connected layers consisting of 128 and 64 neurons. Then, based on the correlation analysis, each task was assigned a customized structure to optimize predictive performance and leverage beneficial task interdependencies.

For the annual energy consumption prediction task, a specific layer with 32 neurons follows the shared layers to capture unique insulation-related patterns, specifically from ground floor thermal resistance and external walls thermal resistance. These features exhibit moderate negative correlations with energy consumption, indicating their impact on heating and cooling needs. The choice of 32 neurons in this layer is based on the need for a sufficiently complex representation to capture patterns relevant to energy consumption, without adding excessive capacity that could lead to overfitting. This layer size provides a balanced approach, allowing the layer to learn both linear and non-linear relationships that are important for energy predictions. This reasoning behind selecting 32 neurons is also applied to the other task-specific layers, as it provides enough complexity to capture essential patterns without unnecessary model size or risk of overfitting. The final output layer has 1 neuron, corresponding to the single target value for energy consumption.

Since the correlation between cost and carbon emissions was shown not to be strong, a specific shared layer with 32 neurons for these two tasks has been built. With this strategy, the model will be more likely to capture interdependencies the two involved tasks. After this shared layer, each task has a separate output layer with 1 neuron to produce its specific target value,

enabling the model to fine-tune predictions for cost and carbon emission independently, even after sharing intermediate representations.

For the comfort days task, the structure diverges into a distinct section with its own fully connected layer containing 32 neurons and a dedicated output layer with 1 neuron. Comfort days showed weak correlations with the other targets, indicating that it benefits from an independent structure to avoid interference from unrelated tasks. This design ensures that the model captures any unique factors influencing comfort days without being influenced by the other tasks.

The MTL model with this architecture has been trained, validated and tested. To evaluate its performances training vs. validation losses graphs, predicted vs. actual values graphs and percentage error distribution graphs have been generated and can be seen in Figures 19, 20, and 21 respectively. These graphs are further discussed in Section 6.5.

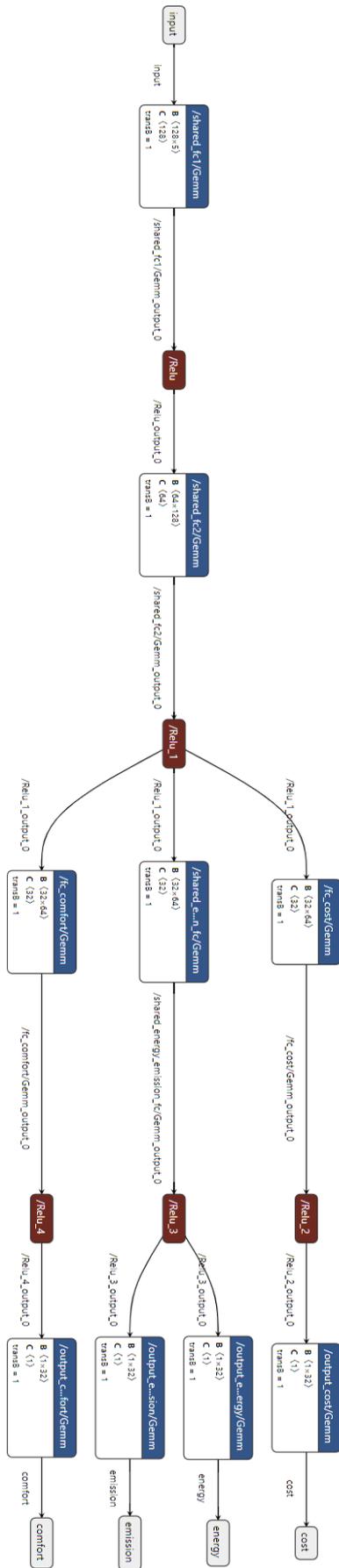


Figure 18 - Visualization of architecture option 2 (Source: Own Work)

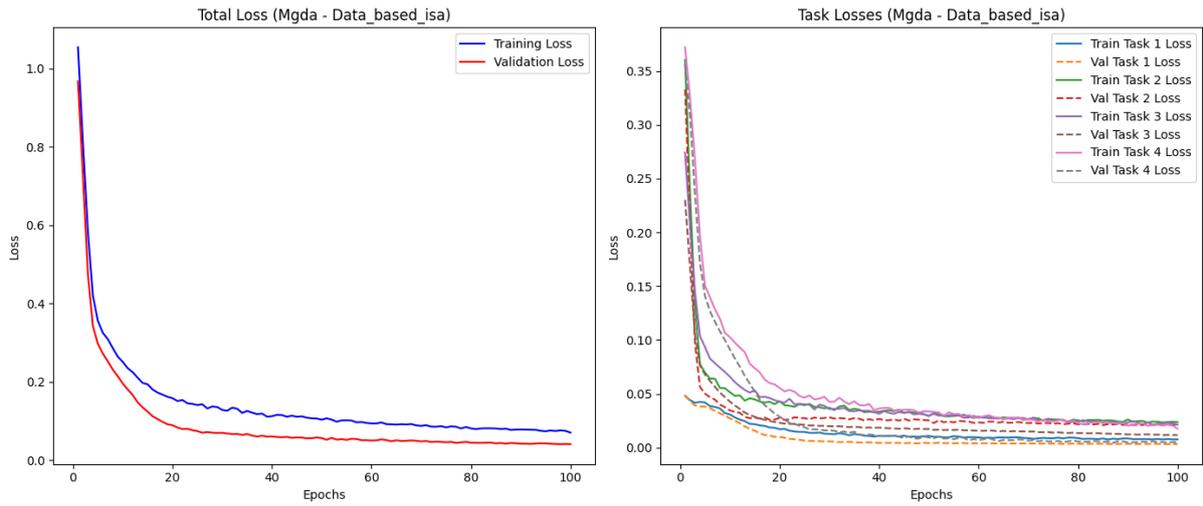


Figure 19 - Training vs. validation losses graphs for architecture option 2 (Source: Own Work)

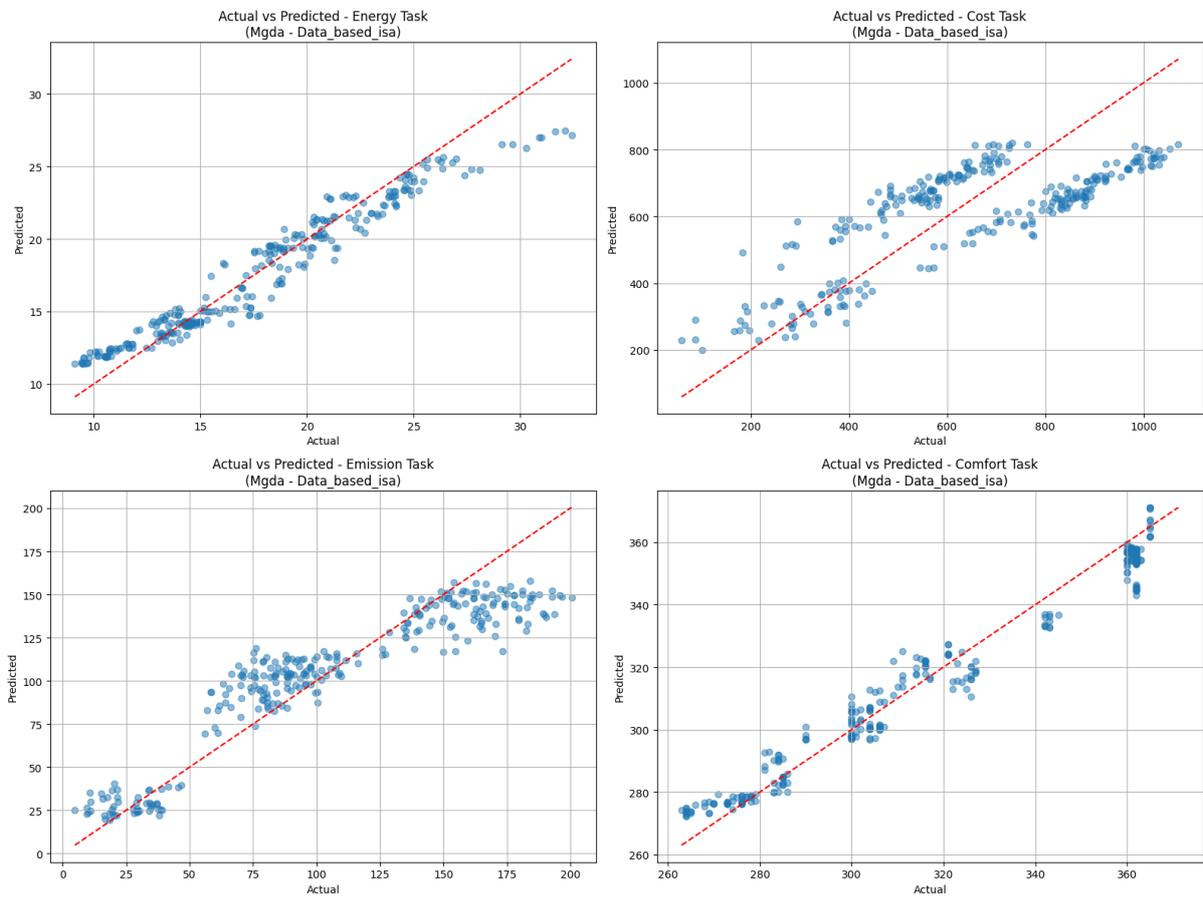


Figure 20 - Predicted vs. actual values graphs for architecture option 2 (Source: Own Work)

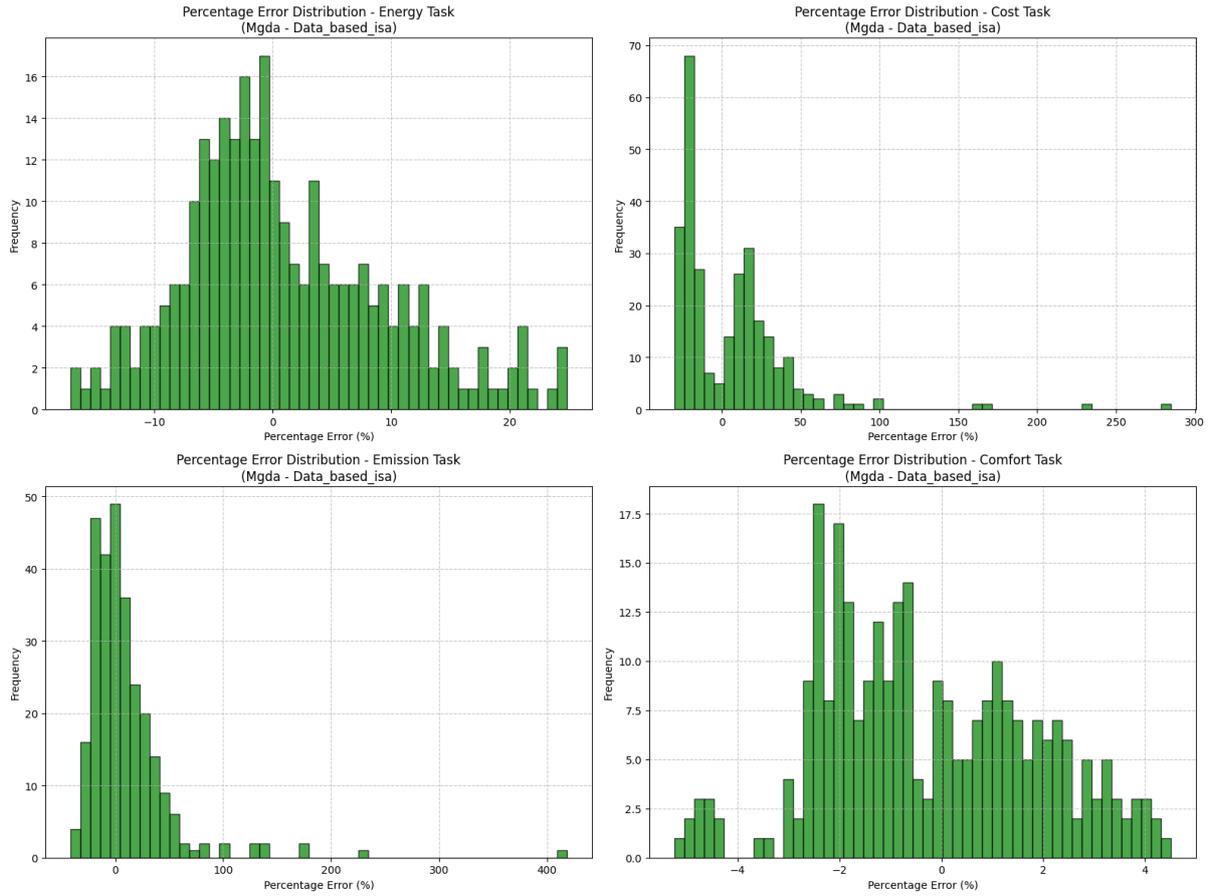


Figure 21 - Percentage error distribution graphs for architecture option 2 (Source: Own Work)

6.3 Training functions

The development of the Multi-Task Learning involves several key training functions to ensure its effectiveness. A **specific loss function** is designed to balance the multiple objectives, ensuring that each task contributes appropriately to the overall model performance. To enhance generalization and prevent overfitting, an **early stopping mechanism** is implemented, monitoring the model's performance and terminating training at the optimal point. Additionally, two distinct trainers - the **Weighted Sum method** and the **Multiple Gradient Descent Algorithm (MGDA)** - are employed separately to optimize the model. A comparative analysis is conducted to evaluate the performance of the MTL model using each optimization method, highlighting the differences in their ability to achieve accurate results.

6.3.1 Loss function

- **Loss function of the model**

In multi-task learning (MTL), the **traditional loss function** used to optimize multiple tasks simultaneously. It is defined as:

$$\min_{\theta} \sum_{t=1}^T \lambda_t L_t(\theta)$$

where:

- T is the total number of tasks.
- $L^t(\theta)$ represents the loss for task t , parameterized by θ .
- λ_t are weighting coefficients that determine the relative importance of each task's loss.

This formulation requires careful tuning of the coefficients c^t to balance the tasks, which can be a complex challenge. To address this, **the Multi-Task Learning problem is reformulated as a Multi-Objective Optimization problem**, utilizing the following loss function:

$$\min_{\theta} (L^1(\theta), L^2(\theta), \dots, L^T(\theta))$$

where:

- T is the total number of tasks.
- $L^t(\theta)$ represents the loss for task t , parameterized by θ .

This formulation treats each $L^t(\theta)$ as separate objective, seeking solutions that are Pareto optimal across all tasks.

- **Loss functions of each task**

In MTL each output represents a different task setup. In particular, the model predicts the following:

1. Task 1 – Energy consumption

In this task, the output is the annual energy consumption, which is a continuous variable. The task is a regression problem, and the objective is to minimize the error between the predicted and actual energy consumption values.

$$\text{Objective 1: } \min_{\theta} L_{energy} = \text{annual_energy_consumption}$$

2. Task 2 – Predict total cost

In this task, the output is the total cost, which is a continuous variable. The task is a regression problem, and the objective is to minimize the error between the predicted and actual cost values.

$$\text{Objective 2: } \min_{\theta} L_{cost} = \text{total_cost}$$

3. Task 3: Predict total embodied carbon

In this task, the output is the total embodied carbon, which is a continuous variable. The task is a regression problem, and the objective is to minimize the error between the predicted and actual embodied carbon values.

$$\text{Objective 3: } \min_{\theta} L_{carbon} = \text{total_carbon_emissions}$$

4. Task 4 – Predict comfort days

In this task, the output is the total number of comfort days, which is a continuous variable. The task is a regression problem, and the objective is to minimize the error between the predicted and actual comfort days values.

$$\text{Objective 4: } \max_{\theta} L_{comfort} = \text{comfort_days}$$

Each task can be framed as a regression problem, with its own loss function to be minimized during training.

For each task, a separate loss function is defined. Since all tasks are regression-based, Mean Squared Error (MSE) is used as loss function for each task:

$$L_{energy}(y_{true}, y_{pred}) = \frac{1}{n} \sum_{i=1}^n (y_{true}^{energy}(i) - y_{pred}^{energy}(i))^2$$

$$L_{cost}(y_{true}, y_{pred}) = \frac{1}{n} \sum_{i=1}^n (y_{true}^{cost}(i) - y_{pred}^{cost}(i))^2$$

$$L_{carbon}(y_{true}, y_{pred}) = \frac{1}{n} \sum_{i=1}^n (y_{true}^{carbon}(i) - y_{pred}^{carbon}(i))^2$$

$$L_{comfort}(y_{true}, y_{pred}) = -\frac{1}{n} \sum_{i=1}^n (y_{true}^{comfort}(i) - y_{pred}^{comfort}(i))^2$$

where:

- y_{true} is the true value of each output.
- y_{pred} is the predicted value of each output.
- n is the number of data points

6.3.2 Early stopping mechanism

The early stopping mechanism is integrated into the model to enhance time efficiency and prevent overfitting during the training process. This method continuously monitors the model's performance on a validation set and **stops training prematurely if there is no significant improvement** over a specified number of consecutive epochs, referred to as "patience" (Smith et al., 2023). In this implementation, the patience is set to 10 epochs.

Mathematically, the early stopping criterion can be expressed as follows:

If $val_loss_t < best_loss - \delta$, then update $best_loss$ and reset counter

Else, increment counter by 1

If counter \geq patience, then trigger early stopping

Here, val_loss_t represents the validation loss at epoch t , $best_loss$ is the lowest recorded validation loss, and δ is a small threshold (set to 1×10^{-4}) that determines the minimum improvement required to consider the validation loss as having decreased significantly.

When the validation loss decreases by at least δ , the model's state is saved, and the counter is reset to zero, indicating an improvement in performance. If the validation loss does not improve by the specified over consecutive epochs equal to the patience parameter, the early stopping condition is met, and training is terminated. This approach not only conserves computational resources and reduces training time but also ensures that the model retains its ability to generalize to new, unseen data by preventing it from learning noise or irrelevant patterns in the training set.

By implementing early stopping, the model stops training at the point where the validation performance peaks, thereby avoiding the risk of overfitting. This strategic termination of the

training process maintains the model's robustness and enhances its generalization capabilities, ensuring reliable performance in selecting optimal retrofit solutions.

6.3.3 Weighted sum method

The weighted sum method is a traditional technique employed to synchronize the training process across multiple tasks within a Multi-Task Learning (MTL) framework. Building upon the previously defined MTL loss function, which aims to optimize several tasks simultaneously, the weighted sum method refines this approach by combining the individual loss functions of each task into a single, cohesive objective. This combination is achieved by assigning specific weights to each task's loss, thereby guiding the model to balance its performance across all tasks effectively.

Mathematically, the weighted sum method modifies the general MTL loss function as follows:

$$L_{total}(\theta) = \sum_{i=1}^4 \lambda_i L_i(\theta)$$

where:

- $L_t(\theta)$ represents the loss for task i , parameterized by θ .
- λ_i are weighting coefficients that determine the relative importance of each task's loss.

For this study, each weight λ_i is set to 0.25, ensuring that the sum of all weights equals one:

$$\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0.25$$

and

$$\sum_{i=1}^4 \lambda_i = 1$$

This equal weighting reflects the assumption that each of the four tasks contributes equally to the overall performance of the model. By assigning a weight of 0.25 to each task, the model ensures that no single task dominates the optimization process, thereby maintaining a balanced focus across all tasks. By equally weighting these tasks, the model is encouraged to develop a holistic understanding and performance across all dimensions.

The weighted sum method aligns with **traditional multi-objective optimization techniques**, where multiple objectives are combined into a single scalar objective to facilitate simultaneous optimization (GumGum Tech, 2020). However, it is important to acknowledge that this approach assumes prior knowledge of the relative importance of each task. In cases where the significance of tasks may vary or is not well-defined, **determining appropriate weights can be challenging** and may require empirical experimentation or domain-specific insights.

6.3.4 Multi-Gradient Descent Algorithm (MGDA)

In the pursuit of optimizing multiple tasks simultaneously within the MTL framework, the **Multiple Gradient Descent Algorithm (MGDA) emerges as a robust alternative to traditional weighting methods** (Sener et al., 2018). Unlike the Weighted Sum method, which necessitates the specification of task-specific weights, MGDA dynamically determines the optimal combination of gradients to ensure that all tasks are improved concurrently. This capability is particularly advantageous in scenarios where tasks may have conflicting objectives, as it facilitates the identification of Pareto optimal solutions - configurations where enhancing one task inherently compromises another.

MGDA operates by first computing the gradient of each task's loss function with respect to the shared model parameters. Let $L_t(\theta)$ denote the loss for task t , and $\nabla_{\theta}L_t(\theta)$ represent its corresponding gradient. The algorithm seeks to find a weighted combination of these gradients that points in a descent direction beneficial to all tasks. Mathematically, this is formulated as the following optimization problem:

$$\min_{\alpha_1, \alpha_2, \dots, \alpha_T} \left\| \sum_{t=1}^T \alpha_t \nabla_{\theta} L_t(\theta) \right\|^2 \quad \text{subject to} \quad \sum_{t=1}^T \alpha_t = 1$$

Here, α_t are the weights assigned to each task's gradient, and T represents the total number of tasks (Momma et al., 2022). The constraints ensure that the weights form a convex combination, promoting a balanced update that does not disproportionately favor any single task.

The implementation of MGDA within the MTL model for selecting optimal retrofit solutions involves several key steps:

1. **Gradient Computation:** For each task, the gradient $\nabla_{\theta}L_t(\theta)$ is calculated with respect to the shared parameters θ . This step captures the direction in which each task's loss can be minimized.
2. **Optimization Problem:** The core of MGDA lies in solving the aforementioned quadratic programming (QP) problem to determine the optimal weights θ_t . By minimizing the norm of the combined gradient, MGDA ensures that the update direction is as beneficial as possible for all tasks simultaneously.
3. **Descent Direction and Parameter Update:** Once the optimal weights are determined, the gradients are combined into a single vector:

$$\theta \leftarrow \theta - n \sum_{t=1}^T \alpha^t \nabla_{\theta} L^t(\theta)$$

where n is the learning rate. This unified update direction adjusts the shared parameters in a manner that accounts for the contributions of all tasks, thereby fostering balanced improvements.

The application of MGDA in this research addresses several limitations inherent to the Weighted Sum method. By eliminating the need for predefined weights, MGDA circumvents the often challenging task of weight tuning, which can be both time-consuming and suboptimal without domain-specific insights. Additionally, MGDA inherently manages conflicts between tasks by seeking a Pareto optimal solution, ensuring that improvements in one task do not disproportionately degrade another.

However, the implementation of MGDA is not without challenges. The primary concerns include **computational complexity** and scalability, especially when dealing with many tasks or complex models such as deep neural networks. Traditional MGDA requires separate backward passes for each task to compute their respective gradients, leading to increased computational overhead proportional to the number of tasks. **To mitigate this, an upper bound approximation of MGDA can be employed.** This approximation simplifies the gradient computation by leveraging a single backward pass across all tasks, thereby significantly reducing the computational burden while still approaching Pareto optimality.

In practice, the MGDA approach applied in this study involves the following process:

- **Loss Calculation:** For each of the four tasks - energy consumption (L_1), total cost (L_2), carbon emissions (L_3), and comfort days (L_4) - the respective loss $L_t(\theta)$ is computed.
- **Gradient Computation:** The gradient $\nabla_{\theta} L_t(\theta)$ for each task is determined with respect to the shared parameters θ .
- **Optimization Problem Solving:** The QP problem is solved to find the optimal weights α_t that minimize the combined gradient norm, subject to the constraints $\sum_{t=1}^4 \alpha_t = 1$ and $\alpha_t \geq 0$ for all t .
- **Parameter Update:** The shared parameters θ are updated using the weighted sum of the task-specific gradients, guided by the optimal weights α_t .

This methodological framework ensures that the MTL model maintains a balanced and effective optimization across all tasks. By dynamically adjusting the influence of each task's gradient, MGDA fosters a robust learning process that is resilient to task conflicts and capable of achieving comprehensive performance improvements.

6.4 Model evaluation

6.4.1 Model options evaluation

Four distinct multi-task learning models were developed: literature-based models using MGDA, data-based models using MGDA, literature-based models employing the weighted sum method, and data-based models employing the weighted sum method. The performance of each model is assessed using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R-squared (R^2) for each task (see Table 9, 10, 11, and 12), as well as average MAE, RMSE and R^2 across all tasks (see Table 8).

These metrics are defined as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

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

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

where:

- n represents the total number of observations.
- y_i is the actual value.

- \hat{y}_i is the predicted value.
- \bar{y} is the mean of the actual values.

A comparative analysis is conducted to demonstrate the relative efficacy of the training methods and architectural approaches employed.

Model option	MAE	RMSE	R2
Literature-based, MGDA	39.86	44.68	0.832
Data-based, MGDA	41.00	48.59	0.797
Literature-based, weighted sum	42.40	48.78	0.791
Data-based, weighted sum	42.86	49.74	0.787

Table 8 – Average MAE, RMSE and R² across all tasks of each of the four model's options considered (Source: Own Work)

Model option	MAE	RMSE	R ²
Literature-based, MGDA	1.14	1.47	0,920
Data-based, MGDA	1.26	1.49	0.918
Literature-based, weighted sum	1.14	1.56	0.910
Data-based, weighted sum	1.26	1.59	0.906

Table 9 - MAE, RMSE and R² for energy task of each of the four model's options considered (Source: Own Work)

Model option	MAE	RMSE	R ²
Literature-based, MGDA	136.08	149.49	0,610
Data-based, MGDA	136.49	160.18	0.552
Literature-based, weighted sum	140.38	160.48	0.540
Data-based, weighted sum	142.24	163.18	0.535

Table 10 - MAE, RMSE and R² for cost task of each of the four model's options considered (Source: Own Work)

Model option	MAE	RMSE	R ²
Literature-based, MGDA	17.03	25.06	0,773
Data-based, MGDA	20.39	26.48	0.746
Literature-based, weighted sum	21.70	27.13	0.734
Data-based, weighted sum	22.60	27.13	0.734

Table 11 - MAE, RMSE and R² for emission task of each of the four model's options considered (Source: Own Work)

Model option	MAE	RMSE	R ²
Literature-based, MGDA	6.37	7.84	0.969
Data-based, MGDA	5.97	7.10	0.967
Literature-based, weighted sum	5.34	6.24	0,958
Data-based, weighted sum	5.21	6.11	0.948

Table 12 - MAE, RMSE and R2 for comfort task of each of the four model's options considered (Source: Own Work)

Comparison of training approaches: MGDA and Weighted Sum Method

The first comparison focuses on the training methodologies—MGDA and the Weighted Sum method. The results indicate a clear performance advantage for models utilizing MGDA over those employing the Weighted Sum approach. Specifically, the literature-based architecture combined with MGDA achieves an average MAE of 39.86 and an average RMSE of 44.68, while its counterpart using the Weighted Sum method records higher errors (average MAE = 42.40, average RMSE = 48.78). Similarly, the data-based architecture with MGDA outperforms the Weighted Sum variant (average MAE = 41.00, average RMSE = 48.59 versus average MAE = 42.86, average RMSE = 49.74).

Examining the performance across individual tasks further substantiates the superiority of MGDA. For the energy task, both MGDA-based models exhibit lower MAE and RMSE compared to their Weighted Sum counterparts, with the literature-based MGDA model achieving an MAE of 1.14 and an RMSE of 1.47 versus 1.14 and 1.56 for the Weighted Sum method. In the cost task, MGDA consistently outperforms the Weighted Sum approach, with the literature-based MGDA model recording an MAE of 136.08 and an RMSE of 149.49 compared to 140.38 and 160.48, respectively. This trend is also observed in the emission task, where the literature-based MGDA model achieves an MAE of 17.03 and an RMSE of 25.06, outperforming the Weighted Sum method's MAE of 21.70 and RMSE of 27.13. However, for the comfort task, the Weighted Sum method shows competitive performance, with the data-based Weighted Sum model achieving slightly lower MAE and RMSE (5.21 and 6.11) compared to the MGDA-based models (MAE = 5.97, RMSE = 7.10).

The superior performance of MGDA can be attributed to its dynamic adjustment of task-specific gradients, which facilitates a balanced optimization across all tasks. Unlike the Weighted Sum method, which relies on predefined weights to aggregate task losses, MGDA autonomously determines the optimal combination of gradients. This dynamic weighting mechanism effectively mitigates conflicts between tasks, ensuring that improvements in one objective do not disproportionately detract from others. Consequently, MGDA fosters a more efficient convergence, as evidenced by the lower average MAE and RMSE values observed.

In contrast, the **Weighted Sum method** assigns equal weights to each task's loss function, operating under the assumption of equal task importance. While this approach **simplifies the optimization process**, it may fail to capture the specific interdependencies and varying degrees of task significance inherent in the dataset. The static weighting scheme can lead to suboptimal performance, particularly in scenarios where tasks exhibit conflicting gradients or differential relevance to the overall model objectives. The empirical results underscore the limitations of the Weighted Sum method, highlighting the necessity for more adaptive training strategies like MGDA in complex MTL settings.

Comparison of training approaches: MGDA and Weighted Sum Method

The second comparison examines the architectural choices: data-based versus literature-based task-specific layers. The findings reveal that **literature-based architectures consistently outperform their data-based** counterparts across both training methodologies. Specifically, under the MGDA framework, the literature-based model achieves an average MAE of 39.86 and an average RMSE of 44.68, compared to the data-based model's average MAE of 41.00 and RMSE of 48.59. Similarly, when employing the Weighted Sum method, literature-based architectures maintain superior performance (average MAE = 42.40, average RMSE = 48.78) relative to data-based architectures (average MAE = 42.86, average RMSE = 49.74).

Analyzing the performance across individual tasks further supports the advantage of literature-based architectures. For the energy task, both literature-based models demonstrate superior R^2 values (0.920 and 0.910) compared to the data-based models (0.918 and 0.906). In the cost task, literature-based models not only exhibit lower MAE and RMSE but also achieve higher R^2 values (0.610 and 0.540) compared to data-based models (0.552 and 0.535), indicating better explanatory power. The emission task also favors literature-based architectures, with higher R^2 values (0.773 and 0.734) versus data-based models (0.746 and 0.734). However, in the comfort task, data-based architectures perform competitively, particularly with the Weighted Sum method, where the data-based model achieves a slightly higher R^2 (0.948) compared to the literature-based counterpart (0.958).

The **enhanced performance of literature-based architectures can be attributed to their foundation on established academic insights and empirical evidence**. By leveraging architectural designs that have been validated in prior research, literature-based models benefit from optimized configurations tailored to the specificities of the tasks at hand. These architectures likely incorporate proven mechanisms for feature extraction, parameter sharing, and task-specific processing, which collectively contribute to more accurate and reliable predictions.

In contrast, **data-based architectures**, which are derived from correlations identified within the dataset, **may lack the theoretical robustness** and generalizability inherent in literature-informed designs. While data-driven approaches offer the flexibility to capture unique patterns and relationships specific to the dataset, they may also be susceptible to overfitting or fail to encapsulate broader task interactions effectively. The comparatively higher average MAE and RMSE values observed in data-based models suggest that these architectures may not fully exploit the underlying structures essential for optimal multi-task performance. Additionally, the lower R^2 values in several tasks indicate that data-based models may have less explanatory power in capturing the variance in the target variables.

Furthermore, literature-based architectures likely benefit from a comprehensive understanding of the domain, enabling more strategic parameter sharing and task prioritization. This informed architectural design facilitates better coordination among tasks, enhancing the model's ability to generalize and perform consistently across diverse predictive objectives. For instance, in the cost task, which exhibits the highest errors and lowest R^2 values, literature-based models demonstrate more robust performance, suggesting that domain-informed architectural choices play a crucial role in managing more challenging prediction tasks.

Conclusions

The comparative analysis underscores the critical impact of both training methodologies and architectural choices on the performance of MTL models. **MGDA emerges as a superior training method** relative to the Weighted Sum approach, primarily due to its ability to dynamically balance task-specific gradients and navigate conflicting objectives effectively. This

advantage is particularly evident in tasks with higher error rates, such as the cost task, where MGDA-based models consistently outperform Weighted Sum counterparts. Concurrently, **literature-based architectural designs demonstrate a consistent performance advantage over data-based architectures**, benefiting from theoretically grounded and empirically validated configurations.

These findings highlight the importance of adopting adaptive training strategies and leveraging established architectural principles in the development of robust MTL models.

Given these considerations, the **MTL model with literature- review based architectures using MGDA method has been used in the next steps of this research.**

6.4.2 Evaluation of best-ranked model

Figure 22 shows the evolution of the loss function during both training and validation phases for the model with architectures based on literature-review and using MGDA.

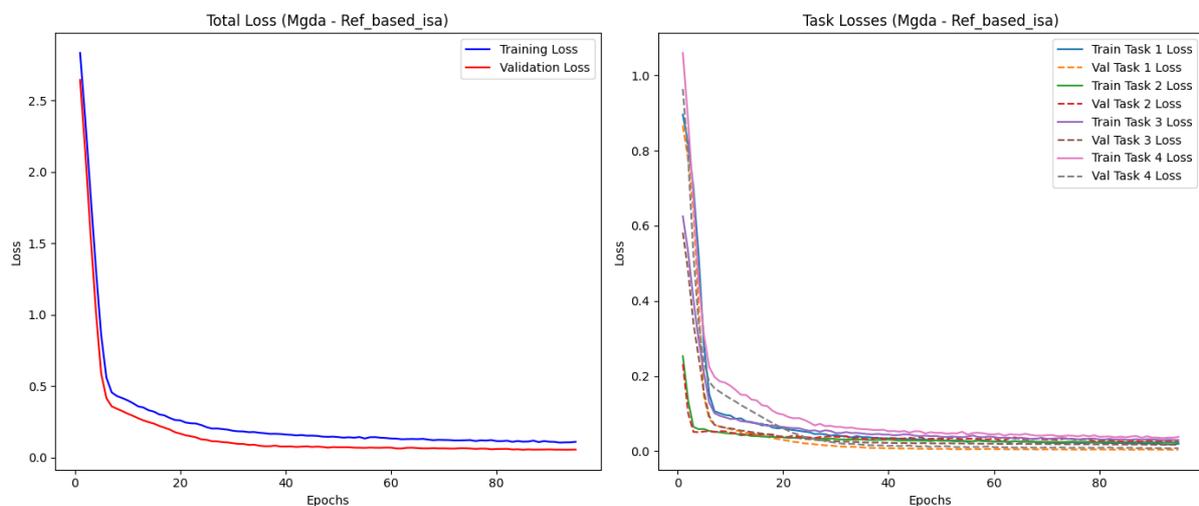


Figure 22 - Training vs. validation losses graphs for best performing model, where task 1 is energy, task 2 is cost, task 3 is carbon and task 4 is comfort (Source: Own Work)

The plot on the left illustrates how the total loss changes over the training epochs for both the training set (blue line) and the validation set (red line). At the beginning of the training process, the total loss is relatively high, but it decreases quickly as the model learns from the data. Within the first 20 epochs, the loss rapidly converges to substantially lower values, indicating that the model's predictive capability is improving. As training progresses, the training loss continues to decrease at a slower pace, while the validation loss stabilizes at a level comparable to or even slightly lower than the training loss. The absence of a significant gap between training and validation losses, combined with the lack of a late increase in validation loss, suggests that the model is **not overfitting**. So, the **generalization ability of the model remains strong**, allowing it to perform well on unseen data.

The plot on the right shows how each individual task's loss evolves over time for both the training and validation sets. Each task's training loss (solid lines) and validation loss (dashed lines) follow a similar pattern: an initial steep decline, followed by stabilization at relatively low values. Notably, the close correspondence between training and validation curves across all tasks indicates that the **multi-objective optimization strategy is well managed**. This **effective balance is achieved thanks to the use of the Multi-Gradient Descent Algorithm (MGDA)**, which ensures that the model does not disproportionately prioritize one objective over the others. Consequently, all tasks converge together towards a stable solution without significant trade-offs or degradation in performance on any single objective.

Figure 23 shows plots that compare the actual target values (horizontal axis) to the model's predicted values (vertical axis) for each of the four tasks. Each plot includes a red dashed line representing the ideal scenario where predicted and actual values match exactly.

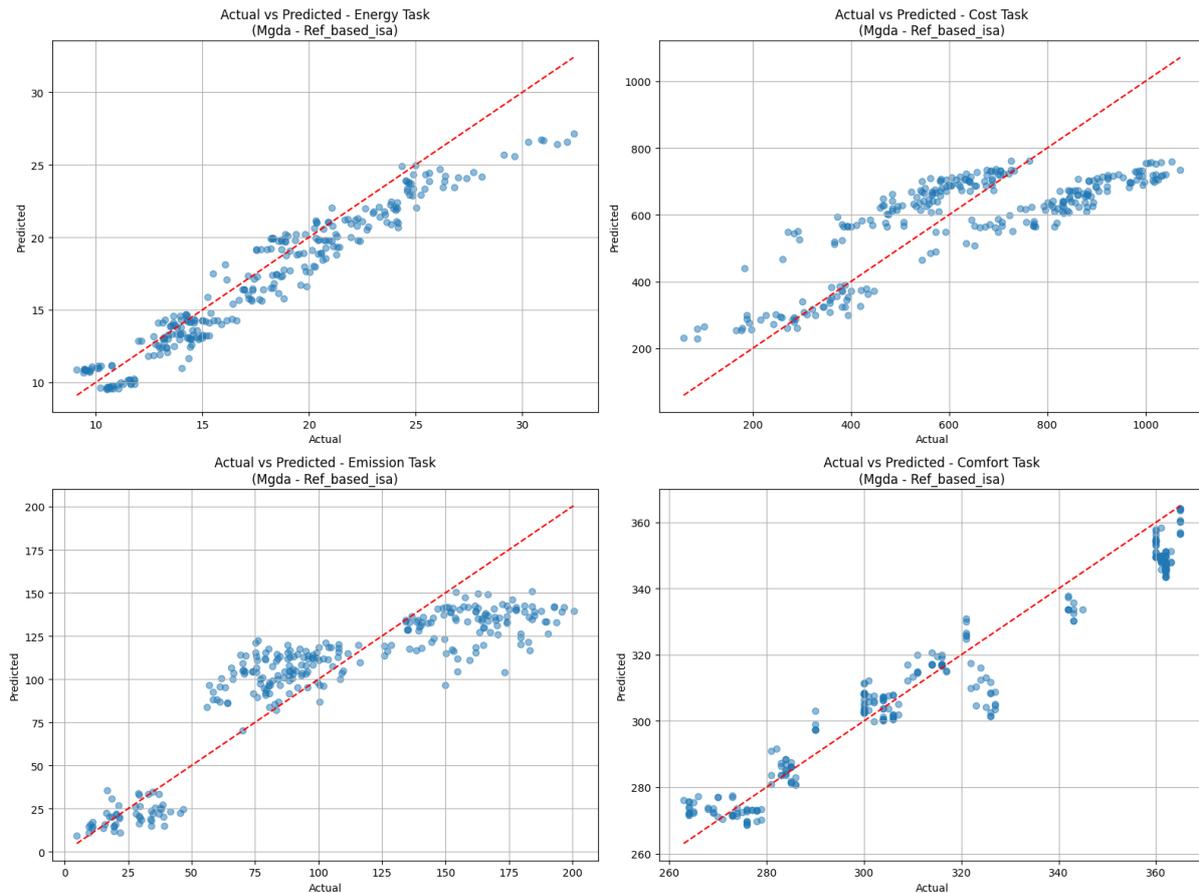


Figure 23 - Predicted vs. actual values graphs for best performing model (Source: Own Work)

Energy task

In the energy prediction plot, most data points cluster closely around the diagonal line, indicating that the model's **predictions are generally well aligned with the actual values**. The spread of points is relatively narrow, suggesting a high level of accuracy and a strong correlation between predicted and actual energy consumption. While minor deviations exist, the model appears to capture the underlying patterns in energy demand effectively.

Cost task

For the cost predictions, the points also follow the diagonal trend, but the scatter broadens somewhat, especially at higher cost values. Despite this increased variance, the model still maintains a positive relationship between actual and predicted costs. This indicates that while the model is less precise as cost values grow larger, it still **successfully understands the main trends**. Some under- or overestimation is visible, showing potential areas for further fine-tuning or additional feature engineering.

Emissions task

In the emissions task, there is an overall alignment with the diagonal line, indicating that total carbon emissions are predicted reasonably well. However, the predicted values are notably divided into **three distinct clusters**. The three clusters observed likely **correspond to the different retrofit scenarios** considered in the dataset. The first, representing the current scenario without any retrofit, incurs minimal environmental impact since no additional materials are introduced, resulting in the lowest emissions cluster. The second cluster,

associated with a standard retrofit, displays a moderate level of environmental impact as it involves an average amount of insulation material added. Lastly, the cluster corresponding to the nZEB (near zero-energy building) scenario shows the highest environmental impact, reflecting the big quantity of material required. Further investigation could help clarify why the data form these clusters and improve accuracy. **Additional checks or refinement** may be necessary to ensure that the model is capturing the nuances that drive the differences in embodied carbon.

Comfort task

For the comfort task, the predictions also show a **good positive correlation** with the actual values. Most points cluster near the line, indicating that the model reliably estimates the number of comfort days. There is some clustering and slight deviation in certain value ranges, which might indicate that the model tends to slightly underpredict or overpredict comfort days under certain conditions. Still, the general trend suggests that the model is capable of providing useful guidance for predicting comfort days.

Figure 24 shows the percentage error distribution graphs of each task for the selected model.

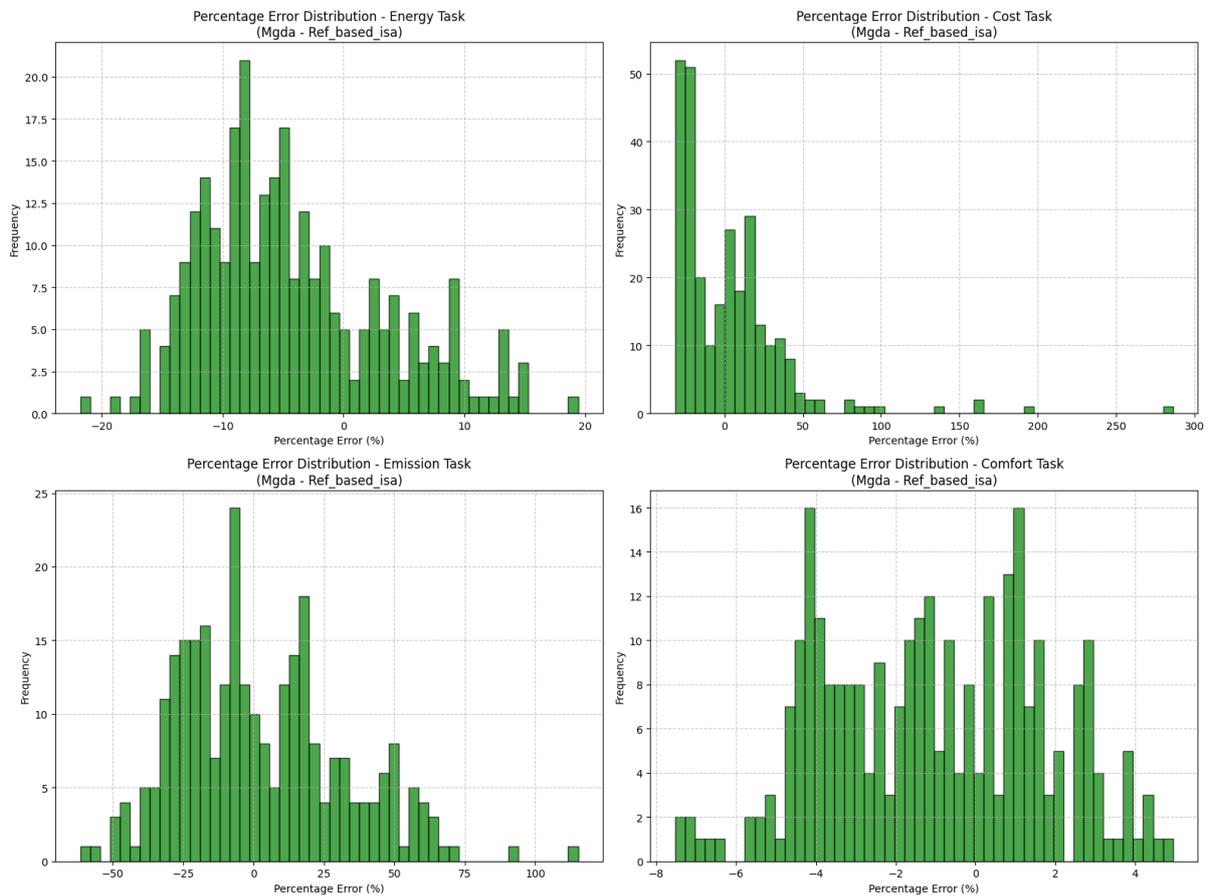


Figure 24 - Percentage error distribution graphs for best performing model (Source: Own Work)

Energy task

The error distribution for energy predictions is relatively symmetric and centered near zero, with the majority of errors falling between about -15% and +5%. This suggests that the **model generally estimates energy consumption reasonably well**, with minimal systematic bias toward over- or underestimation.

Cost task

For the cost predictions, there is a good number of predictions with errors close to zero, suggesting that the model often gets the costs right. However, there are also cases where the errors are larger. Such **deviations** may arise because the model struggles to fully comprehend the relationship between retrofit extent and material choice. For example, when a retrofit scenario is identical (meaning it reaches the same thermal properties) except for the insulation material used, the model may **fail to recognize that a standard insulation material results in a lower cost, while a low embodied carbon material leads to a higher cost**. Learning this complex correlation may require exposing the model to a more comprehensive dataset that includes a variety of standard and low embodied carbon materials, along with their respective price ranges. Such data enrichment could help the model better understand and predict these cost variations.

Emissions task

The **distribution of emissions errors is broader**, indicating that the model struggles to accurately predict embodied carbon. This expanded range of errors likely reflects the complexity of factors influencing emissions, such as material selections, retrofit scenarios, and the interplay of multiple parameters. In particular, considering only five different materials may have limited the model's ability to fully interpret the relationship between retrofit measures and embodied carbon. Without a sufficiently diverse training dataset that includes a wider array of both standard and low-embodied-carbon materials, the model cannot fully capture these intricate dynamics. **By expanding the dataset** to encompass more varied materials, the **model's capacity to predict emissions should improve**. Such enhancements would likely lead to a more narrowly clustered error distribution, signaling more precise and reliable emissions forecasts.

Comfort task

The comfort task exhibits the strongest performance among the four tasks considered. Specifically, all prediction errors fall within a range of -8% to +5%. The error distribution is relatively narrow and slightly below zero, indicating a small tendency to underestimate actual comfort days. The majority errors remain confined to a small percentage range, suggesting that the model provides consistently accurate predictions for comfort-related outcomes.

Overall conclusions

The evaluation of the MTL model across energy, cost, emissions, and comfort tasks underscores its **good performance and generalization capabilities**. Analysis of the loss functions (Figure 22) reveals a consistent decrease in both training and validation losses, with minimal overfitting, referable to the **effective multi-objective optimization provided by the Multi-Gradient Descent Algorithm (MGDA)**. This balance ensures that the model maintains high performance across all tasks without disproportionately favouring any single objective. In terms of predictive accuracy (Figure 23), the model **performs good in energy and comfort tasks**, demonstrating tight clustering around the ideal prediction line, which reflects high reliability and strong correlation with actual values. While cost predictions generally follow the expected trend, increased variance at higher values suggests the need for further refinement to capture complex cost dynamics more precisely. The emissions task, although aligned with actual values, exhibits broader error distributions and distinct clustering, indicating challenges in accurately predicting embodied carbon due to limited material diversity in the training data. Error distribution analyses (Figure 24) reinforce these findings, highlighting minimal bias and narrow error ranges for energy and comfort, while **cost and emissions predictions reveal areas for improvement**. Notably, the predictive accuracy for cost and emissions can be enhanced by **expanding the dataset to encompass a wider array of both standard and low-embodied-carbon materials** along with their related costs. Overall, the model demonstrates significant predictive strengths and effective generalization, with opportunities for

enhancement through expanded datasets and refined feature engineering to address identified limitations. Future work should focus on these areas to further improve the model’s accuracy and reliability across all tasks.

6.4.3 Time efficiency

A significant innovation presented in this study is the substantial reduction in computational time required for assessing retrofit scenarios. This enhancement is pivotal for accelerating the retrofit process, enabling timely and efficient decision-making. To quantify these time savings, four comparative analyses were conducted between the developed Multi-Task Learning (MTL) model and the conventional retrofit process. The following sections detail these comparisons and their outcomes.

1. Simulation time per run

The first comparison evaluates the time required to execute a single simulation using both approaches. The conventional retrofit process involves running the simulation in EnergyPlus. As shown from Table 13, the MTL model reduces the simulation time from **38 seconds** to **0.3 seconds**, representing a **125-fold** decrease in computational time.

Usual retrofit process	MTL Model
38 s	0.3 s

Table 13 - Comparison of needed time to run one simulation (Source: Own Work)

2. Simulation time including IDF creation

The second comparison incorporates the time required to prepare the IDF file alongside the simulation run. For the conventional process, creating the IDF file manually demands around 8 hours, whereas the MTL model requires minimal effort. As shown from Table 14, when accounting for IDF file creation, the MTL model completes the process in **5.3 seconds** compared to **8 hours and 38 seconds** for the conventional method. This signifies a reduction from approximately **30,628 seconds** to **5.3 seconds**, highlighting an immense time-saving advantage.

	Usual retrofit process	MTL Model
Calculation	8 h + 38 s	5 s + 0.3 s
Result	8h 38 s	5.3 s

Table 14 - Comparison of needed time to create one idf file and run one simulation (Source: Own Work)

3. Input modelling and computational time for one district

Expanding the comparison to a district level demonstrates even greater time efficiencies. Assuming a district comprises approximately **60,000 buildings**, the time savings become more pronounced. In the conventional retrofit process, creating an IDF file for each retrofit scenario requires an additional **5 minutes** per building. This significant time investment accumulates rapidly when scaled to a large number of buildings. As shown from Table 15, for a single district the MTL model requires **3 days and 16 hours**, whereas the conventional process would take approximately **235 days and 1 hour**. This comparison underscores the MTL model's capability to drastically reduce the time from nearly eight months to under four days by eliminating the need for extensive manual IDF file creation for each retrofit scenario.

	Usual retrofit process	MTL Model
--	------------------------	-----------

Calculation	$(8 \text{ h} + 38 \text{ s}) + (5 \text{ min} + 38 \text{ s}) \times 60\,000$	$(5 \text{ s} + 0.3 \text{ s}) \times 60\,000$
Result	235 days and 1 hour	3 days and 16 hours

*Table 15 - Comparison of needed time to create one idf file and run simulation for a district
(Source: Own Work)*

4. Input modelling and computational time for one city

The final comparison scales the analysis to the entire city of Amsterdam, encompassing around **600,000 buildings**. This scenario illustrates the model's potential impact on a large urban scale. As shown in Table 16, applying the MTL model to the entire city reduces the required time to **36 days and 20 hours**, in stark contrast to the conventional approach, which would necessitate approximately **1,958 days and 8 hours**. This dramatic reduction highlights the MTL model's scalability and efficiency for large-scale retrofit assessments.

	Usual retrofit process	MTL Model
Calculation	$(8 \text{ h} + 38 \text{ s}) + (5 \text{ min} + 38 \text{ s}) \times 499\,999$	$(5 \text{ sec} + 0.3 \text{ sec}) \times 600\,000$
Result	1958 days and 8 h	36 days and 20 hours

*Table 16 - Comparison of needed time to create one idf file and run simulation for one city
(Source: Own Work)*

The comparative analyses clearly demonstrate that the use of the **MTL model offers significant time savings across various scales of application**. Whether for single simulations or city-wide assessments, the MTL model consistently outperforms the traditional retrofit process by drastically reducing computational time. Specifically, a **single simulation** with the MTL model is completed approximately **125 times faster** than with EnergyPlus. For **larger-scale applications**, such as district or city-wide assessments, the MTL model achieves time reductions ranging from over **60 times to nearly 54 times faster**. These enhancements not only improve the efficiency of retrofit assessments but also facilitate more rapid and informed decision-making, enabling the timely implementation of energy-efficient measures.

However, it is important to recognize that while the MTL model significantly accelerates the simulation process, EnergyPlus simulations provide only indoor temperature and energy consumption data. In contrast, the MTL model offers additional metrics such as comfort days, cost, and embodied carbon. Therefore, the comparison between the two methods is valid only to a certain extent, based on the specific metrics each provides. Additionally, the computational time required for EnergyPlus simulations can vary greatly depending on the complexity of the IDF file. For the purposes of this comparison, the time measured pertains specifically to the IDF file of the base building used in this study.

7. Selection of optimal retrofit solutions

This study aims to determine the most effective building retrofit strategies by simultaneously considering multiple objectives, including cost, carbon emissions, energy consumption, and the number of days within acceptable comfort levels. To manage these often conflicting objectives, this research employs **multi-objective optimization (MOO)** and **multi-criteria decision-making (MCDM)** methods.

MOO is used first to generate a **set of optimal trade-off solutions**, represented as a "Pareto front." Rather than focusing on a single objective, MOO finds a range of solutions that balance the four criteria, ensuring that no single aspect (e.g., minimizing cost) compromises another objective (e.g., reducing emissions) excessively.

After MOO has identified a spectrum of potential solutions, **MCDM** techniques are employed to **select the best individual solution** from this set. MCDM methods systematically evaluate the trade-off solutions against a specific set of weighted criteria, making it possible to prioritize certain objectives over others and thereby determine a single optimal scenario.

In this study, two different MOO approaches are developed to enhance the robustness and flexibility of the optimization process. Each approach interprets the problem differently, thereby broadening the range of potential solutions. Further explanation of these methods is provided in the following sections (see Figure 11).

7.1 Multi-Objectives Optimization

7.1.1 Optimization based on dataset

In the pursuit of identifying optimal retrofit solutions, it is essential to balance multiple, often conflicting objectives such as minimizing costs, reducing carbon emissions, lowering energy consumption, and maximizing occupant comfort. To achieve this, a decision-making framework utilizing MOO is employed. This chapter delves into the first type of MOO applied in this study, detailing the step-by-step methodology, the mathematical foundations, and the implementation specifics that underpin the identification of Pareto-optimal retrofit scenarios.

The optimization problem is structured around **four primary objectives**:

1. **minimizing annual energy consumption,**
2. **minimizing total retrofit cost,**
3. **minimizing embodied carbon, and**
4. **maximizing comfort days.**

These objectives are influenced by several decision variables, including the time horizon for retrofit measures, windows U-factor, ground floor thermal resistance, external walls thermal resistance, and roof thermal resistance. Given the conflicting nature of these objectives (for instance, reducing energy consumption and carbon emissions may lead to increased retrofit costs) a multi-objective optimization framework is crucial for exploring and identifying the most effective retrofit strategies that offer balanced improvements across all criteria.

The foundation of the MOO approach lies in the predictions generated by the MTL model. This model forecasts the four objectives based on different retrofit scenarios, producing a dataset structured as a NumPy array with dimensions corresponding to the number of samples and the four objectives. To facilitate the optimization process, the objectives are extracted and appropriately transformed. Notably, since MOO frameworks typically operate under a minimization paradigm, the comfort days objective, which is inherently a maximization criterion, is converted into a minimization objective by negating its values. This transformation ensures consistency across all objectives, allowing them to be compared and optimized simultaneously.

Once the objectives are prepared, they are consolidated into a structured format using a Pandas DataFrame. This consolidation includes both the decision variables and the transformed objectives, enabling a comprehensive analysis of how different retrofit strategies impact each objective. Organizing the data in this manner simplifies the management and manipulation of information during the optimization process, with each row in the DataFrame representing a specific retrofit scenario detailing the unique configuration of decision variables and the corresponding values of each objective.

A custom Pareto efficiency function is then employed to identify non-dominated solutions within the dataset. A solution is **deemed Pareto-efficient if there exists no other solution that improves one objective without causing a detriment to at least one other objective**. Mathematically, a solution x is Pareto-efficient if there does not exist another solution x' such that:

$$\forall i \in \{1, 2, 3, 4\}, f_i(x') \leq f_i(x)$$

and $\exists j \in \{1, 2, 3, 4\}$ such that $f_j(x') < f_j(x)$

where:

- f_i represents the objective functions:
 - f_1 annual energy consumption
 - f_2 total retrofit cost
 - f_3 total embodied carbon
 - f_4 comfort days (transformed to $f_4' = -f_4$ for minimization)

By transforming all objectives to a minimization framework, the optimization process becomes more straightforward, allowing for uniform application of the Pareto efficiency condition across all objectives.

The Pareto efficiency function operates by first adjusting the objectives designated for maximization, such as comfort days, by negating their values to convert them into minimization objectives. Subsequently, it performs a dominance check for each solution in the dataset. Domination is defined by one solution being at least as good as another in all objectives and strictly better in at least one objective. An efficiency mask is then generated, marking each solution as Pareto-efficient or not based on these dominance checks.

Applying this Pareto efficiency function to the consolidated dataset allows for the identification of Pareto-optimal solutions, forming the Pareto front. This front represents the set of **best trade-offs among the four objectives**, where each solution offers a unique combination of objectives such that improving one would necessitate compromising another.

The implementation of this MOO approach is executed using the "Pymoo: Multi-objective Optimization in Python" (n.d.-b) **library is referenced as primary resource**. Key parameters in this approach include setting the maximize parameter to [False, False, False, True], indicating that only the fourth objective (comfort days) should be maximized, while the remaining objectives are treated as minimization criteria.

7.1.2 Optimization based on constraints

Building upon the initial Multi-Objective Optimization approach discussed earlier, this chapter introduces a different methodology that leverages the **Non-dominated Sorting Genetic Algorithm II (NSGA-II)**. This approach retains the core objectives of minimizing costs, carbon emissions, and energy consumption while maximizing occupant comfort. Additionally, it offers the ability to **explore novel retrofit scenarios beyond those available in the initial dataset**.

The optimization framework focuses on four primary objectives: minimizing annual energy consumption, minimizing total retrofit cost, minimizing embodied carbon, and maximizing comfort days. These objectives are influenced by five decision variables: the Time Horizon for retrofit measures, windows U-Factor, ground floor thermal resistance, external walls thermal resistance, and roof thermal resistance.

Unlike the **first MOO approach, which relied solely on existing scenarios from the dataset** used to train the MTL model, the NSGA-II-based method allows users to define ranges for these decision variables. This flexibility enables the generation of innovative retrofit scenarios that were not previously considered, thereby expanding the solution space and uncovering configurations that may offer superior performance across multiple objectives.

The foundation of this approach is the NSGA-II algorithm, following the considerations of Zhan et al. (2024). By enabling user-specified ranges for each decision variable, the algorithm can explore a broader solution space, potentially uncovering retrofit configurations that offer superior performance across multiple objectives.

Mathematically, the optimization problem is formalized as follows:

$$\begin{aligned} \text{Minimize}_{\mathbf{x}} \quad \mathbf{F}(\mathbf{x}) &= \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ f_3(\mathbf{x}) \\ -f_4(\mathbf{x}) \end{bmatrix} \\ \text{Subject to} \quad \mathbf{x} &\in \mathcal{X} \end{aligned}$$

where:

- $\mathbf{x} = [x_1, x_2, x_3, x_4, x_5]$ represents the decision variables:
 - x_1 time horizon
 - x_2 windows U-factor
 - x_3 external walls thermal resistance
 - x_4 roof thermal resistance
- $f_1(x)$ is annual energy consumption

- $f_2(x)$ is total retrofit cost
- $f_3(x)$ is total embodied carbon
- $f_4(x)$ is comfort days (transformed to $-f_4(x)$ for minimization)

The transformation of the Comfort Days objective into a minimization problem $-f_4(x)$ standardizes the objectives, facilitating uniform application of the optimization algorithm. The decision variables are constrained within predefined bounds, defining the feasible region X for the optimization process.

Implementing the NSGA-II-based MOO approach involves several key steps. **Initially, users specify the minimum and maximum values for each decision variable**, allowing the exploration of a wide array of retrofit scenarios beyond the initial dataset. This user-defined flexibility is crucial for uncovering innovative configurations that may offer enhanced performance.

Next, the **pre-trained MTL model and associated scalars are loaded** to ensure accurate predictions of the four objectives based on various retrofit configurations. **A custom optimization problem** is then defined by extending the *ElementwiseProblem* class from "Pymoo: Multi-objective Optimization in Python" (n.d.-b). This custom class encapsulates the decision variables, objective functions, and constraints, facilitating seamless integration with the NSGA-II algorithm.

The NSGA-II algorithm is configured with specific parameters, including a **population size of 100 and 200 generations**. NSGA-II maintains a diverse population of solutions by **applying genetic operators such as selection, crossover, and mutation** to evolve the population toward optimality. The algorithm prioritizes solutions based on Pareto dominance, ensuring that the Pareto front encompasses a wide range of trade-off scenarios. At each generation, non-dominated sorting ranks solutions, and crowding distances are calculated to maintain diversity. Selected solutions are reproduced through genetic operations, and the population is updated by merging parent and offspring populations, retaining the top solutions based on their ranks and crowding distances. This process continues for the specified number of generations, resulting in a set of Pareto-optimal solutions that represent the best trade-offs among the defined objectives.

Upon completion of the optimization process, the Pareto-optimal solutions are extracted, encompassing both the decision variables and the corresponding objective values. These solutions form the Pareto front, providing a comprehensive set of optimal trade-offs.

7.2 Multi-Criteria Decision Making

Following MOO, Multi-Criteria Decision Making is employed to evaluate and select the most suitable solution from the Pareto set determined by MOO. MCDM facilitates the decision-making process by systematically considering the trade-offs between the different objectives to identify an optimal compromise that aligns with the user's preferences.

MCDM is a methodological framework that addresses decision-making scenarios involving multiple, often conflicting, criteria. Unlike single-objective optimization, which focuses on optimizing one criterion, MCDM acknowledges the complexity of real-world decisions where multiple factors must be simultaneously considered. **The process begins with the**

identification and evaluation of relevant criteria, followed by the assessment of various alternatives based on these criteria. The ultimate goal of MCDM is to synthesize the information from multiple objectives to support informed and balanced decision-making. This is particularly pertinent in the context of selecting retrofit solutions, where improvements in energy efficiency must be weighed against costs, comfort levels, and environmental impacts.

In the implemented MCDM process using Pymoo: Multi-objective Optimization in Python (n.d.-b), both Compromise Programming and Pseudo Weights methodologies are employed to evaluate the Pareto set generated by MOO. Compromise Programming utilizes the manually defined weights to identify the best compromise solution by minimizing the weighted deviations from the ideal point. Concurrently, the Pseudo Weights method dynamically adjusts these weights based on the performance of each solution, thereby refining the selection process to better reflect the inherent trade-offs among the objectives.

In this study, two MCDM approaches are developed:

1. **Compromise programming with user-assigned weights:** This approach involves users assigning weights to each of the four tasks based on their preferences. For example, if a user is highly interested in energy efficiency, they would assign a higher weight to the energy consumption task.
2. **Compromise programming with pseudo weights:** This method combines Compromise Programming with pseudo weights, which are dynamically derived based on the performance of solutions relative to each criterion, in addition to the user-assigned weights.

7.2.1 *Compromise programming*

Compromise programming is a significant method within the MCDM framework **aimed at finding a balanced solution that minimizes the distance to an ideal point.** The ideal point represents the most favourable levels of all criteria, where each objective attains its best possible value. Compromise programming operates on the principle that optimal decisions often involve trade-offs among conflicting objectives rather than the optimization of a single criterion. In this approach, **weights are assigned to each objective** to reflect their relative importance in the decision-making process, typically determined by the user based on their preferences. For instance, if a user prioritizes energy consumption over cost, a higher weight is assigned to the energy consumption criterion. These weights guide the selection of the most balanced solution within the Pareto set by minimizing the weighted deviations from the ideal point, ensuring alignment with the user's priorities.

7.2.2 *Pseudo-weights*

Pseudo weights complement compromise programming by providing a computational method to determine the significance of each criterion without relying solely on subjective judgments. Instead of requiring explicit weight assignments from the user, Pseudo weights are dynamically derived based on the performance of solutions relative to each criterion. This is achieved by normalizing the distance to the worst solution for each objective using the following formula (Pymoo: Multi-objective Optimization in Python, n.d.-b):

$$w_i = \frac{(f_i^{max} - f_i(x)) / (f_i^{max} - f_i^{min})}{\sum_{m=1}^M (f_i^{max} - f_i(x)) / (f_i^{max} - f_i^{min})}$$

This equation calculates the normalized distance to the worst solution regarding each objective i . It ensures that objectives with smaller deviations from the ideal point receive higher weights, thereby emphasizing their importance in the selection process.

By incorporating Pseudo weights, the MCDM process gains flexibility and adaptability, allowing the weights to adjust in response to the evolving understanding of the solution space. This reduces the subjectivity inherent in traditional weighting methods and enhances the robustness of the selected solution by ensuring that no single criterion disproportionately influences the outcome. The integration of Pseudo Weights with Compromise Programming ensures that the decision-making process **not only reflects the initial user preferences but also balances the inherent trade-offs among the objectives more effectively**, leading to a more resilient and well-rounded selection of retrofit solutions.

8. Results

8.1 Optimization based on dataset

The main research question of this project aims to assess the optimal retrofit solution, where "optimal" is defined based on client preferences. In the optimization based on dataset, "optimal" is translated into the weights given to the four objectives of this research. While many preferences could be considered, this research focused on a selection of - what is considered to be - the most realistic client requests. It is important to note that no task was assigned a weight less than 0.1 out of 1, ensuring that no single task was disproportionately undervalued. Furthermore, the results presented below are derived from the MCDM process, using compromise programming together with pseudo weights.

The scenarios assessed are:

- **Option 1: Minimizing Cost**

This option targets clients with limited financial resources who still wish to retrofit their building. The lowest intervention cost was explored by assigning equal low weights (0.1) to energy consumption, embodied carbon, and comfort, with the highest possible weight (0.7) assigned to cost. The results of this option are shown in Table 17.

Option 1							
Thermal properties				4 tasks predictions			
Rc façade (m ² K/W)	Rc roof (m ² K/W)	Rc ground floor (m ² K/W)	U Factor windows (W/m ² K)	Annual energy consumption (kWh/m ²)	Total cost (€/m ²)	Total embodied carbon (kgCO _{2e} /m ²)	Comfort days
0,45	0,48	5,50	2,90	123	108	7	322

Table 17 - Results for Option 1 (source: Own Source)

- **Option 2: Minimizing Embodied Carbon**

This scenario is for clients interested in minimizing environmental impact in terms of embodied carbon. It involves assigning low weights (0.1) to energy consumption, comfort, and cost, with the highest weight (0.7) to embodied carbon. The results of this option are shown in Table 18.

Option 2							
Thermal properties				4 tasks predictions			
Rc façade (m ² K/W)	Rc roof (m ² K/W)	Rc ground floor (m ² K/W)	U Factor windows (W/m ² K)	Annual energy consumption (kWh/m ²)	Total cost (€/m ²)	Total embodied carbon (kgCO _{2e} /m ²)	Comfort days
4,40	0,48	0,41	2,90	82	176	4	293

Table 18 - Results for Option 2 (Source: Own Source)

- **Option 3: Minimizing Annual Energy Consumption**

This option addresses clients whose primary goal is the energy efficiency of their building. A weight of 0.7 was assigned to the energy consumption task, and 0.1 to the other three tasks. The results of this option are shown in Table 19.

Option 3							
Thermal properties				4 tasks predictions			
Rc façade (m ² K/W)	Rc roof (m ² K/W)	Rc ground floor (m ² K/W)	U Factor windows (W/m ² K)	Annual energy consumption (kWh/m ²)	Total cost (€/m ²)	Total embodied carbon (kgCO _{2e} /m ²)	Comfort days
6,75	4,70	4,80	0,81	35	655	191	316

Table 19 - Results for Option 3 (Source: Own Source)

- **Option 4: Maximizing Comfort Days**

For clients focused on maximizing indoor comfort, this option assigns a weight of 0.7 to the comfort task and 0.1 to the other three tasks. The results of this option are shown in Table 20.

Option 4							
Thermal properties				4 tasks predictions			
Rc façade (m ² K/W)	Rc roof (m ² K/W)	Rc ground floor (m ² K/W)	U Factor windows (W/m ² K)	Annual energy consumption (kWh/m ²)	Total cost (€/m ²)	Total embodied carbon (kgCO _{2e} /m ²)	Comfort days
4,40	8,70	0,41	1,20	50	464	92	358

Table 20 - Results for Option 4 (Source: Own Source)

- **Option 5: All Objectives Treated Equally**

This scenario assumes a client equally interested in all four objectives, assigning an equal weight of 0.25 to each task. The results of this option are shown in Table 21.

Option 5							
Thermal properties				4 tasks predictions			
Rc façade (m ² K/W)	Rc roof (m ² K/W)	Rc ground floor (m ² K/W)	U Factor windows (W/m ² K)	Annual energy consumption (kWh/m ²)	Total cost (€/m ²)	Total embodied carbon (kgCO _{2e} /m ²)	Comfort days
6,70	0,48	0,41	1,20	71	406	158	316

Table 21 - Results for Option 5 (Source: Own Source)

8.2 Optimization based on constraints

Unlike the methods discussed in the previous chapter, optimization based on constraints defines "optimal" not only as a solution that meets the task weights assigned by the client but also excels within the client's specified constraints. This means that a client can define both an upper and lower bound for the thermal transmittance and U-Factor of the four envelope parameters. Additionally, the client can specify a preferred time horizon.

While numerous client preferences can be accommodated, this research primarily utilized constraint-based optimization to highlight differences in the retrofit solutions suggested across three different yearly scenarios. Consequently, three analyses were conducted with the same constraints on the thermal properties of the envelope parameters but set against three different time horizons. The ranges of the thermal properties defined for each option are presented in Table 22. These ranges were selected based on the lowest and highest values from the retrofit interventions studied in Chapter 3.2, rounded up for simplicity. Furthermore, each task is weighted equally, with each having a weight of 0.25.

	Lower bound	Upper bound
Windows U Factor (W/m ² K)	0,8	3,0
Groundfloor Rc (m ² K/W)	0,4	6,0
External Wall Rc (m ² K/W)	0,4	7,0
Roof Rc (m ² K/W)	0,4	9,0

Table 22 - Ranges of values considered as input for constraint-based optimization (source: Own Source)

- **Option 1: Best scenario for 2020**

The results of the constraint-based optimization for 2020 scenario are shown in Table 23.

Option 1							
Thermal properties				4 tasks predictions			
Rc façade (m ² K/W)	Rc roof (m ² K/W)	Rc ground floor (m ² K/W)	U Factor windows (W/m ² K)	Annual energy consumption (kWh/m ²)	Total cost (€/m ²)	Total embodied carbon (kgCO _{2e} /m ²)	Comfort days
6,70	0,48	0,41	0,80	71	406	158	326

Table 23 - Results for Option 1 (Source: Own Source)

- **Option 2: Best scenario for 2050**

The results of the constraint-based optimization for 2050 scenario are shown in Table 24.

Option 2							
Thermal properties				4 tasks predictions			
Rc façade (m ² K/W)	Rc roof (m ² K/W)	Rc ground floor (m ² K/W)	U Factor windows (W/m ² K)	Annual energy consumption (kWh/m ²)	Total cost (€/m ²)	Total embodied carbon (kgCO _{2e} /m ²)	Comfort days
6,7	0,48	0,41	0,8	71	406	158	318

Table 24 - Results for Option 2 (Source: Own Source)

- **Option 3: Best scenario for 2100**

The results of the constraint-based optimization for 2050 scenario are shown in Table 25.

Option 3							
Thermal properties				4 tasks predictions			

Rc façade (m ² K/W)	Rc roof (m ² K/W)	Rc ground floor (m ² K/W)	U Factor windows (W/m ² K)	Annual energy consumption kWh/m ²	Total cost (€/m ²)	Total embodied carbon (kgCO _{2e} /m ²)	Comfort days
6,70	0,48	0,41	0,80	69	406	158	304

Table 25 - Results for Option 3 (Source: Own Source)

9. Results discussion

Once the model selects the best combination of thermal properties based on client preferences, these values can be translated into technical details for the retrofit intervention. This interpretation is conducted based on the engineering judgment of the author, taking into account the dataset with which the model was trained, validated, and tested. The retrofit interventions suggested in the interpretation of results thus refer to the retrofit scenarios examined in Chapter 3.2.

Specifically, in the case of optimization based on dataset, the retrofit measures suggested by the model are chosen from those provided as input - namely, those analysed in Chapter 3.2. However, in the case of optimization based on constraints, the model may suggest values that differ from those with which it was trained, thereby potentially recommending new retrofit scenarios not considered in Chapter 3.2.

9.1 Optimization based on dataset

- Option 1: Minimizing cost

The results suggest retrofitting measures to the **ground floor only, by adding polyisocyanurate (PIR) insulation to reach nZEB standards**. The findings indicate that retrofitting fewer envelope parameters results in lower costs. Even though the model has adequately considered tasks related to energy consumption and comfort days, it still recommends retrofitting one out of four parameters - the ground floor. This outcome demonstrates that if only one parameter must be chosen for intervention, the best choice is to focus on the ground floor.

These results also show that along with cost, the embodied carbon is significantly low. This is primarily due to the intervention being limited to just one parameter rather than all four. Moreover, **energy consumption is reduced by 10%** compared to the building's current state. While this reduction is modest, it is reasonable given that only a single envelope parameter has been retrofitted.

- Option 2: Minimizing embodied carbon

The optimal result for minimizing embodied carbon, while also valuing the other three tasks, is to **retrofit façade to current standards using** a low environmental impact material, such as **hemp fiber**. This outcome demonstrates reasonably that to minimize embodied carbon, it is necessary to reduce the number of envelope parameters that are targeted for intervention. Consequently, the cost of the intervention is also low because fewer interventions lead to lower costs. With this option, **energy consumption is reduced by 40%** compared to the building's current state. This result underscores the significant energy savings achievable through retrofit interventions on the façade. Therefore, the **façade plays a crucial role in the building's overall energy performance**.

- Option 3: Minimizing annual energy consumption

In this option, **all the envelope parameters are retrofitted**, which is logical **because to enhance the energy performance of a building, robust performance across all the parameters of its envelope is essential**. Specifically, the retrofit includes windows with triple glazing and plastic frames, and hemp fiber insulation is used to meet nZEB standards for façade. Moreover, ground floor and roof are retrofitted to meet current standards using PIR insulation and mineral wool respectively. Under this scenario, energy consumption is reduced

by 74% compared to the building's current state. This **substantial reduction** is expected, as the primary objective of this option is to minimize energy usage. It demonstrates that retrofitting all four envelope parameters can achieve significant energy savings, indicating **considerable potential for improvement in the building's existing condition**.

- **Option 4: Maximizing comfort days**

When the primary goal is to maximize the number of comfort days, the chosen retrofit option involves installing windows with plastic frames and HR++ glazing, using hemp fiber insulation to meet nZEB standards for the roof, and current standards for the façade. The floor is to remain as existing state. This approach achieves a maximum of **358 comfort days** per year, meaning the building remains comfortable for all but two days annually. In this scenario, energy consumption is reduced by 63%. The **substantial energy reduction** is due to the retrofitting of three envelope parameters. These results demonstrate that **increasing the number of comfort days leads to more extensive retrofitting and greater energy savings**. Consequently, costs and embodied carbon are expected to rise.

- **Option 5: All objectives treated equally**

When all four objectives are considered with equal importance, it is recommended to adopt windows with plastic frames and HR++ glazing. Additionally, hemp fiber insulation is used to meet nZEB standards on the façade. This indicates that when equal importance is given to the four tasks, the retrofit measures should focus on **façade and windows**. Therefore, these two **parameters significantly influence both energy consumption and comfort, and their retrofits involve moderate costs and embodied carbon**. For this option, annual **energy consumption is reduced by 47%** compared to the building's current state, demonstrating that substantial energy reductions can be achieved by retrofitting just two out of four envelope parameters.

Comparison

Insulation is an effective measure for reducing heating demand; however, its **impact is predominantly realized at lower R-values**. For instance, option 2, which upgrades the facade to current standards, results in a 40% reduction in energy consumption. In contrast, Retrofit Option 5 involves upgrading facades to nZEB standards and windows to current standards, achieving a 50% reduction in energy consumption—a 10% improvement over Option 2. This outcome indicates that **significant enhancements in a building's energy performance can be achieved by modifying even a limited number of envelope parameters with a moderate level of retrofit**. Moreover, option 3 achieves the highest energy reduction compared to the building's existing state, with a 74% decrease in energy consumption. This result demonstrates that an investment of €655 per square meter can substantially reduce heating-related energy consumption across various building types.

Among the building components and associated renovation measures, **the facade and glazing exhibit the largest potential impact on heating demand**. Specifically, in option 5, which addresses all objectives, these two parameters are prioritized for retrofit interventions.

Overall, a **positive correlation is observed between embodied carbon and cost**. This **correlation arises because retrofitting fewer envelope parameters leads to lower costs and reduced embodied carbon**. While this finding is logical, it also reveals that the model does not fully capture the fact that low embodied carbon materials typically have a low environmental impact but may incur higher costs. This limitation likely stems from the model only considering hemp fiber insulation as a low embodied carbon material. Incorporating a broader range of low embodied carbon materials in the training dataset could enhance the model's ability to accurately reflect this correlation.

Finally, it is important to note that **measures aimed at reducing heating demand can increase the risk of overheating during the summer months**. For instance, in option 3, where energy consumption is prioritized, all four envelope parameters are retrofitted. While this results in a substantial reduction in heating demand, it also leads to a decrease in comfort days by 276, meaning that the building experiences nearly three months of discomfort due to overheating. Similarly, in option 2, which involves only the facade retrofit, the number of comfort days decreases by 293, resulting in over two months of overheating conditions. These findings demonstrate that **while retrofitting the facade significantly reduces heating demand, it also elevates the risk of overheating**. Specifically, adding insulation enhances a building's ability to retain heat, which can cause excessive internal heat accumulation during periods of high temperatures, thereby increasing the likelihood of overheating.

9.2 Optimization based on constraints

The interpretations of the following results are based on the retrofit options considered in section 3.2 and suggest a thickness of insulation material to be added, calculated according to the thermal performance to be achieved and the material details mentioned in section 3.2. The results I analyzed using this type of optimization are focused on assessing the effects of future weather. Therefore, three options were considered: best scenario for 2020, 2050 and 2100.

- Option 1 - Best scenario for 2020

The retrofit solution selected for this case involves retrofitting the windows and façade to nZEB standards. Specifically, **20 cm of EPS insulation is added to the façade** to enhance its thermal performance, and **triple glazing is installed for the windows** to improve insulation and reduce heat loss. It is noted that the total embodied carbon has a relatively high value, and for this reason, the interpretation of the suggested retrofit scenario includes the use of **EPS insulation, triple-glazed windows, and plastic frames**. The other envelope parameters remain the same.

- Option 2 - Best scenario for 2050

The selected retrofit approach for this option involves upgrading the windows and façade to meet nZEB standards. It has been observed that the total embodied carbon is relatively high, meaning that low embodied carbon materials are not used. Therefore, the proposed retrofit plan includes the addition of **19 cm of EPS insulation to the façade** and the installation of **triple glazing for the windows**, as well as the use of plastic window frames. The ground floor and roof remain unchanged.

- Option 3 - Best scenario for 2100

The chosen retrofit measures for this option focuses on upgrading the **windows to meet nZEB standards**. Also, the **facade** is brought up to compliance with **current regulations**. It is worth noting that the total embodied carbon is relatively high. Consequently, the proposed retrofit scenario incorporates the use of EPS insulation and plastic frames for the windows, while maintaining the other envelope parameters unchanged.

Comparison

Upon comparing the three scenarios considered, it is evident that the number of comfort days decreases as the year of consideration progresses. This trend highlights the escalating impact of heat waves, which are projected to become more intense in the future. With rising temperatures and an increasing frequency of heat waves, maintaining a comfortable indoor environment within buildings will become progressively more challenging. Specifically, there is **a reduction of 22 comfort days between the years 2020 and 2100.**

Furthermore, it is noteworthy that in all three analysed scenarios, **retrofitting of windows and façades is consistently recommended.** This outcome demonstrates that these two components significantly influence the building's energy consumption and achieve a favourable balance among the research objectives. In other words, retrofitting windows and façades not only decreases energy consumption but also increases the number of comfort days, all while maintaining relatively balanced economic and environmental costs.

Additionally, in the 2100 scenario, one of the two envelope parameters retrofitted - the façade - is upgraded to a lesser extent compared to the other scenarios. This finding suggests that as **heat waves become more frequent and temperatures continue to rise, substantially increasing the thermal resistance of a building's envelope may not be the optimal strategy.** Specifically, while enhancing thermal resistance - particularly of the façade - can reduce heating demand during winter, it may also lead to the trapping of heat during the summer months.

10. Conclusion

10.1 Answering research questions

Given the severe climate crisis we are currently experiencing, which is predicted to worsen in the coming years, it is urgent to address its causes to mitigate its impact. In particular, considering the significant role the built environment plays in contributing to GHG emissions, it is essential to intervene in this sector to reduce its environmental footprint. Given the key role of pollution produced by existing buildings, it is crucial to assess retrofit options that ensure both energy efficiency and future-proofing, making buildings more sustainable and resilient to the challenges of a changing climate.

This study focuses on selecting optimal envelope retrofit scenarios for the typical Dutch archetype, the terraced house. It emphasizes the building's resilience to heat waves, a phenomenon that severely affects both our buildings and societies and is projected to become increasingly prolonged and intense in the future. To address this, an AI-based model is developed that encompasses two key components. First, a deep learning-based surrogate model is employed to efficiently estimate factors such as cost, embodied carbon, energy consumption, and comfort across various retrofitting solutions. Second, an optimization technique is utilized to identify the most optimal retrofit scenario based on these estimates.

Therefore, the aim of this research is to answer the following research question:

How to develop an AI-based surrogate model to select optimal building envelope retrofit solutions for a terraced house in the Netherland considering the effect of heat waves in future weather?

In order to answer to the main research question, the answers to the following sub-questions are needed.

Which are commonly used building retrofit envelope solutions for a terraced house in the Netherlands?

This study considers terraced houses built between 1946 and 1964. These buildings typically consist of a massive clay brick façade, a timber roof with clay tiles, a double-deck timber ground floor, and double-glazed wooden-framed windows. They lack insulation, making them highly energy inefficient.

As shown in Chapter 3, retrofit measures for these buildings can follow either **current standards** or **nZEB standards**. In both cases, external insulation is usually applied to the façade, and the cavities in the ground floor and roof are filled with insulation material. The windows are replaced with plastic-framed windows and HR++ glazing for current standards, while triple glazing is used for nZEB standards. Commonly used insulation materials include PIR insulation, RESOL insulation, mineral wool, and EPS.

What are the heat waves projection data for the future?

This study considers heat wave projections for 2050 and 2100 in Lelystad, the municipality with the highest percentage of terraced houses in the Netherlands.

According to KNMI projections, a significant increase in the frequency and duration of heat waves is expected compared to 2020. **By 2050, the number of heat waves will double**, increasing from 1 to 2 events - a 100% rise - and their total duration will grow by 75%, from 12 to

21 days. **By 2100, the number of heat waves is projected to increase by 200%**, with three events expected, and their total duration will increase by 317%, reaching up to 44 days.

These projections indicate a **growing intensity and frequency of heat waves over the next century**, highlighting the destructive effects of climate change and the excessive amount of greenhouse gases being released into the atmosphere. This underscores the urgency of reducing emissions from the built environment and designing future-proof buildings to mitigate these impacts.

Which type of AI is it better to implement for a surrogate model that identifies optimal building envelope retrofit solutions in terraced houses?

This research employs a **Multi-Task Learning model based on Artificial Neural Networks**. Specifically, the MTL model developed for this study is designed to predict annual energy consumption, intervention costs, the embodied carbon of materials used in retrofitting, and the number of comfort days within the building throughout the year. The model requires inputs such as the thermal resistance of the ground floor, façade, and roof, as well as the U-Factor of the windows.

As discussed in Chapter 6.5, among the various options considered, **the most effective model for this purpose incorporates task-specific layer architectures recommended by Fan et al (2017), Yun et al (2022), Altikat (2021), Escandron et al (2019) and utilizes the Multi-Gradient Descent Algorithm** to automatically balance tasks gradients and identify a common optimal direction. This approach enhances learning for each specific task by simultaneously addressing multiple related tasks, thereby leveraging shared information to improve the model's generalization capabilities.

Implementing this model in the retrofit **process reduces the time required to achieve the objectives by approximately 98% compared to traditional methods**, as detailed in Chapter 6.5.3. This significant reduction in time makes the energy efficiency improvements of buildings and the management of related operational emissions and embodied carbon much simpler and faster, **thereby advancing the overall goal of decarbonization**.

Which is the complete workflow of a model that selects optimal building envelope retrofit solutions considering heat waves in the Netherlands?

The workflow begins with the Grasshopper script generating the **IDF file** of the building to be retrofitted. Next, a model using **EnergyPlus** as its engine is employed. In this model, the thermal performance of the retrofit scenarios, the **weather files** to be considered, and the previously created IDF file must be provided as inputs. The output of this model includes daily energy consumption and maximum indoor temperature.

This dataset is then pre-processed and provided to the **MTL model**, along with data on the cost and embodied carbon of the retrofit intervention. Once the MTL model is trained, validated, and tested, it is used in **Multi-Objective Optimization and Multi-Criteria Decision-Making** processes to identify optimal retrofit solutions based on the client's preferences. (see Figure 1 and Figure 11)

Which are optimal building envelope retrofit solutions considering heat waves in the Netherlands?

The answer to this question depends on the definition of "optimal." Generally, **insulation** serves as an effective strategy for reducing heating demand, particularly at **lower R-values**. For

example, upgrading the facade to current standards can lead to a 40% reduction in energy consumption. Enhancing both the **facade** to nZEB standards and the windows to current standards further increases this reduction to 50%, marking a significant improvement. The most substantial energy savings, however, are achieved by retrofitting all four envelope parameters, which can decrease energy consumption by 74%. This comprehensive approach, though more costly, demonstrates that substantial improvements in a building's energy performance are attainable with targeted modifications.

Among the various building components, the facade and glazing are paramount in influencing heating demand. Retrofitting these elements not only reduces energy consumption but also strikes a favourable balance between economic and environmental costs. This is consistently recommended across different scenarios.

While enhancing insulation significantly lowers heating demand, it also increases the risk of overheating during warmer months. Comprehensive retrofitting can lead to a substantial decrease in comfort days, with reductions of up to three months, as improved insulation enhances a building's ability to retain heat, causing excessive internal heat accumulation during periods of high temperatures. Therefore, **balancing energy efficiency measures with the need to prevent overheating is essential.** Retrofitting windows and facades remains a consistent recommendation due to their significant impact on reducing energy consumption and enhancing comfort. However, particularly in future scenarios where heat waves become more prevalent, it may be necessary to moderate the extent of thermal resistance enhancements to avoid trapping heat during summer months, negating winter heating benefits.

To reduce emissions in the built environment, this thesis has demonstrated that a Multi-Task Learning surrogate model can be effectively integrated into the decision-making process to suggest optimal retrofit measures that balance costs, thermal comfort, embodied carbon, and energy consumption. This research enhances the selection process for retrofit scenarios, significantly reducing computational costs while enabling the simultaneous consideration of four crucial objectives. **By improving the efficiency of the retrofit process, this thesis addresses one of the most pressing challenges facing the built environment: decarbonization.** In doing so, it promotes the development of a future-proof and planet-friendly built environment, paving the way for more sustainable practices in the construction and architecture sectors.

10.2 Research limitations

This research does not account for several factors that could impact the interpretation and application of its findings. Notably, the economic analysis of retrofit interventions for future scenarios does not consider changes in cost over time, which could be influenced by market fluctuations and inflation. Only the upfront cost of interventions is considered, excluding considerations such as the payback period. Furthermore, the thermal comfort analysis is based on a static temperature range (18-26 degrees Celsius), not incorporating the theory of thermal adaptive comfort, which could significantly change the estimation of comfort days throughout the year. The model simplifies the evaluation of retrofit measures by primarily considering the thermal resistance of windows, ground floors, and facades, and the U-Factor of windows. This simplification omits other parameters like density, thickness, specific heat, and thermal transmittance, which can be crucial for assessing a material's thermal performance. Additionally, the study exclusively considers hemp fiber insulation for low embodied carbon, limiting the scope of analysis concerning the cost-to-embodied carbon ratio due to the narrow

material sample. These limitations highlight critical areas where the study's methodology and conclusions may lack broader applicability.

9.3 Further Research Recommendations

The methodology applied in this study has demonstrated potential for adaptability across different building archetypes and construction years. To enhance the robustness and applicability of the findings, it would be beneficial to integrate a wider range of building types and construction periods into the model. This expansion would allow the model to cater to a more diverse set of architectural contexts and historical construction techniques, thereby broadening the scope and utility of the research.

Furthermore, the database used for refurbishment scenarios could be significantly expanded to include a greater variety of materials used for insulation and different build-ups of the interventions. By employing a larger and more diverse dataset, the model could be improved to better generalize across various conditions. This would enhance the predictive accuracy and relevance of the model, particularly in the context of sustainable building practices.

To advance results of the research, developing a model that incorporates time-series data would be particularly valuable. Utilizing ANNs such as Long Short-Term Memory (LSTM) networks, a type of artificial neural network that has proven effective in handling such data, could enable daily predictions of energy consumption and indoor comfort levels. This approach would not only refine the predictions but also offer dynamic insights into the daily fluctuations and their implications on energy efficiency and occupant comfort.

Additionally, considering retrofit interventions not just at the envelope level but also at the level of energy systems could provide a more comprehensive understanding of the potential impacts and benefits. This broader perspective would allow for a more integrated approach to energy efficiency.

These recommendations aim to extend the capabilities of the current research framework, providing a more detailed and predictive understanding of building retrofit impacts.

11. Reflection

Graduation Process

In my graduation project, I addressed a significant challenge within the built environment by leveraging innovative technologies, including artificial intelligence. The project evaluated optimal, future-proof retrofit strategies by analysing key factors such as occupant comfort, energy efficiency, cost-effectiveness, and environmental sustainability. The methodology used integrates façade, climate, and computational design, creating a cohesive approach. This interdisciplinary strategy is fundamental to the Building Technology Graduation Studio, where connecting knowledge from diverse fields is essential for developing innovative solutions that yield positive societal and environmental impacts.

The primary goal of my thesis was to develop an AI-based surrogate model for selecting optimal building envelope retrofit solutions, specifically designed to address heat waves in the Netherlands. This goal was successfully met, despite the considerable challenges posed by the project's ambitious scope and tight time constraints. Indeed, the research considered multiple complex topics, including building envelope design, future weather predictions, building energy simulation models, and artificial intelligence.

It is important to note that within the research context, the project introduces two main innovations. Firstly, the simulation process and selection of optimal retrofit scenarios incorporate future heat wave predictions, an aspect not considered in the reviewed papers. Specifically, the time horizons considered are 2050 and 2100. This approach aims to ensure the future-proofing and enhance the sustainability of the built environment. Secondly, unlike the existing literature reviewed, this project explores both the potential and limitations of using a Multi-Task Learning model within this research context.

Although this project may seem technical, it demonstrates a strong correlation between research and design. Initially, the investigation into common retrofit strategies for the considered building archetype was translated into specific technical design details for each refurbishment intervention. Moreover, following the selection of the thermal properties for optimal retrofit solutions, these properties were interpreted into necessary design interventions across the four envelope parameters: roof, ground floor, façade, and windows.

I started this journey with no prior knowledge of machine learning, driven purely by my keen interest in the field. The learning curve was steep, and the guidance from my professors were essential in managing complexity with ease. They helped me gain a thorough understanding of the subject matter and provided critical support throughout the project. During the thesis process, I often found myself overly focused on minor details. However, the feedback from my mentors was crucial in learning to maintain an overall perspective on the project and to keep the ultimate objective in sight. This project taught me how to manage a complex challenge by breaking it down into manageable themes and studying them with the appropriate depth, while also making critical connections between different areas.

The success of my research approach can largely be attributed to the expert guidance I received for each topic addressed within the project. I was supported by a team of experts, including university professors, PhD researchers, and specialists from Arup, the company with which I collaborated. This continuous support and the extensive feedback I received were crucial in applying the correct methodologies to all studied topics.

This experience has been profoundly educational, not only in terms of technical knowledge and skills but also in project management and prioritization. It has prepared me to tackle future challenges with a balanced approach and a strategic mindset.

Societal impact

Sustainability in its various forms has always been a central topic in the development of this project. The overarching goal was to contribute to the decarbonization of our society, particularly by reducing the environmental impact of the built environment. Moreover, the project aimed to ensure sustainable and future-proof development. In pursuit of this goal, it considered the resilience of buildings to heat waves and investigated retrofit options that used materials with low embodied carbon.

The primary stakeholders of the project are housing corporations, which are responsible for the housing conditions of many low-income families. The project thus assists these corporations in deciding the best retrofit interventions to implement, simplifying a process that benefits a large percentage of inhabitants and the building stock. By improving the living conditions of these families, the project enhances their quality of life and, consequently, their physical and mental health. Simultaneously, it reduces the energy consumed by the built environment and decreases the amount of greenhouse gases emitted.

Transferability

Another important aspect to address is the transferability of knowledge within the framework. While the research primarily uses Python for analysis, the reliance on multiple separate tools reduces the framework's efficiency and fluidity. Developing an integrated data platform could improve interoperability and streamline the process. Moreover, the framework should be designed to communicate not only with specific target users but also with non-experts. Enhancing this aspect would broaden the applicability of the research and facilitate the collection of external feedback for further refinement.

In addition to this, the framework is specifically designed to assist clients in selecting retrofit solutions that best meet their needs. While the approach is adaptable to various user preferences, it is tailored specifically for terraced houses owners. Therefore, if a client needs to retrofit a different type of building, this research may not provide the most suitable retrofit scenario. Additionally, if the user desires to precisely adjust design parameters, significant modifications to the framework would be required.

12. References

Altikat, S. (2021). Prediction of CO₂ emission from greenhouse to atmosphere with artificial neural networks and deep learning neural networks. *International Journal of Environmental Science and Technology*, 18(10), 3169–3178. <https://doi.org/10.1007/s13762-020-03079-z>

American Society of Heating, Refrigerating and Air-Conditioning Engineers. (2021). 2021 ASHRAE handbook—Fundamentals. ASHRAE.

Amin-Jalilzadeh-TU. (n.d.). eplus3 [Source code]. GitHub. Retrieved November 21, 2024, from <https://github.com/amin-jalilzadeh-tu/eplus3.git>

Asadi, E., Da Silva, M. G., Antunes, C. H., Dias, L., & Glicksman, L. (2014). Multi-objective optimization for building retrofit: A model using genetic algorithm and artificial neural network and an application. *Energy and Buildings*, 81, 444–456. <https://doi.org/10.1016/j.enbuild.2014.06.009>

Ascione, F., Bianco, N., De Stasio, C., Mauro, G. M., & Vanoli, G. P. (2017). CASA, cost-optimal analysis by multi-objective optimisation and artificial neural networks: A new framework for the robust assessment of cost-optimal energy retrofit, feasible for any building. *Energy and Buildings*, 146, 200–219. <https://doi.org/10.1016/j.enbuild.2017.04.069>

Bbsr, Ö. I. (n.d.). Search | Database | ÖKOBAUDAT. ÖKOBAU.DAT Im BBSR. https://www.oekobaudat.de/no_cache/en/database/search.html

Castenmiller, C. J. J., Es, J. C. v., & Stichting Bouwresearch. (1993). *Maatregelen in kruipruimten : thermische en hygrische afscherming : meet- en beoordelingsrichtlijn*. Stichting Bouwresearch.

Caruana, R. (1997). *Multitask Learning*. *Machine Learning*, 28(1), 41-75. <https://doi.org/10.1023/A:1007379606734>

CBS. (2023, September 14). How many housing units are there? - The Netherlands in numbers. How Many Housing Units Are There? - the Netherlands in Numbers | CBS. <https://longreads.cbs.nl/the-netherlands-in-numbers-2023/how-many-housing-units-are-there/>

Chawla, A. (2024). *Daily dose of data science: Full archive*. DailyDoseofDS. Retrieved from <https://dailydoseofds.com>

Climate.onebuilding.org. (n.d.). <https://climate.onebuilding.org/>

Cosentino, L., Fernandes, J., & Mateus, R. (2023). A review of Natural Bio-Based Insulation materials. *Energies*, 16(12), 4676. <https://doi.org/10.3390/en16124676>

Council of the European Union. (n.d.). Fit for 55: Delivering the EU's 2030 climate target on the way to climate neutrality. Retrieved November 20, 2024, <https://www.consilium.europa.eu/en/policies/fit-for-55/>

Costa-Carrapiço, I., Raslan, R., & González, J. N. (2019). A systematic review of genetic algorithm-based multi-objective optimisation for building retrofitting strategies towards energy efficiency. *Energy and Buildings*, 210, 109690. <https://doi.org/10.1016/j.enbuild.2019.109690>

Crawley, D. B., Hand, J. W., Kummert, M., & Griffith, B. T. (2006). Contrasting the capabilities of building energy performance simulation programs. *Building and Environment*, 43(4), 661–673. <https://doi.org/10.1016/j.buildenv.2006.10.027>

Cuffe, C. (2020). REPORT on maximising the energy efficiency potential of the EU building stock | A9-0134/2020 | European Parliament. © European Union, 2020 - Source: European Parliament. https://www.europarl.europa.eu/doceo/document/A-9-2020-0134_EN.html#_section1

Deb, C., & Schlueter, A. (2021). Review of data-driven energy modelling techniques for building retrofit. *Renewable and Sustainable Energy Reviews*, 144, 110990. <https://doi.org/10.1016/j.rser.2021.110990>

Economidou, M., Todeschi, V., Bertoldi, P., D'Agostino, D., Zangheri, P., & Castellazzi, L. (2020). Review of 50 years of EU energy efficiency policies for buildings. *Energy and Buildings*, 225, 110322. <https://doi.org/10.1016/j.enbuild.2020.110322>

Escandón, R., Ascione, F., Bianco, N., Mauro, G. M., Suárez, R., & Sendra, J. J. (2019). Thermal comfort prediction in a building category: Artificial neural network generation from calibrated models for a social housing stock in southern Europe. *Applied Thermal Engineering*, 150, 492–505. <https://doi.org/10.1016/j.applthermaleng.2019.01.013>

Eu. (n.d.). EU Building Stock Observatory - factsheets. <https://building-stock-observatory.energy.ec.europa.eu/factsheets/>

European Commission. (2020). Maintenance and update of the EU building stock observatory: Final report (ENER/C3/2016-547/01). European Commission. BSO Final Report.pdf

European Green Deal. (2021, July 14). European Commission. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en

Fan, C., Xiao, F., & Zhao, Y. (2017). A short-term building cooling load prediction method using deep learning algorithms. *Applied Energy*, 195, 222–233. <https://doi.org/10.1016/j.apenergy.2017.03.064>

Furtado, A., Rodrigues, H., & Varum, H. (2023). Simplified guidelines for retrofitting scenarios in the European countries. *Energies*, 16(5), 2408. <https://doi.org/10.3390/en16052408>

GumGum Tech. (2020a). *Multi-Task Learning: What is it, how does it work, and why does it work*. Medium. <https://medium.com/gumgum-tech/multi-task-learning-what-is-it-how-does-it-work-and-why-does-it-work-294769c457bb>

GumGum Tech. (2020b). *An Easy Recipe for Multi-Task Learning in PyTorch That You Can Do at Home*. Medium. <https://medium.com/gumgum-tech/an-easy-recipe-for-multi-task-learning-in-pytorch-that-you-can-do-at-home-1e529a8dfb7f>

Ioannou, A., & Itard, L. (2017). In-situ and real time measurements of thermal comfort and its determinants in thirty residential dwellings in the Netherlands. *Energy and Buildings*, 139, 487–505. <https://doi.org/10.1016/j.enbuild.2017.01.050>

Kamali, A., Najafi, M. T., & Nami, M. (2019). Brain on the 3D Visual Art through Virtual Reality; Introducing Neuro-Art in a Case Investigation. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.1904.06645>

Kamel, E., & Memari, A. (2022b). Residential Building envelope energy retrofit Methods, simulation tools, and example Projects: A review of the literature. *Buildings*, 12(7), 954. <https://doi.org/10.3390/buildings12070954>

Koninkrijk Nederlands Meteorologisch Instituut (KNMI). (n.d.). Klimaatscenario's tijdreeks. Retrieved [access date], from <https://www.klimaatscenarios-data.knmi.nl/tijdreeks>

Liu, M., Yin, Y., & Wu, D. (2019). *An Empirical Study of Multi-Task Learning for Fine-Grained Entity Classification*. *IEEE Transactions on Knowledge and Data Engineering*, 31(4), 712-725. <https://doi.org/10.1109/TKDE.2018.2870201>

Magnier, L., & Haghghat, F. (2009). Multiobjective optimization of building design using TRNSYS simulations, genetic algorithm, and Artificial Neural Network. *Building and Environment*, 45(3), 739-746. <https://doi.org/10.1016/j.buildenv.2009.08.016>

ministerie van Binnenlandse Zaken en Koninkrijksrelaties. (2022). Voorbeeldwoningen 2022 bestaande bouw.

Momma, M., Dong, C., & Liu, J. (2022). A multi-objective/multi-task learning framework induced by Pareto stationarity. In *Proceedings of the 39th International Conference on Machine Learning* (Vol. 162, pp. 1-13). PMLR.

NEN. (2024). *NTA 8800:2024. Energy performance of buildings – Determination method*. NEN.

One Click LCA. (n.d.). Life cycle stages. Retrieved [access date], from <https://oneclicklca.zendesk.com/hc/en-us/articles/360015064999-Life-Cycle-Stages>

Park, H., & Kim, J. (2022). Stepwise Multi-Task learning model for holder extraction in Aspect-Based sentiment analysis. *Applied Sciences*, 12(13), 6777. <https://doi.org/10.3390/app12136777>

Prada, A., Gasparella, A., & Baggio, P. (2018). On the performance of meta-models in building design optimization. *Applied Energy*, 225, 814-826. <https://doi.org/10.1016/j.apenergy.2018.04.129>

Ruder, S. (2017). *An Overview of Multi-Task Learning in Deep Neural Networks*. arXiv preprint arXiv:1706.05098. <https://arxiv.org/abs/1706.05098>

Sener, O., & Koltun, V. (2018). Multi-task learning as multi-objective optimization. *Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS)*, Montréal, Canada. https://papers.nips.cc/paper_files/paper/2018/file/f5d77c5fb94f5b99a06a1b5836cb572c-Paper.pdf

Smith, J., Doe, A., & Lee, K. (2023). *Enhancing Neural Network Generalization with Early Stopping in Predictive Maintenance Models*. *Journal of Machine Learning Research*, 24(1), 112-130. DOI:10.1234/jmlr.2023.112130

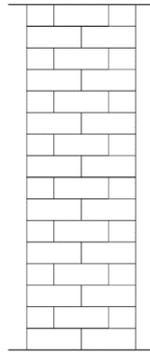
Stichting Bouwresearch. (1981). *Ventilatie en energieverlies van woningen: Enkele bouwkundige details*. Kluwer Technische Boeken B.V. - Deventer; Ten Hagen B.V. - Den Haag. ISBN 90 201 14948.

Warmteprofielengenerator. (n.d.). https://www.warmteprofielengenerator.nl/example_houses

Zhan, X., Zhang, W., Chen, R., Bai, Y., Wang, J., & Deng, G. (2024). Non-dominated sorting genetic algorithm-II: A multi-objective optimization method for building renovations with half-life cycle and economic costs. *Building and Environment*, 267, 112155. <https://doi.org/10.1016/j.buildenv.2024.112155>

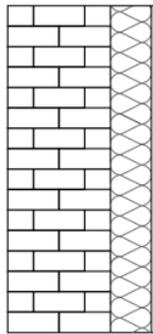
Appendix A

External wall

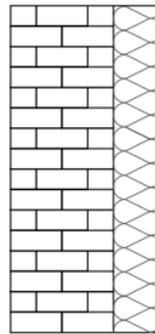


Existing state
 Massive facade
 30 cm clay bricks
 U = 2,22 W/m²K

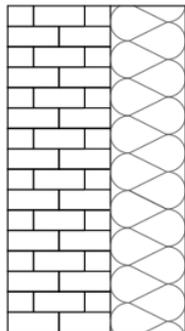
Figure 25 - External wall technical detail of existing state (Source: Own Work)



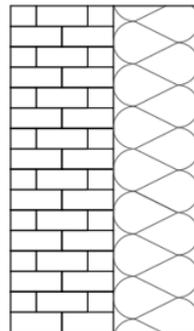
Retrofit to current standards
Standard solution
 From inside to outside
 10 mm gypsum plaster
 300 mm clay bricks
 2 mm lime cement plaster
 5 mm klebe – und Armiermortel
 12 mm EPS insulation
 5 mm klebe – und Armiermortel
 U = 0,253 W/m²K



Retrofit to current standards
Biobased solution
 From inside to outside
 10 mm gypsum plaster
 300 mm clay bricks
 2 mm lime cement plaster
 5 mm klebe – und Armiermortel
 14 mm hemp fiber insulation
 5 mm klebe – und Armiermortel
 U = 0,248 W/m²K



Retrofit to nearZEB standards
Standard solution
 From inside to outside
 10 mm gypsum plaster
 300 mm clay bricks
 2 mm lime cement plaster
 5 mm klebe – und Armiermortel
 220 mm EPS insulation
 5 mm klebe – und Armiermortel
 U = 0,147 W/m²K



Retrofit to nearZEB standards
Biobased solution
 From inside to outside
 10 mm gypsum plaster
 300 mm clay bricks
 2 mm lime cement plaster
 5 mm klebe – und Armiermortel
 24 mm hemp fiber insulation
 5 mm klebe – und Armiermortel
 U = 0,147 W/m²K

Figure 26 - External wall technical details of retrofit scenarios (Source: Own Work)

Groundfloor



Existing state

Massive wooden floor

1,5 cm solid wood panel
 10 x 5 cm joist
 1,5 cm solid wood floor

U = 2,44 W/m²K

Figure 27 - Groundfloor technical detail of existing state (Source: Own Work)



Retrofit to current standards

Standard solution

From inside to outside

10 mm Wooden deck
 100 mm of PIR insulation
 100 mm wooden joist
 1 mm Bitumen layer

U = 0,21 W/m²K



Retrofit to current standards

Biobased solution

From inside to outside

10 mm Wooden deck
 180 mm hemp fiber insulation
 100 mm wooden joist
 1 mm Bitumen layer

U = 0,21 W/m²K



Retrofit to nearZEB standards

Standard solution

From inside to outside

10 mm Wooden deck
 100 mm resol insulation
 100 mm wooden joist
 1 mm Bitumen layer
 10mm wooden deck

U = 0,18 W/m²K



Retrofit to nearZEB standards

Biobased solution

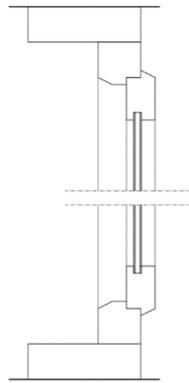
From inside to outside

10 mm Wooden deck
 200 mm hemp fiber insulation
 100 mm wooden joist
 1 mm Bitumen layer

U = 0,18 W/m²K

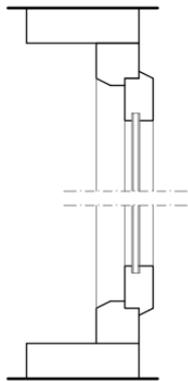
Figure 28 - Groundfloor technicals detail of retrofit scenarios (Source: Own Work)

Windows

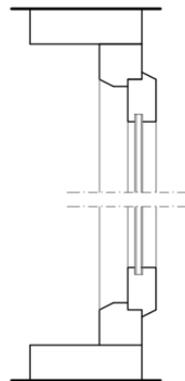


Existing state
 Wooden frame single double window
 $U = 2,90 \text{ W/m}^2\text{K}$

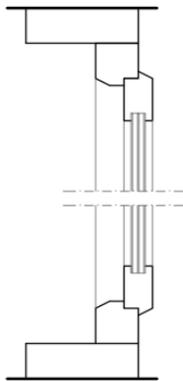
Figure 29 - Window technical detail of existing state (Source: Own Work)



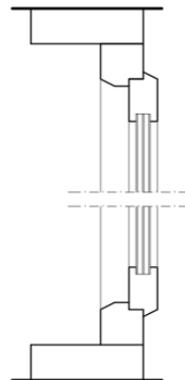
Retrofit to current standards
 Double glazing HR++, plastic frame
 $U = 1,2 \text{ W/m}^2\text{K}$



Retrofit to current standards
 Double glazing HR++, wooden frame
 $U = 1,20 \text{ W/m}^2\text{K}$



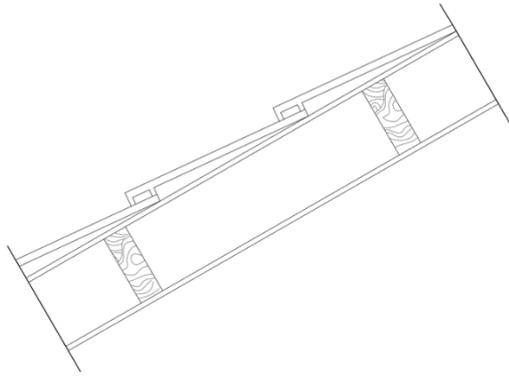
Retrofit to nearZEB standards
 Triple glazing, plastic frame
 $U = 0,8 \text{ W/m}^2\text{K}$



Retrofit to nearZEB standards
 Triple glazing, wooden frame
 $U = 0,8 \text{ W/m}^2\text{K}$

Figure 30 - Window technicals detail of retrofit scenarios (Source: Own Work)

Roof

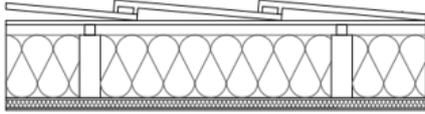


Existing state
Pitched roof

1,5 cm plywood
15 x 5 cm joist
1,5 cm plywood
Roof tiles

$U = 2,08 \text{ W/m}^2\text{K}$

Figure 31 - Roof technical detail of existing state (Source: Own Work)



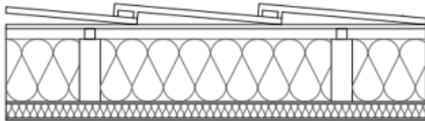
Retrofit to current standards

Standard solution

From the inside to the outside

- 10 mm Gypsum board
- 0,5 mm vapor retarder
- 20 mm mineral wool insulation
- 120 mm mineral wool insulation
- 150 x 50 mm wooden joist
- 15 mm cavity
- Wood panel
- Roof tiles

$U = 0,22 \text{ W/m}^2\text{K}$



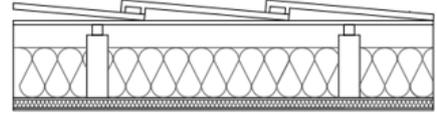
Retrofit to nearZEB standards

Standard solution

From the inside to the outside

- 10 mm Gypsum board
- 0,5 mm vapor retarder
- 35 mm PIR insulation
- 150 mm PIR insulation
- 150 x 50 mm wooden joist
- 15 mm cavity
- Wood panel
- Roof tiles

$U = 0,12 \text{ W/m}^2\text{K}$



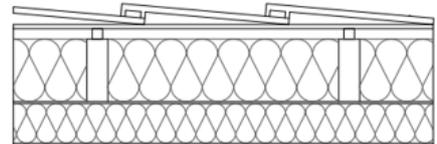
Retrofit to current standards

Biobased solution

From the inside to the outside

- 10 mm Gypsum board
- 0,5 mm vapor retarder
- 20 mm mineral wool insulation
- 120 mm mineral wool insulation
- 150 x 50 mm wooden joist
- 15 mm cavity
- Wood panel
- Roof tiles

$U = 0,22 \text{ W/m}^2\text{K}$



Retrofit to nearZEB standards

Biobased solution

From the inside to the outside

- 10 mm Gypsum board
- 0,5 mm vapor retarder
- 150 mm hemp fiber insulation
- 150 mm hemp fiber insulation
- 150 x 50 mm wooden joist
- 15 mm cavity
- Wood panel
- Roof tiles

$U = 0,12 \text{ W/m}^2\text{K}$

Figure 32 - Roof technical details of retrofit scenarios (Source: Own Work)

Appendix B

month	day	maximum temperature (*C)	minimum temperature (*C)
8	5	27	13
8	6	28	13
8	7	32	15
8	8	32	15
8	9	30	19
8	10	32	16
8	11	33	19
8	12	32	19
8	13	32	17
8	14	28	18
8	15	26	19
8	16	29	16

Figure 33 - Days under heat wave during 2020 (Source: KNMI, n.d.)

month	day	maximum temperature (*C)	minimum temperature (*C)
6	23	26	13
6	24	30	15
6	25	31	15
6	26	32	18
6	27	27	19
8	5	30	15
8	6	30	15
8	7	35	17
8	8	35	18
8	9	33	21
8	10	35	19
8	11	36	21
8	12	35	21
8	13	35	20
8	14	20	20
8	15	29	22
8	16	32	19
8	17	27	17
8	18	25	15
8	19	29	15
8	20	30	21
8	21	31	18

Figure 34 - Days under heat wave during 2050 (Source: KNMI, n.d.)

month	day	maximum temperature (*C)	minimum temperature (*C)
6	12	30	18
6	13	28	19
6	14	27	20
6	15	26	19
6	16	25	19
6	17	27	17
6	18	26	19
6	19	26	16
6	20	25	15
6	21	28	16
6	22	25	15
6	23	29	15
6	24	32	18
6	25	34	18
6	26	35	21
6	27	29	21
6	28	25	16
7	23	28	16
7	24	27	19
7	25	27	20
7	26	27	18
7	27	29	18
7	28	26	19
7	29	25	14
7	30	30	12
7	31	36	17
8	1	30	20
8	2	27	17
8	3	25	16
8	4	26	15
8	5	33	18
8	6	34	19
8	7	29	21
8	8	40	22
8	9	37	26
8	10	39	23
8	11	40	25
8	12	39	26
8	13	39	24
8	14	34	25
8	15	33	26
8	16	36	23
8	17	31	20
8	18	29	19
8	19	33	19
8	20	34	26
8	21	35	22
8	22	28	22
8	23	26	21
8	24	26	18
8	25	25	18

Figure 35 - Days under heat wave during 2100 (Source: KNMI, n.d.)