A Digital Twin framework for multi-objective optimization

Application of predicative machine learning and genetic algorithms for cost and energy performance optimization

TRON ERLEND TRONSEN

SUPERVISOR
Haidar Hosamo

Universitetet i Agder, 2023
Fakultet for Teknologi og Realfag
Institutt for Ingeniørvitenskap

Obligatorisk egenerklæring

Den enkelte student er selv ansvarlig for å sette seg inn i hva som er lovlige hjelpemidler, retningslinjer for bruk av disse og regler om kildebruk. Erklæringen skal bevisstgjøre studentene på deres ansvar og hvilke konsekvenser fusk kan medføre. Manglende erklæring fritar ikke studentene fra sitt ansvar.

1.	Jeg erklærer herved at min besvarelse er mitt eget arbeid, og at jeg ikke har brukt andre kilder eller har mottatt annen hjelp enn det som er nevnt i besvarelsen.	☒
2.	Jeg erklærer videre at denne besvarelsen: - ikke har vært brukt til annen eksamen ved annen avdeling/universitet/høgskole innenlands eller utenlands. - ikke refererer til andres arbeid uten at det er oppgitt. - ikke refererer til eget tidligere arbeid uten at det er oppgitt. - har alle referansene oppgitt i litteraturlisten. - ikke er en kopi, duplikat eller avskrift av andres arbeid eller besvarelse.	☒
3.	Jeg er kjent med at brudd på ovennevnte er å betrakte som fusk og kan medføre annullering av eksamen og utestengelse fra universiteter og høgskoler i Norge, jf. Universitets- og høgskoleloven §§4-7 og 4-8 og Forskrift om eksamen §§ 31.	☒
4.	Jeg er kjent med at alle innleverte oppgaver kan bli plagiatkontrollert.	☒
5.	Jeg er kjent med at Universitetet i Agder vil behandle alle saker hvor det forligger mistanke om fusk etter høgskolens retningslinjer for behandling av saker om fusk.	☒
6.	Jeg har satt meg inn i regler og retningslinjer i bruk av kilder og referanser på biblioteket sine nettsider.	☒

Publiseringsavtale

Fullmakt til elektronisk publisering av oppgaven

Forfatter har opphavsrett til oppgaven. Det betyr blant annet enerett til å gjøre verket tilgjengelig for allmennheten (Åndsverkloven. §2).
Alle oppgaver som fyller kriteriene vil bli registrert og publisert i Brage Aura og på UiA sine nettsider med forfatter(ne)s godkjennelse.

Oppgaver som er unntatt offentlighet eller tausehetsbelagt/konfidensiell vil ikke bli publisert.

Jeg gir herved Universitetet i Agder en vederlagsfri rett til å gjøre oppgaven tilgjengelig for elektronisk publisering:	☒JA	☐NEI
Er oppgaven båndlagt (konfidensiell)? (Båndleggingsavtale må fylles ut)	☐JA	☒NEI
- Hvis ja: Kan oppgaven publiseres når båndleggingsperioden er over?	☒JA	☐NEI
Er oppgaven unntatt offentlighet? (inneholder taushetsbelagt informasjon. Jfr. Offl. §13/Fvl. §13)	☐JA	☒NEI

Preface

This master's thesis is created at the engineering and science faculty as part of the civil and structural engineering masters course at the University of Agder (UiA). I would like to take the opportunity to thank my supervisor Haidar Hosamo for his support and guidance in the creation of this thesis.

On a slightly more personal note, I would thank my friends and family. Especially my fellow "lab-rats" for their company and support throughout all the long hours spent at the civil engineering lab. I would also like to take the opportunity to thank my former supervisor during my BSc degree and specialisation project, Paul Ragnar Svennevig, whose support and advice have been indispensable for me during my time at the University of Agder.

This thesis aims to develop a framework implementing machine learning, genetic algorithms and digital twin. The goal is to use predictive machine learning models, perform a multi-objective optimization in terms of costs and energy consumption and create an easily visually interpretable model for accurate and efficient decision-making.

Summary

This thesis represents the culmination of the Msc civil engineering course at the University of Agder. This thesis aims to attempt to define a framework for implementing digital twins in an investment cost/energy consumption optimization process. The methodology applied is a complex software hierarchy. The original dataset rests on randomly generated values of thermal transmittance, which are analysed in IDA ICE simulations, and compared to existing materials identified in the Norsk Prisbok for cost estimation. The results are optimized using a combination of Artificial Neural Networks and a multi-objective optimization algorithm, the elitist non-dominated sorting algorithm NSGA-II.

The research question this thesis attempts to answer is:

How can digital twins be implemented to reduce energy-consumption and costs in buildings?

This thesis concludes that "A digital twin may be implemented to translate energy consumption and cost-optimization into an easily interpreted result that serves as a foundation for efficient decision-making." This conclusion is based on the functionality of the various steps in the framework: Accuracy of ANN models, NSGA-II performance and visual presentation.

The thesis presents a functional framework with a high degree of automation. Furthermore, applying said framework to a case study identified a potential energy consumption reduction of 35 % and a reduction in investment costs by 5 %.

Table of contents

List of Figures

List of Tables

List of Equations

1 Introduction

We are currently experiencing the fourth industrial revolution, the digitalization era, where digital tools have become an increasingly integrated part in all aspects of our everyday life [1]. In the Architecture, Engineering and Construction (AEC) industry, this is referred to as Construction 4.0 [2]. Reflective of this era is the adaption and development of new technology. However, the AEC industry is notoriously sluggish at adapting to new technologies [3].

Despite this, the later decade has seen an increased adaption of emerging technology. The construction industry appears to be embracing the benefits that these technologies may provide [4]. Examples may include Digital Twins (DT) for efficient decision-making and building performance optimization [2], and Artificial Intelligence (AI) for its ability to analyse and process large quantities of data efficiently and allow for accurate predictions and optimization purposes [5].

With the emergence of artificial intelligence, so has the application and development of methodologies such as genetic algorithms [5]. These algorithms are well known for identifying optimal design parameters and minimising trade-offs between seemingly contradictory objectives [6]. Such objectives may be contextualised with the increased focus on resource efficiency and green technologies in the AEC. Given that financial barriers may serve as a deterring source for sustainable building design [7].

One may argue that reducing investment costs may serve as an incentive for future investment. Reducing these costs will likely increase sustainability in the construction and operational phases. One may further argue that as the level of digital tools increases, so would the knowledge and expertise required to operate them. In conjunction with new fields of interest in the AEC, it seems safe to presume that future digital construction models will become more complex. Artificial intelligence may help automate processes and outsource human interaction. In such a context, one may consider artificial intelligence a tool allowing practitioners to catch up to the development and allow for efficient decision-making by non-experts.

This thesis proposes a cost and energy consumption optimization framework using machine learning regression predictive models combined with multi-objective optimization genetic algorithms. The goal is to optimize the tradeoff between investment cost and energy consumption.

The chosen case is a classroom at Tvedestrand Videregående Skole, which undergoes building performance simulations in terms of energy consumption based on various values of thermal transmittance. The same design parameters are used to perform a cost analysis. Both the simulation and cost analysis is based on the same decision variables.

The goal is to create a framework which implements digital twin and visualization in a BIM environment based on multi-objective optimisation. A visual model may help stakeholders identify influential parameters intuitively, not depending on numbers or code. It is worth stressing that such a framework is not limited to the chosen application area. It may be expanded to cover other areas where building performance may be measured by two objectives which share decision variables.

2 Societal Impacts

It is common knowledge that we are experiencing climate change, and its consequences are evident. The Paris Agreement in 2015 formulated a legally binding international treaty and entered into force in November 2016. The goal is to limit global warming. In order to reach the goal of limiting global warming to 1,5 degrees compared to the pre-industrial era, greenhouse gas emissions must peak by 2025 and be reduced by 43% by 2030 [8]. We are currently at a critical stage to reach this goal.

In the context of AEC, it has been found that the building sector is responsible for over 40% of energy consumption and 32% of resource consumption worldwide [9]. Additionally, heating needs account for 40-60% of energy use in cold climates such as Norway [10]. Considering that the energy production sector is currently one of the most environmentally damaging sectors in the world [11]. It becomes apparent that energy performance optimization will become increasingly important to keep shifting the AEC in a sustainable direction.

Although energy efficiency is essential in sustainable building design, the concept of costs cannot be ignored [12]. Achieving a high degree of cost-efficiency in buildings whilst maintaining a high degree of energy efficiency will likely be an essential incentive for achieving sustainable and performance-based building design. In such a context, it is natural to consider multi-objective optimization, reducing costs and energy consumption based on the same decision variables. The use of new technologies, such as artificial intelligence, may be a means towards this end.

From a historical perspective, the degree of energy efficiency has increased in the later years. It has improved annually by approximately 1.9 %. However, to reach the energy efficiency target, this amount needs to average 3.2 % [13]. This is addressed in the UN sustainability goal 7·3, "*Double the improvement in energy efficiency*". In this context, it is also natural to consider the value and benefits of sharing technology advancement with others. Promoting access to energy-efficient and cost-effective design research will likely advance the benefits and results on a broader scale. This is directly linked to the UN sustainability goal 7·A. "Promote access to research, technology and investments in clean energy".

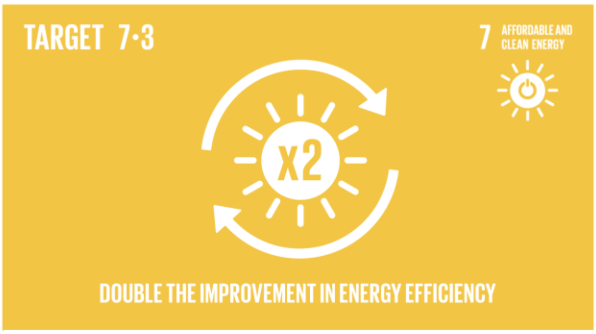


Figure 2.2 UN Sustainability Target 7·3 [14].

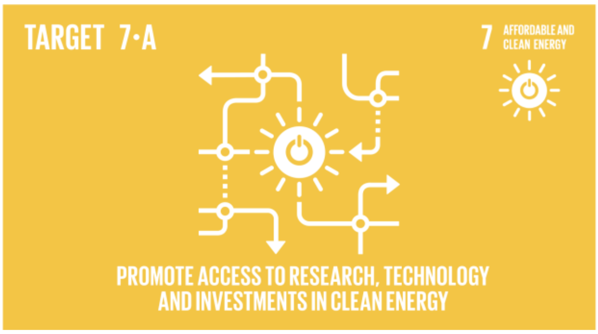


Figure 2.1 UN Sustainability Target 7·A [14].

This thesis influences these two goals regarding societal impact. Applying artificial intelligence and digital twins may allow for effective decision-making for increasing energy efficiency in both design and potential retrofitting stages. The application of such a technology may have a large societal impact in years to come, especially since such a framework is not exclusively limited to the theme of this thesis.

3 Theoretical Background

The theoretical background of this thesis is based on a literature review, attempting to identify relevant papers and provide the necessary background for the theme in question. Many of these papers are bibliometric studies which include current research status and present relevant case studies.

3.1 Building Information Modelling (BIM)

In the context of digitalization in the AEC industry, BIM is essential. BIM allows for efficient collaboration among stakeholders in AEC and is applied in all parts of building life-cycle and building life-cycle management [2]. The term BIM arose in the early 90s [2] and started emerging in AEC in the early 2000s [15]. It has gained increased traction in the last decades. BIM is considered to be widely adopted in AEC post-2010 [2].

The use of BIM and BIM processes has revolutionized the AEC industry. It has been found that benefits include 61% reduced errors and emissions, 20-30% reduced construction costs, 17% faster approval cycles, 35% reduced rework, and 20% reduced project duration [16].

As the industry develops, new areas of research and technology emerge, and consequently, so do the scope and application of BIM. This highlights the need for digital tools combined with BIM. This holds especially true for Building Energy Modelling (BEM) [17].

There are, however, fields in which BIM is currently lacking or not fully adapted. Examples include areas such as risk management and incorporation of IoT (Internet of Things) in areas such as construction performance and progress monitoring [18] or energy efficiency and environmental optimization design [6]. In recent years there has been an increased discussion on the limitations of BIM [19].

3.2 Digital Twin

Digital Twin (DT) is a concept that has been developed to overcome some of the shortcomings of BIM. Digital Twins first arose in the aerospace industry in the 1960s with two simulators to mirror space conditions to prepare flight training for the National Aeronautics and Space Administration (NASA)'s Apollo 13 Program [20]. The term "Digital Twin" first arose in Michael W. Grieve's Product Life-Cycle Management model (PLM) in 2002 [21].

Digital Twins is primarily applied for engineered products, production lines or production machines [16]. Recently, it has gained increased traction in AEC and is now considered an integrated part of Construction 4.0 [2].

However, its application in the AEC industry is still in its infancy, as illustrated by a bibliometric study by Naderi & Shojaei [21]. The figure below displays the frequency of annual digital twin-related publications in the AEC field.

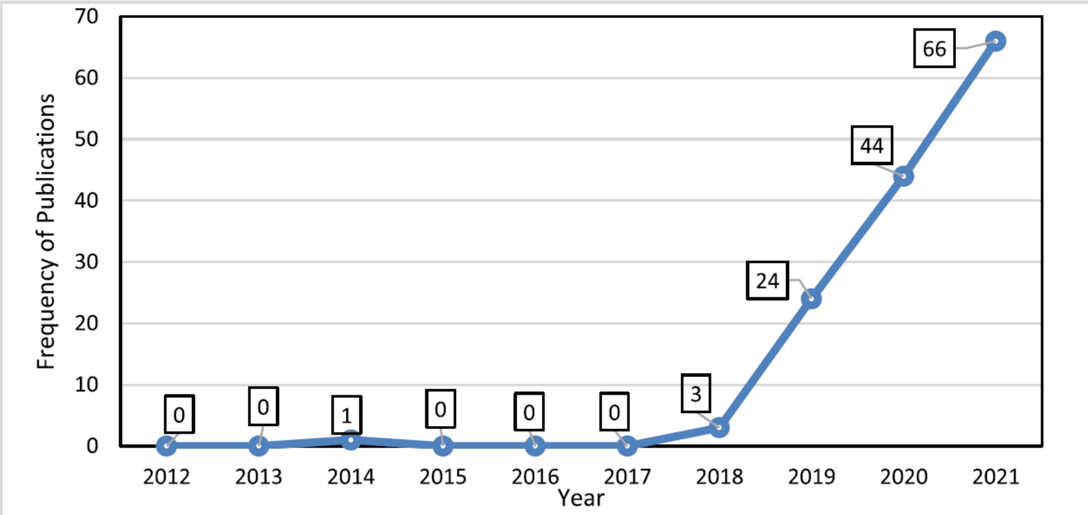


Figure 3.1 Digital Twin related publications in AEC [21].

As may be expected from such a new field of research, there is still a lacking consensus on the definition of a digital twin. A study by Camposano et al. investigated different metaphors for understanding digital twins. Their study was performed using semi-structured interviews with project managers and C-level executives in the Finnish AEC industry. Camposano et al. found that from a practitioner's point of view, a digital twin was considered a complete replica of an asset that may substitute for its physical counterpart [15]. However, its definitions vary by application. For example, Camposano et al. found that the digital twin was defined as a: life-cycle representation, visualisation tool or cyber-physical system [15].

Yitmen, Alizadehsalehi et al. published a paper introducing subforms of digital twins, hybrid twins and cognitive twins [22]. According to Yitmen, Alizadehsalehi et al., Cognitive computing is a machine's ability to mimic the human ability to think, sense and make decisions in any given situation [22]. Based on this definition, It may as such be considered that cognitive computing is closely connected to Artificial Intelligence (AI).

Their study identified a broad range of applications for digital twins: Real-time data presentation by presenting real-time physical entity data. Analytics, for analyzing stored data to provide helpful insight. Simulations to run various data. Visualization, by overlaying real life and live 3D BIM models, automation, by a bi-directional dataflow that may influence the behaviour of the physical asset. Predictions to predict future behaviour based on analytics and historical data.

Yitmen, Alizadehsalehi et al. argues that a digital twin contains three components in a practical loop, a physical entity, a digital entity and a data link [22]. Further, there are two methodologies for mapping information to the digital models, inspection data and simulations or predictions and optimization. Inspection may include IoT integration, predictions and optimization achieved by learning data.

A cognitive twin, according to Yitmen, Alizadehsalehi et al., is a digital twin with the ability to detect anomalies and learn with the power to improve measures of its state and function [22]. Further, cognitive digital twins apply optimization to aid decision-making for the physical twin. In such a context, it is challenging to differentiate the two. However, Yitman, Alizadehsalehi et al. argue that a cognitive digital twin is an evolution of the digital twin framework and allows for visualization in a dynamic process that allows the physical assets to be modified [22].

Another definition arose from a bibliometric study conducted by Nour El-Din et al., which defined Digital Twins as "the connection between technologies in the form of a digitized model of a physical asset, transmitting data in at least one direction, and monitoring the physical asset in real-time" [20]. Yet another definition from Deng et al. argues that Digital Twin is an evolution of BIM, BIM level 5, "Which provides seamless interaction between the virtual and real-world environment" [2].

Camposano et al. argues that a digital twin is a "complex software ecosystem that emerges from the increased expectations that AEC/FM stakeholders place upon BIM and other related technology" [15]. Yet another definition was proposed by Zhao et al., who argue that Digital Twin is "a realistic digital representation that combines data describing the physical assets, processes and system in a digital format" [23].

In terms of the development of digital twins. A bibliometric study by Deng et al. describes the evolution of Digital Twin from BIM in the built environment [2]. Their study defines five generations of BIM, where digital twins are at the pinnacle, defined as 5D BIM.

Generation 1 BIM constitutes design, construction and scheduling. Generation 2 combines BIM with simulations for facility management purposes, including operations, estimations and sim-based predictions. Generation 3 consists of real-time visualizations, associating BIM with sensory information technology such as IoT devices. Generation 4 combines BIM and AI for decision-making and data-based prediction purposes. And finally, 5D BIM is based on automatic feedback from the environment and allows for automatic decision-making based on optimized results, typically achieved through AI and Machine Learning algorithms. In other words, a digital twin is considered a level of maturity where AI is an integral part of BIM concepts and purpose [2].

Furthermore, the border between BIM and digital twins appears blurred [21]. On the one hand, the digital representation of a model in a 3D environment, even with implemented tools such as the Internet-of-Things (IoT) for observational purposes, still is a digital 3D model. On the other hand, incorporating this information makes it arguably more advanced.

The lack of a standard definition poses a problem in the industry, as pointed out by Camposano et al. [15]. Establishing a common definition may help reduce the hype and practical implications of the concept as a "buzzword" and help the industry understand what is economically and technically feasible from a practitioner's point of view. To illustrate, "Digital Twin" was ranked among the top 10 strategic technology trends in 2017, 2018 and 2019 [15]. Its market value is expected to reach 73,5 billion USD by 2027 [20]. Furthermore, it has reached the pinnacle of the so-called "hype circle curve" of emerging technology [15].

It is worth noting that the research on Digital Twin is by and by large published in a limited amount of journals, where automation of construction is overrepresented. Naderi et al. found automation in construction to be the primary core journal on Digital Twin research [21]. Furthermore, the research field of Digital Twin currently suffers from a severe degree of "inward-looking", with co-citations limited to its own domain of research. Indicating that the knowledge and implementation of Digital Twin in AEC is isolated from other relevant fields of research from other domains and disciplines [21]. These findings also correspond with the findings of Darko et al. [5] and Maureira et al. [24] in the application of artificial intelligence in AEC.

In this context, it is appropriate to introduce a contemporary international project, SPHERE (Service Platform to Host and Share residential data) is an EU-funded 4-year horizon project. The project aims to provide citizens, AEC stakeholders, administrators, and urban developers with an integrated ICT (Information Communication Technology) for better evaluation and assessment of energy performance at the start of the construction process [16]. SPHERE integrates digital twin in the construction and maintenance phases, provides a collaborative platform between stakeholders, and enables efficient data interoperability between digital twins with BIM software tools.

SPHERE aims to reduce 15 % of energy demand in the residential building's operational phase, 25 % reduction in construction time, and 25 % reduction in GHG emissions in the construction and operational phases [16]. It is argued that the SPHERE platform, with its digital twin implementation, will improve the building's energy performance from the initial stages of the construction project, allowing different stakeholders to use this technology [16]. It may be considered an effort to establish a framework and application of digital twin technology on a large scale with numerous stakeholders involved. In such a context, the SPHERE project may serve as an incentive to adapt further and implement digital twin technology.

Despite different definitions, there appears to be a consensus that a digital twin framework consists of 3 parts. The physical twin, the digital twin and the decision-making process. One may consider interconnectivity displayed in the figure below:

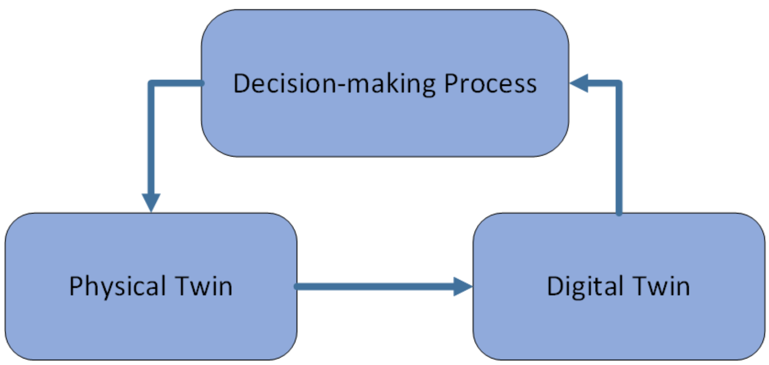


Figure 3.2 Digital Twin Framework.

The physical twin represents the actual physical building, providing the basic semantic data in a BIM environment and potential real-time data utilizing IoT devices. The digital twin compiles this information, providing a framework for fault prediction, performance predictions and simulations/optimization. These two form the base for decision-making, which may be applied to the physical model.

Per this thesis, a digital twin is "a digital asset representing a physical construct with third-party information flow". This information serves as a baseline for data analysis and aids decision-making using AI applications.

As such, it is not considered a cognitive model as described by Yitmen, Alizadehsalehi et al., as it does not allow for constant modification of the physical assets' behaviour. Also, it does not correspond with Deng et al.'s definition of a generation 5 BIM or Nour el-din et al., as it does not have seamless interaction between virtual and real-world environments. Rather it represents the information required to do so. Further, it differs from Zhao et al., as it does not describe the processes or system of a physical asset. And finally, It differs from a digital model by the definition of receiving semantic information from a third party and visualizing it.

One could argue that one key benefit of digital twins is that it allows for efficient decision-making based on semantic data. This may be accomplished either by human decision-making or by implementing artificial intelligence.

3.3 Artificial Intelligence in AEC

Artificial intelligence has been around since the 1950s, whilst its application to the AEC has been an area of research since the 1970s [5]. The area has experienced a growing interest since the early 2000s, as can be identified in the following figure:

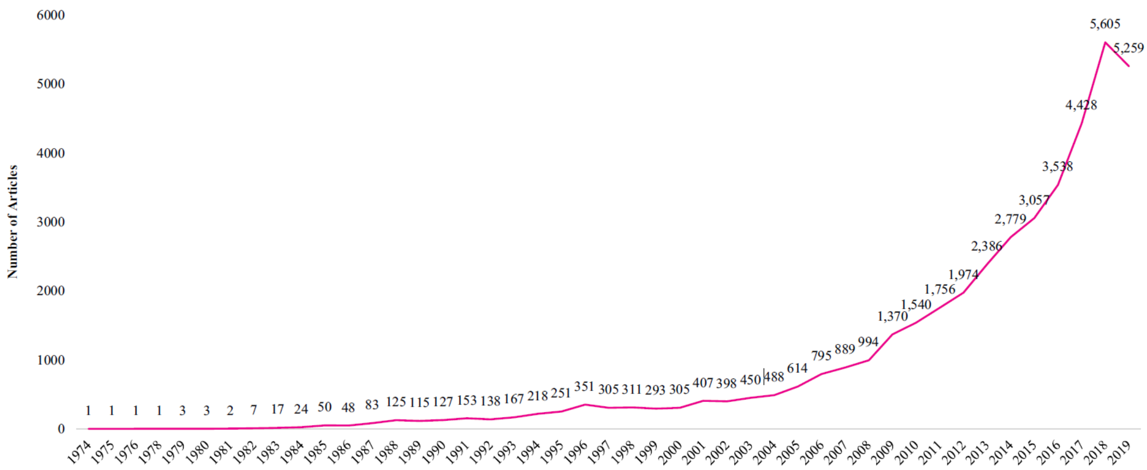


Figure 3.3 Trend of research publications on artificial intelligence in AEC [5].

This figure originates from a study by Darko, Chan et al. [5]. Their study conducted a mapping of research publications through a bibliometric analysis. The figure shows a decline in research in 2019. This is due to the date of the study itself, which was completed in May 2019.
The study identified that some of the leading research areas had been optimization, machine learning, genetic algorithms, and construction management. The most common application of artificial intelligence in AEC was found to be genetic algorithm optimizations [5].

In their bibliographic study, Darko, Chen et al. identified a list of the relative influence of existing research on artificial intelligence in AEC based on a co-occurrence keyword analysis. The findings reflected that the top four categories were: Optimization, Genetic Algorithm, Neural Networks, and Simulation [5].

By the list of relative influence on a scale from 1-100, the study identified areas in which artificial intelligence is under-represented, indicating a need for additional research. Some key examples are Life cycle costs (89), Energy (87), and Energy Efficiency (58). Additionally, the study identified that there had been limited research on the application of machine learning for optimization in these fields [5]. These are the fields that are the focus of this thesis.

A bibliometric study conducted by Zhang, Chan et al. found that two main topics were predominantly covered in recent BIM and AI research literature [4]. Firstly, the application of artificial intelligence by means such as facility management, fault detection, safety management, diagnosis etc. Secondly, the AI techniques covered in AEC, such as machine learning, artificial neural networks and genetic algorithms for optimization [4]. In terms of application, they found that AI may be integrated with BIM and provide benefits in the whole life cycle of projects [4]. Examples follow below:

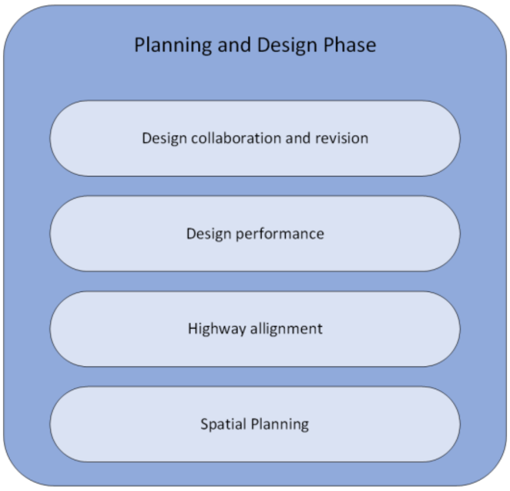


Figure 3.4 Zhang, Chan et al. - AI in AEC - Planning and Design Phase.

Design collaboration may benefit from AI by extending functionality, and algorithms may be developed to propose alternative designs or optimize designs based on multiple objectives.

Changelogs within the BIM model may monitor design performance. This allows project managers to assess the design performance of given teams and provide predictive means to improve the modelling process. Highway alignment design may benefit from AI integrated with BIM using Machine Learning methodology to predict pavement performance.

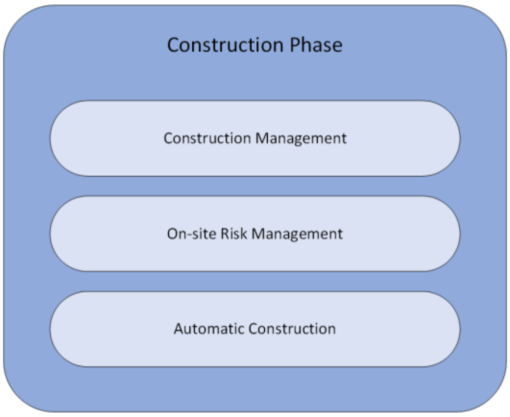


Figure 3.5 Zhang, Chan et al. - AI in AEC - Construction Phase.

Automated construction is a field that is used for repetitive tasks and working in hostile environments. Some tasks that may be automated are brick assembly, welding, material dispatching and routine fabricating. AI may assist in optimizing construction management by site planning, minimizing moving distance for workers, simulating interference by workers during construction, safety hazards, automated assistance in construction scheduling and choice and placement of cranes, to name a few. Like construction management, life cycle costs may be achieved by optimising different variables, such as cost-optimal maintenance and replacement of building components, design schemes and building process. In particular, a sub-form of machine learning supervised learning may be applied to predict costs from the BIM model, depending on the level of detail included.

Figure 3.6 Zhang, Chan et al. - AI in AEC - Operational and Maintenance Phase.

AI in the operational and maintenance phase covers localization utilizing machine learning by comparing images or augmented reality with 3D models (digital twins). The same process may enhance fire safety by recognising and developing dynamic evacuation plans depending on location. Building maintenance may be achieved by gathering real-time data forming a basis for predictive means by comparing datasets and performance. Structural health monitoring works similarly, where dynamic monitoring may help identify the structure's health using machine learning algorithms.

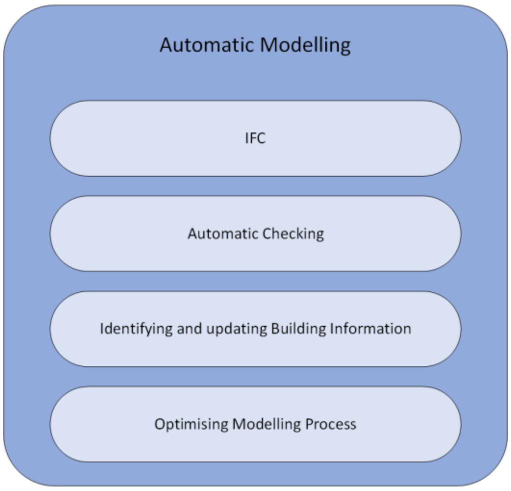


Figure 3.7 Zhang, Chan et al. - AI in AEC - Automatic Modelling.

In the case of automatic modelling. Artificial intelligence may be applied to compare real-life measurements with the original model, forming a basis for an accurate digital twin and optimizing this information flow and comparing in-situ geometric data with the digital twin. Another application is reducing data loss in the IFC interpretation caused by misclassification. For example, lacking data from the BIM model may be interpreted or predicted using AI techniques. Machine learning algorithms may be applied to automatically identify and classify objects, increasing the semantic integrity of the model. Furthermore, given the generic nature of traditional geometric design, an algorithm may produce suggestions to the modeller in the early phases of the project. Zhang, Chan et al. argue that this will increase productivity significantly [4].

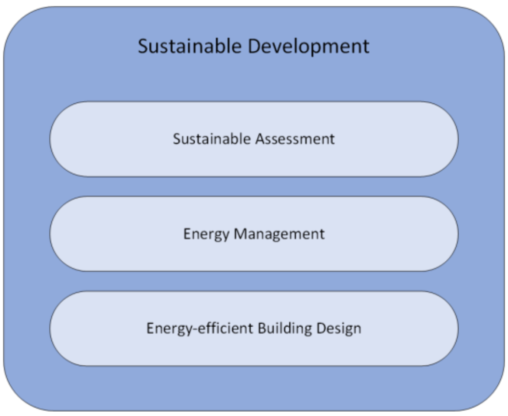


Figure 3.8 Zhang, Chan et al.- AI in AEC - Sustainable Development.

Artificial intelligence and BIM play an essential role in the sustainable development of the AEC industry. AI may be applied to assess building environmental performance compared to relevant certifications such as LEED or BREAAM etc. In terms of energy management and energy efficiency, AI may be applied for energy management, provide energy-saving suggestions (both in the operational and design phases), and reduce energy consumption by optimizing design for envelope or internal illumination.

A similar bibliometric study to Zhang, Chan et al., by Pan & Zhang identified that there had been a shift in research areas of AI combined with BIM [25]. From the years prior to 2017, the research was mainly placed at BIM as a tool, with areas focused on architectural design, information theory, semantics, ontology and interoperability. Between 2017 and 2018, the focus of research shifted to decision-making, life cycle, structural design and sustainable development. Whilst recent studies have focused on automation, machine learning, optimization, and risk assessment, to name a few [25]. This indicates that the focus of research on BIM-AI has shifted from theoretical BIM-focused studies to the practical application of artificial intelligence.

Pan & Zhang argue that AI is an ideal data-retrieval and analysis solution when combined with BIM in all life cycles [25]. In their study, they identify examples in 3 steps of the life cycle: The planning and design phase, the construction phase and the operation/maintenance phase.

BIM facilitates planning and design in large parts. BIM as a tool helps reduce costs, duration and workflow. Integrating AI with BIM has the potential for more efficient design through optimized information exchange between project stakeholders. Further applications may be automated design and drafting, code compliance, and validation of design standards.

During the construction phase, BIM provides benefits such as reducing rework, errors and conflicts on site. Further benefits may be progress monitoring for hazard detection, optimizing 4D BIM for construction schedules and logistics or implementing robotics to replace workers and worker exposure to hazardous environments. The application of AI may assist in predicting risk issues and allow for prevention measures by using tools such as Augmented Reality (AR), Unmanned Area Vehicles (UAV) and Internet-of-Things (IoT), to name a few.

The operation and maintenance phase accounts for approximately 60% of the total project budget [25]. BIM in itself is still insufficient for the operation and maintenance of buildings, according to Pan & Zhang [25]. It is causing error-prone decision-making by facility managers. AI may help in the operation phase through the applications of IoT and data mining for operational monitoring or visualizing building performance. This will allow for more efficient decision-making, risk/error detection and adjusting day-to-day services efficiently, economically, and reliably.

Although the level of application of BIM and AI is increasing, there is still a reduced level of adoption from the industry. The reasons for this are diverse but likely include high costs, trust, talent shortage, internet connectivity and project uniqueness [25].

One should keep in mind that AI method development is complex. As such, it is essential to establish the value of applying AI as a methodology, reduce the implementation costs, to further the interest in artificial intelligence as a field.

3.4 Visual Programming

Another area of research which has received increased interest over the last few years is the combination of BIM and visual programming/parametric modelling in the AEC [26]. Programming has a large variety of application areas, also within the field of AEC. However, it has been found that professionals in the AEC generally have limited to no programming experience [27].

Visual language is commonly defined as "*a formal language with visual syntax and semantics*" [27]. Combining this with programming, a Visual Programming Language (VPL) may be interpreted as a tool to make a program based on well-defined syntax and semantics through visual aids. In the construction industry, visual programming language is primarily used to generate geometric and semantic data or to check existing models [27]. One of the benefits of visual programming is that it is easier to understand and interpret than regular coding, which allows non-experts to write and create advanced programs.

Visual programming interfaces are usually divided into two components: the canvas and the library, as illustrated in the figure below. The canvas, displayed on the right, is used for the workflow, illustrating the processes and their order. The library, displayed on the left, contains standardized functions, such as geometric functions, math operators, scripts, and export/import processes, amongst other things.

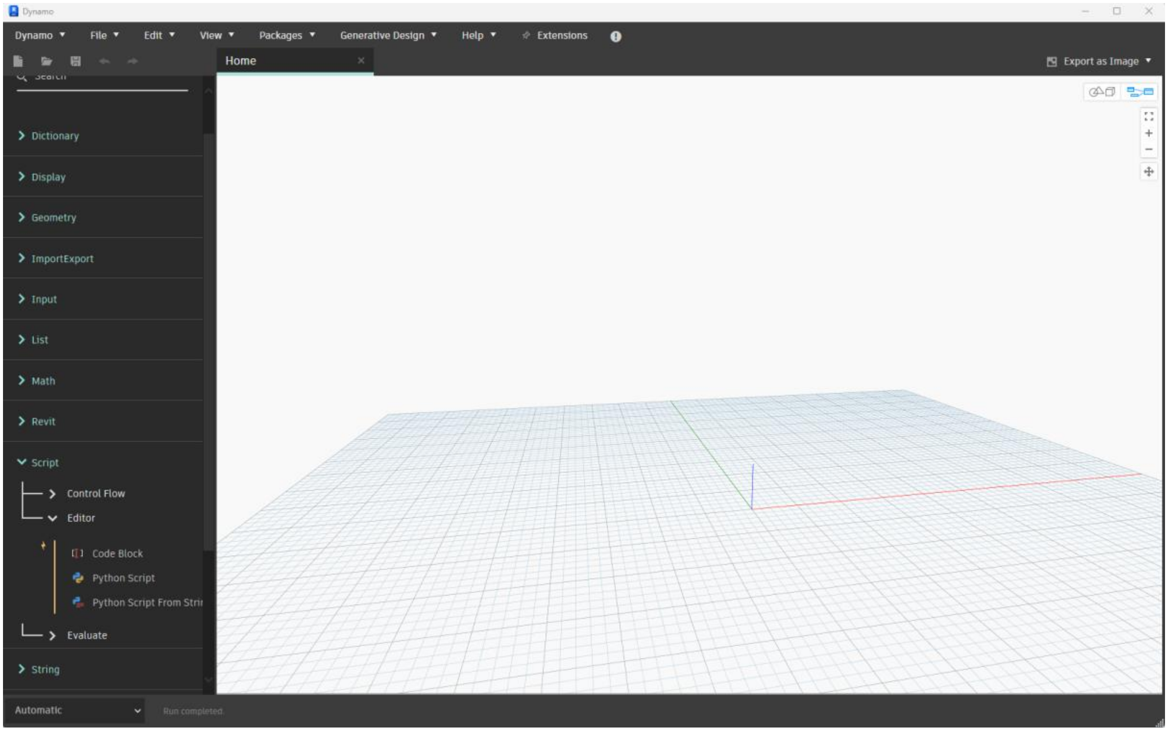


Figure 3.9 Visual Programming Interface.

A typical way of working with visual programming is illustrated in the figure below. This workflow is inspired by design research methodology, initially a way of structuring a research process [28]. The process may be described as identifying the purpose and the problem to be solved and separating it into subgoals and increments of processes. Once the goal is defined, one can design a solution. When applying this solution, one can evaluate its function and success.

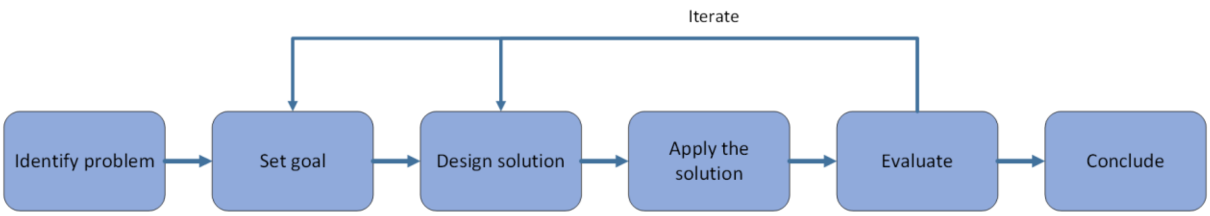


Figure 3.10 Visual Programming Workflow.

For semantic data, this may be achieved by previewing the data as a list generated by the node (in this case, the *range* function), as illustrated in the following figure:

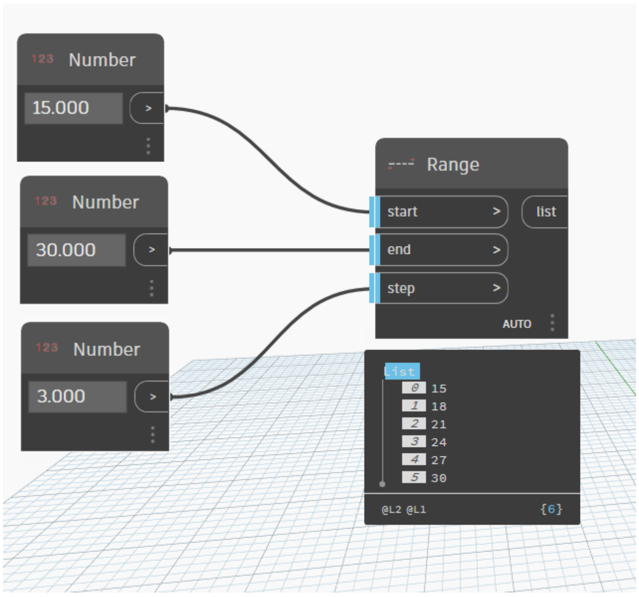


Figure 3.11 Visual Programming data preview.

AEC's most common visual programming tools are Dynamo for Revit and Grasshopper for Rhinoceros 3D [27]. The program functions by placing operators displayed as boxes from the library and connecting them by wires. The dataflow functions as inputs on the left of the box and output on the right. The library may also be expanded by textual programming tools such as C+ or Python scripts to expand its functionality further [28].

The application of VPL is not problem-free, however. It is generally considered that debugging or maintaining a VPL script is impractical for anyone other than the programmer himself. As the program becomes increasingly complex, it may cause high latency and be hard to interpret for others than the author [29]. As illustrated by the visual program in the figure below:

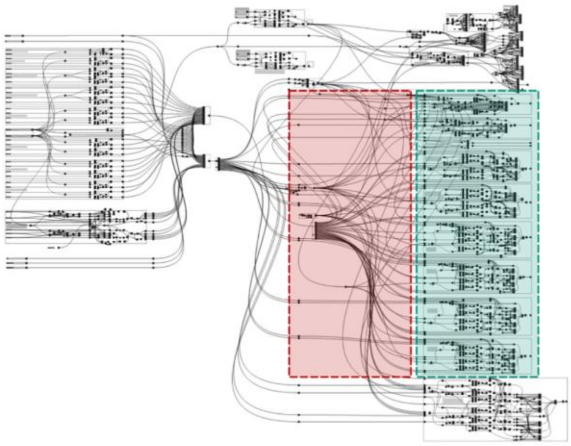


Figure 3.12 Example – Complicated Visual Program [27].

The software providers have attempted to reduce this risk and challenge by allowing for the creation of functional units, which is a simplified presentation of a given function, respectively called "*node*" for Dynamo and "*cluster"* for grasshopper.

An example of semantic data which may be imported is third-party information complied in Excel. When this data is combined with the integrated VPL in BIM software, it allows for visualization of this data. Depending on dataflow and potential incorporation with IoT, this visualization may also be real-time. There have been conducted several studies integrating BIM and VPL. A selection of case studies is introduced below:

Notably, the work by Hosamo, Tingstveit et al. visualize predicted mean vote (PMV), representing the mean value of thermal comfort reported by a given population in a BIM environment [6]. Another study by Hosamo, Hosamo et al. streamed an optimal solution for energy and thermal comfort optimization solution through a multi-objective optimization algorithm back to BIM [30]. Desogus, Quaquero et al. streamed real-time semantic data collected from IoT devices into a Digital Twin [31].

It would appear, however, that this field is underutilized, given its significant benefits and integration in conjunction with Digital Twins. This thesis will stream the degree of optimization for costs and energy efficiency and may be considered an innovative application of visual programming.

3.5 Machine Learning

For this thesis, Artificial Intelligence is defined as the ability of a computer to imitate human-like decision-making. In this context, Machine Learning is defined as the ability of a computer to make human-like decision-making based on *experience*. Machine Learning may, as such, be considered to be a field of study that allows a computer the ability to learn without giving explicit information on how to do so. Generally, we separate machine learning into three categories: Supervised learning, Unsupervised learning, and Reinforcement learning [32]. The focus of this thesis lies in supervised learning.

Supervised learning is mapping inputs to an output, where the goal is inferring a function from the information [4]. Neural networks, like Artificial Neural Networks (ANN), are a part of this category which learns sample patterns by tuning neural parameters. It has been widely successful in approaching non-linear problems in the AEC industry [5]. One of its key applications is optimization, for example, architectural design [33] and structural engineering [5]. It is worth noting that ANN is not an optimization tool in itself but is used for classification or regression analysis of semantic data which may then be optimized using a variety of methodologies.

Artificial neural networks is a network where each neuron layer is interactively connected with one input layer and one output layer [4]. We separate between 3 layers: the input layer, which is the input parameters, one or more hidden layers, which interpret inbound information and the output layer. ANN is also versatile, as it can be used for classification and regression tasks and handle complex non-linear problems [4].

The process may be illustrated in the figure below. In this setting, one may consider that by providing the input and the output value, the machine learning process infers a connection or correlation between them in one or more steps. The input layer serves as the starting point, listing all variables connected to the network. So, an artificial neural network equals nodes (*Neurons*) and *Network* (the connection between them).

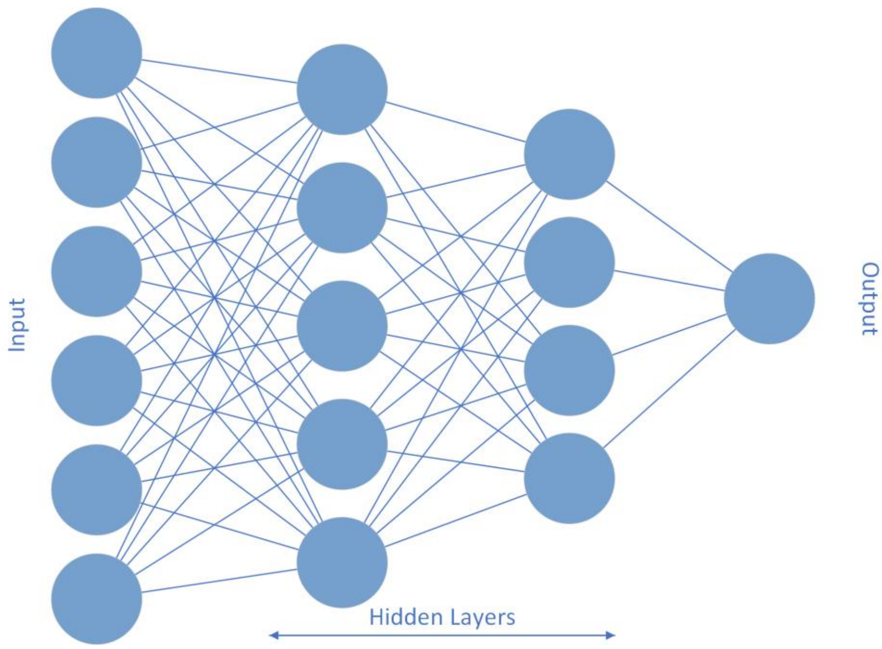


Figure 3.13 Neural Network.

Each parameter in the original input is connected to all neurons in the following layer. The neurons in this layer (in this case, the first hidden layer) apply a mathematical evaluation of the input variables and an added bias to the layer itself to define whether or not to provide an output. The bias may be described as how "hard" it is to activate the neurons of the following layer.

Each neuron has several inputs: the nodes' input and the bias value. This is then transformed using various mathematical functions, for example, the sigmoid function, the step function, the rectified linear unit (Relu) function or the Levenberg-Marquart function [34, 35]. These functions will not be further described in this thesis. However, one may consider it a mathematical function treating inputs to define whether the node is "activated" by assigning a value between 0 and 1. This value is then treated similarly with weights in the following layers. As information flows one way from the input layer to the output layer, weights are applied to the connections, which are usually random. Once the data has passed through the model, the model then evaluates and sends information back by adjusting these weights. This process is known as backward propagation. A complete run back to the inputs is called an epoch.

A simplified illustration is shown in the figure below, depicting connected variables (x_n), weights (w_n), bias (b), and activation function, $f(x)$:

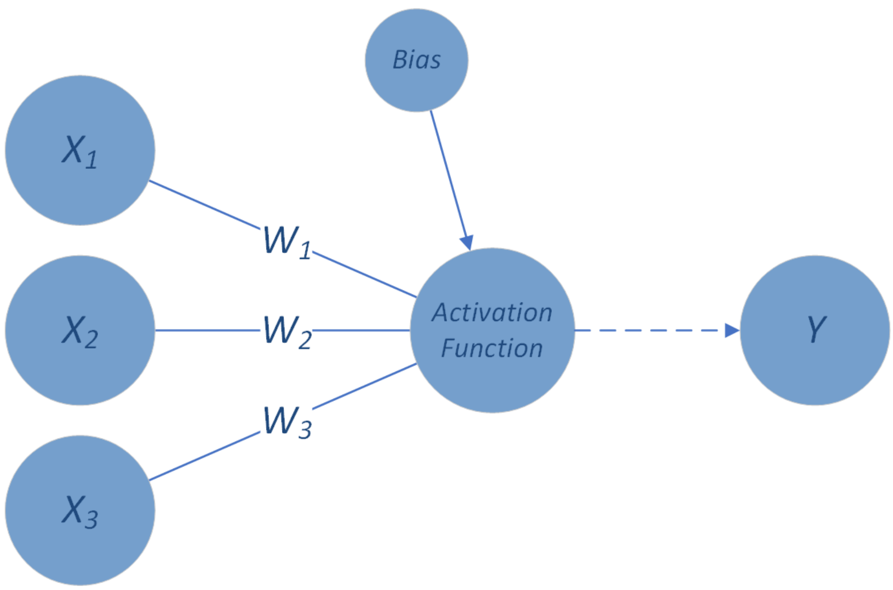


Figure 3.14 Neural Node.

A neuron calculates the weights of the inputs combined with its values with an added bias.

One may consider inputs:

$$x_1, x_2, x_3 \dots x_n$$

With corresponding weights

$$w_1, w_2, w_3 \ldots . w_n$$

With an added bias

$$b$$

And activation function

$$f(x)$$

The output of a neuron may be described as:

$$y = f(w_1 x_1 + w_2 x_2 + w_3 x_3 + b)$$

Equation 3.1 Artificial Neural Network Equation.

For practical purposes in the creation of an Artificial Neural network, one may define five parameters mentioned below:

1. Bias
2. Weight of network connections
3. Activation function
4. Number of Neurons per layer
5. Number of Layers

Artificial neural networks have several benefits: They can create their own features (combination of nodes), and given their flexibility, they can model complex non-linear hypotheses, and it is also very accurate. One of the drawbacks is that it requires a significant amount of training data, and the function it creates is very complex. Interpreting and its weights are complicated, as the value of the input transmitted to different neurons and producing output is not intuitively discernible [32].

We separate training, validation, and testing data in the context of artificial neural networks, which are separate datasets from the original data pool. The training data, as the name might suggest, is used to train the model itself to create a fitting function. It identifies a correlation between a large dataset and the given output. The validation data is a set of data that are then used to validate the models' accuracy and tune the model's hyperparameters to fit the validation dataset. The test dataset is applied to test the model adjusted by the validation set on a new dataset.

In terms of the performance of a machine learning model, we apply two terms, overfitting and underfitting. An overfitted model may be described as a model that works well on the training data but poorly on the validation data. In practical terms, this may be considered a model that is overly accurate on the testing data and cannot generalize to other datasets. An underfitted model, however, is a model that does poorly on both training data and the validation model. Which, briefly explained, may be considered to be a poor model. An example of overfitting and underfitting is identified in the figures displayed below:

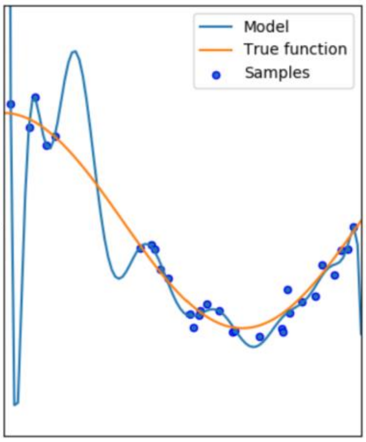


Figure 3.15 Overfitted Model [36].

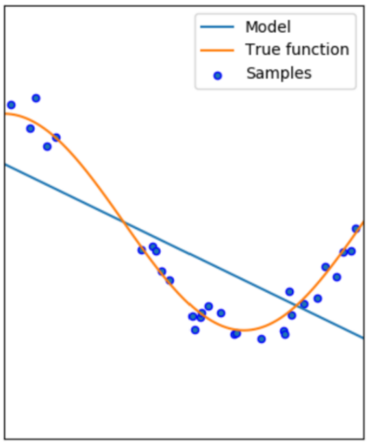


Figure 3.16 Underfitted Model [36].

A common way of estimating the error of a machine learning model is the mean square error (MSE) and the correlation coefficient value (r) [37]. In essence, the mean square error is the mean distance between the forecast and the actual value from the regression line squared, the lower the mean square error, the more accurate the prediction model. One may consider the graph below for illustrative purposes. Where the distance error is depicted for each value plot outside the generated regression model. In this context, the optimization of the ANN training process is worth mentioning.

The Levenberg-Marquardt algorithm is a variation of the newtons method, where the goal is to minimize functions where the sums of squares of nonlinear functions. This makes it suitable for ANN networks, where the performance index is based on MSE [38, pp. 12-19].

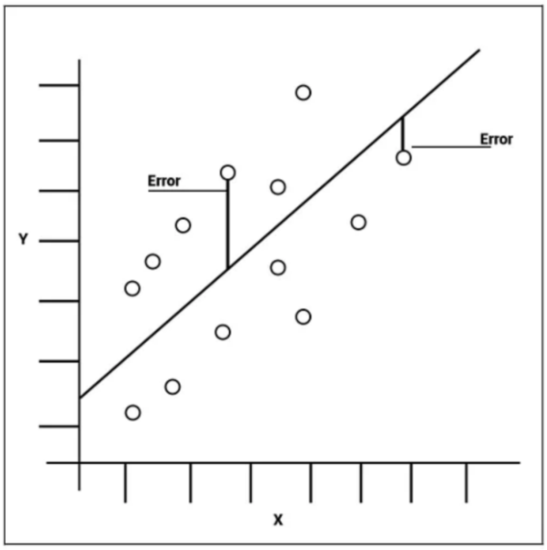


Figure 3.17 Regression plot with errors [37].

If one considers the regression plot with errors in the figure, the MSE is the mean value of this distance squared.

The formula for mean square error may be described as:

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

Equation 3.2 Mean Square Error.

Where:
y_i – Actual value
$\hat{y}_i$ – Predicted value
n – Number of errors
i – increments

A more intuitive value for evaluating the model's accuracy is the r value (the Pearson correlation coefficient). The r value ranges between +1 and -1. A value of + 1 describes a perfectly positive correlation, - 1 is a perfectly negative correlation. A value of 0 describes no correlation between data points.

The equation of r may be described as:

$$r = \frac{n * (\sum(X, Y) - (\sum(X) * \sum(Y)))}{\sqrt{(n * \sum(X^2) - \sum(X)^2) * (n * \sum(Y^2) - \sum(Y)^2)}}$$

Equation 3.3 Correlation coefficient.

Where:
r – Correlation coefficient
n – Number of observations

In the context of the practical application of machine learning and artificial intelligence, it is appropriate to introduce the new OpenAI tool, ChatGPT 3.5, which may be found on the OpenAI webpage [39]. ChatGPT is developed using reinforcement learning provided by human feedback, where potential answers are ranked or a lack of follow-up questions, indicating that the answer is satisfactory. The tool functions as a chatbox, where one is given an output from an input question.

In terms of application, ChatGPT has been found as a powerful tool in the context of creating scripts [40]. However, the tool has weaknesses identified and stated on the OpenAI webpage [39]. For example, occasionally, ChatGPT provides plausible but incorrect answers, and it is sensitive to tweaks on input or even the same question. Given the nature of these weaknesses, it is essential to quality-check the answers.

3.6 Energy Performance

Energy performance is a field of research that has gained increased interest in the last few years. A bibliometric study by Azima & Seyis, published in 2023, aimed to investigate the current state of energy performance research in the AEC industry [41]. The findings reflect that the number of publications in the last decade has increased drastically, as can be identified in the following figure:

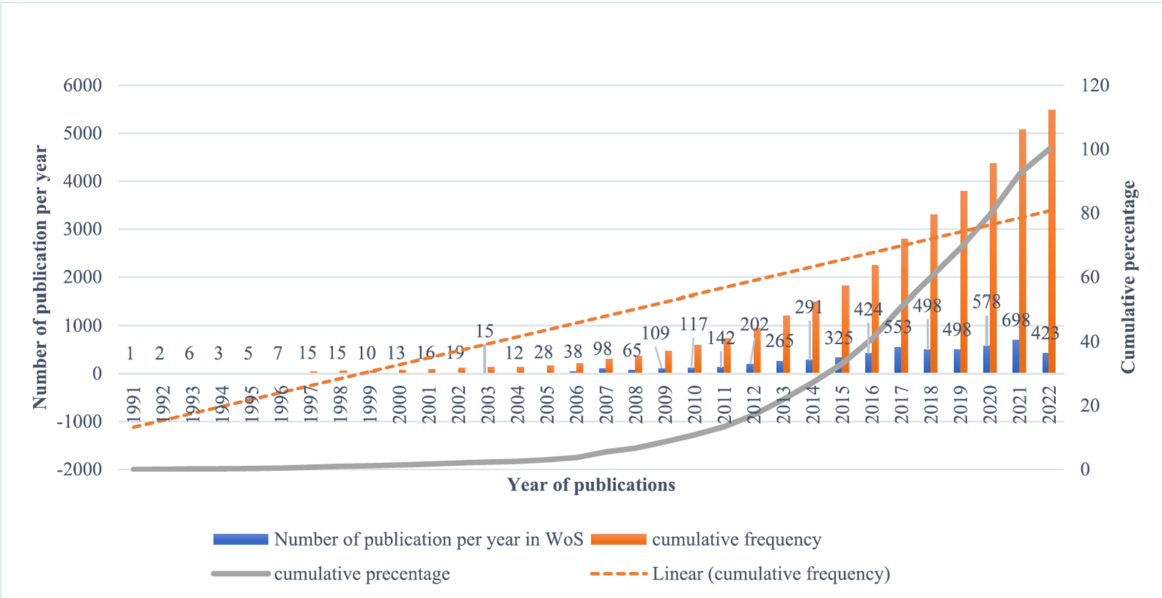


Figure 3.18 Number of publications on energy performance between 1991 and 2023 [41].

Interestingly, however, through the co-keyword analysis of the study, Azima & Seyis found very limited research had been performed on energy performance combined with artificial intelligence and information technology [41]. This is the focus of this thesis.

Energy performance in itself is a highly dynamic and complex subject. For the purpose of this thesis, we focus on the energy heat loss of envelopes and argue that this accounts for a large part of the energy consumption of a building.

The thermal performance of a building is arguably the most crucial parameter for energy demand calculations as it is directly influenced by its environment, and particularly the U-Value (thermal transmittance) is important [42]. The U-value is defined in ISO 7345 as "the steady-state heat flow divided by the area of a system and the temperature difference between the surroundings on each side" [43]. The methodology for the theoretical calculation of the U-value is provided in ISO 6946 [44].

For this thesis, we limit ourselves to the simplified calculation methodology for thermally homogenous layers. The simplified calculation methodology for thermal transmittance is defined in the following equation:

$$U = \frac{1}{R_{tot}}$$

Equation 3.4 Thermal Transmittance.

Where:
U – thermal transmittance [W/(m^2·K)]
R_{tot} – total thermal resistance [(m^2·K)/W]
For homogenous layers, the thermal resistance is defined in the following equation:

$$R = \frac{d}{\lambda}$$

Equation 3.5 Thermal Resistance.

Where:
d – thickness of the material component [m]
λ – the design thermal conductivity of the material [W/(m·K)]

Thermal conductivity (λ) is calculated through ISO 10456 or gathered through tabulated data. According to the standard, thermal resistance should be calculated to at least three decimal places.

The total thermal transmittance for a building component of homogenous layers is calculated by:

$$R_{tot} = R_{si} + R_{se} + R_1 + R_2 \ldots R_n$$

Equation 3.6 Total Thermal Transmittance.

Where:
R_{tot} – total thermal transmittance [(m^2·K)/W]
$R_1, R_2 \ldots R_n$ – refer to the design thermal resistance in the respective layer [(m^2·K)/W]
R_{si} and R_{se} – internal and external surface resistance [(m^2·K)/W]

Surface resistance is a measurement intending to account for airflow in contact with a surface. According to the standard, there are predefined values which may be applied. We separate horizontal and vertical airflow values.

Table 3.1 Surface Resistance [44].

Surface Resistance (m^{2}K)/W	Direction of Airflow		
	Upwards	*Horizontal*	*Downwards*
R_{si}	0,10	0,13	0,17
R_{se}	0,04	0,04	0,4

Surface resistance is a value calculated by the following equation:

$$R_s = \frac{1}{h_c + h_r}$$

Equation 3.7 Surface Resistance.

Where:
R_s – is the surface resistance [(m^2·K)/W]
h_c – is the convective coefficient [W/(m^2·K)]
h_r – is the radiative coefficient [W/(m^2·K)]

And

$$h_r = \varepsilon \cdot h_{r0}$$

Equation 3.8 Radiative Coefficient.

$$h_{r0} = 4\sigma \cdot {T_{mn}}^4$$

Equation 3.9 Radiative Coefficient for a black-body surface.

Where:
h_r – is the radiative coefficient [W/(m^2·K)]
ε – is the hemispherical emissivity of the surface
h_{r0} – is the radiative coefficient for a black-body surface [W/(m^2·K)]
σ – is the Stefan-Boltzmann constant: $5{,}67\ x\ 10^{-8}$ [W/(m^2·K^4)]
T_{mn} – is the mean thermodynamic temperature of the surface and its surroundings [K]

The values in Table 3.1 Surface Resistance are recommended as design values for plane surfaces in lack of the information described in the formulas above and are calculated under presumption for internal surfaces: $\varepsilon = 0{,}9$ and h_{r0} evaluated at $20°$, and for external surfaces $\varepsilon = 0{,}9$ h_{r0} evaluated at $10°$.

Material science is an integrated part of energy-efficient design. For interpretation purposes, one may consider that a low value of heat transfer (thermal conductivity) equals an energy-efficient material with good insulation properties, which is valuable for energy-efficient design. Materials considered to have suitable thermal insulation parameters are, for example, extruded polystyrene (EPS) with a thermal conductivity of 0,031 W/mK [45], or ultrafine glass wool with a thermal conductivity of 0,027 W/mK [46]. Currently, the material with arguably the best insulation performance is aerogel. Silica and organic aerogel have been found to have a thermal conductivity of approximately 0,003 to 0,012 W/mK [47].

Energy efficiency requirements for Norway are described in Byggteknisk Forskrift TEK 10 [48]. The most recent version of Byggteknisk Forskrift is TEK 17. However, energy-efficiency requirements are not covered in this version.

For example, school buildings have a maximum energy consumption requirement of 110 [kWh/m²] per year [48]. Furthermore, there have been defined energy measures that are category-specific functional requirements. There are no specific energy measures described for school buildings, however for comparison purposes, one may consider the following energy measures requirements for small houses and residential blocks:

Table 3.2 Energy Measures per TEK10, reworked from [48].

Energy Measures		
Category	Small House	Residential Block
U-Value External Wall [W/(m²K)]	≤ 0,18	≤ 0,18
U-Value Roof [W/(m²K)]	≤ 0,18	≤ 0,13
U-Value Floor [W/(m²K)]	≤ 0,10	≤ 0,10
U-Value Windows and Doors [W/(m²K)]	≤ 0,8	≤ 0,8
Fraction window door of heated Internal Floor Area (IFA) [-]	≤ 25%	≤ 25%
Annual average temperature efficiency for heat recovery in ventilation systems	≥ 80%	≥ 80%
Specific fan effect in ventilation systems [Kw/(m³/s)]	≤ 1,5	≤ 1,5
Air leakage figures per hour at 50 Pa pressure difference	≤ 0,6	≤ 0,6
Normalized thermal bridges, where m² defined as Internal Floor Area (IFA) [W/(m²K)]	≤ 0,05	≤ 0,07

The minimum value for energy efficiency is further described in § 14-3 (1a).

Table 3.3 Minimum level for energy efficiency per TEK10, reworked from [48].

Minimum Value for Energy Efficiency	
U-Value External Wall [W/(m²K)]	≤ 0,22
U-Value Roof [W/(m²K)]	≤ 0,18
U-Value Floor [W/(m²K)]	≤ 0,18
U-Value Window and Door (including frame) [W/(m²K)]	≤ 1,2
Air leakage figures per hour at 50 Pa pressure difference	≤ 1,5

Various floor area definitions (such as IFA) used for these requirements and values are described in the EN 15221 standard. Facility Management Part 6: Area and Space Management Measures in Facility Management [49].

It is worth noting that the unit for U-value and thermal transmittance is defined in degree Kelvin. This value is generally considered applicable for scientific purposes and not necessarily everyday use, which is usually degrees Celsius. However, since the difference between one kelvin and one Celsius is the same, one may use the two units interchangeably.

3.7 Life Cycle Costs

Life cycle cost is a standardized methodology to calculate construction's economic implications. Although many studies on energy consumption do not cover costs, they cannot be ignored in the context of sustainable building design [12]. There are numerous standards on the subject, but for this thesis, we limit ourselves to the European Standard for assessing the economic performance of buildings in terms of sustainability, the NS-EN 16627:2015 standard [50]. The standard specifies calculation methods based on Life Cycle Costs (LCC). The standard may be applied for various purposes: budgeting, tendering, estimating life cycle costs and waste streams, and specific economic analysis.

The standard process may be illustrated in the figure below. The following paragraphs will cover the separate steps displayed in the flowchart.

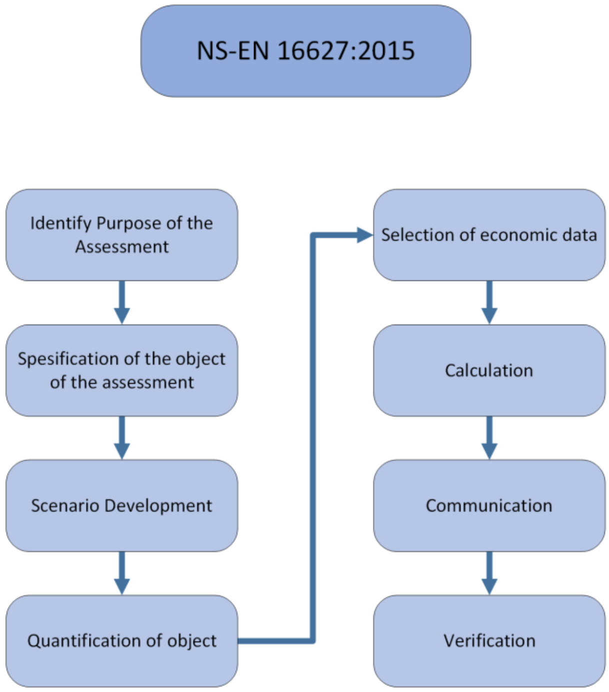


Figure 3.19 NS-EN 16626:2015 Flowchart, inspired by [50].

The overall purpose of using the NS-EN 16627:2015 standard may be argued to quantify economic performance.

The first step in identifying the purpose includes establishing the analysis's objectives, criteria, scope, and extent.

The specification of the object of the assessment, scenario development and quantification of the object steps refers to the process of identifying parameters and analysis requirements as well as confirming the project and asset requirements and options for analysis. These steps include identifying the period of analysis and methodology for economic assessment and identifying the need for potentially additional analysis (such as risk and sensitivity analysis). Identify project and asset requirements. And finally, identifying what options to be included in the analysis and which items to consider.

Selection of economic data. Covers assembly of costs and performance data and finalizes parameters for analysis. This step includes actions such as assembling costs and time-based asset performance to be used in the analysis, verifying values of financial parameters and analysis period. Also, potential risk strategy and uncertainty analysis is covered in this step.

Calculation. Include the calculations and the economic analysis and its results. Actions covered in this step are performing the economic evaluation, potentially the detailed risk and uncertainty analysis, and sensitivity analysis.

Communication. This step covers reporting and interpretation. Actions in this step are the interpretation of results and presenting them in the required format and preparing a final report.

The final step is verification, which covers the interpretation and reporting verification of the results.

3.8 Building Performance Optimization

Building performance is a highly dynamic concept and challenging to define. One may consider performance to be the building's ability to perform its task or intention, the degree of control of its delivery process or its success as a presentation [51]. The optimization process may be argued to be the selection of the best solutions from a range of alternatives for a given design problem according to the set criteria [52].

Building energy performance is a field that has gained increased interest in the later years. The earliest work was published in 1976 and has increased drastically in the last decade [51]. This is likely due to the emergence of evaluation systems worldwide, such as BREEAM (UK), LEED (USA), CASBEE (Japan), GreenStar (Australia), HQE (France) and DGNB (Germany) [51].

It is important to remember that optimizing variables does not equal optimizing the solution. For example, reducing energy usage is not enough to optimize the design, as the design must also cover other criteria [6]. This becomes increasingly important in instances with arguably contradictory goals, like the relationship between environmental and economic criteria [53]. In other words, sustainable building design may be argued to seek a balance between controlling negative environmental impacts, such as energy consumption and reducing the degree of non-renewable energy sources *and* their implementation costs [51]. This is an example of multi-objective design criteria.

Currently, building performance optimization is generally done in the design phase. However, studies have shown that it can be equally successful in later life stages [52]. For example, Hosamo, Hosamo et al. study on HVAC optimization [30] or Hosamo, Tingstvedt et al. study on model predictive control strategies [6].

A recent bibliometric study by Li, Liu et al. reviewed performance-oriented architectural design and optimization in the context of sustainability [51]. An apparent problem of energy optimization is its complexity. Li, Liu et al.'s study found that most current energy efficiency and optimization research suffered severe limitations [51]. For instance, simplified virtual buildings accounted for 86% of the current research on sustainable building design [51]. Due to its complex nature, there is no guarantee that simulations can simulate perfect solutions. The difference between a building's performance and its simulated performance is called the "performance gap".

There is also the concern of optimization and the different roles in sustainable design. According to the bibliometric study conducted, Li, Liu et al. found that engineers and computer technicians were currently leading the research development of sustainable building design [51]. The lack of inclusion of architects comes with a severe disadvantage, as there is little doubt that aesthetically pleasing design combined with optimization would serve as a driving force for sustainable development.

These objectives may be considered to be a multi-objective design criterion. Furthermore, the emerging concept of green buildings and green building design has shifted the framework. Architects and engineers no longer have clear boundaries. This highlights the need for cross-disciplinary collaboration and the necessity of allowing non-experts to make well-informed decisions.

Li, Liu et al. argue that the primary obstacle to achieving performance-oriented design and optimization in sustainability is the lack of middle-ware solutions [51]. For this thesis, visual programming is intended to serve as the middle-ware gap between the contradicting goals of lowest cost and energy consumption these in terms of envelope constructional design.

Another study by Attia, Hamdy et al. argues that the most significant challenge to building performance optimization is the lack of resources, appropriate tools, and the ability to define the problem sufficiently [52]. For example, defining the desired combination of objectives like lowest LCC, highest comfort or energy consumption [52]. Attia, Hamdy et al. further argues that future tools within optimization should include real-time solutions integrated within a BIM model, couple simulation with optimization, and connect real physical components to optimization models for better cost information [52].

3.9 Genetic Algorithms

According to Amos, Chana et al. application of Genetic Algorithms (GA) for optimization purposes has received the most attention in the AEC [5]. It has repeatedly been found to be one of the most successful tools in optimization [54].

A genetic algorithm is a form of an evolutionary algorithm that seeks to imitate Darwins' evolutionary theory of *"survival of the fittest"* but with numbers. The algorithm was first developed by Holland in

1975 and highlighted that there must be a sorting mechanism that selects the fittest members of one generation to pass to the next [54].

Genetic algorithms may be divided into single-objective and multi-objective genetic algorithms (MOGA). The differences should be apparent. Multi-objective optimisation aims to find an acceptable trade-off that satisfies the objectives and plots this on an optimal curve between the two objectives. Solutions on this curve have the characteristic that none of the objectives may be improved without affecting the other. This curve is called the Pareto Front [39].

One of the drawbacks of genetic algorithms, in general, is that they cannot preserve population variety [30]. The non-dominated sorting genetic algorithm was developed in 1994 to address this concern (NSGA). However, its complexity the optimization purposes are not sufficiently improved compared to GA. Therefore, the elitist non-dominated sorting genetic algorithm (NSGA-II) was proposed. The NSGA-II algorithm retains the superior individuals from the parent generation and children. Allowing for additional non-dominated solutions [6].

Drawing parallels to Darwin's evolutionary theory, genetic algorithms rely on three fundamental concepts: crossover, mutation and selection.

A genetic algorithm may be considered to include the steps illustrated in the figure below:

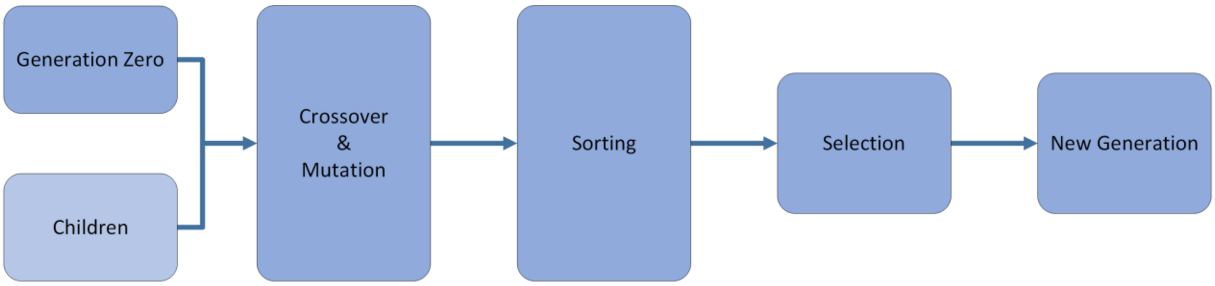


Figure 3.20 Genetic Algorithm.

Generation zero consists of a random number of solutions with a defined upper and lower boundary for each variable. Generation zero then goes through selection based on Darwin's three operators: selection, mutation and cross-overs. Better solutions are chosen from the mating pool for the next generation [54]. There are several techniques in which this may be achieved, for example, roulette wheel selection, rank selection and tournament selection [54].

Tournament selection, as the name suggests, is the concept of selecting a random number of individuals from a population and letting them compete with each other. Generally, there are two against two, resulting in the fittest being able to continue the process and being selected for crossover and mutation [54]. This selection process filters how many individuals remain before crossovers to the next generation and may be considered the selection process, "winners" are assigned a higher rank.

We separate between types of crossovers and the crossover rate. The crossover rate reflects how many genes are transported to the next generation, whilst the crossover type reflects how the children are created. There are several cross-over types: single-point, double-point, and intermediate. A brief illustration of these types is illustrated below. Parents are displayed on the left, whilst children are displayed on the right:

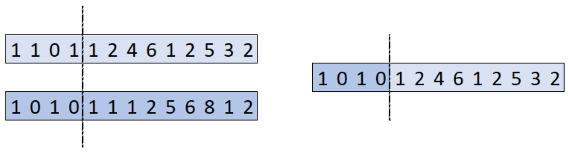


Figure 3.22 Single-Point Crossover.

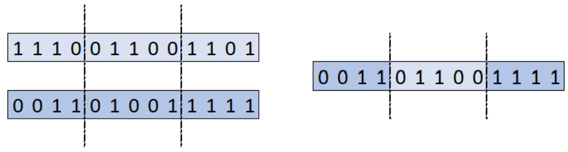


Figure 3.21 Double-Point Crossover.

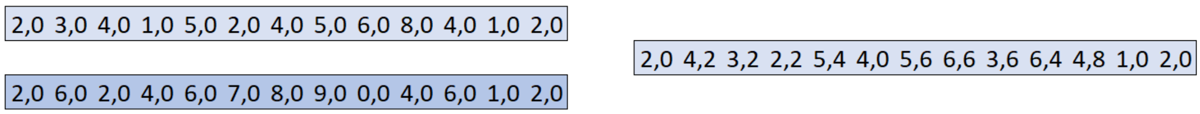


Figure 3.23 Intermediate Crossover.

As may be interpreted from the figures, the single-point crossover is performed by replacing values from a single point to the new generation. Double-point crossover is similar, only with two points. Intermediate Crossover is more complicated, as it is calculated by the formula displayed below:

$$Child = Parent_1 + Rand * Ratio(Parent_2 - Parent_1)$$

Equation 3.10 Intermediate Crossover

The child of this crossover is performed by taking the value of parent one and adding a random number between 0 and 1, multiplied by a ratio (usually 1) times the value difference of parent two with parent one. In a practical context, the different values of the decision variables serve as genes (chromosomes).

In the crossover process, mutation also occurs, which serves as a means to create new solutions by creating genes not included in parent generations and passing them on to the children.

After the crossover and mutation process combined with the previous generation and superior individuals from the parent generation, we have additional solutions. These solutions go through a sorting process, where NSGA-II separates itself from other forms of genetic algorithms. Based on two different criteria, crowding distance and rank. Consider the figure below:

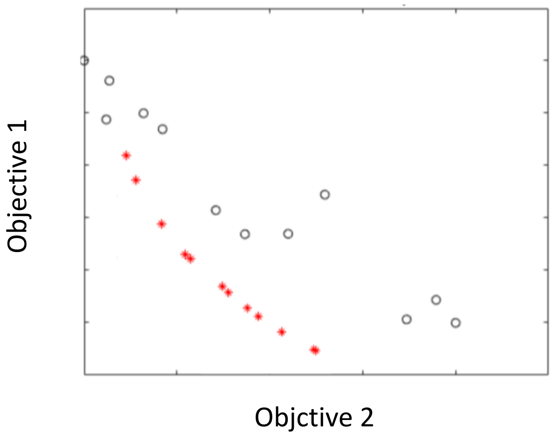


Figure 3.24 Pareto front.

The point closest to the intersecting axis (point 0,0) reflects the lowest value of objective one and objective two. This point is thus more optimal compared to the other points. Such a point is automatically selected by rank. The closer to optimal, the higher the rank. In the figure above, one may consider the red points closer to the intersecting axis than the circles. Thus, the red points are selected.

Secondly, in the cases where two values with the same rank (solutions marked with red). The selection of the winner is based on the relative crowding distance. This process is repeated per generation. As the selection process is completed, a higher degree of optimization will occur each generation [12].

The algorithm repeats itself until specific criteria are met. Examples include a max number of generations, convergence, time limit, or spread in Pareto solutions.

Genetic algorithms may be considered an iterative tool for evolving solutions into new and better solutions. This process is repeated per generation until the least adaptive individuals are eliminated [54].

This process is not without challenges, however. According to a study published by An, Li et al., uncertainties connected to genetic algorithms are divided into five subcategories [55]:

- Unclear on selection representations of chromosome operators, as the selection significantly influences the efficiency.
- Varying length with population size, small population size may cause the genetic algorithms to converge towards a sub-optimal solution, causing a situation where it is required with larger population sizes, but this again causes more computing efforts.
- The selection operators implicitly entail the probability that chromosomes of the latter generations have better fitness. This may cause loss of valuable information with chromosomes with lower fitness, lowering diversity and possibly converging prematurely. Furthermore, the choice of selector operator rests on user subjectivity.

- Possible derivatives of the first operation, the derivation of chromosomes, depending on the cross-over operator. This operator may influence the results. Several cross-over options exist, like geometrical, arithmetic, random, single-point, double-point and intermediate cross-over.

- Unimplicated determination principles for mutation operators. The determination of mutation will have a significant influence on chromosome health. A lower mutation may cause premature convergence, whilst a higher mutation may result in the probability of destroying good chromosomes.

4 Research Question

The research question formulated for this thesis is:

How can digital twins be implemented to reduce energy consumption and costs in buildings?

Answered by the following sub-questions:

- How effective are ANN models in predicting building energy performance indicators (e.g., U-value/kWh, U-value/costs), and what are the practical implications of using these models in terms of energy efficiency and cost savings for school buildings?

- How to use NSGA-II to reduce energy consumption and costs?

- How to use visual programming to stream results to BIM?

5 Case

This chapter will describe and explain the case for this thesis, the starting point from which to answer the research question and the problem area. It is built on a digital model of a physical asset for simulation purposes, serving as the digital twin.

5.1 Tvedestrand Videregående Skole

The case chosen for this study is Tvedestrand Videregående Skole, a high school in Tvedestrand, Norway. The building was erected in 2020. There was a clear focus on reducing carbon emissions in the design and construction phases [56]. It holds room for 690 students, 420 employees and has a gross area of 14 500 m^2 [57].

Figure 5.1 Tvedestrand Videregående Skole - Entrance.

The school was built to achieve the futurebuilt plus-house certification. It is worth noting that the definition and methodology for plus-house per futurebuilt was updated by Version 2.2 19.09.22, published after the construction of Tvedestrand Videregående Skole was completed.

Figure 5.2 Tvedestrand Videregående Skole - Entrance Area.

There were numerous efforts put in place to achieve the certification. For instance, of the 8 420 m^2 of the gross roof area, 4050 m^2 is covered by solar panels [58]. Construction of the asphalt road achieved a 37% reduced carbon footprint than its concrete alternative, and concrete stairways were replaced with wooden alternatives. In total, the building achieved a 78 % carbon footprint emission reduction compared to its reference building. Making it one of the most environmental-friendly schools in Norway [56].

5.2 Classroom 3006

The case for this thesis is limited to classroom 3006, located on the third floor of the building. The location is illustrated by the two figures below, which are gathered from the floorplan available in the Revit Model included in *Appendix A – TVGS.rvt*

Figure 5.3 Tvedestrand Videregående Skole - Room 3006 Floorplan.

The pictures that follow display the classroom. Firstly, its entrance and following perspectives are noted by symbols.

Figure 5.4 Tvedestrand Videregående Skole - Room 3006 Entrance.

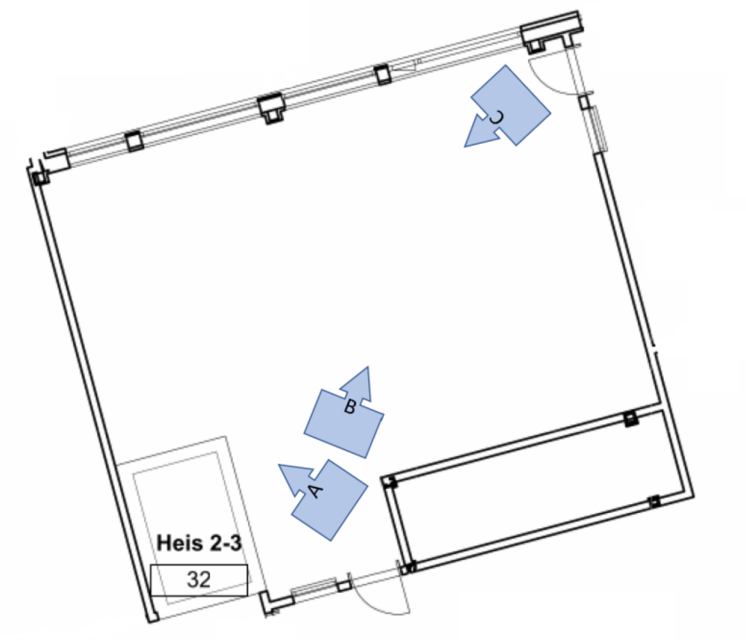


Figure 5.5 Perspective Overview.

Figure 5.6 Room 3006 - Perspective A.


Figure 5.7 Room 3006 - Perspective B.


Figure 5.8 Room 3006 - Perspective C.

The particular case was chosen on the following basis:

- Basic walls covering three sides. The final side is an external wall.
- Concrete bearing for the elevator shaft is connected to the room's southwest corner.
- Two exit doors with windows attached.
- Windows cover a significant fraction of the external wall.

A room containing these different construction elements allows for a more extensive range of calculations: Standard internal walls, external walls and a concrete-bearing elevator shaft. Once including doors, roof, floor and windows, we have six material-specific variables that may be analysed and optimized.

6 Methodology

This thesis is designed to include new technological advancements within digitalization and optimisation in AEC. It seeks to cover energy performance calculations, machine-learning models, multi-objective optimization and digital twin. As such, it covers a wide range of software solutions. In this chapter, the strategy and choice of methodology will follow. Firstly, the literature review part of this thesis intends to cover the necessary theoretical background and identify relevant case studies presented in the results. The formatting process of this thesis will follow, and finally, a description and workflow of the various software solutions that are applied in answering the research question.

6.1 Literature review

This chapter will cover the methodology for collecting the literature for this thesis. The literature review is performed to establish a sufficient theoretical background and case studies related to the research question. This thesis selects a two-pronged approach consisting of two different search strategies, a *search module (*or *module search)* and a *supporting search*.

A search module (or module search) intends to cover the theme in broader terms. The first ten articles of a literature search were all selected and then abstract-reviewed for relevance. Detected keywords not connected to the AEC industry caused the article to be discarded. Geographically limited case studies were ignored, leaving a large part of bibliometric studies. These articles were all full-text reviewed and form the basis of the different fields of research presented in this thesis.

The primary motivation behind a search module is to find several sources for each respective term or theme, as this ensures a higher degree of academic accuracy. The chosen methodology allows for more accurate coverage of the theme in question. Furthermore, the different themes occasionally overlap each other. For example, Artificial Intelligence in AEC and Machine Learning in AEC cover the same theme. The hits from both searches may help identify overlapping or relevant articles not initially detected in a singular search module.

A supporting search is a specific search intended to expand on relevant terms identified in the search module, such as the definition of artificial intelligence and the digital revolution. Articles were selected within the first ten hits of each search based on keywords in the heading. Followingly, articles were reviewed for keywords connected to the search term. Selected articles were not full-text reviewed. Instead, relevant information was extracted from the introduction, results or conclusion.

The relevant case studies were collected through these searches. In general, one could argue that a supporting search holds lower academic value than a module search due to its higher degree of specification and limited selection of articles. Furthermore, support searches do not require a full-text review to be applicable. A simplified illustration of the literature review methodology is illustrated below:

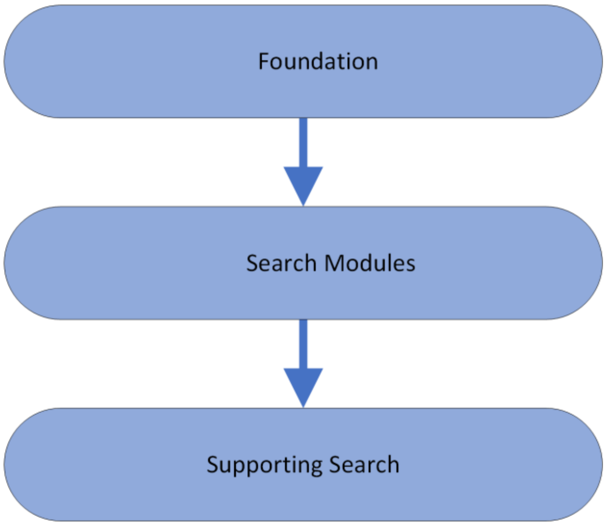


Figure 6.1 Litterature Review, A two pronged approach.

This thesis is intended to be an expansion of previous work within the field of using machine learning in combination with genetic algorithms for optimization purposes. Therefore, the thesis holds three articles as a foundation. These articles identified keywords relevant to the research question, serving as terms for module searches. Supporting searches were placed on specific key terms and themes, such as definitions like IoT, ANN, NSGA, and Genetic Algorithms.

The three initial articles serving as a foundation are based on previous work published by Hosamo et al. Namely, *"Digital Twin of HVAC System for multi-objective optimization of energy consumption and thermal comfort based on BIM framework with ANN-MOGA"* by Hosamo, Nielsen et al. [30]. *"Multi-objective optimization of building energy consumption and thermal comfort based on integrated BIM framework with machine learning – NSGAA II"* by Hosamo, Tingstveit et al. [6]. *"Digital Twin Framework for automated fault source detection and prediction for comfort performance evaluation of existing non-residential Norwegian buildings"* by Hosamo, Nielsen et al. [59].

A fourth article was selected as the foundation of the other aspects of this thesis, cost analysis, namely "*Minimizing delivered energy and life cycle cost using Graphical script: An office building retrofitting case"* by Rabani, Madessa et al. [10]. An illustration depicting the module searches and the amount of articles extracted that were full-text reviewed from each search module is depicted in the following figure:

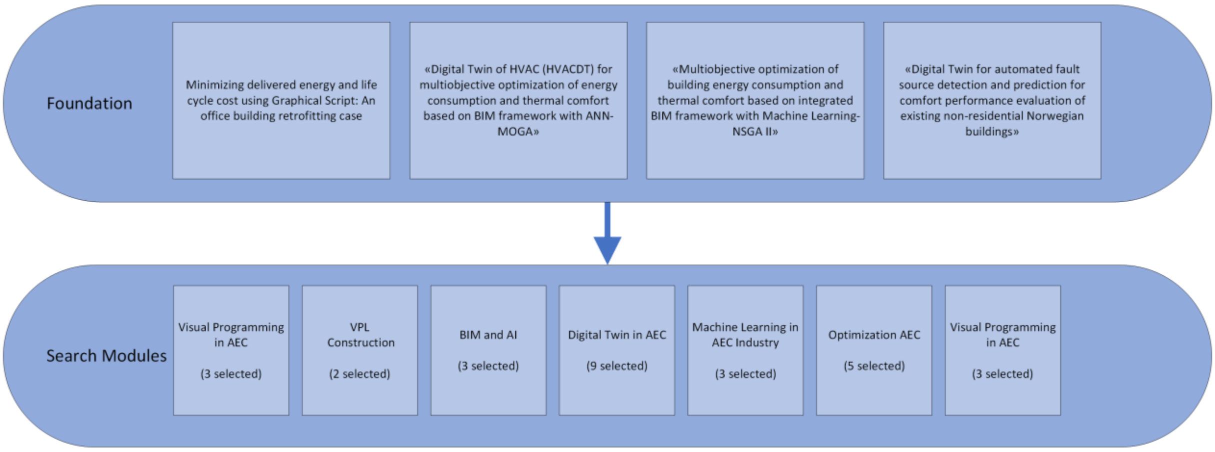


Figure 6.2 Literature Review Methodology.

Criteria for the selection of articles were defined as follows:

- Available in the Oria Database.
- Published between 2012 and 2023.
- Written in English.
- Peer-reviewed articles.
- Within the first ten hits from the search results.

Criteria for rejection were defined as follows:

- Articles containing keywords not associated with civil engineering.
- Geographically limited case studies, this holds true for all search modules but not supporting searches.

The choice of using the Oria Database in favour of other search engines, such as Google Scholar, was made for two reasons. Firstly: Functionality, using the Oria database, one can save searches for use in the search log, find relevant information for citations and also supply an intuitive user interface to exclude terms from searches. Secondly: Validity, using the Oria database, enables a direct search to interpret and choose relevant articles swiftly and accurately. The search engine is also precise to the point where one can add terms to limit search hits more efficiently, ensuring the search results are relevant for their intended purpose. Although Google Scholar arguably has the largest number of articles, it may be considered inappropriate as an academic search engine because it cannot provide sufficient replicability [21].

A detailed search log with keywords, search strategy (module searches/supporting searches) and motivation is included in this thesis in *Appendix B – Logbook for literature search.pdf*

As of the time of writing this thesis, machine learning and numerical optimization using genetic algorithms is not an integrated part of the Msc course in civil engineering at the University of Agder. In an attempt to overcome this weakness, the chapter on Machine Learning and Genetic Algorithms leans heavily on external sources and courses from other universities.

For Machine Learning, the lecture material from the class "*Introduction to Artificial Intelligence and Machine Learning for Engineers*" was taught at UCLA (University of California, Los Angeles) by Mathieu Bauchy in 2018 [34]. The "Learning from Data – Introductory Machine Learning Course" also from UCLA, presented by Yaser Abu-Mostafa in 2012 [60]. The course *"CS229 -Machine Learning"* from Stanford presented by Andrew Ng in 2018 [32]. And finally, the course *"Introduction to Deep Learning"* by the Technical University of Munich (TUM) in 2020 presented by Matthias Niessner [61].

Whereas the chapter on Genetic Algorithms is supported by the course *"Selected Topics in Decision Modelling"* by Biswajit Mahanty, taught at the Indian Institute of Technology in 2018 [62]. Followed by the "Practical Genetic Algorithms in Python and Matlab", presented by S. Mostafa Kalami Heris, phd for the Yarpiz Project [63]. These courses are available online and may be identified using a search engine.

In the lack of supporting literature for these courses, popular sciences were also applied for relevant terms identified in the courses or articles relevant to the theme. These were applied for illustrations and figures, while the respective courses covered the theoretical background. The identification of these sources is not included in the logbook.

"Grey sources", such as the relevant standards, were identified through keywords in the respective articles. They are also not represented in the search log, as they are clearly referenced in the relevant chapters.

6.2 Formatting

Unless specified otherwise, the figures and tables in this thesis are created by the author. Specific figures that strongly resemble the original source are referenced in the figure text. Most figures for illustration purposes are created using Microsoft Visio [64]. Pictures of the chosen case were gathered using the camera of an iPhone SE. For illustrations or pictures needing editing, GIMP version 2.10.32 was applied [65]. The necessary proficiency for this software was already obtained by the start of this thesis.

6.3 Strategy and Software Hierarchy

The overall strategy for this thesis is to include different digitalization tools to illustrate different functionality and application areas. The goal is to implement these tools for building performance optimization and stream the results in a digital twin. The process and software hierarchy is illustrated in the figure that follows:

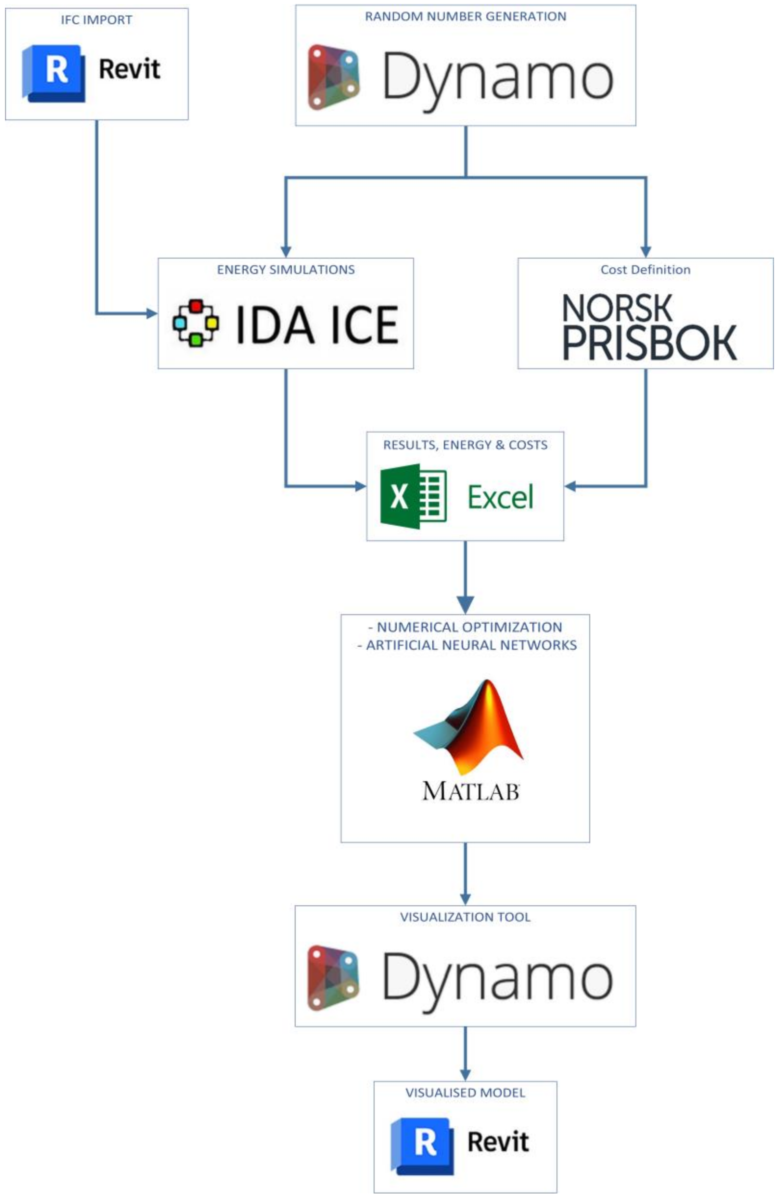


Figure 6.3 Software Hierarchy.

At the time of writing this thesis, there was only a limited amount of knowledge of Revit, Excel, Matlab and Dynamo. There was no previous experience with the application of IDA ICE. At all levels, it was necessary with third-party sources to gain the experience and knowledge required to achieve sufficient proficiency within the various software. Sources that were applied were: The IDA ICE getting started document available at Equa.se [66]. Tutorial videos are available at LinkedIn Learning [67]. For Matlab, online Machine Learning Courses were applied, in particular, "*Introduction to Artificial Intelligence and Machine Learning for Engineers*" by Mathieu Bauchy in 2018 [34] and tutorials and machine learning descriptions are available at the Mathworks website [35]. For code troubleshooting during the scripting process, the AI bot ChatGPT was applied [39].

The simulations are based on three available decision variables from the chosen case: Internal Walls, External Walls and Windows. The motivation for this choice was made in order to quantify as much of the room as possible. Other categories and variables were excluded due to challenges in determining a suitable methodology and challenges related to operating the IDA ICE software.

6.4 Visual Programming

Visual programs in this thesis are created in Dynamo and applied in two main areas, one for generating random variables serving as inputs for the IDA ICE simulation software and cost estimation, in addition to serving as a visualization tool for Revit, the digital twin.

6.4.1 Generation of Random Variables

For this thesis, the input for the machine learning model was the analysis of 100 combinations of randomly generated numbers based on inputs created in Dynamo. This number was chosen to establish a sufficient dataset for adequate accuracy for the neural networks, which will be described in later chapters.

The code is displayed below:

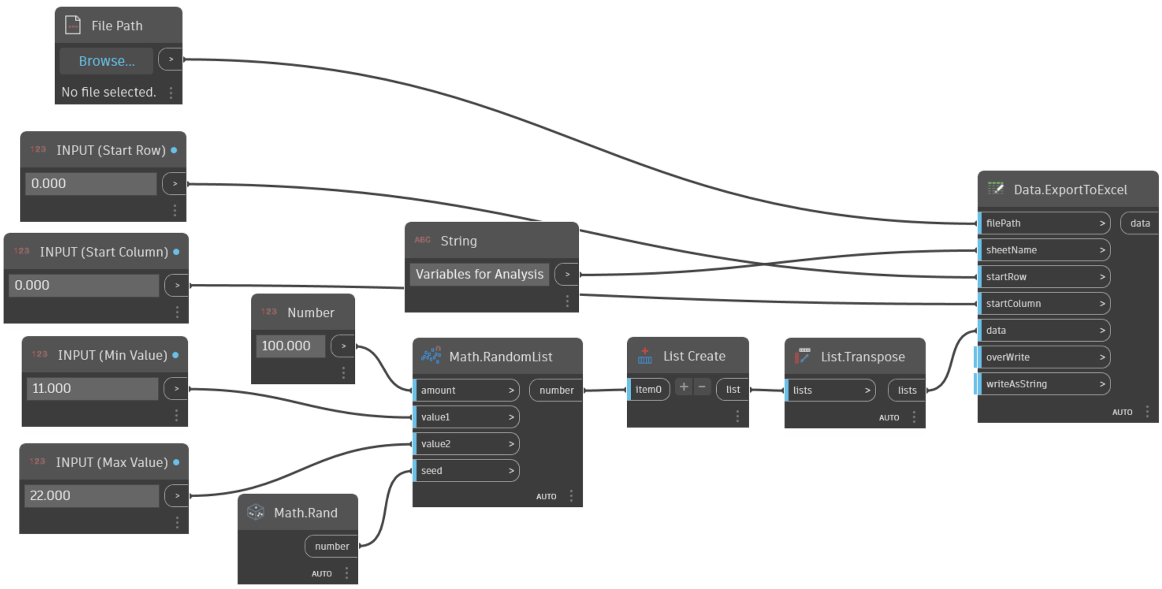


Figure 6.4 Random Number Generator.

The program may be explained as follows; it rests on five inputs. INPUT (*Min Value)* and INPUT *(Max Value)* reflect the domain of the numbers we wish to generate. INPUT (*Start Column)* and INPUT *(Start row)* refer to the Excel sheet for which the numbers are exported and the *File Path* for the

Excel sheet file. It is worth noting that visual programming works differently than the regular numbering we are usually accustomed to. The initial value for this form of programming is 0 and not 1. So start Column 0, and Start Row 0, translates to A1 in the Excel sheet.

Number as an input to *Math.Randomlist* describes the number of times the *seed* is repeated. The input for *seed* is the specific function, in this case, *Math.Rand,* which creates a random value between 0 and 1. The *String* is located above the *Math.Randomlist* function creates a name for the Excel sheet from which the data is to be exported.

The output from *Math.Randomlist* creates a list which is initially horizontally exported to Excel. For convenience, this is flipped to a vertical list through the *List.Transpose* function. It is worth stressing the importance of booleans in this case. *OverWrite* is set to false by default. Should this value be changed to true, it would cause the list to be randomized each time the script is applied, which is set to automatic by default. The completed visual program is included in *Appendix C – Random Number Generator.dyn*.

The function is repeated for each variable we wish to include in future simulations. This thesis is based on three categories illustrated in the table below:

Table 6.1 Range of thermal transmittance for each category.

Category	Min U-value [W/(m²K)]	Max U-value [W/(m²K)]
External Wall	0,12	0,22
Window	0,7	1,6
Internal Wall	0,22	0,42

The decision variables' range is based on technical requirements in the Norwegian Byggteknisk Forskrift 10, §14-2 "krav til energieffektivitet" [48]. And the materialist available at Norsk Prisbok [68].

Internal walls are not covered in heat loss calculation, as there is presumed no heat transfer between two heated zones. As such, the original literature does not cover the range of values for internal walls. The values of internal walls were calculated from a lower limit, set as a composition satisfying the minimal soundproofing requirement of a school building in Norway of 48 dB noise reduction, which composition provided in the 524.325 manual in Byggforskserien [69], to a construction with a maximum width of 200 mm.

The composition of the two alternatives is displayed in the table below. The total value for the thermal transmittance of each alternative is summarized in the final row. These values were calculated using IDA ICE, using the predefined material constants in the material library.

Table 6.2 Alternative Compositions of Internal Wall.

Alternative 1			Alternative 2		
Material	Thickness [mm]	λ [W/mK]	Material	Thickness [mm]	λ [W/mK]
Gypsum (x2)	25	0,22	Gypsum (x2)	25	0,22
Insulation	150	0,036	Insulation	70	0,036
Gypsum (x2)	25	0,22	Gypsum (x2)	25	0,22
U-Value[W/(m²K]	0,22		U-Value[W/(m²K]	0,42	

The list of randomly generated numbers is included in *Appendix D – Datasheet.*xlsx under the "*Variables for analysis*" sheet.

6.4.2 Visualization of Digital Twin

The process of creating a visual program for Digital Twin illustrative purposes in Autodesk Revit is separated into several different steps:

1. Create a digital model of the case room in Revit.
2. Define each component as a solid in Dynamo.
3. Export these components from Dynamo to display in Revit.
4. Import optimization degree from Excel.
5. Visualize components with a colour range depicting the degree of optimization.

The visual program to achieve the digital twin rests on several different inputs. Each input goes through midway processes and modifications. As such, it may be considered complex. To cover all levels of the process in writing is deemed impractical. Instead, a rough illustration of the program is displayed below. Further information about the program may be identified from the included file in *Appendix E – Digital Twin Script.dyn*

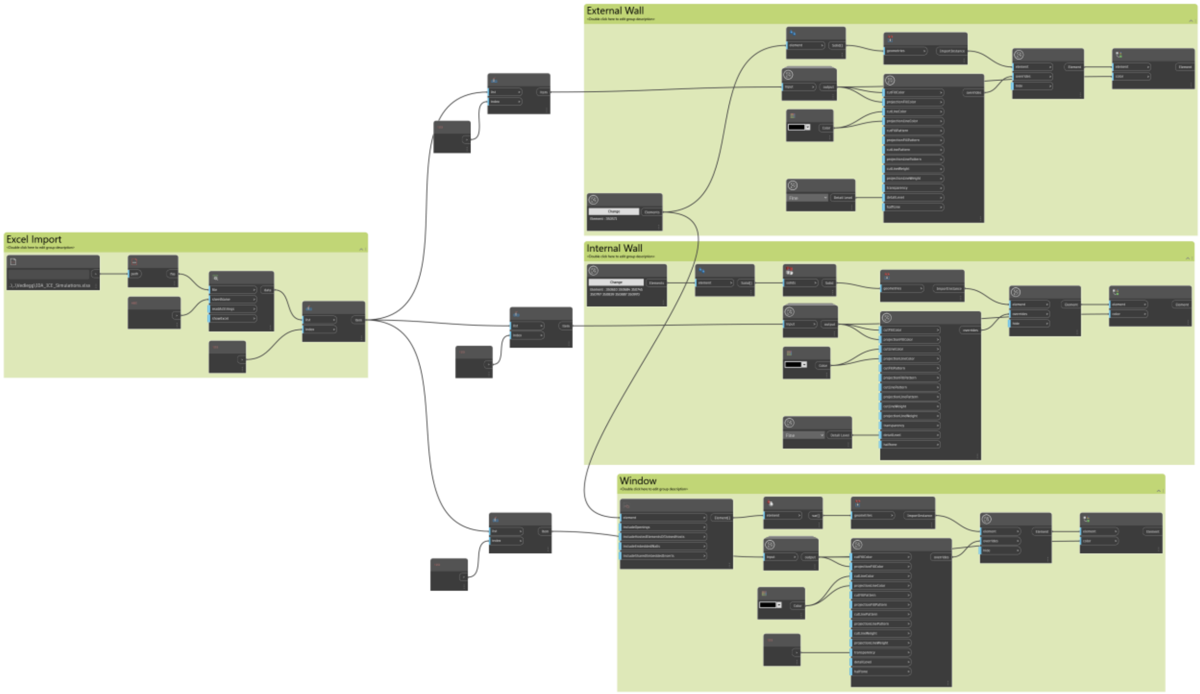


Figure 6.5 Visual Program - Visualization.

In order to avoid repetitive tasks, a custom node for optimization degree was created. The functionality of this node is to return a colour between green and red reflective of the optimization degree in the range between 0 and 1, where 1 illustrates red (large degree of change), and 0 illustrates green (low degree of change). An illustration of the node is shown in the figure below. The node is also exported to a separate file in *Appendix F – Coloring.dyf*. The finalized model is included in *Appendix G – Digital Twin.rvt*

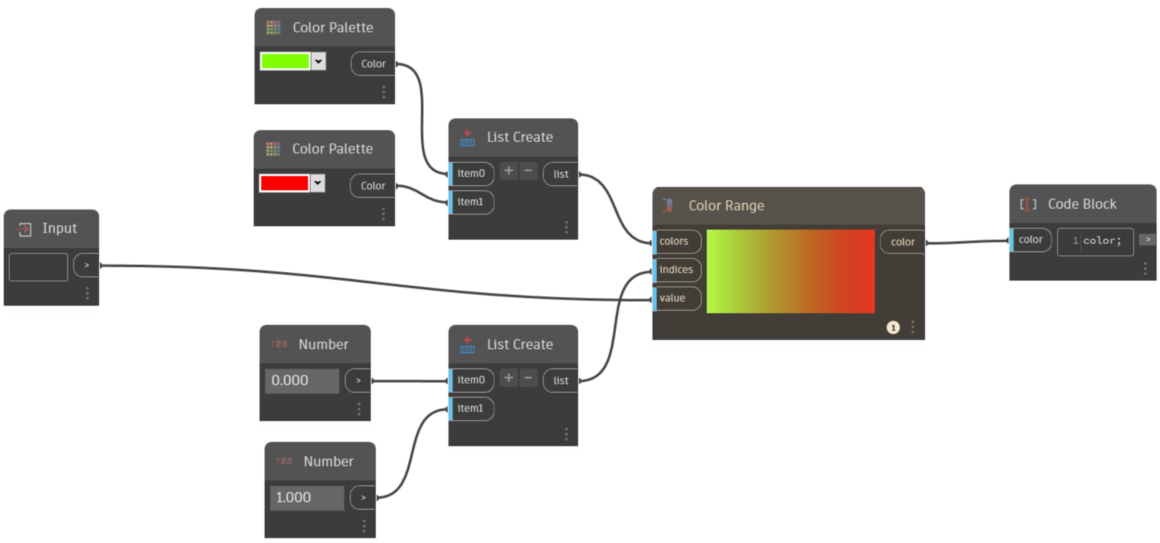


Figure 6.6 Custom Dynamo Node - Degree of Optimization.

6.5 IDA ICE

IDA ICE is applied in this thesis to establish values for energy consumption by running simulations based on variables established with a random number generator. Three decision variables are chosen for simulation purposes Internal Walls, External Walls, and Windows.

The process of using IDA ICE in this thesis may be described in the figure below. The various steps will be covered in later chapters.

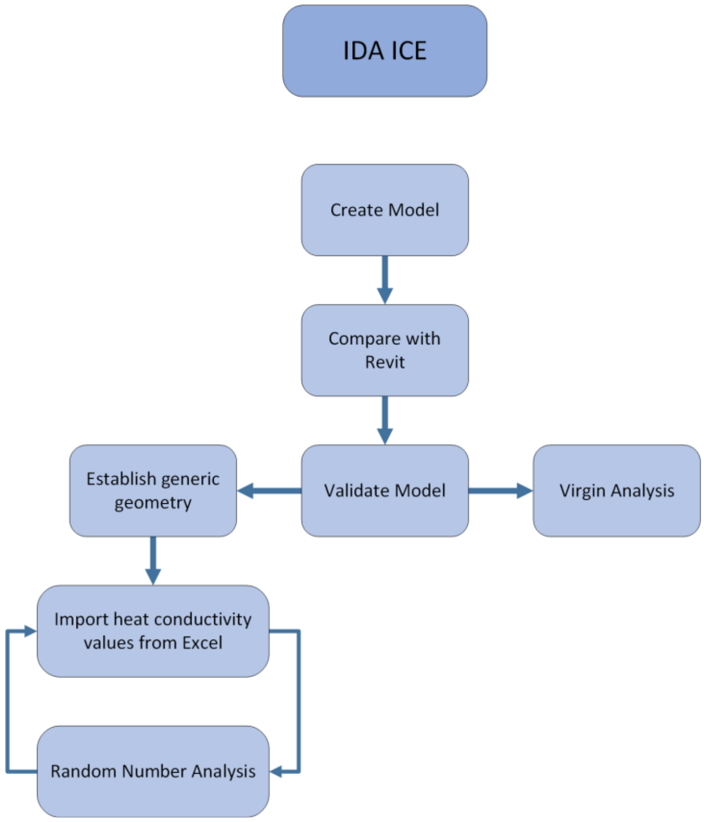


Figure 6.7 IDA ICE Modelling Process.

In the context of this thesis, IDA ICE serves as the baseline for simulation purposes to determine energy consumption. The simulations here are based on the particular case, with randomly generated thermal transmittance values created by visual programming.

The model validation process, adjustments and simulations conducted in this thesis is self-taught. At the start of this thesis, there was no previous knowledge of the use of IDA ICE. However, online support material proved helpful, particularly the "IDA ICE 4: Getting Started" document available at the Equa.se website [66]. The relevant simulation result, in this case, is the value of kWh/m^2. The relevant theoretical background is covered in Chapter 3.6.

6.5.1 Model Validation

All simulations in IDA ICE are done on an IDA ICE project file with the file extension IDM. For future reference, IDM will be applied to describing a particular file. Tvedestrand Videregående Skole has undergone a similar simulation in the past. The simulations were conducted by erichsenhorgen by commission of Aust-Agder Fylkeskommune. An IDM file with a readily imported IFC model and attached climate data was provided. Project-specific data, such as vacation hours or material compositions, was entered from the report by Erichsenhorgen. This report is not open for republishing. As such, it will not be covered in depth.

IFC and Revit files are described interchangeably in this thesis. One may consider the following: The software Revit applied the IFC file for IDA ICE modelling purposes.

In order to perform a simulation in IDA ICE, one needs to establish a "zone", which describes the area one wishes to simulate. By using the default function in IDA ICE for the definition of the zone for classroom 3006, it was evident that the model was not sufficiently accurate, as displayed in the figure that follows below:

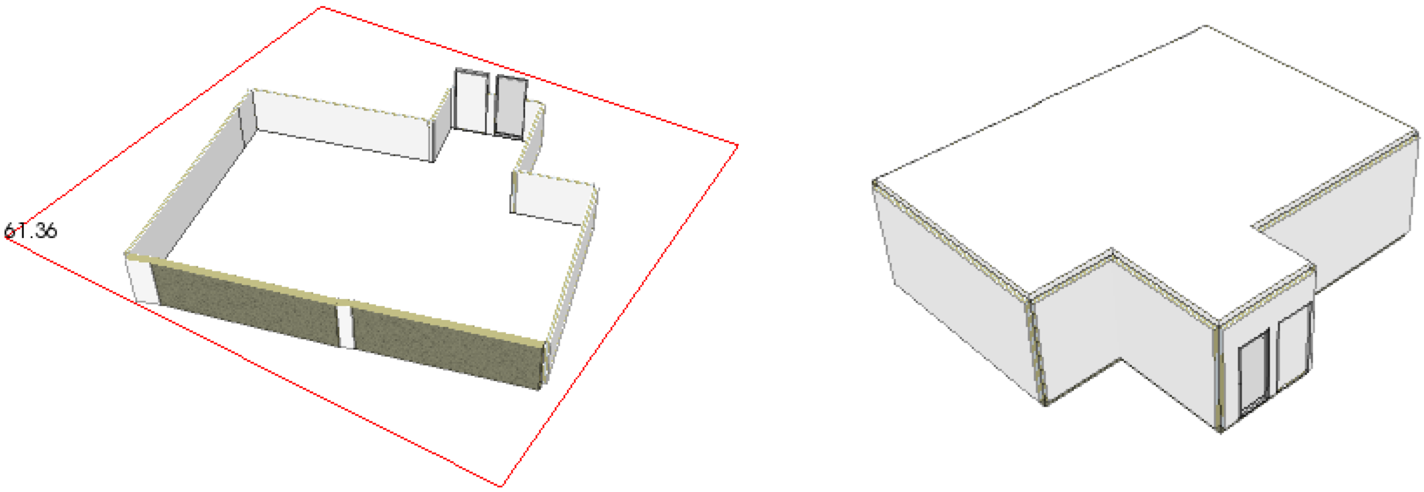


Figure 6.8 Original IDA ICE Model.

In preparation for the modelling process, an excursion to Tvedestrand Videregående Skole was conducted with permission granted by the acting principal. The excursion aimed to establish a sufficient background for the case chapter, gain illustrative pictures for this thesis and also aid in formulating a suitable methodology.

By comparing the generated model to visual observations, it was clear that essential parameters were missing from the model, particularly the southernmost windows, the north-eastern door, the eastern door, and the side-panel windows next to the doors.

In order to address this, the digital model was compared with the IFC file imported Revit. The geometry of this Revit file used as inputs are illustrated in the figures below:

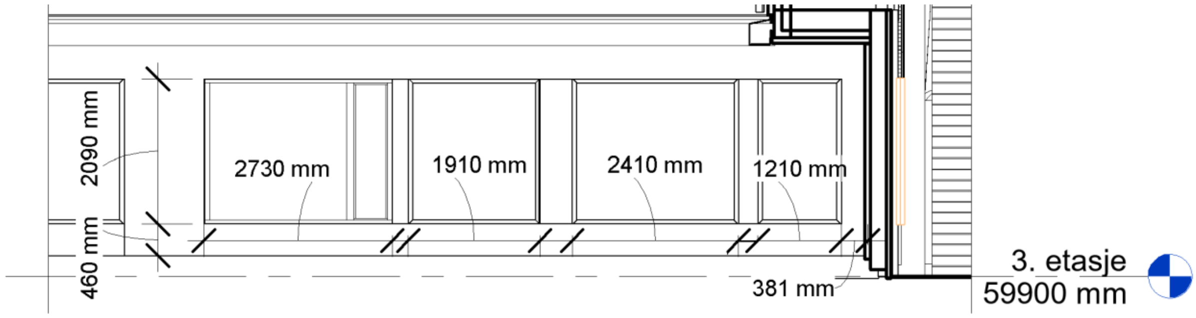


Figure 6.9 Window Geometry External wall.

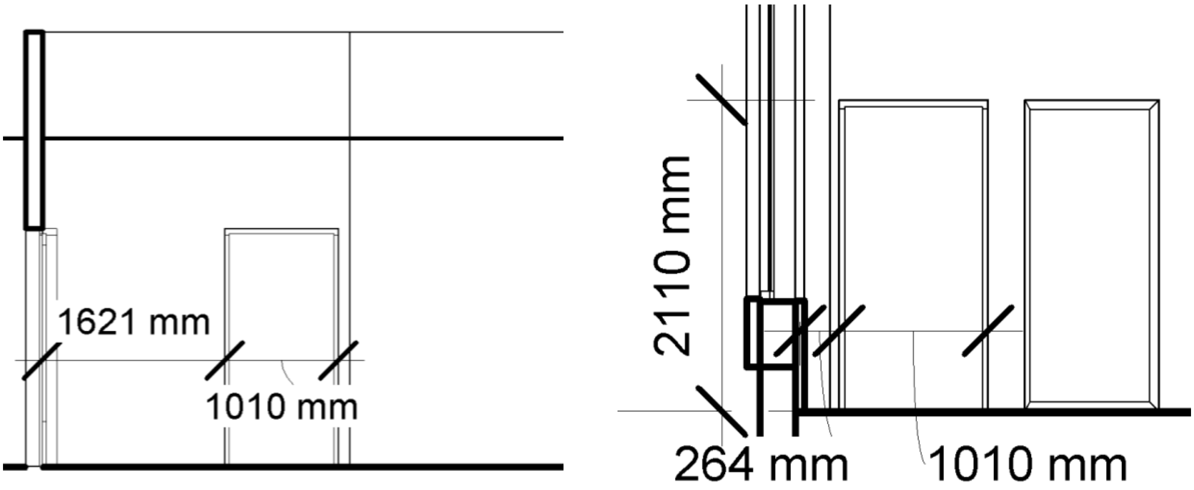


Figure 6.10 Door Geometry.

As described per methodology in the previous chapter. The simulation methodology does not include internal windows and doors. As such, the side panels for classroom 3006 were removed. Doors are considered the same as defined walls unless the zone is connected to another zone. As such, the doors remained on the model.

Furthermore, the mapping and material composition of windows and doors from the IFC file is not translated correctly to the IDA ICE model. They needed to be redefined by extracting information from the IFC file for the virgin analysis or by including material composition and material parameters included in the report by Erichsenhorgen.

By default, IDA ICE does not calculate energy performance to internal walls unless it is specified as an external zone. In the IDM file, the façade wall and roof are defined as external, which means, theoretically, no heat loss to the internal walls or the floor construction.

In the southern part of the room, visual observations identified an elevator shaft of pure concrete, highlighted in the figure below:

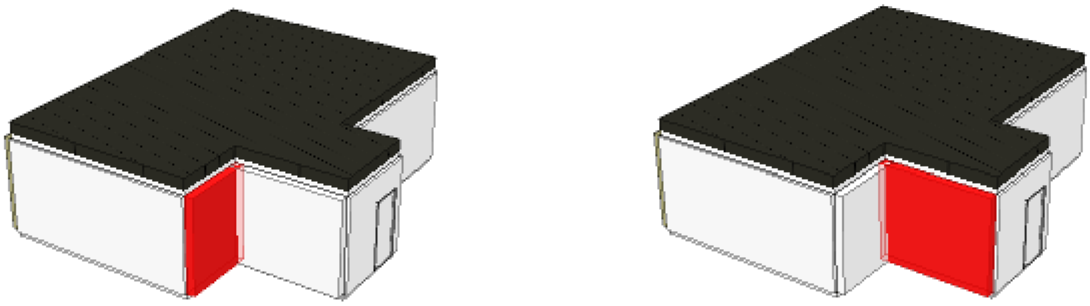


Figure 6.11 Elevator Shaft.

These walls are considered static and unalterable for optimization purposes. By extracting information from Revit, these walls were identified with a thickness of 0.25 m. By applying the concrete material already defined in the model, the following thermal transmittance value is defined:

Table 6.3 IDA ICE - Elevator shaft.

Elevator shaft	
Concrete	0,25 [m]
U-Value	3,154 [W/(m^{2}K)]

Once all data was exported and edited, the validated model is displayed in the following figure and is also included in *Appendix H – IDA ICE SIM.idm*

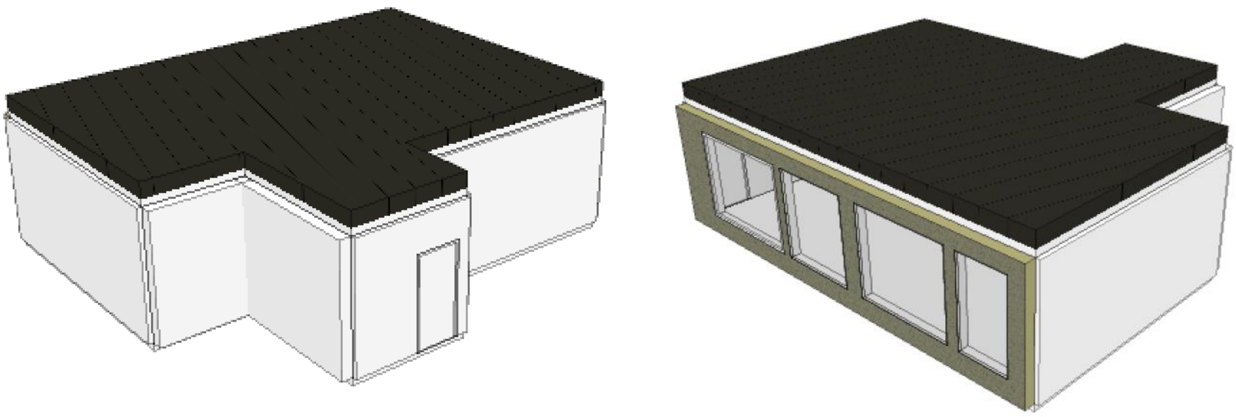


Figure 6.12 Validated IDA ICE model.

6.5.2 Presumptions

The IDA ICE software requires input on other variables for the respective room, mainly information about the occupants, lighting, HVAC system and usage patterns to perform an accurate analysis.

Since the case room chosen in this study is not included in the original IDM file, this data was not readily available. However, the previous simulations conducted by Erichsenhorgen also included a classroom. Data from this report was plotted in IDA ICE. Furthermore, material specifications for doors and windows were considered similar, except for thermal transmittance values for simulation purposes. Other data were calculated from the occupants and square meters ratio by provided presumptions included in the original simulation report by Erichsenhorgen. As described earlier, this report is not open for republishing. As such, it will not be covered other than the constants chosen for these calculations.

The tables below list the room specifications serving as data for simulations:

Table 6.4 IDA ICE - General Data Specifications.

General Data	
Climate	Kristiansand
Location	Kristiansand/Kjevik
Room occupancy	Monday-Friday 08:00-16:00
Internal Heating load	Laptop 30 W per occupant Digital board 200 W
Lighting Period	Monday-Friday 07:30-15:30
Digital Board Operating time	Monday-Friday 08:00-15:30
Equipment Operating time	Monday-Friday 08:00-15:30
Vacation Period	Weeks: 26-32 and 52

Table 6.5 IDA ICE - Window Specifications.

Window Specifications	
Solar Heat Gain Coefficient	0,34 [-]
Solar Transmittance	0,29 [-]
Visibility Transmittance	0,63 [-]
Internal Emissivity	0,837 [-]
External Emissivity	0,837 [-]
Fraction Frame/Window	0,10 [-]
Opening	None
Azimuth degree	48,38

Table 6.6 IDA ICE - Vacation Period.

Vacation Period	
26 June – 14 August	Summer Holidays
25 December – 1 January	Christmas Holidays

Table 6.7 IDA ICE - Air Handling Unit Specifications.

Air Handling Unit Specifications	
Central Air Handling Unit	Variable Air Volume, temperature and CO2 monitoring
Supply Air	20 [m^3/hm^2]
Set point Heat	21 °C
Set point Cooling	24 °C
External Shading	None
Opening of Windows	None
Loss factor for thermal bridges	3,3349 °C
Leak area -Central Air Heating Unit	0,00209 [m^2]
Operating Time	Monday-Friday 08:00-16:00

By using the presumptions described earlier in this chapter, the final room specifications plotted into IDA ICE are repeated in the table below:

Table 6.8 IDA ICE - Data input.

IDA ICE Data Input	
Room size	80,85 [m^2]
Occupants	30 [No.]
Ceiling Height	2,9 [m]
Occupants	0,371 [No./ m^2]
Supply Air	5,55 [l/s.m^2]
Lights	4 [W/m^2]
Digital Board	200 [W]
Laptop	30 [W/No.]

6.5.3 Virgin Analysis

The Virgin analysis is intended to form a basis for comparison for the model and verify approximate values from previous simulations in similar conditions. However, there is contradicting information when comparing the Revit IFC file and IDA ICE.

From the Revit model, the following semantic data were extracted:

Table 6.9 Revit - External Wall Composition.

External Wall composition	
Tiles	8 [mm]
Air	36 [mm]
Insulation	200 [mm]
Plasterboard	22 [mm]

By comparing this value to the default component composition already defined in IDA ICE, we have:

Table 6.10 IDA ICE - External Wall Composition.

YV-Bindingsverkvegg	
Render	0,01 [m]
Light Insulation	0,23 [m]
Gypsum	0,01 [m]
U-Value	0,1511 [W/(m^{2}K)]

There are discrepancies between the Revit Model and the IDA ICE model information.
Although one could argue that the Revit file contains more detailed information than IDA ICE, there is a challenge in using this data for simulation purposes. Namely, the lack of material information, specifically values such as density and thermal conductivity, are missing. Using the Revit model data is therefore considered impractical, and the data in the IDM material library was applied and considered valid. The values are displayed in the following tables:

Table 6.11 IDA ICE - Internal Wall Composition.

Internal Wall composition	
Gypsum	0,026 [m]
Air in vert.air gap	0,032 [m]
Light Insulation	0,030 [m]
Air in vert.air gap	0,032 [m]
Gypsum	0,026 [m]
Total U-Value	0,6187 [W/(m^{2}K)]

Table 6.12 IDA ICE - Window specification.

Window Specifications	
Solar Heat Gain Coefficient	0,34 [-]
Solar Transmittance	0,29 [-]
Visibility Transmittance	0,63 [-]
Internal Emissivity	0,87 [-]

External Emissivity	0,87 [-]
Fraction Frame/Window	0,10 [-]
Opening	None
Azimuth degree	48,38 °
U-Value Frame	2,0 [W/(m^{2}K)]
U-Value Glazing	0,6 [W/(m^{2}K)]
U-Value Total	0,74 [W/(m^{2}K)]

The windows are not all fixed and include the opportunity to open them. However, establishing an estimate of how much a window will be opened is difficult and prone to error. As such, it is presumed that the windows are permanently closed.

For the floor and the roof, the following definitions were applied:

Table 6.13 IDA ICE - Floor Composition.

Floor Composition	
Floor Coating	0,005 [m]
L/W Concrete	0,02 [m]
Concrete	0,15 [m]
Total U-Value	2,385 [W/(m^{2}K)]

Table 6.14 IDA ICE - Roof Composition.

Roof Composition	
Light Insulation	0,35 [m]
Concrete	0,15 [m]
Total U-Value	0,1002 [W/(m^{2}K)]

6.5.4 Random Number analysis

The Random number analysis is based on the generation of the random values for thermal transmittance for the respective categories. For the IDA ICE simulations, the process was simplified by defining a generic geometry. Each wall is defined as consisting of a homogeneous material with a thickness of 100 mm. By this presumption, one can input the variables as thermal conductivity values rather than the thermal transmittance by rewriting the equation described in Chapter 3.6. The equation is repeated below:

$$U_i = \frac{1}{R_{tot}}, \qquad R_{tot} = R_{si} + R_{se} + R_i\,, \qquad R_i = \frac{d_i}{\lambda_i}$$

By rewriting this formula, we have:

$$U_i = \frac{1}{R_{si} + R_{se} + {}^{d_i}/_{\lambda_i}}$$

Which equals

$$U_i \left(R_{si} + R_{se} + {}^{d_i}/_{\lambda_i} \right) = 1$$

$$R_{si} + R_{se} + {}^{d_i}/_{\lambda_i} = {}^{1}/_{U_i}$$

$${}^{d_i}/_{\lambda_i} = {}^{1}/_{U_i} - R_{si} - R_{se}$$

$$d_i = \lambda_i ({}^{1}/_{U_i} - R_{si} - R_{se})$$

Thus:

$$\lambda_i = \frac{d_i}{{}^{1}/_{U_i} - R_{si} - R_{se}}$$

As described in Chapter 3.6 Energy Performance, R_{si} and R_{se} Vary depending on airflow, respectively vertical for walls and horizontal for floors. If we define a single homogeneous layer with a thickness of $d_i = 0{,}1\ meters$ we get the following expression:

$$\lambda = \frac{0.1\ m}{{}^{1}/_{U} - \sum R}$$

By putting the various thermal transmittance values generated by the visual programming into this formula expressed in Excel, we end up with a single variable for thermal conductivity (λ) for each material which would equal the generated U-Value.

A new material with a new thermal conductivity value may be generated per variable for each combination of values entered in IDE ICE. This methodology is applied to internal and external walls. It is presumed that the values in Table 6.12 IDA ICE - Window specification are valid for simulation purposes. Beyond these values, the required inputs are the frame/window fraction, the frame u-value, and the glazing u-value. In order to simplify calculations, it is presumed that the frame covers 1

% of the total window area and has a U-value of 2 W/m^{2}K. The value of the frame will affect the total U-value of the window itself. However, by reducing the area to 1 %, it is presumed that this value does not significantly disrupt the results. As such, only the glazing U-value will be altered for each simulation. As this data may be directly input, no new window material is defined.

In IDA ICE, materials have other inputs than thermal conductivity, namely specific heat and density. For this thesis, these are considered equal to concrete. A density of 2300 kg/m^3, and a heat of 880 J/KgK, respectively. Although these values will affect the results, they will not disrupt the comparison for random number simulations as the values are constant. They may, however, disrupt the comparison.

IDA ICE can display the U-value for a specific combination of materials. Per the standard, which is described in Chapter 3.6, Energy Performance. The R-value should equal 0,14 for vertical airflow:

$$R_{si} + R_{se} = 0{,}10 + 0{,}04 = 0{,}14$$

However, when comparing the resulting U-Value from the input thermal conductivity, we have a discrepancy with this value. Through manual experimentation, the following R-value is defined.

$$R_{si} + R_{se} = 0{,}175$$

Calculations of this value and the degree of error are included in *Appendix D – Datasheet.*xlsx under the "*IDA ICE"* sheet. The error margin varies between 0,18 % to 1,4 % resulting in the assumption that the methodology is considered valid. Thus, we have the following equation for the input of thermal conductivity in IDA ICE from the randomly generated thermal transmittance value through visual programming:

$$\lambda = \frac{0.1\ m}{{}^{1}\!/_{U} - 0{,}175}$$

Equation 6.1 Thermal conductivity in IDA ICE.

This value is used for definitions of various materials. In the IDM file included in *Appendix H – IDA ICE SIM.idm.* The following system for the custom materials is defined:

Table 6.15 IDA ICE - Custom material library.

Material categorization	
x.IntW	Internal Wall
x.ExtW	External Wall

Where x reflects the combination number, as described earlier, the thermal transmittance value for windows is not included in this table, as there is no direct need for altering the composition to generic geometry. In these cases, the U-value of the glazing was altered directly.

For evaluating energy performance, it is considered valuable to calculate energy consumption per square meter per year. Values reflected in later parts of this thesis will operate with this value, although given the simulation lasts for one year, the unit displayed will be kWh/m^2.

6.5.5 Validation

The IDA ICE software is also applied for validation purposes. Similar to random number analysis, identified optima values are entered and simulated. This methodology will serve as a tool for validating the optimized results.

Further materials are entered with the following material definitions:

Table 6.16 NSGA-II Optima validation.

NSGA-II Optima Validation Simulations	
Optima.x.IntW	NSGA-II Optima, Internal Wall
Optima.x.ExtW	NSGA-II Optima, External Wall

6.6 Cost Analysis

The limitations of Life cycle costs in this thesis are defined as investment costs. It consists of material costs and installation costs. Using the "Life Cycle Cost Analysis" definition would be misleading. However, one may argue that the limitations seem appropriate as we assess the elements' *performance* over a short timeframe for comparison purposes with energy consumption and not in a life cycle timeframe. For this thesis, we define cost calculation as a cost analysis.

The cost analysis is performed in 7 steps, displayed in the figure below:

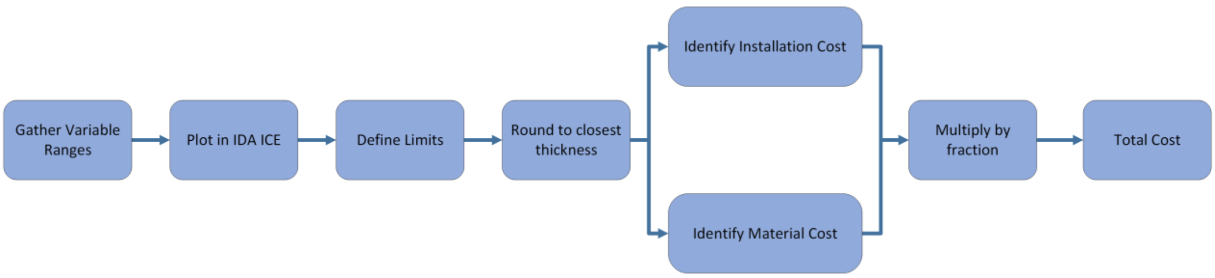


Figure 6.13 Cost Analysis Process.

Firstly: Identify materials corresponding to the U-value ranges defined by the randomly generated numbers. In this thesis, we limit ourselves to the effect of insulation. In real-life, energy efficiency is dependent on several other factors. Various thickness of insulation was gathered through the Norsk Prisbok website [68], corresponding to the respective element, External Wall (Translated to Norwegian: Klimavegg), insulation in internal walls (Translated to Norwegian: isolation i innervegg), and finally, windows. The values are identified under "prislinjer per fag", which describes the activity. In the case of insulation, this is identified by category 12 "Tømrerarbeider". In the case of windows, the values are depicted under category 14, "Vindusarbeider".

Secondly, in the insulation case, these materials were plotted in IDA ICE to calculate the U-Value for the combined element. We apply the element composition corresponding to the Virgin Model. An example of an Insulation External Wall is provided in the table below:

Table 6.17 Example, thermal transmittance - Insulation External Wall.

Name	U-Value [W/(m^{2}K)]
Isolasjon i klimavegg, mineralull, t=150 mm, 0,035 W/mK	0,2276
Isolasjon i klimavegg, mineralull, t=170 mm, 0,035 W/mK	0,2020
Isolasjon i klimavegg, mineralull, t=200 mm, 0,035 W/mK	0,1729
Isolasjon i klimavegg, mineralull, t=225 mm, 0,035 W/mK	0,1544
Isolasjon i klimavegg, mineralull, t=250 mm, 0,035 W/mK	0,1394
Isolasjon i klimavegg, mineralull, t=250 mm, 0,035 W/mK	0,1168

By following these simulations, definite, quantifiable materials are identified. Whereas windows have described U-values from the product itself. It is important to note that the walls are considered a homogenous insulation material per this methodology.

Thirdly: Identify limits. The domain for each material is defined by ranges of the mean value between two different materials and the actual value of the simulations.

$$Limit_i = \left(\frac{Sim.Material_i + Sim.Material_{i-1}}{2} \middle| \frac{Sim.Material_i + Sim.Material_{i+1}}{2}\right)$$

Equation 6.2 Calculation of thermal transmittance limits.

This methodology is applied to all construction elements. Ranges are included in *Appendix D – Datasheet.*xlsx under the "*Cost Definitions.*" sheet

Fourth step: Rounding materials to the closest limit, the randomly generated thermal transmittance value is rounded to the closest range depicting thickness (or product in the case of windows) identified in step 3. This process is repeated for each variable in all combinations. In order to speed up this process, a logic test script for Excel was designed.

Fifth step: The material cost is then extracted from the Norsk Prisbok, displayed under details for the material identified in step one. For this thesis, we limit ourselves to installation costs and material costs. As such, only these two are combined. Therefore, the displayed unit price may not be applied directly, as this also includes yearly costs.

The sixth step is: Multiply by a fraction. The goal is to identify costs per m^2 IFA. As such, each element (external Wall, window, and internal wall) are divided by its surface area compared to the indoor floor area. This remaining fraction "φ_i" Illustrates the fraction of the category quantity [m^2] compared with the indoor floor area of 80,85 m^2.

$$\varphi_i = \frac{Category_i[m^2]}{IFA\ [m^2]}$$

Equation 6.3 Element quantity per m^2.

The final step is defining the total cost. The total cost of the material may be described in the equation below:

$$Costs = \sum_i (I_{c,i} \cdot \varphi_i + M_{c,i} \cdot \varphi_i)$$

Equation 6.4 Total Costs.

Where:
I_c – Installation Cost
M_c – Material Cost
φ_i – Fraction of element area divided by gross floor area
i – Respective material, window or insulation in the case of external and internal walls.

In terms of the costs for the virgin model, the costs are calculated from the material compositions described in *6.5.2* Presumptions. The values are repeated below:

Table 6.18 Virgin Simulation U-Values.

Category	U-Value [W/(m^{2}K)]
External Wall	0,1511
Windows	0,74
Internal Walls	0,6187

The total dataset and calculations are included in *Appendix D– Datasheet.*xlsx under the "*Cost calculation"* sheet.

6.7 Machine Learning and Optimization

This thesis's machine learning models and optimization algorithms were created in Mathworks Matlab, version 2023b, available on the Mathworks Webpage [70].

At the start of this thesis, there was minimal experience with Matlab. The experience only included basic mathematic operations, addition, subtraction, multiplication, fractions and defining constants. There was no experience in the application of machine learning or general coding.

Creating machine learning models and genetic algorithm scripts in a foreign code language within the timeframe of this thesis is ambitious, especially when combined with the other foreign software applied in this thesis, IDA ICE and the lacking proficiency in Dynamo.

The goal of the scripting process was to create a sufficiently accurate model based on relevant sources, up to a termination process where the model was interpreted to be satisfactory. In order to create these scripts, a process consisting of 4 different layers was applied. The different layers are illustrated in the following figure:

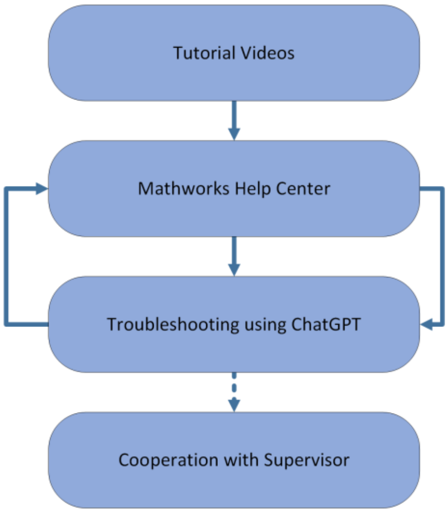


Figure 6.14 Script Creation Process.

The performance of the models was based on different criteria. For the case of the ANN models, a plot illustrating the different input variables and an r-value above 0,90. The optimization algorithm underwent a similar interpretation process. The input values were correctly identified and plotted, the various functions sufficiently identified, and the final results measured out of credibility within the expected range of optimization based on previous research.

Necessary sources for creating models were gathered through online sources, in particular, the course "*Introduction to Artificial Intelligence and Machine Learning for Engineers*" by Mathieu Bauchy [34] and the course *"Introduction to Deep Learning"* by Matthias Niessner from the technical university in Munich in 2020 [61]. General functionality for Matlab, such as file import, was gathered through the Mathworks online help center [35]. In the case of NSGA-II, sources were identified using Mathworks online help center, various videos available at the Learnwithpanda webpage [71], and troubleshooting using ChatGPT [39]. This process was iterative. The last resort was cooperation with the supervisor of this thesis.

The models will not be described in detail except for the main functions, which contain parameters introduced in Chapter 3.5 for the neural network models and Chapter 3.9 for NSGA-II. In the case of ANN models, this function is the *net.DivideParam* combined with the *train* function. Whilst in NSGA-II, *optimoptions* combined with *gamemultiobj* function.

6.7.1 Artificial Neural Network

The artificial neural network model was created by importing data from the sims from Excel into Matlab. The imported data is gathered in *Appendix D – Datasheet.*xlsx under the sheet "*Matlab_Import.*" In this context, the data used was the thermal transmittance, not the thermal conductivity.

The goal is to create two artificial neural networks based on the three separate inputs, the U-values for windows, external wall and internal wall, to one output per model. These two models will then be analysed in terms of r-value. One is defined through the IDA ICE simulations, the energy consumption of room 3006 in kWh/m^2, and the value from the cost analysis NOK/m^2 by values gathered by Norsk Prisbok and Excel functions.

The different models are included in *Appendix I – Matlab Code.mlx.* The two models are created with a single hidden layer due to the limited amount of dataset (100 combinations). As described in Chapter 3.5, a larger sample of data would most likely provide a more accurate result.

The code in Matlab was then created to interpret this data and create an ANN model. The code is included in *Appendix I – Matlab Code.mlx*. The process of creating the code is displayed in the figure below. For the purpose of this chapter, only the key elements are introduced. The main functions described earlier in this chapter and potential additional methodology applied for the creation of the final models.

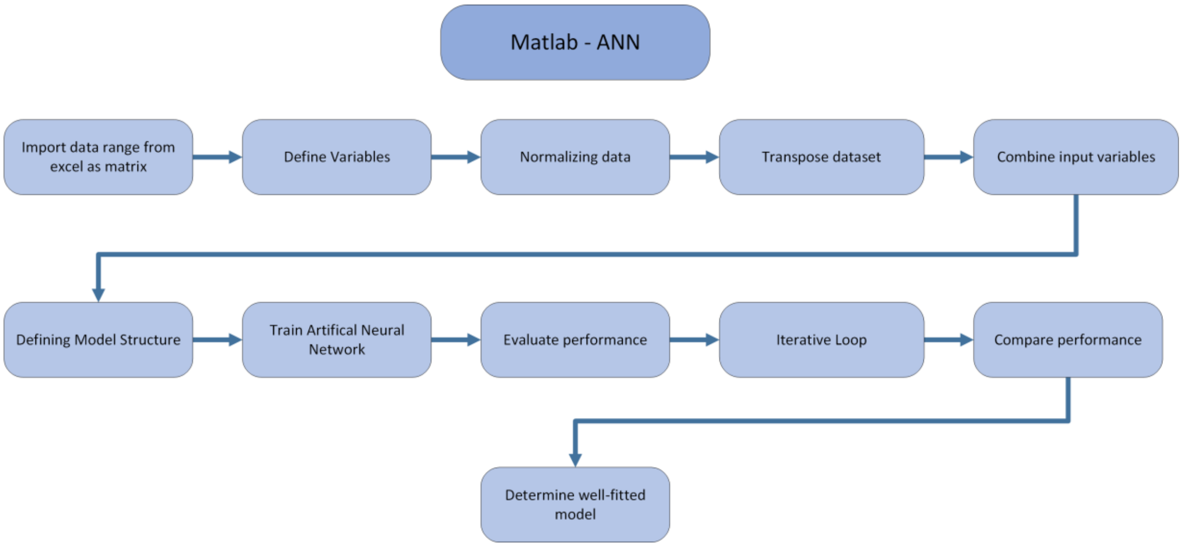


Figure 6.15 Matlab - Artificial Neural Network Workflow.

Normalizing data is performed to ensure the data has an equal range. In this case, the range of variables varies between the various features. For example, U - values range between 0,7 to 1,6 in the case of Windows and 0,12 and 0,22 in the case of External Wall, as described in chapter 6.4.1

By default, the larger values will significantly impact the output value [34]. As such, normalizing the value will ensure that each variable has the same initial weight in the neural network. For a given dataset, the process may be described in the equation below:

$$Normalized\ Dataset = \frac{(Dataset - \min value\ dataset)}{(Max\ value - Min\ value)}$$

Equation 6.5 Normalized Dataset.

Where the minimum value in the given dataset is removed, the remaining variables are divided by the range of the dataset. This provides a number between 0 and 1 for all lists of variables ensuring all values are within the same range.

The model structure is defined by the "*hiddenlayersize* [x y z]" function, where [x y z] here refers to layers, whilst the value of these variables describes the number of neurons in each layer. For these models, only a single layer was selected.

Separating the dataset is done through the *"net.divideParam.(Train/Val/Test)Ratio*". For this thesis, the dataset was separated into 70 % for the training set and 30 % for validation. This is due to the limited dataset of 100 combinations. Although a testing set would be beneficial, it is considered that this data is required in tuning the model rather than testing. The *train(x,y,z)* function takes three inputs (x,y,z). In this case, x reflects the *fitnet* function, which returns a neural network with a defined number of layers and neurons. y reflects input variables, whilst z refers to the output variable.

The following values are defined in the creation of the ANN models:

Table 6.19 ANN Inputs.

ANN Inputs	
Training Set	70 [%]
Validation Set	30 [%]
Testing Set	-
Number of hidden layers	1
Number of neurons in hidden layers	Iterative evaluation
Propagation algorithm	Levenberg-Marquardt algorithm

The model's performance was calculated from Matlab's correlation coefficient (R-value) at the end of the ANN training process. In order to ensure a good fit model and avoid overfitting and underfitting described in Chapter 3.5, an iterative loop was coded, with the number of neurons as variables. The MSE for both the training and validation data was calculated and plotted. This thesis defines a good fit as a point where the MSE for the validation and training sets are relatively low. A comparison of R-values serves as a confirmation.

At this point, the final model with a suitable amount of neurons. Data relevant to this model in this thesis are the FIT function and the R-value. It is worth noting that the accuracy Matlab operates with for ANNs is r which is further described in Chapter 3.5. However, it is notated as R.

6.7.2 NSGA-II

This thesis applies the elitist non-sorted dominated genetic algorithm (NSGA-II) to optimise output based on the three decision variables. Different ANN models were created for the NSGA-II model, with the same number of neurons identified in the optimization step for the previous ANN models. The methodology for application also rests on simulations and calculations performed in Matlab.

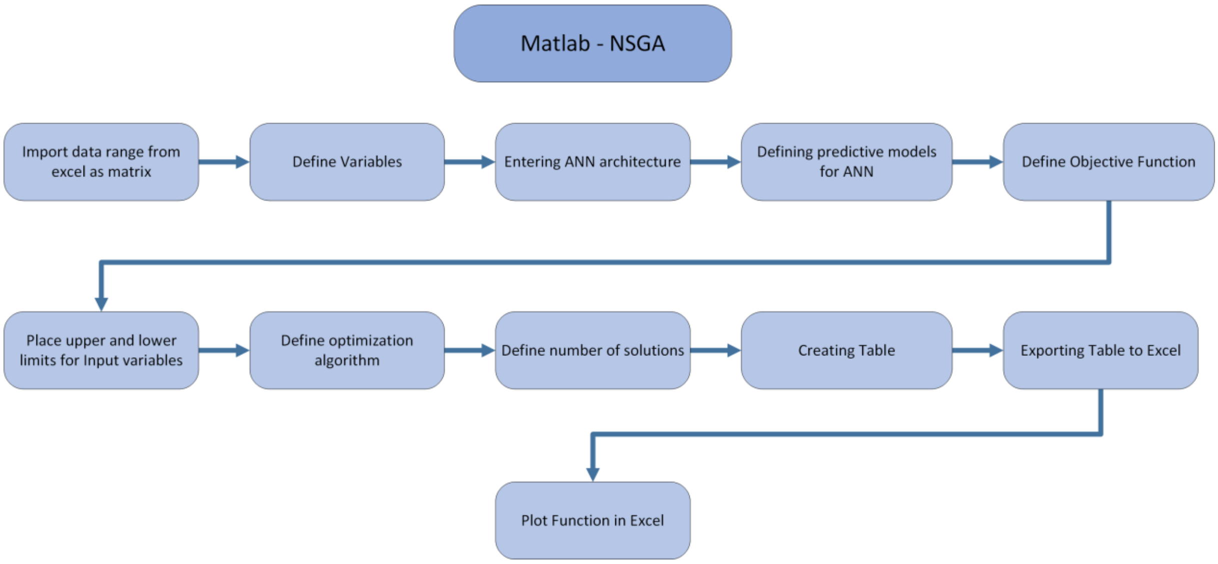


Figure 6.16 Matlab - NSGA-II Workflow.

In the case of this process, two key functions will be described further. The *optimioptions*, and *gamultiobj* function.

Optimioptions takes five inputs, defined with numbers 1 through 5. The first input refers to the function for solving multiple objectives. The second input refers to options to plot the specific function used for the Pareto Front. The third input is the function handle, in this case, the default for Pareto fronts in Matlab. The fourth input specifies the level to display to show under the optimization process, and the final input refers to a value of the display function defining what information should be displayed during the optimization process.

The gamultiobj function has a total of 10 inputs, defined with numbers 1 through 10.
The first input (1) refers to the objective function defined. In this case, the ANN models for costs and energy consumption are combined into a multiobjective function. The following input (2) refers to the number of variables in the objective function. In our case, 3 U-values. The third to sixth inputs (3-6) are potential constraints for the optimization problem, which is not defined in this case. The 7 and 8 input refers to the upper and lower boundaries of the different decision variables (U-Values). The ninth input refers to the potential linear constraint function, which is not defined in this particular case. The final input refers to the previously defined function of *optimioptions.*
In the applied code, the following terms are defined:

Table 6.20 NSGA-II Inputs.

NSGA-II Inputs	
Initial Population	100
Mutation Rate	0,02
Cross-over type	Intermediate crossover
Cross-over rate	0,8
Selection process	Tournament Selection

Termination Criteria	Average spread of Pareto solutions less than 1e-4

In order to identify improvement potential for the decision variables, the following formula is applied:

$$Improvement\ Potential_i = \left\| \left(\frac{Virgin\ value - NSGA\ II}{NSGA\ II} \right) * 100 \right\|$$

Equation 6.6 Improvement Potential.

Where:

$Optima$– is the results of the NSGA-II process from the optimal point identified in the Pareto front
$Virgin\ Analysis$ – is the result of the virgin analysis
i – Decision variable (External wall, Window, Internal Wall)

It is worth noting that by this equation, a negative value describes that the virgin value is lower than the identified improvement potential. By placing an absolute value on this function, we may describe how the virgin value needs to be adjusted to reach the identified optimization value. The higher the number, the more adjustment is required, thus, area for improvement.

The following equation calculates potential improvement:

$$Improvement = (NSGA\ II\ Optima_i - Virgin\ Value_i)$$

Equation 6.7 Simulation Improvement.

Where:
$NSGA\ II\ Optima_i$ – Identified optimized objectives through simulations
$Virgin\ Value_i$ – Current objective performance
i – Objective (Costs or energy consumption)

As described in Chapter 3.9, NSGA-II returns a Pareto front. The selection of the given optimum is based on identifying the "knee-point", a subjectively chosen point that may be considered closest to the intersecting axis (0,0). This value serves as the identified optima used for calculations. Also, this optimum point returns identified optimized decision variables that will be applied to IDA ICE to validate the predicted performance identified in the algorithm.

6.7.3 Validation

The IDA ICE software is also applied for validation purposes. Similar to random number analysis, identified optima values are entered and simulated. This methodology will serve as a tool for validating the optimized results.

Further materials are entered with the following material definitions:

Table 6.21 NSGA-II optima validation.

NSGA-II optima validation simulations	
Optima.x.IntW	NSGA-II Optima, Internal Wall
Optima.x.ExtW	NSGA-II Optima, External Wall

The methodology described in Chapter 6.6 is applied to the identified decision variables through the NSGA-II process for cost analysis validation.

6.8 Digital Twin

The Digital Twin in this thesis is based on the IFC BIM file that was provided. However, illustrating the results of the optimization process in such a large model is considered impractical. The Digital Twin model was created as a separate Rvt file to reflect the given case, classroom 3006.

The creation of this model is based on a combination of extracting data from the IDA ICE model, which identified the general geometry of the case, whilst other data that were not sufficiently defined, such as the digital visualization of the windows, was directly copied from the IFC model.

This serves as the digital model of the physical construct to qualify as a digital twin per the definition described in Chapter 3.2. The digital twin aims to enable decision-makers to identify improvement potential and determine which category would be most beneficial to address swiftly and accurately.

The model of room 3006 with the implemented visual program described in Chapter 6.4.2 is included in *Appendix G – Digital Twin.rvt*

7 Results

This thesis combines different methodologies and results before compiling them into a digital twin, as such results may seem segregated. In order to address this concern, one may consider the illustration in the figure. This chapter is divided into similar steps as Figure 6.3 Software Hierarchy, presented in Chapter 6.3. Random number generation serves as the basis for energy and cost analysis. The datasets from this analysis serve as the foundation for creating artificial neural network models, which NSGA-II then optimizes. The final result is visualised in BIM. The process as a whole is considered the process of creating a digital twin.

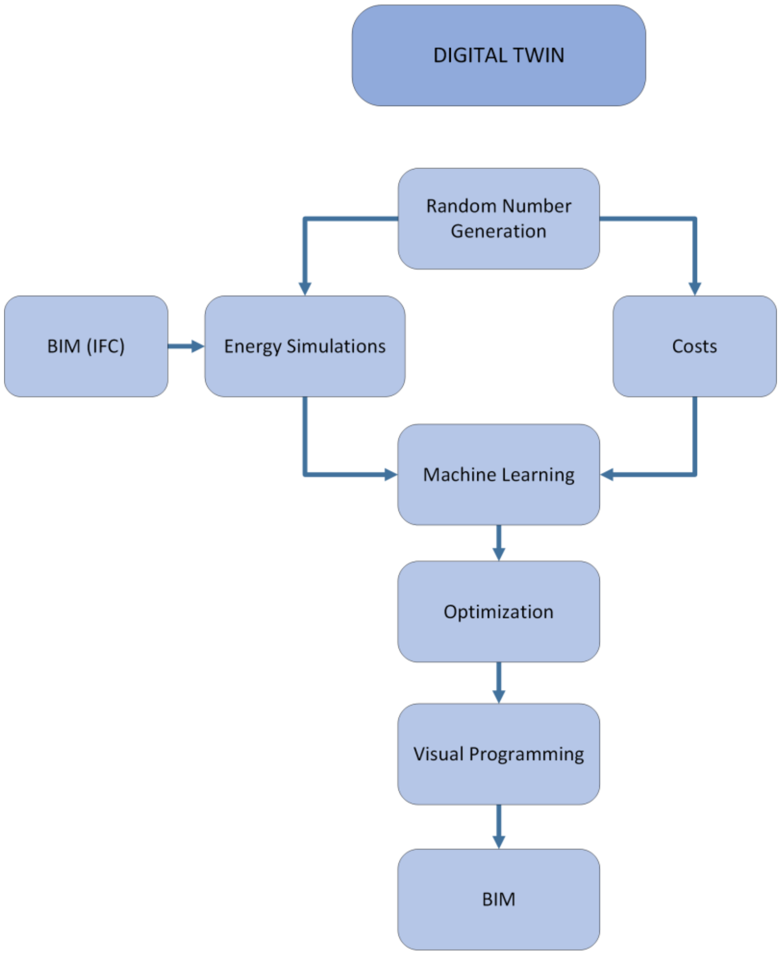


Figure 7.1 Result Presentation.

The final part of this chapter will present relevant case studies identified by the literature review methodology.

7.1 Random Number Generation

We get the following distributions by plotting the random variables generated by visual programming for internal walls, external walls, and windows. The full range of combinations is included in *Appendix*

D – Datasheet.xlsx under the *"Variables for analysis"* sheet. The code applied is illustrated in Figure 6.4 Random Number Generator.

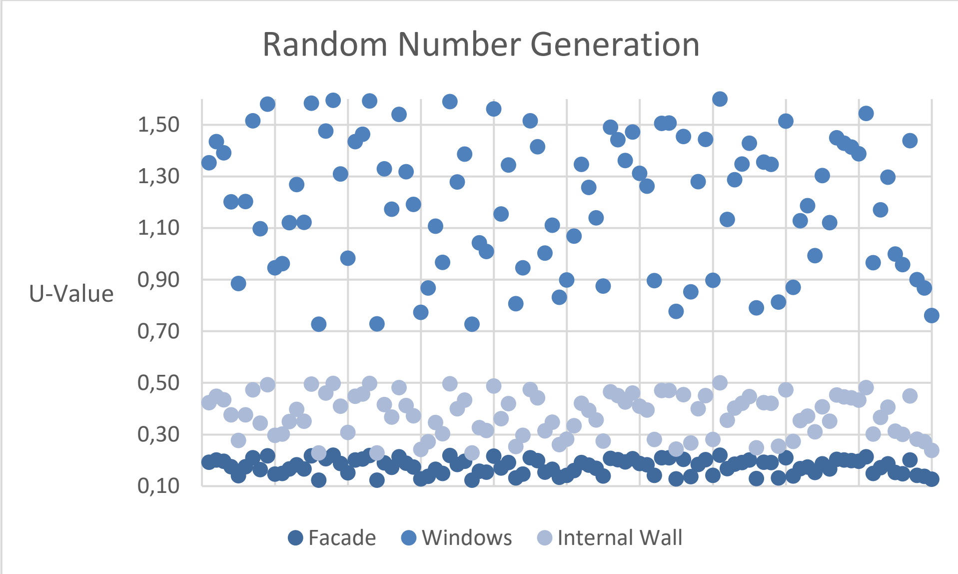


Figure 7.2 Random Number Generation.

7.2 Energy Consumption Analysis

The simulation results in this thesis reflect the results achieved through the IDA ICE methodology, described in Chapter 6.5.

The Virgin analysis is based on the following values for thermal transmittance, which is a repeat of Table 6.18 Virgin Simulation U-Values

Table 7.1 Virgin Simulation U-values, repeat from Table 6.18.

Category	U-Value [W/(m²K)]
External Wall	0,1511
Windows	0,74
Internal Walls	0,6187

As described previously, IDA ICE is a complex software that allows for results with significant detail. As such, only key areas are provided in this table. A complete analysis is included in a separate analysis thesis provided as *Appendix J – Virgin analysis.xlsx*

Table 7.2 Virgin Analysis Result.

Virgin Analysis Results		
Energy	Energy consumption [kWh/year]	Energy consumption [kWh/ (m².year)]
Zone Heating	935,7	11,57

HVAC aux	479,7	5,93
Hot water	-	-
Fans	275,9	3,41
Pumps	-	-
Lighting	341,7	4,23
Technical equipment	1160,30	14,35
Room Cooling	606,1	7,5
∑ (approx.)	3800	47

Using the random number analysis described in Chapter 6.5.4, 100 variables for each category were created, leaving 100 combinations. The results of the IDA ICA energy efficiency simulation varied from 35,83 to 55,18 kWh/m^2. A list of these variables and their results from simulations in IDA ICE is included in *Appendix D – Datasheet.xslx* under the *"IDA ICE"* sheet.

The choice of adding 100 simulations was made in order to achieve a good database for the ANN models. The results of the simulations may be identified in the plot below. The results from the virgin analysis included in *Appendix J – Virgin Analysis.xlsx* is also included in this figure.

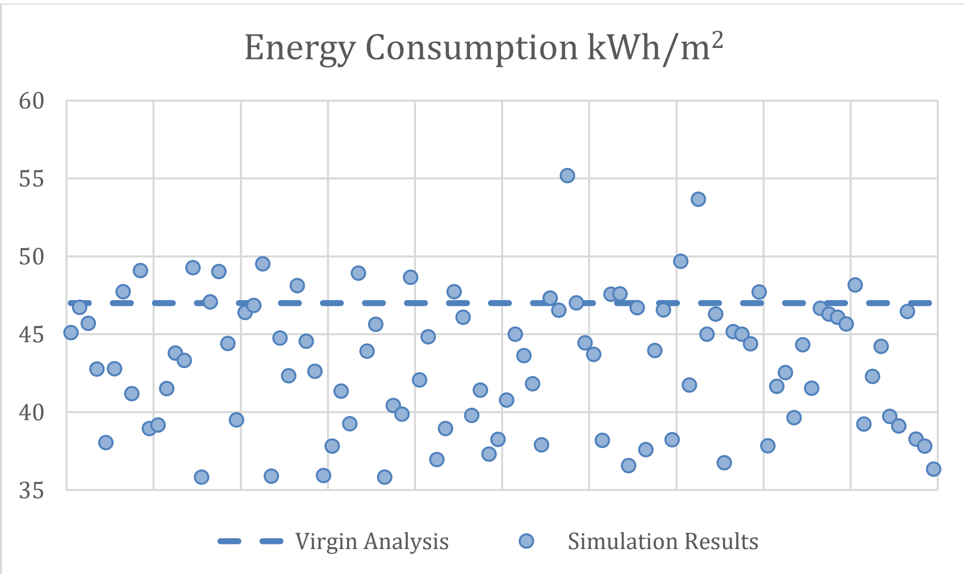


Figure 7.3 Energy Consumption Simulation Results.

7.3 Cost Analysis

As per the NS-EN 16627:2015 flowchart covered in illustrated in Figure 6.13 Cost Analysis Process. We define the following:

Table 7.3 Cost Analysis Overview.

Cost Analysis overview	
Purpose of the assessment	Identify costs for classroom 3006
Specification of the object of the assessment	Limits to internal walls, windows and external Wall
Scenario Development	Immediate, Non-lifecycle
Quantification of the object	Calculation per fraction of area, see *Appendix D – Datasheet*.xlsx "*Cost Definitions"* sheet.
Selection of economic data	Installation Costs & Material Costs
Calculation	Excel, see *Appendix D – Datasheet*.xlsx "*Cost Calculation.*" sheet
Communication	N/A
Verification	N/A

By applying Equation 6.3 Element quantity per m^2/IFA, we have:

Table 7.4 Element Quantity per IFA.

Element Quantity per IFA		
Category	Area	φ_i [m^2/IFA]
External Wall	13,6 [m^2]	0,168
Window	64,075 [m^2]	0,792
Internal Wall	16,994 [m^2]	0,21

Per the definition of the Virgin Model, the following costs are identified:

Table 7.5 Cost Calculation - Virgin Model.

Virgin Model Costs			
Category	External Wall	Window	Internal Wall
U-Value	0,1511	0,74	0,6187
Defined Material	Insulation 225 mm	Window Type = 0,7	Insulation 50 mm
Maintenance cost [NOK]	65,84	1087,2	41,68
Material Cost [NOK]	203,68	2879,03	52,59
φ_i [m^2/IFA]	0,168	0,792	0,21
Costs [NOK/m^2]	45,34	833,67	74,71
Grand Total Costs [NOK/IFA]	**953,72**		

By rounding the randomly generated numbers to existing materials per methodology described in Chapter 6.6 and identifying costs per Equation 6.4 Total Costs, we have the following plot:
All cost calculations are included in *Appendix D – Datasheet.xlsx* under the "*Cost Calculation"* sheet.

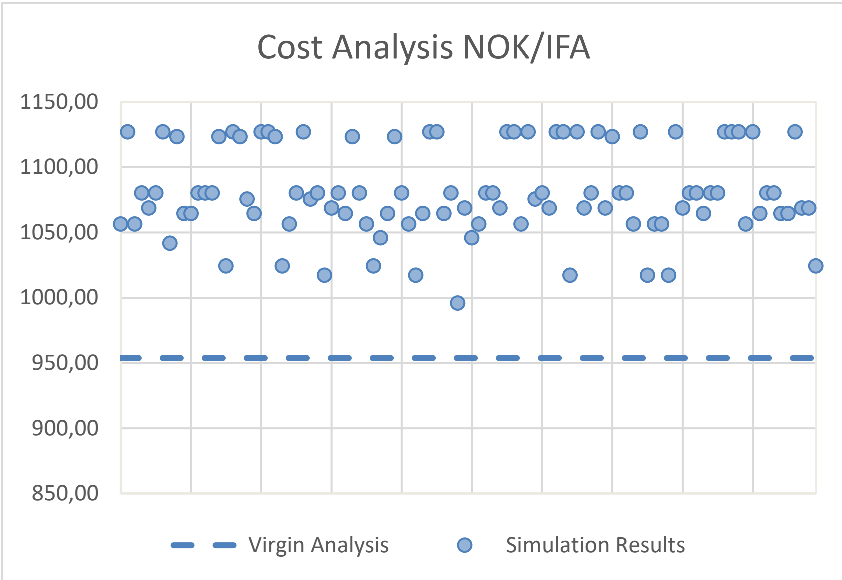


Figure 7.4 Cost Analysis Results.

For further reference, NOK/IFA is translated to NOK/m^2 for ease of interpretation.

7.4 Artificial Neural Network

This thesis creates two artificial neural network models, one for energy consumption defined as kWh/m^2 and one for NOK/m^2. The models are evaluated with a single hidden layer, where the optimal number of neurons is identified by evaluating the performance of 30 models. They are further evaluated on two criteria. Firstly, the performance is identified through a well-fitted model, identified as a model with low MSE for both the training set and the validation set, and its R-value. This chapter will present the results for the two neural networks separately, firstly the cost analysis prediction model NOK/m^2), followed by the energy consumption prediction model (kWh/m^2). Finally, it will present a graph with R values for all 30 models.

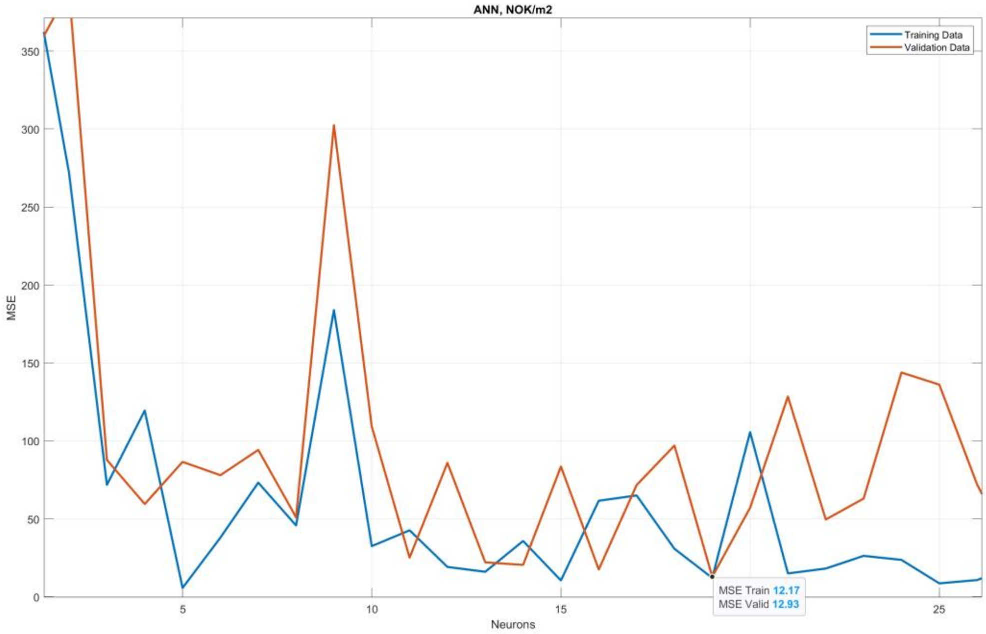


Figure 7.5 ANN, MSE/Neurons Cost analysis.

As may be identified from the graph above, the place with the lowest degree of MSE for both the training set and the validation set is when the neural network contains 19 neurons.

The identified model with 19 neurons is displayed in the figure that follows:

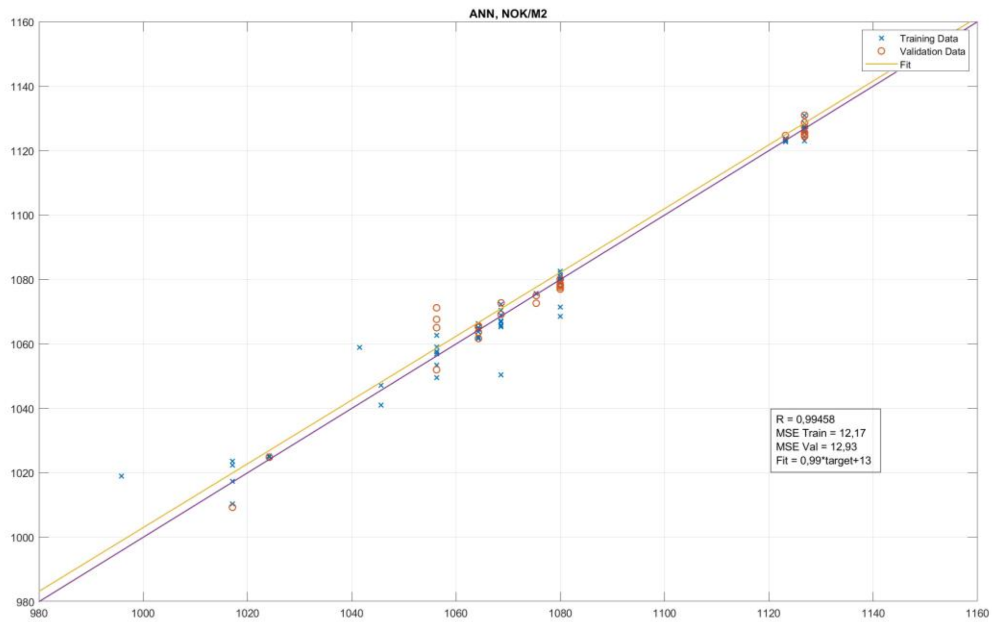


Figure 7.6 ANN, Cost Analysis - 19 Neurons.

Followingly, the values are repeated in the table below:

Table 7.6 Artificial Neural Network, Cost Analysis - 19 Neurons.

Artificial Neural Network- Cost analysis	
R-Value	0,99458
MSE Training Set	12,17
MSE Validation Set	12,93
Fitness	0,99*target+13

Followingly the results of the ANN model for energy consumption will be presented.

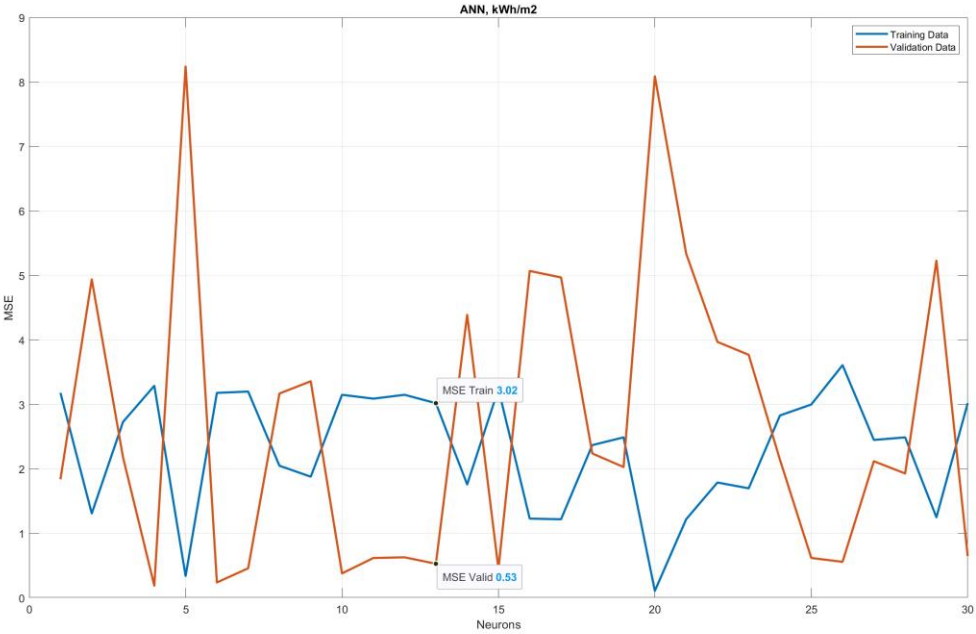


Figure 7.7 ANN, MSE/Neurons Energy Consumption.

By applying the same methodology to the energy consumption model, we may identify that a suitable number of neurons is 13, where the MSE of the validation and training data is low (Note; by comparing the R-value, 13 neurons perform better than 18 and 19).

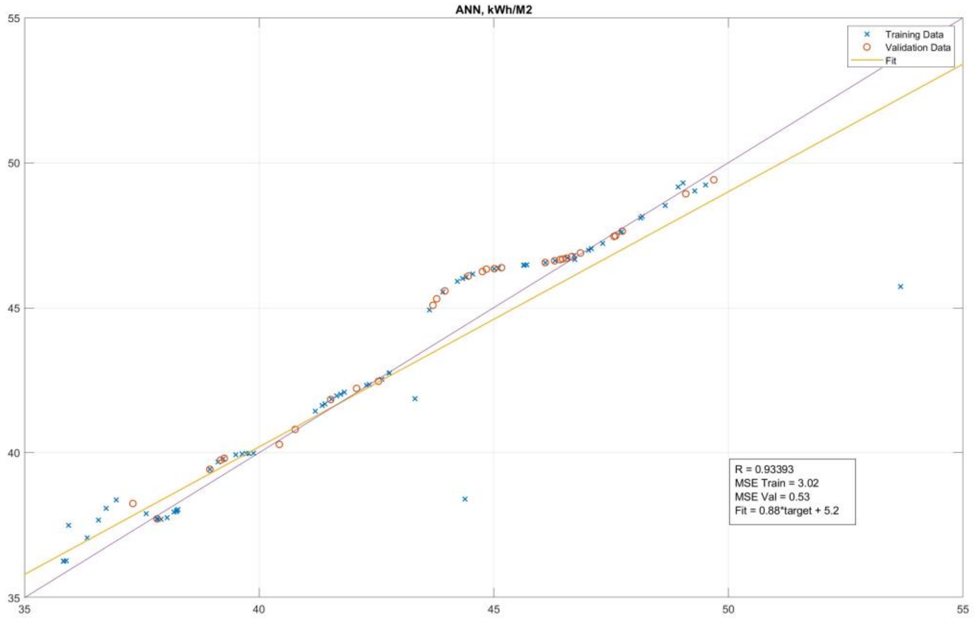


Figure 7.8 ANN, Energy Consumption - 13 Neurons.

The specific values are repeated in the table below:

Table 7.7 Artificial Neural Network, Energy Consumption - 13 Neurons.

Artificial Neural Network – Energy Consumption	
R-Value	0,93393
MSE Training Set	3,02
MSE Validation Set	0,53
Fitness	0'88*target+5.2

The overall performance of the models is depicted in the figure below, which displays the R-value identified in the simulation process.

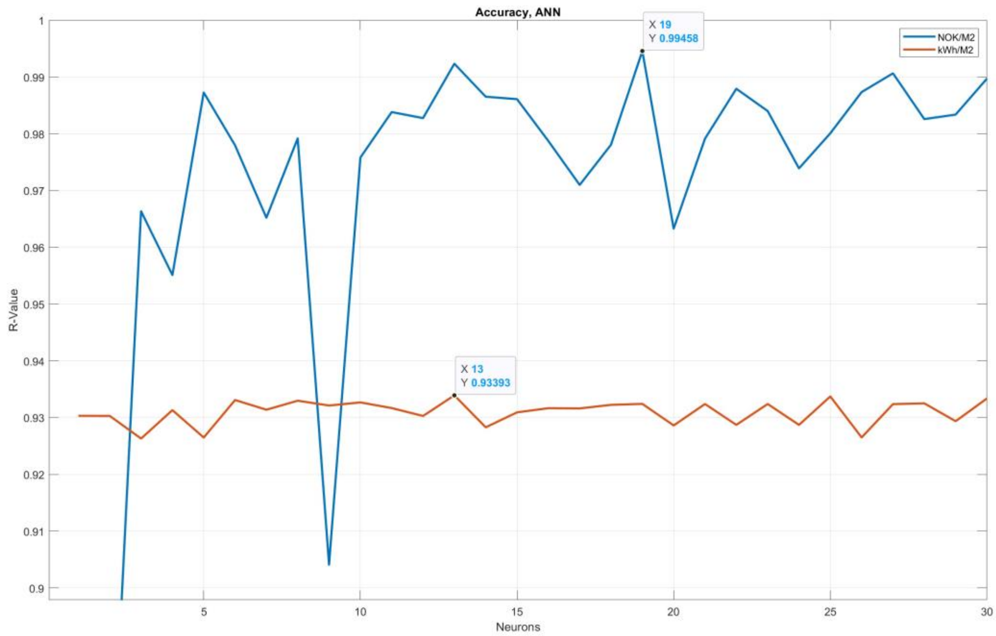


Figure 7.9 R-Accuracy ANN models.

7.5 NSGA-II

The Pareto front below illustrates the findings from the NSGA-II simulation process. The genetic algorithm was based on the following parameters:

Table 7.8 NSGA-II Parameters.

NSGA-II parameters	
Crossover type	Intermediate crossover
Crossover rate	0.8
Selection	Tournament selection
Mutation Rate	2,5 %
Ending Criteria	The average spread of Pareto solutions less than 1e-4

The algorithm created 187 generations before the termination criteria were met. The Pareto Front is displayed in the figure below, with a chosen optimum with approximate values of 30,3 kWh/m^2 and 905,23 NOK/m^2.

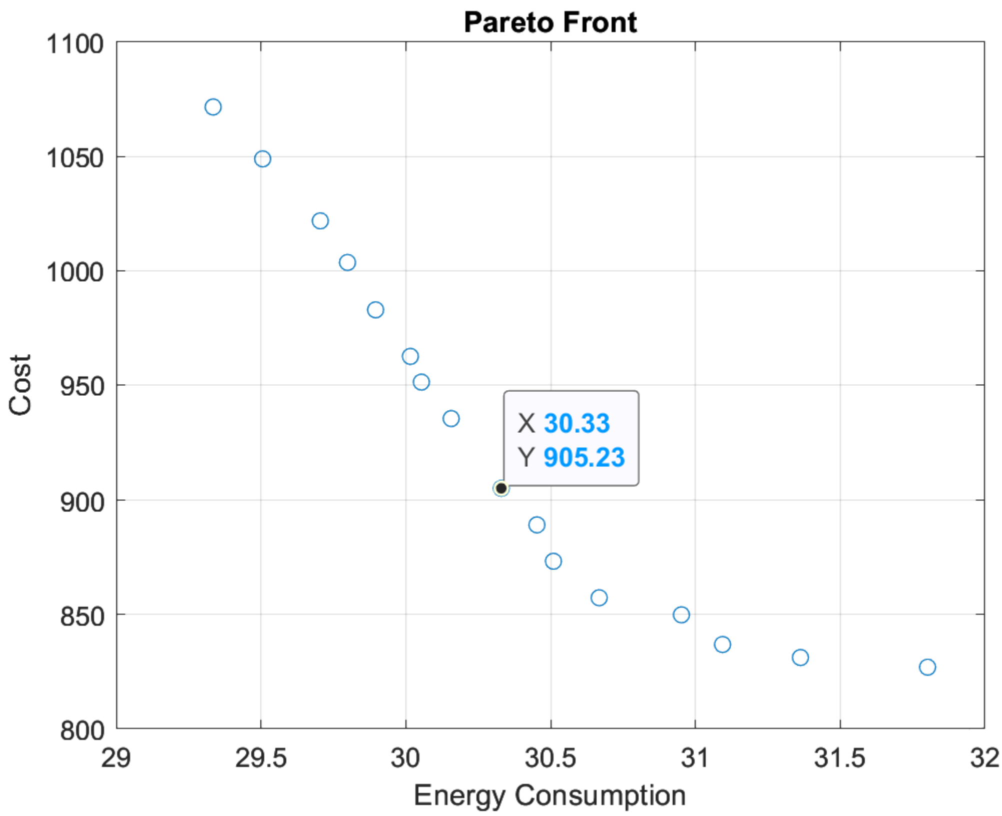


Figure 7.10 NSGA-II Pareto Front.

The identified decision variables from the simulation variables are identified in the figure below, where the chosen optimum is highlighted in orange.

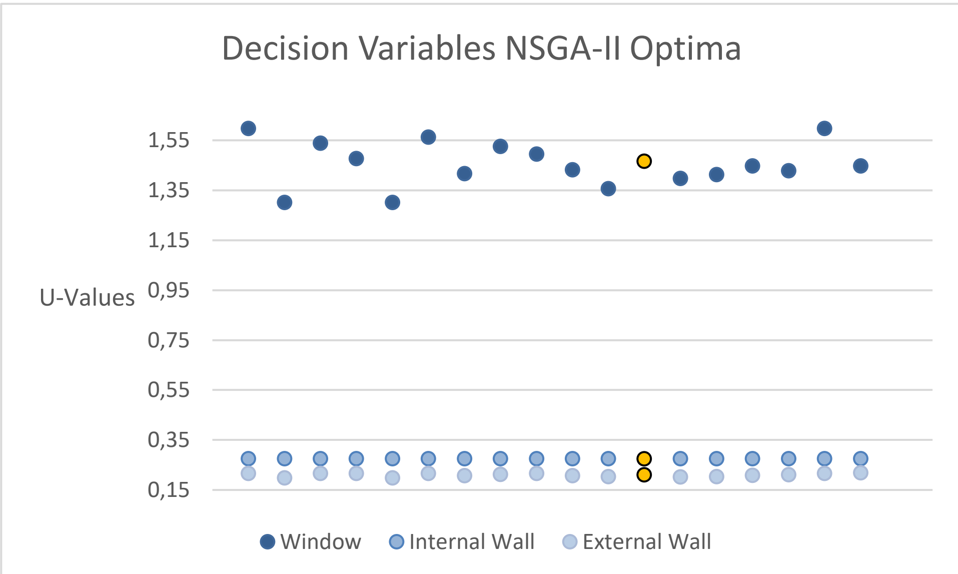


Figure 7.11 Optima U-Values.

By extracting these values combined with the optimised decision variables, the total performance is illustrated in the table below:

Table 7.9 NSGA-II Results.

NSGA-II Optima				
Energy Consumption [kWh/m^2]	Costs [NOK/m^2]	U-Value External Wall [W/(m²K]	U-Value Window [W/(m²K]	U-Value Internal Wall [W/(m²K)]
30,33	905,23	0,21024	1,4656	0,27458

Table 7.10 NSGA-II optima, Decision variables.

NSGA-II Optima, Decision Variables		
Insulation External Wall	Windows type	Insulation Internal Wall
170 mm	Window = 1,6	100 mm

By applying Equation 6.6 Improvement Potential we have:

Table 7.11 Improvement Potential.

Improvement Potential			
Category	External Wall	Window	Internal Wall
Optima	0,21024	0,21024	0,27458
Virgin Analysis	0,151	0,74	0,619
Improvement Potential	39 %	98 %	56 %

While applying Equation 6.7 Simulation Improvement we have:

Table 7.12 Simulation Improvement.

Simulation Improvement		
Category	Energy Performance [kWh/m^2]	Cost Performance [NOK/m^2]
Virgin Analysis	47	953,72
Optima	30,33	905,23
Simulation Improvement	16,67	48,49
Simulation Improvement [%]	35 %	5 %

If we translate the simulation improvement with Equation 6.6 Improvement Potential, we may identify an improvement of 35 % in Energy Performance and a cost improvement of 5 %.

By performing new simulations with the identified decision variables for energy consumption, we have the following distribution with the optima marked with orange:

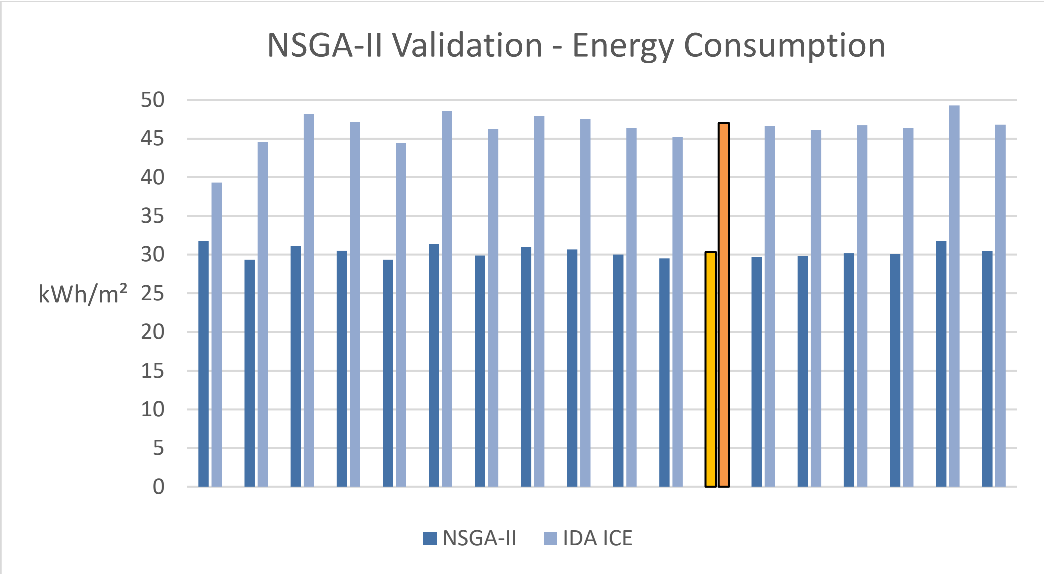


Figure 7.12 NSGA-II Validation - Energy Consumption

Whilst new calculations per methodology described in Chapter 6.6, we have the following distribution:

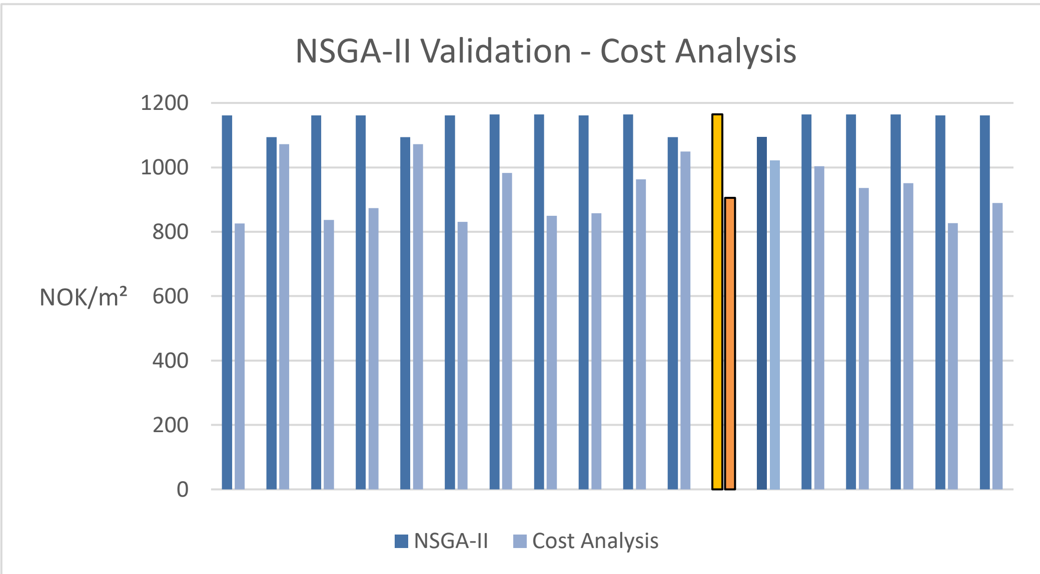


Figure 7.13 NSGA-II Validation - Cost Analysis

7.6 Digital Twin

The digital Twin is displayed in the figure below:

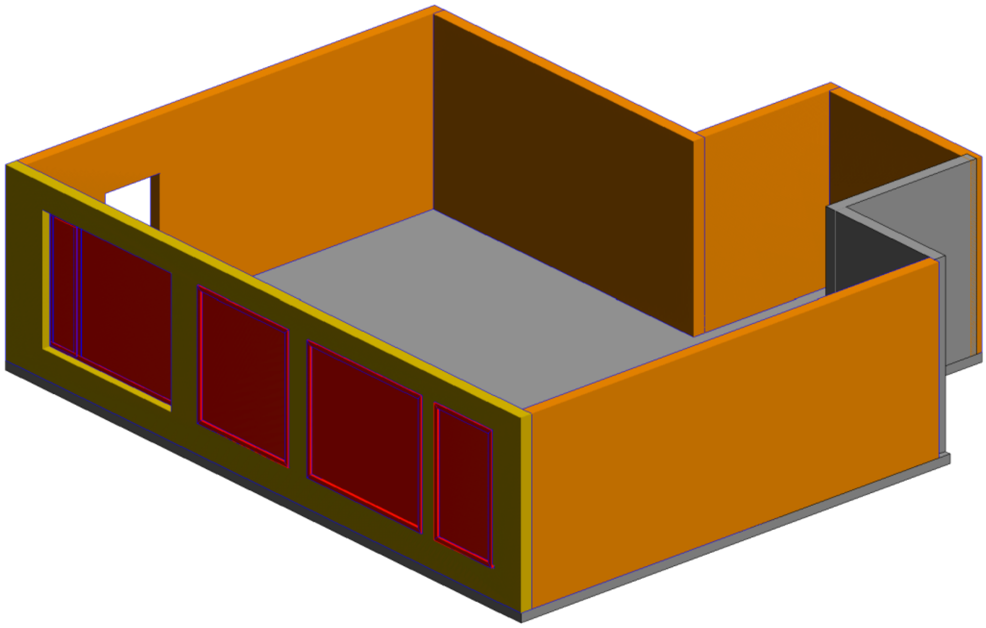
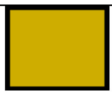
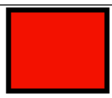


Figure 7.14 Digital Twin

This figure covers the given case room with the visual program. The colours reflect the degree of optimization calculated by Equation 6.6 , comparing U-values from the virgin analysis with the optimized variables identified by NSGA-II. Visualised by the program illustrated in Figure 6.5 Visual Program - Visualization. As described per methodology: closer to green, less effort and degree of change will be required to reach the optimized value. It is worth remembering that the grey walls reflect the elevator shaft, which is not included in the simulations.

The values and colour scheme are displayed in the table below:

Table 7.13 Digital Twin Visualization.

Digital Twin Visualization		
Colour	Category	Improvement Potential
	External Wall	39 %
	Window	98 %
	Internal Wall	56 %

7.7 Previous research

Similar research has been conducted in the field of implementing digital twins for energy and cost performance improvement. However, in the chosen methodology, the research identified covers other parameters than this thesis. For example, the work of Hosamo, Nielsen et al. implementing digital twin of HVAC systems (HVACDT) and multi-objective algorithms to identify optimal energy consumption and thermal comfort solutions [30]. Hosamo, Nielsen et al. Implementing digital twin for fault detection and predictive maintenance [59]. Hosamo, Tingstevdt et al. stream visual information to BIM using digital programming to identify parameters for reducing energy-building consumption and optimizing thermal comfort [6].

The study of Nour El-din, Pereira et al. [20] covers digital twin applications in areas such as structural health monitoring, concept frameworks, construction management, facilities management and circular construction. Further, the bibliometric study by Zhao, Feng et al. [23] covers additional applications of digital twins in AEC in areas such as facility management and predictive maintenance. However, due to the chosen methodology, these studies are not expanded upon in further detail.

Limited research has been conducted in implementing a digital twin framework with building performance optimization to reduce energy consumption and investment costs as of the time of writing this thesis. Identifying studies based on the same subjects as this thesis has proven to be largely unsuccessful by the chosen methodology. The following paragraph will introduce one case study that aims to explore cost-effective and energy-efficient design using NSGA-II and BEM simulation data, which strongly resembles the methodology in this thesis. It does not, however, implement Digital Twin as defined in this thesis.

An article published by Yaolin Lin and Wei Yang published in 2018, *"Application of Multi-Objective Genetic Algorithm Based Simulation for Cost-Effective Building Energy Efficiency Design and Thermal Comfort Improvement"* [12]. Explore the trade-offs between energy consumption, initial construction costs, life-cycle costs, and the number of thermal discomfort hours using an improved multi-objective algorithm combined with building simulations [12]. The case study was conducted using a single reference design building. It simulated the cold/warm weather effects in 5 different areas of China, all of which experienced hot summers and cold winters. The study aimed to explore cost-effective building design with four objective functions, calculating energy consumption, initial construction costs, life cycle costs, and indoor thermal comfort [12].

The variables investigated covered: Building orientation, window-to-wall ratio, heating temperature setpoint, cooling temperature setpoint, external window shading, glazing type for windows, external wall type and roof type. Firstly, an analysis of energy consumption was investigated using building energy modelling (BEM) software.

The optimization process was performed using NSGA-II, and the factors for the algorithm were set as a population size of 200, 0.95 crossover rate, 0.02 mutation rate and 100 generations. The objective functions were separated into construction costs + building energy consumption and discomfort hours + life cycle costs. They found 244 distinct Pareto solutions through their analysis.

Lin and Yang conclude that an average energy-saving potential of 29,08 % can be achieved at the design stage. Furthermore, based on the case in 5 different environments, the findings reflect an average energy consumption reduction of 38,6 % with an increase of 3,18 %,

The study, however, did suffer limitations: It was a single-story concrete residential building focusing on the building envelope and cooling heating setpoints, so it could not optimize the HVAC system. It is set in specific summer/cold winter regions in China and is thus not generalizable. The study considered a constant window-to-wall ratio, which in practice varies considerably. The study only considered overhang shading. The building envelope design parameters were limited to ceramic tiles, concrete bricks and plasterboard for external wall types. For ceilings, the design parameters were limited to light concrete, waterproofing, extruded polystyrene (EPS) and concrete slabs.

8 Discussion

The discussion chapter seeks to expand and discuss the results found by the given methodology. Given the length of this thesis, certain results from Chapter 7 are repeated for ease of interpretation and discussion. This chapter is organized as follows: firstly, the findings are compared to existing research and literature identified in the literature review. Followingly the various steps in the framework are covered in the order they are applied in the framework. At the end of the chapter, a brief summary of the various discussion chapters is presented, leading up to the conclusion of this thesis.

8.1 Relation to Previous Research

As described earlier, Digital Twin implementation is a relatively new field in the AEC. One may argue that this thesis is innovative and novel if one considers digital twins combined with neural networks and genetic algorithms for optimising energy consumption and investment cost. However, this serves as a challenge, as there is limited existing literature for comparative purposes. This thesis only identified one per literature review methodology described in Chapter 6.1. Namely, the study of Lin & Yang presented in the previous chapter.

Their study identified an average energy-saving potential of approximately 38,6 % with a cost increase of 3,18%. Comparingly the findings of this thesis identify an energy-saving potential of 35 % and a cost *reduction* of approximately 5 %.

There are, however, significant differences in methodology. For the NSGA-II algorithm, one may identify the following differences.

Table 8.1 Lin & Yang NSGA-II parameters.

Lin & Yang NSGA-II parameters	
Initial Population	200
Mutation Rate	0,02
Cross-over rate	0,95
Termination Criteria	100 generations

Other information, such as cross-over and selection types, is not disclosed. Comparingly, this thesis applies the following parameters:

Table 8.2 NSGA-II parameters, repeat from Table 6.21.

NSGA-II Inputs	
Initial Population	100
Mutation Rate	0,02
Cross-over rate	0,8

Cross-over type	Intermediate crossover
Termination Criteria	Average spread of Pareto solutions less than 1e-4
Selection process	Tournament Selection

When comparing the results, they appear to correspond in terms of energy. The discrepancy is only at 3 %. Costs, however, differ significantly (8 %). Reasons are unclear but likely include additional decision variables (parameters for analysis) and more products.

The study by Lin & Yang covers eight decision variables, whilst this thesis only applies three. Lin & Yang apply 35 different discrete values (products), whilst this thesis applies 13. Additionally, the case geometry differs significantly. Lin & Yang applies a square single-story building of 10x10 meters in 5 different environments. Whilst this thesis is based on a classroom in a single environment.

Although the values are comparingly equal in terms of energy, there is a level of uncertainty in the findings of this thesis, as will be discussed in later chapters.

8.2 Random Number Generation

The results of the random number generation are intriguing and worth a mention. By inspecting these numbers further and plotting them in a graph diagram instead of a dot diagram, we have the following:

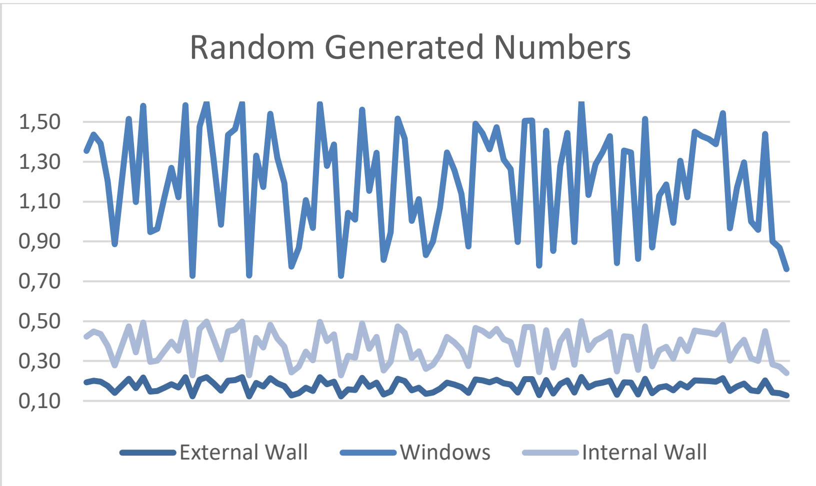


Figure 8.1 Random Generated Numbers, Graph plot.

By inspecting these graphs, one can quickly identify that the skewness of the various simulated numbers coincides with other categories. This reveals an apparent weakness in the random number generation programming conducted by Dynamo. This indicates that the random number generation is limited to a specific count within a certain range rather than an arbitrary value within each range. To illustrate this, when comparing the values from external and internal walls in a histogram, we have the following distribution:

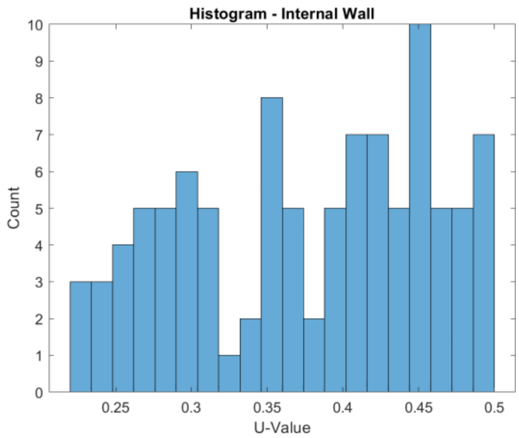


Figure 8.2 Histogram - Internal Wall.

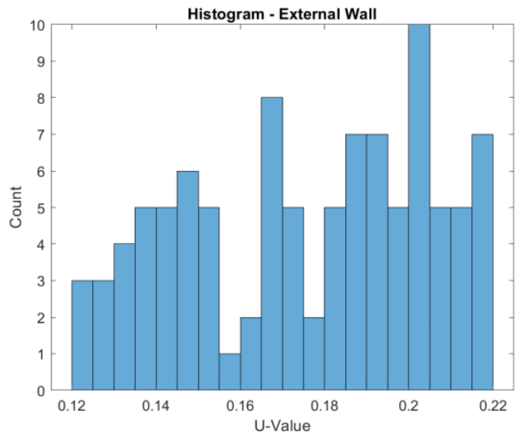


Figure 8.3 Histogram - External Wall.

These histograms reveal the exact amount numbers in the ranges for both categories. In effect, one gets a lesser degree of combination diversity, causing more extreme values rather than small incremental changes by tuning the different variables to a higher or lesser degree.

However, regarding simulations and machine learning techniques for identifying a correlation between energy consumption and U-values, the numbers may be considered satisfactory, although not ideal, as may be identified by the R-value of the created ANN models presented in Chapter 7.4.

Further, the range of the variables chosen per methodology is unfortunate. The values generated surpass the performance of the virgin model. This results in the energy performance simulations will return more energy-efficient results. Correspondingly, the cost analysis will return a higher value than the virgin model.

8.3 IDA ICE Simulations

The IDA ICE simulations in this thesis are based on three separate decision variables: Internal walls, External walls and Windows. However, by inspecting the correlation between them, it is clear that internal walls do not influence the result in a significant way. This is because one does not calculate heat loss for internal zones as described in Chapter 3.6 Energy Performance. In fact, in inspecting two additional simulations with the maximum and minimum range chosen per methodology. The energy consumption differs by 0.1 kilowatt hours, efficiently making up for 0.2 % of the final value.

The choice to include it in the simulations is made to quantify as much of the room as possible with common decision variables. From the viewpoint of material choice and envelope design, this is based on material performance. Thus thermal transmittance was considered to be the most appropriate choice.

This process is not without weaknesses. For example, the IDA ICE model is simulated on a generic geometry of wall composition, with randomly generated values in terms of thermal conductivity. These values do not correspond to real-life measurements, affecting the decision-making's reliability and credibility.

As described in Chapter 3.6, thermal efficient materials used for insulation are, for example, ultrafine glass wool with a thermal conductivity of 0,027 W/mK, extruded polystyrene at 0,031 W/mK, or one of the most thermally efficient materials in the world, aerogel with a thermal conductivity of 0,003-0,012 W/mK. In this thesis, many values surpass ultrafine glass wool and extruded polystyrene. For instance, the external wall input in IDA ICE has values as low as 0,0126 W/mK, which is over twice as thermal efficient as ultrafine glass wool and nearly on the level of aerogel. Furthermore, the custom materials have been defined with unrealistic density and heat storage capacity.

It is unclear what impact this has on the actual prediction and optimization process. It is, however, safe to assume that the physical construct these simulations are based on does not correspond with the assets of the physical construct.

There is also the consideration that IDA ICE is complex software. The application of this software is self-taught in this thesis. This serves as a considerable weakness when compared to accurate results. For instance, by comparing the results of the virgin analysis (47 kWh/m^2) with the definition of criteria for energy efficiency measures of (110 kWh/m^2), one may easily identify that the analysis returns a value more than twice energy efficient as the requirement. However, Tvedestrand Videregående Skole has achieved the plus house standard and is considered one of Norway's most energy-efficient schools, as described in Chapter 5.1. It may, as such, be considered to be realistic. However, considering an energy-consumption simulation returning a value of 35 kWh/m^2 raises doubt.

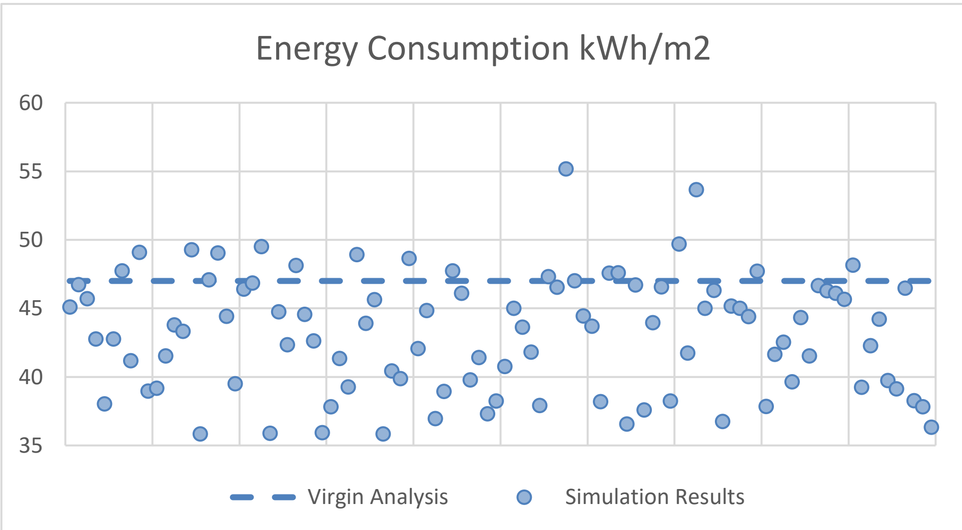


Figure 8.4 IDA ICE Simulations, Energy-Consumption, repeat of Figure 7.3.

When inspecting the results of the simulations illustrated in the figure above, we may identify that a significant amount of the results are lower than the virgin analysis, indicating that the range of the randomly generated variables is not ideal. Ideally, these values should be spread evenly surrounding the virgin analysis for a more fair expectation. It illustrates and arguably sets the expectation that the room may be optimized.

One may also identify two outliers, with values exceeding 50 kWh/m^2. This indicates that the simulations are wrong in these two cases. The reasons are unclear but likely include a human error in entering input or correctly reading output.

Another concern we may interpret from the figure is the distribution of results compared to the virgin analysis. Most of the simulations conducted perform better than the original model.

8.4 Cost Analysis

The result of the cost analysis is illustrated in the figure below:

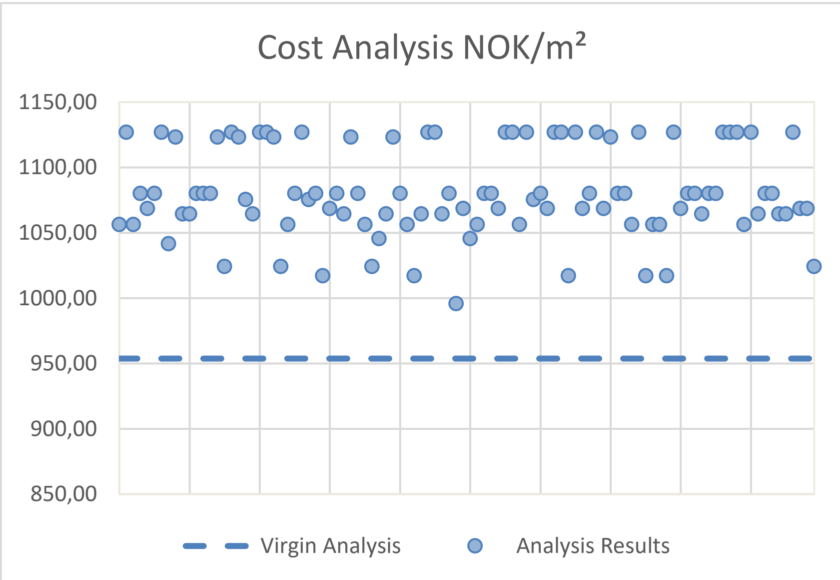


Figure 8.5 Cost Analysis Results, repeat from Figure 7.4.

As may be interpreted by the figure, the results clearly emphasise the concern mentioned in previous chapters, the original dataset values. One may identify that the results are far beyond the virgin analysis. It is especially prominent in the cost analysis as all values are more costly than the original model. This indicates that the range of decision variables is insufficient. It is worth remembering that the ANN models intend to cover a good regression plot. The models may sufficiently interpret and predict values from the analysis and simulations if accurate.

8.5 Artificial Neural Networks

The creation of artificial neural networks in this thesis is primarily based on tutorials from the identified courses and troubleshooting described in Chapter 6.7. The process of creating the ANN models has, in large parts, been successful, as may be identified for accuracy. Furthermore, defining a well-fit model with low MSE for both the training and validation set corresponds to the expectation of reaching a high R-value. However, the findings reflect that a low MSE for the validation set is an important factor that affects the accuracy of the ANN model.

The chosen methodology for the selection of the model corresponds to the expected results. The models presented in this thesis may be considered accurate. In particular, the cost model presented in Figure 7.6 with an R-value of 0,9948. This model may be considered to be extremely accurate. In the case of the energy-consumption model presented in Figure 7.7, the R-value equals 0.9339, which may also be considered an accurate model. One may also interpret that one of the reasons the R-value is reduced is a relatively high degree of MSE on the training set data. Another reason may be the erroneous data points included in the simulation results.

There are, however, challenges with the chosen methodology and application of ANNs on a broader scale. Firstly, the identified models are not the ones applied for the NSGA-II algorithm, only the amount of neurons in the hidden layers. This methodology is unfortunate. Each time a model has a unique training and validation set, the values in creating the model differ in performance, despite having the same amount of neurons. Generally, this is one of the drawbacks of using neural networks themselves. They require a significant amount of data.

Secondly, by applying ANN to the dataset, the different variables are considered to be continuous. This is, however, not accurate. A material is discrete, with a fixed price and fixed performance. This is especially prominent with the material rounding methodology for cost calculation. The methodology chosen in this thesis is only accurate if the price and performance of different materials have a linear relationship (under the presumption of a sufficiently large dataset). One may consider this as displayed in the figure below:

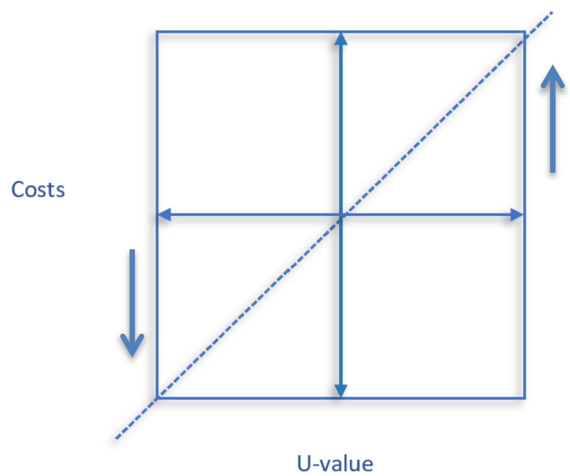


Figure 8.6 Linear relationship of regression.

This figure summarizes the presumption of the current methodology, which considers the square to represent a given material. Using regression and rounding values work as the arrows displayed above. The closest part of the square is selected, which identifies the corresponding price. Values located in the range of the left half of the square are rounded to the lower boundary of costs, whilst the right half of the square is rounded to the top boundary of the costs. The square illustrates that the U-values have the same ratio as a range of costs.

This is, however, not the case. The larger the range of the U values, the more significant error. All values on the left side are considered lower, while all values on the right side are considered higher. Thus, a rounding error occurs corresponding to the regression.

One may consider the level of error will be reduced based on two parameters:

- More products - (which leads to fewer product squares, thus less rounding error)
- More decision variables – The error level is spread amongst a more extensive range of products.

If we consider this to be the case between many different decision variables, one could argue that the level of error will decrease as it will be distributed over a more extensive range, and the rounding methodology will reduce the error. It will, however, always be present.

We may consider this a performance gap or a weakness in the methodology. Regardless, it reflects the consequences of applying purely theoretical concepts to a "real-life" scenario.

8.6 Optimization

Applying the NSGA-II algorithm identifies an optima with an energy consumption performance of 30,33 kWh/m^2 and a cost performance of 905 NOK/m^2. When comparing these values to the ones from the virgin analysis, a cost reduction of 5% and an energy consumption reduction of 35% is identified. These numbers show great promise in the energy consumption category, although the cost reduction is significantly smaller. These values also raise concerns. One would consider that an energy-efficient solution rests on more energy-efficient materials with higher thermal resistance. These materials may be considered to be "better performing". Such materials are usually more expensive than "less performing".

The following paragraph will discuss the validation process, with the additional simulations and calculations of the identified optima. The two figures below are a repetition of the results in Chapter 7.5.

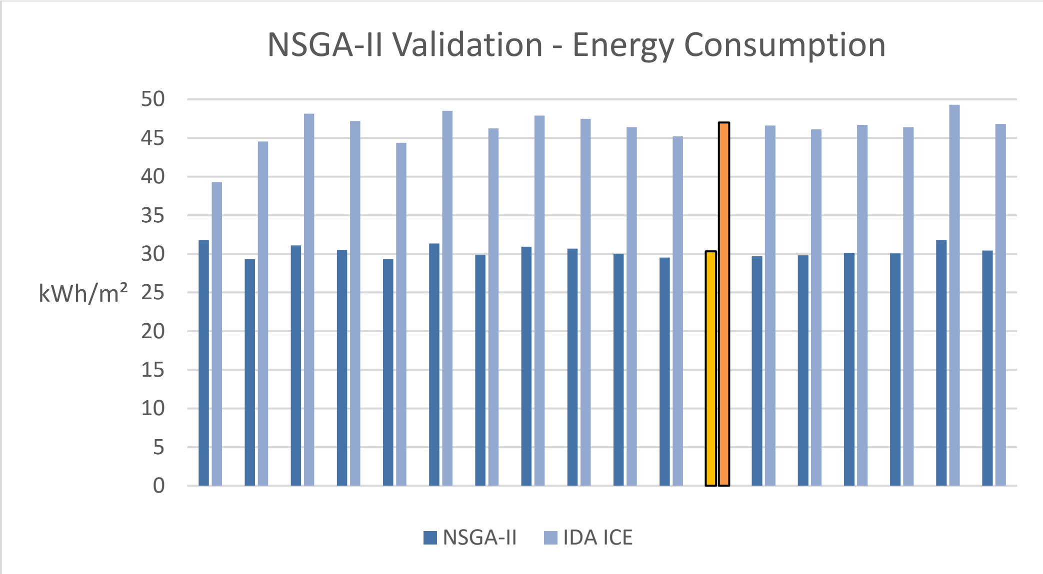


Figure 8.7 NSGA-II validation - Energy Consumption, repeat from Figure 7.12.

This figure clearly indicates weaknesses during the validation process. The chosen optima has an energy consumption value of 30,33 kWh/m^2. By entering the identified parameters in IDA ICE, the simulation results return a value of approximately 47 kWh/m^2. This is a discrepancy of 54% compared to the NSGA-II value.

One can easily interpret that the optimization process identifies a correlation and that the optima values share a common trend with values of approximately 30 kWh/m^2, as may be expected. However, the discrepancy is substantial. The reasons behind this are unclear but most likely an inaccurate ANN model for energy consumption. This is surprising given the accuracy of 0.93393 identified in Chapter 7.4. Another point to consider is the methodology of applying NSGA-II. Currently, it rests on applying the identified number of neurons from a similar model. This methodology has obvious weaknesses. As described earlier, each model is unique.

When inspecting the cost analysis results repeated in the figure below, one may also identify a discrepancy between the calculation and the identified NSGA-II solution. Although not as substantial.

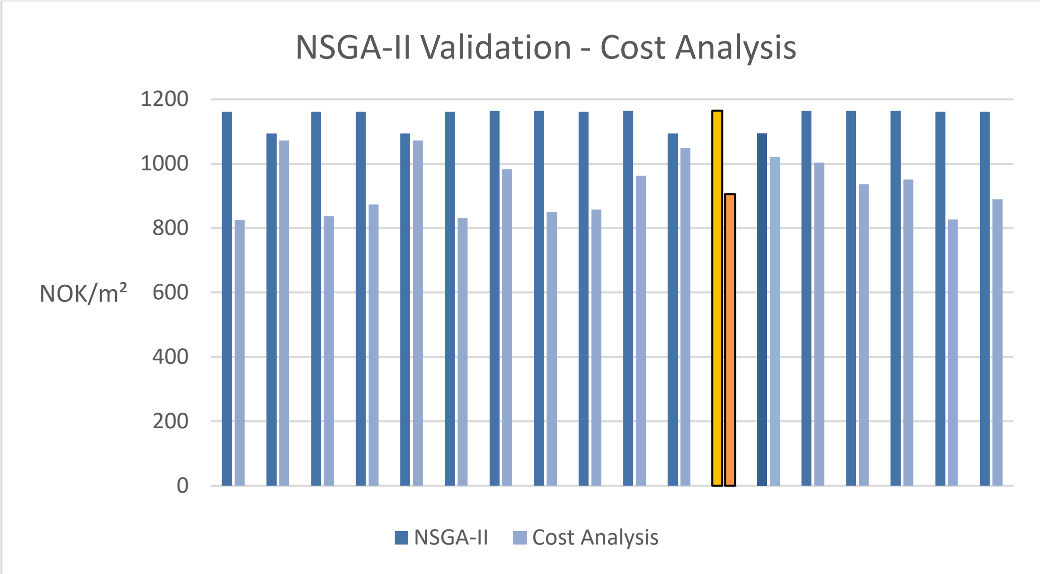


Figure 8.8 NSGA-II validation - Cost Analysis, repeat from Figure 7.13.

From this figure, one may also interpret that the optimization process is not evenly distributed at a given value. In some cases, there is a cost discrepancy of almost 100 NOK/m^2. These results are also surprising. The NSGA-II algorithm should provide an optimized value based on linear regression and should be approximately the same given the Pareto fronts accuracy. One factor that should be considered is that the chosen methodology for creating random variables does not cover the material in the virgin model. The lowest range of insulation for internal walls is defined as 50 mm, whilst the one applied in the virgin model is 30 mm.

Further, when one considers the decision variables displayed in the figure below, a repeat from (Figure 7.11 Optima U-Values)

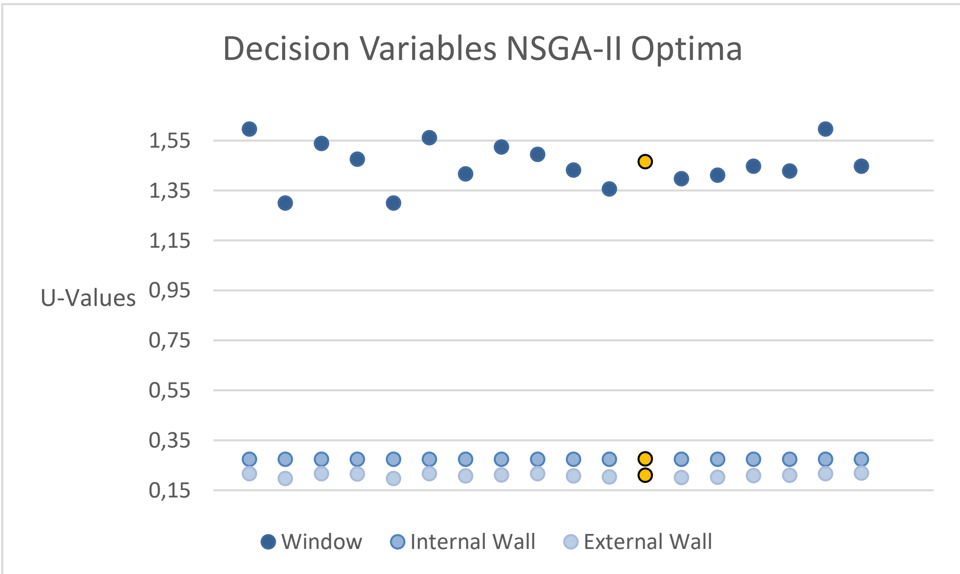


Figure 8.9 Optima U-Values, repeat from Figure 7.11.

One may identify that the main change in performance comes from altering the U-values of the windows. This intuitively appears correct, as windows are essential to energy and cost performance, given the large price and significant effect on heat loss. It also appears correct, given its extensive range of values created by the randomly generated numbers.

However, the thermal transmittance value of internal walls is relatively constant at a low value. Indicating that that optimized value is a well-performing thermal efficient internal walls. This is surprising. Internal walls do not influence heat loss; one may expect this to be the cheapest alternative (insulation of 50mm). However, as described in Chapter 8.6, The identified combination is an insulation of 100 mm, corresponding to a cost of 112,5 NOK/m^2 compared to 50 mm with a cost of 74,71 NOK/m^2. This result indicates an error.

Furthermore, the spread is relatively low, and values appear to be approximately constant at 0,27 W/(m^{2}K). These results are surprising and likely have two possible explanations:

- The case room modelled in IDA suffer severe weaknesses.
- Inaccuracy in the ANN model.

Given the discrepancy in the NSGA-II validation values identified in Figure 7.12 and Figure 7.13, the latter seems likely.

8.7 Digital Twin

In terms of the results, the digital twin illustrates a considerable value of variety. In fact, the given room holds a degree of 35 % optimization in terms of energy and a 5 % reduction in costs.

The figure and table below repeat the results presented in Chapter 8.7 for ease of discussion.

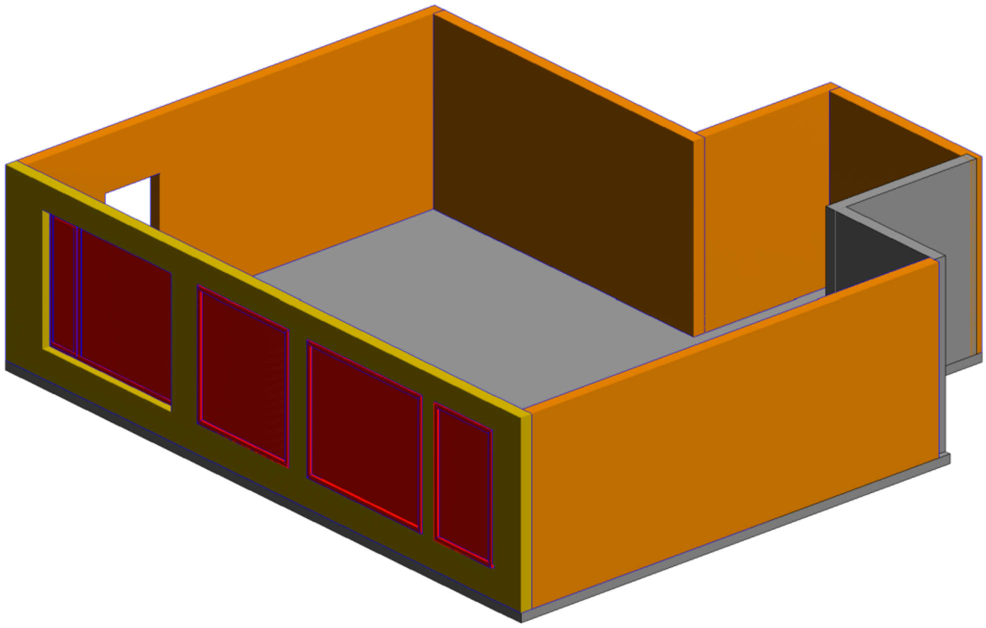


Figure 8.10 Digital Twin, repeated from Figure 7.14.

Table 8.3 Digital Twin Visualization repeated from Table 7.13.

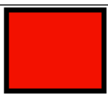
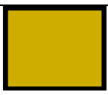
Digital Twin Vizualisation		
Colour	Category	Improvement Potential
	Window	98 %
	External Wall	39 %
	Internal Wall	56 %

Table 8.4 Improvement Potential, repeat from Table 7.11.

Improvement Potential			
Category	External Wall	Window	Internal Wall
Optima	0,21024	1,4656	0,27458
Virgin Analysis	0,151	0,74	0,619
Improvement Potential	39 %	98 %	56 %

When inspecting this model, one may identify that the area with the highest degree of improvement potential is the windows. By further inspecting the identified U-values in the table above, one may determine that the best solution would be to change windows from a window with a U-Value of 0,7 to a window with a U-value of 1,6. External walls should increase their U-Value to 0,2 whilst the U-value of the internal walls should be lowered to 0,27. In essence, lowering the energy efficiency of external walls and windows to benefit internal walls.

As discussed earlier, the results themselves seem counterintuitive. Lowering the U-value refers to a more energy-efficient solution. The optimization process identified a potential savings of energy performance of 16,67 kWh/m^2 and 48,49 NOK/m^{2,} which is a substantial amount.

The results indicate that increasing energy efficiency for inner walls should increase thermal and cost performance. Despite that, there is no theoretical heat loss for internal walls. In fact, previous simulations have proven the heat loss factor for internal walls to equal a total of when comparing a U-Value of 0,1 kWh/m^2.

One may claim, however, that regardless of results, the digital twin serves as a good visualization tool.

8.8 Scripts and Coding

It is essential to mention that the results are interactive based on a limited dataset. This dataset undergoes several different iterations during the scripting process for optimization, which will alter the final results. One may consider the figure of the process displayed below:

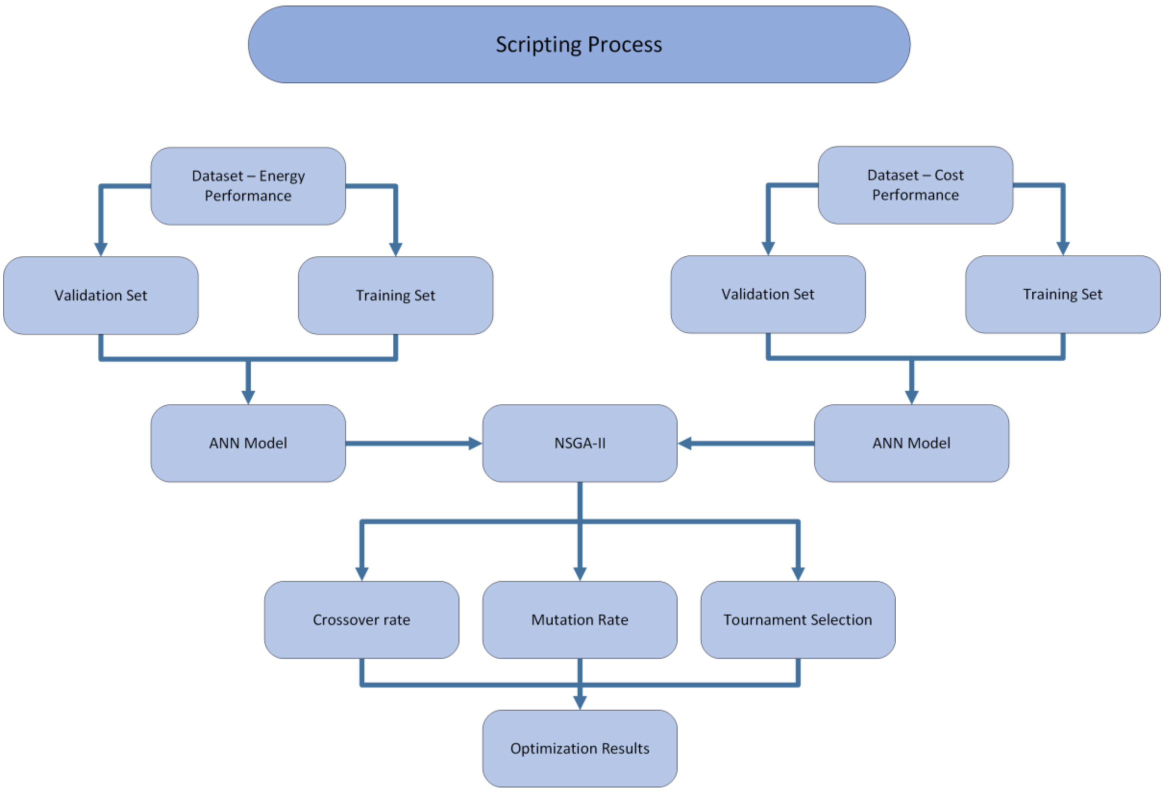


Figure 8.11 Scripting Process.

Firstly, the energy and cost performance datasets are separated into two datasets - training and validation. This process is done arbitrarily based on the 100 combinations. These are then applied to train the different ANN models that serve as input for the NSGA-II algorithm. This algorithm undergoes three processes which further adapt the dataset, the crossover rate, the mutation rate and the results of the tournament selection. This is also based on the predictive values provided by the ANN models.

There are also parameters that may be adjusted in the process, which will alter the final optimization results, such as the degree of distribution from the dataset into validation and training sets. The architecture of the ANN models, in terms of layers and neurons and the activation function. For the NSGA-II algorithm, different values for crossover rate, mutation and crossover type may be altered, as well as ending criteria. All of these parameters will significantly influence the results.

Given this complex hierarchy of dataflow and the arbitrary nature of many of these processes, the results are not fully replicable, as each simulation provides different parameters for the optimization process. In this context, it is essential to emphasize that the final result reflects an *indication* and not a perfectly defined optimal value.

One may argue that it provides a sufficient foundation for effective decision-making. However, the problem remains as the foundation for this decision is not static but up for interpretation. Additionally, the results are not fully replicable.

8.9 Weaknesses

One of the largest weaknesses in this thesis is the original defined dataset. For this thesis, one may identify that the original dataset is skewed. As described in Chapter 8.2 the histogram reveals the same exact same amount of numbers within the various ranges

Another valid concern is the limited methodology with respect to replicating a functional ANN model. The current methodology effectively presumes that two ANN models with the same number of neurons will perform similarly. As identified through the validation processes, this appears to be false.

There is also a concern with the model validation process. The results from the NSGA-II algorithm raise doubt about the accuracy of the IDA ICE model. This is primarily based on the results identified in this thesis: improving energy performance for internal walls will significantly impact the room's energy efficiency. This has no foundation in the existing literature.

Further, the performance of the virgin model is not comparable with simulation results. All cost calculations have returned a higher value than the cost analysis of the virgin model. Only 17 energy simulations have returned an energy consumption larger than the virgin model. Indicating that the original range of thermal transmittance for the randomly generated decision-variables returns design options is unrealistic. Also, in quantifying these variables, only 13 materials are identified.

A clear goal of this thesis was to quantify and analyse as much of the room as possible, which was the motivation for the choice of the case room and methodology of covering internal walls in the energy consumption analysis despite its limited impact on energy performance. However, of the six potential variables presented in Chapter 5.2 that was available for cost and energy performance optimization, only three were chosen, serving as decision variables. As such, it does not cover optimization of the case room as a whole but only design options for external walls, windows and internal walls. Quantifying these variables may be argued to be an essential part of cost and energy performance.

Furthermore, the energy performance of a room depends on more parameters than the limited amount of components defined in this thesis, with effects as mentioned earlier for heat storage. The energy performance is influenced by all factors of the room's envelope, including roofs, floors, walls, internal doors and windows. Another value to consider is the use of the windows, as it drastically affects the heat loss from the given room. In this thesis, the windows are considered constantly closed. The doors are not covered at all. Despite the promising results, one may thus argue that the limitations in methodology serve as a source of insecurity in the optimization process and the final results.

Generally, the weaknesses of this thesis rest on two main concerns. Unsuitable methodology in creating the dataset and improper use of software and coding.

It is worth mentioning that IDA ICE and coding are both very complex themes. Weaknesses in software use and coding may be a result of the limited experience at the start of this thesis.

8.10 Concluding Remarks

In general, the challenge of this thesis is the threshold between purely theoretical concepts and applications physical world.

Identifying costs based on U-Values work well for a single entity such as insulation. However, quantifying costs based on U-values alone would prove inaccurate for more extensive design options. Even by applying the chosen methodology by rounding randomly generated variables to a pre-existing solution, applying this value is not intuitive.

This is primarily due to the combination of materials that make up thermal efficiency. This thesis limits itself to a generic homogeneous layer that is not replicable in real-life situations, as each category contains numerous materials, depending on different solutions. Applying a methodology that attempts to quantify these categories and variables solely on thermal transmittance would suffer severe weaknesses and inaccuracy. However, in such a context, one may consider applying validation calculations. The Digital Twin may help identify contributing elements theoretically, allowing for further calculations.

The findings of this thesis show great promise within the optimisation field, highlighting the possibilities and value of performing optimization algorithms and comparing the effects to other optimized solutions on each other. Although there are weaknesses, the framework itself appears satisfactory.

The framework benefits from its high degree of automation by the scripts and coding applied in this thesis. Consider the figure below:

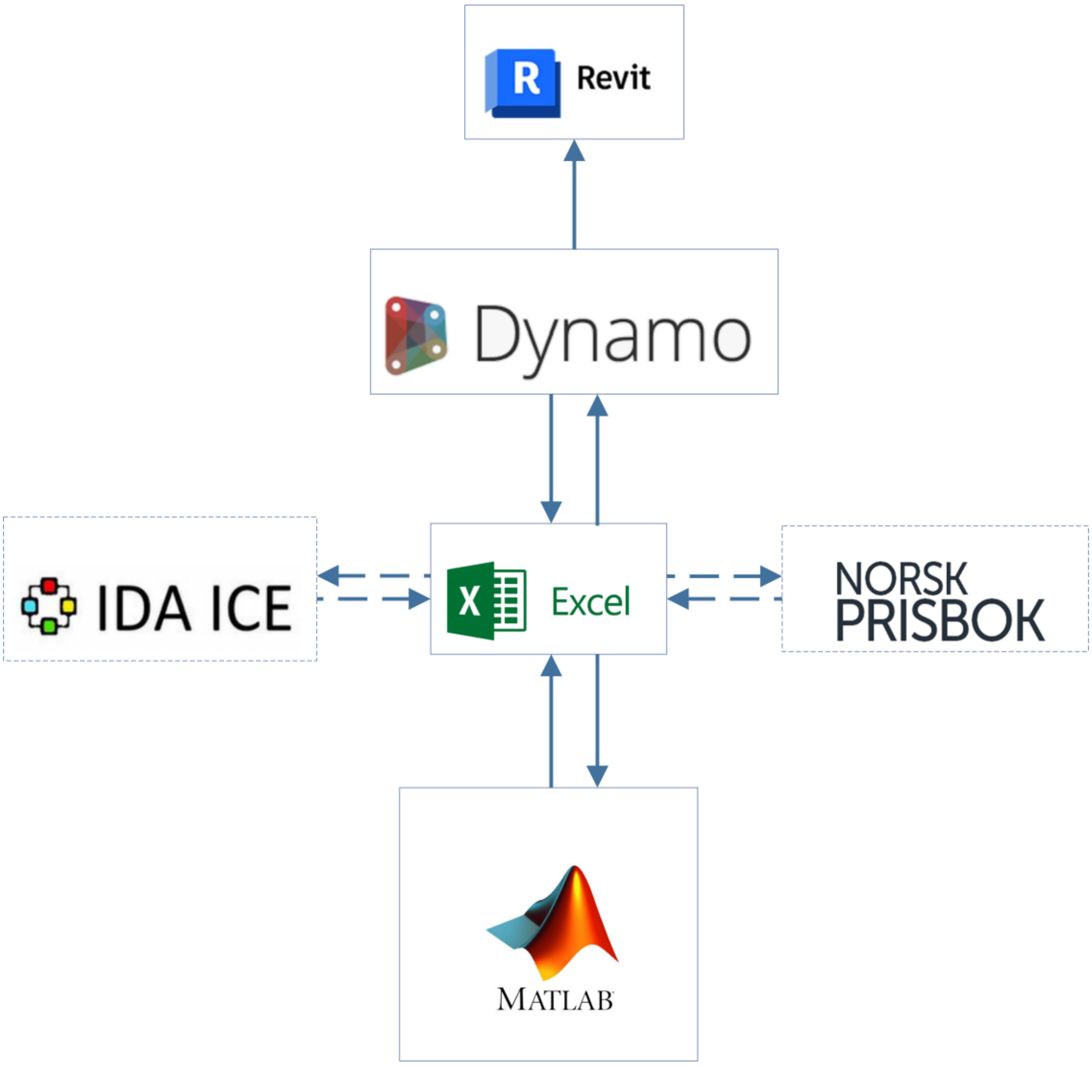


Figure 8.12 Digital Twin Framework.

This figure illustrates the framework itself and the automation process. The dashed lines are illustrative of non-automated processes. And the arrow direction illustrates dataflow. Data from the simulation process is connected to an automated rounding methodology using Excel logic test codes to calculate the corresponding costs for the selected material. This data, combined with the results from the simulation process, is further automated into Matlab using the import data from the Excel sheet itself.

Furthermore, the results of Matlab optimization are exported back to Excel. They are serving as a basis for visualization program in dynamo, streaming optimization degree to the digital twin. The framework loop is closed after data has been gathered. There are, however, weaknesses in the various steps that should be addressed.

The framework may be considered satisfactory for illustrating optimization degrees in an easily visually interpreted model. There is little doubt that this field is of great interest, not only for educational value but also from a practical point of view, as it lays the foundation for effective decision-making.

9 Conclusion

The research question this thesis attempts to answer is:

How can digital twins be implemented to reduce energy consumption and costs in buildings?

Answered by the following sub-questions:

- How effective are different ANN models in predicting building energy performance indicators (e.g., U-value/costs, kWh/costs), and what are the practical implications of using these models in terms of energy efficiency and cost savings?
- How to use NSGA-II to reduce energy consumption and costs?
- How to use visual programming to stream results to BIM?

The following section will describe the conclusion of this thesis, firstly the sub-questions, followed by the main research question.

- How effective are different ANN models in predicting building energy performance indicators (e.g., U-value/costs, kWh/costs), and what are the practical implications of using these models in terms of energy efficiency and cost savings?

The ANN models defined and illustrated in this thesis hold an R-rate of 0,934 for energy consumption and 0,995 for investment cost. The practical implication may be argued to be a theoretical application of regression learning to fit real-life measurements. When applied in NSGA-II, they significantly impact the results of the optimization process.

- How to use NSGA-II to reduce energy consumption and costs?

NSGA-II may be used in combination with ANN models to optimize energy consumption and costs.

- How to use visual programming to stream results to BIM?

Visual programming applied through Dynamo may translate the results from an optimization process to BIM in an easily visually interpreted result.

These combined answer the main research question:
How can digital twins be implemented to reduce energy consumption and costs in buildings?

By stating that "A digital twin may be implemented to translate energy-consumption and cost-optimization into an easily interpreted result that serves as a foundation for efficient decision-making."

By applying the said framework, the results reflect a potential energy consumption reduction of 35 % and a cost reduction of 5 % compared to the original model.

10 Recommendations

Recommendations for future work may be divided into two main strategies, either improving the existing framework presented in this thesis in the existing problem area. Alternatively, expand the framework to cover new areas of interest. This chapter will cover these two strategies.

The existing framework suffers numerous weaknesses described in Chapter 8.9 and should be addressed in future work. One may summarize it as follows:

- Additional decision variables.
- More extensive range of thermal transmittance values for the randomly generated numbers.
- Additional design options in terms of materials.
- Validating the IDA ICE model by comparing it to other analyses and/or verification by a third party.
- Additional simulations for an increased dataset.
- Create a script that exports ANN models to a separate file for future import to the NSGA-II algorithm.

Furthermore, this thesis holds significant limitations as it only focuses on a single room as a case. Expanding the case to include other rooms or even floors will enable design options for the identification of effective means to improve energy consumption and costs on a broader scale. This may have lasting effects, as one can base future design on an already defined database.

For other applications, it is worth stressing that in the context of multi-objective optimization, energy efficiency and cost performance in total form do not form the basis for the operation of the room itself. A room may be both energy-efficient and cost-efficient, but it holds little value if it goes at the expense of the users.

As such, it may be worthwhile to include predicted person dissatisfied (PPD) in the algorithm. These numbers were also extracted from simulations and are included in *Appendix D – Datasheet* under the "IDA ICE SIMULATIONS" sheet. These were not applied in this thesis, given the complexity and limited timeframe.

One may, for example, try to optimize the efficiency of the room in terms of costs and thermal comfort. For instance, problem areas like "How to apply NSGA-II to optimize the relation between cost and comfort?" may be worth exploring.

Another valid application of the framework introduced in this thesis is the application of multi-objective optimization with LCA studies compared to costs. This will help identify design options with the maximum effect on a given room in an easily interpreted model. Problem areas like "How to optimize environmental performance whilst reducing costs using NSGA-II?" may also be worth exploring.

This thesis rests on existing data, essentially showing the value of potential what-ifs and how-tos. However, sufficient application of this methodology to a large number of cases will allow for a dataset depicting cause-effect situations. Meaning, showing value in the design phase of a building itself, granting a large amount of value in terms of occupant satisfaction or environmental impacts compared to price. In essence, how to gain the most value for the least cost.

As described earlier, the presented framework has a wide array of application areas. The challenge lies in identifying common decision variables between two different objectives, which may then be measured by a common denominator transferred further to results for comparison.

11 References

[1] H. Begić, M. Galić and Z. Dolaček-Alduk, "Digitalization and automation in construction project's life-cycle: a review," *Journal of Information Technology in Construction,* pp. 441-460, 22 April 2021.

[2] M. Deng, C. C. Menassa and V. R. Kamat, "From BIM to digital twins: A systematic review of the evolution of intelligent building representations in the AEC-FM industry," *Journal of Information Technology in Construction,* pp. 58-82, March 2021.

[3] K. Wang, F. Guo, C. Zhang and D. Schaefer, "From Industry 4.0 to Construction 4.0: Barriers to the digital transformation of engineering and construction sectors," *Engineering, Construction and Architectural Management,* 31 August 2022.

[4] F. Zhang, A. P.C. Chan, A. Darko, Z. Chen and D. Li, "Integrated applications of building information modelling and artificial intelligence techniques in the AEC/FM industry," *Automation in Construction,* 4 May 2022.

[5] D. Amos, A. P. Chana, M. A. Adabrea, D. J. Edwards and R. M. Hosseinic, "Artificial intelligence in the AEC industry: Scientometric analysis and visualization of research activities," *Automation in Construction,* 20 April 2020.

[6] H. Hosamo, M. S. Solvang Tingstveit, H. Kofoed Nielsen, P. Ragnar Svennevig and K. Svidt, "Multiobjective optimization of building energy consumption and thermal comfort based on integrated BIM framework with machine learning-NSGA II," *Energy & Buildings,* 12 October 2022.

[7] P. Schneider-Marin and W. Lang, "A Temporal Perspective in Eco2 Building Design," *Sustainability,* 16 May 2022.

[8] "The Paris Agreement | UNFCCC," [Online]. Available: https://unfccc.int/process-and-meetings/the-paris-agreement.

[9] C. Cavalliere, G. R. Dell'Osso, F. Favia and M. Lovicario, "BIM-based assessment metrics for the functional flexibility of building designs," *Automation in Construction,* 8 August 2019.

[10] M. Rabani, H. Madessa Bayera, O. Mohseni and N. Nord, "Minimizing delivered energy and life cycle cost using graphical script: An office building retrofitting case," *Applied Energy,* 28 March 2020.

[11] "Management of Energy Sources and the Development Potential in the Energy Production Sector—A Comparison of EU Countries," *Energies,* 21 January 2021.

[12] Y. Lin and W. Yang, "Application of Multi-Objective Genetic Algorithm Based Simulation for Cost-effective Building Energy Efficiency Design and Thermal Comfort Improvement," *frontiers in Energy Research,* 5 April 2018.

[13] United Nations, "Goal 7: Affordable and clean energy | Joint SDG Fund," [Online]. Available: https://www.jointsdgfund.org/sustainable-development-goals/goal-7-affordable-and-clean-energy. [Accessed 5 May 2023].

[14] Global Goals, "https://www.globalgoals.org/resources/," [Online]. Available: https://www.globalgoals.org/resources/. [Accessed 05 May 2023].

[15] J. C. Campasano, K. Smolander and T. Ruippo, “Seven Metaphors to Understand Digital Twins of Built Assets,” *IEEE Access,* pp. 27167-27181, 18 February 2021.

[16] R. Alonso, M. Borras, R. H.E.M Joppelaar, A. Lodigiani, E. Loscos and E. Yöntem, “Sphere: BIM Digital Twin Platform,” *Proceedings,* 22 July 2019.

[17] M. H. Elnabawi, “Building Information Modeling-Based Building Energy Modeling: Investigation of Interoperability and Simulation Results,” *Frontiers in Built Environment,* 3 December 2020.

[18] S. Tanga, D. R. Sheldena, C. M. Eastmana, P. Pishdad-Bozorgib and X. Gaob, “A review of building information modelling (BIM) and the internet of things,” *Automation in Construction,* 2 February 2019.

[19] I. Yitmen, S. Alizadehsalehi, I. Akiner and M. E. Akiner, “An Adapted Model of Cognitive Digital Twins for Building Lifecycle Management,” *Applied Sciences,* 7 May 2021.

[20] M. Nour El-Din, P. F. Pereira, J. P. Martins and M. M. N. Ramos, “Digital Twins for Construction Assets Using BIM,” *Buildings,* 7 December 2022.

[21] H. Naderi and A. Shojaei, “Civil Infrastructure Digital Twins: Multi-level Knowledge Map, Research Gaps, and Future Directions,” *IEEE Access,* 23 November 2022.

[22] I. Yitmen, S. Alizadehsalehi, I. Akıner and M. E. Akıner, “An Adapted Model of Cognitive Digital Twins for Building Lifecycle Management,” *Applied Sciences,* 9 May 2021.

[23] J. Zhao, H. Feng and Q. G. d. S. B. Chen, “Developing a conceptual framework for the application of digital twin technologies to revamp building operations and maintenance processes,” *Journal of Building Engineering,* 13 January 2022.

[24] C. Maureira, H. Pinto, V. Yepes and J. García, “Towards an AEC-AI Industry Optimization Algorithmic Knowledge Mapping: An adaptive Methodology for Macroscopic Conceptual Analysis,” *IEEE Access,* 3 August 2021.

[25] Y. Pan and L. Zhang, “Integrating BIM and AI for Smart Construction Management: Current Status and Future Directions,” *Archives of Computational Methods in Engineering,* 22 October 2022.

[26] Z. Wu, C. Chen, Y. Cai, C. Lu, H. Wang and T. Yu, “BIM-Based Visualization Research in the Construction Industry: A Network Analysis,” *International Journal of Environmental Research and Public Health,* 18 September 2019.

[27] C. Preidel, S. Daum and A. Borrmann, “Data retrieval from building information models based on visual programming,” *Visualization in Engineering,* 11 October 2017.

[28] R. Kabede, A. Moscati, H. Tan and P. Johansson, “Integration of manufacturer's product data in BIM platforms using semantic web technologies,” *Automation in Construction,* 28 October 2022.

[29] N. Neythalath, A. Søndergaard, B. Kumaravel, J. A. Bærentzen and R. Naboni, “Applying software design patterns to graph modelled robotic workflows,” *Automation in Construction,* 15 September 2021.

[30] H. Hosamo, H. Mohsen, H. K. Nielsen, P. R. Svennevig and S. Kjeld, “Digital Twin of HVAC system (HVACDT) for multiobjective optimization of energy consumption and thermal comfort based on BIM framework with ANN-MOGA,” *Advances in Building Energy Research,* 27 September 2022.

[31] G. Desogus, E. Quaquero, G. Rubiu, G. Gatto and C. Perra, "BIM IoT Sensors Integration: A Framework for Consumption and Indoor Conditions Data Monitoring of Existing Buildings," *Sustainability,* 17 April 2021.

[32] A. N. -. S. University, *CS 229: Machine Learning.*

[33] T. C. J. Wortmann and C. Waibel, "Simulation-based optimization in architecture and building engineering- Results from an international user survey in practice and research," *Energy & Buildings,* 22 January 2022.

[34] M. U. o. C. L. A. Bauchy, *Introduction to Artificial Intelligence and Machine Learning for Engineers.*

[35] Mathworks, "Documentation - MATLAB & Simulink - MathWorks Nordic - Help Center," [Online]. Available: https://se.mathworks.com/help/index.html. [Accessed 20 March 2023].

[36] A. Opperman, "Underfitting and Overfitting in Deep Learning | MLearning.ai," 18 July 2021. [Online]. Available: https://medium.com/mlearning-ai/underfitting-and-overfitting-in-deep-learning-687b1b7eb738. [Accessed 1 April 2023].

[37] Simplilearn - Online Certification Training Course Provider, "Mean Squared Error: Overview, Examples, Concepts and more | Simplilearn," [Online]. Available: https://www.simplilearn.com/tutorials/statistics-tutorial/mean-squared-error. [Accessed 28 March 2023].

[38] M. T. Hagan, H. B. Demuth, M. Hudson Beale and O. De Jesús, Neural Network Design, Martin, Hagan, 2014.

[39] OpenAI, "Introducing ChatGPT," 22 November 2022. [Online]. Available: https://openai.com/blog/chatgpt. [Accessed 19 April 2023].

[40] A. Kolides, A. Nawaz, A. Rathor, D. Beeman, M. Hashmi, S. Fatima, D. Berdik, M. Al-Ayyoub and Y. Jararweh, "Artificial intelligence foundation and pre-trained models: Fundamentals, applications, opportunities, and social impacts," *Simulation Modelling Practice and Theory,* 22 March 2023.

[41] M. Azima and S. Seyis, "Science mapping the knowledge domain of energy performance research in the AEC industry: A scientometric analysis," *Energy,* 3 February 2023.

[42] S. Kim, J. Seo, H. Joeng and K. Jonghun, "In situ measurement of the heat loss coefficient of thermal bridges in a building envelope," *Energy & Buildings,* 18 January 2022.

[43] The European Committee for Standardization, *ISO 7345:2018 - Thermal performance of buildings and building components - Physical quantities and definitions,* 2018.

[44] European Commitee for Standardization, *EN ISO 6946 - Building components and building elements - Thermal resistance and thermal transmittance - Calculation Methods (ISO 6946:2017),* European Standard, 2017.

[45] A. Urzędowski, Wójcicka-Migasiuk, Dorota and J. Styczeń, "Analysis of Thermal Properties and Heatloss in Construction and Isothermal Materials of Multilayer Building walls," *Advances in Science and Technology Research Journal,* 2 June 2017.

[46] F. Wang, Z. Chen and C. Wu, "Modeling and estimation of thermal conductivity of ultrafine glass wool mats by artificial neural network and correlation," *The Journal of the Textile Institute,* 19 May 2019.

[47] N. Shukla, A. Fallahi and J. Kosny, "Aerogel Thermal Insulation — Technology Review and Cost Study for Building Enclosure Applications," *ASHRAE Transactions,* 2014.

[48] Direktoratet for Byggkvalitet, "Byggteknisk Forskrift TEK10".

[49] European Commitee for Standardization, "EN 15221-6 - Facility Management - Part 6: Area and Space Measurement in Facility Management," 2011.

[50] European Committee for Standardization, *NS-EN 16627:2015 Sustainability of construction works, Assessment of economic performance of buildings, Calculation Methods.*

[51] S. Li, L. Liu and C. Peng, "A Review of Performance-Oriented Architectural Design and Optimizations in the Context of Sustainability: Dividends and Challenges," *Sustainability,* 14 February 2020.

[52] S. Attia, M. Hamdy, W. O'Brien and S. Carlucci, "Assessing gaps and needs for integrating building performance optimization tools in net zero energy buildings design," *Energy and Buildings,* 19 January 2013.

[53] D. Zhuang, X. Zhang, Y. Lu, C. Wang, X. Jin, X. Hou and X. Shi, "A performance data integrated BIM framework of building life-cycle energy efficiency and environmental optimization design," *Automation in Construction,* 11 April 2021.

[54] M. El Menshawy and M. Marzouk, "Automated BIM schedule generation approach for solving time-cost trade-off problems," *Engineering, Construction and Architectural Management,* 28 December 2021.

[55] Y. An, H. Li, T. Su and Y. Wang, "Determining Uncertainties in AI Applications in AEC Sector and their corresponding Migitation Strategies," *Automation in Construction,* 16 August 2021.

[56] Agder Fylkeskommune, "Tvedestrand Videregående skole - Agder Fylkeskommune," Agder Fylkeskommune, 2020. [Online]. Available: https://agderfk.no/vare-tjenester/utbygging/eiendom-og-skolebygg/tvedestrand-videregaende-skole/. [Accessed 20 February 2023].

[57] Byggeindustrien, "Tvedestrand vgs og idrettspark," Bygg.no, 30 June 2020. [Online]. Available: https://www.bygg.no/tvedestrand-vgs-og-idrettspark/1437910!/. [Accessed 20 February 2023].

[58] A. Jortveit, "Tvedestrand: En miks av solceller, biogass, flisfyr, energibrønner og brukte bilbatterier," 12 December 2022. [Online]. Available: https://energiogklima.no/nyhet/tvedestrand-en-miks-av-solceller-biogass-flisfyr-energibronner-og-brukte-bilbatterier/. [Accessed 20 February 2023].

[59] H. Hosami, H. K. Nielsen, D. Kraniotis, P. R. Svennevig and K. Svidt, "Digital Twin framework for automated fault source detection and prediction for comfort performance evaluation of existing non-residential Norwegian Buildings," *Energy & Buildings,* 17 December 2022.

[60] C. I. o. T. Yaser Abu-Mostafa, *Learning from Data - Introductory Machine Learning Course.*

[61] T. U. o. M. Matthias Niessner, *Introduction to Deep Learning.*

[62] I. I. o. T. Biswajit Mahanty, *Selected Topics in Decision Modeling.*

[63] Yarpiz, *Introduction to Genetic Algorithms.*

[64] Microsoft, "Flowchart Maker and Diagramming Software | Microsoft Visio," [Online]. Available: https://www.microsoft.com/en/microsoft-365/visio/flowchart-software.

[65] "GIMP - GNU Image Manipulation Program," [Online]. Available: https://www.gimp.org/. [Accessed 14 February 2023].

[66] EQUA, "IDA ICE Getting Started - ICE Getting Started," January 2018. [Online]. Available: https://www.equa.se/deliv/ICE4GettingStartedEng.pdf. [Accessed 02 February 2023].

[67] LinkedIn, "LinkedIn Learning: Online Courses for Creative, Technology, Business Skills," [Online]. [Accessed March 2023].

[68] Norconsult Informasjonssystemer AS, "Norsk Prisbok," 2022. [Online]. Available: https://www.norskprisbok.no/. [Accessed 8 March 2023].

[69] SINTEF, "524.325 Lydisolasjon for innerbegger av bindingsverk - Byggforskserien," April 2020. [Online]. Available: https://www.byggforsk.no/dokument/2956/lydisolasjonsegenskaper_til_lette_innervegger. [Accessed 2 February 2023].

[70] "Mathworks - Makers of MATLAB and Simulink - MATLAB & Simulink," 2023. [Online]. Available: https://se.mathworks.com/. [Accessed 28 February 2023].

[71] S. D. Dao, "Solving Optimization Problems Archives -Learn with Panda," [Online]. Available: https://learnwithpanda.com/. [Accessed 24 March 2023].

12 Appendices

Appendix A – TVGS.rvt
Appendix B – Logbook for literature search.pdf
Appendix C – Random Number Generator.dyn
Appendix D – Datasheet.xlsxl
Appendix E – Digital Twin Script.dyn
Appendix F – Coloring.dyf
Appendix G – Digital Twin.rvt
Appendix H – IDA ICE SIM.idm
Appendix I – Matlab_Code.mlx
Appendix J – Virgin Analysis.xlsx
Appendix K – A3 Poster.pdf