



UNIVERSITÀ DEGLI STUDI DI PALERMO

PhD in Energy - Low Carbon Energetics and Innovative Nuclear Systems
Dipartimento di Ingegneria
SSD IIND-07/B Fisica Tecnica Ambientale

DIGITAL TWINS OF SMART ENERGY SYSTEMS: MODELS AND OPTIMIZATION ALGORITHMS

PHD CANDIDATE
TANCREDI TESTASECCA

PHD COURSE COORDINATOR
PROF. ELEONORA RIVA SANSEVERINO

SUPERVISORS
PROF. GIUSEPPINA CIULLA
PROF. MARCO BECCALI

CO-SUPERVISORS
ENG. MARCO FERRARO
ENG. MARILENA LAZZARO
ENG. DIEGO ARNONE

CYCLE XXXVII
YEAR OF QUALIFICATION 2024

Table of Content

Abstract	4
Chapter 1 Introduction	5
<i>References</i>	<i>10</i>
Chapter 2 Digital Twins of Buildings and Energy Systems	12
2.1 Digital Twins Architecture	13
2.2 From Digital Models to Digital Twins	14
2.2.1 Kritzinger Classification	14
2.2.2 Digital Twin Consortium Classification	15
2.3 Digital Twins Maturity Level	16
2.3.1 Institution of Engineering and Technology's Approach	16
2.3.2 Cooper's Approach	18
2.4 Buildings Digital Twins	19
2.4.1 Digital Twins to Balance Indoor Comfort and Improve Energy Efficiency	20
2.4.1 Digital Twins to Detect Faults and Enhance Building Monitoring Capabilities	21
2.5 Digital Twins of Smart Energy Networks	23
2.5.1 Digital Twin for Real Time Management of Smart Grids	25
2.5.1 Digital Twin for Cost and Energy Optimization of Smart Grids	26
2.5.2 Digital Twins to Improve District Heating Networks Design and Operation	27
2.6 Digital Twins of Energy Systems Components	28
2.6.1 Physics-Based Digital Twins	29
2.6.2 Artificial Intelligence-Based Digital Twins	30
2.7 Discussion and Limits of Current Research about DTs	30
<i>References</i>	<i>32</i>
Chapter 3 Methodology for Developing Digital Twins	41
3.1 Data Integration and Processing Strategies	41
3.2 Data-Driven Models	46
3.2.1 Supervised Machine Learning Algorithms	46
3.2.2 Supervised Deep Learning Algorithms	48
3.2.3 Supervised Algorithms Validation Methods	49
3.2.3 Unsupervised Algorithms	51
3.3 Physics-Based Modelling	52
3.3.1 Hydraulic Pipe – Type 496	53
3.3.2 Absorbption Chiller	54
3.3.3 TRNSYS DHN Model Components	56
3.4 Optimization Algorithms for Digital Twins	59
3.4.1 Fundamentals of Optimization Problems Theory	59
3.4.2 Optimization Algorithms	61
3.4.3 Optimization Algorithms in Python	63
3.5 Graphic User Interface	65
<i>References</i>	<i>69</i>
Chapter 4 Digital Twins and Models of District Heating Networks	74
4.1 Overview of District Heating Networks	74
4.2 Digital Twin of Rio Vena Third Generation Network	76

4.2.1 Methodology	76
4.2.2 Case Study	80
4.2.3 Results and Findings	82
4.3 Digital Model of a Fourth Generation District Heating Network	90
4.3.1 Methodology	90
4.3.2 Case Study	93
4.3.3 Results and Findings	97
4.3 Discussion on energy savings and environmental impact	106
4.4 Digital Model of a Fifth Generation District Heating Network	107
4.4.1 Methodology	108
4.4.2 Case Study	111
4.4.3 Results and Findings	115
4.5 Final Discussion and Future Directions	123
References	127
Chapter 5 Digital Twin of a Microgrid	131
5.1 Overview of Smart Grids Optimal Management through Electric Vehicles	131
5.2 Digital Twin of Emotion Microgrid	133
5.2.1 Digital Twin Development Methodology	134
5.2.2 Case Study and Results	140
5.3 Final Discussion and Future Directions	150
References	152
Chapter 6 Digital Twin of a Solid Oxide Fuel Cell	154
6.1 Overview of Solid Oxide Fuel Cells	154
6.2 Data-Driven Models for Solid Oxide Fuel Cell Digital Twin	155
6.2.1 Methodology	156
6.2.2 Case Study and Framework	158
6.2.3 Results and Findings	160
6.3 Final Discussion and Future Directions	165
References	167
Chapter 7 Buildings Digital Twins and Models	170
7.1 Overview of Thermal and Visual Comfort in Buildings	171
7.2 SolarLab Case Study	172
7.2.1 SolarLab Energy Model	174
7.3 Digital Twin of SolarLab to Improve Visual Comfort	175
7.3.1 Methodology	176
7.3.2 Results and Findings	181
7.4 Visual Comfort and Energy Efficiency Evaluation Using Shadings	182
7.4.1 Methodology, Models and Scenarios	182
7.4.2 Results and Findings	187
7.4.3 General Discussion	195
7.5 A Methodology to Enhance Thermal Comfort Monitoring in Buildings	196
7.5.1 Methodology	197
7.5.2 Results and Findings	201
7.6 Final Discussion and Future Directions	207
References	210
Chapter 8 Conclusions	214
List of Figures	219

List of Tables	222
List of Research Products	224
Appendix	227
<i>List of Acronyms</i>	227
Acknowledgements	230

Abstract

In 2022, urban areas in Europe accounted for 80% of energy consumption [1], and, as the global population grows and living standards improve, the worldwide energy demand will not diminish in the coming years. Renewable Energy Sources (RES) are part of the solution toward decarbonisation, however, due to the irregular availability of renewable energy from solar and wind, achieving a balanced and resilient energy system is a challenge. Therefore, there is a critical need for innovative methods and models to create an economically sustainable energy sector which includes Smart Energy Systems (SESS) where various energy sources and consumers are integrated. For these purposes, The European Union actively promotes the digitalisation of these systems via various initiatives [2], and Digital Twins (DTs) represent the opportunity to combine scientific progress with real-world application. Using these synchronised replicas, operators may acquire real-time simulation results by monitoring the system. This would allow to determine the most optimal actions for improving system operation and safety, as well as reducing costs and emissions.

The focus of this PhD research project is the development of DT-oriented algorithms to model and optimize different energy systems. The current state-of-the-art (SoA) of DTs in the energy field is presented, emphasizing the applications of this technology for various objectives and describing how to assess the maturity of a DT or its differences from other models. All the methodologies used are then explained, including data-driven and physics-based modelling, optimization algorithms, data processing, and visualization. Specifically, studies on District Heating Networks (DHNs), microgrids, Solid Oxide Fuel Cells (SOFCs) and buildings were conducted, detailing each methodology and the results in distinct chapters. These studies address various gaps in the existing literature, demonstrating the benefits and potential of DT technology. For instance, simulations on fourth and fifth-generation DHNs revealed significant energy and emissions reductions compared to third-generation networks or autonomous consumers. The twin of Rio Vena DHN, a third-generation network, confirmed that integrated optimization of boilers could reduce the environmental impact even when operating with natural gas. In the context of microgrid optimization, data from the Emotion facilities case study enabled the integration of a DT capable of maximizing Photovoltaic (PV) self-consumption by optimizing Electric Vehicles (EVs) charging and providing real-time suggestions. Therefore, by analysing over six months of data of a SOFC provided by the CNR, an "ageing" DT was developed to maintain an accurate Machine Learning (ML) model for optimizing the operation of four different SOFCs in a larger context. Finally, different studies on SolarLab laboratory were conducted to create a building DT capable of enhancing indoor comfort.

Chapter 1

Introduction

International organisations such as the International Energy Agency (IEA) and the European Union (EU) promote the development of a solid and efficient energy infrastructure. Based on the IEA's report [3], digitalization has the potential to enhance decarbonisation, resilience, and efficiency by utilising demand-side flexibility and integrating and optimising RES through the use of Internet of Things (IoT) technology. Contemporaneously, in October 2022, the EU adopted an action plan [4] that prioritises the digitalization of energy systems to reach the target of a 45% share of renewable energy and a 55% decrease in greenhouse gas (GHG) emissions by 2030.

The increase in energy prices, intensified by global conflicts, has highlighted the urgent requirement for reliable techniques. SESs enable optimal management techniques, often focused on diminishing energy use or costs, through the utilisation of smart materials, devices, or technology [5]. Hence, as a result of the reduced expenses associated with sensors or actuators, digitalization is increasingly involving both energy users (such as companies and buildings) and producers. DTs are a promising technology that will aid in the green transition of SESs by following the path towards digitization. Within this approach, DTs serve as digital representations of tangible entities, such as objects, systems, or processes, which provide the continuous monitoring, analysis, and enhancement of their performance. This technology may facilitate the integration of Distributed Energy Sources (DESS) and EVs into the power grid, [6]. Additionally, DTs have been previously utilised to optimise smart grids [7], improve fault detection [8], and boost monitoring capabilities [9]. Having a digital copy of a real system enables building managers and energy operators to make well-informed decisions throughout both the design and operational phases by using predictive models or optimization algorithms.

These digital replicas, synchronized with a real system, were first used in the manufacturing domain to enhance the efficiency and economic sustainability of various processes and, just recently, it was extended to energy applications. Over the past three years, publications on DTs in the energy field have grown almost exponentially, underscoring the increasing interest and importance of this research. Several Italian and European research projects, such as NEST [10],

DigiBUILD [11], and Dedalus [12], are involved in developing DTs. It must be specified that this industrial PhD was co-financed by the “Consiglio Nazionale delle Ricerche” (CNR) of Italy, ITAE section, and the company Engineering Ingegneria Informatica S.p.A. This collaboration between these two entities and the University of Palermo facilitated a research path oriented towards developing DTs validated through real case studies. Part of the studies were conducted in the framework of the DigiBUILD European project, grant agreement no. 101069658, coordinated by Engineering. One of the aims of this project is to apply optimization and efficient management methods to SESs by the use of innovative data-driven services in different contexts (Fig. 1).

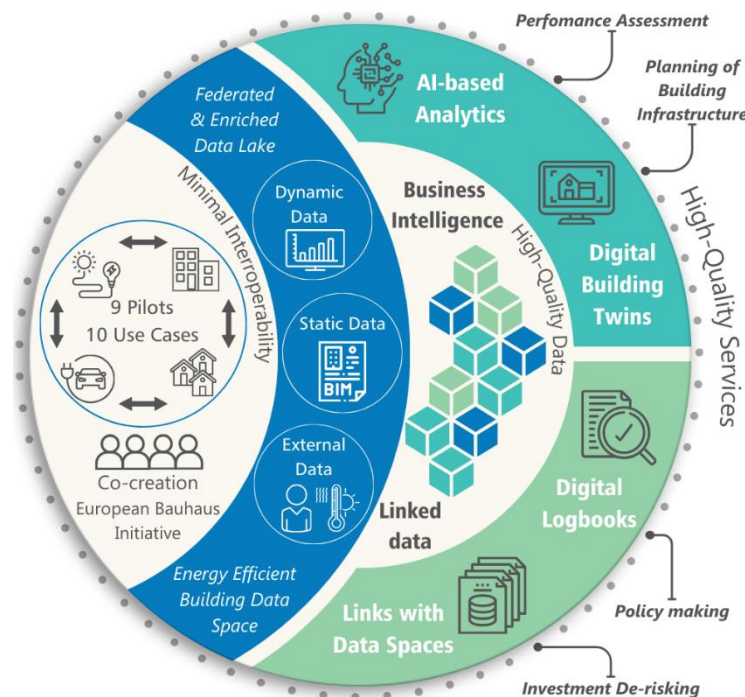


Fig. 1. Summary Schema of the DigiBUILD Project [11]

Within the project, eight demonstrators, referred to as “pilots”, implement algorithms and models developed by technical partners such as universities (e.g., University College London, National Technical University of Athens, Università Politecnica delle Marche) and research groups such as Cartif (Spain) and Engineering S.p.A. These data-driven models utilize historical or real-time data from the pilots to optimize energy consumption, reduce carbon footprints, improve comfort, and enhance real-time operation. Specifically, the DTs of the pilots presented in this thesis include “Emotion”, an Italian company that produces charging stations for EVs, and Veolia-Rio Vena, a third-generation DHN located in Burgos (Spain). In the former, optimization algorithms and predictive models of EV consumption are utilized to maximize PV self-consumption while accommodating user needs. In the latter, data-driven models of boiler efficiency were developed and integrated into an optimization algorithm to suggest the optimal

boiler management strategy for cost and emission reduction. During this PhD, the participation in the project led to the writing of contributions in different technical deliverables such as D3.1 [13] whose aim was to present the status of the architecture and interdependencies between the models (services) within the Work Package (WP) 3. Then, the activities on the development and definitions of DTs in charge of Engineering S.p.a. were presented in two different progressive deliverables D4.1 [14] and D4.2 [15]. The implementation on the Pilots of both models and DT are discussed in D5.3 [16] which is a deliverable Pilots-oriented as the main goal of the WP5. Finally, the models implementation in the beta and final release are reported respectively in D3.2 [17], published in December 2023, and D3.3 (to be submitted at the end of October).

Before the collaboration within the DigiBUILD project, which started in December 2022, research activity on DHNs was conducted, focusing on fourth and fifth-generation network models developed using TRNSYS. As reported by the IEA in [18], DHN technology is anticipated to expand in the coming years, with fourth and fifth-generation DHNs representing the most advanced network types. Historically, traditional networks operated with pressurized hot water above 90 °C. These novel networks operate at temperatures below 60 °C (4th Generation) or below 30 °C (5th Generation) to minimize heat losses and easily integrate heat supplied by RESs. The former can directly supply heating and cooling to users, while the latter requires substations with heat pumps to deliver either cool or hot water. However, operating at nearly ambient temperatures allows for minimal heat losses and simultaneous heating and cooling operations. For completeness, this thesis investigates these two generations through two distinct case studies, including prosumers and solar cooling technology.

In the current literature, DTs are also frequently developed for single components, which may be more feasible than installing a large array of sensors for larger energy systems, thus reducing the costs and emissions of the entire system. SOFCs are a type of fuel cell suitable for both heat and electricity production due to their high operating temperatures and, in addition to hydrogen, can also be fuelled by direct air and other fuels, including biogas or natural gas. In this context, research on an existing SOFC provided by CNR was undertaken to develop a DT capable of “aging” and calculating optimal management strategies.

This thesis aims to demonstrate how DT technologies, through the development of actual DTs or advanced models, can support decarbonization efforts in alignment with the EU's digitalization objectives [4]. When a sensor infrastructure was available during the PhD, real DTs were developed to reduce gas consumption and emissions in a DHN. Additionally, DTs were used to maximize PV self-consumption in a microgrid and improve visual comfort in the SolarLab laboratory at the University of Palermo. For other cases, such as the SOFC system

and the SolarLab, where sensor networks were not yet fully in place, models were used to form the foundation for self-evolving DTs as well as the architecture for the SOFC's DT. For the laboratory, while awaiting new sensor installations, model simulations were conducted. Algorithms were developed for improving indoor comfort using novel methodologies for monitoring thermal comfort parameters and for proposing optimal shading control strategies that could be easily implemented in real building management systems. Lastly, the models of fourth and fifth-generation DHNs, though based on hypothetical case studies, provide valuable insights for effectively managing new-generation DHNs with a high integration of RESs.

The structure of this PhD thesis follows the progression of the research activities conducted over the three years. After a brief introduction in this first chapter, **Chapter 2** reviews the current SoA of DTs in the energy domain. Given the novelty of this technology, four definitions of DTs are initially presented, followed by a discussion on the differences between a DT and traditional models or dashboards according to two classification methods. Once the concept of maturity levels for DTs is introduced, applications of DTs are presented for buildings and SESs.

Chapter 3 focuses on methodologies for developing the DTs of this research project. It was necessary to present a general overview of all the modelling methods. These include both data-driven approaches (e.g., Artificial Neural Networks, Random Forest, Gradient Boosting) and physics-based modelling using TRNSYS software [19]. Additionally, insights about data processing and visualisation through dashboards were also provided.

Subsequent chapters focus on different applications of DTs or Digital Models (DMs) developed during this PhD. DHNs were the first and one of the most significant fields of research studied, and for this reason, **Chapter 4** presents the results of three different case studies of DHNs. The first Case Study is the DT of the Rio Vena DHN where the twin is used to optimise boiler efficiency by suggesting strategies for fine-tuning the boilers setpoints. ML models and optimization algorithms are involved to reach this goal. The second case study is a ring-shaped 4th generation network model developed in TRNSYS. The thermohydraulic behaviour, as well as potential energy and environmental savings, were investigated. Finally, scenarios incorporating prosumers were compared with those using fourth- and third-generation networks to evaluate their performance. The third case study focuses on a Solar District Heating and Cooling Network capable of providing heating and cooling even to a fifth-generation network. This application required the modelling of substations with heat pumps and was later compared with a 4th generation Solar DHN and a baseline scenario.

Chapter 5 discusses the microgrid of the Emotion case study, one of the most common energy systems for which DTs are developed. Following a brief overview of Smart Grids and Microgrids in the context of DTs, the chapter focuses on the microgrid itself, presenting the

methodology used. The ML models of EVs and the optimization algorithms were described accompanied by the initial results achieved within the DigiBUILD project. Precisely, the DT of Emotion is used to minimize interactions with the electric grid through accurate management of EV charging, considering user preferences. Furthermore, this study was conducted in collaboration with partners from the National Technical University of Athens throughout the entire project. This collaboration intensified during the period abroad in Athens from June 2023 to January 2024, part of which was conducted online.

Toward the integration within an energy hub that includes PV systems, an electrolyser, and smart buildings, **Chapter 6** introduces the development of the DT of a SOFC. Having a DT of this component enables a constantly updated model of the SOFC, capable of estimating its efficiency based solely on current electric power generation and ramp power increment. Additionally, an algorithm for optimising the operation of 4 SOFCs in parallel to supply the building loads was developed to partially meet the future research objectives.

In **Chapter 7**, the potential of implementing a building DT will be explored, with a focus on applications in the University of Palermo's SolarLab. By developing various models and utilizing the existing sensor infrastructure, investigations into enhancing visual and thermal comfort, even in real-time, were conducted. Specifically, a DT for monitoring and providing suggestions to improve visual comfort was developed. Additionally, different strategies were examined to evaluate the effects of advanced control for shading systems on energy consumption, as well as visual and thermal comfort. Finally, a novel methodology for estimating local Mean Radiant Temperature (MRT) and Predicted Mean Vote (PMV) based on ML algorithms and view factor calculations is presented.

Chapter 8, the Conclusions, aims to set the point of the presented thesis, summarizing the findings of each application investigated. It discusses the challenges encountered in developing such systems and suggests novel approaches or guidelines for other researchers in this field. Finally, it outlines the future steps for this research project, providing additional information on its potential developments.

References

- [1] European Union, “Inforegio - Sustainable urban development.” Accessed: Mar. 30, 2023. [Online]. Available: https://ec.europa.eu/regional_policy/policy/themes/urban-development_en
- [2] European Union, “EU initiatives for smart energy systems.” Accessed: Mar. 30, 2023. [Online]. Available: https://energy.ec.europa.eu/topics/research-and-technology/eu-initiatives-smart-energy-systems_en
- [3] International Energy Agency, “Digitalization & Energy,” 2017. Accessed: Mar. 24, 2023. [Online]. Available: <https://www.iea.org/reports/digitalisation-and-energy>
- [4] European Commission, “Digitalising the energy system - EU action plan,” 2022. doi: 10.2833/492070.
- [5] T. Testasecca, S. Stamatopoulos, M. Lazzaro, and E. Sarvas, “Recent advances on data-driven services for smart energy systems optimization and pro-active management,” in *2023 IEEE International Workshop on Metrology for Living Environment (MetroLivEnv)*, 2023. doi: 10.1109/MetroLivEnv56897.2023.10164056.
- [6] P. Skaloumpakas *et al.*, “A Multi-Criteria Approach for Optimizing the Placement of Electric Vehicle Charging Stations in Highways,” *Energies (Basel)*, vol. 15, no. 24, Dec. 2022, doi: 10.3390/en15249445.
- [7] M. Jafari, A. Kavousi-Fard, T. Chen, and M. Karimi, “A Review on Digital Twin Technology in Smart Grid, Transportation System and Smart City: Challenges and Future,” 2023, *Institute of Electrical and Electronics Engineers Inc.* doi: 10.1109/ACCESS.2023.3241588.
- [8] Z. Yuan *et al.*, “Fault data generation of lithium ion batteries based on digital twin: A case for internal short circuit,” *J Energy Storage*, vol. 64, Aug. 2023, doi: 10.1016/j.est.2023.107113.
- [9] H. Edtmayer *et al.*, “Modelling virtual sensors for real-time indoor comfort control,” *Journal of Building Engineering*, vol. 67, May 2023, doi: 10.1016/j.job.2023.106040.
- [10] “NEST – Network 4 Energy Sustainable Transition.” Accessed: Jun. 27, 2024. [Online]. Available: <https://fondazione nest.it/>
- [11] “DigiBUILD Project - High-Quality Data-Driven Services for a Digital Built Environment towards a Climate-Neutral Building Stock.” Accessed: Jun. 27, 2024. [Online]. Available: <https://digibuild-project.eu/>
- [12] “Dedalus Project - Data-driven Residential Energy Carrier-agnostic Demand Response Tools and Multi-value Services.” Accessed: Jun. 27, 2024. [Online]. Available: Data-

driven Residential Energy Carrier-agnostic Demand Response Tools and Multi-value Services

- [13] E. Sarmas *et al.*, “D3.1: ‘First wave’ of DigiBUILD AI-based data-driven services for the built environment,” 2023.
- [14] Antonino Sirchia *et al.*, “Deliverable 4.1: Digital Twin Suite and cloud-based DigiBUILD data toolbox (Alpha version),” 2023.
- [15] “D4.2: Digital Twin Suite and cloud-based DigiBUILD data toolbox (Beta version),” 2024. [Online]. Available: <https://digibuild-project.eu/>
- [16] Serena Serroni *et al.*, “D5.3: Pilots’ execution documentation – Pre-pilot phase ,” 2023.
- [17] “D3.2: ‘Second wave’ of DigiBUILD AI-based data-driven services for the built environment,” 2024. [Online]. Available: <https://digibuild-project.eu/>
- [18] IEA, “Heating.” [Online]. Available: <https://www.iea.org/reports/heating>
- [19] University of Wisconsin--Madison: Solar Energy Laboratory, “TRNSYS 17, a transient simulation program,” 1975, *Madison, Wis. : The Laboratory, 1975*. Accessed: Oct. 05, 2023. [Online]. Available: <https://www.trnsys.com/>

Chapter 2

Digital Twins of Buildings and Energy Systems

The term "Digital Twin" was first introduced by Grieves in 2003 [1] and later adopted by the National Aeronautics and Space Administration in 2010 [2]. It refers to a simulation of a system or vehicle that is based on probabilistic and physics-based models. This simulation replicates the entire lifespan of the real counterpart by incorporating real-time and historical data. The DT is characterised by its ultra-realistic nature, incorporating high-fidelity physical models, historical sensor data, and real-time observations. According to the more recent definition of Singh et al. [3], a DT is a digital representation or simulation that accurately replicates a physical item, process, or system, including its tangible behaviours and attributes, in real time. The DT is continuously updated with both current and historical data, enabling a self-evolving and dynamic representation of the physical twin. The link between the DT and its physical counterpart is reciprocal, indicating that any modifications made to the DT are mirrored in the physical twin and vice versa. According to this definition, the DT consists of three main components: the physical twin, which is the actual system that faces changes over time; the DM, which updates and modifies itself by receiving real-time data from the physical twin and running simulations; and a linking mechanism, typically comprised of sensors, actuators, and dashboards for user interaction. The domain of DT definitions continues to evolve, with regularly updated and improved definitions. One of the most reliable and up-to-date definitions [4] states that a DT is unique because it is connected to its physical counterpart, setting it apart from other DMs or replicas. This 'linkage' indicates a specific level of connection and correlation between the physical and digital elements. More precisely, the level of development of this twin might differ, spanning from a simple 2D or 3D representation of a particular section to a full and extremely accurate representation of a whole asset, facility, or even a nation. Every component in this model is intricately linked to data from engineering, construction, and operations. Finally, the Digital Twin Consortium (DTC) defines a DT as a virtualized representation of real-world things and activities that is synchronised with a given frequency

and quality [5]. This information emphasises that a twin can be related either to a tangible item, such as a building or a power plant, or simply a process. This definition also asserts that synchronisation is crucial in a genuine DT, which necessitates the monitoring and managing of devices.

2.1 Digital Twins Architecture

Once the DT is defined, it is fundamental to describe the architecture of a generic twin, describing how it should be composed. Amazon Web Services, Microsoft Azure, and IBM, among other information technology companies, have already provided a detailed description of the architecture of DTs [5]. Based on the latest information, as seen in Fig. 2, the primary components of a DT consist of the following layers: the IoT stack, a data layer, an information management layer, modelling and simulation, an analytics layer, visualisation, and a management layer. Collectively, these mechanisms manage the bidirectional transmission of data from hardware and sensors to edge platforms and data centres, considering integration, governance, and security as the main elements. The complete system additionally enables the utilisation of numerous simulation and analytics models, as well as the visualisation of data through a dashboard or virtual and augmented reality technologies.

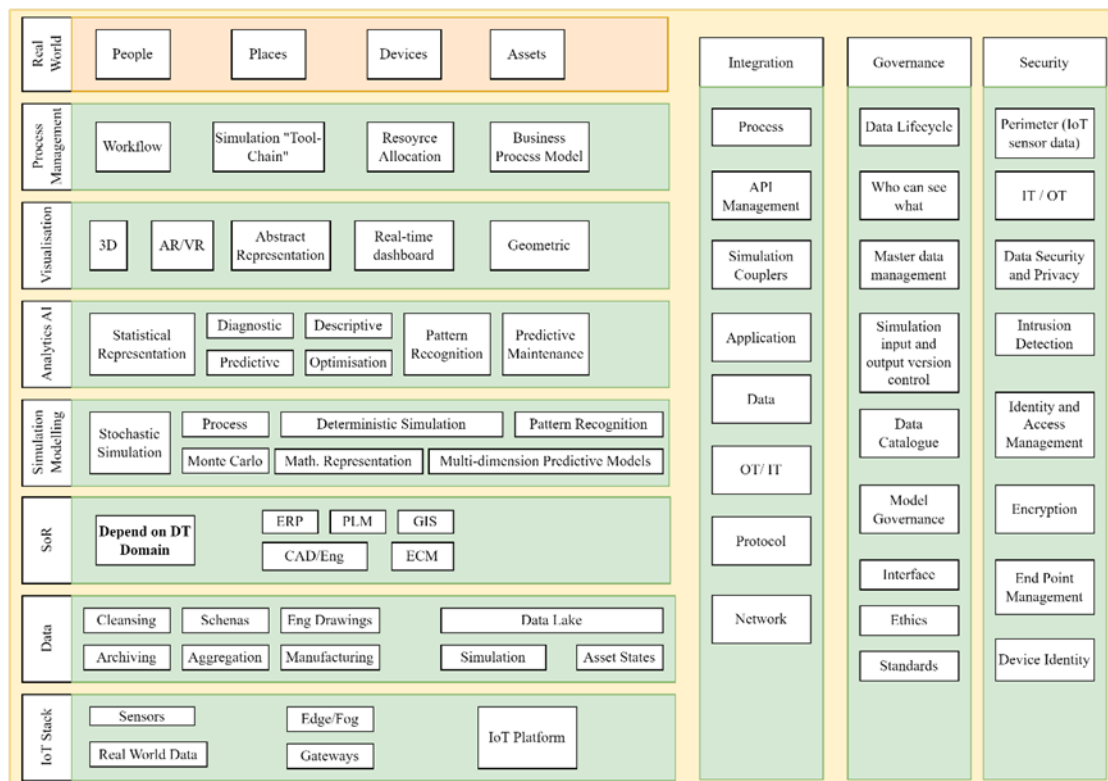


Fig. 2. Example of DT Reference Architecture according to IBM [5].

2.2 From Digital Models to Digital Twins

Fig. 3 presents the publishing patterns of publications relevant to DTs from 2013 to 2023 using Scopus as the research database. In January 2024, an analysis of publishing patterns shows that the growth rate of energy-related DT publications in 2023 surpassed that of general DT publications, which saw a decrease from 2022. Currently, there is a noticeable trend of faster growth in publications related to the energy industry compared to other fields of study. The increase in energy-related publications highlights the importance of having a precise definition to determine if a produced model can be appropriately labelled as a DT or not.

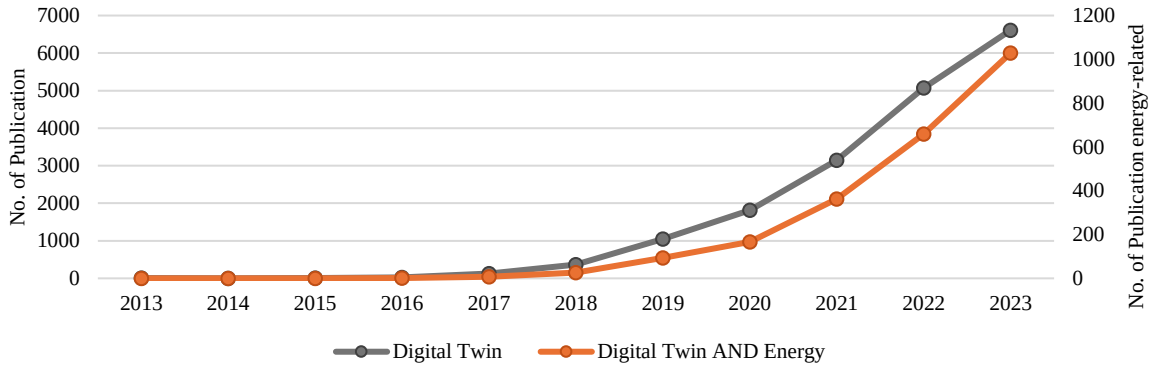


Fig. 3. Paper Publication Trend.

2.2.1 Kritzingner Classification

In 2018, Kritzingner et al. [6] observed a rise in the use of the word DT in scientific literature, often in an improper manner. As a result, they revised the definition of DT, differentiating it from a DM and a Digital Shadow (DS), as seen in Fig. 4. The authors defined the DM as "a digital representation of a physical object, whether existing or planned, that does not involve any automated data exchange between the physical object and the digital object". In this scenario, the model can be either data-driven or physics-based. However, any immediate changes in the physical item do not have an impact on the data model, and vice versa. If the physical system updates the digital object in a one-way interaction, such as in a monitoring system, then this satisfies the definition of a DS. A DT, for the authors, must involve a bidirectional automatic connection between the physical system and the digital object.

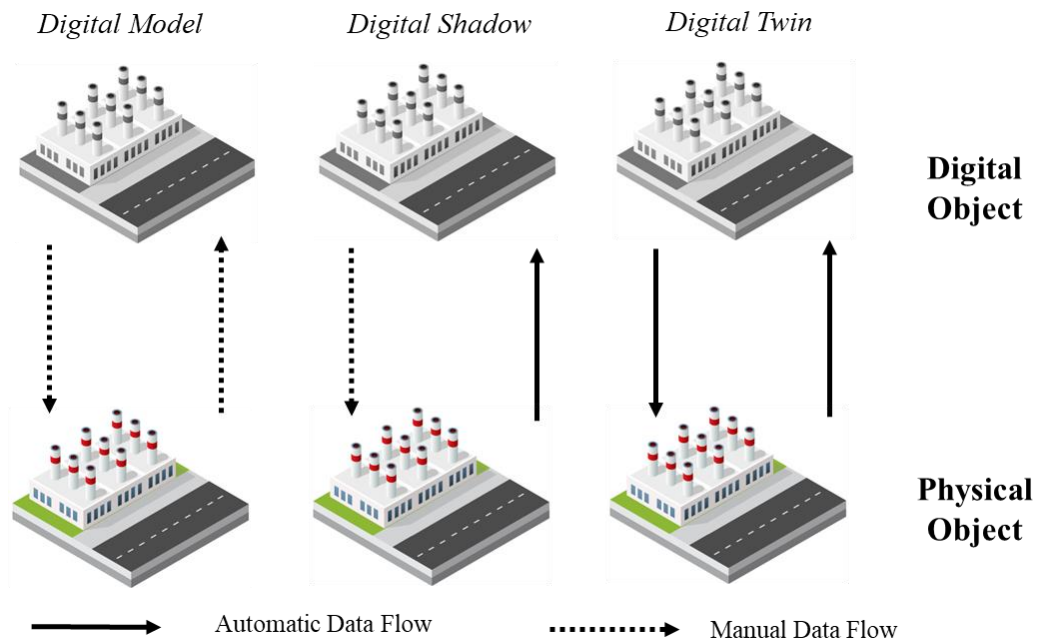


Fig. 4. Kritzinger Classification revised schema [6].

This categorization emphasises the idea that only relying on historical data to improve or establish a model is insufficient to produce a DT, supporting the concepts given in the first chapter. Under such circumstances, the result would be a resilient and precise model with the ability to replicate real-life behaviour and to be connected to the real-world system. Moreover, this categorization emphasises that the presence of a link where the sole digital entity's ability is updated using real-world data does not automatically confer the status of a twin. The essential necessity is a two-way flow of data between the two.

2.2.2 Digital Twin Consortium Classification

The DTC is a governing organisation within the context of DT applications, which unites stakeholders from industry, government, and academia. In their white paper, which serves as an introductory guide toward DTs architectural framework, specific criteria are outlined for identifying whether a system qualifies as a DT, (Table 1) [5]. The same document provides details on the Information-Technology/Operational-Technology infrastructure, virtual representation, service interfaces, and techniques for synchronising the real and virtual environment.

Table 1. DT Classification according to DTC.

Criterion	Model	Simulation	Dashboard (or other Application)	Digital Twin	Digital Twin System
Virtual (Stored) Representation	Yes	No	No	Yes	Yes
Virtual (Computational) Representation / Simulation & Forecasting	No	Yes	No	Yes	Yes
Built on Data (Synchronization for Initialization)	No	No	Yes	Yes	Yes
Synchronized (including Feedback & Control)	No	No	Yes	Yes	Yes
System Integration	No	No	Yes	No	Yes

Table 1 describes five different possibilities for the classification of DT-related works, including also both the terms "DT" and "Digital Twin System (DTS)". According to the DTC, a DTS is described as "a system-of-systems that implements a digital twin", an example being a microgrid that incorporates various DTs within a single system. This classification is somewhat more intricate than the one proposed by Kritzinger, but it underscores again a fundamental point: simply possessing a model or a dashboard is insufficient for a system to be classified as a DT

2.3 Digital Twins Maturity Level

Publications on DTs have proliferated globally in recent years, with China emerging as a prominent contributor to research on energy-focused DTs. Notably, China accounts for a substantial 25% share of articles on this subject. Although utilising these categorization techniques might aid in identifying a DT, it is crucial to recognise that not all DTs exhibit the same degree of complexity. The level of maturity is related to the particular models, services, predictive analytics, and visualisation tools employed, which impact their accuracy, reliability, and level of user engagement. Therefore, several researchers have suggested different standards for categorising DTs, largely focused on their level of advancement, as provided in the coming sections.

2.3.1 Institution of Engineering and Technology's Approach

The Institution of Engineering and Technology (IET) report [4] introduces a maturity spectrum to define the components and prerequisites of a DT. It also highlights three primary applications that are believed to offer the greatest overall advantage to the built environment. Additionally, it presents the most recent advancements in the United Kingdom while also

recognising advancements in other industrial areas and nations.

Table 2 resumes the classification of maturity levels, for DTs related to the built environment, which consists of six distinct levels. It is important to note that the elements required for each level do not necessarily need to be in a specific order. For example, a third-level DT may or may not include a Building Information Model (BIM), which is an element typically found in a second-level DT. The authors also specified that many technology and service providers consider level 3 as the beginning point of a 'genuine' digital transformation because it involves real-time interaction with sensors, linked devices, and IoT technologies.

Table 2. DT Maturity Level according to IET [4].

	Defining principle	Outline usage
0	Reality capture (e.g. point cloud, drones, photogrammetry, or drawings/sketches)	Brownfield (existing) as-built survey
1	2D map/system or 3D model (e.g. object-based, with no metadata or BIM)	Design/asset optimisation and coordination
2	Connect model to persistent (static) data, metadata and BIM Stage 2 (e.g. documents, drawings, asset management systems)	4D/5D simulation Design/asset management BIM Stage 2
3	Enrich with real-time data (e.g. from IOT, sensors)	Operational efficiency
4	Two-way data integration and interaction	Remote and immersive operations Control the physical from the digital
5	Autonomous operations and maintenance	Complete self-governance with total oversight and transparency

Level 0 in this categorization represents the initial stage of DT development. It functions as an abstract depiction of the real system, used for a basic conceptualization. When this depiction is incorporated into a two-dimensional or three-dimensional model, it provides further understanding throughout both the operational and design stages, enhancing the maturity of the DT to Level 1.

Level 2 DTs incorporate simulation models developed using static data obtained from the physical system. This level allows for the use of the twin for activities such as data validation and the enhancement of design and operational efficiency. The next step towards Level 3 entails the incorporation of real-time data from interconnected devices, sensors, and the Internet of Things (IoT). Currently, there is a unidirectional flow of data that connects the physical system to the DM. This flow allows for the real-time presentation of operational information and its utilisation to improve the model. As previously said, this is the minimum criteria for an authentic DT for many stakeholders.

Upgrading from Level 3 to Level 5 DTs requires the installation of sensors and equipment to actively collect data, which involves a significant financial commitment. At Level 4,

bidirectional integration and interaction are introduced, allowing operators to actively interact with the physical system via the DM. This involves the utilisation of actuators in conjunction with model outputs, such as optimisation and predictive maintenance, to improve the efficiency and dependability of the system. Level 5 is an improvement upon Level 4 since it has automated control methods that facilitate communication between the twin and the real system. The twin can autonomously use rule-based algorithms or optimisation methods that were derived via autonomous simulations. This enables automated system control in real time without requiring human involvement.

2.3.2 Cooper's Approach

In [7] DTs are defined as a method of working with interoperative services in an interactive environment to comprehend, monitor, inform, optimize, or simulate an asset, system, process, or organization from conception to completion. A twin differs from a Supervisory Control and Data Acquisition (SCADA) or a geographic information system by the maturity of the integration with the system. In detail, the approach presented distinguishes four maturity levels of a DT, as shown in Table 3.

Table 3. DT Maturity level according to Cooper et al. [7].

	Name	Description
0	Digital Twin Ready	A virtual representation of a physical asset
1	Informational Twin	A virtual representation of an asset that uses historical data to enhance a utility function
2	Operational Twin	A virtual representation of an asset that uses real-time data analytics to enhance a utility function
3	Connected Twin	A virtual representation of connected assets that uses real-time data analytics to coordinate functions across domains

Level 0 refers to the initial level of becoming "Digital Twin Ready", in which a virtual model of a physical system is used to improve the design and planning process. A dynamic model of a generic system can be classified as Level 0 if it does not include static historical data to improve the model. Level 1, often known as the "Informational Twin", involves using historical data to improve a utility function such as the optimisation of operation, the execution of realistic hypothetical situations, and the strengthening of the accuracy of the models.

Level 2, also known as the "Operational Twin", is achieved when real-time data is transferred from the physical world to the digital. This level enables the optimisation of processes in real time and performs the dual function of detecting faults and offering suggestions. Level 3 introduces the concept of the "Connected Twin", which refers to DTs of interconnected systems. These systems employ up-to-the-minute data to synchronise operations across the physical and

digital realms, collaborating to achieve goals such as maximising resource efficiency and mitigating risks.

2.4 Buildings Digital Twins

In the future, DTs will play a significant role in the development of smart buildings, which will be achieved by incorporating sensors, smart meters, and IoT technologies. This is evident from a noticeable pattern in the literature, as shown by the increasing number of articles that discuss the uses of DTs in buildings [8]. Table 5 presents a comprehensive summary of the results obtained from research concerning DTs in the context of buildings. The data shown in Table 4 clearly shows that only six studies have used real-time connection to enhance building utilities. This highlights the importance of additional research in this field. However, the techniques used in these works have the potential to be replicated and adapted in order to improve and progress building management operations. In this section, some of the most novel studies presenting Building DTs will be presented, offering insights in the current SoA.

Table 4. Overview on Buildings-related DTs.

Ref.	Application	Modelling	Tools and Algorithms	Real-Time Connection	Optimization Goal	Results	Strengts	Weaknesses
[9]	Building air leakage model	Physical and Empirical relation	Energy+, DOE	No	-	Accuracy of air leakage models presented	Strong model and well validated, promising results in simulations air infiltration using measured data such as CO ₂ , temperature etc.	Not a real twin since no real time application for the moment. No optimization hypotized
[10]	Building energy monitoring and KPIs calculations	BIM and KPIs calculations from monitoring data	BIM	Yes	-	Good visualization and monitoring of KPIs and building information about energy consumption	Good DT, connected, detailed info about monitoring system and chosen KPIs	Leak of simulation environment for fault detection or energy optimization
[11]	Thermal monitoring	Physical and geometric relation	gmapping, MoPAD	No	-	Robot for monitoring and represent through 3d model the thermal behaviour inside the building	Robust modeling and infrastructure, well detailed robot components and interaction	Not sure if it can be considered a digital twin
[12]	Indoor IAQ and Energy efficiency in public toilets	Physical and CFD	Revit, AnsysFluend and Design Builder	No	Maximizing IAQ and Energy Savings	Optimal ventilation design scheme that can reduce the indoor pollutant concentration, improve the indoor air quality, and save energy.	Strong Modelling phase, accurate scenarios individuation, multipurpose approach	Leaks of validation and not a real twin since no real time implementation for the moment
[13]	Building management system	ML-Based	Clustering	No	-	Reduce malfunctioning and breakdowns in residential systems	Promising frameworks and objectives of the twin	Not real implementation, lacks of modelling details and optimization
[14]	Building Energy Flexibility	Physical and ML	Matlab and TRNSYS	No	Reduce building energy costs and PV tracking	Omnicomprehensive Building Energy Management Model	Promising Results, well validated, strong model	Not a real twin since no real time application for the moment.
[15]	Building Sensors	Physical	Ansys, IDA ICE, Matlab	Yes	-	Incrementing monitored variables by integrating real monitoring system with CFD and BEM models	High accuracy with short computational time, strong architecture and modelling also in real time.	To be a full twin it misses Dashboard for visualization of results and control strategies implementation of HVAC
[16]	Building Footprint KPIs	Mathematical and ML	ANN	No	-	A strong ANN to guide hotel manager to improve the efficiency and reduce carbon footprint	Good Accuracy, promising results, good approach in KPI estimation.	Not a real twin since no real time application for the moment. No optimization hypotized
[17]	Building HVAC system	Black/Grey-box modeling of components using SAREF Ontology	SAREF Ontology	No	-	A robust framework which could be use in future DT of buildings	Developed in python, object oriented and easily to replicate, promising results	Not a real twin since no real time application for the moment. No optimization hypotized
[18]	HVAC System	Physics and ML-based	BIM, ANN, Matlab for GA	Yes	Comfort and HVAC efficiency maximization	Real-time infrastructure that can reduce HVAC consumption while preserving comfort	Well validated model, Promising results, strong architecture and well detailed, real-time	No detail about how optimization is implemented in real-time

Ref.	Application	Modelling	Tools and Algorithms	Real-Time Connection	Optimization Goal	Results	Strengths	Weaknesses
[19]	Comfort and Maintenance in buildings	Physics and ML-based	XGBoost, MLP, Revit	Yes	-	DT for comfort investigation and HVAC fault detection	implementation, description of each architecture component and facility Promising fault detection and predictive maintenance potential, strong architecture and well detailed, real-time implementation,	No automated interaction from the twin to the real system
[20]	Energy storage PCM Wall	Physics based	Thermal relations	No	-	Model for PCM-based building and load optimization	Physical model description, constraints and objective function detailed, promising results	Only use of historical data and low detail about optimization algorithm
[21]	University classroom lighting	Probabilistic	-	No	-	Useful tool for quantifying the lighting energy-saving potential of different strategies	A lot of scenarios considered, promising results	Not real implementation, lacks of modelling details and optimization
[22]	Data Center	ML and BIM	LSTM	Yes	-	Improve monitoring and predictive management of datacenter	DT including temperature predictive and real-time implementation through a dashboard or VR	Lacks of optimization
[23]	Building	BIM	Revit	No	-	Energy Audits of buildings	Interesting integration of BIM and Energy Audits	No twin since no real time application, no optimization
[24]	Buildings HVAC system	ML Based	-	No	-	Accurate estimations of building energy demand	Models with an R2 up to 97.9% capable to replicate gas and electrical demands	No synchronization, no optimization
[25]	University campus	Physics Based	OpenCitiesPlanner, CitySim	Yes	-	Building consumption calculation and visualization	Real Digital Twin of a university campus, historical and estimation on consumptions	More details are required

2.4.1 Digital Twins to Balance Indoor Comfort and Improve Energy Efficiency

Modelling energy in buildings frequently involves using software such as Revit Plugins, Energy+, Ansys, and DesignBuilder. Hence, these techniques are employed to construct the digital representation of a structure during the creation of a DT. Zhao et al. constructed a DT of a four-storey bathroom building in Nanjing, China. The purpose of this DT was to replicate various ventilation systems and parameters. To do this, the researchers utilised a variety of software tools including Revit, Ansys Fluent, and DesignBuilder [12]. The objective of this study was to determine the most effective ventilation design technique for reducing indoor pollutant levels, improving Indoor Air Quality (IAQ), and achieving energy efficiency. The authors analysed the effectiveness of several scenarios by considering indicators such as pollutant concentration, ventilation efficiency, air age, PMV, Predicted Percentage of Dissatisfied (PPD), and energy use. By implementing an all-air air conditioning system with an exhaust height of 400 mm, it was feasible to achieve a 20.4% reduction in energy consumption, resulting in a decrease of 30.6 tonnes of CO₂ emissions.

Recently, there has been a growing trend of using data-driven strategies to handle problems in the energy domain, thanks to the decreasing cost of sensors and improvements in Artificial Intelligence (AI). An approach is introduced to improve the Building Energy Flexibility (BEF) using Thermostatically Controlled Loads (TCL) thanks to a DT-based data-model fusion dispatch technique [14]. The method integrates the benefits of data-driven and analytical optimisation in a novel optimisation framework, which guarantees fault tolerance and noise

reduction of parameters while enhancing the BEF. In [16], the authors introduced a new method for evaluating the energy, environmental, and economic efficiency of hotels in tropical areas of Mexico, employing Artificial Neural Networks (ANNs) and energy audits of 9 hotels. The study showcased the ANN's ability to find the most effective technology replacement configurations, resulting in a possible reduction of energy consumption per room-year by 9-12%, a decrease in CO₂ emissions by 19.46%, and a decrease in yearly energy expenses for the hotels analysed in this study. Manfren et al. [24] employed several regression models to assess the electricity and gas usage of a renovated building equipped with gas absorption heat pumps. The predictors used in the twin resulted in an R² value of up to 97.9, with a MAPE of 4.9%, indicating a significant reduction in energy consumption of up to 25.8% as a result of the renovation. Arabali et al. [26] suggest comparing Genetic Algorithms (GAs) with discontinuous nonlinear programming optimisation methods to achieve optimal balance in building HVAC demand, energy storage, PV, and wind turbine production. Their strategy was able to deliver cost savings and efficiency gains in many circumstances by describing limitations and objective function details.

DTs are being used to improve the thermal and visual comfort evaluation inside the buildings. In this framework, Tsolkas et al. [27] provided an illustrative study in which they propose a new model to improve thermal comfort in buildings by utilising interior temperature and humidity predictions. In addition, in [28], a combination of a Deep Neural Network (DNN) method and a Particle Swarm Optimisation (PSO) algorithm was used to enhance the efficiency of HVAC systems in a commercial building. The authors showcased the efficacy of DNNs in accurately representing the intricate non-linear correlation between adjustable parameters and in achieving a 7% reduction in energy usage throughout the duration of the research. Kusiak et al. [29] used a combination of data-driven ANN forecasting models and a Strength-Multi-Objective PSO method to calculate several HVAC management techniques for office buildings. Multiple solutions were offered due to the intrinsic nature of multi-objective analysis, which considers both comfort and energy consumption. Pezeshki et al. [30] used a GA to achieve optimal design in their investigation. The suggested model successfully identified the most effective positions for fan coils and radiators minimising both temperature discomfort and energy usage.

2.4.1 Digital Twins to Detect Faults and Enhance Building Monitoring Capabilities

In order to address the constraints associated with the installation of numerous sensors in interior areas, Edtmayer et al. developed a model-based DT using a combination of a building energy simulation model and a Computational Fluid Dynamics (CFD) model to simulate real

test facilities [15]. The models are connected by a simulation control data interface (SCDI) implemented in Matlab, and actual sensors are combined with virtual sensors within the model. Validation is accomplished by comparing the outputs of the virtual sensors with extensive measurement data obtained from the test facility revealing that the system is capable of delivering real-time and precise data for controlling HVAC systems. The authors of [9] assessed the EnergyPlus infiltration model "DesignFlowRate" for a residential high-rise building DT by using field measurements. In order to estimate the amount of air that was escaping, they performed blower door tests and used tracer gas measurements. Lin et al. [25] introduced an urban-scale DT model of a university campus (Fig. 5). The authors have combined historical data with predicted energy consumption projections using CitySim to create and release the first version of the twin online via OpenCitiesPlanner.

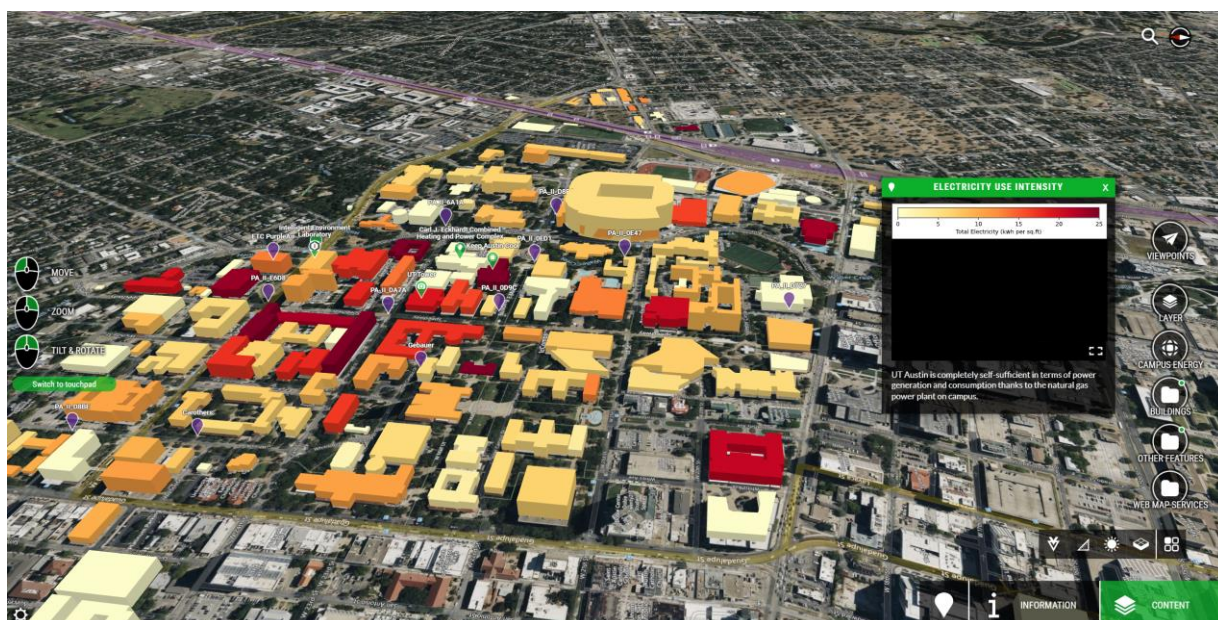


Fig. 5. Digital Twin of University of Texas Campus [25]

Regarding the advancement of DTs for HVAC systems, many studies have employed a hybrid methodology that integrates data-driven algorithms with energy modelling to minimise energy usage and maintenance costs. Bjørnskov et al. established a framework for developing a DT by employing gray-box and black-box models to mimic common building components, including HVAC systems [17]. The SAREF ontology was utilised to demonstrate the advantages and possibilities of the framework in providing services of performance monitoring, strategy creation, and operational optimisation. In two distinct investigations, Hosamo et al. focused on enhancing the thermal comfort of individuals in buildings and optimising maintenance operations in HVAC systems [19], while also prioritising energy savings [18]. Their methodology included several ML methods, including ANN and Extreme Gradient Boosting (XGBoost), and conducted energy assessments using Revit simulations. The outcome

of their efforts led to the development of two DTs, both attaining an IET Level 4 maturity indicating significant advancements in improving the thermal comfort of individuals within buildings and in optimising HVAC systems. Moreover, Ni et al. created a cloud-based digitalization framework to gather and analyse real-time data for a DT of a room in a historical building [31]. The purpose of this data is to train ML algorithms in order to forecast and enhance energy usage and human comfort. Finally, the DT presented in [32] is used to monitor, manage, and predict carbon dioxide emissions in an actual residential unit. The work provides a detailed description of the infrastructure of this twin integrating BIM, IoT, and AI.

Various studies in the existing literature have demonstrated the advantages of developing a DT for enhancing the efficiency of maintenance operations and improving the ability to detect faults in buildings, including historical ones. In this context, aiming to increase the quality of teaching and the maintenance costs, Han et al. [33] provided a comprehensive DT of a university campus with the goal of enhancing teaching quality and reducing maintenance expenses. The entire campus was designed based on the Unity paradigm, and it could be thoroughly explored utilising the platform. The virtual model presented versatile applications, such as facilitating simulations for students to engage in virtual hands-on activities, or for monitoring reasons, such as tracking vehicle traffic or identifying contaminants in sewage sources in rivers. The research authored by Hosamo et al. [34] effectively demonstrated the benefits of implementing a DT for the building. The proposed model can detect problems in buildings, such as issues with HVAC systems, lighting, etc., in real time. Several algorithms will be utilised to identify any malfunctioning systems based on the monitoring of occupants' comfort, including thermal, visual, and acoustic aspects. Focusing on historical buildings, Angjeliu et al. [35] developed a DM specifically for Milan Cathedral. The purpose of this model was to demonstrate the advantages of DT technology in improving monitoring and maintenance activities in ancient masonry structures. The findings indicated that the dynamic measurements simulation effectively mirrored the structural characteristics of the actual structure and recommended the substitution of iron ties as a component of the maintenance procedure.

2.5 Digital Twins of Smart Energy Networks

The proliferation of DESs, such as PV panels, micro-turbines, electrolyzers, and thermal prosumers, alongside the increasing adoption of EVs, has driven research interest toward the development of smart grids and smart energy networks. Within these complex systems, characterized by bidirectional flows of various energy vectors, Internet of Things (IoT) technologies, including DTs, provide transmission and distribution system operators with novel possibilities to enhance network management. This interest is highlighted in Table 5, which

summarizes the findings from 20 smart grid-focused studies, representing one of the most prevalent categories within the papers analysed.

Table 5. Overview on smart grid-related DTs.

Ref.	Application	Modelling	Tools and Algorithms	Real-Time Connection	Optimization Goal	Results	Strengts	Weaknesses
[36]	Two cities covered by smart meters per HP, PV e buildings	Basing on big data, they modelled a normalized profile PV, HP e batterie	Statistic/Mean	No	Maximize self consumption	Study effects on future RES expansion basing on collected data	Interesting results, good collection of data	Too simple modelling, not enough validated, simple optimization basing on casual choices and selecting best value, NO real time
[37]	Distributed Generation Microgrid	Real Time digital Simulator using GTNET card	-	No	-	A theoretical framework which could be use in future DT of microgrid	Useful description of a twin for a microgrid	No modelling and real application
[38]	Distributed Energy Source	Physical, ML, Rule-Based	-	No	-	A theoretical architecture which could be use in future DT of microgrid	Valid description of an architecture for a DER including digital twin	No modelling and real application
[39]	Simulated MicroGrid and DER	Physical and ML	NN, Matlab Simulink and CPLEX Solver and AMPL for optimization	Yes	Frequency, Energy Saving, Resilience	Strong DT outperformed base case in terms of energy efficiency and frequency management.	Strong DT, good application also launched in real-time for optimization purposes, well described architecture, promising results	The real-time values are generated by a well validated simulation tool, no real implementation, but the architecture presented could be easily scaled in a real case
[40]	Island Smart Grid	Physical and ML	RNN, IPSO, Matlab	No	Cost and Pollution reduction	Optimal management of DER in a Island Smart Grid	Detailed Optimization structure, well explained RNN model, interesting application	Digital Twin only mentioned, the main contribution of the work is the optimization algorithm.
[41]	Solar Power Prediction in Microgrid	ML-Based	XGBoost, Random Forest	Yes	-	Prediction of PV production	New metrics for evaluate model, in deep analysis of differend ML model	No use of prediction for optimizing the grid
[42]	Microgrid	ML-Based	NN for HVAC, GA and GUROBI Solver	Yes	Optimize consumption of thermal loads	Optimization through Virtual model of a grid in laboratory	Promising results of optimization, real testing	Simulated data for operation in real time, lacks of modelling detail and dashboard
[43]	Regional Multi-energy system	ML and Stathistic	PSO for optimization	Yes	Yes but No details	Smart methods for the management of multi-energy systems,	Robust framework for monitoring, dynamic planning, and real-time optimization	No enough detail of how the DT is implemented in real time or about optimization
[44]	Smart Grids with RES	No detail	LBAs, PSO, HMO	No	Optimize resource allocation	Optimized smart-grid operation	Honybee-mating optimization algorithm well-explored, potential real-time implication	Lacks of detail on load modelling
[45]	Multi vector energy network	ML Based	Q-Learning	No	Cost Minimization	Potential Real-time optimization	Robust algorithm to optimal manage future smart energy networks	No detail about data since confidentials, not synchronized
[46]	Power Grids with RES	Mathematical	WOA, ES, PSO	No	Maximize efficiency and minimize response time	Optimized smart-grid operation	Detailed optimization problem and algorithms	The "DT" here defined just represent connected PV
[47]	Distribution Electric Network	ML Based	ANN and TRR Method	No	-	Improve control of the network assessing dynamic security	A strong framework which, despite still not implemented, could support DSO.	Not a real application
[48]	Electric Network	Physics Based	GWO, Mca	No	Peak Load Reduction	Improve performance of the electric network	Successful hybrid optimization using Min-Conflict Algorithm and Grey Wolf Optimization	Despite being a stong model, it is not a DT
[49]	Microgrid	ML Based	ANN, GWO	No	Minimize electricity costs	Balance different energy sources and storages	Interesting combination on ANN and Grey wolf optimization	No real twin since it lacks of real physical system
[50]	Distribution Grid with Evs, storages and PV	Mixed	OpenDSS and GA	Yes	Maximize self-consumption	Reduction in reverse energy flows and increase in self-consumption	Detailed architecture, robust DT and optimal results	The bidirectionality of the model will be implemented soon
[51]	Microgrid	Mathematical	GA, PSO, PIOA	No	Minimize costs	Balance different energy sources and storages	Novel approach basing on Pigeon-inspired optimization algorithm	Not synchronized since no real implementation
[52]	Micro Grid with RES	No detail	-	No	Minimize consumption and PV injections	New operation strategies to optimal manage the network	Accurate model of the real facility and promising operation improvements	DT modelling should be more detailed
[53]	Microgrid	ML Based	PSO and EPSOLW	No	Minimize costs of energy production	Balance different energy sources and storages	Strong Optimization algorithms description and new approach using PID controllers	Lack of informations about the real system of the twin

Ref.	Application	Modelling	Tools and Algorithms	Real-Time Connection	Optimization Goal	Results	Strengths	Weaknesses
[54]	Microgrid	Physics Based	BOA	No	Maximize Profits	Model to investigate quantitative energy trading strategies	Optimization problem detailed and novel approach using Bat Optimization Algorithm	Not synchronized with a real application
[55]	Microgrid	Stochastic and Physics based	Shark Opt.	No	Maximize Profits	Model to propose optimized strategy for managing microgrids	Optimization problem detailed and novel approach using Shark Optimization Algorithm	Not based on a real system

2.5.1 Digital Twin for Real Time Management of Smart Grids

The digitalization of electric networks is now a highly debated research issue in the fields of business, industry, and government. This transformation needs the most updated technology to enhance network operations, optimise predictive maintenance, and integrate RES. Most publications reviewed focuses on the use of DTs in Smart Grids since, due to the extensive use of RES, DTs are additionally utilised for predicting renewable energy generation. This can be noticed in [41], where the authors employed several ML algorithms to anticipate the PV output in a microgrid. Furthermore, [37] outlines a framework for developing a DT of a microgrid that incorporates a substantial proportion of RESs. The authors proposed the use of a real-time digital simulator (RTDS) that relies on a communication electronic card to accurately model and simulate the microgrid system. Kovalyov [38] introduced a DT framework for the development and control of distributed energy resources, which relies on the concept of a DT. The platform utilises Industry 4.0 technology to provide customers with access to emerging market prospects, cost reductions, and enhancements in power supply quality. The essential elements of a twin are then proposed including novel mathematical tools based on category theory to formalise and automate the process of building a DT. In order to address the issue of monitoring the stability of electric networks and provide early warning, Cai and Westermann propose a DT [47] where the Trust-region-reflective optimisation approach was utilised to maximise the accuracy of the model in estimating real parameters. Finally, an ANN was applied for state estimation and dynamic security evaluation.

The deployment of DT in smart grids has achieved a significant level of advancement, as demonstrated by the real-time applications found in this literature study. The DT created by Bragatto et al. [50] accurately replicated the functioning of a distribution network in Terni, Italy, with operation that closely resembles real-time conditions. This DT enabled advanced capabilities such as gathering data, estimating the system's status, and assessing its adaptability. Hong et al provided a detailed explanation of the procedure for creating DTs in Networked Microgrids (NMGs). This process includes the use of several modelling techniques such as physics-based, data-driven, and hybrid models involving even real-time data connection [39]. The main contributions include the development of a DM that allows real-time interaction with

DERs, auxiliary services, and predictive scheduling. Additionally, an investigation was conducted on the use of DT architecture in both isolated and grid-tied modes for NMGs. Bai et al. [43] developed a DT for the regional multi-energy system of Hong Feng Lake in China. This DT had the capability to automatically and continuously optimise an active distribution network, handling several types of RES, as well as charging stations and thermal and electric loads. Moreover, Rizeakos et al. [56] proposed a study that provided a data-driven system for identifying the location of faults and classifying their types in low voltage distribution grids. The algorithm obtained a detection accuracy of over 91% in recognising defects and accurately locating them by using synthetic data and implementing a CNN.

2.5.1 Digital Twin for Cost and Energy Optimization of Smart Grids

The proliferation of DESs will make demand response technologies crucial for their efficient control in the entire transmission and distribution networks. In this framework, Cao et al. used the DT technology in a micro-grid consisting of various RESs and Plug-In Hybrid Vehicles [57]. The simulations yielded optimal operational strategies for cost minimization by effectively using available resources, as well as including algorithms for detecting cyber-attacks. Pan et al. constructed a DM of a micro-grid consisting of an electrical storage system, a wind turbine, 5 PV panels, and a building [58]. By employing reinforcement learning, it was possible to devise a scheduling strategy that optimises the annual expenses associated with power, battery use, and wind generator energy.

At the same time, several works have employed DT frameworks to optimise microgrids for different goals, without being yet deployed in real-time or connected to an actual network. Bayer et al. [36] created a DM for an energy system that includes various buildings, PV systems, heat pumps, and batteries. The researchers utilised almost 8000 smart metres in two small towns located in northern Germany, and three years' worth of hourly data to develop simulation models for studying the impact of PV systems, batteries, and Heat Pumps (HPs) on the power grid. The penetration of PV systems led to a 24% increase in renewable power generation, a 13% reduction in energy imports, and the study found that battery installations were unable to cut peak energy imports above than 1%. The use of HPs for heating purposes would only lead to an increase in energy consumption, as the use of the sole PV plants will not be sufficient to meet the demand, especially on the coldest days. An optimised smart island (SI) energy system was developed in [40], which combines energy storage with various RESs such as micro gas turbines, PV plants, fuel cells, and wind turbines. The optimization algorithm depends on an improved particle swarm optimisation technique for efficient power scheduling, and it employs Recurrent Neural Networks (RNNs) for RES forecasts. Furthermore, the concept of

interruptible load demand response is introduced, which involves the use of battery storage to shift energy consumption and hence decrease operational expenses while maintaining system stability. In [46], Fan and Li presented the positive results of utilising a DT for a smart grid, which was effectively controlled across both Cloud and Fog layers. A comparative analysis was performed to evaluate the effectiveness of state-of-the-art optimisation algorithms in comparison to whale optimisation, emphasising the benefits of the latter technique. Using a similar approach, [44] employed fog-cloud technology in a smart grid system, improving it using a honeybee-mating optimisation algorithm. The DT infrastructure efficiently models 35 clusters of buildings, with each cluster being controlled through the fog-cloud interface to improve the efficiency of resource allocation. Yang et al. [48] proposed a system that combines a DT with a hybrid optimisation algorithm, namely the Min-Conflict Algorithm and Grey Wolf Optimisation. The goal of this method is to efficiently distribute supply and demand resources on the electric network. Further instances of microgrid optimisation using DT can be seen in [54] and [55], where, respectively, the bat optimisation algorithm and the shark optimisation method were utilised to maximise profit in various conditions.

2.5.2 Digital Twins to Improve District Heating Networks Design and Operation

Optimisation techniques are crucial for enhancing the design and efficiency of future DHNs, since these systems will interface with various heating or cooling sources, and an effective management strategy is necessary to identify the most efficient mix of heat sources, such as electricity to heat, solar collectors, and boilers, in order to satisfy the heating demand while minimising costs and Green House Gas (GHG) emissions. Research conducted by Reynolds et al. [59] analysed a multi-vector energy where the optimal management techniques were determined by employing a GA and an ANN for PV production and building consumption forecasting. The findings revealed a substantial 44.88% increase in profits as compared to the first scenario. In the study carried out by Cox et al. [60], a data-driven methodology was employed to investigate the most effective control method for an ice storage system in a District Cooling Network to minimise operational expenses while accommodating fluctuating loads and energy prices. Lin et al. [61] conducted a study on another DT application in DHN. The authors presented an overview of the practical uses and future prospects of smart heating systems. They then showcased the efficiency of their model in properly predicting the heat delay time and pressure at each node within the district. Through the use of effective management strategies, which were trialled over a two-day period, they achieved a significant reduction in the consumption of natural gas by up to 31% when compared to traditional operational methods.

Given that DHNs might be impacted by many typologies of faults, Bahlawan et al.

introduced a DT that can be employed for fault identification in an actual case study [62]. Through the implementation of an innovative diagnostic methodology, the simulation results demonstrated the successful detection and localization of both single and multiple defects. Furthermore, the diagnostic tool accurately calculated the temperature and pressure of each individual node, with an impressively low error rate of less than 0.1%. Research conducted by Kim et al. [63] examined ML approaches to determine their efficacy in identifying defects in DHNs. The suggested model, utilising a Gradient Boosting Regressor, successfully predicted the operational behaviour of a DHN substation by accurately differentiating between high-performing and defective substations. Using training data from a Swedish DH system, the authors were capable of develop an accurate model.

2.6 Digital Twins of Energy Systems Components

This section examines research focused on the single elements inside broader energy systems. Indeed, the utilisation of DT technology extends beyond intricate systems like as power plants, smart grids, or buildings, since implementing a DT for a single device could greatly improve the productivity and efficiency of the entire system. Furthermore, by concentrating solely on each individual component, it becomes feasible to optimise their specific operation and maintenance plans, as well as to proactively detect future faults and downtimes periods. Precisely, Table 6 provides a comprehensive summary of various studies pertaining to energy components identified in the course of this review section.

Table 6. Overview on energy component-related DTs.

Ref.	Application	Modelling	Tools and Algorithms	Real-Time Connection	Optimization Goal	Results	Strengths	Weaknesses
[64]	Li-On battery	Thermo-Electrical coupled problem (COMSOL + Simulink)	FEM + Physycs	No	-	Model validation for Internal Short Circuit detection basing on enhancing temperature moniroring	Strong Model and well validated, faster results than experimental, enhance of measured variable, good fault detection approach	No real time application for the moment
[65]	Blade of a tidal turbine	FEM with a in-house developed software so-called BadeComp	FEM	No	-	Effect of seawater and structural degradation in a blade of a tidal turbine	Good Accuracy, detailed FEM model, the model can simulate in different from test environment (seawater instead of normal air)	Not a real twin since no real time application for the moment
[66]	Battery Energy Storage System	Physics based	RSCAD	Yes	-	Strong control strategy fot achieving ramp-limit violation reduction	Well validated control strategy and simulation through historical and real-time data. Methods and devices used strongly explored	It lacks of optimization and PV modelling, but authors will consider it in future
[67]	Blade of wind turbine	Physics based	Augmentek Kalman Filter	No	-	Structural health monitoring of a wind turbine blade	Robust modelling framework, well-validated through experimental test	No real time application, no optimization
[68]	Ultrahigh-Power Li-On Battery	Physics Based	Comsol	No	Accurate Design	Model which replicate a real battery enabling simulations	Accurate modelling phase including battery power capability optimization	Few details of the optimization process
[69]	Large-scale battery system	Physics based	-	No	-	Model of a MMh scale grid battery modeled to study thermal and electrical behaviour	Useful and detailed physical model for a large battery. Strong mathematics and physics relations. Impact of thermal effects and degradations are investigated	Not implemented in real time

2.6.1 Physics-Based Digital Twins

The components being mentioned in Table 6 are mostly modelled using physics-based methods. This results in accurate and resilient models that not only simulate but also enable users to do virtual fault testing, therefore enhancing our understanding of preventative actions.

This study of the SOA identified two primary types of DTs used in energy systems: batteries and wind turbine blades. All of these models utilised a physics-based approach, employing Finite Element Method (FEM) software such as Comsol Multiphysics. DT of batteries are utilised to increase the monitored parameters, hence allowing improved simulations or the identification of defects. To enhance design process, Zhang et al. introduced a DT for an Ultrahigh-Power lithium battery in [68] where an accurate model was developed using Comsol for a LiFePO₄/graphite battery. In their study, Yuan et al. [64] created a DT model of an experimental battery to reproduce faults data and establish a new threshold for detecting internal short circuits. Unlike the actual duration of almost eight months, the twin completed 78 experimental trials within a few weeks and provided precise data on the temperature of each cell. In order to mitigate the occurrences of ramp-rate limit breaches in battery electric energy storage systems Alam et al. [66] created a DT for a control system of a PV power plant located in Australia which was evaluated in real-time for a 5-hour period. Their control approach and system, consisting of two computers that performed real-time simulation utilising a MATLAB-RSCADRTDS co-simulation platform, were verified using actual historical data.

Despite many applications of physics-based algorithms involve the development of DT of batteries, application on diesel generator and wind turbine blades have been even encountered. Bondarenko et al. [70] introduced a DT model for a diesel engine that relies on a physical analytical model incorporating energy and mass conservation restrictions, together with the Wiuebe combustion model. The authors examined the difficulties and current techniques for modelling diesel engines and introduced their own methodology, which showed encouraging outcomes in faithfully reproducing the engine's performance. The study in [65] examines the structural performance of a fiber-reinforced composite blade for a floating tidal turbine device with a 1 MW rotor blade. The model demonstrated a high level of accuracy and was validated through the use of experimental data, indicating that, despite some deterioration in material properties, the blade can safely function at the full intended load for the whole planned lifespan, even when exposed to saltwater. Finally, in [71], a DT for a heat pump is presented whose model could enhance its precision and computing speed by employing a particle swarm optimisation technique that integrates real-time observed data. In addition, the suggested optimal control approach, following the prediction of future building demand, shown a decrease in energy usage of up to 15% in comparison to the manual control technique.

2.6.2 Artificial Intelligence-Based Digital Twins

Dynamic modelling can be achieved by the application of physical equations, data-driven algorithms, or a mix of both approaches. In this framework, to optimise the efficiency of a 200 MW biomass boilers Spinti et al. developed a DT based on a combined approach [72]. The DT integrates a science-based model with data-driven algorithms, precisely deep learning Bayesian analysis, to achieve optimal outcomes in simulations that accurately represent the behaviour of the actual system. Additionally, by utilising acquired data, the twin will autonomously update itself to ensure it remains up-to-date and relevant at all times [73]. Blume et al. [74] used a data-driven approach to model an industrial cooling tower by investigating various modelling techniques, including linear regression, basic regression tree, and ANN. Natgunanathan et al. presented a study [75] in which they used ML models to forecast the electricity output of a 7 MW solar farm in Australia. The authors employed data obtained from the site to examine the efficacy of several forecasting models, such as Support Vector Regression and Random Forest (RF), and presented two innovative criteria for evaluation. Their findings demonstrated that the ANN attained the greatest R^2 value of 0.94, while, when evaluating the recently presented criteria, the RF approaches shown greater performance. Finally, A DM was developed by Akhlaghi et al. to enhance the performance of a dew point cooler [76]. This was achieved by utilising genetic optimisation techniques, feed-forward NN, and physics correlations. The results showed a significant improvement in COP by 72.75% and a reduction in power consumption by 49.41%.

2.7 Discussion and Limits of Current Research about DTs

The field of DT is gaining greater recognition for its capacity to transform numerous applications by enabling the possibility of simulations and predictive analytics. An investigation on 64 reviewed papers revealed linkages within the categorization rules, emphasising regions of similarity and divergence. According to the results, the categorical method might not include as a twin some works that should be recognised as DTs, especially those that integrate an existing SCADA with models to create a Decision Support System (DSS) [77]. This integration increases the maturity level to 3, as classified by IET, highlighting the significant potential of these applications. This finding implies that utilising a maturity level-based approach, based on either IET or Cooper's categorization, might provide a more reliable method of characterising twins. Moreover, it is still important to discern an authentic twin from other DM and when considering maturity level frameworks, it is recommended to position Digital Twins at a minimum of level 3 according to the IET standards and level 2 according to Cooper's criteria. This not only improves their usefulness but also aligns with the expectations

of the EU [78] and IEA [79] for DT technology. As illustrated by presented works, a DS can function as a DSS, offering optimisation strategies, warnings for fault detection, and suggestions for operators who interact with the dashboard. This solution helps facility managers achieve particular goals, such as reducing fuel usage in DHN or maximising RES utilisation through intelligent EV charging management [80]. This alignment highlights the crucial significance of digital technologies in fulfilling upcoming technical and operational requirements. Based on the criteria of Sections 2.2 and 2.3, as presented in Tables 7, which focus on the 64 papers subject to a systematic review, only 36% of the works achieved a maturity level higher than 3. This analysis indicates a significant gap in the development of comprehensive DTs. Addressing this gap requires first clarifying the distinctions between DTs and DMs. Secondly, research on these technologies should concentrate on developing synchronized models that can be utilized by facility managers, ranging from Smart Grids and DHNs to Smart Buildings, to enhance energy savings and system reliability.

Table 7. Overview of classification results.

Ref.	Field of Application	Kritzinger	DTC	IET	Cooper	Ref.	Field of Application	Kritzinger	DTC	IET	Cooper
[36]	Smart Grid	DM	Simulation	2	1	[20]	Building	DM	Simulation	2	1
[64]	Energy Component	DM	Simulation	2	1	[81]	Power Generator	DM	Simulation	2	1
[65]	Energy Component	DM	Simulation	2	1	[67]	Energy Component	DM	Simulation	2	1
[10]	Building	DS	MD	3	2	[82]	Smart Grid	DT	DTS	4	3
[17]	Energy Component	DM	Simulation	2	1	[82]	Power Generator	DT	DT	4	3
[83]	Power Generator	DM	Simulation	2	1	[84]	Power Generator	DM	Simulation	2	1
[66]	Energy Component	DT	DTS	4	2	[42]	Smart Grid	DT	DTS	4	3
[77]	Power Generator	DS	DT	3	2	[85]	Building	DS	DT	3	2
[86]	Vehicle	DM	Simulation	2	1	[87]	Smart Grid	DM	Simulation	2	1
[88]	Power Generator	DM	Model	2	1	[13]	Building	DM	Simulation	2	1
[89]	Industry	DS	DT	3	2	[21]	Building	DM	Simulation	2	1
[9]	Building	DM	Model	2	1	[22]	Building	DS	DT	3	2
[11]	Building	DS	MD	3	2	[43]	Smart Grid	DT	DTS	5	3
[90]	Power Generator	DM	Simulation	2	1	[23]	Building	DM	Simulation	2	1
[16]	Building	DM	Model	2	1	[91]	Vehicle	DM	Simulation	2	1
[12]	Building	DM	Model	2	1	[44]	Smart Grid	DM	Simulation	2	1
[92]	Industry	DM	Model	2	1	[93]	Power Plant	DS	DT	3	2
[39]	Smart Grid	DT	DT	4	3	[45]	Smart Grid	DM	Simulation	2	1
[94]	Energy Component	DS	DT	3	2	[68]	Component	DM	Simulation	2	1
[14]	Building	DM	Model	2	1	[46]	Smart Grid	DM	Simulation	2	1
[15]	Building	DS	DT	3	2	[47]	Smart Grid	DM	Simulation	2	1
[95]	Power Generator	DT	DT	4	2	[48]	Smart Grid	DM	Simulation	2	1
[40]	Smart Grid	DM	Model	2	1	[24]	Building	DM	Simulation	2	1
[73]	Power Generator	DT	DT	4	2	[49]	Smart Grid	DM	Simulation	2	1
[96]	Vehicle	DM	Simulation	2	1	[50]	Smart Grid	DS	DT	3	2
[41]	Power Generator	DS	Simulation	3	2	[97]	Power Plant	DM	Model	2	1
[98]	Smart Grid	DS	Dashboard	3	2	[51]	Smart Grid	DM	Simulation	2	1
[18]	Building	DS	DT	4	2	[52]	Smart Grid	DM	Simulation	2	1
[99]	Power Generator	DM	Simulation	2	1	[25]	Building	DS	DT	3	2
[69]	Energy Component	DM	Simulation	2	1	[53]	Smart Grid	DM	Simulation	2	1
[100]	Power Generator	DM	Simulation	2	1	[54]	Smart Grid	DM	Simulation	2	1
[19]	Building	DS	DT	4	2	[55]	Smart Grid	DM	Simulation	2	1

References

- [1] M. Grieves, “Digital Twin: Manufacturing Excellence through Virtual Factory Replication,” 2015, [Online]. Available: <https://www.researchgate.net/publication/275211047>
- [2] M. Shafto, M. C. Rich, D. E. Glaessgen, C. Kemp, J. Lemoigne, and L. Wang, “DRAFT Modeling, Simulation, information Technology & Processing Roadmap,” 2010.
- [3] M. Singh, E. Fuenmayor, E. P. Hinchy, Y. Qiao, N. Murray, and D. Devine, “Digital twin: Origin to future,” 2021, *MDPI AG*. doi: 10.3390/asi4020036.
- [4] Institution of Engineering and Technology, “Digital twins for the built environment - An introduction to the opportunities, benefits, challenges and risks,” 2019. Accessed: Feb. 04, 2024. [Online]. Available: <https://www.theiet.org/impact-society/factfiles/built-environment-factfiles/digital-twins-for-the-built-environment/>
- [5] D. Mckee, “Platform Stack Architectural Framework: An Introductory Guide - A Digital Twin Consortium White Paper,” 2023.
- [6] W. Kritzinger, M. Karner, G. Traar, J. Henjes, and W. Sihn, “Digital Twin in manufacturing: A categorical literature review and classification,” in *IFAC-PapersOnLine*, Elsevier B.V., Jan. 2018, pp. 1016–1022. doi: 10.1016/j.ifacol.2018.08.474.
- [7] J. P. Cooper, S. Jackson, S. Kamojjala, G. Owens, K. Szana, and S. Tomic, “Demystifying Digital Twins: Definitions, Applications, and Benefits,” *J Am Water Works Assoc*, vol. 114, no. 5, pp. 58–65, Jun. 2022, doi: 10.1002/awwa.1922.
- [8] R. Bortolini, R. Rodrigues, H. Alavi, L. F. D. Vecchia, and N. Forcada, “Digital Twins’ Applications for Building Energy Efficiency: A Review,” Oct. 01, 2022, *MDPI*. doi: 10.3390/en15197002.
- [9] G. Bastos Porsani, N. Casquero-Modrego, J. B. Echeverria Trueba, and C. Fernández Bandera, “Empirical evaluation of EnergyPlus infiltration model for a case study in a high-rise residential building,” *Energy Build*, vol. 296, Oct. 2023, doi: 10.1016/j.enbuild.2023.113322.
- [10] P. Spudys, N. Afxentiou, P. Z. Georgali, E. Klumbyte, A. Jurelionis, and P. Fokaides, “Classifying the operational energy performance of buildings with the use of digital twins,” *Energy Build*, vol. 290, Jul. 2023, doi: 10.1016/j.enbuild.2023.113106.
- [11] A. Adán, A. López-Rey, and A. Ramón, “Robot for thermal monitoring of buildings,” *Autom Constr*, vol. 154, Oct. 2023, doi: 10.1016/j.autcon.2023.105009.
- [12] L. Zhao *et al.*, “Digital Twin Evaluation of Environment and Health of Public Toilet Ventilation Design Based on Building Information Modeling,” *Buildings*, vol. 12, no. 4, Apr. 2022, doi: 10.3390/buildings12040470.
- [13] S. Agostinelli and F. Cumo, “MACHINE LEARNING APPROACH FOR PREDICTIVE

MAINTENANCE IN AN ADVANCED BUILDING MANAGEMENT SYSTEM,” in *WIT Transactions on Ecology and the Environment*, WITPress, Aug. 2022, pp. 131–138. doi: 10.2495/EPM220111.

- [14] Y. Song, M. Xia, Q. Chen, and F. Chen, “A data-model fusion dispatch strategy for the building energy flexibility based on the digital twin,” *Appl Energy*, vol. 332, Feb. 2023, doi: 10.1016/j.apenergy.2022.120496.
- [15] H. Edtmayer *et al.*, “Modelling virtual sensors for real-time indoor comfort control,” *Journal of Building Engineering*, vol. 67, May 2023, doi: 10.1016/j.job.2023.106040.
- [16] S. G. Mengual Torres *et al.*, “Analysis of Energy and Environmental Indicators for Sustainable Operation of Mexican Hotels in Tropical Climate Aided by Artificial Intelligence,” *Buildings*, vol. 12, no. 8, Aug. 2022, doi: 10.3390/buildings12081155.
- [17] J. Bjørnskov and M. Jradi, “An ontology-based innovative energy modeling framework for scalable and adaptable building digital twins,” *Energy Build*, vol. 292, Aug. 2023, doi: 10.1016/j.enbuild.2023.113146.
- [18] H. Hosamo, M. H. Hosamo, H. K. Nielsen, P. R. Svennevig, and K. Svidt, “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*, vol. 17, no. 2, pp. 125–171, 2023, doi: 10.1080/17512549.2022.2136240.
- [19] H. H. Hosamo, H. K. Nielsen, D. Kraniotis, P. R. Svennevig, and K. Svidt, “Improving building occupant comfort through a digital twin approach: A Bayesian network model and predictive maintenance method,” *Energy Build*, vol. 288, Jun. 2023, doi: 10.1016/j.enbuild.2023.112992.
- [20] Z. Lv, C. Cheng, and H. Lv, “Digital twins for secure thermal energy storage in building,” *Appl Energy*, vol. 338, May 2023, doi: 10.1016/j.apenergy.2023.120907.
- [21] H. Seo and W. S. Yun, “Digital Twin-Based Assessment Framework for Energy Savings in University Classroom Lighting,” *Buildings*, vol. 12, no. 5, May 2022, doi: 10.3390/buildings12050544.
- [22] P. Pinmas, P. Panyadee, M. Kaewmoracharoen, S. Khueankhan, and A. Wannachai, “Towards Digital Twin Data Center Using Building Information Modeling and Real-Time Data Sensing,” in *ITC-CSCC 2022 - 37th International Technical Conference on Circuits/Systems, Computers and Communications*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 908–911. doi: 10.1109/ITC-CSCC55581.2022.9895024.
- [23] P. Spūdys, A. Jurelionis, N. Afxentiou, P. A. Fokaides, and P. Konatzii, “Employment of digital twins for the implementation of energy audits.”
- [24] M. Manfren, P. A. James, V. Aragon, and L. Tronchin, “Lean and interpretable digital twins for building energy monitoring – A case study with smart thermostatic radiator valves and gas

-
- absorption heat pumps,” *Energy and AI*, vol. 14, Oct. 2023, doi: 10.1016/j.egyai.2023.100304.
- [25] C. Lin, T. Y. Dai, A. D. Dilsiz, D. Crawley, D. Niyogi, and Z. Nagy, “UTwin: A digital twin of the UT Austin campus,” in *BuildSys 2023 - Proceedings of the 10th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, Association for Computing Machinery, Inc, Nov. 2023, pp. 282–283. doi: 10.1145/3600100.3626261.
- [26] A. Arabali, M. Ghofrani, M. Etezadi-Amoli, M. S. Fadali, and Y. Baghzouz, “Genetic-algorithm-based optimization approach for energy management,” *IEEE Transactions on Power Delivery*, vol. 28, no. 1, pp. 162–170, 2013, doi: 10.1109/TPWRD.2012.2219598.
- [27] C. Tsolkas, E. Spiliotis, E. Sarma, V. Marinakis, and H. Doukas, “Dynamic energy management with thermal comfort forecasting,” *Build Environ*, vol. 237, p. 110341, Jun. 2023, doi: 10.1016/j.buildenv.2023.110341.
- [28] Jiahao Deng and Haoran Wang, “Modeling and Optimizing Building HVAC Energy Systems Using Deep Neural Networks,” *2018 International Conference on Smart Grid and Clean Energy Technologies (ICSGCE)*, 2018.
- [29] A. Kusiak, G. Xu, and F. Tang, “Optimization of an HVAC system with a strength multi-objective particle-swarm algorithm,” *Energy*, vol. 36, no. 10, pp. 5935–5943, 2011, doi: 10.1016/j.energy.2011.08.024.
- [30] Zahra Pezeshki, Ali Soleimani, and Ahmad Darabi, “GA-FBHCL: A method for the best HVAC location,” *2020 Zooming Innovation in Consumer Technologies Conference (ZINC)*, 2020, doi: <https://doi.org/10.1109/ZINC50678.2020.9161819>.
- [31] Z. Ni, P. Eriksson, Y. Liu, M. Karlsson, and S. Gong, “Improving energy efficiency while preserving historic buildings with digital twins and artificial intelligence,” in *IOP Conference Series: Earth and Environmental Science*, IOP Publishing Ltd, Oct. 2021. doi: 10.1088/1755-1315/863/1/012041.
- [32] A. Arsiwala, F. Elghaish, and M. Zoher, “Digital twin with Machine learning for predictive monitoring of CO2 equivalent from existing buildings,” *Energy Build*, vol. 284, p. 112851, Apr. 2023, doi: 10.1016/j.enbuild.2023.112851.
- [33] X. Han *et al.*, “Intelligent Campus System Design Based on Digital Twin,” *Electronics (Switzerland)*, vol. 11, no. 21, Nov. 2022, doi: 10.3390/electronics11213437.
- [34] H. H. Hosamo, 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 Build*, vol. 281, Feb. 2023, doi: 10.1016/j.enbuild.2022.112732.
- [35] G. Angjeliu, D. Coronelli, and G. Cardani, “Development of the simulation model for Digital Twin applications in historical masonry buildings: The integration between numerical and
-

-
- experimental reality,” *Comput Struct*, vol. 238, Oct. 2020, doi: 10.1016/j.compstruc.2020.106282.
- [36] D. Bayer and M. Pruckner, “A digital twin of a local energy system based on real smart meter data,” *Energy Informatics*, vol. 6, no. 1, Dec. 2023, doi: 10.1186/s42162-023-00263-6.
- [37] P. Yu *et al.*, “Framework design and application perspectives of digital twin microgrid,” *Energy Reports*, vol. 9, pp. 669–678, Sep. 2023, doi: 10.1016/j.egyr.2023.04.253.
- [38] S. P. Kovalyov, “Distributed Energy Resources Management: From Digital Twin to Digital Platform,” in *IFAC-PapersOnLine*, Elsevier B.V., 2022, pp. 460–465. doi: 10.1016/j.ifacol.2022.07.080.
- [39] Y. Y. Hong and G. F. D. Apolinario, “Ancillary Services and Risk Assessment of Networked Microgrids Using Digital Twin,” *IEEE Transactions on Power Systems*, vol. 38, no. 5, pp. 4542–4558, Sep. 2023, doi: 10.1109/TPWRS.2022.3215144.
- [40] L. P. Gao and Z. Gao, “An optimal management architecture based on digital twin for smart solar-based islands incorporating deep learning and modified particle swarm optimization,” *Solar Energy*, vol. 262, Sep. 2023, doi: 10.1016/j.solener.2023.111872.
- [41] I. Natgunanathan, A. Anwar, S. Rajasegarar, and V. Mak-Hau, “Error Spectrum Analysis of Solar Power Prediction for Deakin Microgrid Digital Twin,” in *Asia-Pacific Power and Energy Engineering Conference, APPEEC*, IEEE Computer Society, 2022. doi: 10.1109/APPEEC53445.2022.10072286.
- [42] G. Fernández *et al.*, “Development and Initial Testing of a Virtual Laboratory for the Build-up and Testing of Microgrid Management Algorithms,” *Renewable Energy and Power Quality Journal*, vol. 20, pp. 644–649, Sep. 2022, doi: 10.24084/repqj20.390.
- [43] H. Bai *et al.*, “Automatic Modeling and Optimization for The Digital twin of a Regional Multi-energy System,” in *Proceedings - 2022 Power System and Green Energy Conference, PSGEC 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 214–219. doi: 10.1109/PSGEC54663.2022.9881075.
- [44] H. Zhang and R. Sun, “A novel optimal management method for smart grids incorporating cloud-fog layer and honeybee mating optimization algorithm,” *Solar Energy*, vol. 262, Sep. 2023, doi: 10.1016/j.solener.2023.111874.
- [45] D. Wang, R. Fan, Y. Li, and Q. Sun, “Digital twin based multi-objective energy management strategy for energy internet,” *International Journal of Electrical Power and Energy Systems*, vol. 154, Dec. 2023, doi: 10.1016/j.ijepes.2023.109368.
- [46] X. Fan and Y. Li, “Energy management of renewable based power grids using artificial intelligence: Digital twin of renewables,” *Solar Energy*, vol. 262, Sep. 2023, doi: 10.1016/j.solener.2023.111867.
-

-
- [47] H. Cai and D. Westermann, "Enhancing dynamic security assessment in converter dominated distribution networks through digital twin-based small signal stability analysis," *At-Automatisierungstechnik*, vol. 71, no. 12, pp. 1028–1039, Dec. 2023, doi: 10.1515/auto-2023-0119.
 - [48] H. Yang, S. Zhang, J. Zeng, S. Tang, and S. Xiong, "Future of sustainable renewable-based energy systems in smart city industry: Interruptible load scheduling perspective," *Solar Energy*, vol. 263, Oct. 2023, doi: 10.1016/j.solener.2023.111866.
 - [49] Q. Li, Z. Cui, Y. Cai, and Y. Su, "Multi-objective operation of solar-based microgrids incorporating artificial neural network and grey wolf optimizer in digital twin," *Solar Energy*, vol. 262, Sep. 2023, doi: 10.1016/j.solener.2023.111873.
 - [50] T. Bragatto *et al.*, "Near real-time analysis of active distribution networks in a Digital Twin framework: A real case study," *Sustainable Energy, Grids and Networks*, vol. 35, Sep. 2023, doi: 10.1016/j.segan.2023.101128.
 - [51] J. Xu and J. Gong, "Novel sustainable urban management framework based on solar energy and digital twin," *Solar Energy*, vol. 262, Sep. 2023, doi: 10.1016/j.solener.2023.111861.
 - [52] M. Dellaly, S. Skander-Mustapha, and I. Slama-Belkhodja, "Optimization of a residential community's curtailed PV power to meet distribution grid load profile requirements," *Renew Energy*, vol. 218, Dec. 2023, doi: 10.1016/j.renene.2023.119342.
 - [53] P. Pei, Z. Chen, and Y. Wang, "Production analytics for renewable energy portfolio management: A comparative study of active and passive strategies in digital twin," *Solar Energy*, vol. 262, Sep. 2023, doi: 10.1016/j.solener.2023.111923.
 - [54] G. Yang, L. Zhang, S. Li, and X. Wu, "Quantitative energy trading strategies in cooperative microgrids in electricity market: A multi-dimensional analysis of risk and return," *Solar Energy*, vol. 262, Sep. 2023, doi: 10.1016/j.solener.2023.111860.
 - [55] S. Shen and Y. Yuan, "The economics of renewable energy portfolio management in solar based microgrids: A comparative study of smart strategies in the market," *Solar Energy*, vol. 262, Sep. 2023, doi: 10.1016/j.solener.2023.111864.
 - [56] V. Rizeakos, A. Bachoumis, N. Andriopoulos, M. Birbas, and A. Birbas, "Deep learning-based application for fault location identification and type classification in active distribution grids," *Appl Energy*, vol. 338, May 2023, doi: 10.1016/j.apenergy.2023.120932.
 - [57] H. Cao, D. Zhang, and S. Yi, "Real-Time Machine Learning-based fault Detection, Classification, and locating in large scale solar Energy-Based Systems: Digital twin simulation," *Solar Energy*, vol. 251, pp. 77–85, Feb. 2023, doi: 10.1016/j.solener.2022.12.042.
 - [58] M. Pan, Q. Xing, Z. Chai, H. Zhao, Q. Sun, and D. Duan, "Real-time digital twin machine learning-based cost minimization model for renewable-based microgrids considering

-
- uncertainty,” *Solar Energy*, vol. 250, pp. 355–367, Jan. 2023, doi: 10.1016/j.solener.2023.01.006.
- [59] J. Reynolds, M. W. Ahmad, Y. Rezgui, and J. L. Hippolyte, “Operational supply and demand optimisation of a multi-vector district energy system using artificial neural networks and a genetic algorithm,” *Appl Energy*, vol. 235, pp. 699–713, Feb. 2019, doi: 10.1016/j.apenergy.2018.11.001.
- [60] S. J. Cox, D. Kim, H. Cho, and P. Mago, “Real time optimal control of district cooling system with thermal energy storage using neural networks,” *Appl Energy*, vol. 238, pp. 466–480, Mar. 2019, doi: 10.1016/j.apenergy.2019.01.093.
- [61] X. Lin, W. Zhong, Y. Mao, and N. Zhang, “Smart Integrated Energy System Based on Digital Twin: Research, Application and Outlook,” in *2022 IEEE 2nd International Conference on Digital Twins and Parallel Intelligence, DTPI 2022*, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/DTPI55838.2022.9998890.
- [62] H. Bahlawan *et al.*, “Detection and identification of faults in a District Heating Network,” *Energy Convers Manag*, vol. 266, Aug. 2022, doi: 10.1016/j.enconman.2022.115837.
- [63] Y.-G. Kim, K. Heo, G.-E. You, H.-S. Lim, J.-I. Choi, and J.-S. Eom, “A Study on the Improvement of Thermal Energy Efficiency for District Thermal Energy Consumer Facility based on Reinforcement Learning,” 2018, doi: 10.20944/preprints201805.0353.v1.
- [64] Z. Yuan *et al.*, “Fault data generation of lithium ion batteries based on digital twin: A case for internal short circuit,” *J Energy Storage*, vol. 64, Aug. 2023, doi: 10.1016/j.est.2023.107113.
- [65] Y. Jiang, W. Finnegan, F. Wallace, M. Flanagan, T. Flanagan, and J. Goggins, “Structural analysis of a fibre-reinforced composite blade for a 1 MW tidal turbine rotor under degradation of seawater,” *J Ocean Eng Mar Energy*, vol. 9, no. 3, pp. 477–494, Aug. 2023, doi: 10.1007/s40722-023-00279-w.
- [66] M. Rezaul Alam, M. J. E. Alam, T. K. Saha, and M. Sohrab Hasan Nizami, “A PV variability tolerant generic multifunctional control strategy for battery energy storage systems in solar PV plants,” *International Journal of Electrical Power and Energy Systems*, vol. 153, Nov. 2023, doi: 10.1016/j.ijepes.2023.109315.
- [67] S. Vettori, E. Di Lorenzo, B. Peeters, M. M. Luczak, and E. Chatzi, “An adaptive-noise Augmented Kalman Filter approach for input-state estimation in structural dynamics,” *Mech Syst Signal Process*, vol. 184, Feb. 2023, doi: 10.1016/j.ymssp.2022.109654.
- [68] H. Zhang *et al.*, “Digital Twin Enables Rational Design of Ultrahigh-Power Lithium-Ion Batteries,” *Adv Energy Mater*, vol. 13, no. 1, Jan. 2023, doi: 10.1002/aenm.202202660.
- [69] J. M. Reniers and D. A. Howey, “Digital twin of a MWh-scale grid battery system for efficiency and degradation analysis,” *Appl Energy*, vol. 336, Apr. 2023, doi: 10.1016/j.apenergy.2023.112000.
-

10.1016/j.apenergy.2023.120774.

- [70] O. Bondarenko and T. Fukuda, “Development of a diesel engine’s digital twin for predicting propulsion system dynamics,” *Energy*, vol. 196, Apr. 2020, doi: 10.1016/j.energy.2020.117126.
- [71] K. Zhang *et al.*, “Intelligent Optimal Control Strategy of Heat Pump System Based on Digital Twins,” *J Phys Conf Ser*, vol. 2452, no. 1, p. 012029, Mar. 2023, doi: 10.1088/1742-6596/2452/1/012029.
- [72] J. P. Spinti, P. J. Smith, and S. T. Smith, “Atikokan Digital Twin: Machine learning in a biomass energy system,” *Appl Energy*, vol. 310, Mar. 2022, doi: 10.1016/j.apenergy.2021.118436.
- [73] J. P. Spinti, P. J. Smith, S. T. Smith, and O. H. Díaz-Ibarra, “Atikokan Digital Twin, Part B: Bayesian decision theory for process optimization in a biomass energy system,” *Appl Energy*, vol. 334, Mar. 2023, doi: 10.1016/j.apenergy.2022.120625.
- [74] C. Blume, S. Blume, S. Thiede, and C. Herrmann, “Data-driven digital twins for technical building services operation in factories: A cooling tower case study,” *Journal of Manufacturing and Materials Processing*, vol. 4, no. 4, 2020, doi: 10.3390/JMMP4040097.
- [75] I. Natgunanathan, V. Mak-Hau, S. Rajasegarar, and A. Anwar, “Deakin Microgrid Digital Twin and Analysis of AI Models for Power Generation Prediction,” *Energy Conversion and Management: X*, p. 100370, Apr. 2023, doi: 10.1016/j.ecmx.2023.100370.
- [76] Y. Golizadeh Akhlaghi *et al.*, “A constraint multi-objective evolutionary optimization of a state-of-the-art dew point cooler using digital twins,” *Energy Convers Manag*, vol. 211, May 2020, doi: 10.1016/j.enconman.2020.112772.
- [77] A. Haghshenas, A. Hasan, O. Osen, and E. T. Mikalsen, “Predictive digital twin for offshore wind farms,” *Energy Informatics*, vol. 6, no. 1, Dec. 2023, doi: 10.1186/s42162-023-00257-4.
- [78] European Commission, “Digitalising the energy system - EU action plan,” 2022. doi: 10.2833/492070.
- [79] International Energy Agency, “Digitalization & Energy,” 2017. Accessed: Mar. 24, 2023. [Online]. Available: <https://www.iea.org/reports/digitalisation-and-energy>
- [80] T. Testasecca, M. Lazzaro, and A. Sirchia, “Towards Digital Twins of buildings and smart energy networks: Current and future trends,” 2023. doi: 10.1109/MetroLivEnv56897.2023.10164035.
- [81] M. Vaccari, G. Pannocchia, L. Tognotti, and M. Paci, “Rigorous simulation of geothermal power plants to evaluate environmental performance of alternative configurations,” *Renew Energy*, vol. 207, pp. 471–483, May 2023, doi: 10.1016/j.renene.2023.03.038.
- [82] F. Arrano-Vargas and G. Konstantinou, “Modular Design and Real-Time Simulators Toward Power System Digital Twins Implementation,” *IEEE Trans Industr Inform*, vol. 19, no. 1, pp. 52–61, Jan. 2023, doi: 10.1109/TII.2022.3178713.

-
- [83] Y. Liu, J. M. Zhang, Y. T. Min, Y. Yu, C. Lin, and Z. Z. Hu, “A digital twin-based framework for simulation and monitoring analysis of floating wind turbine structures,” *Ocean Engineering*, vol. 283, Sep. 2023, doi: 10.1016/j.oceaneng.2023.115009.
 - [84] C. S. B. Clausen, Z. G. Ma, and B. N. Jørgensen, “Can we benefit from game engines to develop digital twins for planning the deployment of photovoltaics?,” *Energy Informatics*, vol. 5, Dec. 2022, doi: 10.1186/s42162-022-00222-7.
 - [85] Y. Tan, P. Chen, W. Shou, and A. M. Sadick, “Digital Twin-driven approach to improving energy efficiency of indoor lighting based on computer vision and dynamic BIM,” *Energy Build*, vol. 270, Sep. 2022, doi: 10.1016/j.enbuild.2022.112271.
 - [86] L. Schäfers, D. Silber, R. Savelsberg, and S. Pischinger, “Efficient determination of driving resistance through system identification based on driving route information and weather data,” *Expert Syst Appl*, vol. 232, Dec. 2023, doi: 10.1016/j.eswa.2023.120755.
 - [87] E. Söderäng, S. Hautala, M. Mikulski, X. Storm, and S. Niemi, “Development of a digital twin for real-time simulation of a combustion engine-based power plant with battery storage and grid coupling,” *Energy Convers Manag*, vol. 266, Aug. 2022, doi: 10.1016/j.enconman.2022.115793.
 - [88] J. Zhang and X. Zhao, “Digital twin of wind farms via physics-informed deep learning,” *Energy Convers Manag*, vol. 293, Oct. 2023, doi: 10.1016/j.enconman.2023.117507.
 - [89] M. Guo, X. Fang, Q. Wu, S. Zhang, and Q. Li, “Joint multi-objective dynamic scheduling of machine tools and vehicles in a workshop based on digital twin,” in *Journal of Manufacturing Systems*, Elsevier B.V., Oct. 2023, pp. 345–358. doi: 10.1016/j.jmsy.2023.07.011.
 - [90] X. Zhao, M. H. Dao, and Q. T. Le, “Digital twining of an offshore wind turbine on a monopile using reduced-order modelling approach,” *Renew Energy*, vol. 206, pp. 531–551, Apr. 2023, doi: 10.1016/j.renene.2023.02.067.
 - [91] N. Vasilikis, R. Geertsma, and A. Coraddu, “A digital twin approach for maritime carbon intensity evaluation accounting for operational and environmental uncertainty,” *Ocean Engineering*, vol. 288, Nov. 2023, doi: 10.1016/j.oceaneng.2023.115927.
 - [92] X. Yu, Y. Yang, M. Du, Q. He, and Z. Peng, “Dynamic Model-embedded Intelligent Machine Fault Diagnosis without Fault Data,” *IEEE Trans Industr Inform*, 2023, doi: 10.1109/TII.2023.3245677.
 - [93] Y. Y. Hong and R. A. Pula, “Diagnosis of PV faults using digital twin and convolutional mixer with LoRa notification system,” *Energy Reports*, vol. 9, pp. 1963–1976, Dec. 2023, doi: 10.1016/j.egy.2023.01.011.
 - [94] F. J. Folgado, I. González, and A. J. Calderón, “PEM Electrolyser Digital Twin Embedded within MATLAB-Based Graphical User Interface †,” *Engineering Proceedings*, vol. 19, no. 1,
-

2022, doi: 10.3390/ECP2022-12676.

- [95] Z. Lei, H. Zhou, W. Hu, G. P. Liu, S. Guan, and X. Feng, “Toward a Web-Based Digital Twin Thermal Power Plant,” *IEEE Trans Industr Inform*, vol. 18, no. 3, pp. 1716–1725, Mar. 2022, doi: 10.1109/TII.2021.3086149.
- [96] M. O. Genc, “Development of a digital twin model for electric bikes based on urban experimental driving in full electric mode,” *Journal of Mechanical Science and Technology*, vol. 37, no. 8, pp. 4349–4359, Aug. 2023, doi: 10.1007/s12206-023-0750-9.
- [97] M. T. Qaiser, J. Ejaz, O. Osen, and A. Hasan, “Digital twin-driven energy modeling of Hywind Tampen floating wind farm,” *Energy Reports*, vol. 9, pp. 284–289, Oct. 2023, doi: 10.1016/j.egyr.2023.09.023.
- [98] S. Agostinelli, “DIGITAL TWIN MODEL FOR ZERO-ENERGY DISTRICTS: THE CASE STUDY OF ANZIO PORT, ITALY,” in *WIT Transactions on Ecology and the Environment*, WITPress, Dec. 2022, pp. 357–363. doi: 10.2495/SC220291.
- [99] Y. Li, X. Shen, and C. Zhou, “Dynamic multi-turbines spatiotemporal correlation model enabled digital twin technology for real-time wind speed prediction,” *Renew Energy*, vol. 203, pp. 841–853, Feb. 2023, doi: 10.1016/j.renene.2022.12.121.
- [100] X. Ma, C. Yang, X. P. Zhang, Y. Xue, and J. Li, “Real-Time Simulation of Power System Electromagnetic Transients on FPGA Using Adaptive Mixed-Precision Calculations,” *IEEE Transactions on Power Systems*, vol. 38, no. 4, pp. 3683–3693, Jul. 2023, doi: 10.1109/TPWRS.2022.3199181.

Chapter 3

Methodology for Developing Digital Twins

DTs are a recent technology that has primarily developed over the last five years, especially in the energy sector. At the beginning of this PhD course, the research literature on the topic was limited, with only a few studies related to energy applications. It was unclear how a DT differed from other models, and no established methodology for their development existed. Even recently, as mentioned in the previous chapter, confusion persists in many works reviewed during the state-of-the-art analysis, where DTs are often mistaken for DMs. Chapter 3 of this thesis, therefore, elaborates on major components designed for DTs as part of this research project. Given the shared framework among the applications presented, such as the use of ML algorithms and specific physical models developed in TRNSYS, the aim is to present a general methodology for developing DTs in SESs. It starts from data collection and management infrastructure through sections on modelling and services (i.e. optimization algorithms) and ends with brief references to user interfaces. The objective of this work is to guide future researchers and IT developers to create solid and robust DTs to provide a real-world benefit to the energy infrastructure. However, it is important to note that each application has its unique features, including the hyperparameters used in various ML and optimization algorithms, as well as the composition of the TRNSYS model. Therefore, specific details for each application will be provided in the respective chapters.

3.1 Data Integration and Processing Strategies

Data integration and processing form the bottom of a DT, without which it may be only considered a DM, even if based on data. This integration enables DTs to replicate the physical systems with a high degree of accuracy and it could provide suggestions or high-precision control. This section focuses on the simplest wired connection mechanisms to highly sophisticated architectures with context brokers, middleware APIs, and data preprocessing techniques.

One efficient approach through which data can be integrated into a DT system is using wired connections among the sensors, data acquisition systems, and the computing system on which

the DTs are run. The low latencies and high data fidelities can be very noticeable, along with an almost insignificant interference and minimal signal degradation because of the wired connection. Moreover, this approach enables easy debugging as the physical connections are visible and easily traceable. However, this is applicable in small-scale systems since a large number of sensors deployed over a wide geographical area exponentially increases the system complexity and costs. Furthermore, it would need an accurate and frequent maintenance of cables, connectors, and contacts. In this context, for developing a DT of one of the laboratories in Building 9 at the University of Palermo, an existing wired infrastructure was utilized [1].



Fig. 6. SolarLab Experimental Facility in University of Palermo.

Fig. 6 displays sensors for monitoring illuminance (4 luxmeter probes Delta Ohm) and luminance which are connected to a data acquisition system provided by Yokogawa. The data are then imported into a Windows PC connected to the system and collected and visualized using the Labview [2] infrastructure. Additional details on the experimental setup are available in [3]. Using this system, it was possible to fine-tune a DM using Dialux [4] to understand the illuminance distribution inside the room virtually. Furthermore, it enabled the creation of a DT whose dashboard was used to provide suggestions and insights for improving the visual comfort of occupants as will be presented in Chapter 7.

In contrast, when dealing with more complex systems, an IoT-based infrastructure is needed as it simplifies data gathering and processing. One approach involves using a context broker such as Kafka [5], which is employed within the DigiBUILD project and in this thesis for the Emotion and Veolia pilots [6]. Kafka is a distributed streaming platform that handles large volumes of data with high throughput and low latency, making it well-suited for processing real-time data streams. This capability is fundamental for applications requiring immediate responses to changes in the dataset, such as DTs. These message brokers can be integrated with Extract, Transform, and Load (ETL) processes to enhance data handling capabilities. Data from various sources is extracted, transformed into the appropriate format, and then loaded into

databases or data lakes. This process ensures the collected data is cleaned and structured, making it suitable for analysis or modelling. Moreover, these online databases provide a centralized digital repository easily accessible to DTs. Within the DigiBUILD project [7], WP2 is responsible for ensuring the collection of high-quality data and making it available to the developed services. Another approach could involve integrating data streams from software provided by device manufacturers, as seen in the SOFC case studied in this PhD thesis. Here, historical data gathered from proprietary software was used to train the SOFC ML models by directly accessing a dashboard. However, it should be noted that integrating models and algorithms with third-party applications is challenging, as it often requires user interaction, making automated control difficult.

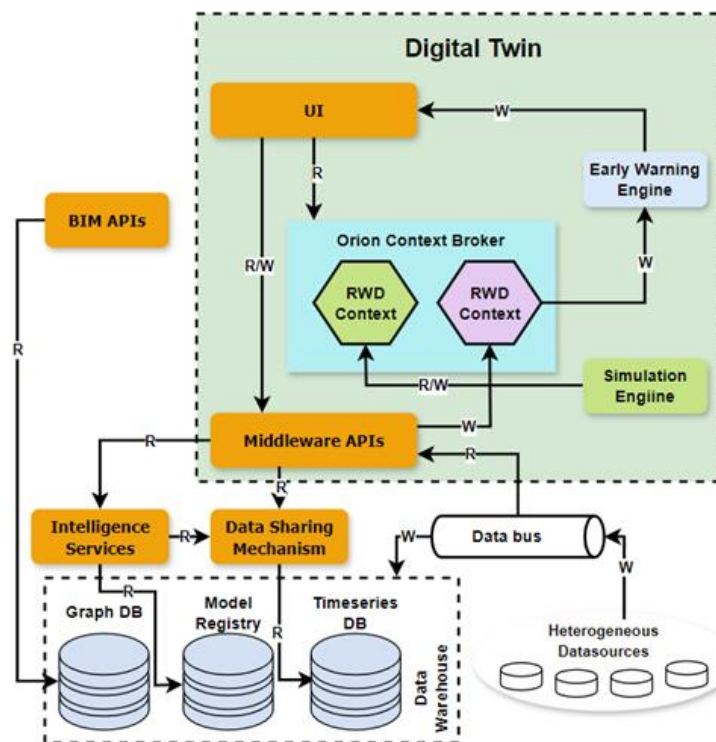


Fig. 7. DigiBUILD DT Architecture.

In Fig. 7 is represented an example of DT architecture expected in the DigiBUILD project [8]. The Middleware Application Programming Interface (API) is in charge of interacting with the underlying layer of DigiBUILD architecture, the Intelligence Services and the Data Sharing mechanism. Therefore, when dealing with DT which have different functions and include different algorithms the middleware APIs are used to manage the exchange of data between different services and data sources. It is most commonly done with the programming language Python because of its strong libraries, covering even data manipulation up to ML libraries. Middleware APIs can make data flow easily through different parts of a proposed DT architecture: from sensors to processors and back into the DT models in an effective mode. Precisely, being able to effectively do data preprocessing in making the loaded data into DT

models results in a reliable, quality dataset.

Preprocessing could be defined as the group of activities involving cleaning and transforming raw data into analysable and modellable data. As previously stated, when dealing with ML modelling it is necessary to process data for avoiding that bad entries affect the model reducing its reliability. Indeed, sensors can register anomalous data which are very far from the real one due to many factors, leading to the generation of “outliers”: extreme values that highly differ from other measurements. Analytical and ML approaches to detecting and filtering outliers from the dataset are collectively applied, assuming for instance that correct data belong to a certain known distribution. For simplifying this process, many python libraries are available and were used for developing the ML models in this thesis, for instance Scikit-learn library [9]. In this library, different algorithms and functions are used to identify both outliers (such as `covariance.EllipticEnvelope`, `ensemble.IsolationForest` and `neighbors.LocalOutlierFactor`). In the first, outliers are removed if data do not belong from a known distribution [10], while in the second the isolation forest algorithm [11] is based on outliers removal based on data pattern. Similarly, the Local Outlier Factor algorithm [12] measured the local density deviation of a certain measurement compared to other ones, following the theory of the k-nearest neighbours. Finally, once the data is filtered, it is necessary to exclude entries from the dataset that are physically implausible, which could indicate an issue with meter registration.

Data can come in varying scales and units, for this reason, when training a ML algorithm, it may need normalization or standardization. The first scales the data to a range, usually $[0, 1]$ or $[-3, 3]$. On the other hand, standardization transforms it so that it has a mean of zero and a standard deviation of one. Therefore, all data inputs to the algorithm should be equal contributors from their different sources and equally scaled, thereby avoiding biases which may lead toward the identification of dependencies from one quantity instead of another. In cases where a time series is needed for training a model (such as in the case of Long Short-Term Memory algorithms), it is necessary to detect and correct missing data. This is one of the major problems concerning the sensor network, which remains caused by reasons ranging from sensor malfunctions and transmission errors to the phenomena themselves. In these cases, techniques of interpolation, regression, or more advanced methods of ML were utilized to estimate the intermediate values, filling in missing values in a dataset, so that it is complete and ready for analysis.

Feature management process marks the most relevant features that contribute majorly to the optimum performance of the data-driven models and it is related to the selection of best inputs and outputs of the model. Precisely, feature extraction creates new features from the existing data to capture essential patterns or relationships between features that one can notice. For

instance, in the context of this project, the efficiency of DHN boilers or SOFC was calculated, the same for energy consumption or speed related to EVs. On the contrary, feature selection is the process of picking out only relevant features and discarding irrelevant or redundant ones in the whole dataset. This procedure enhances the efficiency, accuracy and training speed of the DT models, ensuring only necessary data need to be processed. However, to select the correct features, it is fundamental to understand correlations among different variables in the data before developing DT models. Physical and linear dependencies between the measured data are pointed out with the help of a correlation analysis. Physical correlations are based on known scientific or engineering principles, and linear correlations can be identified through the use of statistical methods such as Pearson's correlation coefficient [13]. Therefore, in the current SoA, a common practice is to develop a correlation matrix which considers Pearson's coefficient r_{xy} , in python directly implemented in Pandas library, which is defined as:

$$r_{xy} = \frac{\sum_i x_i y_i - n \bar{x} \bar{y}}{\sqrt{\sum_i x_i^2 - n \bar{x}^2} \sqrt{\sum_i y_i^2 - n \bar{y}^2}} \quad (1)$$

Where:

- n is the sample size.
- x_i, y_i are the individual sample points indexed with i .
- $\bar{x}$ is the sample mean of x vector.
- $\bar{y}$ is the sample mean of y vector.

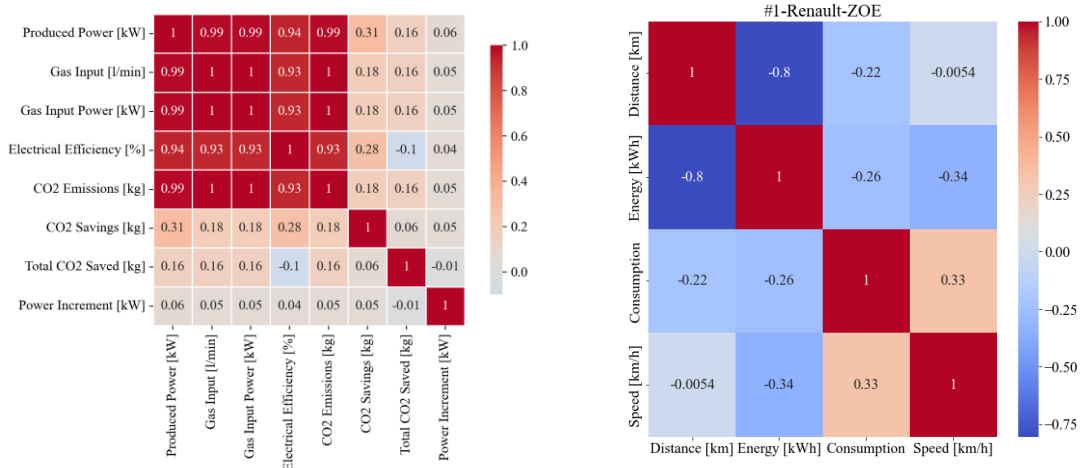


Fig. 8. Example of Correlation Matrix Heatmap for SOFC (left) and EV (right).

This correlation matrix was generated for boilers, EVs and SOFC in the context of this project and the output, represented as a heatmap, is represented on the left of Fig. 8 for SOFC and right for an EV of Emotion Pilot. With the knowledge of these correlations, accurate

models for DTs were developed by selecting accordingly input features and output features. However, details on model feature selection and extraction will be presented in the next chapters of this thesis.

3.2 Data-Driven Models

Using white or grey-box approaches involves the knowledge of the entire physical system of the considered energy system which is difficult to be accurately estimate in cases of complex systems. In fact, to correctly model the physical behaviour it is necessary to follow a multi-physics approach, which is difficult even coupling different tools such as COMSOL [14] or TRNSYS [15]. A black-box approach, utilizing validated and accurate models, allows for the estimation of one or more outputs based on correlated inputs. Within this research project framework, several ML algorithms were employed and compared, including ANNs, Gradient Boosting (GBoost), XGboost, RF, Long Short Term Memory (LSTM) and Polynomial Regression, with the operational logic of each explained in subsequent sections.

Nevertheless, as mutual correlations may depend on the magnitude of quantities, in these studies, data was standardized using the scikit-learn Python library, which standardizes features by removing the mean and scaling to unit variance. Although this preprocessing step enables the evaluation and enhancement of the chosen ML approach, the presence of outliers should be minimized as much as possible, as they can affect the quality of the modelling process. However, in this chapter, the focus is on the fundamentals of used algorithms, while details on hyperparameters chosen will be found in the following chapter of this thesis.

3.2.1 Supervised Machine Learning Algorithms

In the framework of ML algorithms, Polynomial Regression represents an extension of linear regression that fits a polynomial equation to the dataset. This approach allows the model to capture non-linear relationships between the predictor (independent) and response (dependent) variables, which is a significant advantage over linear models. Empirical studies have demonstrated that Polynomial Regression can effectively model curved trends in data, making it a valuable tool in various scientific and engineering fields where linear models are inadequate [16]. The benefits of its use are related to model complex, non-linear relationships with very low computational effort.

Random Forest, GBoost and Xgboost algorithms belong all together to the family of Decision Trees. This supervised learning method is used to predict the output by imposing simple decision rules inferred from the input data. The rooted tree is a graph where all the branches and leaves start from a defined origin. Precisely, the input space is divided into smaller

portions of the space based on the different hypotheses carried out by the model. GBoost is a ML technique first introduced in 1999 [17] and used for both regressions and classifications. As stated in [17] it was first defined by Kearns as “an efficient algorithm for converting relatively poor hypotheses into very good hypotheses”. The model, composed of a weak hypothesis, the first group of decision trees, is updated with additional decision trees capable of reducing model loss. The final model will be composed of an optimized sequence of decision trees capturing even unknown and complex patterns within the data. It must be noted that this algorithm tends easily to overfitting so the estimators and the random states inside the algorithm should be selected accurately.

XGBoost, is one of the most used and up-to-date ML algorithms, as the evolution of GBoost algorithm, where multiple ML algorithms are combined to increase accuracy and reduce computational effort. Precisely, it uses parallel and dispersed computing of decision trees and GBoost algorithms. Decision trees are built sequentially to correct errors in the previous step and, to avoid overfitting, it uses a clever penalization of trees and a proportional shrinking of leaf nodes.

```
X_train, X_test, y_train, self.y_test = train_test_split(X,
                                                         y,
                                                         test_size=0.2,
                                                         random_state=2)

# Standardizing features
self.scaler = StandardScaler().fit(X_train)
X_train_scaled = self.scaler.transform(X_train)
X_test_scaled = self.scaler.transform(X_test)

# Building the XGBoost regressor
self.model = XGBRegressor(n_estimators=150,
                           subsample=1,
                           learning_rate=0.2,
                           max_depth=7,
                           random_state=rst_model)

# Training the model
self.model.fit(X_train_scaled, y_train)

# Making predictions
self.y_pred = self.model.predict(X_test_scaled)
```

Fig. 9. Example of code for XGBoost model for the SOFC

In Fig. 9, the typical hyperparameters that can be selected for an XGboost algorithm are presented. Specifically, the number of estimators corresponds to the number of trees iteratively added to the model. The learning rate acts as a shrinkage factor applied to new trees to prevent overfitting. Moreover, although increasing tree depth could enhance model precision by adding complexity to each tree, [17] recommends using values between 4 and 8.

In RF algorithm, multiple decision trees are used to improve the accuracy of the model and avoid overfitting [18]. This method, belonging to the family of ensemble learning methods, overcomes the limitations of Decision Tree algorithms in having low bias and high variance. In pure RF, the trees are built randomly while, to improve the quality of predictions, in the most recent versions the trees could be generated based on input data. This results in a model, suitable for regression tasks, even in scenarios where input data are noisy or complex. Moreover, in cases of multiple inputs, RF could be used to assess the relevance of each variable in order to better understand the problem [18]. It is notable that the number of trees and their depth, particularly the latter, are usually higher than those used in GBoost and XGBoost. This happens because, in boosting processes, new trees are incrementally added, and it is preferable to have simpler individual trees to avoid a sure overfitting. Conversely, in the RF algorithm, a single forest composed of diverse individual trees is used, allowing each tree to possess a more complex structure.

3.2.2 Supervised Deep Learning Algorithms

One of the most used classes of regression models based on Deep Learning (DL) are ANNs which emulate the functional characteristics of human brain neurons and their synaptic connections. The simplest form of an ANN is the single-layer perceptron which comprises only an input layer and an output layer. Focusing on research purposes, multilayer perceptrons (MLPs) are widely used, where, as suggested by the name, one or more hidden layers of neurons are modelled between the input and output layers to enhance the model's ability to capture the underlying patterns of the data points it is trained to estimate.

An iteration and optimized process of finding different layers is carried out to find the correct value of hyperparameters, which in ANN represent the number of hidden layers and neurons for layers. This process is necessary to finetune the network to maximize the accuracy of the model while reducing the time spent in the iterative process. Additionally, since using internal optimization process of DL algorithms, is affected by the randomness of parameters, different processes should be carried out for different random states.

```

# Splitting the dataset into training and testing sets
X_train, X_test, y_train, self.y_test = train_test_split(X, y,
                                                         test_size=0.2,
                                                         random_state=42)

# Standardizing features
self.scaler = StandardScaler().fit(X_train)
X_train_scaled = self.scaler.transform(X_train)
X_test_scaled = self.scaler.transform(X_test)

# Building the MLP Regressor
self.model = MLPRegressor(hidden_layer_sizes=(10, 62, 100, 50, 10),
                           max_iter=100000000,
                           verbose=False,
                           solver='adam',
                           learning_rate='adaptive',
                           random_state=rst_model,
                           activation='relu')

# Training the model
self.model.fit(X_train_scaled, y_train)

# Making predictions on the test set
self.y_pred = self.model.predict(X_test_scaled)

```

Fig. 10. Example of code for ANN model for the SOFC

LSTM networks, an advanced variant of Recurrent Neural Networks (RNN), are used to individuate long-term dependencies, a capability encountered across various fields such as thermal comfort or energy demand, and predictive maintenance. The distinctive architecture of LSTM includes memory cells that replace the standard neuron units in the hidden layer of RNNs. Given its capability to handle long and short-term dependencies, LSTM is particularly suited for tasks requiring the prediction of time-series data. The model architecture presented in this thesis was developed using Keras library.

3.2.3 Supervised Algorithms Validation Methods

Once the models were defined, to verify the accuracy of each model and its robustness it is needed to select the parameters for evaluating it. In these studies, it was chosen to consider the following parameters:

- Mean Absolute Error (MAE)
- Mean Absolute Percentage Error (MAPE)
- Mean Squared Error (MSE)
- Root Mean Square Error (RMSE)
- R^2 Score
- Weighted Score (WS)

The MAE is used to measure the average of the absolute errors, providing a linear score without exaggerating outliers. In particular, it measures the average absolute differences between actual and predicted values.

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

The MSE measures the average of the squares of the errors, giving more weight to larger errors.

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

The MAPE then calculates the average absolute percentage difference between actual and predicted values.

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (3)$$

The RMSE is the square root of MSE, providing error in the same units as the target variable, which is useful for interpretability.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4)$$

R^2 Indicates the proportion of the variance in the dependent variable that is predictable from the independent variables.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (5)$$

WS is a score proposed in the SOFC study which try to assess the issue of considering different metrics by combining all in a single one in order to be evaluated autonomously by the DT when it must chose the best model. Since all the quantities must be comparable, the first steps include the normalization of each result basing from best and worst possible value (for R^2) and from best and mean value measured in output data (for other parameters). Once these are calculated, the mean of all the values will give the WS. The equation is represented in (5).

$$\text{WS} = \frac{1}{2k} \left(k \cdot (1 - R^2) + \sum_{i=1}^k \frac{M_i}{M_{\text{mean},i}} \right) \quad (6)$$

In these equations:

- n is the total number of observations or data points.
- y_i represents the actual observed values in the dataset.
- $\hat{y}_i$ represents the predicted values corresponding to y_i .
- $\bar{y}$ is the mean of the observed values (y_i)
- k is the number of metrics other than R^2
- $R_{\min}^2$ and $R_{\max}^2$ are the minum and maximum possible values for R^2 (respectively 0 and 1)
- $M_{\text{mean},i}$ is the measured values for metric k applied to the mean value of the dataset
- M_i is the metric k score

3.2.3 Unsupervised Algorithms

Depending on the case study considered, investigation of data correlations could reveal that data are not highly correlated. In these cases, if even decision tree-based algorithms fail to detect dependencies between input and output features, it may be useful to change the approach and use unsupervised algorithms. This kind of algorithm, more than finding input-output relationships, aims to find different patterns in data that can be read by the data analyst for further investigation [19]. Indeed, while in supervised learning human supervision is needed since it must label input data and related output data, in unsupervised learning the machine tries to find any patterns in the data without any human intervention. Clustering techniques, such as K-means clustering are one of the most used unsupervised algorithms in the current research field. Starting from selecting the number of clusters n , the algorithm will individuate n point (centroid) in a space x -dimensional where x is the number of features whose patterns want to be investigated. The position of these centroids will change iteratively until the distance between centroids and individual data points in a cluster is minimized (Fig. 11).

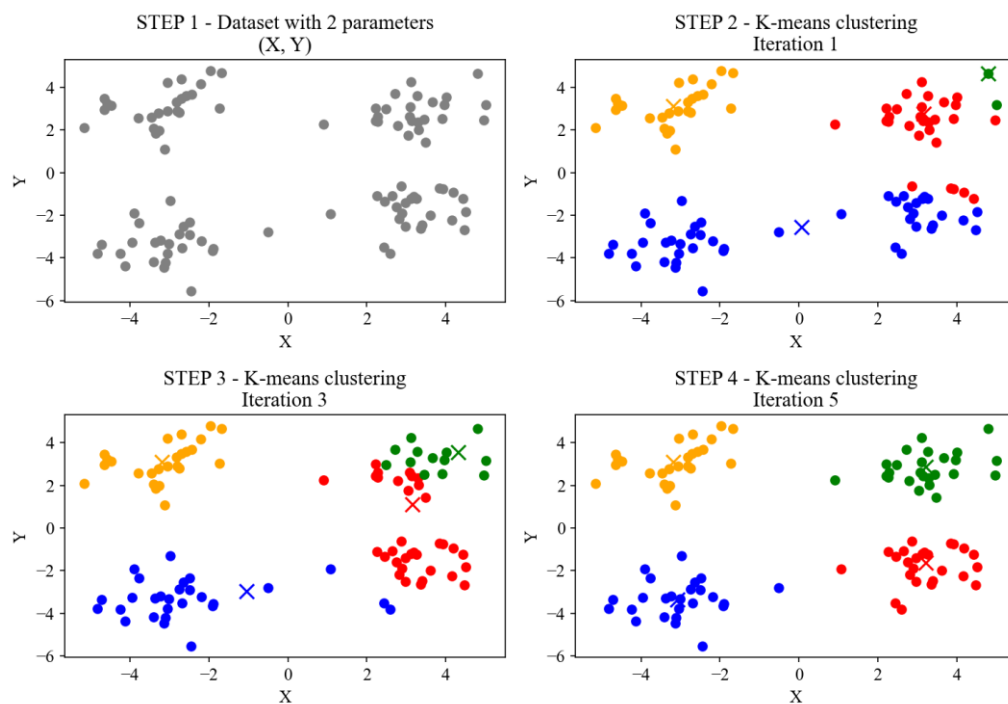


Fig. 11 Process of Clustering Based on k-means

In this thesis, clustering techniques were used to investigate patterns in the energy consumption of EVs since no strong dependencies were noted between available features. However, using k-means some consumption patterns were identified between consumption and vehicle speed enabling the possibility of associating a certain consumption depending on vehicle typology and speed.

Finally, the most used techniques for the validation process of clustering include the elbow

rule, silhouette score, and Davies-Bouldin Index (DBI). The elbow method is a methodology for finding the correct balance in choosing the number of clusters in a data set. Plotting the explained variation as a function of the number of clusters, the elbow represents the point where the rate of decrease changes. That point will mean a good trade-off between the minimization of the number of clusters and the maximization of variance explained by them since increasing the number of clusters will decrease the within-cluster variance. However, at some point, the marginal gain will decrease, and this forms an "elbow" in the plot. The silhouette score is a measure of the quality of clustering and quantifies the similarity of an object to its cluster (cohesion) compared to the other clusters (separation). The Silhouette Score (SS) lies within the range $[-1, 1]$, and high values indicate that it is very similar to its cluster but quite dissimilar from other neighbouring clusters. A silhouette value near zero indicates that the object is on or very close to the decision boundary between two neighbouring clusters. The DBI evaluates the performance of the clustering algorithm by calculating the average similarity ratio of every cluster with its most similar cluster. Hence, the lower the DBI values, the better the clustering performance as this indicates lower intra-cluster distances and higher inter-cluster distances, which represents the dissimilarity between clusters.

3.3 Physics-Based Modelling

Before the widespread adoption of sensors and tools for metering real systems, models related to energy systems, including buildings, relied solely on physics-based relationships. These models considered principles from thermodynamics, detailed according to the type of application, or fundamental electrical engineering concepts. Even though these models could be developed using state-of-art equations, implemented in calculation sheets or software such as Matlab [20], or even directly coded in programming languages, during those years, the majority of models in the energy field utilized software such as Modelica [21], TRNSYS [15], or EnergyPlus [22]. These software packages, equipped with user interfaces, were capable of creating complex models that integrated different components and properties. They formed the foundation of the current state-of-the-art simulation tools, allowing researchers to investigate various scenarios and advance research on energy systems. Each component in these software packages (e.g., a PV panel, a solar collector, a boiler, a wall) includes referenced equations that provide strong reliability for simulation purposes.

In this thesis, different approaches were followed depending on the case study considered. For modelling complex energy systems, such as DHNs, TRNSYS software was predominantly used, while for buildings, EnergyPlus was often employed. These software tools allow for the input of properties for each component, the control logic linking the components, and a weather

file to obtain reliable results. However, due to limitations in the availability of components within these tools, new models based on correlations or thermohydraulic equations were also implemented to simulate novel components in the studied DHNs, such as absorption chillers or pipes.

3.3.1 Hydraulic Pipe – Type 496

The choice of the pipes is one of the most relevant parts of the design of a DHN. In general, all the pipes used in a DHN are pre-insulated pipes [23] with a layer of insulation around a pipe made of steel or, more rarely, plastic. The research on the optimum between the minimization of pressure drops and the reduction of investment costs is still a hot topic in the scientific literature [24]. Precisely, once the flow rates in each branch of the network have been determined during the design phase, a hydraulic study must be carried out to determine the appropriate pipe diameters. This involves an iterative process of adjusting the diameters until the desired pressure drops are achieved in the most stressed branch.

Regarding analysis during the operation of a DHN, no TRNSYS type can calculate the pressure trend in a network with not-smooth pipes, so basic hydraulic relations (including flow velocity and Reynolds calculations) and equations 7 and 8 were implemented in a new component, Hydraulic Pipe – Type 496, created using the Fortran language [25].

The specific pressure drops $\Delta P_{s,i}$ in the branch i can be calculated using the Darcy-Weisbach relation [26]:

$$\Delta P_{s,i} = \frac{\Delta P_i}{L_i} = \frac{\lambda_i \cdot \rho_i \cdot v_i^2}{2 \cdot D_i} \quad (7)$$

In the previous equation, ΔP_i is the total pressure drop, L_i is the length of the pipe, λ_i the Darcy factor, ρ_i the density of the fluid and v_i and D_i are respectively the speed of the water and the diameter.

By the calculation of the Reynolds number Re_i on each branch i and the by the knowledge of the roughness of the pipe ε_{pipe} the Darcy factor is calculated as the same as [27] using approximations of the Colebrook-White relation:

$$\lambda_i = \begin{cases} \frac{64}{Re_i}, & Re_i < 2000 \\ 4.36E(-6) \cdot Re_i + 0.0233, & 2000 \leq Re_i < 4000 \\ \frac{0.25}{\left[\log \left(\frac{\varepsilon_{pipe}}{3.7 \cdot D_i} + \frac{5.74}{Re_i^{0.9}} \right) \right]^2}, & Re_i \geq 4000 \end{cases} \quad (8)$$

Utilizing inputs such as interior diameter, inlet pressure, and inlet flow rate, and parameters including density, viscosity, pipe roughness, and length, the model generates as outputs: velocity, Reynolds number, Darcy Factor, pressure drop, and outlet pressure. The choice of using the diameter as an input permits the use of this pipe also for future optimization purposes. Specifically, the developed type initially calculates the fluid velocity based on the provided diameter, inlet mass flow rate, and fluid density [26]. Then, the Reynolds number is determined using parameters such as density, interior diameter, velocity, and viscosity. Once the Reynolds number, roughness, and diameter are known, the Darcy factor is computed using Equation (8). Finally, the pressure drop in the branch is calculated as explained in Equation (7), and the outlet pressure is determined by simply subtracting the calculated pressure drop from the inlet pressure. By incorporating well-validated equations into a single type, the dynamic simulation of the hydraulic behaviour of DHNs on TRNSYS could be simplified. However, to also include the study of the thermal behaviour of the network, this type is used in combination with the available Type 31, so-called “Pipe or Duct”. This type can calculate, given geometric data, inlet flow rate and temperature, and ground temperature, the heat losses and consequently the outlet temperature.

A preliminary study of thermal behaviour was conducted at first. This revealed that in the worst case, the difference in temperature between the first and the last user is below 6 °C, and this happens when the flow rate is at its minimum (on the night) leading to minimum pressure drops. In this situation, considering the difference between real values and the calculated mean temperature and pressures, an analysis carried out on Engineering Equation Solver (EES) revealed that the viscosity change is below 3 % while for density this change is below 0.2 %. For this reason, in the first release of this Type, the density and the viscosity are assumed constant.

3.3.2 Absorption Chiller

Solar cooling technology is based on the use of absorption chillers, capable of generating cooling energy through high-temperature heat. In the fifth generation DHN developed in this

thesis, the absorption chiller was modelled to improve the functioning of Type 107 - Single Effect Absorption Chiller, which models the behaviour of this device based on a database file referring to manufacturers' catalogues. This model requires as inputs the temperature and flow rate of: the absorber's supply fluid, coming from the storage; the fluid to be cooled, coming from the network return piping; the fluid that removes heat from the absorber, necessary for its proper operation. Additionally, the machine size, chosen based on the cooling demand of the users and the desired setpoint temperature, i.e., the temperature of the refrigerated fluid exiting the absorber that circulates in the network's supply piping, needs to be set.

As for Type 107, the amount of heat $\dot{Q}_{set}$ that must be removed from the fluid to be cooled to change from the inlet temperature $T_{ch,in}$ to the setpoint temperature $T_{ch,set}$ is described by Eq. (9) [28]:

$$\dot{Q}_{set} = \dot{m}_{ch} C_p (T_{ch,in} - T_{ch,set}) \quad (9)$$

Where $\dot{m}_{ch}$ is the fluid flow rate, and C_p is the specific heat of the fluid.

The minimum generator temperature, contingent upon the desired chilled water temperature, was first determined. This calculation involved considering the performance curves outlined in the study by Cefarin [29], which modelled a single-stage water-ammonia absorption chiller, similar to the model available in [30]. Knowing the value of this temperature, it is possible to calculate the maximum theoretical amount of heat extractable from the fluid feeding the generator $\dot{Q}_{gen}$ through Eq. (10):

$$\dot{Q}_{gen} = \dot{m}_{gen} C_p (T_{gen,in} - T_{gen,set}) \quad (10)$$

Where $T_{gen,in}$ is the temperature of the hot fluid entering the chiller, $T_{gen,set}$ is the minimum temperature of the hot fluid compatible with chiller operation and $\dot{m}_{gen}$ is the flow rate of the fluid feeding the generator.

The machine's Coefficient of Performance COP_{chill} is estimated as suggested by the same study by knowing the temperatures of entering fluids and using Type 581 - Data Interpolator. Having obtained COP_{chill} and the amount of heat $\dot{Q}_{gen}$, it is possible to calculate the maximum energy removable to the network side fluid $\dot{Q}_{chill,max}$ using Eq. (3):

$$\dot{Q}_{ch,max} = \dot{Q}_{gen,ex} \cdot COP_{chill} \quad (11)$$

At this point, knowing the size of the machine, the maximum heat that can be removed from the generator side, and the amount of energy to be subtracted from the fluid to bring it to the setpoint temperature, the amount of heat removed from the fluid to be chilled will be the minimum of the three aforementioned quantities.

3.3.3 TRNSYS DHN Model Components

The two models presented in the previous sections, combined with the ones that will be described in the next subsections or detailed in Chapter 4, constitute the DHN model of a ring-shaped network developed on the TRNSYS environment. Precisely, The ring topology is notably prevalent in fourth and fifth-generation networks due to its facilitation of interaction with prosumers. The challenges, such as increased heat losses in the final pipe of the ring, are mitigated by the lower operating temperature. In this context, contemporary research is exploring new methodologies for optimizing the topology of fourth and fifth-generation DHNs [31]. Instances of the application of ring-shaped networks, where Combined Heat and Power Plants (CHPs) are supported by waste heat from datacentres, are documented in [32].

3.3.3.1 Energy Balance

For validating the DHN model developed, energy and mass balances in all the nodes of the network must be monitored. For this reason, details on the equations considered in this study will be presented. At each node of the network, it must be applied the mass conservation law and the first conservation of energy principle:

$$\sum_{i \in \text{inlets}} \dot{m}_i = \sum_{j \in \text{outlets}} \dot{m}_o \quad (12)$$

$$\sum_{i \in \text{inlets}} \dot{m}_i \cdot h_i = \sum_{j \in \text{outlets}} \dot{m}_o \cdot h_o \quad (13)$$

where $\dot{m}_i$ is the flow rate that enters the node and h_i is the specific enthalpy of the inlet flow rate. In addition, the subscript “o” refers to the outlet flows. In this case, it is assumed that the node is adiabatic and there is not any contribution of kinetic or gravimetric energy.

Different energy balances must be then considered for solving the system of equations.

The energy balance for each user j is:

$$\dot{Q}_{user;j} = \dot{Q}_{pros;j} - \dot{Q}_{cons;j} - \dot{Q}_{stor;j} = \dot{m}_j \cdot c_w \cdot (T_s - T_r) \quad (14)$$

where $\dot{Q}_{pros;j}$ is the hourly production of heat (if user j is a prosumer), $\dot{Q}_{cons;j}$ is the hourly demand of heat (for Domestic Hot Water, Space Heating or Absorption chiller), $\dot{Q}_{stor;j}$ the power that is stored, c_w specific heat of the water, $\dot{m}_j$ mass flow rate through the user j , and T_s and T_r the supply and return temperatures respectively.

For the producers:

$$\dot{Q}_{pro;j} = \dot{m}_j \cdot c_w \cdot (T_r - T_s) \quad (15)$$

For storages in the district:

$$\dot{Q}_{sto;j} = \dot{m}_j \cdot c_w \cdot (T_s - T_r) \quad (16)$$

For the main power plant:

$$\dot{Q}_{gen} = \dot{m}_{gen} \cdot c_w \cdot (T_s - T_r) \quad (17)$$

Where $\dot{m}_{gen}$ is the mass flow rate which flows through the main power plant.

The energy balance for the entire network is:

$$\begin{aligned} \dot{Q}_{gen} + \sum_{j=1}^{No. \text{ users}} (\dot{Q}_{user;j}) - \sum_{j=1}^{No. \text{ producers}} (\dot{Q}_{pro;j}) + \sum_{j=1}^{No. \text{ storages}} (\dot{Q}_{sto;j}) + \dot{Q}_{losses} \\ = 0 \end{aligned} \quad (18)$$

Where $\dot{Q}_{losses}$ are the environmental heat losses of the buried pipes calculated by using a TRNSYS component and its equations. This balance is a general equation so it should be modified for each different application, for instance, the producer member and storage member could be considered zero to be more conservative in the design phase.

3.3.3.2 Main power plant and 4th Generation User Substation

The main power plant should satisfy the needs of each building in the cluster [33]. It must produce heat at the highest temperature required by the buildings' heating systems, so the correct production system must be carefully chosen (gas turbine, alternative engine, solar collectors' plant, etc.). Since the main interest of this first analysis is to focus on the DHN, the CHP unit and the loads are modelled just through basic equations and datasheets: The CHP will provide hourly thermal energy to satisfy all the user needs, considering microturbine efficiency as in [24] for calculating the fuel and assuming that the flow rate passing through the CHP strictly depends on the loads.

Focusing on user substation, it is necessary to choose the type of connection between the network and the users. In the case of prosumers, the grid connection requires a specific connection and substation scheme study. For instance, it can be return-supply if the temperature of the source is hot enough or, vice-versa it can be a return-return configuration in which the water is taken from the return pipe and fed back into the return pipe after being heated. Lennermo et al. [34] suggested different substation configurations to achieve proper flow and pressure. To simulate the DHN of the case study, the flow rate needed by each user is first calculated for each timestep by fixing the variation of temperature between the supply and return ring and by knowing the thermal demand of the building. Controlled flow diverter valves and tee pieces TRNSYS types are used to manage the flow rates between hot and cold rings.

In the case of consumers, the tee piece is connected with the return loop and the diverter valve

with the supply one while in the prosumer substations, there are also diverter valves in the supply side and tee pieces in the return one.

Since the valve needs as input a control value (where 1 means that all the flow is directed to the user and 0 to the next branch), an algorithm has been developed:

$$Control = GT(\dot{Q}_{cons}, 0) \cdot \frac{\dot{m}_j}{\dot{m}_{in}} = 0 \quad (19)$$

Where $\dot{m}_{in}$ is the flow rate coming from the previous branch. In this equation if there is a consumption of heat from the user j , the flow rate $\dot{m}_j$ will be drawn from the hot side. For prosumers the same algorithm is applied, using $\dot{Q}_{prod}$ in case of heat production and water injection on the supply side.

3.3.3.3 5th Generation User Substation

in a Fifth Generation District Heating (5GDH) network, every user has a substation where a reversible HP is located. In the case study of a fifth-generation network, a water-water type HP was considered interfacing with the network (source) and the user, drawing feed fluid from the supply piping and, after exchanging heat with the user, releasing it into the network's return piping.

To achieve more accurate results a macro function was developed to simulate the behaviour of a reversible water-water HP based on the performance curves available in scientific publications [35] and manufacturer catalogues of some HP models [36]. The TRNSYS macro for the HP is expanded and detailed in Fig. 12. For operation in heating mode, the study by Hüsing et al. [35] was considered for calculating the Coefficient of Performance (COP). Using Type 581 - Data Interpolator, based on the required power, the system will operate at a different frequency and finally, the COP is calculated also considering the evaporator outlet temperature and condenser inlet temperature. To calculate the EER in cooling mode operation, a water-to-water chiller power curves provided in [36] were referred to and, as in the previously described case, Type 581 is used to identify the EER, using as input the partial load ratio and the temperature of the fluid entering the condenser. The obtained coefficients of performance, the energy required by the user, the temperature, and the flow rate on the source side are inputted into the calculator named "Exchanger", which calculates the temperature of the fluid exiting the HP and injected into the return piping, employing applications of the first law of thermodynamics [38].

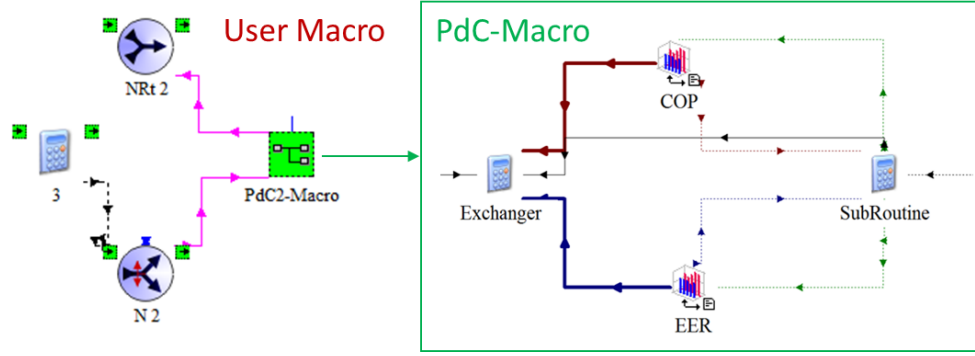


Fig. 12. Macro of the fifth-generation user substation.

3.4 Optimization Algorithms for Digital Twins

The increasing complexity of future energy networks, involving distributed heat source in DHNs, smart grids, and prosumers, presents significant challenges in terms of management. In order to achieve decarbonization goals while simultaneously reducing costs, it is essential to implement optimal-pursuing algorithms. In this framework, according to [37], ML based on big-data collection is a promising solution for speeding up the process of optimum research and for increasing efficiency in DHNs. In general, the search for optimal solutions involves finding the minimum or maximum value of an objective function (or multiple objective functions in the case of multi-objective optimization). In the context of energy, this function is typically related to the minimization of costs, energy consumption, or emissions. The objective function represents the final step of the optimization problem, which is subject to various constraints that limit the solution space [38].

In this thesis, the optimization process will be a key part of DT frameworks, since developing, according to user need, an optimization problem will guide toward the objective of decarbonization or cost reduction. Two different optimizations were provided for the DTs of two very different case studies: the first involve the optimization of heat production in a 3rd Generation DHN, and the second involve the optimization of a microgrid. However, as details on these applications will be provided in the next chapters, this section will introduce the libraries and the theory related to the optimization algorithms that were considered in the context of this thesis.

3.4.1 Fundamentals of Optimization Problems Theory

An optimization problem represents a scenario where an objective function is defined, which is a quantity to be maximized or minimized, and is normally denoted by f [39]. As a function, its value depends on certain variables, referred to as "decisional variables", which are the ones to be determined. These are symbolically represented as $(x_1, x_2, \dots, x_n)$. For example, if a

thermal storage tank is being designed, the decision variables might include the radius, height, overall heat transfer coefficient, and generally a series of variables for each of which the optimal value needs to be identified.

The decision variables are physically structured in the form of a vector or "array", where each variable is placed in a specific position that must be remembered and maintained fixed throughout the optimization process. Compactly, arrays/vectors or matrices will be written in bold. The optimization problem can be compactly written as:

$$\text{minimize/maximize } f(\mathbf{X}) \quad (20)$$

Where $\mathbf{X}$ is the vector of decision variables, expressed as:

$$\mathbf{X} = (x_1, x_2, \dots, x_n) \quad (21)$$

The goal is to minimize or maximize this function by varying all values within the array $\mathbf{X}$. This means choosing the optimal values for each x_i while respecting a set of constraints, which may be physics-based, economic, etc. Constraints are usually divided into two families:

- **Inequality Constraints:**

$$g_i(\mathbf{X}) \leq 0, \quad \text{for } i = 1, 2, \dots, m \quad (22)$$

where g_i is a generic constraint function. This function g_i imposes a condition on some of the variables (not necessarily all). Inequality constraints are typically written in the ≤ 0 form, as software like MATLAB or Python recognizes only this form. However, variables usually have "upper" and "lower" bounds, which are the maximum and minimum limits.

- **Equality Constraints:**

$$h_i(\mathbf{X}) = 0, \quad \text{for } i = 1, 2, \dots, p \quad (23)$$

It may happen that a specific combination of values must equal a certain number. However, sometimes the use of equality constraints must be evaluated conveniently since they can limit the feasible region which is the ensemble of variables that follow the imposed constraints. Indeed, the software performs calculations without ratio, thus constraints guide the solution space to feasible regions.

In the energy field, optimizations often involve "multilevel" problems, recognizing different levels, each deserving appropriate attention: synthesis, design and operation. The first, **synthesis**, concerns the plant layout, identifying which components should be included in the system and which should not. When planning to install a plant, not all possible components

may lead to an efficient system in every aspect. It is possible that only certain components considered for installation are actually necessary to achieve the desired configuration. Therefore, it is needed to determine which components to include and which to exclude. This is done through a synthesis operation, identifying a plant layout that will be the basis for subsequent decision levels.

This level can also consider components like connectors, heat recovery units, or piping that connects in series or parallel: anything recognized as having a degree of freedom in its inclusion or exclusion. For plants with a large number of components, it is necessary to initially include all components in a "superstructure", a layout containing all possible plant solutions. Each component in the superstructure is associated with a Boolean synthesis variable (0/1), indicating its inclusion (1) or exclusion (0) post-optimization. Secondly, **design** optimization involves determining the optimal size of the components. After deciding on component inclusion in the layout, the size must be determined. The software should provide the size or a selection parameter for the component. For instance, for a DHN, it could be the size of the pipes, defined by their diameter. Variables are not entirely independent; if the decision variable excludes a thermal storage tank, the program should not suggest a specific size for it. Therefore, design-level constraints must be operational only if the component is included, ensuring rational solution choices. Finally, **operation** level provides guidance on the operational modes of each component over time. The software needs to understand the elementary time unit for operational indications. If studying a rapid transient phenomenon, a time unit of a few milliseconds might be set. For energy plants with seasonal variations, an hourly time step is common, as electricity prices also vary hourly. The synthesis, design, and operation decision variables together form the vector $\mathbf{X}$.

3.4.2 Optimization Algorithms

After defining the optimization problem, constituted by inequality and equality constraint, the variables and the objective function, it is essential to introduce the solution methodologies that can be employed ranging from mathematical approaches to more modern, nature-inspired algorithms. Optimization algorithms are diffused in various applications in engineering, economics, ML, and industrial research. Basically, these algorithms are computational methods for finding the best feasible solution among a set of available solutions, often maximizing or minimizing a particular (or multiple in case of multi-objective optimization) objective function. Optimization plays an important role since it applies to many domains, providing gains in effectiveness, cutting costs, and bringing better performance. These methods can, with some generalities, be divided into two large categories: deterministic and stochastic [39].

Deterministic optimization methods are both predictable and consistent in their outcome; that is, from any specific starting point, these procedures will take a set path to converge to the same solution each time. The predictability of these methods comes in because they rely on precise mathematical models and algorithms that do not contain an element of randomness, and are majorly useful for convex problems or when dealing with problems that have well-defined mathematical structures under which exact solutions are preferred. Examples of these deterministic optimization methods include:

- **Linear Programming:** It has the purpose of optimizing a linear objective function subject to a set of equality and inequality constraints which are also linear.
- **Quadratic Programming:** This is the generalization of linear programming where the objective function is quadratic but the constraints are still linear.
- **Newton's Method:** This involves the use of the gradient and Hessian of an objective function in finding a local optimum. It is mostly used in unconstrained optimization.

In general, such techniques are quite efficient and very accurate provided that the problem structure is well known. However, they may be bad options for very complicated problems or problems that are nonlinear or in large-scale settings. These methods are typically efficient and provide high accuracy when the problem structure is well understood. However, they may struggle with complex, non-linear, or large-scale problems due to computational constraints.

In contrast, stochastic heuristic and meta-heuristic optimization methods [40] incorporate randomness into their search processes enabling them to explore a broader solution space and potentially escape local optima, which is particularly useful in complex problems. These methods are surely necessary when the problem studied is mathematically complex, involves uncertainty, or when a global rather than a local optimum is desired. In this thesis, as highlighted in the study of the SoA, for solving energy-related problems these types of algorithms are often used, precisely one of the family of GA (such as Evolutionary Strategy or Non Sorting GA) or other stochastic methods [41]. Precisely, for the development of the optimization algorithm for the case study presented, these stochastic algorithms were considered and compared:

- **Genetic Algorithms:** GAs are one of the most used algorithms in energy efficiency application. This algorithm is based on the principles of natural selection and genetics: using a population of potential solutions to a problem and applying genetic operators (such as selection, crossover, and mutation) the population of solution will improve itself (evolving) over time [39].
- **Non-dominated Sorting Genetic Algorithms (NSGA).** NSGA were firstly designed for multi-objective optimization problems [42]. They sort and rank solutions based on dominance criteria, ensuring a diverse set of Pareto-optimal solutions. NSGA-II, an

improved version, incorporates fast non-dominated sorting, elitism, and crowding distance mechanisms to maintain solution diversity and convergence

- **Particle Swarm Optimization (PSO):** PSO models the social behaviour of birds flocking or fish schooling [43]. Each particle represents a potential solution and adjusts its position based on personal experience and the collective experience of the swarm. This collaborative approach allows PSOs to effectively explore the solution space and converge towards optimal solutions.
- **Pattern Search (PS) Optimization:** PS optimization is a direct search method that does not require gradient information [44]. It explores the solution space by iteratively evaluating points in a predefined pattern around the current best solution. PS is particularly useful for optimization problems with non-smooth or discontinuous objective functions.
- **Grey Wolf Optimization (GWO):** GWO mimics the leadership hierarchy and hunting strategy of grey wolves in nature [45]. It uses alpha, beta, delta, and omega wolves to simulate the leadership hierarchy and encircle prey. This method effectively balances exploration and exploitation, making it suitable for a wide range of optimization problems.
- **Whale Optimization Algorithms (WOA):** The WOA is inspired by the bubble-net hunting strategy of humpback whales [46]. It models the optimization process through encircling prey, bubble-net attacking, and searching for prey behaviours. WOA is effective for solving complex optimization problems due to its ability to perform both local and global searches.

These methods are powerful tools for tackling problems with complex landscapes, however, their inherent randomness can lead to variability in the results, necessitating multiple runs to ensure reliability. In the family of stochastic problems, heuristics are problem-solving approaches that use practical methods to find satisfactory solutions within a reasonable timeframe.

3.4.3 Optimization Algorithms in Python

To solve optimization problems, researchers have employed various approaches over the years, including the use of software such as Homer Pro [47], the TRNOPT module for TRNSYS, or writing code directly in Matlab or Python environments. In this thesis, Python was chosen because it was compatible with the DT architecture of DigiBUILD, which included the development of these algorithms. Additionally, Python offers access to numerous open-source libraries with pre-implemented and refined optimization algorithms, thereby avoiding

unnecessary effort. However, the problem must be redefined each time, with all constraints and objective functions specified but choosing an existing library provides access to a vast array of optimization algorithms. During this PhD, three different libraries were primarily explored, each with its advantages and limitations.

The Pymoo library [48] was the first used in this thesis, specifically for the development of the optimization algorithm for CP Rio Vena, the DHN. At that time, one of the most commonly used libraries was SciPy, which is considered an alternative to Matlab for Python and allows for the modelling of complex optimization problems. However, using this library required the objective function to be defined as a mathematical or physical function, whereas in the DHN application (as detailed in Chapter 4), the objective function initially included an ANN and later a GBoost algorithm. Given this limitation, Pymoo was selected due to its user-friendly approach, flexibility in defining constraints and functions, and availability of advanced algorithms such as GA and NSGA-2. Using this library, it was only necessary to define the objective function and the constraints as different functions in Python, and the algorithm could manage all of them, considering only feasible solutions and finding the optimum value.

In cases of a highly constrained problem and a sparse solution matrix, such as the optimization problem of the Emotion microgrid, the approach using Pymoo encountered difficulties in finding feasible solutions. Consequently, PyGAD, another open-source library that primarily relies on GA optimization [49], was considered. This library offered a more accessible optimization algorithm, allowing for the customization of parameters related to generation successions, such as crossover and mutation functions, parent mating ratios, and elitism selection. Additionally, it was possible to provide initial solutions for the algorithm to start with, leading to a faster and more efficient solution achievement process.

```
# Create a new GA instance with the parameter on_generation
ga_instance = pygad.GA(num_generations=500 ,
                      fitness_func=fitness_func,
                      initial_population=init_pop, # Leave it as it is since there are many zeros we like it
                      num_parents_mating=int(len(init_pop)*0.9), # How many individuals are selected as parents
                      gene_type=int, # How the gene should be
                      random_seed=4,
                      parent_selection_type="tournament", # it was sss before
                      K_tournament=3,
                      keep_elitism= int(len(mhanz)/10),
                      crossover_type="uniform",
                      crossover_probability=0.1,
                      mutation_type=mutation_func,
                      mutation_percent_genes=80,
                      mutation_probability=0.9,
                      allow_duplicate_genes=True,
                      stop_criteria=["saturate_150"],
                      on_generation=on_generation)

# Run the genetic algorithm
ga_instance.run()
```

Fig. 13. Example of Optimization Algorithm Script using PyGAD

In November 2023, the latest version of MEALPY was released [50], described on its

Tancredi Testasecca

website as “the largest Python library in the world for most of the cutting-edge meta-heuristic algorithms” including random search optimization, continuous optimization, and nature-inspired algorithms. During the analysis of the current scientific literature, it was noted that the most recent applications of optimization for DTs of microgrids relied on nature-inspired algorithms such as bat optimization or GWO [51]. Observing that the GA of PyGAD was sometimes slow in finding the optimal solution, even with additional problem constraints, the problem was implemented in MEALPY. Specifically, the most novel algorithms (i.e., WOA, GWO, and a hybrid WOA-GWO) were tested, revealing that the hybrid algorithm could achieve an optimal solution, even without starting solutions, in less time than the PyGAD algorithm.

3.5 Graphic User Interface

The Graphic User Interface (GUI) is one of the most important components that enables interaction, monitoring, and management of digital replicas within DTs. A well-designed GUI enhances user experience through intuitive and insightful visualizations that facilitate better decision-making, thus allowing seamless control. This section describes the purpose of GUIs in DTs, their functionalities, and how they were implemented using Grafana [52], React [53], and LabVIEW [2] in the framework of this thesis.

GUIs in DTs are the medium through which users can interact with the DT of the physical asset. They are quite essential for use in visualization because they offer both visualization of real-time data and historical trends that facilitate state monitoring and performance in the physical system. Besides, within them, users may issue commands or configure parameters that would help manage the digital system and the physical system. Additionally, it would provide diagnostics, predictability of maintenance, and operations optimization with a set of varied analytic tools. Grafana is an open-source platform for effective visualization. It is very well suited for implementing GUIs for DTs, since it can process real-time data streams and provide comprehensive customizable dashboards for the continuous monitoring of a DT. Grafana was initially selected for the DT of CP-Rio Vena, the DHN provided by the Veolia Pilot in DigiBUILD.



Fig. 14. Veolia Digital Twin Dashboard in Grafana

Fig. 14 presents a typical screen of the Grafana dashboard for CP Rio Vena, displaying the most recent registered values (such as the maximum supply and return temperature) as well as historical data. When accessing a different page, it is also possible to run implemented Python scripts, enabling the execution of boiler optimization algorithms and then visualizing the results on another page. However, despite being very user-friendly and easy to develop, Grafana has limitations regarding user interaction, making it challenging to use for simulation purposes that may be required by a DT. In brief, to develop GUI using Grafana, it is needed the connection to data sources such as databases, IoT platforms, and streaming services (such as MQTT broker) providing real-time data from the physical system. After that, dashboards are designed to show relevant metrics and visualizations with controls, all of which are bespoke to the needs of the DT. Finally, it is even possible to set up alert rules and notifications allowing for enhanced monitoring of system health and performance.

To overcome the limitations of Grafana related to user interaction, React was selected for the final release of the Rio Vena DT. React is another excellent solution to implement advanced GUIs for DTs because of its flexibility and performance, considering that it is a library of JavaScript targeting the construction of dynamic and interactive user interfaces. Moreover, React enables the development of reusable components of the user interface, while displaying real-time insights of the real system. It also provides high interactivity since the complex operations and visualizations can be carried out smoothly by the users thanks to Java programming language. Before Rio Vena, React was used for Emotion DT, providing even in its first releases deeper user interaction.

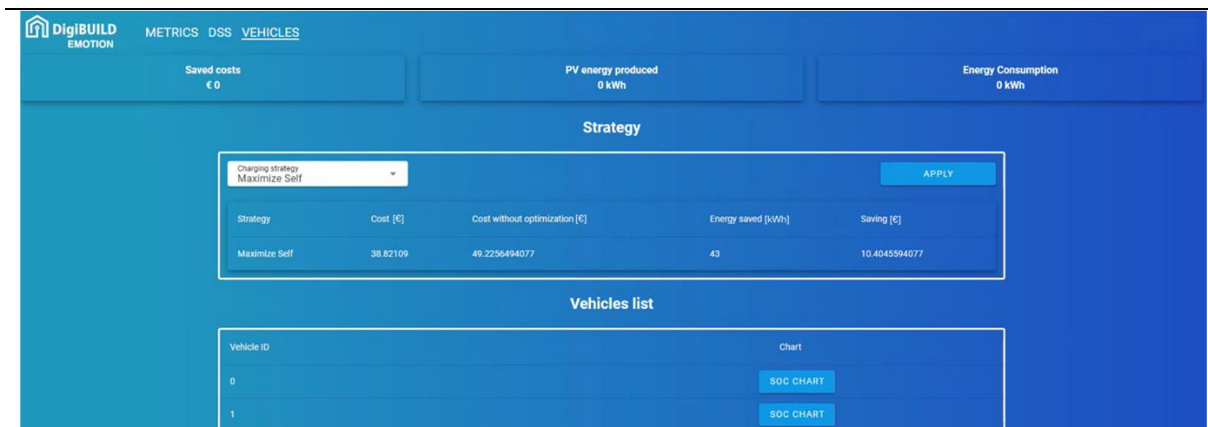


Fig. 15. A Section of the Emotion Digital Twin Dashboard in React

In the dashboard shown in Fig. 15, users can select different optimization strategies, each running distinct optimization processes to provide customized solutions. The DT incorporates two optimization strategies through the Middleware API, offering day-ahead and real-time recommendations for facility managers on which vehicle to charge, the optimal charging times, and the power modulation for each plug. Additionally, by clicking on each connected EV (at the bottom), the expected State of Charge (SoC) based on the optimization results will also be displayed. In the final release, another pop-up will allow users to input their preferences in terms of the vehicle needed, time of need, and distance to travel, updating the optimization process with these preferences. To implement a DT GUI in React, a React project must be set up, typically initialized with Create React App or any similar tool. Then, it follows with the development of components for all parts of the interface: data visualizations, control panels, and dashboards. After, state management libraries like Redux or Context API will be included to bind real-time data.

LabVIEW is a system design and development environment developed by National Instruments with a specific orientation to the graphical programming approach. It efficiently works with applications that need exact hardware integration and control. Its graphical language for programming, based on blocks, enables engineers and scientists to build easily complex systems. The possible integration with a wide array of hardware devices makes LabVIEW ideal for monitoring and controlling physical systems via real-time data acquisition and processing. A GUI for a DT in LabVIEW is implemented by developing blocks that model different parts of a DT while handling the whole process from data acquisition to processing, and finally visualization. Despite not being an open-source software, Labview allows the creation of user-friendly interfaces for real-time data and control through the front panel design tools. Next, there is the hardware connection and configuration of the required hardware parts that would allow proper communication between them and the LabVIEW environment. Finally, deploy the application to accurately model the physical system and provide all specified control and

monitoring capabilities.

In this framework, an example of GUI for a DT used for enhanced monitoring is represented in Fig. 16. This DT implemented algorithms for the calculation of visual comfort parameters such as Uniformity of Illuminance or Daylight factor. Moreover, through this GUI it was possible to monitor visual comfort variables and to visualize possible basic suggestions in case of discomfort for the user.

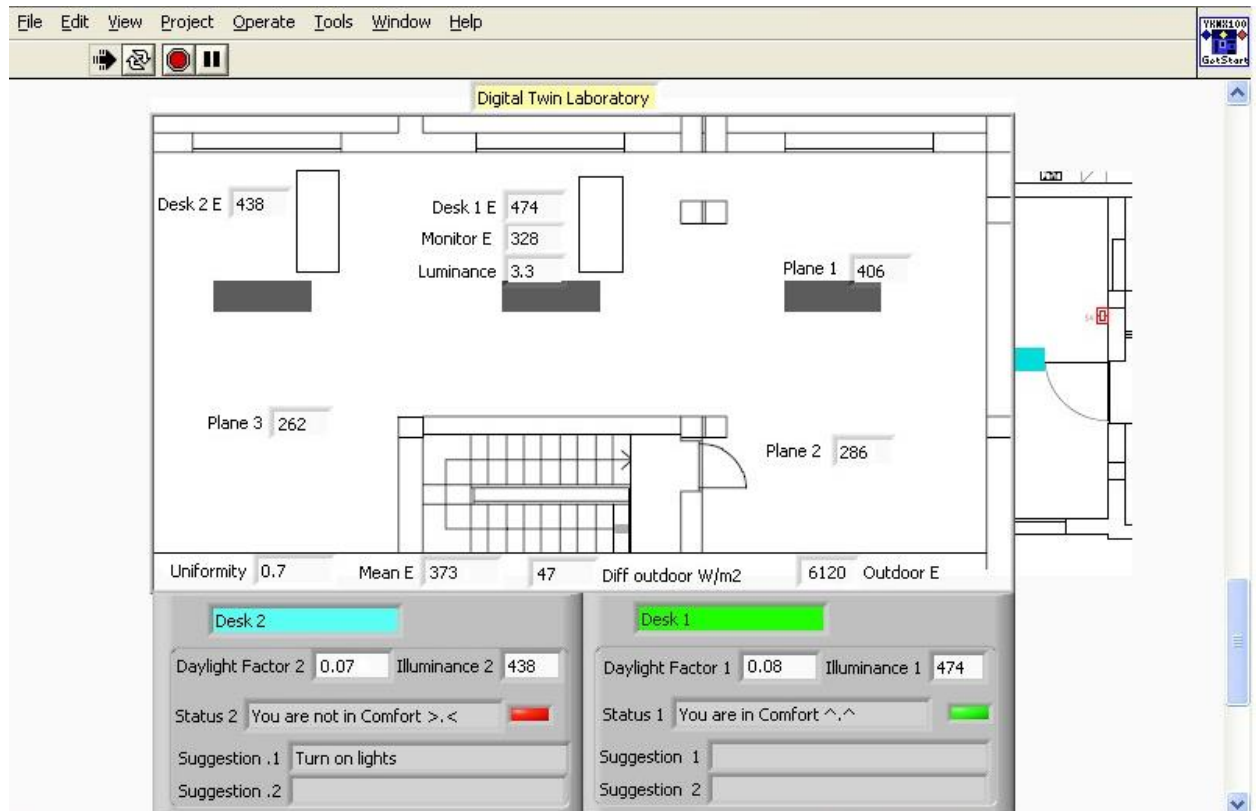


Fig. 16. Dashboard of DT of SolarLab for Visual Comfort Monitoring

References

- [1] M. Bonomolo, T. Testasecca, A. Buscemi, F. L. Alberto Munafò, and M. Beccali, “Digital Twin-Based Assessment Framework For Monitoring Visual Comfort,” in *2024 IEEE 22nd Mediterranean Electrotechnical Conference (MELECON)*, IEEE, Jun. 2024, pp. 1139–1145. doi: 10.1109/MELECON56669.2024.10608560.
- [2] R. Bitter, T. Mohiuddin, and M. Nawrocki, *LabVIEW: Advanced programming techniques*. Crc Press, 2006.
- [3] M. Bonomolo, G. Zizzo, M. Beccali, B. Di Pietra, A. Barberi, and S. Guarino, “Analisis on the feasibility to modulate lighting systems load: the case study of Palermo,” in *2022 IEEE International Conference on Environment and Electrical Engineering and 2022 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)*, IEEE, Jun. 2022, pp. 1–6. doi: 10.1109/EEEIC/ICPSEurope54979.2022.9854724.
- [4] DIAL GmbH, “DIALux.” Accessed: Aug. 08, 2024. [Online]. Available: <https://www.dialux.com/it-IT/dialux>
- [5] The Apache Software Foundation, “Apache Kafka.” Accessed: Aug. 08, 2024. [Online]. Available: <https://kafka.apache.org/>
- [6] “D4.2: Digital Twin Suite and cloud-based DigiBUILD data toolbox (Beta version),” 2024. [Online]. Available: <https://digibuild-project.eu/>
- [7] “DigiBUILD Project - High-Quality Data-Driven Services for a Digital Built Environment towards a Climate-Neutral Building Stock.” Accessed: Jun. 27, 2024. [Online]. Available: <https://digibuild-project.eu/>
- [8] T. Testasecca *et al.*, “Implementing Digital Twins for Enhanced Energy Management in Three Case Studies,” in *2024 IEEE International Workshop on Metrology for Living Environment (MetroLivEnv)*, IEEE, Jun. 2024, pp. 343–348. doi: 10.1109/MetroLivEnv60384.2024.10615496.
- [9] F. Pedregosa *et al.*, “Scikit-learn: Machine learning in Python,” *Journal of machine learning research*, vol. 12, no. Oct, pp. 2825–2830, 2011.
- [10] P. J. Rousseeuw and K. Van Driessen, “A Fast Algorithm for the Minimum Covariance Determinant Estimator,” *Technometrics*, vol. 41, no. 3, pp. 212–223, Aug. 1999, doi: 10.1080/00401706.1999.10485670.
- [11] Z. Ding and M. Fei, “An Anomaly Detection Approach Based on Isolation Forest Algorithm for Streaming Data using Sliding Window,” *IFAC Proceedings Volumes*, vol. 46, no. 20, pp. 12–17, 2013, doi: 10.3182/20130902-3-CN-3020.00044.
- [12] Vijay Kotu and Bala Deshpande, *Data Science - Concept and Practice*, 2nd ed. Elsevier, 2019.

doi: 10.1016/C2017-0-02113-4.

- [13] J. Brownlee, *Statistical Methods for Machine Learning: Discover how to Transform Data into Knowledge with Python*. Machine Learning Mastery, 2018. [Online]. Available: <https://books.google.it/books?id=386nDwAAQBAJ>
- [14] C. Multiphysics, “Introduction to COMSOL multiphysics®,” *COMSOL Multiphysics, Burlington, MA, accessed Feb*, vol. 9, p. 2018, 1998.
- [15] University of Wisconsin--Madison: Solar Energy Laboratory, “TRNSYS 17, a transient simulation program,” 1975, *Madison, Wis. : The Laboratory, 1975*. Accessed: Oct. 05, 2023. [Online]. Available: <https://www.trnsys.com/>
- [16] V. Palomba, M. Ferraro, A. Frazzica, S. Vasta, F. Sergi, and V. Antonucci, “Dynamic simulation of a multi-generation system, for electric and cooling energy provision, employing a SOFC cogenerator and an adsorption chiller,” *Energy Procedia*, vol. 143, pp. 416–423, Dec. 2017, doi: 10.1016/j.egypro.2017.12.705.
- [17] Brownlee Jason, *XGBoost with Python: Gradient Boosted Trees with XGBoost and scikit-learn*. 2018.
- [18] G. Louppe, “Understanding Random Forests: From Theory to Practice,” 2014. [Online]. Available: <http://arxiv.org/abs/1407.7502>
- [19] Douglas. Steinley, “K-means clustering: A half-century synthesis,” *British Journal of Mathematical and Statistical Psychology*, vol. 59, no. 1, pp. 1–34, May 2006, doi: 10.1348/000711005X48266.
- [20] The Math Works Inc., “MATLAB. Version 2020a,” 2020, *The Math Works, Inc.* Accessed: Oct. 14, 2022. [Online]. Available: <https://www.mathworks.com/>
- [21] Open Source Modelica Consortium, “OpenModelica.” Accessed: Oct. 20, 2022. [Online]. Available: <https://openmodelica.org/>
- [22] “Energy Plus,” 9.5.0. [Online]. Available: <https://energyplus.net/>
- [23] J. Danielewicz, B. Śniechowska, M. A. Sayegh, N. Fidorów, and H. Jouhara, “Three-dimensional numerical model of heat losses from district heating network pre-insulated pipes buried in the ground,” *Energy*, vol. 108, pp. 172–184, 2016, doi: 10.1016/j.energy.2015.07.012.
- [24] P. Catrini, T. Testasecca, A. Buscemi, and A. Piacentino, “Exergoeconomics as a Cost-Accounting Method in Thermal Grids with the Presence of Renewable Energy Producers,” *Sustainability (Switzerland)*, vol. 14, no. 7, Apr. 2022, doi: 10.3390/su14074004.
- [25] T. Testasecca, P. Catrini, M. Beccali, and A. Piacentino, “Dynamic simulation of a 4th Generation District Heating network with the presence of prosumers,” 2023.
- [26] Cengel Yunus and Cimbala John M., *Fluid Mechanics: Fundamentals and Applications*. 2014.
- [27] T. Lickleder, T. Hamacher, M. Kramer, and V. S. Perić, “Thermohydraulic model of Smart

-
- Thermal Grids with bidirectional power flow between prosumers,” *Energy*, vol. 230, 2021, doi: 10.1016/j.energy.2021.120825.
- [28] S.A. Klein *et al.*, *TRNSYS 17 - Mathematical Reference*. 2012. [Online]. Available: <http://sel.me.wisc.edu/trnsys>
- [29] Marco Cefarin, “Design of NH₃-H₂O absorption chiller for low grade waste heat recovery,” 2015. Accessed: Aug. 08, 2024. [Online]. Available: <https://hdl.handle.net/11390/1132442>
- [30] K. E. Herold, R. Radermacher, and S. A. Klein, *Absorption Chillers and Heat Pumps*. CRC Press, 2016. doi: 10.1201/b19625.
- [31] J. von Rhein, G. P. Henze, N. Long, and Y. Fu, “Development of a topology analysis tool for fifth-generation district heating and cooling networks,” *Energy Convers Manag*, vol. 196, no. March, pp. 705–716, 2019, doi: 10.1016/j.enconman.2019.05.066.
- [32] M. Tunzi, M. Ruyschaert, S. Svendsen, and K. M. Smith, “Double loop network for combined heating and cooling in low heat density areas,” *Energies (Basel)*, vol. 13, no. 22, 2020, doi: 10.3390/en13226091.
- [33] ASHRAE, *DISTRICT HEATING GUIDE - Complete Design Guide for District Heating Systems*. 2013. [Online]. Available: www.ashrae.org
- [34] G. Lennermo, P. Lauenburg, and L. Brand, “Decentralized heat supply in district heating systems - implications of varying differential pressure,” *14th International Symposium on DH and Cooling*, p. 6, 2014, [Online]. Available: http://www.svenskfjarrvarme.se/Global/Konferenser/DHC14/Proceedings/_5.2 Gunnar Lennermo DECENTRALISED HEAT SUPPLY IN DISTRICT HEATING SYSTEMS IMPLICATIONS OF VARYING DIFFERENTIAL PRESSURE.pdf
- [35] F. Hüsing, “Modelling of Inverter Heat Pumps in TRNSYS,” in *Proceedings of the ISES EuroSun 2020 Conference – 13th International Conference on Solar Energy for Buildings and Industry*, Freiburg, Germany: International Solar Energy Society, 2020, pp. 1–8. doi: 10.18086/eurosun.2020.04.08.
- [36] Johnson Controls-Hitachi Air Conditioning, “Centrifugal Chiller High Efficiency GFG-S/SIT Series GXG-S/SIT Series GSG-S/SIT Series.” Accessed: Aug. 08, 2024. [Online]. Available: <https://www.jci-hitachi.com/>
- [37] C. Ntakolia, A. Anagnostis, S. Moustakidis, and N. Karcianas, “Machine learning applied on the district heating and cooling sector: a review,” Feb. 01, 2022, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1007/s12667-020-00405-9.
- [38] M. You, Q. Wang, H. Sun, I. Castro, and J. Jiang, “Digital twins based day-ahead integrated energy system scheduling under load and renewable energy uncertainties,” *Appl Energy*, vol. 305, Jan. 2022, doi: 10.1016/j.apenergy.2021.117899.
-

-
- [39] M. J. Kochenderfer and T. A. Wheeler, "Algorithms for Optimization," 2019. [Online]. Available: <http://mitpress.mit.edu>.
 - [40] S. Jalili, "Introduction to Stochastic Optimisation," 2022, pp. 3–16. doi: 10.1007/978-981-19-4633-2_1.
 - [41] T. Testasecca, S. Stamatopoulos, M. Lazzaro, and E. Sarvas, "Recent advances on data-driven services for smart energy systems optimization and pro-active management," in *2023 IEEE International Workshop on Metrology for Living Environment (MetroLivEnv)*, 2023. doi: 10.1109/MetroLivEnv56897.2023.10164056.
 - [42] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, Apr. 2002, doi: 10.1109/4235.996017.
 - [43] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, IEEE, pp. 1942–1948. doi: 10.1109/ICNN.1995.488968.
 - [44] R. Hooke and T. A. Jeeves, "'Direct Search' Solution of Numerical and Statistical Problems," *Journal of the ACM*, vol. 8, no. 2, pp. 212–229, Apr. 1961, doi: 10.1145/321062.321069.
 - [45] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, Mar. 2014, doi: 10.1016/j.advengsoft.2013.12.007.
 - [46] S. Mirjalili and A. Lewis, "The Whale Optimization Algorithm," *Advances in Engineering Software*, vol. 95, pp. 51–67, May 2016, doi: 10.1016/j.advengsoft.2016.01.008.
 - [47] HOMER Software, "HOMER Pro." Accessed: Aug. 09, 2024. [Online]. Available: <https://homerenergy.com/products/pro/index.html>
 - [48] J. Blank and K. Deb, "Pymoo: Multi-Objective Optimization in Python," *IEEE Access*, vol. 8, pp. 89497–89509, 2020, doi: 10.1109/ACCESS.2020.2990567.
 - [49] A. F. Gad, "PyGAD: an intuitive genetic algorithm Python library," *Multimed Tools Appl*, 2023, doi: 10.1007/s11042-023-17167-y.
 - [50] N. Van Thieu and S. Mirjalili, "MEALPY: An open-source library for latest meta-heuristic algorithms in Python," *Journal of Systems Architecture*, vol. 139, Jun. 2023, doi: 10.1016/j.sysarc.2023.102871.
 - [51] H. Yang, S. Zhang, J. Zeng, S. Tang, and S. Xiong, "Future of sustainable renewable-based energy systems in smart city industry: Interruptible load scheduling perspective," *Solar Energy*, vol. 263, Oct. 2023, doi: 10.1016/j.solener.2023.111866.
 - [52] Grafana Labs, "Grafana: The open observability platform." Accessed: Aug. 09, 2024. [Online]. Available: <https://grafana.com/>
 - [53] Meta Open Source, "REACT - The library for web and native user interfaces." Accessed: Aug.

09, 2024. [Online]. Available: <https://react.dev/>

Chapter 4

Digital Twins and Models of District Heating Networks

Aligning with the objectives of the European Green Deal [1], it is necessary to establish a harmonious relationship between building efficiency, the incorporation of RESs, and the enhancement of energy production systems. The "Heating" study of the IEA [2] emphasises the urgent need to expand the proportion of ultra-low-carbon technology in building heating systems. District Heating and Cooling (DHC) networks are crucial in mitigating building emissions among the many solutions being considered. Since 2010, the number of new connections to DHNs has been growing at an annual rate of 3.5%, with the aim of reducing emissions [2]. China, accounting for almost 25% of global heat demand, has the highest rate of growth in DHC capacity worldwide. In European nations like Finland, Denmark, Sweden, and the Baltics, DHC networks fulfil more than 8% of the total heat demand [3]. Technology in Italy is concentrated mostly in the northern half of the peninsula. However, its coverage has increased fourfold from 2000 to 2019, while energy consumption has tripled [4].

In this framework, fossil fuels continue to be the primary energy source, while RES (primarily biomass) accounted for less than 8% of the total energy used [3]. When considering the sole Europe, Iceland has successfully reached a renewable energy usage of almost 100% thanks to its exceptional geothermal resources. Meanwhile, Denmark is at the forefront of incorporating solar thermal energy into District Heating (DH), with over three-quarters of the world's thermal installed capacity by the end of 2018. In 2019, 16% of the energy supplied to DH networks in Italy was derived from RES such as biomass and geothermal energy [4]. Countries including Denmark, China, Sweden, Saudi Arabia, and Austria have installed DHNs with extensive solar collector plants, resulting in a global installed capacity of 1.6 GWth [5].

4.1 Overview of District Heating Networks

From a technological point of view, there has been a noticeable decline in the operating temperature and pressure of DHNs throughout the years. A reduction in network temperature

not only minimises heat dissipation but also enables the utilisation of RES such as solar, geothermal, and waste heat from industrial operations, data centres, and supermarkets. In the literature, the process of this development has been elucidated by proposing the notion of "DHN generation" [6]. The majority of currently installed DHNs are of the 3rd Generation. These districts operate using pressurized water and temperatures ranging from 80 °C to 100 °C. Typically, these systems are supplied by large-scale centralised CHP that emit 150-300 grams of carbon dioxide (CO₂) per thermal kilowatt-hour [2]. Currently, third-generation DHNs with an operating temperature below 100 °C are the most spread while the research is focusing on fourth and fifth-generation DHNs. Specifically, fourth-generation networks function at peak temperatures ranging from 60 °C to 70 °C. Generating heat at this temperature will enable the incorporation of RES and the CHPs will produce heat with greater efficiency [6]. Sorknaes et al. [7] showed that, in a 4th generation network, there is a drop in grid losses and overall yearly energy system costs, while the efficiency of the heat-producing systems improves compared to the 3rd generation base case.

Seasonal storage systems, like aquifer Thermal Energy Storage (TES), are commonly utilised to enhance the flexibility of the overall system. These systems store heat throughout the summer months for later usage in the winter, and vice versa [8]. Within the context of fourth-generation networks the concept of the "prosumer" was also presented [9]. A prosumer is an individual who not only consumes but also produces energy in a decentralised manner, utilising RESs to create heat on-site. For example, a residence that has solar thermal panels can send excess heat to the grid during periods of low self-consumption, which can then be utilised by an adjacent user. In situations when there is a high level of RES being utilised, it is feasible to utilise the excess electricity from the grid to power "Power to Heat" technologies such heat pumps (HPs). These heat pumps may then deliver heat to the grid, which can be used by other consumers. In this situation, there would be a merging of the electricity and heat grids, where the HPs would help to decrease the disparity between electricity production and demand leading to advantages for managing the electricity grid itself [10]. In the meantime, 5th generation networks would be complementary to the fourth ones [6]. Typically, a single-loop or double loop network layout is used, where the water temperature ranges from 20 °C to 30 °C. This temperature range ensures minimal heat loss and allows for the use of any available heat source in the system. However, it is necessary to install booster heat pumps at each user's location to achieve the desired water temperature. This technology allows the simultaneous generation of cooling and Domestic Hot Water (DHW) without the need for a 4-pipe system to transport hot and cold water simultaneously.

4.2 Digital Twin of Rio Vena Third Generation Network

Optimisation services play a crucial role in enhancing the configuration and efficacy of energy infrastructures such as DHNs. Upcoming DHC systems will engage with various heat or cool sources, including renewable energy generators, necessitating novel approaches for estimating heat unit costs. To fulfil heating demands while minimizing expenses and Greenhouse Gas (GHG) emissions, optimal management is essential for determining the most effective combination of heat sources (e.g., power to heat, solar collectors, and boilers). VEOLIA, as an Energy Service Company, is one of the global leaders in optimised resource management and aims to achieve a sustainable development of cities and industries; in DigiBUILD framework [11], a DT was developed for CP Rio Vena, a third-generation DHN located in Spain. Therefore, the developed DT is a tool for the DHN manager which optimizes the operation of heat generators at the main plant so that both emissions and costs are reduced. Several novel services are proposed in this methodology being implemented in the DT of CP Rio Vena based on ML and optimization algorithms. Once predictions of the future heat demand are fed via an API developed by another project partner, the optimizer and models developed during this PhD research will recommend the optimal setpoint temperatures for the boilers.

Data-driven models of boiler efficiency were developed, one for winter operation and one for summer operation, using ML algorithms (i.e. ANNs and GBoost). Once the model development was done, the optimization problem was formulated and implemented in Python by the use of the Pymoo library. The goal of the optimization was the minimization of gas consumption, with the hourly boiler setpoint temperature being the decision variable. The resulting models were then exposed as an API that was to be integrated into the Middleware API as imposed by the DT architecture in the DigiBUILD project.

The final step was to build the dashboard of the DT through which historical and predicted data are accessed to be used as a monitoring tool by the manager. The second dashboard page can use the optimizer API to display the optimal setpoint temperatures for the control heat generation. Tests on historical data show that the developed tool can save up to 15% on gas consumption depending on the day. This intention also aligns with the goals of the IEA [12] and EU [13] on digitalising energy systems to increase their efficiency.

4.2.1 Methodology

The methodology followed for this DT is in correspondence with the steps outlined in Chapter 3 of this thesis. Because ML algorithms were used to model the boilers in the network, it was necessary to make a deep analysis of the available data. Subsequently, ML models were studied and developed based on the collected data, with periodic updates incorporating new

data. Finally, the optimizer was developed in parallel to the Dashboard of the DT. The core of Veolia DT is represented in Fig. 17, showing the service s3.2.2 developed during DigiBUILD.

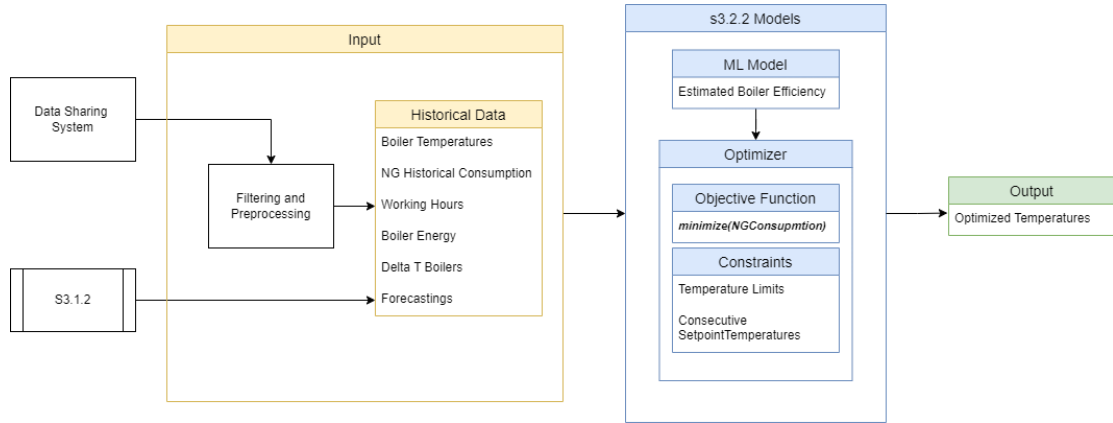


Fig. 17. I/O Scheme for Service s3.2.2 for CP – Rio Vena (VEOLIA)

Throughout this thesis and the DigiBUILD project, algorithms and models evolved from alpha to beta versions that reached the final release recently in Deliverable D3.3. For clarity purposes, this section will focus on the most updated methodology that led to the final release of the DT.

4.2.1.1 Data-Processing and ML models

The first step involved injecting the raw data into a data-processing script to create a DataFrame. This was cleaned by removing unnecessary columns and then resampling so that each row represented a specific time interval. As the data were not collected at the same frequency for all variables, the resampling process was required.

After that, new variables were computed to extend the available parameters list. Most of the variables were cumulative quantities, which had to be converted into dynamic ones at 30-minute intervals (the lowest frequency in this dataset). The gas consumption V_{NG} , expressed in Sm^3 was then used to calculate the thermal power associated in kW using equation (23):

$$P_{NG} = \frac{V_{NG} \cdot LHV_{NG}}{1800} \quad (23)$$

Where LHV_{NG} is the design Lower Heat Value (LHV) provided by Pilot leader (39,748 kJ/Sm^3) and 1800 is related to V_{NG} measurements collected with a frequency of 30 minutes (1800 s).

Once P_{NG} was calculated, boiler efficiency was computed as the ratio between thermal power on the demand side and the thermal power of NG fed into boilers. The script then filtered the dataset to include only values of efficiency not exceeding 102% (value detailed in the boiler datasheets). For training, the input data were scaled using the StandardScaler from the Scikit-Learn Python library [14].

The selected models for estimating boiler efficiency were two of the most widely used ML algorithms in energy applications: ANNs and Gboost. In the alpha version, only ANN was implemented, but before the final release, various models were compared, with GBoost demonstrating superior performance.

Focusing on winter functioning, the ANN model was structured as an MLP regressor with five hidden layers, consisting of 20, 100, 300, 100, and 20 neurons, respectively. The learning rate was set to 'adaptive,' and the solver used was 'adam'. For the GBoost model, optimal accuracy was achieved with 200 estimators, a learning rate of 0.1, a minimum of 2 samples at the leaf nodes, and a maximum tree depth of 4. In both cases, after filtering, more than 8700 data entries were available. Given the high availability of data, 10% of the data was reserved for testing, allowing the algorithms to be trained on the remaining 90%.

A similar approach was conducted for operation during summer, when only one out of three boilers is functioning due to lower thermal demand. In this case, the ANN regressor consisted of five layers consisting of 15, 130, 230, 130 and 10 neurons. The GBoost model included 120 trees, with a learning rate of 0.1, a maximum of tree depths equal to 5 and a minimum of 4 samples for each leaf node.

4.2.1.2 Optimization Problem and Algorithm

The optimization of heat generation in the DHN of Rio Vena was a primary requirement set by the pilot leader for the development of the DT. Consequently, all models and even the dashboard were specifically tailored to achieve this objective. Through various discussions with the pilot, it was determined that remote modulation of the gas entering the boiler was not feasible. The only controllable parameter was the boilers' setpoint temperature, which refers to the temperature of the water exiting the boiler post-heating. Therefore, this temperature was selected as the objective variable for optimization. Specifically, a sequence of 48 consecutive setpoint temperatures was chosen as variables, as it was deemed technically difficult to change the setpoint more frequently than every 30 minutes. This limitation was related to the fact that the process could not be automated, thus requiring human intervention. Furthermore, the predictions for future thermal energy consumption provided by the forecasting service were given at 30-minute intervals, consistent with the sampling frequency of the registered data.

Regarding the constraints of the optimization problem, the upper and lower bounds for the variables were determined based on the maximum and minimum values recorded in the historical data available. This approach seemed to be appropriate ensuring realistic boundary conditions for the optimization. However, due to thermal inertia of thermal exchanges and processes, accentuated by the 2000 l thermal tank separating the plant loop and district loop,

the difference between two consecutive temperatures in 30 minutes must be limited. Therefore, the maximum allowable difference was calculated from the historical data, and this value was set as a limit in the constraint that is included directly as a penalty in the objective function, as highlighted in equation 24.

$$\sum_{i=1}^{48} (T_{set,i} - T_{set,i-1}) \leq \max(\Delta T_{\text{historical}}) \quad (24)$$

in this equation:

- $T_{set,i}$ represents the boiler setpoint temperature at time step i . This variable denotes the desired temperature target that the boiler should reach during the specific time interval within the optimization period.
- $T_{set,i-1}$ denotes the boiler setpoint temperature at the preceding time step $i-1$.
- $\max(\Delta T_{\text{historical}})$ signifies the maximum allowable temperature difference between consecutive setpoint temperatures, derived from historical data. This value sets the limit for temperature fluctuations over a 30-minute period, accounting for the thermal inertia present in the system.

Finally, the objective function was defined to minimize the gas consumption, in terms of thermal energy, including the constraint penalty pen , on the next 24 hours of operation of the district. It considers as input the prediction of future thermal energy consumption by the entire district in a frequency of 30 minutes for the next n timestep, and the ML model for the efficiency of boilers $f_{\eta,i}$ as expressed in equation 25:

$$\min \left(\sum_{i=1}^n \frac{E_{d,i}}{f_{\eta,i}(T_{set,i}, E_{d,i})} + pen \right) \quad (25)$$

It was also decided not to extend the optimization period to cover all 48 successive timesteps, as this would be computationally expensive, particularly for timesteps where no heat was demanded by the district, and thus no heat generation was required from the boilers. By limiting the number of objective variables, the computational effort was significantly reduced.

The objective function and the associated constraints were implemented in Python using the Pymoo library [15]. Fig. 18 illustrates the code used to define the objective function and constraints, including the "errors" variable. Specifically, the vector of objective variables ('x' variable) is first reshaped into a matrix (when multiple types of decision variables could be involved). Then, the variation of consecutive temperatures, 'deltatemp', is calculated, and the number of instances where this exceeds the set constraint is determined. Next, the 'x_matrix' variable is integrated with $E_{d,i}$ values, representing 'INSTANTANEOUS ENERGY (15 minutes)'. After scaling, the efficiency vector is calculated using the previously developed ML

model. Finally, the objective function is computed in the penultimate line of the code, and penalties are added, multiplied by 10^5 to ensure the optimization algorithm, a NSGA-II type, properly accounts for feasibility when evaluating potential solutions.

```
def f(x, n=n_no, df_o=df_o_no):
    # Reshape the decision variables into a matrix with n rows and X columns
    x_matrix = x.reshape((n, X))

    # Calculate the differences between the rows of x_matrix
    deltatemps = np.diff(x_matrix, axis=0)

    # Count the number of times the difference is lower/higher than constraints
    errors = np.sum((deltatemps < delta_neg) | (deltatemps > delta_pos))

    x_matrix = np.hstack((df_o['INSTANTANEOUS ENERGY (15 minutes)'].values.reshape((n, 1)),
                          x_matrix))

    x_matrix = pd.DataFrame(x_matrix,
                           columns=['INSTANTANEOUS ENERGY (15 minutes)',
                                    'BOILER TEMP (15 minutes)'])

    # Apply the scaler transformation to the decision variables matrix
    x_matrix_scaled = scaler.transform(x_matrix)

    # Calculate the sum of the model predictions for all timesteps
    eta = regressor.predict(x_matrix_scaled)

    # Check if 'INSTANTANEOUS ENERGY (15 minutes)' is zero
    energia_instantanea = df_o['INSTANTANEOUS ENERGY (15 minutes)'].values

    f = energia_instantanea / eta

    return (np.sum(f) + 10000 * errors)
```

Fig. 18. CP Rio Vena Optimization Function

4.2.2 Case Study

The DHN of CP Río Vena is located in Burgos, the provincial capital of the autonomous community of Castile and León in Spain. The city's climate is characterized by cold winters, with air temperatures frequently falling below 1°C , and frequent occurrences of snowfall. Additionally, due to its geographic location (42.34, -3.69), in December Burgos experiences reduced daylight hours, up to 9 hours per day, which limits the influence of solar radiation. More detailed information about the climate in Burgos is shown in Fig. 19, which presents average values for temperature and humidity as reported in [16].

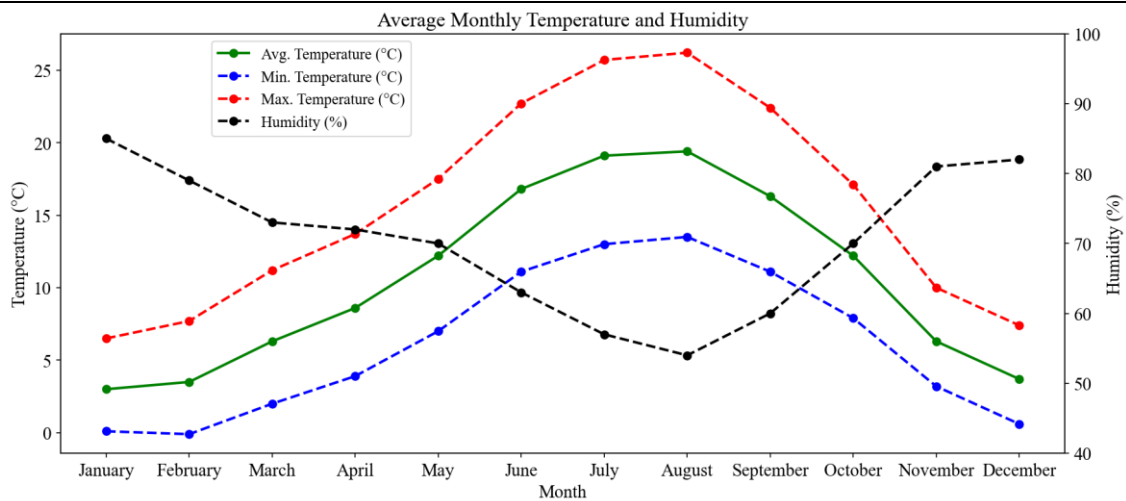


Fig. 19. Average and Extreme Values of Temperature and Humidity

At Río Vena location (Fig. 20), the DHN has 23 distribution substations serving SH and DHW for 706 households, supplied by three natural gas boilers with a total capacity of 13 MW. Thermal energy is generated, in fact, by three boilers: two 4.5 MW units that work alternatively according to schedule and thermal request, while a 4 MW unit works only for the DHW production during the summer period. Except for the coldest periods, heating energy for both SH and DHW is supplied according to specific schedules to optimize district management and reduce heat losses. This is facilitated by storage tanks of varying volumes in each building, adjusted according to individual consumption patterns.



Fig. 20. Sky View of Río Vena District Heating Network Cluster of Buildings [17]

Focusing on sensor and data availability, temperature sensors monitor all heat production flows including supply temperature from boilers. Temperature sensors on the district side monitor supply and return temperatures; on the user side, the DHW tank temperature and

setpoint temperature are monitored. In addition, valve actuation and alarms are captured as control parameters in real time. It measures and records the flow rate between the exchanger-tank and the district, while only the activation of the pump on the boiler side. These datasets, except for the communication between the distinct devices where the time represented is the timestamp in which the device has sent that communication, have a minute and timestamp frequency.

At the beginning of this project, more than two years of historical data for the Río Vena DHN had been provided by VEOLIA. Data included cumulative values of thermal energy production, gas consumption, and number of hours that the boiler was running. New data collection began in April 2023 in order to have additional variables potentially required for ML model and optimization algorithm development. Testing of the ML models was allowed with more recent data, even though the sample size was reduced compared to the first two years of data collection. This new dataset includes variables regarding the supply and return temperatures of the boilers and district, flow rates of the district, and thermal power exchanged on the district side. All these variables are acquired with a 15-minute period. Due to limitations on storage capacity and processing power, it was impossible to store the whole dataset being monitored in the existing SCADA system, according to the pilot leader.

Located in the boiler room are three natural gas-fired boilers, each serving a different purpose in the heating process. Boiler 1 serves as the main unit and provides the majority of required heat. Boiler 2 is used as an auxiliary unit when peak times need supporting. Boiler No. 3 is a different model than previous ones and, in addition to the summer period, it would be kept as a reserve for maintenance or emergencies. The system is combined with an inertial water tank of 2000 litres, which ensures effective separation of the hydraulic circuit of the boiler and the distribution ring. During the winter, supply and return temperatures were approximately 80°C and 60°C, respectively, ensuring continuous coverage of heating and DHW demands.

4.2.3 Results and Findings

4.2.3.1 Data Analysys

A thorough data analysis was carried out before the models were developed. For a correct development, only data collected from week 17 of 2023 was relevant. The first analysis was conducted to determine the temperature behaviour of the district, as shown in Fig. 21. For clarity, the temperatures were resampled to daily averages. It was observed that, during winter, boiler supply temperatures were higher compared to the summer maximum of 65°C, though they never exceeded 85°C. The inlet temperature of the boiler, referred to as the Boiler Return

Temperature, is typically 15-20°C lower than the supply temperature. This differential allows for more efficient heat generation during summer, as the return temperature drops below 57°C, enabling the capture of latent heat. The district's supply temperatures range from approximately 70°C in winter to 65°C in summer, with an average daily temperature difference of 10°C between the supply and return sides. Lastly, it should be noted that gaps in the data were present due to sensor malfunctions or issues related to data acquisition and storage processes.

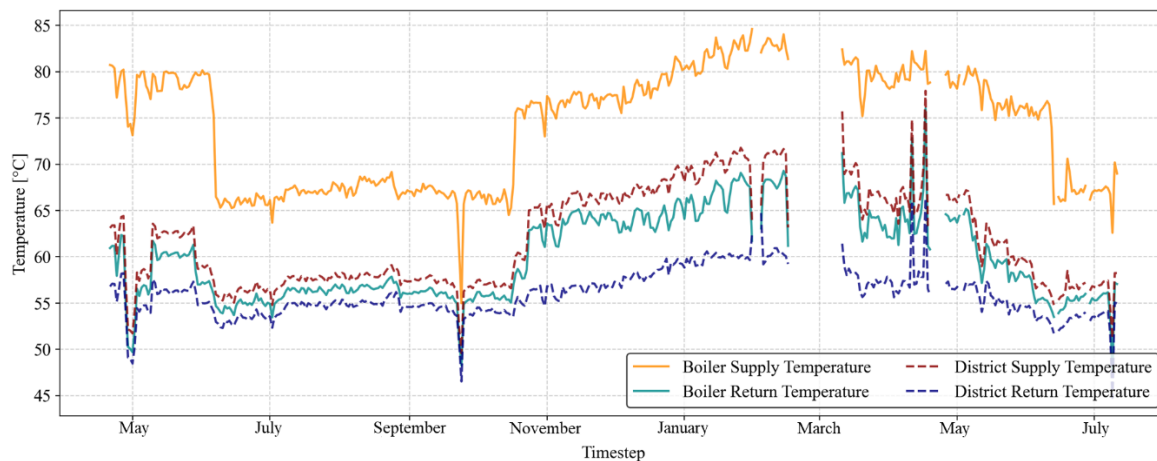


Fig. 21. Daily behaviour of average temperature in Rio Vena network

Nevertheless, it is interesting to evaluate the hourly behaviour since in daily resampling details on the management could be lost. Starting for the summer solstice, in Fig. 22, it can be noticed that the heat is produced in a discontinuous way since the temperature alternates times where a peak is shown followed by a slope. Focusing on this peak, at 6:00 and at 21:00, the supply temperatures and the boiler return temperature are all around 73°C, which means thermal energy demand and production are aligned, because the 2000 l tank temperature stays nearly constant (as the boiler return temperature). During the rest of the day, a variation of the temperature between the supply and return temperature of both the district and boiler sides can be noticed. In Winter, the peak of the thermal request is 2.4 times more relevant than in summer, overcoming more than 4.5 MW. However, in the morning, two different peaks of energy request have been detected for which district supply temperature is respectively 79.8°C and 70.1 °C. After noon, the district operates continuously and the boiler supply temperature is equal to the district supply, revealing an efficient thermal exchange.

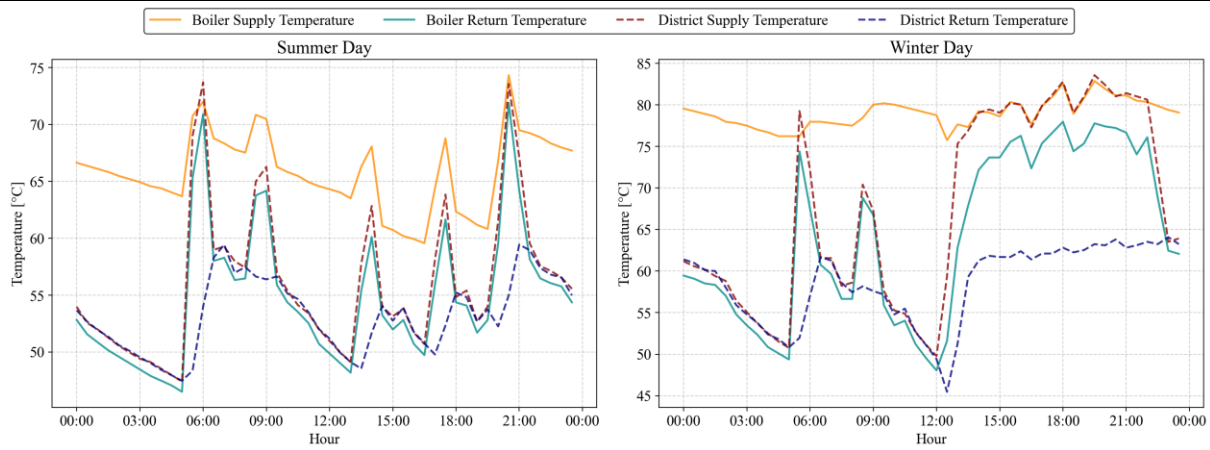


Fig. 22. Hourly behaviour in summer and winter solstice of temperatures and power

Following the analysis of the available data, Equation 23 was applied to the entire dataset to calculate boiler efficiency in heat generation and examine its correlation with other available variables. However, prior to this step, it was important to assess whether boiler inertia significantly affected the discrepancy between the thermal energy produced and the thermal energy demanded on the district side. To address this, the TRNSYS [18] model represented in Fig. 23 was developed. Historical real-world data were input into the model using a data reader, including boiler-side and district-side supply and return temperatures, with the flow rate applied only on the district side. The 2000-liter tank at Rio Vena, which allows thermal exchange between the boilers and the district loop, was modelled using Type 158 – Vertical Tank. Additionally, Type 22 – Iterative Feedback Controller was implemented to regulate the pump (Type 3b) flow rate, simulating the flow rate in the boiler room and ensuring that the controlled variable, the return temperature on the boiler-side tank, remained as close as possible to the measured data.

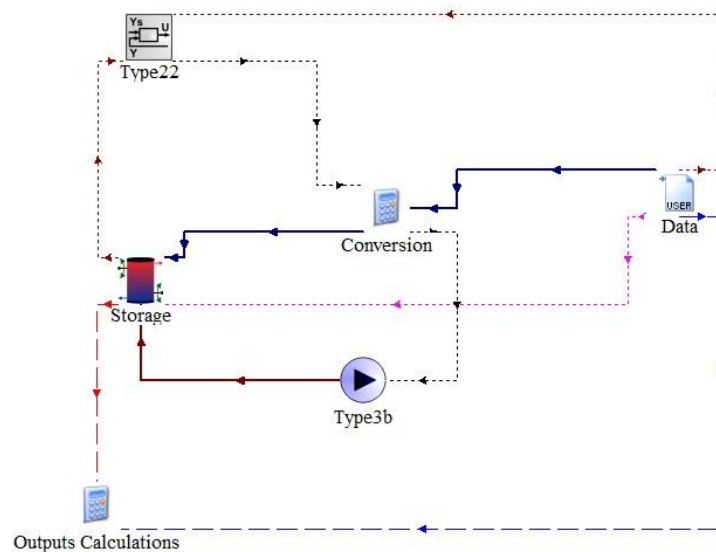


Fig. 23. TRNSYS Tank Model for Rio Vena.

The results of the model indicated that it was possible to achieve the correct flowrate while keeping the real and simulated return temperatures closely aligned. Specifically, the MAE between the real and simulated temperatures was 0.89, with a MAPE of 1.67%. When analysing the thermal energy exchanged in the tank during operational hours and comparing the heat removed by the district with the heat supplied by the boiler, a MAE of 0.03 and an R^2 score greater than 0.99 were observed. These findings highlight the feasibility of utilizing thermal energy data from the district side to calculate boiler efficiency.

As previously mentioned, the Rio Vena system operates under two different boiler modes: one for summer, where only Boiler 3 is active, and one for the rest of the year, when Boilers 1 and 2 are used in combination. Therefore, it was necessary to divide the input dataset into separate datasets for each operational mode, as the system efficiency and corresponding model differ between these modes. The dataset, which contained over 21,000 entries, was split accordingly, and finally efficiency was calculated. However, as previously said, data entries where efficiency exceeded the threshold of 102% were excluded. These accounted for 37% of the entries during Boiler 3 operation and 36% during the operation of Boilers 1 and 2. These discrepancies were consistent with data collection issues previously reported by the pilot leader, and an internal investigation is underway to address these problems.

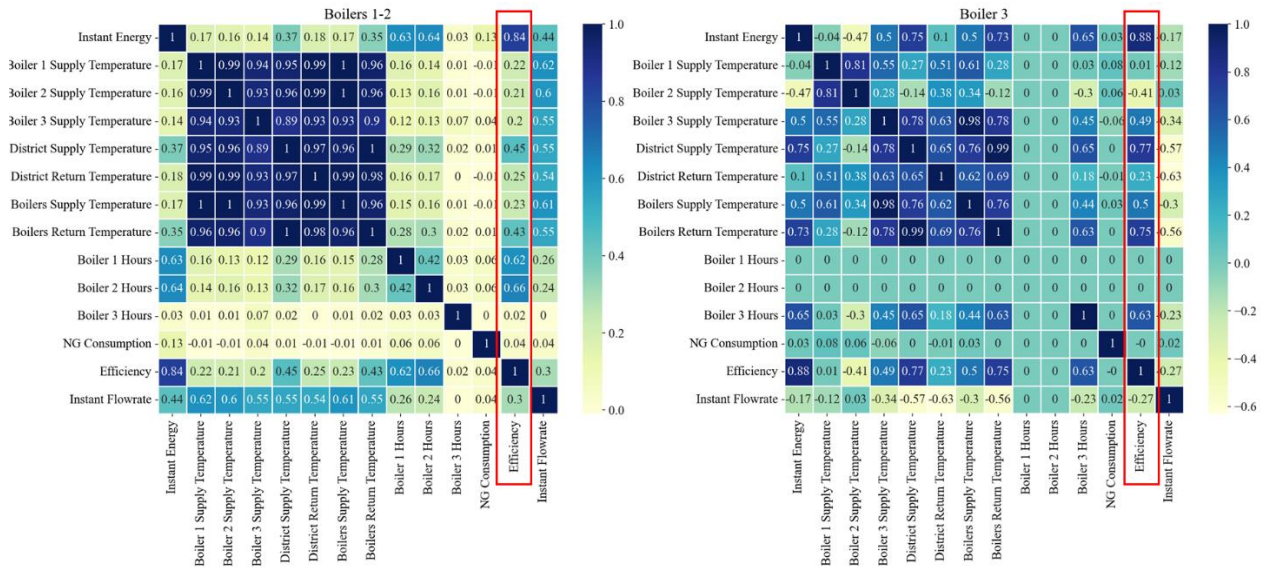


Fig. 24. Correlation Heatmap for Boilers 1-2 (left) and Boiler 3 (right) operation.

To determine the appropriate input features for the ML models, the correlations between variables were analysed, as shown in Fig. 24. Since the primary objective of the model is to identify the most suitable regressor for estimating boiler efficiency and optimizing district operation, the selected variables must be technically adjustable or predictable. For instance, the highest correlation was observed when using Instantaneous Energy (the thermal energy exchanged district side) to estimate efficiency. This feature was also aligned with another task

in the DigiBUILD project, which aimed to predict district thermal consumption, making it the first input feature selected. However, the second-highest correlation values (varying by operational mode) were found with Boiler Hours or Boiler Return Temperature. As these variables could not be directly controlled by the DHN operator and could not be physically calculated due to the lack of data on boiler-side flow rate, they were not used as inputs. Therefore, the second feature chosen for the regressor was Boiler Supply Temperature, as it is the only variable the DHN manager can control during normal operation and could also be used in optimization algorithms.

4.2.3.2 Data-driven modelling validation

In the context of ML modelling, the estimation of boiler efficiency is performed by using GBoost and ANNs according to the methodology described before, considering Instantaneous Energy, the thermal energy demanded by the district, and Boiler Supply Temperature as input features. Comparing the two models presented, different metrics were calculated for both the test set and the entire dataset that is shown in Table 8. As can be noticed, both models were quite accurate, with a coefficient of determination above 0.92 and a MAE below 0.03. In this application, GBoost outperforms the NN model in every metric, most notably in the full R^2 score of the dataset, showing up to 1.03 times higher. It therefore means that for the DT and optimizer developed, only GBoost was applied to predict boiler efficiency for both Boiler 1-2 Operation and Boiler 3 Operation.

Table 8. Metrics Comparison Between ANN and Gboost Models

Metric	Boilers 1-2 Model				Boiler 3 Model			
	Gboost		ANN		Gboost		ANN	
	Test	Full	Test	Full	Test	Full	Test	Full
MAE	0.021	0.015	0.026	0.025	0.014	0.009	0.021	0.020
MSE	0.004	0.002	0.005	0.004	0.003	0.001	0.004	0.003
RMSE	0.063	0.048	0.068	0.066	0.054	0.035	0.061	0.055
R^2	0.937	0.966	0.927	0.933	0.942	0.977	0.926	0.944

For the sake of completeness, graphical comparisons of the predictions made by ANNs and GBoost with real data for Boilers 1-2 and Boiler 3 over different time periods are presented in Fig. 25 and Fig. 26, respectively. Both models were able to estimate boiler efficiency accurately, even at higher and lower values, with the exception of a few peaks where efficiency dropped below 0.2, as shown in the right chart of Fig. 25 for Boiler 1-2 operation.

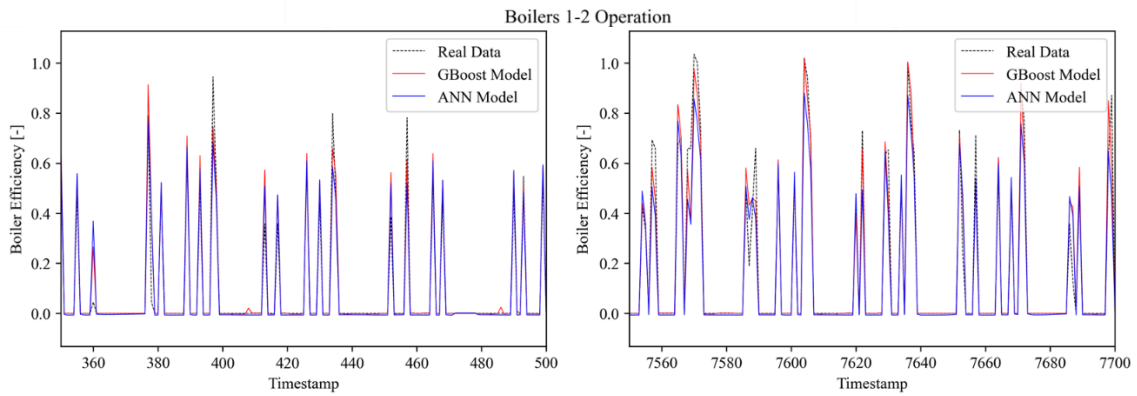


Fig. 25. Boiler Efficiency ML Model Behaviour for Boilers 1-2 Operation

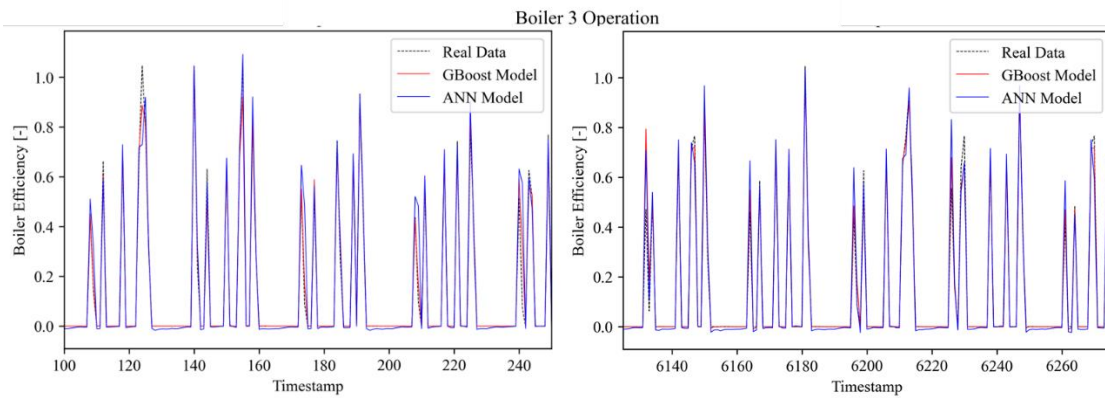


Fig. 26. Boiler Efficiency ML Model Behaviour for Boiler 3 Operation

4.2.3.3 Optimization results

The optimisation process in the DT aims to determine a setpoint temperature strategy that maximizes efficiency over an entire period for the next day. To test the algorithm, several simulations were conducted for historical data to assess the benefit of temperature optimization. In these simulations, once historical data were filtered excluding instances where the efficiency was more than 102% and periods with no demand, the model developed had the potential to save more than 32 MWh of natural gas during winters and over 128 MWh during summers. This translates into energy savings of 6.7% and 31.8%, respectively, for the data available in the year 2023. The efficiency of the algorithm was assessed through a comparison of optimized gas consumption with the gas consumption derived from the application of the model on historical data with unoptimized temperatures. As shown in Fig. 27, through optimal temperature control, optimized gas consumption was always lower than historical consumption, with gas savings ranging from 3% to 20%, depending on the day.

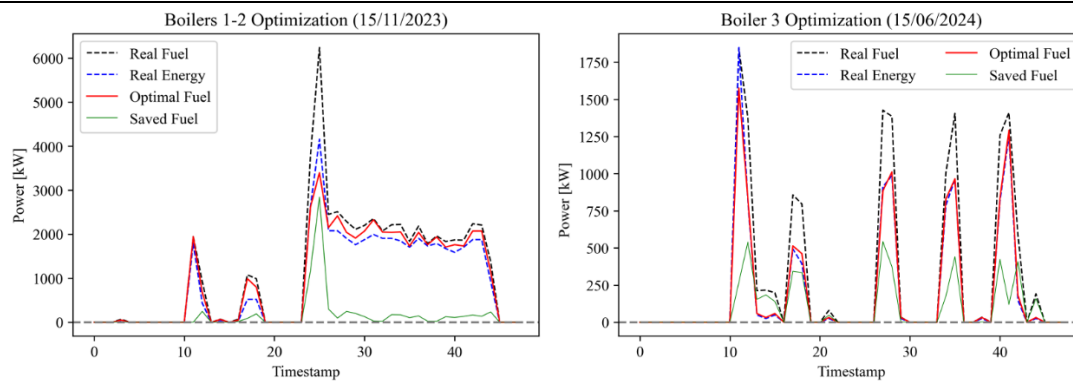


Fig. 27. Optimization Results During a Generic Day of Boiler 3 Operation (right) and Boiler 1-2 (left).

For a single day, the optimization process, which runs over 48 timesteps (the next 48 half-hours after the optimization is initiated), takes between 19 and 23 seconds to complete (tested on a notebook PC with an i7-12650H processor and 16 GB of RAM), converging in fewer than 150 generations.

4.2.3.4 The Dashboard of Rio Vena Digital Twin

The dashboard development is a critical phase in developing a DT because it ensures proper interaction with the end user and the models that are implemented within the twin itself. In this particular case, the Rio Vena DT started to develop from the gathering of user preferences concerning data that should be visualized, and in what format model outputs should be presented. Once these preferences had been clarified, and the models had been conceptualized, partly developed in the alpha version, a mock-up of the interface was presented to the pilot and feedback was gathered. Following this, the interface development began, initially using Grafana [19] for the first release and React [20] for the final version, as described in Chapter 3. Finally, these models were organized as an API that could be accessed from the different pages of the dashboard.

The DT would enable facility managers to enhance operations and maintenance through a dedicated dashboard that displays real-time data, recommended strategies, and data quality reports. This DT enables detailed monitoring of more than 177 parameters at boiler and dwelling levels such as fuel consumption, flow rates, pressures, and temperatures.

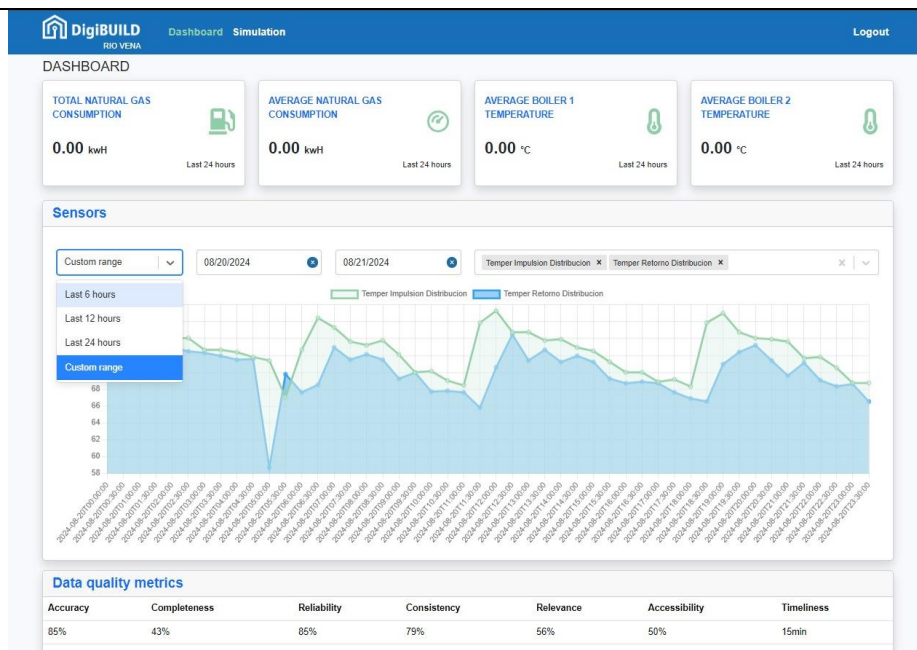


Fig. 28 Digital Twin Dashboard for Data Monitoring and KPIs Visualization.

Fig. 28 shows the first page of the DT's dashboard, illustrating how users can select different timeframes to visualize data and its trends. At the bottom of the page, details about data quality Key Performance Indicators (KPIs) for Rio Vena are displayed. These indicators, produced as part of another task in the project, assess the reliability, completeness, and accuracy of the data. Various boxes contain summary parameters that are useful for DHN managers, offering a clear overview of the system's status and highlighting any potential issues.

On the simulation page (Fig. 29), the manager can run the optimization algorithm, which incorporates ML models developed, to determine the optimal boiler management strategy at the main plant. This page displays a graph showing the setpoint temperature behaviour of the boilers throughout the day, as generated by the optimization process, to provide a comprehensive view. For added precision, temperature details for each hour are presented on the right side of the chart. Finally, at the top of the page, key results, including potential gas and cost savings, are displayed to help the manager understand the benefits of following the proposed strategy.

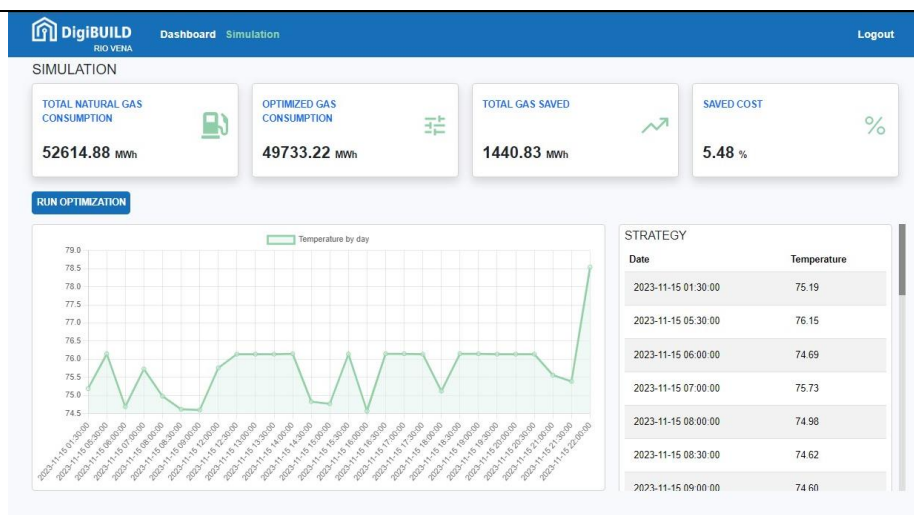


Fig. 29 Rio Vena Digital Twin Simulation Page

4.3 Digital Model of a Fourth Generation District Heating Network

The review conducted in these three years revealed that research on the dynamic behaviour of the primary physical factors in DHC systems is limited and incomplete [21]. Therefore, there is a need for deeper knowledge to support the advancement of next-generation DHNs that incorporate prosumers. It is both intriguing and essential to examine the impact on temperatures, flow rates, and pressures when a prosumer introduces heat into the network, taking into consideration dynamic factors such as the thermal inertia of the DHN. The objective of the study of this section is to develop a dynamic model for a DHN that can produce more accurate results for the operation of a 3rd and 4th generation network. This model will incorporate different RES as heat sources and it aims to address the gap of knowledge regarding the thermohydraulic behaviour of a grid that operates with several prosumers. The presence of these prosumers significantly would impact the grid, necessitating a different approach to its design and operation. By emphasising the pressure analysis, the power of the primary pumps when a prosumer injects hot water into the supply loop will be analysed. A double loop ring-shaped network is simulated in the TRNSYS environment to conduct dynamic analysis and observe the thermohydraulic variables under various operating situations. Additionally, a preliminary design phase was carried out to determine the appropriate dimensions for the pipes, select the appropriate units for the thermal power station, and choose the most suitable pump by analysing the pressure changes in the circuit.

4.3.1 Methodology

In Chapter 3 the general methodology for physics-based modelling was detailed. For this reason, in this section, the TRNSYS model of the case study will be presented. It is worth noting

that the model here proposed relies on elements that are readily accessible in the TRNSYS libraries and have been previously authenticated [22]. In addition, the hydraulic analysis was carried out by using equations found in handbooks [23] or previously presented by Lickleder et al. [24].

4.3.1.1 Network design on TRNSYS

TRNSYS is a software that allows transient systems to be simulated, and it is mainly used for studies of energy and electrical systems as well as DHC systems [25]. The software consists of two macro-parts: the first is the system of algorithms that processes the input file, studies, and verifies the convergence of the equations system, solves it, and provides the required output. The second part is an extensive library of energy and electrical components (called Types) available for the composition of the system, and it is also possible, using a calculator type, to insert different equations (made up of input and output) that can compensate for the lacks.

Components (types) capable of simulating the operation of the grid correctly have been chosen. The model is made by equations, implemented as calculators, and these types:

- Type 11f: Controlled flow diverter. This is used to divert the flow rate from the supply ring to the return one (the opposite for the producers). This is controlled by the equations on the substation and simulates a 2-way valve installed at each user.
- Type 11h: Tee-piece. This is used to mix the flow rate from each substation to the one in the cold ring.
- Type 31: Pipe/Duct. It is used to simulate the heat losses with the ground.
- Type 496: Hydraulic Pipe. Type developed in this work for hydraulic calculations.

Using these types in addition to other data readers and equations, it is possible to determine the hourly time series of flow rates and temperatures in the network. To consider the heat losses in the pipes the loss coefficient has been calculated using the data from the brochure previously noted and the ground temperature has been set as the environment temperature.

The initial pressure on the supply side is imposed at 9 bar, while on the return side, it is set at 7 bar. These values were selected to facilitate the visualization of pressure drops in each branch under various operating scenarios.

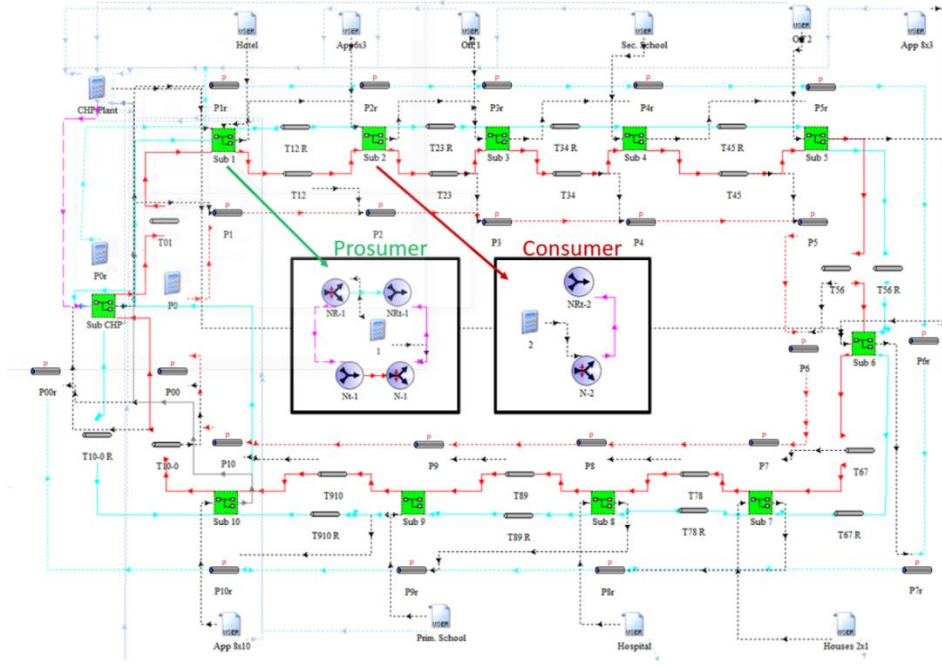


Fig. 30. Case study DHN model developed on TRNSYS

The model's primary focus was on thermal analysis. TRNSYS types including Type 31 "Pipe or Duct", Type 11f "Controlled flow diverter", and Type 11h "Tee piece" were initially used for the creation of the dual-loop model. The calculation of the thermal side's behaviour during the year-long simulation in terms of flow rates, temperatures, and energy balances was made possible by the integration of these three types, together with calculators for controls, datasheets, and weather files. The thermal model is inextricably tied to the hydraulic computation because Type 496 uses the flow rates from Type 31 as inputs. Finally, this type provides a dynamic stationary analysis of the pressure drops in each branch and the pressure at all nodes based on the equations previously described in section 2.2.

4.3.1.2 CO₂ emissions estimation

As previously mentioned, 4th generation DHNs are more environmentally sustainable due to the use of RES such as solar collectors or PV panels for power-to-heat technology. According to this, the following equation is used to calculate the hourly emissions m_{CO_2} associated with the operation of the DHN:

$$m_{CO_2} = W_{pump} \cdot f_{el} + \frac{Q_{gen}}{\eta_{chp} \cdot LHV_{NG}} \cdot f_{ng} \quad (26)$$

Where W_{pump} is the energy consumption of the pumps, LHV_{NG} the LHV of the Natural Gas (NG) used by the CHP, f_{el} and f_{ng} are respectively the conversion factor for CO₂ of the electricity and the natural gas. The efficiency η_{chp} of the generation system depends on the load

factor and the dry bulb temperature of the air. According to the method used in [26], utilizing data derived from an accessible gas turbine [18], an efficiency curve was modelled assuming air temperature and load factor as the most influencing variables. This allows for the computation of hourly natural gas consumption.

4.3.2 Case Study

During the operation of a DHN, the needs of each building change continuously and so do the flow rates. Consecutively, the pressure increases when the heating demand becomes greater. The TRNSYS model developed in this study aims to evaluate the main thermohydraulic properties of the grid on an hourly basis. The main parameters considered are pressures, flow rates, temperatures, and energies.

4.3.2.1 Case Study Description

For the case study, a cluster of 10 users has been considered. The cluster is in Palermo, in the South of Italy where cooling loads prevail over the heating ones. In particular, Palermo belongs to the B climatic zone according to Italian legislation [27] and it is characterized by 751 heating degree days. In Fig. 31 there is a representation of the network and its users. In a clockwise direction: power plant, hotel, six small 3-storey apartments, first office, secondary school, second office, four big 3-storey apartments, 20 small houses, hospital, primary school and 3 10-storey apartments.

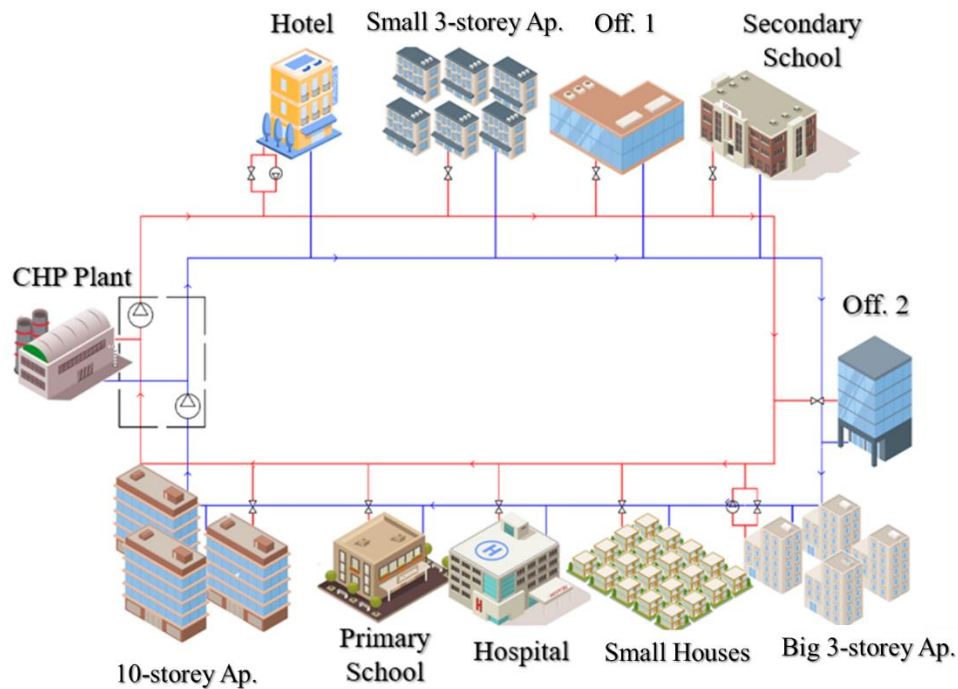


Fig. 31. Case Study Cluster

It is clear from Fig. 32 that in Palermo, building cooling loads are greater than building heating

loads. This is primarily because of the local climate, where the outside air temperature rarely drops below 10 °C. Contrarily, when solar radiation is more intense in the summer, there is a corresponding temperature rise, which necessitates a greater cooling load in order to maintain a comfortable environment within the buildings.

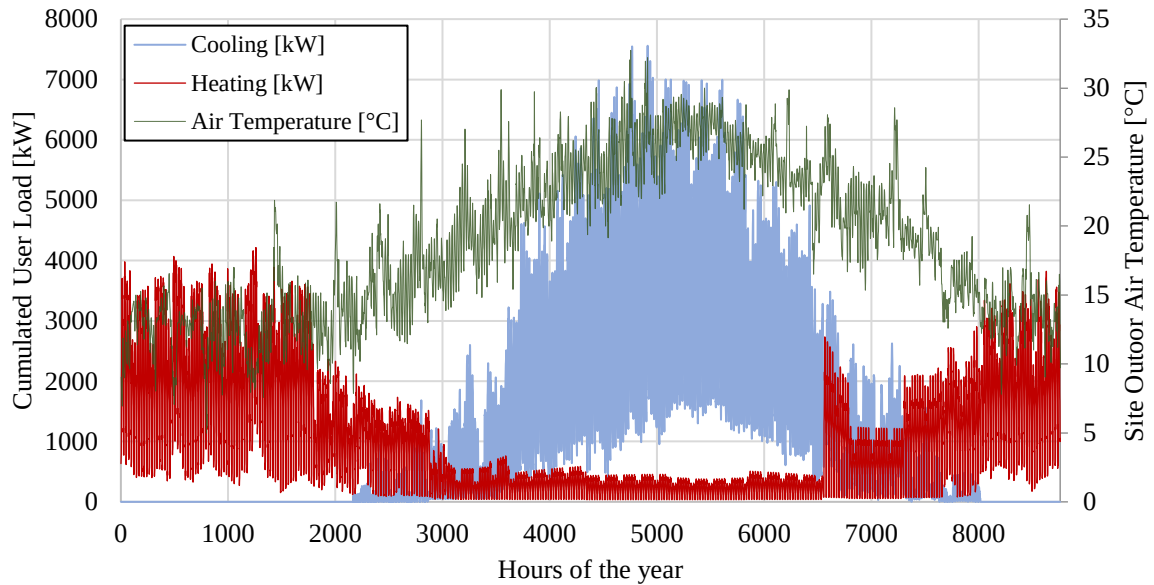


Fig. 32. Cumulative user loads and outdoor air temperature in the case study

The demands of hotels, offices, and the hospital have been previously estimated in another research work [28], where ad hoc energy audits have been performed by the authors. Regarding other groups of buildings, demand profiles were obtained from the simulation performed on Energy Plus [34] using the archetypes of buildings available in [29] considering the buildings for a warm-humid climate, which represents the climate zone of Palermo [30] and considering climate data available in [31]. Each user is indirectly connected to the DHN via a substation. The power plant will generate heat in all seasons, it directly feeds the exchangers in heating mode, and it supplies the absorption chillers which are installed in each user in summer.

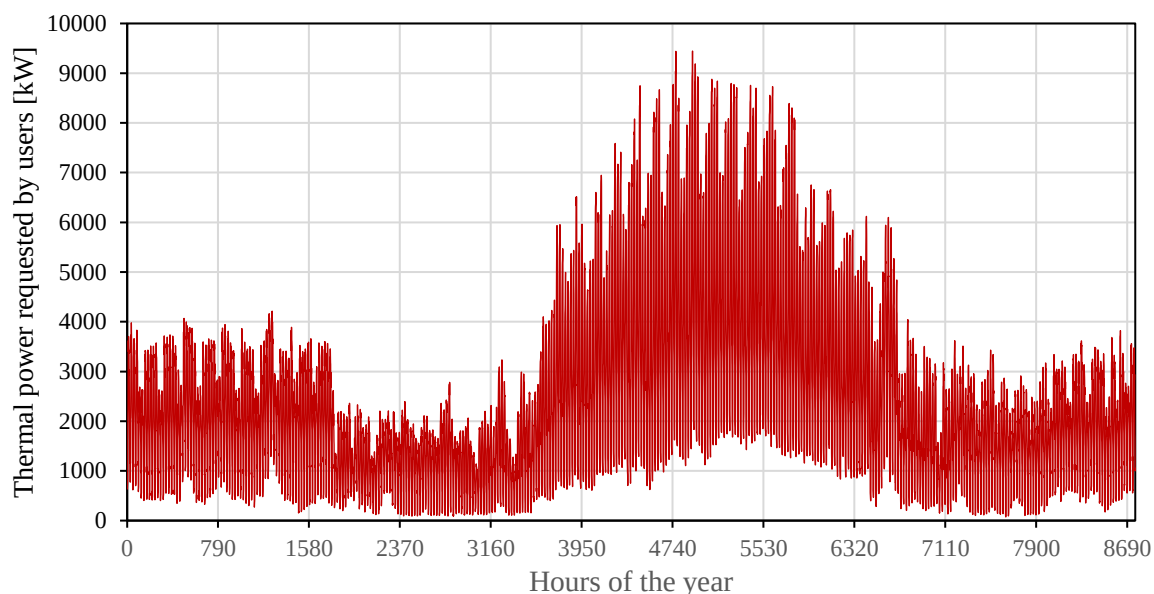


Fig. 33. Overall thermal load for all the buildings in the case study

In this case study, the maximum contemporaneous demand of heat is 9.44 MW on July 18th during the cooling season for feeding the absorption chillers as can be seen in Fig. 33. It was chosen to consider a central plant made of two gas turbines. More specifically, the first one with an electrical nominal capacity equal to 3 MW and the second one 500 kW. Thanks to this configuration, it is possible to partialize the thermal capacity from 10 MW to below 100 kW, so all the loads can be covered. To simulate the gas turbine CHP, efficiency curves available in [26] have been utilized to calculate the required fuel for generating the desired heat.

Table 9 provides a detailed examination of the thermal requirements of each user within the district. Despite its small size, the district interacts with a diverse range of end-users, cumulatively necessitating a concurrent thermal demand of 9.4 MW. Certain facilities, such as offices and schools, exhibit a low load factor due to reduced demand during nighttime hours. Conversely, residential buildings and hotels demonstrate a higher load factor, reaching up to 36 %, reflecting their consistent thermal needs.

Table 9. User and cumulative thermal load

	Total Thermal Energy	Maximum Thermal Power	Mean Thermal Power	Load Factor
	[MWh]	[MW]	[MW]	[-]
Hotel	6839.78	2.19	0.78	36 %
Small 3-storey apartments	1664.80	0.62	0.19	30 %
Office 1	1215.69	1.18	0.14	12 %
Secondary school	1456.61	1.07	0.17	16 %
Office 2	868.35	0.84	0.10	12 %
Big 3-storey apartments	1494.51	0.78	0.17	22 %
Small houses	773.54	0.42	0.09	21 %
Hospital	3477.77	2.05	0.40	19 %
Primary school	798.60	0.43	0.09	21 %
10-storey apartments	2727.34	1.48	0.31	21 %
Total	21316.99	9.45	2.43	26 %

Cooling energy is supplied by absorption chillers installed at each user. In this way, the buildings just need heat from the network to supply both heating and cooling. The chiller chosen has a COP of 0.7 which was assumed to be constant.

The flow rate of each user is controlled by a two-way control valve. In the case of prosumers, the connection with the grid requires, to supply water from the return to the supply side, local variable speed pumps following the scheme suggested by Lickleder et al. [24]. The distances between each user are fixed at 200 m to avoid any divergence in the results that could depend on the distances.

All the equations mentioned before can be solved by imposing a value for the speed or the diameter. In this case, it has been proceeded iteratively using the commercial diameters found in a technical brochure. Both loops have been designed with the same single diameter to avoid section reduction and because the location of maximum flow rates could change because of the presence of prosumers.

4.3.2.2 Description of the investigated scenarios

Different scenarios were assumed to study the trend of the thermohydraulic parameters of the DHN under different operating conditions. Differences in energy and emissions between all the scenarios will be highlighted.

Base Case. It represents a 3rd generation network in which all users are consumers. The design temperatures are 100 °C on the hot side. The high temperatures are chosen to be compatible with a 3rd generation DHN but also with the production of cooling energy through absorption chillers

installed at each user. A plant, made by three CHPs operating in parallel, generates the heat needed by all the users.

Scenario “1”. In the first scenario, the potential of 4th generation DHNs is investigated. During the autumn, winter, and spring months, the supply side of the grid operates at temperatures of either 80 °C or 60 °C to minimize heat loss, so called respectively scenario 1a and 1b. However, during the cooling period, the grid must operate at 80 °C due to limitations with absorption chillers at lower temperatures. In this scenario, a field of evacuated tube solar collectors with a total power output of 7 MW_p in order to increase the share of RES in the grid. This scenario aims to demonstrate the advantages of lowering the temperature of a DHN such as reduction in thermal losses and integrating solar collectors into the system.

Scenario “2”. In this scenario, the supply temperature is 60 °C in the heating period and, to support the solar field and the CHP, two consumers become prosumers in the period between December 1st and 31 March 31st. In the context of high-RES penetration, these users will be producers during the solar peak hours (11 am to 2 pm). Each of them has an HP, with a capacity of their maximum load, that they use for self-consumption and during peak solar hours the HPs work at full load, feeding the surplus of heating energy into the grid. This scenario will show the difference in terms of pressure when a prosumer helps the CHP and there will be energy savings.

4.3.3 Results and Findings

All scenarios have been simulated and the results will follow in this section. For each of them, the analysis will focus on pressure, temperature, energy, and emissions saving.

4.3.3.1 Results for the Base Case

A preliminary study in stationary has been carried out for designing the DHN. The Euroheat Guidelines [32] recommend operating pressures of 6 bar for districts operating at 90°C and 16 bar for districts operating at 110 °C. For this reason, modern districts are kept between 6 and 10 bar. The choice of the diameter of the pipes is linked to the specific pressure drops in every branch of the network [33], for instance, the Swiss guide for the DHC system suggests specific pressure drops between 150-250 Pa/m [34].

The most stressed branch, in this case, is the one in the supply loop after the main production plant and the iteration results, considering the maximum loads of each user, are:

- $D_{01} = DN\ 250\ (263,7\ mm\ D_i)$
- $\Delta P_{m;01} = 140.2\ \frac{Pa}{m}$

The preliminary study's results were used to fix the diameter in the types used for thermal and hydraulic analysis, respectively type 31 and type 496. In order to avoid an anomalous

increase in pressure drops when a prosumer injects heat into the network, the same diameter for each pipe in the network has been considered. Since the same length of 200 m has been chosen for each pipe, the total length of one ring of the district considered measures 2.2 km.

In the case study considered, as can be seen in Fig. 34, the typical trend of pressures in a double loop DHN is achieved [35], guaranteeing a differential pressure of 1 bar at the most stressed user.

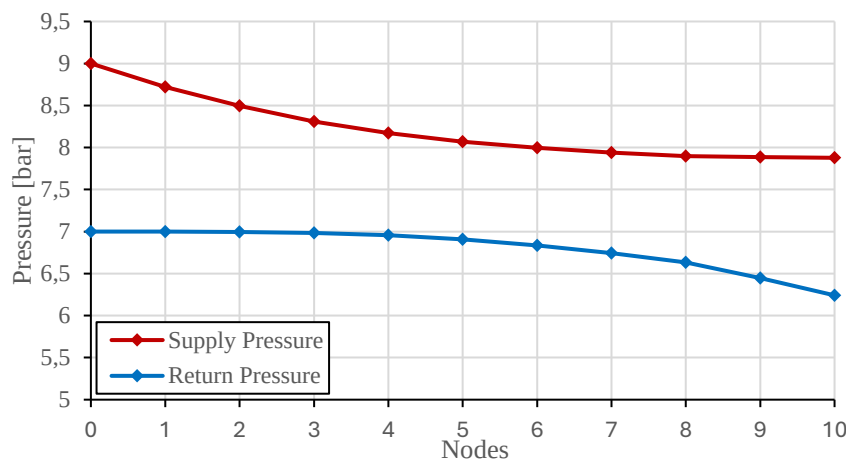


Fig. 34. Trend of pressure for each node of the network in the design phase

At first, the network has been simulated under standard conditions. A comparison will be made, as possible, to the other cases. In this scenario, it was assumed that the presence of no other producer than the power plant, and the temperature of the supply ring is the highest possible (100 °C).

As illustrated in Fig. 35, the pressures just before the pumps, represented by P_{o*} and $P_{o*} R$, are lower during the summer than during the winter. This is due to the higher flow rates during the summer months when the cooling load exceeds the heating load. The maximum pressure drop, measured before and after the pump, is 0.13 bar for the supply side and 0.26 bar for the return side during winter. During the cooling season, it is 0.91 bar in the hot ring and 0.85 bar in the cold ring.

The difference in pressure drops between the two rings is more pronounced during winter than during summer. This is because users at the end of the network, primarily residential buildings, do not require as much heating energy. As a result, less water flows through the last part of the supply side compared to the return side, where flow rates reach high values in branches near the CHP system.

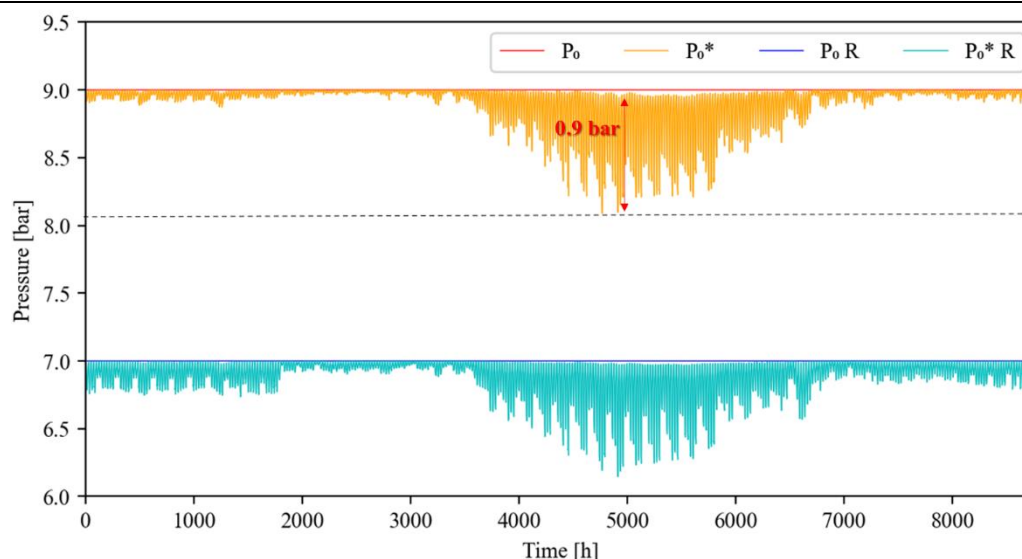


Fig. 35. Hourly pressures, on the sides of the pump, during a year in the DHN

The total electric energy requested for pumping in winter is 2.06 MWh and it is 21.5 % of the summer. It is also interesting to notice that during winter the supply pumps need 49.3 % of electricity than the other side.

Table 10. Electric Energy requested by the pumps

	Winter Energy [MWh/y]	Total Energy [MWh/y]	Max power in winter [MW]	Max power during the year [MW]
Supply pumps	0.681	5.772	0.002	0.012
Return pumps	1.384	5.88	0.003	0.010

Focusing on the trend of pressures in Fig. 36 it is possible to notice that the pressure drops are almost negligible at night. In the cluster considered, four users are apartments and there are two schools and two offices, which are closed at night. During winter, this situation leads to a reduction in total demand whose average value is 141.8 kW for nights and 501.2 kW for the rest of the day. The DHN is designed to operate in the worst condition when the load reaches its maximum. For this reason, when the flow rate decreases the pressures stay close to the operating pressure. In summer this difference is more emphasized since in the nighttime, the mean flow rate processed by the CHP is just 23.7 % of the mean flow rate during the daytime.

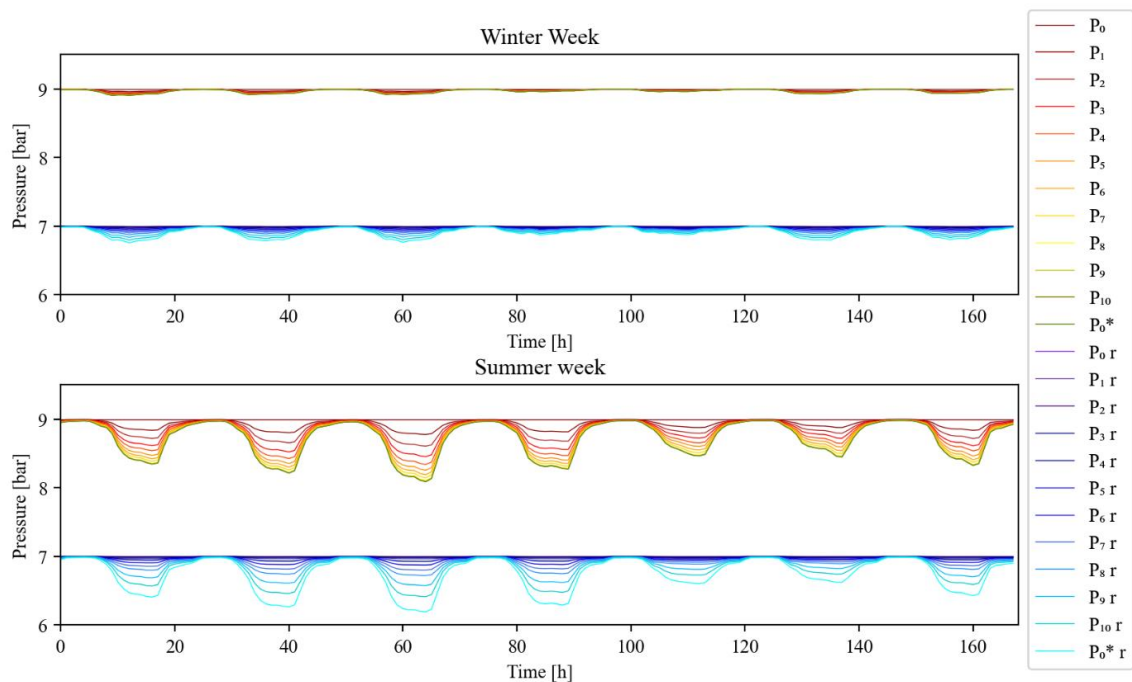


Fig. 36. Hourly pressures during standard weeks in the DHN

Table 11 shows the changes in pressure decreases for two representative hours on a typical winter and summer weekday for each branch of the district. The table shows that on the supply side, the first branch is the most stressed while, on the return side, this happens for the last one. It's important to notice that through the last branch of the supply side and the first branch of the return side, the pressure drops are practically negligible, as only a minimal flow rate passes through these branches. These results also highlight the fact that the total pressure drops during summer are almost 32 times larger for the supply side and 10 times higher for the return than those during winter.

Table 11. Pressure drops in network branches

		Total	0-1	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10	10-0*
		[kPa]	[kPa]	[kPa]	[kPa]	[kPa]	[kPa]	[kPa]	[kPa]	[kPa]	[kPa]	[kPa]	[kPa]
Supply Side	Wint.	2.59	1.28	0.31	0.17	0.17	0.17	0.17	0.15	0.13	0.01	0.01	0.00
Return Side	Week	7.86	0.00	0.40	0.58	0.58	0.58	0.58	0.62	0.65	1.09	1.09	1.72
Supply Side	Sum.	82.82	20.99	15.53	14.32	10.21	7.44	5.30	4.15	3.74	0.74	0.39	0.00
Return Side	Week	78.05	0.00	0.57	0.84	2.30	3.93	5.77	7.07	7.61	14.21	15.87	19.89

The temperatures in each node of the grids change mostly with the flow rate. For instance, at night, when the flow rates are low, these begin to drop, while during the day the drops are lower (Fig. 37). During the year the heat losses are stable and, since they depend on temperature, the

losses on the supply calculated are 582.4 MWh and these are higher than the return side of factor 1.3.

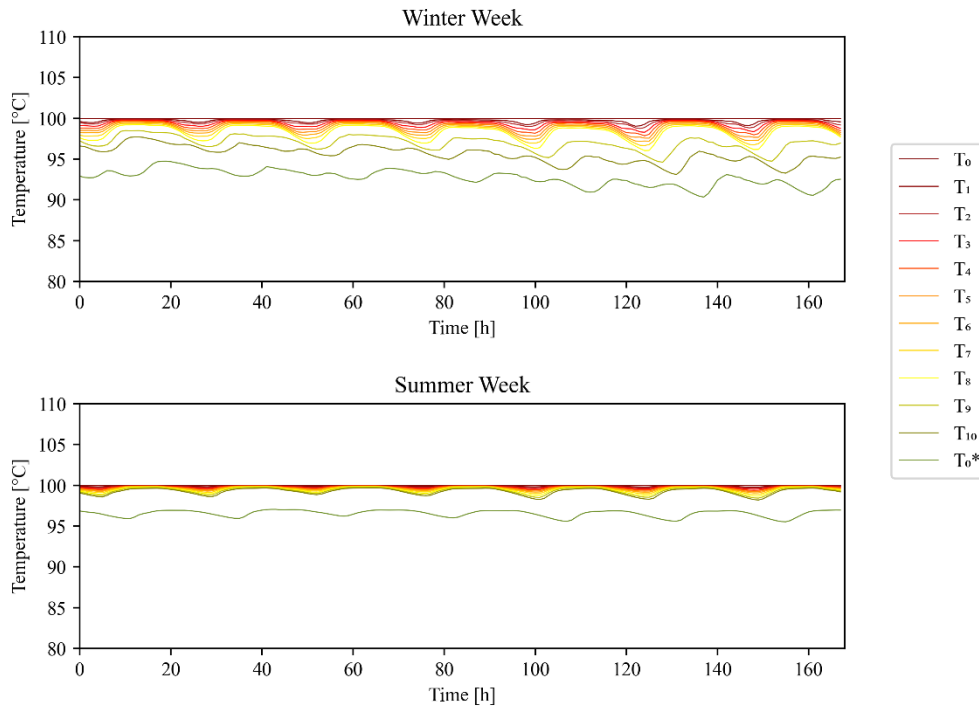


Fig. 37. Hourly temperatures during standard weeks in the DHN

The variation of temperature between the first and last node is higher in winter than in summer. This happens because, as said before, in winter the flow rates are lower and especially in the last branches. In this situation, despite the high variation of temperature in a branch, the losses are not the highest, because they depend strictly on the flow rates.

To ensure thoroughness, a comparative analysis of temperatures and pressures during standard hours was performed between the TRNSYS and stationary models developed on EES [36].

As illustrated in Fig. 38 and Fig. 39, the discrepancy between results from the dynamic and stationary models is minimal. The maximum relative error for pressures is 0.7 %, while for temperatures it is 2.1 %. It should be noted that the greatest error in pressure behaviour occurs during periods of high demand (e.g., summer), whereas the maximum temperature error occurs when the load is very low (e.g., spring). This can be attributed to the fact that as the load increases, the flow rate and pressure drop increase, while the temperature variation in each node decreases (as depicted in Fig. 39). In both cases, the error becomes higher with distance from the CHP, making it more noticeable in larger DHNs.

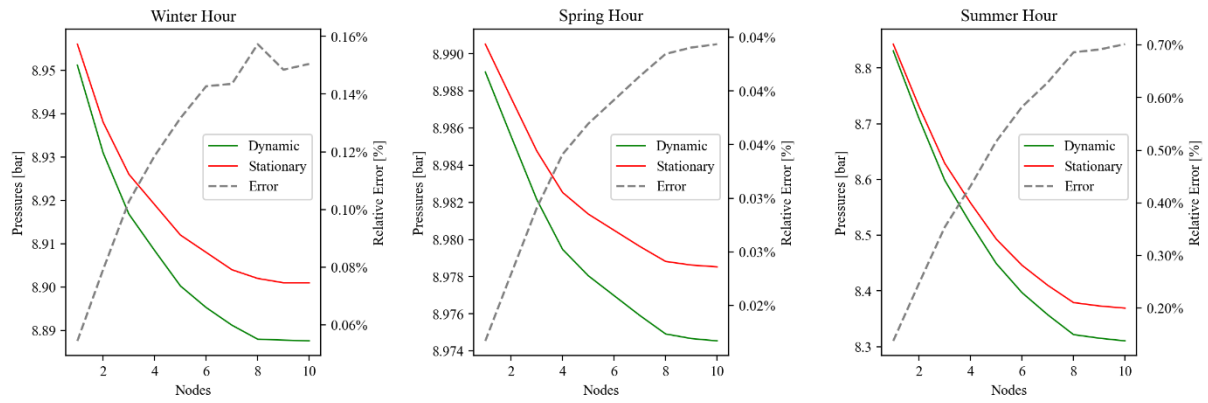


Fig. 38. Pressures comparison during standard hours

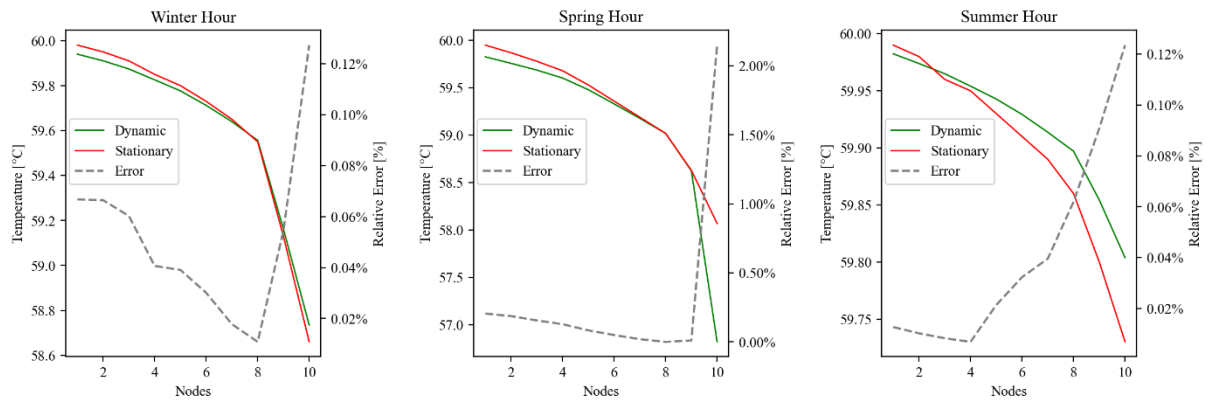


Fig. 39. Temperatures comparison during standard hours

4.3.3.2 Results for Scenario 1

Scenario 1 is characterized by the transition to lower temperatures. The network starts operating at 80 °C and then at 60 °C. In this way, it is possible to feed the users only for heating, because at 60 °C the grid can't feed the absorption chillers installed in each user. The network will be managed with a supply temperature of 80 °C only in the conventional cooling period in Palermo, from June 1st to September 30th. Lowering the temperature of the net, it is easier for the prosumers to assist the power plant, a hypothesis that has been analysed in the third scenario.

The variation of temperature during the day will decrease too when lowering the temperature (Fig. 40), indeed, the mean variation for all the heating period decreases from 7.26 °C to 3.83 °C in the case of 60 °C.

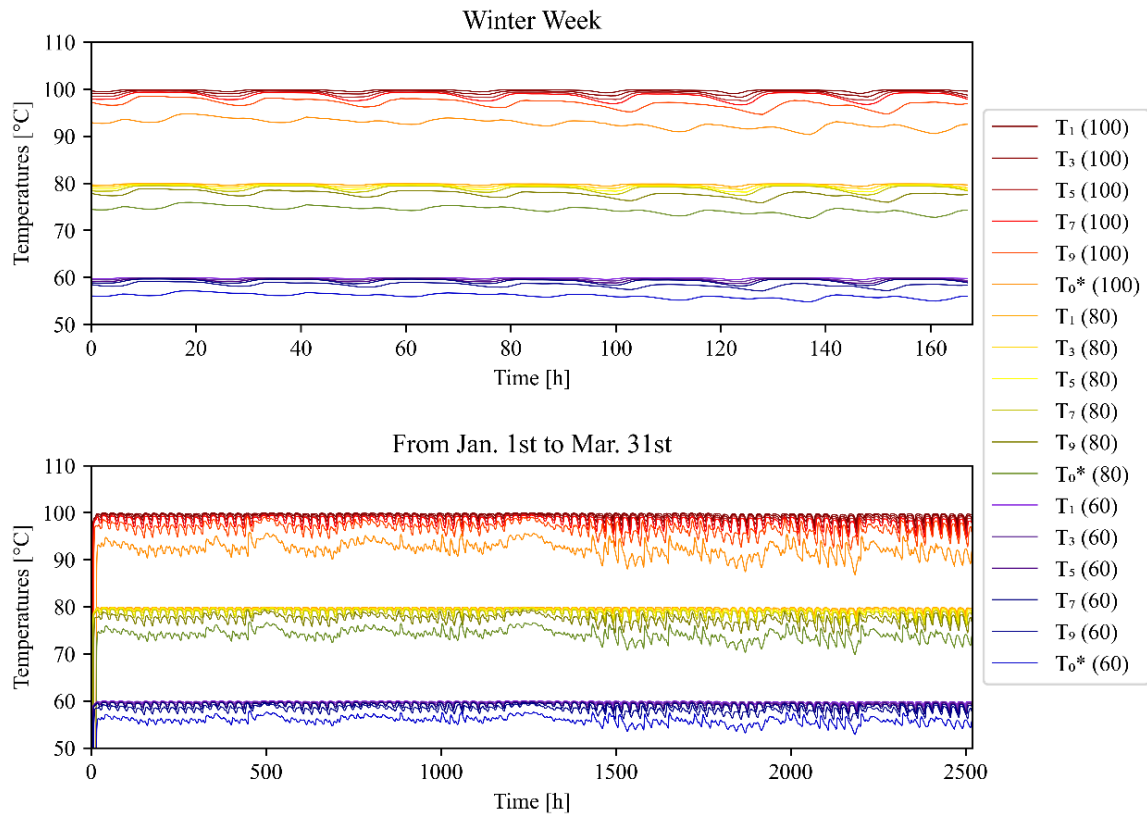


Fig. 40. Temperatures of supply-side nodes operating at 3 different temperatures (100 °C, 80 °C, 60 °C)

In Table 12 it is possible to see how lower the temperature brings considerable energy savings. In this case study, looking only at the heating period, the losses are reduced, compared to the Base Case, by a factor of 1.36 working at a temperature of 80 °C and 2.12 using a temperature of 60 °C. The yearly energy losses vary between 1040 MWh (Base Case) and 475 MWh (Scenario 1b) corresponding to only 4.9-2.2 % of the total heating energy requested by the users (Table 9).

Table 12. Heat losses in the pipes during winter

	Supply Heat Losses [MWh]	Return Heat Losses [MWh]	Mean ΔT [°C]	Maximum ΔT [°C]
Base Case (100 °C)	395.21	313.64	7.24	13.80
Scenario 1a (80 °C)	302.02	219.50	5.53	10.60
Scenario 1b (60 °C)	208.83	125.34	3.83	7.40

In Scenario 1, thanks to the reduction of supply temperature, the main central plant is also supported by a 7 MW_p solar field composed of evacuated tube collectors connected with the district by a 5 m³ storage/heat exchanger. The flow rate drawn from the return ring is pre-heated in the exchanger and if the temperature is below the setpoint the water reaches the CHP which produces

the necessary heat for reaching the desired temperature. Using this coupling, in scenario 1b, it is possible to save 5564 MWh of heating energy which is produced by the solar collector reducing the total energy produced by the CHP by 25.5 %. It is interesting also noticing that, while in full demand periods (such as summer or winter) the solar collectors only pre-heat the water. In low-demand periods (such as spring or autumn) during the solar peak hours the DHN is fully fed by the solar field leading to the previously cited energy savings.

4.3.3.3 Results for Scenario 2

As mentioned before, in this scenario, in addition to the solar collectors, the hotel and the big 3-storey apartments become prosumers. The logic of energy production is simplified: the electric network operator asks these users to use their HPs at full load, from 11:00 AM to 2:00 PM, to absorb the electricity surplus due to a high-RES penetration. Specifically, the system incorporates a high-temperature air-to-water heat pump which uses the outside air as the source. This heat pump draws water from the return side, heats it to the predetermined setpoint temperature, and subsequently injects the heated water into the supply side. In this case, the supply temperature must be settled at 60 °C, so their contribution can not be evaluated in the cooling period (from the 1st of June to the 30th of September in Palermo). During these three hours, they will generate hot water when the capacity of the HP, designed on the maximum heating load, is higher than their load. In summer they just don't receive heat from the network, so their loads are set to zero in those hours.

When the generation is distributed the pressure level will change among the network. The pressure drops decrease when the prosumers inject because the flow rate gets distributed more regularly (Winter Week in Fig. 41), but it also decreases when the two consumers don't request heat, as in the summer case (Summer Week in Fig. 41). The differences are highlighted by comparing Fig. 41 with Fig. 36.

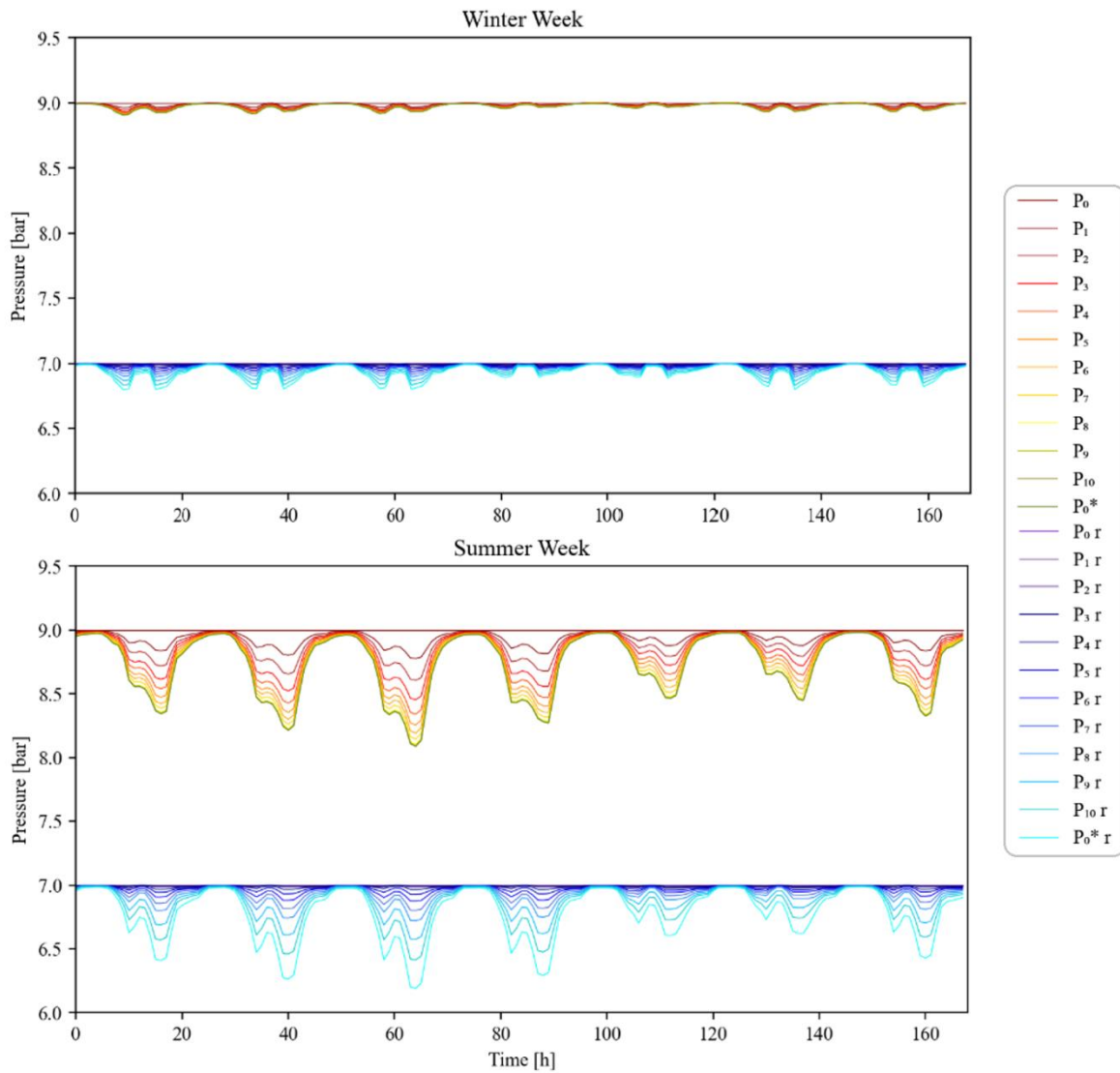


Fig. 41. Pressures distribution on a winter and on a summer week in Scenario 2

During the winter months, the energy required for pumping is 73 % of that required in the Base Case for the same period. For the remainder of the year, assuming that two users do not require heat in the grid for 3 hours per day, the energy savings amount to 1.5 MWh, or 15.5 % of the Base Case energy. When comparing the results of Scenario 2 with those of Scenario 1, it can be seen that the heat production avoided by the main plant due to the presence of prosumers amounts to 1265 MWh which represents 7.8 % of the total for Scenario 1 and 31.3 % for the Base Case. Focusing solely on the heating period when prosumers are active and the network operates at 60 °C, the energy savings compared to Scenario 1 reach 11 %.

It is crucial to keep in mind that the situation under consideration is influenced by the location and size of the prosumer, as well as the length of the connections. Therefore, in the proposed methodology, the prosumers are positioned at two distinct points within the network, while keeping the pipe length and diameters constant.

4.3 Discussion on energy savings and environmental impact

Table 13 summarizes the improvements observed in each scenario. Only the results of scenario 1b are included in the table, since the difference in energy losses between scenario 1a and 1b is negligible, being about 2 % of the energy generated by the CHP. As expected, Scenario 2 is the optimal scenario in terms of energy saving, as the CHP needs to produce 31.3 % less thermal energy than the base case. The comparison between Scenario 1b and Scenario 2 shows that both have similar heat losses since the operative temperature is the same, but prosumers interaction in Scenario 2 enables saving 1266 MWh of thermal energy. The electric energy required for pumping is the same between scenario 1b and Base Case since no prosumers are considered and the network is managed in the same way. As discussed in the last section, the injection of prosumers leads to an overall energy saving of 2 MWh.

Table 13. Main results for each scenario

	Energy generated by CHP [MWh]	Energy generated by RES [MWh]	Heat Losses [MWh]	Energy for Pumping [MWh]
Base Case	21793	0	1040	12
Scenario 1b	16229 (-25.5 %)	6040	475	12
Scenario 2	14963 (-31.3 %)	7305	475	10

In terms of the efficiency of the system, Scenario 2 is the best configuration. To analyse the generation of CO₂ emissions, both the electricity consumed by the pumps and the fuel consumption of the CHP must be considered.

The avoided impact of CO₂ can be calculated, knowing that in Italy the amount of greenhouses gas burdening the production of electricity amounts to 307.7 gCO₂/kWh [37]. Thanks to the benefits in terms of a more regular pressure level, in Scenario 2 the energy saving for pumping operation amounts to 2.05 MWh which corresponds to an avoided production of CO₂ of 0.63 t, which is 17.6 % less than the Base Case.

In Scenario 2 the fuel consumption of the main plant will be the lowest of all the scenarios. This happens for 2 reasons: because the network operates at a lower temperature, 60 °C instead of 100 °C, and this reduces the heating losses. The most impactful reason is due to the 2 users who became prosumers. They feed the DHN in winter and reduce the consumption for the rest of the year, reducing the quantity of heat that the plant must generate.

The assumption is that all the Primary Energy (PE) is equal to the energy of the natural gas used, so it has been considered a factor of 227.31 gCO₂ for 1 kWh of natural gas according to [38]. The findings indicated a reduction in CO₂ emissions by 2803 t, which constitutes 31.3 % of the emissions in the Base Case scenario.

Table 14. CO₂ emissions in the main cases of the case study

Cases		CO ₂ for pumping [t/y]	CO ₂ for heat production [t/y]	CO ₂ hourly emissions [kg/hr]
Base Case	Winter	0.64	4526.34	0.78
	Total	3.59	8952.96	1.02
Scenario 2	Winter	0.46	3101.12	0.53
	Total	2.95	6149.44	0.70

4.4 Digital Model of a Fifth Generation District Heating Network

According to the IEA, novel technologies of DHNs, along with the reduction of operating temperatures, the improvement of pipe insulation, and digitalization, could significantly reduce heat losses, potentially containing them below 10% [39]. For this reason, the goal of this work is to study and develop a model of a modern DHC network associated with a thermal power plant consisting of a solar collectors field, an absorption chiller, and seasonal thermal storage. The choice of this model configuration is based on the desire to reduce dependence on traditional energy sources, favouring instead the use of RES, such as solar energy. The presented methodology includes details on the modelling of the network components and the main power plant. Once the concept of a DHN was defined, a model of a fifth-generation network was developed in the TRNSYS (TRaNsient System Simulation tool) software [18] to conduct a dynamic analysis, to highlight the network's behaviour during the operating period. The model is a double loop network, i.e. a network configuration that more simply allows the insertion of two potential prosumers, associated with a thermal power plant that provides water at a temperature of about 25 °C in the winter period and at a temperature between 20 °C and 8 °C in the summer period. The network serves a cluster of 8-10 users that simulate a residential neighbourhood in a metropolitan area, each building is associated with a substation characterized by a water-water HP and, in some scenarios, a heat exchanger. Furthermore, the fifth-generation scenarios were compared with two additional cases: “Scenario 0” where the DHC network is not present and the production of heating is carried out independently, and “Scenario 1” which represents a fourth-generation network powered by a field of solar collectors and seasonal storage. Finally, in the fifth-generation network, two prosumers were introduced to evaluate how the integration of additional producers into the network influences the uniformity of temperatures. The main goal aims to determine the benefits derived from the use of DHC technology, comparing the results obtained among the three different case studies and evaluating the advantages of developing the latest generation networks.

4.4.1 Methodology

To study the behaviour of the fourth and fifth-generation DHN in dynamic conditions, a network model was developed using the TRNSYS software. Specifically, at first, the portion of the district named the "solar loop" was modelled, whose aim is heating the heat carrier fluid in winter or cooling it in summer, secondly, the Fourth Generation District Heating (4GDH) network presented firstly in [21] was modified to enable the operation of a 5GDH network, including user substations which allow interaction between a water-to-water HP or chiller and the network.

4.4.1.1 Description of Solar Loop

The solar loop section, where the heat carrier fluid is processed to feed the substations, meeting the heating or cooling demand of the users throughout the year, is represented in Fig. 42. This section of the model includes: a field of solar collectors modelled in TRNSYS through Type 71; a Pit Thermal Energy Storage (PTES), modelled using Type 4, which link the solar collectors and the network during heating period, and the absorption chiller during cooling period; an absorption chiller, fundamental in Solar Cooling networks, which is powered by the heat produced by the collector field and stored in the thermal storage; an auxiliary air-to-water chiller which support the absorber and finally valves and calculators for model management.

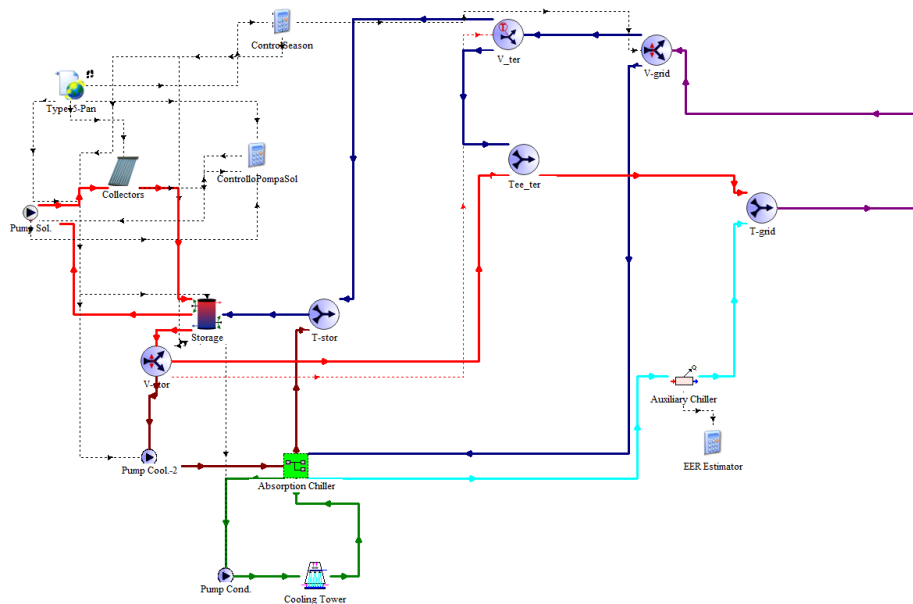


Fig. 42. Solar Loop Model Developed on TRNSYS

The correct functioning of this part of the network is managed and regulated by a series of valves and pumps that ensure that the supply piping circulates hot or cold fluid based on user demand and operation as shown in Fig. 43.

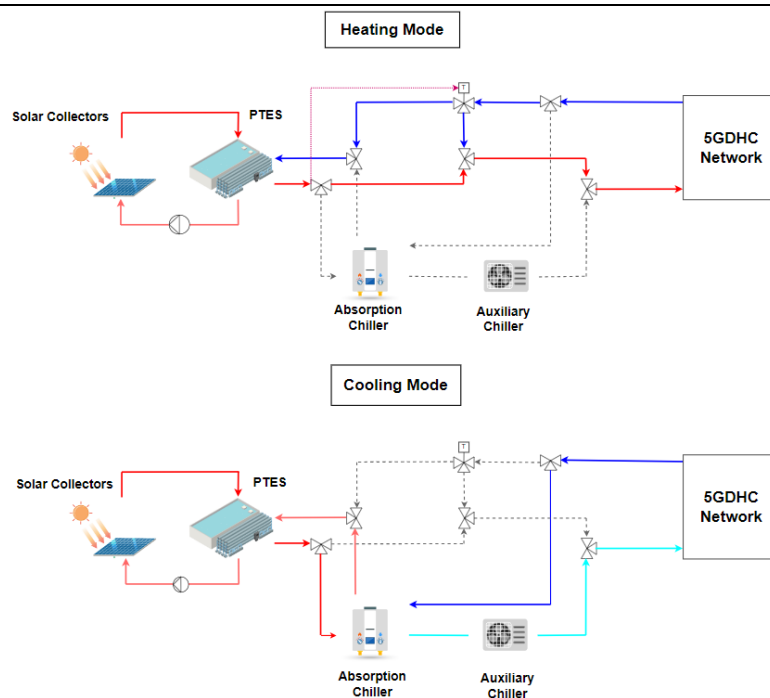


Fig. 43. Solar Loop Operation Modes

In the model implemented on TRNSYS, the thermostatic valve monitors the fluid temperature exiting thermal storage and ensures that the temperature sent to the district does not exceed the setpoint temperature, which for the study of a 5GDH network was set at 25 °C as it is compatible with the winter operation of hydronic HPs.

The field of solar collectors, as anticipated, was modelled in TRNSYS through Type 71 – Evacuated Tube Collector, which requires, to calculate collectors efficiency, several input parameters such as ambient temperature, total radiation incident on the panel surface, solar azimuth and zenith angles, and the inclination and azimuth angle of the panels. These parameters are input using Type 15 – Weather Data Processor, which functions to read data from an external file of meteorological data. Connected to the solar plant is a circulation pump modelled with Type 3d – Single Pump model, controlled by a calculator that allows its operation when the fluid temperature exiting the plant is higher than the fluid temperature entering and when there is solar radiation incident on the surface of the collectors.

For TES, a PTES system located near the solar plant was considered. It was modelled in TRNSYS through Type 4 – Stratified Storage Tank, simulating the behaviour of a sensible energy storage tank filled with a fluid subject to thermal stratification modelled by assuming the tank consists of 6 segments of equal volume mixed and filled with fluid from the solar plant. In winter, the heat carrier fluid exiting the storage is sent directly to the DHN, while in the summer period, it is used to power the absorption chiller to produce the cold fluid that feeds the network. The outside boundary is the temperature of the ground available through the Weather Data Processor, while the transmittance of the walls of the PTES was calculated considering a

concrete wall insulated with 1 m of EPS as recommended in [40].

Finally, an auxiliary chiller was used to support the absorption chiller (whose model was presented in Chapter 3) in cases where, in the presented 5th generation network model, the heat directly produced by the solar collector or collected in the storage is not enough to feed the absorption chiller. In this case, the absorber is bypassed, and the flowrate coming from the district is chilled by this component. In particular, the chiller is modelled with Type-92 - Auxiliary Cooling Device in combination with a calculator which is used to calculate, at the desired temperature and considering outdoor air temperature, the EER of the machine based on a catalogue of an industrial Air-to-water chiller [41].

4.4.1.2 District Heating Network Model

The modelled network consists of eight users and two prosumers connected through a ring network topology, a network configuration that includes two main rings: one for supply and one for return. Unlike traditional networks, where the return line starts from the last user and returns to the central plant, in this case, both lines start directly from the thermal plant and return to it [42]. The ring network also lends itself better to the connection of potential prosumers as it is more challenging for flow reversals to occur. The network model, shown in Fig. 44, developed in TRNSYS is mainly composed of six types of Types, found at each user node, and the user substations macro (previously detailed in Chapter 3). Type 11f: Controlled diverter valve: used in the supply piping to divert the flow requested by the consumer (or produced by the producer) from the flow that continues to circulate in the network; Type 11h - Tee valve: the valve used in the return piping to merge two flows, the one circulating in the pipe and the one exiting the user, into a single flow; Type 9e - Data Reader: it reads an external file to use it in the model. It is used to import user demands; Type 65a - Online Plotter with File: it is used to obtain the output data both graphically and in document form to analyse in a spreadsheet; Type 31: Pipe/Duct, used to simulate the heat losses with the ground, and Type 496: Hydraulic Pipe which was previously developed in [21] for hydraulic calculations.

4.4.2 Case Study

4.4.2.1 Description of Cluster of Buildings

Tancredi Testasecca

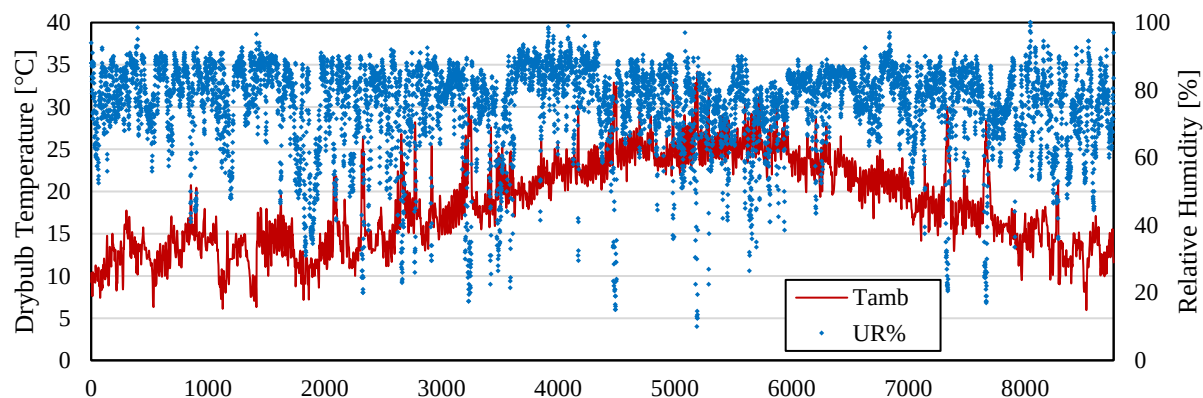


Fig. 45. Hourly Trend of Ambient Temperature and Relative Humidity of Palermo

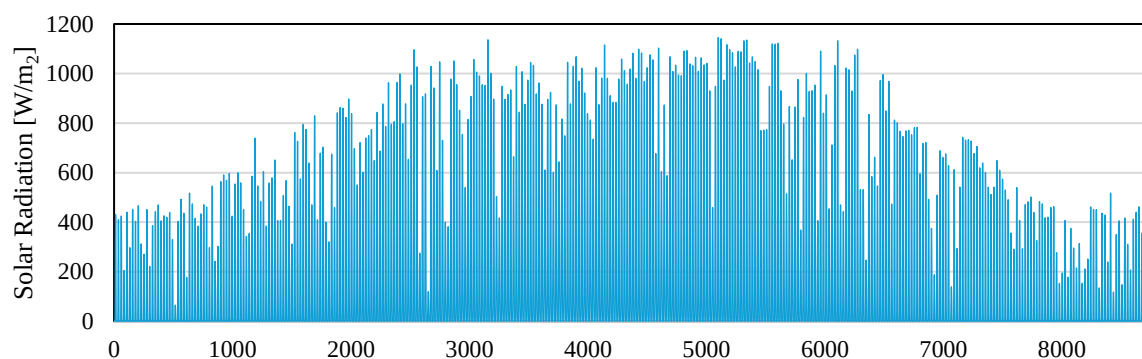


Fig. 46. Trend of Global Solar Radiation on the Horizontal Plane in Palermo

Fig. 47 shows a representation of the network and its users. Specifically, clockwise: The first prosumer (Sit Restaurant), a medium-sized office, a group of three high-rise apartments, a hospital, a supermarket, the second prosumer (the hotel), 8 mid-rise apartments, a primary school and high-density office and 4 high-rise apartments

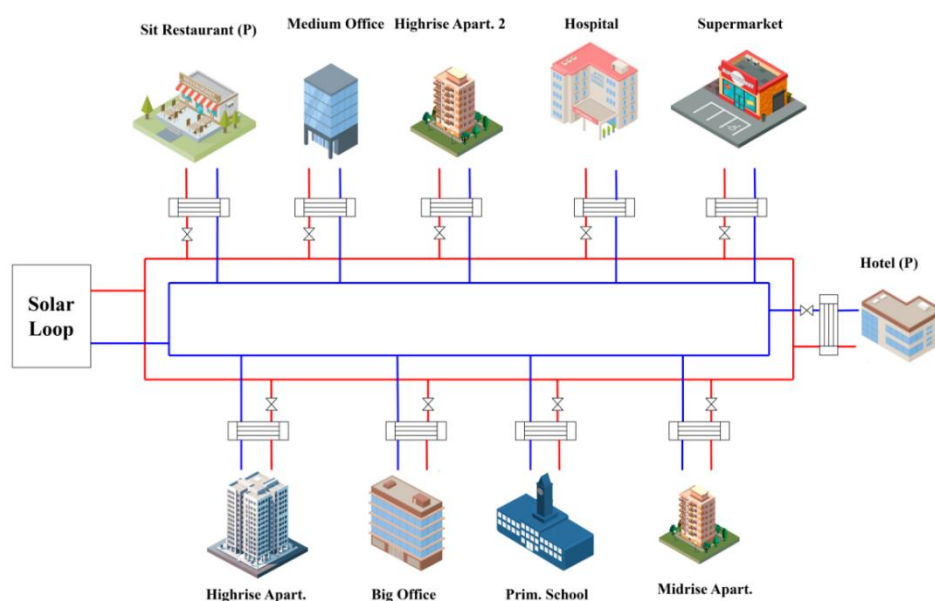


Fig. 47. Representation of buildings cluster of the case study.

To determine the hourly power profiles for heating and cooling of each of these users the demands of the big office, hotel and hospital were estimated previously in another research work [28]. For the other groups of buildings, the demand profiles were obtained from the simulation carried out on Energy Plus [43] using the building archetypes available in [29]. Precisely, the midrise apartments were simulated as in [21] while the remaining were deeply customized and revised for Palermo climate and building category following Italian legislation as shown in [44].

Table 15. Annual energy demand and maximum power of each user in the building cluster

User	Heating Peak [kW]	Cooling Peak [kW]	Heating Energy [MWh]	Cooling Energy [MWh]
Sit Restaurant (P)	0.00	157.44	0.00	200.88
Medium Office	194.93	191.15	54.63	91.21
Highrise Apart. 2	366.86	493.91	280.23	628.86
Hospital	513.67	1114.72	935.12	907.65
Supermarket	301.59	235.88	240.14	117.92
Hotel (P)	1190.51	0.00	590.16	0.00
Midrise Apart.	482.78	408.10	627.79	447.99
Prim. School	640.84	267.81	157.97	177.62
Big Office	474.76	656.24	320.28	196.01
HighRise Apart.	479.56	657.44	394.80	884.90
Total	2777.38	3200.79	3010.97	3452.18

In Table 15, the heating and cooling energy demands, and the maximum power demand of all users are reported. The "Total" row for total demand identifies the hourly sum of loads for all users, while the maximum, refers to the maximum of the sum of all users. As evident, the peak demand for heating and cooling of the hospital is significantly higher than that of all other users. Moreover, offices and the school have downtime periods (for example, weekends or the summer period for schools), while the other users exchange heat and cool with the network during the entire year. Finally, Fig. 48 shows the hourly trend of the entire cluster of buildings.

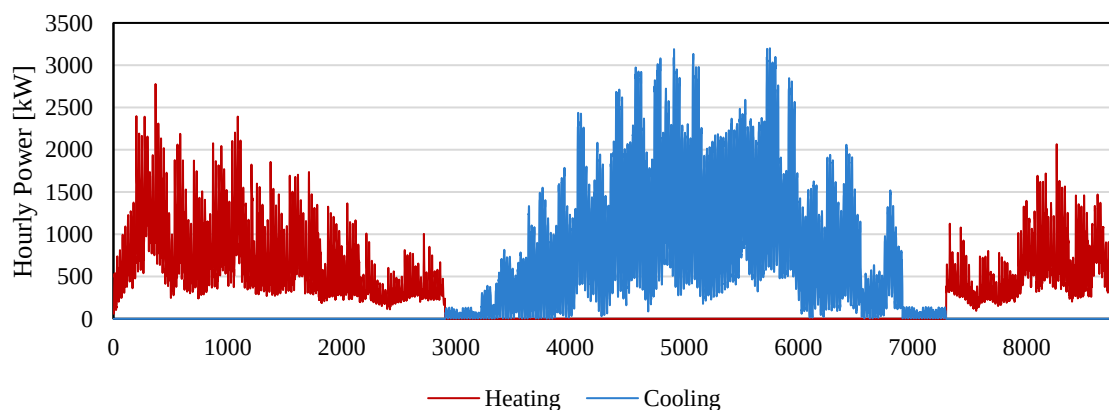


Fig. 48. Hourly Heating and Cooling demands of the entire building cluster

4.4.2.2 Design of Network Components

After determining the heating and cooling needs of the users, the components of the solar loop and the DHN were sized. The solar collector field is tasked with generating hot water year-round, which is stored in a seasonal storage unit. To appropriately size the field, guidelines for solar DH suggest, as per [45], that 1.4-2.4 m² of flat-plate collector surface per MWh of annual heat demand is optimal for central solar heating plants with seasonal storage. Based on these guidelines and the calculated thermal demand of the building cluster, which includes the absorption chiller's cooling demand, a solar collector field with a surface area of 12,000 m² was determined to be adequate. Furthermore, the guidelines specify that the thermal storage should have a volume of 1.4-2.1 m³ per m² of flat-plate collector surface. Therefore, given the thermal demand of the users and the selected collector surface area, a PTES of 20,000 m³ was deemed suitable for this case study. To ensure realistic heat loss behaviour from the storage, insulation with a 1-meter thickness of Extruded Polystyrene was chosen, in alignment with [40].

The design of the network was chosen by steady-state simulations conducted in EES, employing a methodology similar to that presented previously by the authors in [21]. A conservative approach was adopted by assessing the scenario where all users simultaneously demand their peak load, consequently, given the network's composition of 10 users, a simultaneity factor of 0.95 was applied to account for concurrent usage. The analysis prioritized the cooling demand overheating due to its predominance in determining the worst-case scenario. In the context of the fourth-generation network, where each user is equipped with an absorber operating with supply and evaporator temperatures set at 80°C and 8°C respectively, a constant COP of 0.55 for the absorber was assumed, based on the performance curve outlined in [46]. The temperature differential between the supply and return sides was set at 20°C for the fourth-generation DHN and 5°C for the fifth-generation network to calculate the flow rates. With a uniform distance of 200 meters between users and consistent branch diameters, the findings indicated that DN 200 pipes were necessary to maintain a specific pressure drop between 150 Pa/m and 250 Pa/m for the fourth-generation network, whereas DN 300 pipes were required for the fifth-generation network.

4.4.2.3 Case Study Scenarios

To assess the performance of a fifth-generation DHN, different energy production scenarios required by the connected users were hypothesized. Specifically, the fifth-generation network was compared with a scenario where each user meets their demand independently, and with an example of a fourth-generation DHN.

In Scenario 0, the base case, there is no district. Users are not interconnected, and energy

production is not centralized but carried out on-site. For each building, a natural gas-fed boiler for heating in winter and an air-to-water chiller for cooling in summer were hypothesized. The energy generated by the fuel powering the boiler in winter is calculated and considered a fixed efficiency of 0.95. Regarding the air-to-water chiller, the EER calculation referred to data provided in datasheets [41], showing how the performance coefficient varies with the evaporator outlet fluid temperature and the condenser temperature, which depends on the ambient temperature.

Scenario 1 represents a fourth-generation network, powered by a solar collector field and seasonal storage with supply temperatures ranging between 60-80°C, so it is not necessary for a temperature boost at the user substation making feasible the use of a plate exchanger. Since the solar collector field cannot provide an adequate supply temperature at all times of the year, a natural gas boiler was modelled. To satisfy winter heating needs, water heated to 60°C circulates in the supply network, while for cooling, absorption chillers at user substations use water at 80°C.

Scenario 2 delves into the performance of the fifth-generation network previously outlined, which integrates a solar thermal plant and an auxiliary chiller. To evaluate its response across different operational states, Scenario 2a presents a situation where water-to-water HPs and chillers interact with the supply side throughout the year, operating at 25°C in winter and 20°C in summer to enhance HPs efficiency. Scenario 2b modifies this arrangement by directly supplying users with chilled water at 8°C, aiming to assess the advantages and drawbacks of direct cooling provision. Scenario 2c introduces two prosumers into the system: a restaurant seeking to mitigate latent heat during winter and a hotel requiring heating for DHW production in summer. This scenario mirrors the operational strategy of 2a, as a supply temperature of 8°C poses challenges for hot water production during the summer months.

4.4.3 Results and Findings

In this chapter, the dynamic simulation results for the various considered network scenarios will be presented and examined. In particular, a detailed analysis of the results for the "5GDH Scenario" will be conducted, as this represents the baseline model on which subsequent modifications were made to develop the other scenarios.

4.4.3.1 Results for "Scenario 0"

In the outlined scenario, as previously mentioned, users operate independently without a connection to a DHN and are treated as separate entities. This case serves as a baseline for comparisons with potential enhancements through integration into either fourth or fifth-

generation DHNs. For each user, estimates of PE consumption and CO₂ emissions were conducted to provide a comprehensive analysis of this scenario using the same methodology previously presented in [21]. The energy sources utilized by the users include natural gas for boilers that generate heat for space heating and DHW, along with electricity powering air-to-water HPs for cooling purposes.

Fig. 49 illustrates the annual PE consumption and CO₂ emissions per user within a cluster of buildings. Notably, the Hospital emerges as the largest energy consumer, with 1312 MWh, followed by the Midrise Apartments at 895 MWh, representing 27.6% and 18.9% of the total consumption, respectively. Analysing the CO₂ emissions per MWh of PE reveals a nearly uniform ratio across all buildings, indicating comparable electric-to-gas energy consumption ratios. However, the high-rise apartments exhibit the lowest emissions ratio at 0.171 t/MWh, whereas the supermarket records the highest at 0.185 t/MWh, indicating variations in energy efficiency among the buildings.

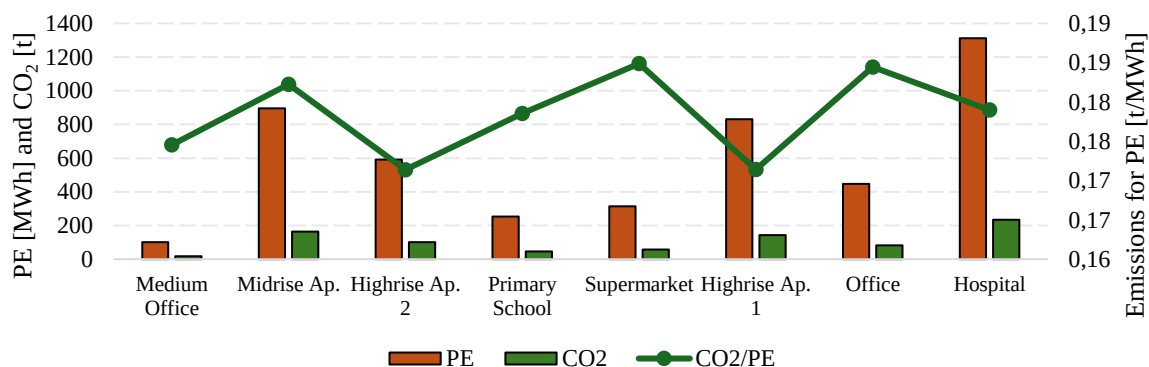


Fig. 49. PE and CO₂ Emissions by User in Scenario 0.

Table 16. Overall Results for Scenario 0.

		Boilers	A-W Chillers	Total
PE	[MWh/y]	3224.63	1520.97	4745.61
CO ₂	[t/y]	619.13	226.12	845.25

In Table 16, the base case results are summarized. In this, it can be noticed that, despite cooling energy being more prevalent than heating in Palermo, the PE used for cooling (provided by air-to-water chillers) is lower than that from boilers. This is because the system's overall EER averages at 4.34, allowing for reduced PE consumption and emissions associated with cooling energy production.

4.4.3.2 Results for "Scenario 1"

This scenario introduces a fourth-generation DHN powered by evacuated solar collectors, supplemented by a natural gas boiler. The network provides heating throughout the year.

Specifically, during the heating season, the supply temperature is maintained at 60°C to minimize heat losses, whereas, in the summer, it operates at 80°C to meet the minimum temperature requirement of the absorption chiller, which produces chilled water at 8°C. When solar collectors cannot meet the network's heating demand, the gas boiler compensates to maintain the necessary supply temperature. Notably, the designed solar collector system and PTES ensure that auxiliary heating from the gas boiler is mainly required during the summer, thereby achieving a reduction in both PE consumption and CO₂ emissions.

Table 17. Overall Results for Scenario 1.

		Pumps	Boiler	Total
PE	[MWh/y]	11.55	2204.23	2215.79
CO₂	[t/y]	1.72	423.21	424.93

Table 17 presents a comprehensive summary of the outcomes from this scenario, highlighting that the energy required for pumping constitutes a minor drawback of this DHN, contributing merely 0.52% to the total PE usage and 0.4% to overall emissions. Despite the inherent heat losses from the network's pipes and storage, these losses represent only 6.4% of the thermal energy generated by the boiler. This efficiency is attributed to the lower supply temperatures relative to those in second and third-generation networks. Consequently, the implementation of this network has led to a reduction in PE consumption exceeding 2.5 GWh and the prevention of 420 tons of CO₂ emissions, achieving an energy saving of 53%.

An in-depth hydraulic analysis revealed that selecting DN200 pipes, to attain a specified pressure loss in the most challenging branch of between 150 and 250 Pa/m, results in pump electrical consumption on the supply and return sides of 3.46 MWh and 2.47 MWh, respectively. In Fig. 50, which describes the pressure trend between each side of the pump for the supply and return loop, it is evident how in summer, the pressure drops are higher than in winter. This happens since in Palermo demands for cooling are higher than heating needs leading to an increase in flow rates in the district.

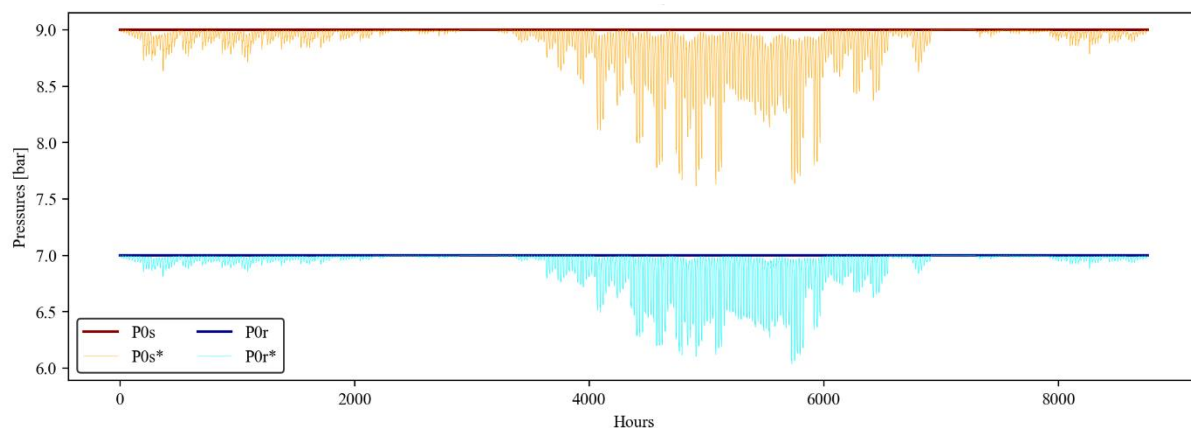


Fig. 50. Yearly pressures trend in Scenario 1

4.4.3.3 Results for "Scenario 2a"

"Scenario 2" describes a fifth-generation network connected to a cluster of eight users. In this scenario, consumption relates solely to the electrical energy used to power the HPs in each user substation and to power the auxiliary chiller which supports the absorption one. In this scenario, the supply temperature of the district is 25°C in winter and 20°C in summer to enable water-to-water HPs and Chillers installed in each user substation to work at maximum efficiency. Data on electrical energy consumption were extracted from TRNSYS simulations and reported in PE and CO₂ emissions using the respective conversion factors and then summarized in Table 18. This scenario performs better than the first two cases, leading to only 2057 MWh of PE consumption and 305.8 tons of CO₂ emissions. In particular, compared to Scenario 1, where a fourth-generation network was considered, the reduction in PE amounts to only 7.2% while in emissions this percentage rises to 28% due to the use of electricity as a unique input vector and avoiding consequently burning natural gas for feeding boiler. Additionally, calculating the SCOP_s of the system as the ratio between the total energy required, in absolute, and the total electric energy consumed, this is equal to 6.13 while SCOP and SEER for the users are respectively 5.33 and 9.62.

Table 18. Overall Results for Scenario 2a.

		Pumps	Users' HPs & Chillers	Auxiliary Chiller	Total
PE	[MWh/y]	35.83	1790.97	207.53	2034.33
CO₂	[t/y]	5.33	266.26	30.85	302.44

Focusing on the pressure trend, which is depicted in Fig. 51, the energy needed for pumping is 3.1 times the one of Scenario 1. In fact, despite the pressure trends being comparable, thanks to the choice of DN300 pipes for the 5GDH network, the mass flow rates are about 5 times the one in the 4GDH network, since the difference in temperature between the supply and return side of the network must be reduced due to the operative limits of the heat exchanger in the HPs and chillers.

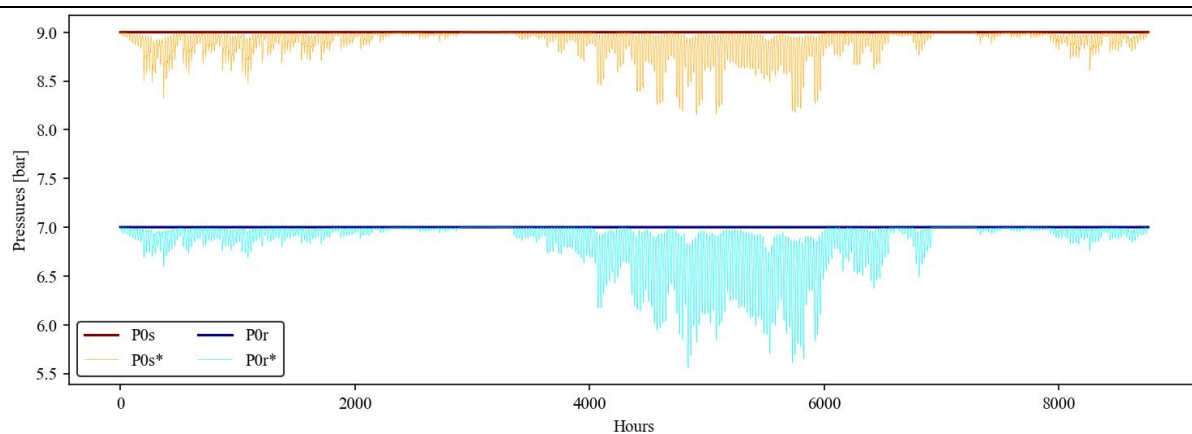


Fig. 51. Yearly pressures trend in Scenario 2a

In Fig. 52, a comparison between heat losses in fourth and fifth-generation network scenarios (1 and 2) demonstrates the main benefits of the fifth one. The heat losses in Scenario 2 are 46% of Scenario 1 while, considering only heat losses in the pipes, this value reduces to 24%. For example, despite heat losses in the PTES being comparable, the heat losses for the entire supply loop, during the year, are never above 11 kW while for Scenario 1 they almost reach 40 kW during summer.

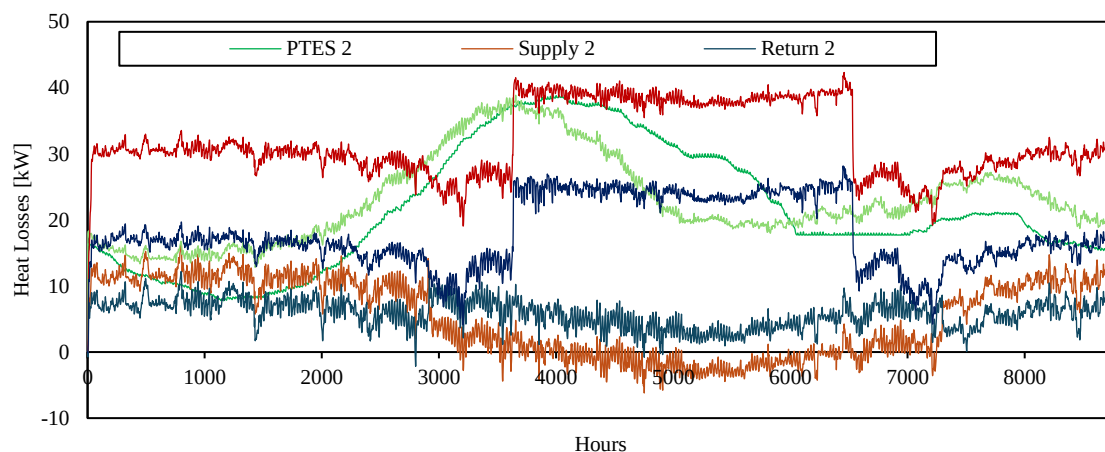


Fig. 52. Heat losses trend in Scenario 1 and 2a

4.4.3.4 Results for "Scenario 2b"

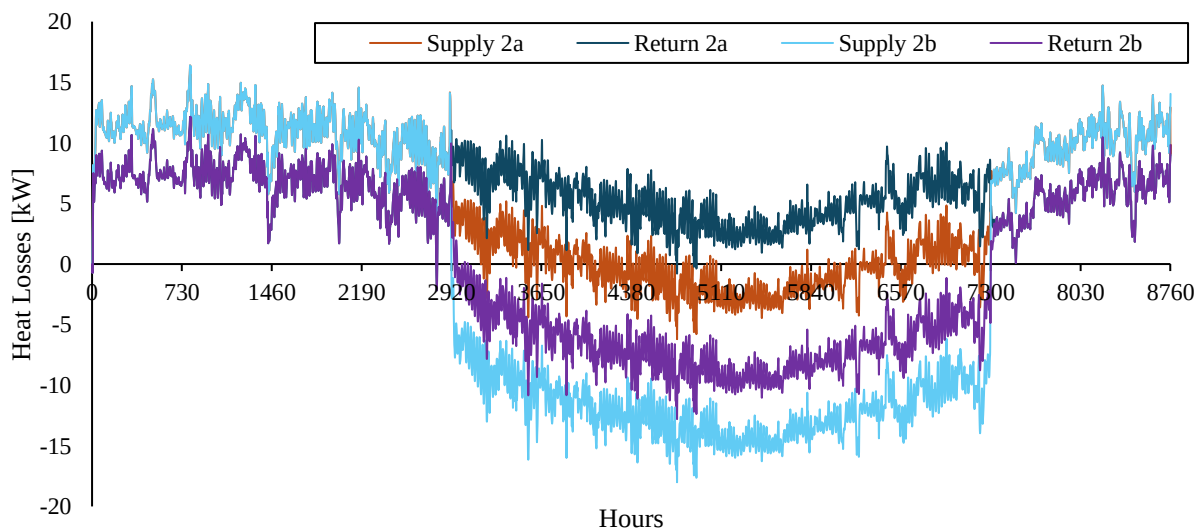
In this analysis, the focus shifts to a setup identical to Scenario 2a, with the notable adjustment that, during summer, an auxiliary chiller and an absorption chiller are employed to produce chilled water at 8°C. This is then distributed to users via a heat exchanger, circumventing the necessity for localized water-to-water chillers while winter operation mirrors Scenario 2a. However, due to the reduced evaporator temperature, both chillers operate at diminished COP and EER, with the absorption chiller's minimum operational temperature requirement increasing from 60°C to 80°C.

Table 19. Overall Results for Scenario 2b.

		Pumps	Users HPs & Chillers	Auxiliary Chiller	Total
PE	[MWh/y]	35.83	1091.92	853.04	1980.78
CO₂	[t/y]	5.33	162.33	162.33	294.45

Despite these operational constraints, as evidenced in Table 19, this scenario outperforms both the baseline case and Scenario 2a in efficiency. This superior performance is attributable to the centralized chiller approach since in Scenario 2a it needs also to remove the heat generated during compression. A comparative analysis reveals that Scenario 2b yields a PE savings of 54 MWh and nearly 8 tons of CO₂ emissions reduction compared to Scenario 2a. Table 19 also indicates that pump energy consumption remains consistent with Scenario 2a due to identical flow management within the district and the absence of prosumer involvement at this stage.

Heat loss analysis between Scenarios 2a and 2b, as shown in Fig. 53, indicates uniform losses during winter due to similar operations. However, in summer, Scenario 2b experiences greater absolute heat losses due to its lower operational temperature. Across the entire year, operating at 8°C increases heat losses by a factor of 1.48, which further escalates to 2.64 when focusing solely on cooling operations.

**Fig. 53. Pipes heat losses trend in Scenario 1 and 2a.**

4.4.3.5 Results for "Scenario 2c"

The operational dynamics of the DHN, characterized by near-ambient temperatures in both the supply and return loops, facilitate the integration of prosumers equipped with HPs or chillers to increase network efficiency. During the heating season, for instance, a restaurant requiring cooling to lower relative humidity draws water from the return loop, heats it, and then

reintroduces it into the supply loop. Conversely, during the cooling season, a hotel needing heat for DHW production cools the water, transferring it from the return to the supply loop. Given that a high-temperature HP required a higher supply temperature than the 8°C available to operate effectively, Scenario 2c mirrors the structure of Scenario 2a but incorporates two prosumers, adjusting the supply temperature to 25°C for heating and 20°C for cooling periods.

Table 20. Overall Results for Scenario 2c.

		Pumps	Users HPs & Chillers	Auxiliary Chiller	Total
PE	[MWh/y]	38.78	1782.28	0.00	1821.06
CO₂	[t/y]	5.77	264.97	0.00	270.73

Table 20 aggregated data confirms that prosumer involvement significantly elevates this scenario's efficiency above others. Notably, the auxiliary chiller's contribution to total PE consumption, significant in Scenarios 2a and 2b, is equal to zero in this configuration. Moreover, this configuration enables potentially reducing the solar collector field's area from 12,000 m² to 11,000 m² diminishing both the capital expenditure required for additional collectors and maintenance. Comparative analysis showcased in Fig. 40 illustrates that prosumer integration results in an 8.1% and 10.5% reduction in emissions and PE usage, respectively, compared to Scenarios 2b and 2a. When contrasted with the emissions and energy consumption figures from the fourth-generation network depicted in Scenario 1, this scenario achieves markedly lower values, with reductions of 1.57 times in emissions and 1.22 times in energy consumption.

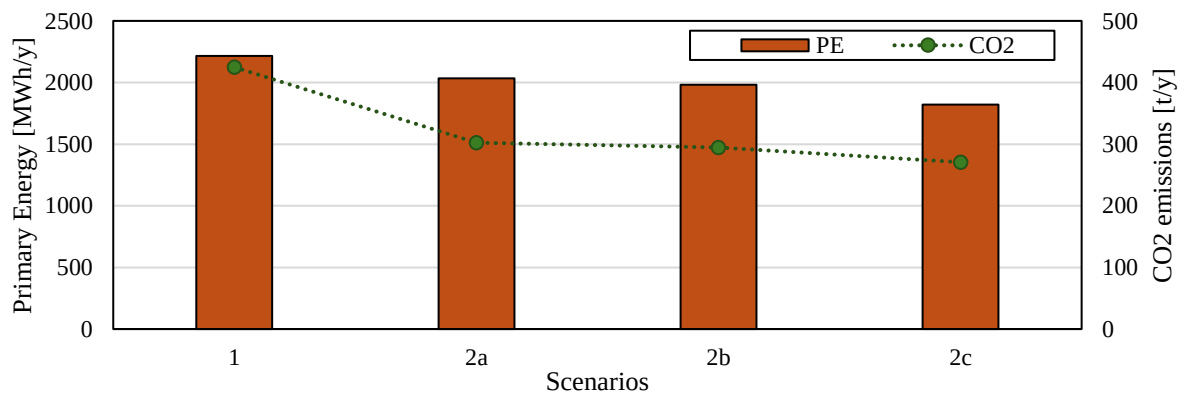


Fig. 54. Comparison between improvement scenarios

Furthermore, while heat losses align with those observed in Scenario 2a, the impact of prosumers on the average temperature of the PTES needs to be examined (Fig. 55). The inclusion of prosumers ensures that the PTES temperature remains above 60°C during the entire year, thereby negating the need for the auxiliary chiller to activate. This observation elucidates the heightened energy consumption associated with the auxiliary chiller in Scenario 2b since

the PTES temperature reaches the 80°C threshold necessary for the absorption chiller to generate chilled water at 8°C, almost 1000 hours before Scenario 2a.

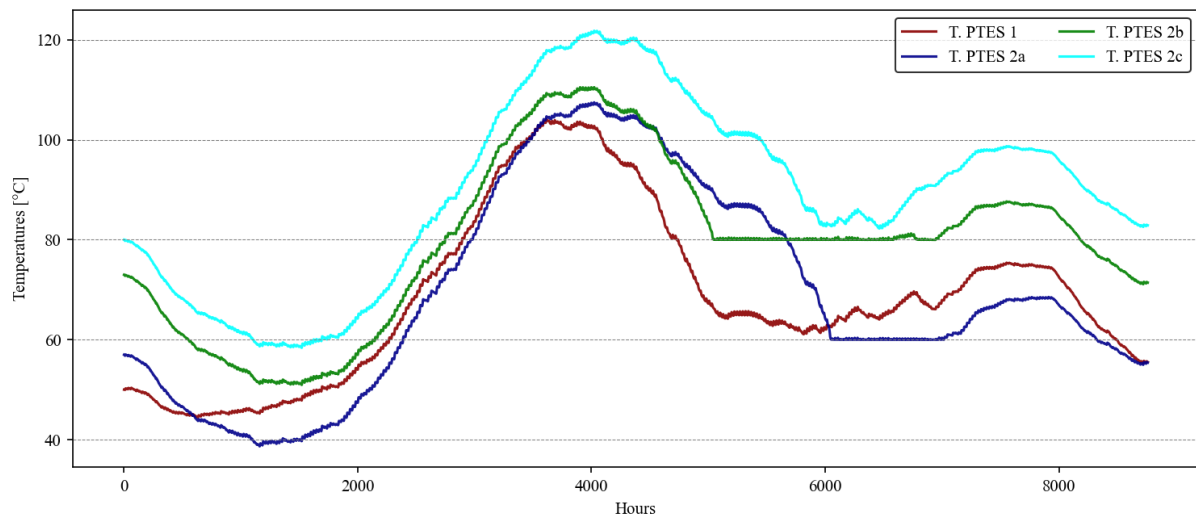


Fig. 55. PTES average temperature in different scenarios

Although Scenario 2c emerges as the most efficient, achieving a system COP of 7.07, it also incurs the highest energy consumption due to pumping. The prosumer contribution positively influences the supply side, as illustrated in Fig. 56, where the pressure immediately before the supply pump in Scenario 2c ($P_{0s} 2c^*$) consistently falls below that of Scenario 2a throughout most of the year. However, this benefit is counterbalanced by increased pressure losses in the return loop, attributed to higher flow rates and resulting in pumping energy requirements nearly 3 MWh greater than in Scenario 2a.

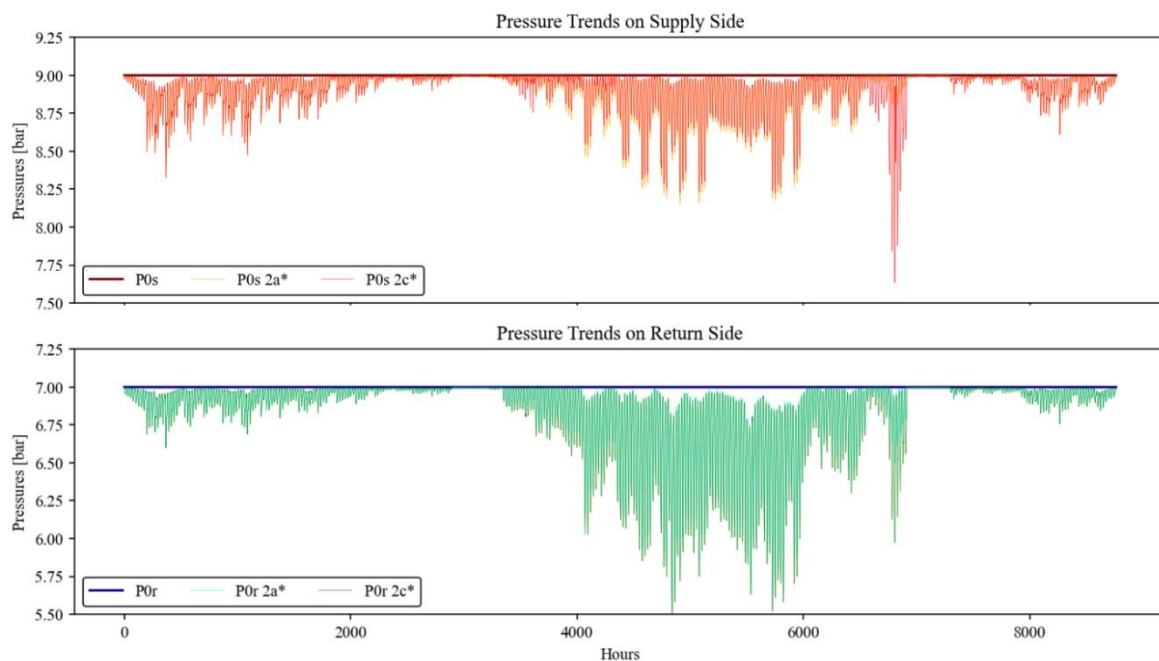


Fig. 56. Comparison of pressure losses between Scenario 2a and Scenario 2b

Furthermore, summer periods exhibit elevated pressure losses due to increased energy

demand and corresponding flow rates within the district, as highlighted in Fig. 57 which represents weekly winter and summer pressure trends. Specifically, pressure drops are minimized during nighttime when offices or schools are closed, whereas daytime activities escalate pressure differences, manifesting a discernible disparity between supply and return loop pressures in Scenarios 2c and 2a in July. On the supply side, the pressure is more equally distributed during prosumer injection leading to a reduction in pressure drops in Scenario 2c which becomes consistent when the demand is higher. On the return side, since more flow rate is needed to be pushed from the return side to the supply by the prosumer, this pressure drop is higher.

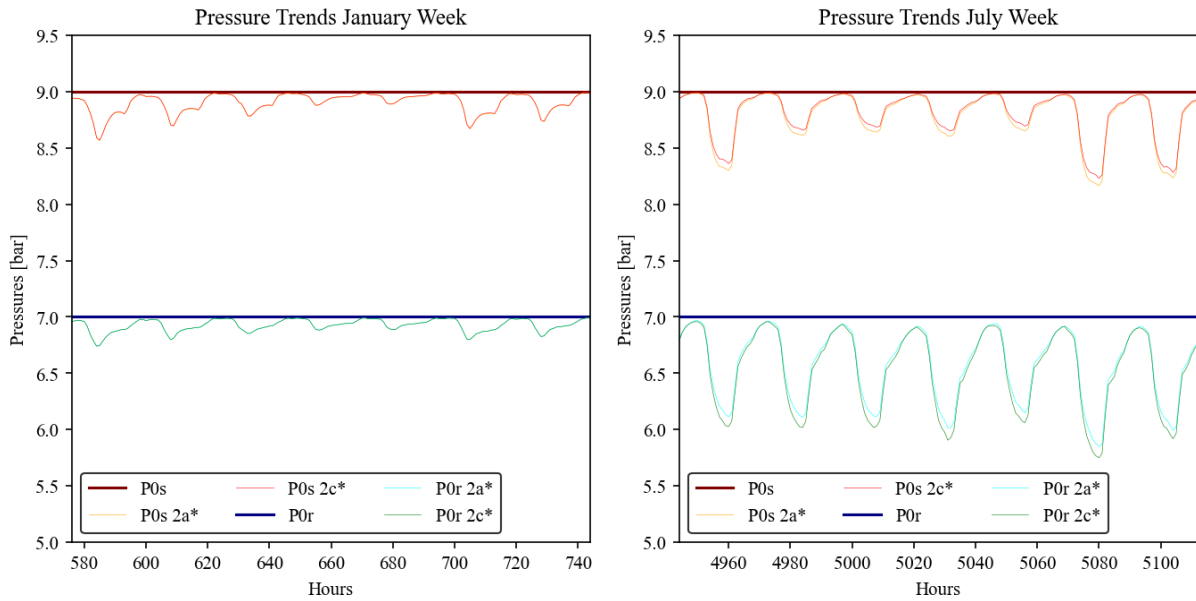


Fig. 57. Comparison of Pressure Trend in the last week of January and July

4.5 Final Discussion and Future Directions

The widespread adoption of efficient, low-carbon heating technologies has become a priority in the context of transitioning to fully decarbonized urban cooling. DHC networks are one of the essential elements for making this transition to more sustainable energy sources for building cooling. However, the DH sector continues to be dominated by the use of fossil fuels, requiring a significant change in managing this technology and extensive research into new strategies. Today, the latest technologies are being developed to optimize resource use, allowing the integration of RES and the valorisation of waste heat from industrial processes. Additionally, the latest generations of DHNs aim to further reduce heat losses by reducing operating temperatures and improving pipe insulation. In this framework, fifth-generation DHC networks are considered an opportunity to increase the efficiency of DHNs due to near-ambient operation temperature which allows the integration of low-temperature heat sources and seasonal TES.

The DT developed for Rio Vena DHN in the context of DigiBUILD project demonstrated the benefits of having a DT for optimal management. Although the system is a third-generation network that relies exclusively on natural gas, the application of this smart technology can improve efficiency, reduce emissions, and lower costs. In addition to monitoring the district's performance, the DT includes KPIs for assessing data quality and models for forecasting thermal energy consumption in the district. Furthermore, as a central aim of this thesis, this DT incorporates an optimization algorithm that utilizes ML models to estimate boiler efficiency. Applications and tests on historical data revealed that the GBoost algorithm was the best-performing model for predicting efficiency, achieving a coefficient of determination greater than 0.97 and an MSE of 0.001. Moreover, NSGA-II optimization run on data from 2023 demonstrated that energy savings of over 30% could be achieved, depending on the day, compared to unoptimized operation. During the development of this DT, some challenges were identified. In the Rio Vena DHN, not all parameters were monitored with consistent frequency and methods, causing inaccuracies that, combined with sensor issues, impacted the performance of the first ML models. Furthermore, operational limitations within the network restricted the optimization process to a single controllable variable, which was used as the decision variable. Regarding the integration, balancing Veolia's preference for a system closely resembling their existing setup with the need to develop a functional DT dashboard was difficult. However, a key advantage of this methodology is its applicability to most third-generation DHNs, which are currently the most prevalent type of network. By identifying correlations between operational parameters, such as boiler setpoint temperatures, and the energy demand of the district, it is possible to maximize heat production efficiency across similar networks.

The analytical model for the Fourth Generation Network presented in this chapter may be utilised to conduct analysis in both dynamic and steady-state scenarios, taking into account the inclusion of prosumers. Additionally, it can be employed for the purpose of designing the DHN. Following several references to the analytical model, the case study is introduced to examine the thermohydraulic factors of the system while it is in operation. The DHN in TRNSYS was modelled by selecting essential components and equations for pressure analysis. Three distinct situations have been simulated. The Base Case refers to a traditional 3rd generation DHN that does not include prosumers. This case has been contrasted to two scenarios that include incremental enhancements. In Scenario 1, the supply temperature of the network was reduced from 100 °C to 80 °C and subsequently to 60 °C. Additionally, a field of evacuated tube solar collectors was incorporated with a CHP. When looking specifically at the heating period, reducing the temperatures has led to a reduction in thermal losses by 26.4% and 52.8% for the supply temperatures of 80°C and 60°C, respectively. Furthermore, the utilisation of solar

collectors resulted in a reduction of heating energy consumption by more than 25% in this particular research. In the second scenario, the supply temperature for the heating period was set at 60 °C and 80 °C. Additionally, two simple prosumers were considered. The prosumers will generate heat using HPs due to the reduced temperatures of the DHN. This system incorporates two distinct energy-saving measures. Firstly, it achieves a decrease in thermal energy demand through the use of CHP. Secondly, it optimises the distribution of flow rate in the network, resulting in lower power consumption by the pumps. When analysing the heating period, the findings indicated that this approach resulted in a 27% reduction in energy consumption for pumping and a 31.3% decrease in the thermal energy required for CHP. To ensure a more precise study, we took into account the CO₂ emissions of the system. This revealed that the overall emissions associated with Scenario 2 are just 68.6% of those in the Base Case. The created model has demonstrated effective performance under various scenarios. The hourly findings provided insights into the network's performance throughout the day and might aid in selecting appropriate management strategies. The same model may be utilised in many circumstances by simply providing inputs such as climatic data, pipe data, and load description. However, replicating this implementation in a DT for a real system poses challenges. Each user requires customized TRNSYS types, making the process complex for DHNs with numerous users. Even if a real system can be replicated using this approach, the concept of a self-evolving model still needs to be fully implemented. While feasible with the latest version of TRNSYS, which allows coupling with Python, this approach requires further investigation. Once achieved, a DT developed with this methodology would provide DHN managers with a powerful tool for testing various operational strategies. This capability could significantly enhance network management, leading to reductions in costs, emissions, and energy consumption.

Finally, this research aimed also to develop and analyse a model of a 5GDH network capable of producing both heating and cooling energy, which is connected to a thermal plant composed of solar collectors, an absorption chiller, and seasonal thermal storage. For each user, dynamic models of HPs and water-water chillers were created to facilitate interaction with the district. On the network side, to guarantee the efficient functioning of the HPs, the operation was set to a maximum of 25°C for heating and a minimum of 20°C in summer for cooling, or 8°C for direct cooling. Accordingly, the design of network components and the thermal plant was tailored to the requirements of the connected building cluster, encompassing the sizing of pipes, thermal collectors, and STES. The model, developed using TRNSYS software, was tested across various scenarios, including one representing a fourth-generation network and three others depicting different fifth-generation configurations, all compared against a baseline

scenario without any DHN. Through the analysis comparing the baseline case with the fifth-generation scenarios, respectively Scenario 2a and Scenario 2b, a significant saving ranging from 2711 to 2924 MWh in PE and from 64% to 68% reduction in CO₂ emissions were identified. When comparing Scenario 1, representing the 4GDH network, with Scenario 2b, which is a variant of the 5GDH employing direct cooling, the savings amount to 10% in PE and 30% in CO₂ emissions. The lower operational temperatures in the 5GDH led to thermal losses being reduced to 76% of those experienced in the 4GDH setup. Moreover, Scenario 2c demonstrates the advantages of incorporating prosumers into the network, achieving an improved equilibrium between heating and cooling demands which resulted in an additional reduction of 160 MWh in PE and 24 tons of CO₂ compared to Scenario 2b. In terms of hydraulic performance, the 5GDH scenarios, specifically 2a and 2c, exhibited a PE consumption that was 3.1 and 3.35 times higher, respectively, than that of the 4GDH network. Nonetheless, in the considered case study, the energy required for pumping constituted only 2.1% of the total energy use in the most stressed scenario. One major limitation of this approach is its associated costs. Emission reductions and operational savings must be maximized to justify both the environmental and economic investments required. Additionally, applying this methodology to districts with numerous users presents challenges for model development. In practical implementation, this model could serve as a foundation for a DT in both the design and operational phases. Once validated, the model could enable real-time synchronization between actual and simulated measurements under identical conditions, facilitating fault detection and identifying areas for improvement.

Climate changes and international conflicts highlight the urgency of finding new alternatives to increase the resilience of the European energy system and reduce dependence on fossil fuels. The models developed in this thesis have demonstrated how the use of a last-generation network powered by RES brings considerable advantages in terms of resource savings and a significant reduction in CO₂ emissions, contributing to the transition towards the decarbonization of the urban sector. However, prosumer injection must still be investigated since, depending on their production rate or position, it could stress the network. For this reason, it is evident that their effect on different case studies and typologies of networks still needs to be investigated. It is, therefore, necessary to find new methodologies and models for the production and distribution of thermal energy for cooling capable of accelerating decarbonization in line with Europe's objectives for 2050.

References

- [1] European Commission, “Building and renovating: The European Green Deal,” no. December, pp. 1–10, 2019, doi: 10.2775/48978.
- [2] IEA, “Heating.” [Online]. Available: <https://www.iea.org/reports/heating>
- [3] IEA, “How can district heating help decarbonise the heat sector by 2024?” [Online]. Available: <https://www.iea.org/articles/how-can-district-heating-help-decarbonise-the-heat-sector-by-2024>
- [4] Gestore dei Servizi Energetici, “Teleriscaldamento e Teleraffrescamento 2019. Nota di approfondimento GSE,” 2020.
- [5] W. Weiss and M. Spörk-Dür, “Solar Heat Worldwide,” 2022.
- [6] H. Lund *et al.*, “Perspectives on fourth and fifth generation district heating,” *Energy*, vol. 227, p. 120520, 2021, doi: 10.1016/j.energy.2021.120520.
- [7] P. Sorknæs *et al.*, “The benefits of 4th generation district heating in a 100% renewable energy system,” *Energy*, vol. 213, Dec. 2020, doi: 10.1016/j.energy.2020.119030.
- [8] O. Todorov, K. Alanne, M. Virtanen, and R. Kosonen, “Aquifer thermal energy storage (ATES) for district heating and cooling: A novel modeling approach applied in a case study of a Finnish urban district,” *Energies (Basel)*, vol. 13, no. 10, May 2020, doi: 10.3390/en13102478.
- [9] D. I. Stanica, M. Bachmann, and M. Kriegel, “Design and performance of a multi-level cascading district heating network with multiple prosumers and energy storage,” *Energy Reports*, vol. 7, pp. 128–139, Oct. 2021, doi: 10.1016/j.egyr.2021.08.163.
- [10] D. Fischer and H. Madani, “On heat pumps in smart grids: A review,” *Renewable and Sustainable Energy Reviews*, vol. 70, no. April, pp. 342–357, 2017, doi: 10.1016/j.rser.2016.11.182.
- [11] “DigiBUILD Project - High-Quality Data-Driven Services for a Digital Built Environment towards a Climate-Neutral Building Stock.” Accessed: Jun. 27, 2024. [Online]. Available: <https://digibuild-project.eu/>
- [12] International Energy Agency, “Digitalisation,” 2023. Accessed: Jul. 23, 2023. [Online]. Available: <https://www.iea.org/energy-system/decarbonisation-enablers/digitalisation>
- [13] European Commission, “Digitalising the energy system - EU action plan,” 2022. doi: 10.2833/492070.
- [14] F. Pedregosa *et al.*, “Scikit-learn: Machine learning in Python,” *Journal of machine learning research*, vol. 12, no. Oct, pp. 2825–2830, 2011.
- [15] J. Blank and K. Deb, “Pymoo: Multi-Objective Optimization in Python,” *IEEE Access*, vol. 8, pp. 89497–89509, 2020, doi: 10.1109/ACCESS.2020.2990567.

-
- [16] “Climate Data - Burgos.” Accessed: Sep. 23, 2024. [Online]. Available: <https://en.climate-data.org/europe/spain/castile-and-leon/burgos-1767/>
 - [17] Google LLC, “Google Earth 7.3.6,” 2023. Accessed: Sep. 23, 2024. [Online]. Available: <https://www.google.it/intl/it/earth/index.html>
 - [18] University of Wisconsin--Madison: Solar Energy Laboratory, “TRNSYS 17, a transient simulation program,” 1975, *Madison, Wis. : The Laboratory, 1975*. Accessed: Oct. 05, 2023. [Online]. Available: <https://www.trnsys.com/>
 - [19] Grafana Labs, “Grafana: The open observability platform.” Accessed: Aug. 09, 2024. [Online]. Available: <https://grafana.com/>
 - [20] Meta Open Source, “REACT - The library for web and native user interfaces.” Accessed: Aug. 09, 2024. [Online]. Available: <https://react.dev/>
 - [21] T. Testasecca, P. Catrini, M. Beccali, and A. Piacentino, “Dynamic simulation of a 4th Generation District Heating network with the presence of prosumers,” 2023.
 - [22] S.A. Klein *et al.*, *TRNSYS 17 - Mathematical Reference*. 2012. [Online]. Available: <http://sel.me.wisc.edu/trnsys>
 - [23] Cengel Yunus and Cimbala John M., *Fluid Mechanics: Fundamentals and Applications*. 2014.
 - [24] T. Lickleder, T. Hamacher, M. Kramer, and V. S. Perić, “Thermohydraulic model of Smart Thermal Grids with bidirectional power flow between prosumers,” *Energy*, vol. 230, 2021, doi: 10.1016/j.energy.2021.120825.
 - [25] F. Calise, F. L. Cappiello, M. Dentice d’Accadia, and M. Vicidomini, “Smart grid energy district based on the integration of electric vehicles and combined heat and power generation,” *Energy Convers Manag*, vol. 234, Apr. 2021, doi: 10.1016/j.enconman.2021.113932.
 - [26] P. Catrini, T. Testasecca, A. Buscemi, and A. Piacentino, “Exergoeconomics as a Cost-Accounting Method in Thermal Grids with the Presence of Renewable Energy Producers,” *Sustainability (Switzerland)*, vol. 14, no. 7, Apr. 2022, doi: 10.3390/su14074004.
 - [27] *DPR 26 agosto 1993, n. 412 - REGOLAMENTO RECANTE NORME PER LA PROGETTAZIONE, L’INSTALLAZIONE, L’ESERCIZIO E LA MANUTENZIONE DEGLI IMPIANTI TERMICI DEGLI EDIFICI AIFINI DEL CONTENIMENTO DEI CONSUMI DI ENERGIA, IN ATTUAZIONE DELL’ART.4, COMMA 4, DELLA LEGGE 9 GENNAIO 1991, N. 10.*
 - [28] A. Piacentino and C. Barbaro, “A comprehensive tool for efficient design and operation of polygeneration-based energy μ grids serving a cluster of buildings. Part II: Analysis of the applicative potential,” *Appl Energy*, vol. 111, pp. 1222–1238, 2013, doi: 10.1016/j.apenergy.2012.11.079.
 - [29] U.S. Department of Energy, “Prototype Building Models.” Accessed: Sep. 26, 2023. [Online].
-

Available: <https://www.energycodes.gov/prototype-building-models>

- [30] ASHRAE and American National Standards Institute, “Climatic Data for Building Design Standards,” 2020. [Online]. Available: www.ashrae.org
- [31] EnergyPlus, “Weather Data by Location.” Accessed: Oct. 05, 2023. [Online]. Available: https://energyplus.net/weather-location/europe_wmo_region_6/ITA/ITA_Palermo.164050_IWEC
- [32] Task Force Customer Installations, “Guidelines for District Heating Substations Approved by the Euroheat & Power Board Prepared by Task Force Customer Installations EUROHEAT & POWER Guidelines for District Heating Substations October 2008,” *Euroheat and Power*, 2008.
- [33] H. I. Tol and S. Svendsen, “Improving the dimensioning of piping networks and network layouts in low-energy district heating systems connected to low-energy buildings: A case study in Roskilde, Denmark,” *Energy*, vol. 38, no. 1, pp. 276–290, 2012, doi: 10.1016/j.energy.2011.12.002.
- [34] Gerhard Oppermann, Othmar Arnold, Joachim Ködel, Marcel Büchler et al, and Martin Jutzeler et al, “Teleriscaldamento / teleraffreddamento,” 2020. Accessed: Sep. 24, 2023. [Online]. Available: <https://www.aelsi.ch/documenti/31.pdf>
- [35] M. Tunzi, M. Ruysschaert, S. Svendsen, and K. M. Smith, “Double loop network for combined heating and cooling in low heat density areas,” *Energies (Basel)*, vol. 13, no. 22, 2020, doi: 10.3390/en13226091.
- [36] F.-C. Software, “EES_manual.pdf,” pp. 210–212, 2011.
- [37] ISPRA, “Fattori di emissione atmosferica di gas a effetto serra nel settore elettrico nazionale e nei principali Paesi Europei,” 2019.
- [38] M. Steen, “Greenhouse gas emissions from fossil fuel fired power generation systems.” [Online]. Available: <http://www.jrc.nl>
- [39] International Energy Agency (IEA), “District Heating,” 2023. Accessed: Sep. 23, 2024. [Online]. Available: <https://www.iea.org/energy-system/buildings/district-heating>
- [40] E. Guelpa and V. Verda, “Thermal energy storage in district heating and cooling systems: A review,” *Appl Energy*, vol. 252, p. 113474, Oct. 2019, doi: 10.1016/j.apenergy.2019.113474.
- [41] “Clivet SCREWLine3 [Online]. Available: https://clivet.org/images/pdf/hava-khonak/Screw/WDAT-SL3/BT14D018GB-00_PRELIMINARY_01.pdf.”
- [42] D. Schmidt *et al.*, “Low Temperature District Heating for Future Energy Systems,” *Energy Procedia*, vol. 116, pp. 26–38, 2017, doi: 10.1016/j.egypro.2017.05.052.
- [43] “Energy Plus,” 9.5.0. [Online]. Available: <https://energyplus.net/>
- [44] G. Ala *et al.*, “A Workflow for Thermal and Electrical Co-Simulation of Energy Systems,” in

2023 Asia Meeting on Environment and Electrical Engineering (EEE-AM), IEEE, Nov. 2023, pp. 1–6. doi: 10.1109/EEE-AM58328.2023.10394689.

- [45] T. Schmidt, D. Mangold, and H. Müller-Steinhagen, “Central solar heating plants with seasonal storage in Germany,” *Solar Energy*, vol. 76, no. 1–3, pp. 165–174, Jan. 2004, doi: 10.1016/j.solener.2003.07.025.
- [46] Marco Cefarin, “Design of NH₃-H₂O absorption chiller for low grade waste heat recovery,” 2015. Accessed: Aug. 08, 2024. [Online]. Available: <https://hdl.handle.net/11390/1132442>

Chapter 5

Digital Twin of a Microgrid

The vehicle sector is a major source of GHG emissions, contributing to global warming and climate change. To reduce emissions, the EU is promoting the electrification of mobility, offering various incentives to buyers who substitute petrol cars with EVs. EVs have lower tailpipe emissions than conventional vehicles and can also reduce the dependence on fossil fuels. However, such environmental benefits would originate from EVs only considering that the sources of electricity used for charging were based on RES like solar and wind power. At the same time, if the generation is based on fossil fuel sources such as coal and natural gas, then EVs will still account for a certain number of emissions of GHGs based on the country's electricity mix.

Moreover, such electrification of mobility poses some challenges even from the side of the electric system, since the demand for electricity increases and it needs to be adequately provided with infrastructure for charging. At the same time, this electrification also leads to opportunities for grid balancing by integrating EVs with RES, since EVs can act as flexible loads or storage devices that can respond to the fluctuations in the supply and demand of electricity. For example, EVs can charge during periods of excess RES production (Grid to Vehicle) and discharge during a shortage (Vehicle to Grid). In this way, the variability and intermittency of RES will be smoothed out for better integration into the electric system.

5.1 Overview of Smart Grids Optimal Management through Electric Vehicles

In the fast-evolving field of EV technology, much interest has been accorded to developing sophisticated algorithms that can optimize EV charging. The need to integrate EVs into the existing power grids with minimal environmental impacts has driven such sudden interest in innovation. A notable example of this trend is [1], where the authors developed a model for optimising the charging schedules of Battery Electric Vehicles (BEVs). The model, adapted to consider the different environmental impacts of the generation mix, demonstrates the

substantial reduction in GHG emissions made possible by smart charging strategies. It considers other factors that pertain to time and location, thus accounting for the dynamic nature of environmental impacts. Furthermore, this work analyses the complexity in balancing multiple environmental impact categories to reach the minimum ecological footprint by BEVs. In a different approach, Bao et al. [2] presented a methodology of mixed fleet scheduling for airport ground service vehicles that takes into consideration both fuel-based and electric tractors in the fleet. It proposed a novel mixed integer model that minimizes the summation of costs concerning time, energy consumption, and emission, which is improved by an enhanced adaptive large neighbourhood search algorithm. Another critical contribution is presented by Yin et al., who introduced a methodology for the safe and efficient operation of electrical grids using the optimum scheduling strategy of EVs [3]. Their methodology included advanced ML models such as Long Short-Term Memory (LSTM) and XGBoost combined with the CPLEX solver to handle complex optimization problems.

By the use of optimization methods, it is possible to determine the optimal charging schedule for EVs to minimize costs while ensuring that they are fully charged when needed. These methods are even necessary for improving the use of RESs by matching the periods of high renewable generation with the charging of EVs. Calise et al. [4] proposed an approach to optimize the charging strategy of a fleet of EVs connected to a DHN served by a cogeneration plant. Once applied, the charging strategy led to a PE saving index of about 32%, which led to the effective reduction of carbon dioxide emissions in the analysed district. In [5], instead, the authors proposed an optimization method based on Multi-Agent Deep Reinforcement Learning. These simulations show that the grid energy consumption can be reduced by up to 40% with appropriate management of EVs and PVs when compared to multi-home systems which are not engaged in energy trading or EVs. Liu et al. have also introduced a study representing the benefits of a multi-objective optimal charging strategy for peak load shaving by comparing various optimization algorithms [6]. With their simulation results, they achieved an excellent reduction in the load fluctuation level and a decrease in the cost of electricity by considering constraints such as the balance of load and capacity limits of EVs and electrical storage systems. Advanced algorithms can be implemented in real-time, enabling management of all energy flows, such as EV charging, grid interaction, and heat pump operations, in real-time to reduce energy drawn from the grid. Examples are found in the work of Mathur et al. [7], and Uzair et al. [8], which presents a relevant example of real-time operational optimisation.

Considering the last IEA report [9], the contribution of the vehicle sector to global emissions exceeds 15%, thus the widespread use of EVs is essential for the goal of abating GHG emissions. This expansion necessitates dual-sided scientific research: one to lower emissions

during EV charging and another to investigate EVs' actual consumption. Zhao et al. [10] have proposed XGBoost and Light GBM-based approaches for the forecast of the remaining driving range based on an extensive EV database: NDANEV. Although these achieved very good accuracy, their models include a large number of features such as braking ratio, acceleration ratio, and motor and battery output energy that make them hard to replicate in simpler scenarios. On the other hand, H.A. Yavasoglu et al. [11] proposed, a range estimation method for EV using data gathered and experimentation. They used a Decision Tree to identify road type and ANN to predict range based on variables such as HVAC usage, vehicle weight, traffic, and road type. Their model resulted in an accuracy of more than 94%. Though models like this are very efficient, the replicability by non-experts or people with less sophisticated equipment is limited.

5.2 Digital Twin of Emotion Microgrid

The highlighted research emphasizes the increasing importance of ML and optimization algorithms in balancing EV charging demands with the stability of electrical grids. Nevertheless, there exists a gap concerning the development of these novel approaches for actual case studies, which requires further development tailored to actual user needs. Moreover, it is important to consider the role of the facility manager, who will be the main stakeholder in benefiting from and managing the vehicles. Since these optimization strategies are normally attached to vehicle usage, it is vital to account for individual vehicle consumption and patterns of use for better management.

This work aims to address these gaps by presenting a DT that integrates ML models and optimization algorithms capable of real-time response. The system recommends EV charging strategies based on user preferences and incorporates adaptive ML models tailored to each vehicle, ensuring continuous performance improvement. This methodology was applied in the framework of the DigiBUILD project [12] to the case of Emotion, an Italian company specializing in EV charging stations. A DT was developed to manage and optimize efficiently the energy consumption of one of its buildings, equipped with 50 kWp PV panels and four EV charging points of 22 kW each. Through the use of optimization algorithms, the DT will provide optimal charge strategies for the company fleet of ten EVs taking into account the forecasts of PV generation, building energy consumption, and user needs. Moreover, the DT could autonomously interact in real time with the charging points to trigger and modulate the charging of the connected EVs to decrease overall energy costs and leverage available PV power. This will contribute not only to energy efficiency but also to grid stability. Finally, the DT dashboard not only displays historical and real-time information but also represents a DSS suggesting manual actions, such as setting setpoint temperature for HVAC and suggesting to connect or

disconnect vehicles. The presented approach meets the most up-to-date focus in scientific literature data and adaptive control for the optimal use of energy in facilities hosting RESs and EV fleets [13].

5.2.1 Digital Twin Development Methodology

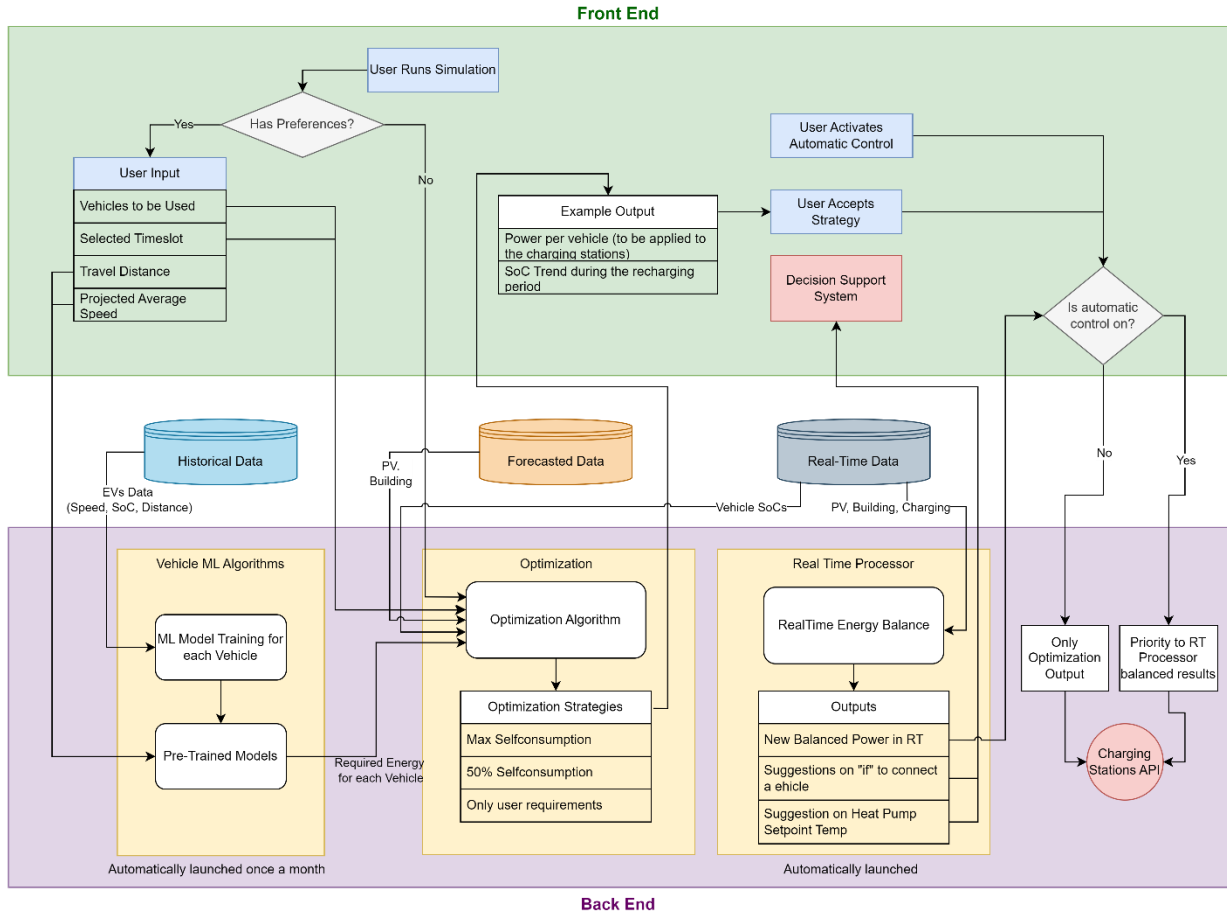


Fig. 58. DT Architecture of Emotion Digital Twin

The DT developed in the context of this thesis presents various models that interact with each other, as shown in Fig. 58. The core of this DT consists of four different types of algorithms:

- ML models, trained on individual vehicles, which are used to estimate EV specific consumption based on average trip velocity;
- An optimization algorithm designed to determine optimal strategies based on different self-consumption objectives and user preferences;
- A real-time processor employed to optimally control charging and provide recommendations;
- Predictive models developed by other project partners, which are used to estimate PV production and building consumption.

Regarding the ML models, an algorithm is executed monthly to update clustering models

based on historical vehicle data. Once updated, the user can access the DT dashboard and input preferences, such as vehicles required, time of need, distance to be travelled, and expected average speed. The last two pieces of information are used by the pre-trained models to calculate specific energy consumption and, consequently, the energy required by each selected vehicle.

For the optimization algorithm, the energy required by each vehicle is used to determine the minimum SoC necessary at the specified time. While user preferences are not strictly essential for the algorithm, essential inputs include the SoC of each vehicle at the start of the optimization process and forecasts of PV production and building consumption for the next 24 hours. Once this data is input, the optimization process begins, and it provides the user with three different strategies based on the level of self-consumption they wish to achieve.

Through the DT dashboard, the user can select one of the suggested strategies or activate the "auto-pilot" mode, which engages the real-time processor module. The output of each strategy specifies the optimal time to connect the EV for charging, as well as the power modulation behaviour for each charging port. The system then regulates the charging power at each port, ensuring the predicted balance between energy import and export. In the "auto-pilot" mode, if real-time data detects an imbalance in grid power import or export, the algorithm automatically adjusts the charging power or provides suggestions to address the issue, functioning as a DSS.

5.2.1.1 EV ML Modelling

The methodology followed for the development of ML models capable of estimating real vehicles consumption was based on the following process:

1. Expand the number of available variables by calculating new features such as velocity and specific consumption.
2. Conduct a data analysis to investigate the mutual correlation between features.
3. Verify the feasibility of using supervised learning for consumption estimation.
4. Apply unsupervised learning such as clustering techniques to identify new relationships.
5. Validate both unsupervised and supervised models.

This approach would guarantee to find the optimal solution leading to the most accurate models. Precisely, while for the supervised approach, it is necessary to choose input features before model training, requiring multiple input features testing based on correlations encountered, the unsupervised approach could be used to understand new relationships by clusters analysis. Additionally, in cases of low correlation in data, the clustering techniques could offer an opportunity for finding simpler relationships which could approximate real behaviour.

To estimate consumption, one of the most used supervised techniques in the current SoA was

used: ANNs. Precisely, multiple tests on hyperparameters were conducted for each vehicle model. In detail, as can be seen in Table 21, each vehicle required a different combination of hyperparameters due to different vehicle models, the amount of data and usage patterns. However, a range of 3 to 5 layers, with between 40 and 300 neurons per layer, was consistently found to be the optimal solution.

Table 21. ANN Model Hyperparameters

Vehicle	Number of Layers	Neurons
#1-Renault-ZOE	4	(80, 150, 70, 100)
#2-Renault-ZOE	5	(100, 200, 100, 200, 100)
#3-Nissan-Leaf	3	(50, 80, 40)
#4-Renault-ZOE	5	(100, 200, 100, 200, 100)
#5-Renault-ZOE	3	(100, 300, 100)
#6-Nissan-Leaf	3	(80, 100, 80)
#7-Renault-ZOE	4	(50, 80, 50, 40)

Since the ANN models demonstrated that even with highly correlated data, depending on the vehicle, some models achieved low R-squared scores, it became necessary to explore a different approach. As a result, unsupervised clustering techniques such as Agglomerative Clustering, K-Means Clustering, and Spectral Clustering were used. To determine the appropriate number of clusters, the elbow method was applied, revealing that four clusters were suitable for all vehicles. The quality of the clusters was subsequently evaluated using the Silhouette Score and the Davies-Bouldin Index.

5.2.1.2 Optimization Algorithm

The optimization algorithm developed for the Emotion Microgrid addresses a gap often noted in the literature: the balance between PV production and electricity consumption for EV charging. Given the practical nature of this study, the end-user was consistently considered to ensure that the findings could be applied tangibly through the DT. Consequently, the optimization problem includes various constraints that account for both user preferences and physics-based relationships. This algorithm aims to exploit various strategies for effectively charging 10 EVs in the Emotion Pilot fleet. It involves the forecasts of the next 24 hours of PV production and building energy consumption, estimating the grid energy purchase (E_{buy}) and potential energy sales (E_{sell}). Other inputs are the real-time vehicle SoCs and user energy requirements (calculated by vehicles models and user preferences). Finally, the 240 decisional variables of this problem are the energy demanded for charging vehicle j in hour i ($E_{i,j}$).

The optimization problem includes seven different constraints which include linear equations and non-linear ones based on if-else states or comparisons between variables. This was possible using Python as a development environment for the optimization problem, guaranteeing both a smooth integration in the DT and interaction between different models by the use of APIs.

The first constraint is applied to incorporate user preferences. Based on the model trained for each vehicle and the user inputs, the required SoC $SoC_{desired}$ and the specified time of use $h_{desired}$ are implemented using Equation (3). Specifically, the sum of the energy consumed in each hour i by vehicle j $E_{i,j}$ between the time of optimization ($i=1$) and the time of use $h_{desired}$ must be greater than or equal to the energy required to reach the desired SoC. This energy is calculated as the difference between the SoC at the time of the optimization run $SoC_{1,j}$ and $SoC_{desired}$, multiplied by the capacity of vehicle j C_j and charging efficiency η_{char} .

For each vehicle j :

$$\sum_{i=1}^{h_{desired}} E_{i,j} \geq C_j * (SoC_{1,j} - SoC_{desired}) * \eta_{char} \quad (27)$$

Secondly, the operation constraints regarding the charging limits are implemented through the equation (28) and (29). For each hour i , the sum of $E_{i,j}$ must be lower than the maximum total plug capacity (44 kW). Simultaneously, the values of $E_{i,j}$ higher than zero in the hour i must be lower than 4, as 4 is the maximum number of plugs available. In detail:

$$\sum_{j=1}^{10} E_{i,j} \leq 44 \quad (28)$$

$$Count.if(E_{i,j} > 0) \leq 4 \quad (29)$$

An additional constraint was set to avoid charging an EV beyond the 100% limit, which depends on the SOC of the vehicle when the simulation starts $SoC_{1,j}$ and its capacity Cap_j , as highlighted in equation (30) which is applied to each vehicle j :

$$\sum_{i=1}^{24} E_{i,j} \leq (100 - SoC_{1,j}) \cdot \frac{Cap_j}{\eta_{char}} \quad (30)$$

To avoid charging too many vehicles unnecessarily (for instance, charging all 10 vehicles in a very short time, leading to minimal energy charging contribution and high user effort in changing vehicles to charge), a new constraint was introduced. First, the constraint counts vehicles set to charge and calculates their potential stored energy from current and 100% SoC. If the excess energy over potential exports is too high, it penalizes solutions, encouraging the

algorithm to minimize the number of vehicles charged.

Finally, to ensure there are no gaps between two non-zero values in the charging sequence, another constraint is introduced to prevent a vehicle from being charged one hour, not charged the next, and then charged again the following hour.

Focusing on output strategies, the only difference is the grade of self-consumption desired, such as maximizing self-consumption or ensuring at least 50% usage of potential exported energy ($Perc_{PV}$). The strategy differentiation is exploited by modifying the self-consumption constraint (31) in the problem.

$$\sum_{i=1}^{24} \left(\sum_{j=1}^{10} E_{i,j} \right) > (Perc_{PV}) \cdot \sum_{i=1}^{24} E_{sell,i} \quad (31)$$

In detail, three different strategies will be provided, considering respectively a percentage $Perc_{PV}$ of 0% (no self-consumption limit), of 50% and 100% (maximize self-consumption).

Even if equation (31) is an inequality constraint, the objective function in equation (32) focuses on minimizing energy costs. Consequently, the decision variables are steered towards equilibrating the energy exported and minimizing charging during periods when PV production is unavailable, as follows:

$$F(X) = \min \left[\sum_{i=1}^{24} \left(\sum_{j=1}^{10} E_{i,j} + E_{buy} - E_{sell} \right) \cdot E_{cost} \right] + Penalties \quad (32)$$

where E_{cost} is the cost of electric energy which differs if the energy is sold or bought, calculated based on average cost associated to kWh using billing details provided by the user. The term *Penalties* in this equation is used to implement each constraint as penalties of the function itself.

In the development of the Emotion Pilot optimization algorithm, after reviewing relevant literature, various algorithms available in python libraries [14–16] were used for solving the optimization problem, including basic GA [16], NSGA-II [17], Chaos Game Optimization (CGO) [18], Improved Grey Wolf Optimization (IGWO) [19], and a hybrid GWO-WOA [20]. In this way, it was possible to find the best algorithm for solving the non-linear problem presented considering the number of constraints met and the time needed for the execution. Different libraries were selected as, for NSGA-II and particularly for GA, the possibility of algorithm customization was higher compared to Mealy which includes the highest number of heuristic and meta-heuristic algorithms.

5.2.1.3 Real-time Processor

The algorithm for real-time management aims to address the limitations of the optimization algorithm charging strategies which rely on predictions that might not always be accurate.

Therefore, the HP energy use is not considered in the previous optimization. Using a simple algorithm, this model uses real-time data, including the power output of charging columns, the SoC of the EVs, as well as the real-time power state of the PV system and the building. Governed by "if-based" rules, the algorithm responds to energy imports from the grid by suggesting either the disconnection of EVs or the adjustment of heat pump setpoint temperatures, lowering them in winter and raising them in summer. Conversely, when there's energy export to the grid, the system recommends connecting additional vehicles (if feasible) or adjusting the setpoint temperature in the opposite direction: increasing in winter and decreasing in summer. This approach would guarantee the use of the thermal capacity of air to store energy when needed.

The algorithm considers the SoCs of all vehicles as input, along with data on the power at each plug of the two charging stations, and the power exchanged with the grid. When there is power export (E_{ex}), the algorithm determines the energy amount that can be distributed among the connected EVs. In the case which the export is less than the maximum available power capacity of the plugs ($Cap_{plug,charge}$), defined as (33), the algorithm recalculates the power assigning the energy in each plug.

$$Cap_{plug,charge} = \sum_{i=1}^{EVs \text{ Connected}} (11 - E_{i,real-time}) \quad (33)$$

In scenarios where the export surpasses the plug's power capacity, the algorithm advises which vehicle to charge first based on the highest EV capacity, defined as (34). Finally, the service suggests adjustments to the heat pump's setpoint temperature, depending on the time period.

$$Cap_{EV} = (100 - SoC_{real-time}) \cdot Cap_{battery} \quad (34)$$

In situations of power import from the grid, the algorithm initially verifies whether the vehicles connected have reached the SoC demanded by the user. Subsequently, the new plug capacity $Cap_{plug,discharge}$ is determined using equation (35) which is specifically tailored to reduce power usage, contrasting with equation (33), in which E_{min} is the minimum charging power that a charging station could reach.

$$Cap_{plug,discharge} = \sum_{i=1}^{EVs \text{ Connected}} (E_{i,real-time} - E_{min}) \quad (35)$$

If the plug capacity is inadequate, the system suggests which vehicle to disconnect verifying whether the vehicles connected have reached the SoC demanded by the user, and it suggests adjusting the setpoint temperature of the heat pump.

5.2.2 Case Study and Results



Fig. 59. The Emotion Microgrid Case Study [21]

The microgrid of this case study is located in the Umbria region, in the centre of Italy, and is hosted by Emotion Srl, a company that produces charging stations and acts as charging point operator and electric mobility service provider. This microgrid infrastructure is composed by:

- 2000 m² tertiary office building
- 50 kWp PV plant
- Two smart charging stations for EVs, equipped with two 22 kW Type 2 connectors.
- A fleet of ten EVs, real-time monitored through OBD systems.
- 45 kW heat pump system, for heating and cooling the office building space

The main devices that will be involved in this study are the fleet of 10 EVs, the two charging columns, the 50 KWp PV panels and the available internal loads. What is notable, is the absence of a battery energy storage, demonstrating the user need to optimize the charge of their vehicles.

The vehicles OBDs were used by the user to collect a wide array of essential data points. This includes metrics such as the connector type used, the maximum power capacity of the battery, its energy capacity, and battery SoC. Moreover, additional operational data monitored included the odometer reading, timestamp, and vehicle autonomy. At the building level, various smart meters are installed in the charging station, building, and PV system. These meters provide measurements of both cumulative quantities, such as the total energy consumed by the building or charging station and the energy generated by the PV system. Additionally, detailed electrical parameters, including power, voltage, and current for each phase, are recorded to offer a comprehensive overview of the system.

5.2.2.1 Data Analysis

In the microgrid presented, the building and EVs consume up to 90 MWh per year, while the 50 kWp PV system produces around 75 MWh. Yet, there is still an issue of matching production

to consumption, with around 40 MWh of energy sold to the grid annually. In 2022, data provided showed that considering the sole energy required for charging EVs, this accounted for 25.5% of the total. Increasing self-consumption is also critical not only for environmental reasons, consuming energy locally minimizes losses of distribution and reduces negative impacts on grid frequency and control, but also for economic ones. For example, the consumption analysis of 4,900 hours from February to August revealed the sale of 22 MWh to the grid and the import of 12 MWh. The self-consumption rate in this period, calculated as the ratio of imported energy to total consumption, is about 45%.

As can be seen from Fig. 60, during a sunny winter day, the building consumption demand and EV charging demand can be totally covered by the PV system. However, due to insufficient production in the morning and evening hours of each day, the required amount of electricity is imported from the grid. The same scenario is repeated on overcast days like Thursday and Friday. In summer, this load is at its minimum value compared with winter, while PV production is maximum. Therefore, more energy is sold to the grid. Even on partly overcast days, the energy yield in the PV system is sufficient to meet the electric load of the building with a high self-consumption rate.

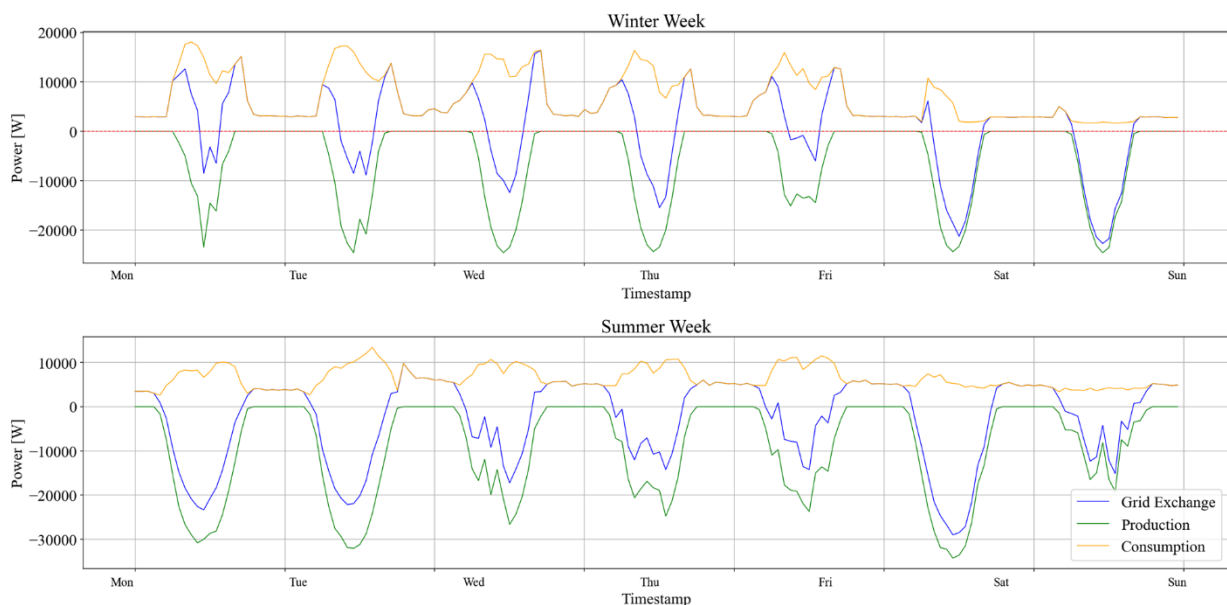


Fig. 60. Microgrid Consumption and Production Behaviour in a Winter and Summer Week

Historical data of EVs provided in this Case Study include information between 2018 and 2020 about 7 EVs. In particular, until January 2020 the data does not comprehend information about vehicle location (latitude and longitude) which is available in the 2020 data. To extend the variables to be used by the ML algorithms, the distance travelled was calculated as difference in the odometer measurement, which has a sensitivity of 1 km, thus not being particularly accurate. The same stands for the energy consumed, which is calculated as the

difference in SoC, multiplied by the capacity of each EV and adjusted based on the time of the interval, which is calculated as difference of two consecutive measurement timestamp. Different criteria were chosen to filter data and consider only relevant measurement, such as taking into consideration only velocities comprised between 2 km/h and the speed limit of each vehicle or considering only negative consumption (to avoid considering the period where vehicle was connected to the charging station). In this way, the periods where vehicle were not used were excluded, passing from 156044 entries to 5845, the 3.75% of the total.

Fig. 61 reveals correlation results between most impactful variables useful for calculating EV consumptions. As expected, the most correlated variables are the energy and the distance, in all of 7 vehicles analysed, ranging from a Pearson score between 0.78 in the case of vehicle #6 (which had only 734 measurements after filtering), and 0.94 in the case of vehicle #2. Focusing on specific consumption, in this dataset it seemed lowly correlated with other variables, suggesting that the unique input features to be used for training the ANN would be the distance for calculating the energy needed to travel that distance. However, as a confirm, it was tested to include also additional input features with the sole results of a degradation of accuracy of the model.

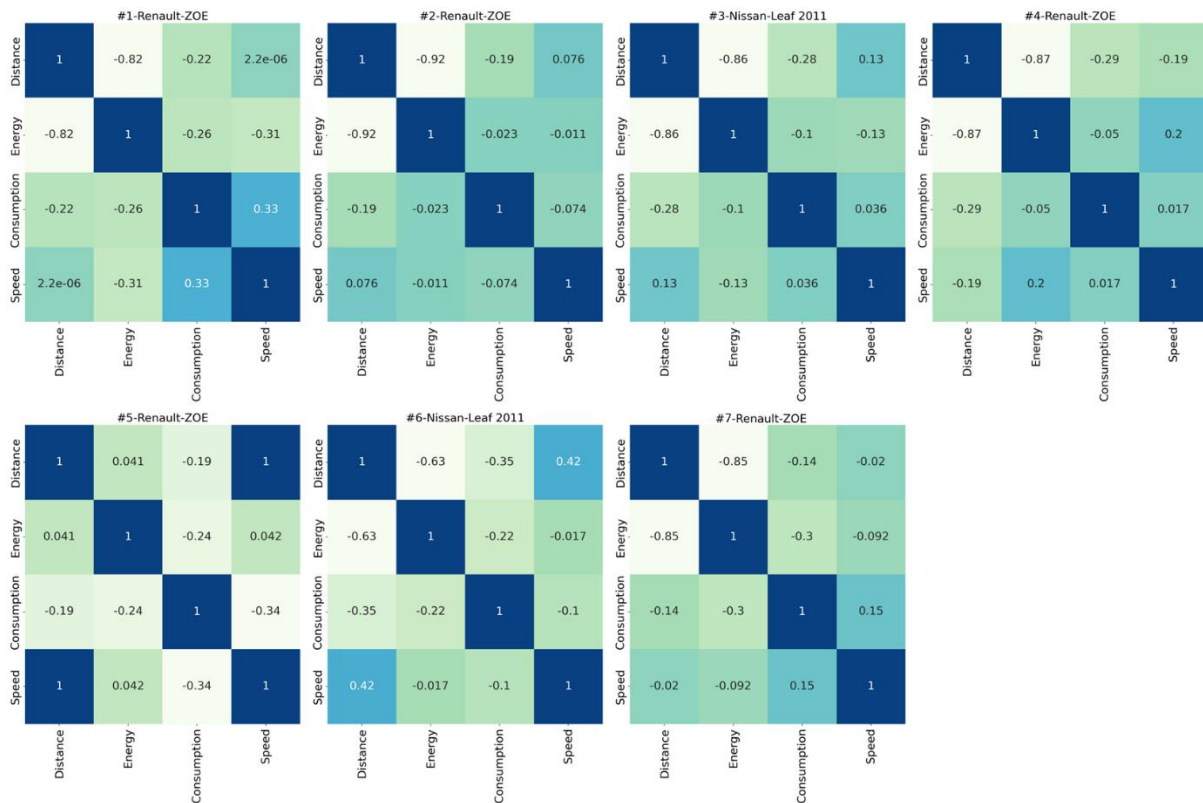


Fig. 61. Correlation Matrixes for 7 Vehicles from Emotion Microgrid

5.2.2.2 EV consumption estimations

Table 22. Metrics of ANN Model for EVs Energy Consumption Prediction

Vehicle	Entries	R ²	MAE	MSE
#1-Renault-ZOE	592	0.69	0.23	0.13
#2-Renault-ZOE	1019	0.82	0.36	0.39
#3-Nissan-Leaf	137	0.75	0.26	0.15
#4-Renault-ZOE	1725	0.76	0.20	0.12
#5-Renault-ZOE	488	0.49	0.13	0.04
#6-Nissan-Leaf	734	0.40	0.20	0.09
#7-Renault-ZOE	157	0.72	0.21	0.07

The results of the ANN model shown in Table 22 demonstrated that, while distance and energy consumption were highly correlated, the R² scores varied between 0.40 and 0.82 depending on the vehicle. This indicates that the ANNs struggled to accurately capture the vehicle's energy consumption behaviour. However, the low MAE and MSE suggest that the predicted and measured consumption values did not differ significantly, showing only slight fluctuations based on the measurements. This implies that using an average consumption value for each vehicle may be a viable approach.

To further explore the potential for improving the reliability of the models for EV consumption estimation, clustering techniques were applied. Specifically, clustering was conducted on available measurements, including speed, consumption, and changes in autonomy. However, many of the clustering results were not particularly informative. For example, when analysing the relationship between consumption, latitude, longitude, and velocity, it became apparent that vehicle speed was the primary factor driving the clustering, as shown in Fig. 62, which compares different clustering techniques based on the consumption behaviour relative to the speed of a Nissan Zoe.

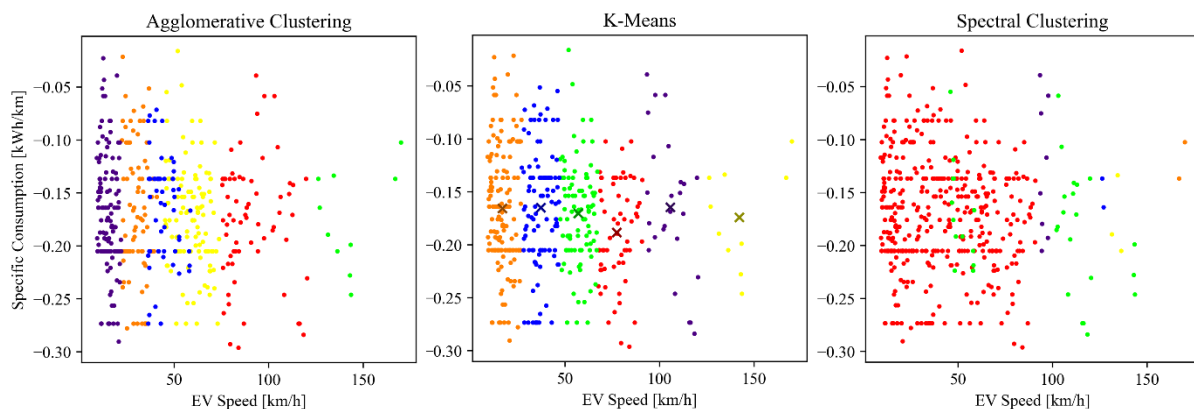


Fig. 62. Analysis of different clustering methods

Three different types of clustering algorithms were investigated but only the K-means and

the Agglomerative clustering led to promising results since the Spectral Clustering individuate 6 different clusters and just one cluster comprehended the 90% of all measurements. Moreover, clustering was also applied considering as features consumption, latitude and longitude revealing that no clustering reason could be clearly identified (Fig. 63). This approach was followed trying to understand if a certain consumption could be applied in a certain place, trying to consider different itinerary such as highways or cities.

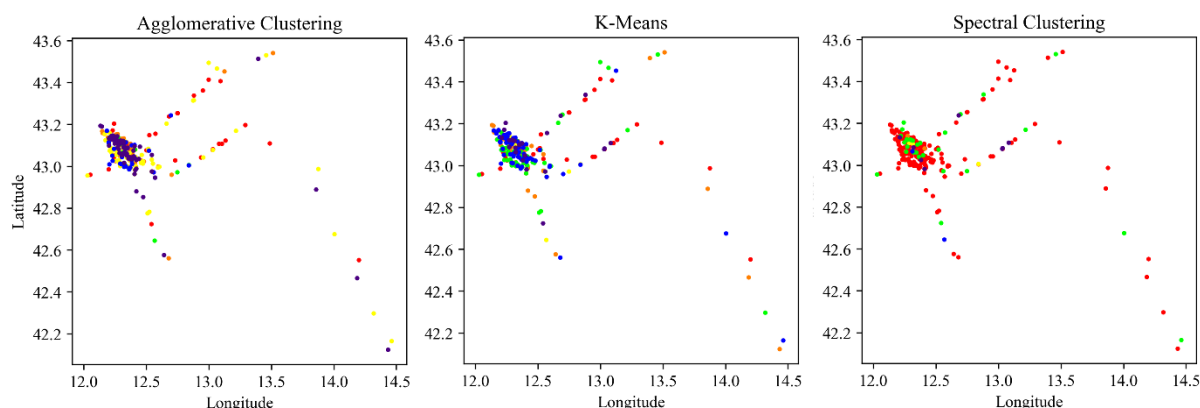


Fig. 63. Analysis of different clustering methods considering position

For this reason, it was decided to consider only velocity to estimate the vehicle consumption since the velocity is a parameter that can be easily calculated given a certain itinerary or set manually by the user. Eliminating the need for latitude and longitude data, allowed the inclusion of data from 2018 to January 2020. Vehicle #4 provided the most collected data set, yielding 1725 entries post-filtering. Utilizing this data, k-means clustering was selected, and the elbow method was employed to determine the optimal cluster count.

Fig. 64 illustrates the elbow method's outcome, including a test with four clusters. To assess the clustering quality, the SS and DBI were utilised, as they are commonly applied in evaluating unsupervised clustering performance. The findings indicated that with four clusters, an SS of 0.75 and a DBI of 0.4 were achieved, signifying a strong and suitable cluster arrangement. By the knowledge of the centroids position, a consumption will be associated to the speed selected by the user in the DT before running the optimization.

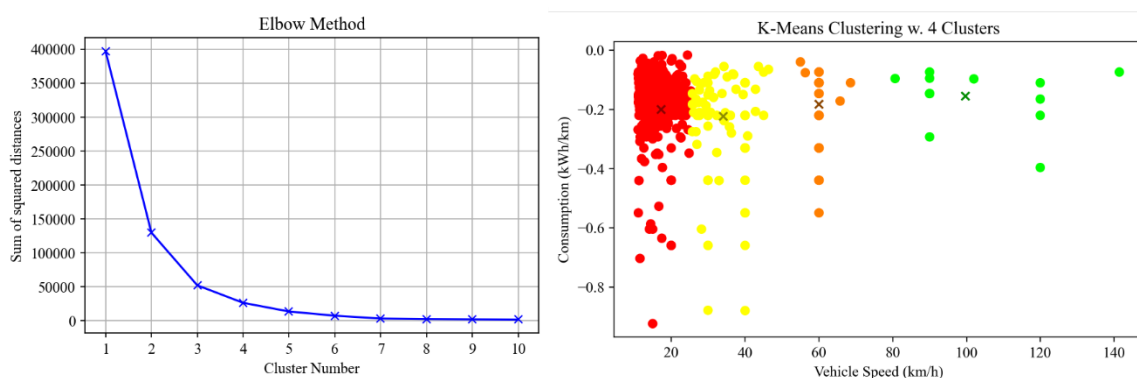


Fig. 64. Elbow method application (left) and k-means result considering only 4 clusters (right)

5.2.2.3 Optimization Algorithm incomes

Given the complexity of the optimization problem, various algorithms and libraries were compared to identify the optimal solution. The algorithms included in this analysis are GA, NSGA-II, CGO, IGWO, and GWO-WOA, as they are among the most commonly used in the current state-of-the-art related to microgrid optimisation [22–25]. To test these algorithms, four different days were selected (two for winter and two for summer) to evaluate their performance under varying conditions. Specifically, for each season, a cloudy day and a sunny day were chosen, as seasonal variations affect not only solar radiation and, consequently, PV production, but also the duration of daylight. However, to maintain consistent boundary conditions, user preferences were assumed to remain constant for each day. Specifically, for Vehicle #2, a distance of 150 km, an average speed of 60 km/h, and an initial SoC of 25% were assumed. For Vehicle #6, a typical mixed route was considered, with an average speed of 30 km/h, a distance of 70 km, and an initial SoC of 10%.

Table 23. Optimization Results in Winter Days

Algorithm	Time of Execution		Convergence Generation		Constraints Violated		Objective Function	
	Sunny	Cloudy	Sunny	Cloudy	Sunny	Cloudy	Sunny	Cloudy
GA	174.59s	99.40s	244	176	0	0	-51.81 €	-45.34 €
NSGA-2	87.28s	81.16s	272	220	0	0	-51.96 €	-45.44 €
CGO	133.4s	63.52s	332	160	0	0	-55.8 €	-45.64 €
IGWO	25.62s	22.71s	283	214	0	0	-51.82 €	-45.34 €
GWO-WOA	22.11s	20.48s	226	189	0	0	-51.86 €	-45.34 €

Simulations were conducted with a Notebook with 16 GB RAM a i7-12750H and a RTX 4060. In all of the cases, similar boundary conditions between different algorithms were applied: a number of individuals per generation of 500, and a number of generations (epochs) equal to 500 to not limit the solution search. On the winter days, all the algorithms provided similar objective function values, with the exception of CGO which suggested charging Vehicle #2 also at 4:00 AM when no PV production was available. Focusing on Sunny Day, as the availability of solar radiation is lower than in summer, it was possible to only charge the two vehicles needed. The charging periods in all the configurations were aligned with PV production, achieving a perfect balance between production and charging power. On the Cloudy Day, only the minimum SoC was achieved due to reduced PV production. As can be noted in Table 23, in both cases, IGWO and GWO-WOA outperformed the other algorithms, reaching the optimal solution up to 0.12% of the time (comparing GA with GWO-WOA) and requiring fewer generations to converge. On the winter days, however, all algorithms succeeded in finding

feasible solutions, demonstrating that when export is limited, all the algorithms converge in fewer than 332 generations.

Table 24. Optimization Results in Summer Days

Algorithm	Time of Execution		Convergence Generation		Constraints Violated		Objective Function	
	Sunny	Cloudy	Sunny	Cloudy	Sunny	Cloudy	Sunny	Cloudy
GA	340s	258s	500	381	1	1	-17.50 €	-23.05 €
NSGA-2	177s	112s	500	376	1	0	-19.43 €	-23.14 €
CGO	204s	197.2s	500	500	1	2	-24.47 €	-34.58 €
IGWO	53s	54.66s	500	500	0	0	-19.35 €	-25.51 €
GWO-WOA	49.5s	29.6s	463	264	0	0	-18.21 €	-23.14 €

During summer, due to the high PV production, it was needed to impose a 5% tolerance on self-consumption constraint for convergence. This approach ensures that the minimum number of vehicles is always charged without violating the constraints for limiting charging vehicle numbers. Additionally, it was observed that exceeding 100% SoC was always avoided except for two different cases. Focusing on the results for a typical sunny summer day (Table 24), even in this case IGWO and GWO-WOA performed the best, consistently converging without violating any constraints in less than 1 minute. Conversely, CGO, which performed well in winter, struggled during summer, failing to find an optimal solution. It exceeded the self-consumption constraint tolerance after 500 generations, requiring over 200 seconds. A similar pattern was observed with GA, which also took 500 generations and slightly exceeded the self-consumption tolerance but achieved a lower objective function value compared to CGO, as some of the energy was sold instead of self-consumed. NSGA-II found an optimal solution that maximized self-consumption; however, it predicted a SoC exceeding 100% for vehicle #10, returning an infeasible solution. As expected, under cloudy conditions, the behaviour resembled that of winter days, with IGWO and GWO-WOA converging more rapidly and producing feasible solutions. In contrast, CGO violated two constraints, with two vehicles exceeding 100% SoC, and failing to achieve self-consumption, leading to non-convergence. NSGA-II achieved the same optimal result as GWO-WOA, but took more time and required more generations. GA implementation violated two constraints: one for the constraints that should limit the occurrence of alternative charging/discharging and slightly exceed the self-consumption tolerance.

An 11-day test period was conducted from 06/12/23 to 16/12/23, where the GWO-WOA algorithm was applied to historical data to assess its efficiency in saving energy for EV charging. This period, falling in winter did not always have excess PV production, resulting in days where the algorithm recommended not charging the EVs. User preferences were excluded

to simplify the testing process since more than one day was simultaneously considered in the simulation. The algorithm suggested strategies to charge EVs with the available PV-generated energy, as illustrated in Fig. 65 confirming that the algorithm effectively could guide users to charge EVs using their self-generated power each day.

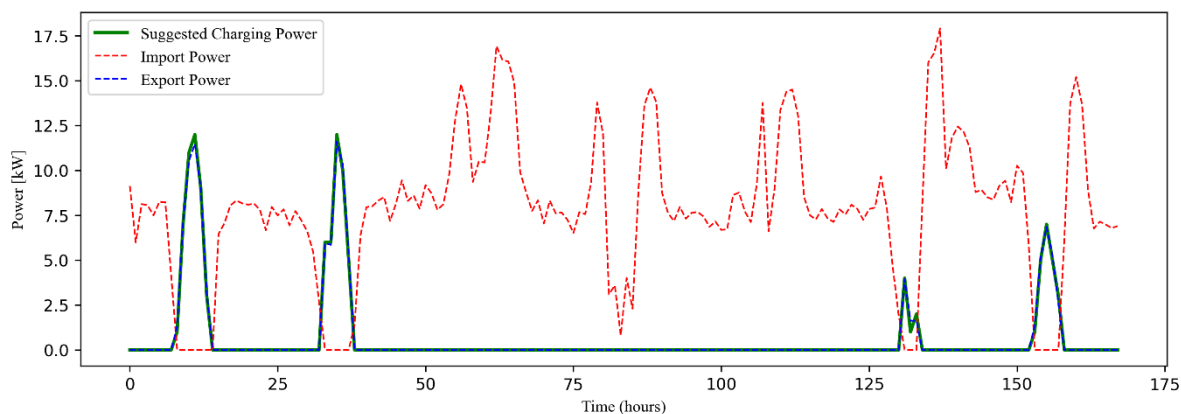


Fig. 65. Weekly Power trends for optimal EV charging compared to power import and export.

Fig. 66 serves as an illustrative example of the dynamic data from a single operational day, showcasing a scenario with overproduction. On these particular days, there was assumed a user requirement to charge Vehicle 4 to 50% by 14:00 and Vehicle 6 to 40% by 18:00. The graph displays the way the vehicles are fully charged during times when there could be potential energy export, optimizing self-consumption. Additionally, the vehicles are not only charged to the required SoC levels but also beyond, according to available PV production. Consequently, the final SoC for EV4 reaches 66%, exceeding the initial target. When the available export capacity is insufficient to fully meet the EV charging requirements, as specified by the user, the algorithm proposes a strategy where a portion of the energy needed for charging EVs is sourced during periods of excess PV production, as shown in Fig. 66. As expected, in this case only the minimum SoC requirement is satisfied.

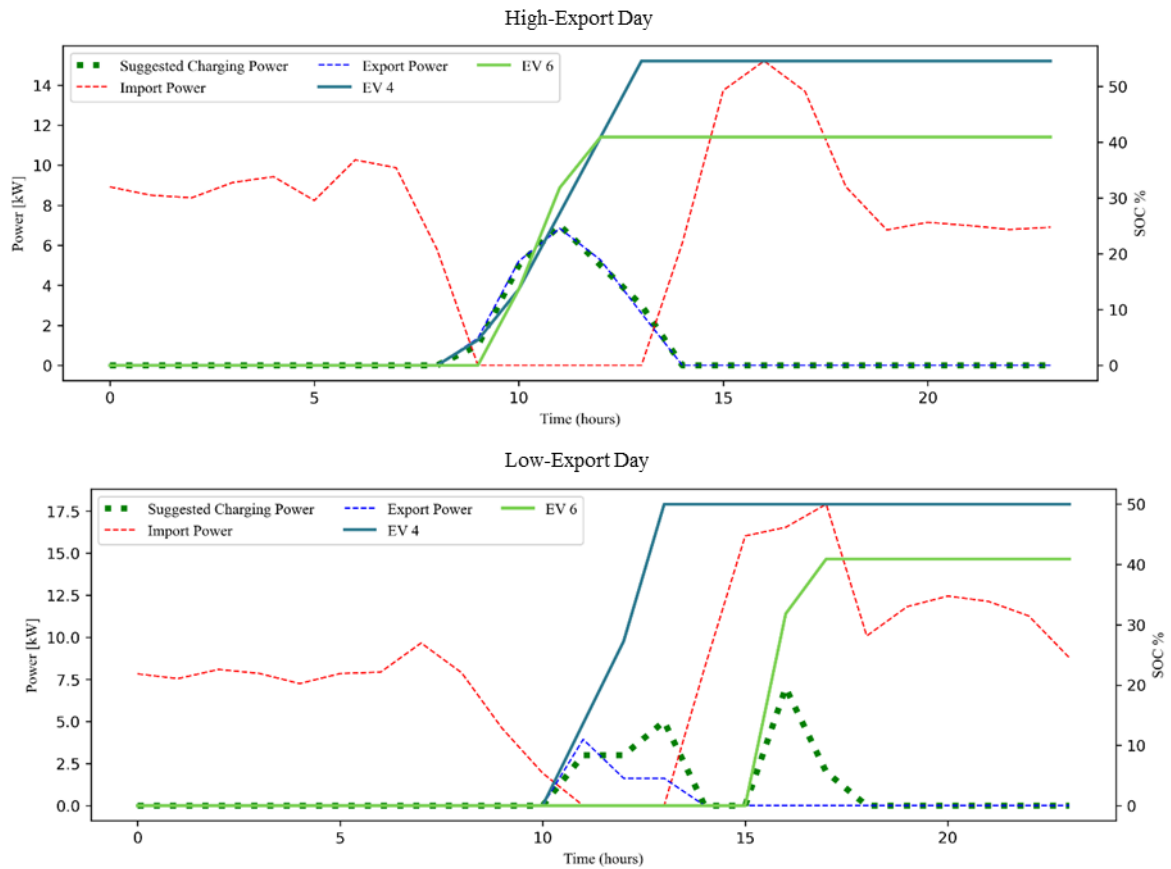


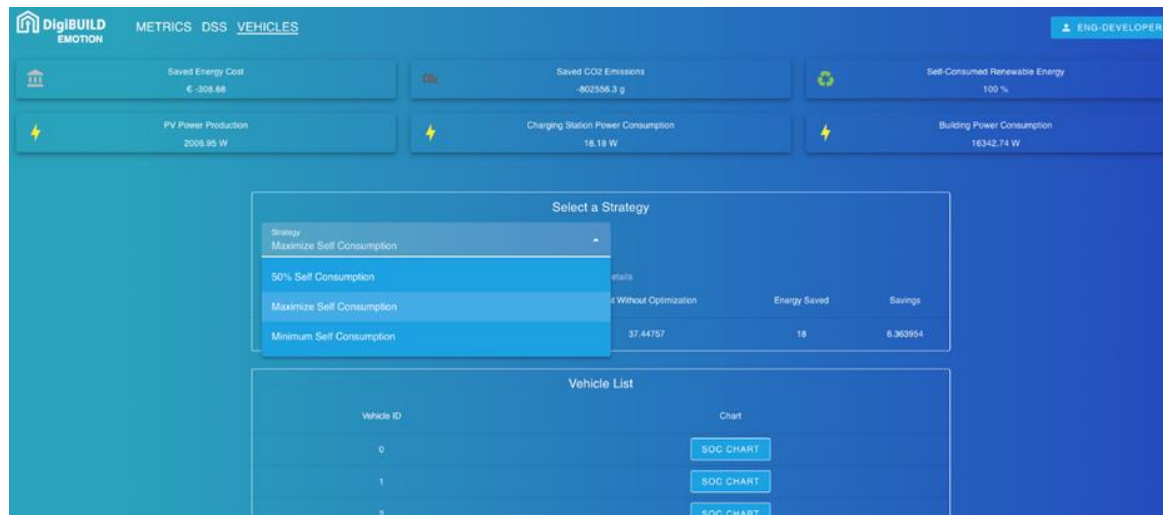
Fig. 66. Power trends and SoCs on winter days considering EV user preferences.

5.2.2.4 Emotion Digital Twin Dashboard

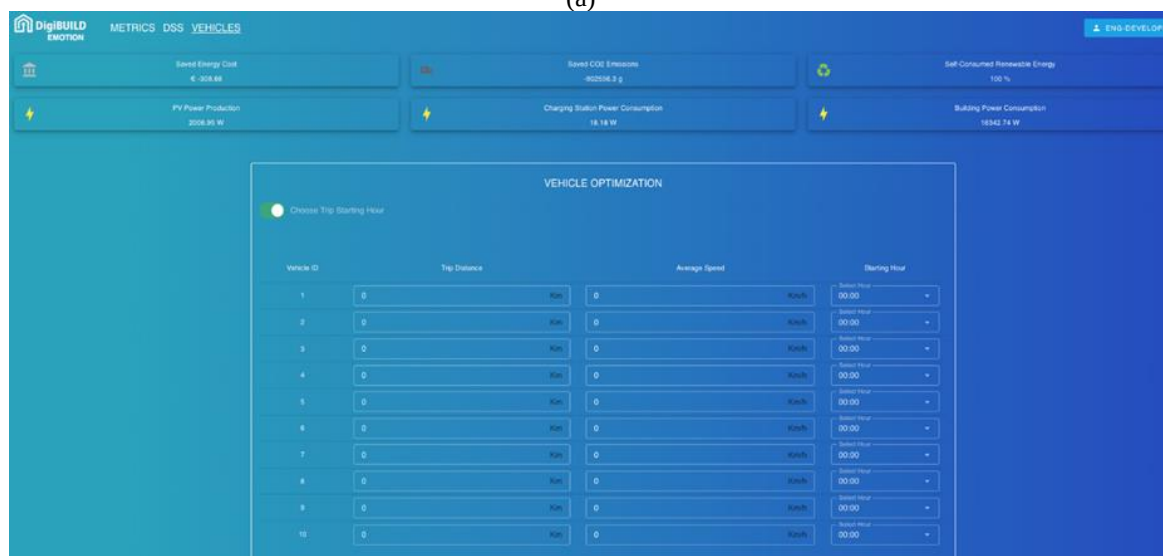
The presented algorithms are used in the DT's DSS, which focuses on optimizing EV charging processes to enhance self-consumption of generated power while also reducing costs. Within this DT, users can access real-time data from the homepage, such as the instantaneous power of each charging station plug, building energy consumption, and power output from the PV plant. Furthermore, as the 10 EVs are fitted with OBD systems, continuous monitoring via the DT is available. This setup also allows for the constant refinement of vehicle consumption ML models through ongoing data collection. Finally, the data collected are utilized to train regressors that forecast future PV production or building energy consumption, both of which can be viewed on the dashboard of the DT.

The DT incorporates three optimization strategies through the Middleware API (Fig. 67a), offering day-ahead and real-time recommendations for facility managers on which vehicle to charge, the optimal charging times, and methods. The day-ahead optimization output includes forecasts of plug power output for the subsequent day and projections of future SoC. Through the DT dashboard, users can specify which EV will be required the next day, the distance to be travelled, and the expected type of itinerary, as shown in Fig. 67(b). After this input, the optimization process is initiated, and the interface presents various charging strategies based on

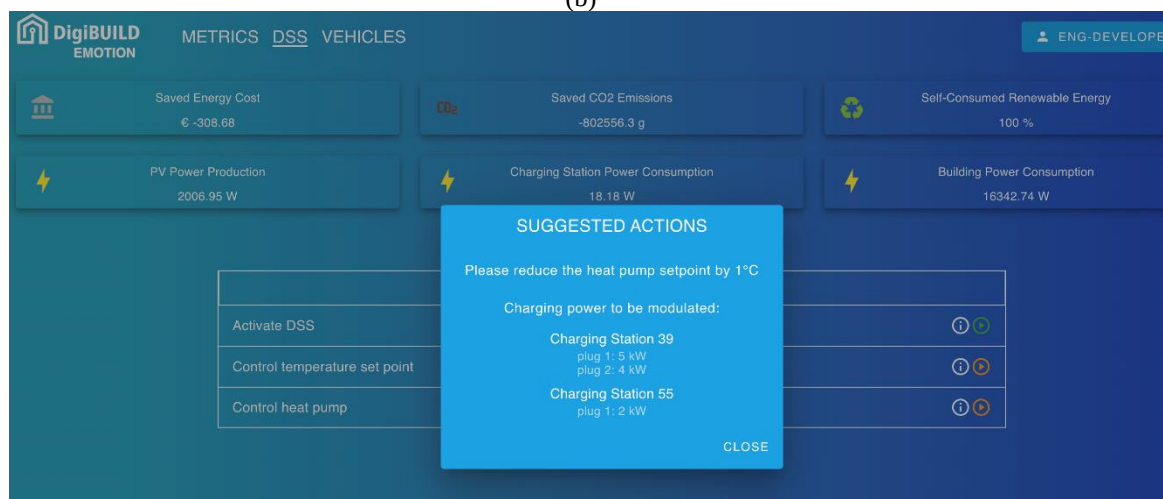
different expected percentages of self-consumption.



(a)



(b)



(c)

Fig. 67. Sections of the Emotion Digital Twin Dashboard: (a) vehicle selection for optimization; (b) results visualization; (c) DSS page.

Once a strategy is selected, the DT autonomously can control the charging station through

an API, after the user connects the corresponding EV, as recommended by the optimization algorithm. The vehicle list displayed in Fig. 67(a) informs the user of the expected future EV SoC via a chart and indicates the remaining hours of charge. Additionally, a real-time optimization, as shown in Fig. 67(c), suggests adjustments to the power output of active plugs to align current PV generation with building energy needs, aiming to reduce energy exports and imports from the grid. If the options for power modulation are limited, such as when only one vehicle is connected, the DSS page guides the user by recommending which vehicle to connect or how to adjust the HP setpoint to alter the building's electrical consumption. This adjustment allows for increasing the heating setpoint during the day so that the energy required for heating in the evening is reduced.

5.3 Final Discussion and Future Directions

The increased complexity of today's electrical grids, and those expected within the near future due to high penetrations of variable renewable generation along with large-scale EV integrations, demand advanced management strategies that balance stability and efficiency. ML and Optimization algorithms have been in wide consideration in recent literature, especially for the charging of EVs, while balancing the challenge of charging demand with grid stability. These theoretically advanced improvements have yet to be joined up with the real deployment of these approaches in actual case studies.

This gap was addressed in this section by discussing a case study based on the DigiBUILD project applied to the company Emotion, a specialist in EV charging stations. The DT developed during this work integrated ML models and optimization algorithms, in a microgrid which includes 10 EVs and a building with 50 kWp PV panels and two EV charging stations of 22 kW each (through 4 plugs in total). The aim is to achieve a DT capable of monitoring and simulating, using forecasted and real-time data the dynamics of energy in the building and, on the other hand, provide recommendations on EV charging strategies, considering user preferences.

Realistically, the system was supposed to deal with a fleet of ten EVs by dynamically modifying its charging schedule with an explicit purpose: cost reduction and maximizing the utilization of available PV power contributing to the energy efficiency increase and stability of the grid. Simulations run on historical data revealed that the optimization algorithms succeed in providing solutions for charging EVs maximizing self-consumption in less than 25s. This algorithm was tested on different days and is being implemented through the DT. Focusing on ML models, the clustering techniques used reveal that it was possible to effectively relate specific electric consumption to speed by a division in 4-5 clusters. Moreover, a user-oriented

DSS was developed as a part of the DT dashboard to include suggestions, such as connecting/disconnecting EVs based on the state of the grid or changing the setpoint for HVAC systems. This real-time processor will also balance PV production and consumption modulating automatically the charging power.

It is worth noting that challenges were encountered during the development of the tools for the DT of Emotion. Data collection from EVs faced communication issues with sensors via MQTT, resulting in limited data availability for training across multiple vehicles. Additionally, smart metering for building energy consumption and PV production began only at the end of 2023, restricting data availability. Since most DigiBUILD services, except optimization, rely on data-driven approaches, addressing this limitation should be a priority in future work. Moreover, while the autopilot mode can directly control the charging station via the API to maximize self-consumption in real-time, it remains uncertain whether end-users will fully trust such a system or prefer suggestion-based guidance. Balancing user needs with algorithmic requirements is a difficult task: for example, optimization processes may require more time to find optimal solutions, whereas users often expect quick responses. Despite these challenges, the proposed approach can be readily implemented in facilities equipped with PV systems and EVs, as it relies only on open-source tools. The primary challenge lies in tailoring the optimization algorithm's constraints to meet specific user requirements, although the presented methodology is sufficiently generalized to accommodate such adaptations.

This work, therefore, contributes to the ever-growing literature on DTs in smart energy management and shows how proper integration of real-time data with an adaptive control strategy can yield a great enhancement in facility operation hosting RESs and EV fleets. While the field is getting wider, the application of such advanced tools will be crucial to achieve broader goals on energy efficiency and sustainability.

References

- [1] Zacharopoulos L, Thonemann N, Dumeier M, Geldermann J. Environmental optimization of the charge of battery electric vehicles. *Appl Energy* 2023;329:120259. <https://doi.org/10.1016/j.apenergy.2022.120259>.
- [2] Bao D-W, Zhou J-Y, Zhang Z-Q, Chen Z, Kang D. Mixed fleet scheduling method for airport ground service vehicles under the trend of electrification. *J Air Transp Manag* 2023;108:102379. <https://doi.org/10.1016/j.jairtraman.2023.102379>.
- [3] Yin W, Ji J, Wen T, Zhang C. Study on orderly charging strategy of EV with load forecasting. *Energy* 2023;278:127818. <https://doi.org/10.1016/j.energy.2023.127818>.
- [4] Calise F, Cappiello FL, Dentice d'Accadia M, Vicidomini M. Smart grid energy district based on the integration of electric vehicles and combined heat and power generation. *Energy Convers Manag* 2021;234. <https://doi.org/10.1016/j.enconman.2021.113932>.
- [5] Kaewdornhan N, Srithapon C, Liemthong R, Chatthaworn R. Real-Time Multi-Home Energy Management with EV Charging Scheduling Using Multi-Agent Deep Reinforcement Learning Optimization. *Energies (Basel)* 2023;16. <https://doi.org/10.3390/en16052357>.
- [6] Liu J, Wang H, Du Y, Lu Y, Wang Z. Multi-objective optimal peak load shaving strategy using coordinated scheduling of EVs and BESS with adoption of MORBPSO. *J Energy Storage* 2023;64. <https://doi.org/10.1016/j.est.2023.107121>.
- [7] Mathur D, Kanwar N, Goyal SK. A Cost-Efficient Energy Management of EV Integrated Community Microgrid, 2024, p. 329–39. https://doi.org/10.1007/978-981-99-4795-9_31.
- [8] Uzair M, Ali Abbas Kazmi S. A multi-criteria decision model to support sustainable building energy management system with intelligent automation. *Energy Build* 2023;301:113687. <https://doi.org/10.1016/j.enbuild.2023.113687>.
- [9] International Energy Agency. Electric Vehicles. 2024.
- [10] Zhao L, Yao W, Wang Y, Hu J. Machine Learning-Based Method for Remaining Range Prediction of Electric Vehicles. *IEEE Access* 2020;8:212423–41. <https://doi.org/10.1109/ACCESS.2020.3039815>.
- [11] Yavasoglu HA, Tetik YE, Gokce K. Implementation of machine learning based real time range estimation method without destination knowledge for BEVs. *Energy* 2019;172:1179–86. <https://doi.org/10.1016/j.energy.2019.02.032>.
- [12] DigiBUILD Project - High-Quality Data-Driven Services for a Digital Built Environment towards a Climate-Neutral Building Stock n.d. <https://digibuild-project.eu/> (accessed June 27, 2024).
- [13] ahmed HA kh, taqi K mohammed, Hassan AMS, khaki M. Joint energy management with

- integration of renewable energy sources considering energy and reserve minimization. *Electric Power Systems Research* 2024;232:110412. <https://doi.org/10.1016/j.epsr.2024.110412>.
- [14] Blank J, Deb K. Pymoo: Multi-Objective Optimization in Python. *IEEE Access* 2020;8:89497–509. <https://doi.org/10.1109/ACCESS.2020.2990567>.
- [15] Van Thieu N, Mirjalili S. MEALPY: An open-source library for latest meta-heuristic algorithms in Python. *Journal of Systems Architecture* 2023;139. <https://doi.org/10.1016/j.sysarc.2023.102871>.
- [16] Gad AF. PyGAD: an intuitive genetic algorithm Python library. *Multimed Tools Appl* 2023. <https://doi.org/10.1007/s11042-023-17167-y>.
- [17] Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 2002;6:182–97. <https://doi.org/10.1109/4235.996017>.
- [18] Talatahari S, Azizi M. Chaos Game Optimization: a novel metaheuristic algorithm. *Artif Intell Rev* 2021;54:917–1004. <https://doi.org/10.1007/s10462-020-09867-w>.
- [19] Kaveh A, Zakian P. Improved GWO algorithm for optimal design of truss structures. *Eng Comput* 2018;34:685–707. <https://doi.org/10.1007/s00366-017-0567-1>.
- [20] Obadina OO, Thaha MA, Althoefer K, Shaheed MH. Dynamic characterization of a master–slave robotic manipulator using a hybrid grey wolf–whale optimization algorithm. *Journal of Vibration and Control* 2022;28:1992–2003. <https://doi.org/10.1177/10775463211003402>.
- [21] Google LLC. Google Earth 7.3.6 2023.
- [22] Li Q, Cui Z, Cai Y, Su Y. Multi-objective operation of solar-based microgrids incorporating artificial neural network and grey wolf optimizer in digital twin. *Solar Energy* 2023;262. <https://doi.org/10.1016/j.solener.2023.111873>.
- [23] Testasecca T, Stamatopoulos S, Lazzaro M, Sarvas E. Recent advances on data-driven services for smart energy systems optimization and pro-active management. 2023 IEEE International Workshop on Metrology for Living Environment (MetroLivEnv), 2023. <https://doi.org/10.1109/MetroLivEnv56897.2023.10164056>.
- [24] Fan X, Li Y. Energy management of renewable based power grids using artificial intelligence: Digital twin of renewables. *Solar Energy* 2023;262. <https://doi.org/10.1016/j.solener.2023.111867>.
- [25] Xu J, Gong J. Novel sustainable urban management framework based on solar energy and digital twin. *Solar Energy* 2023;262. <https://doi.org/10.1016/j.solener.2023.111861>.

Chapter 6

Digital Twin of a Solid Oxide Fuel Cell

The continued rise in carbon emissions, marked by a +1.1% increase in 2023, underscores the ongoing need for intensified efforts to reverse this global trend and achieve the goals outlined in the Paris Agreement [1]. While the expanding deployment of renewable energy capacity worldwide has partially mitigated the rise in emissions, there remains a need for additional systems capable of ensuring consistent and reliable energy production irrespective of weather conditions (such as solar radiation, wind speed, or tidal flows). Consequently, attention is increasingly turning towards systems that can utilise clean fuels or achieve higher efficiencies to facilitate overall emissions reduction. Simultaneously, the green transition involves different key technologies to decarbonise and make more resilient energy production systems. On one hand, hydrogen technologies, such as fuel cells (FCs) or electrolyzers, have a pivotal role in the clean transition in different national strategies to produce, store and use clean hydrogen. On the other hand, the digitalisation wave is also permeating the energy sector, where natural gas (NG) still accounts for almost a quarter of global electricity generation. FCs emerge as a promising solution in this context due to their high efficiency and versatility in operating with clean and renewable fuels, such as hydrogen or syngas produced from waste or biomass, as well as various hydrocarbons [2].

6.1 Overview of Solid Oxide Fuel Cells

Among the different FC technologies, SOFCs operate at the highest temperatures, which range from 600 °C to 1000 °C. This condition, on the one hand, requires a longer start-up time but, on the other hand, can ensure higher efficiency (mostly if used in a cogenerative power system), higher tolerance to corrosion, and can operate with different fuels without the need for apportioning changes to the system [3]. During the last decade, particular attention has also been paid to the development of DMs that can be used to predict cell performance, component degradation, and fault diagnosis, or can act as tools to support component design (thickness, porosity etc.) [4–6]. More recently, the concept of the DM has been extended to the DT which represents a virtual replica created using real-time data collected from sensors embedded in the

physical counterpart. In such a way, DTs can bridge the gap between the physical and digital worlds [7].

Electrochemical parameter identification and optimisation are among the most studied aspects in SOFC modelling, through the application of meta-heuristic algorithms. Yang et al. [8] developed an improved GA for a tubular dynamic SOFC model to enhance output performance based on cell electrical parameters. Gong et al. [9], proposed an adaptive differential evolution algorithm called IJADE (Improved Jingqiao Adaptive Differential Evolution) to fast and accurately identify SOFC parameters. Jiang et al. [10] used a cooperative barebone particle swarm optimisation method to solve different subproblems that regulate the SOFC model based on six internal variables. Wei and Stanford [11] proposed a novel method to evaluate with high stability and accuracy SOFC conditions for steady state and transient operations. Furthermore, in [12–14], the SOFC parameter definition was tackled through different optimisation methods such as Levenberg–Marquardt backpropagation, extreme ML, and the GWO method. Lastly, a DT model with artificial rabbit optimisation was developed by Guo et al. [7] to model a hybrid PV–SOFC system and define the internal unknown parameters.

Three further studies investigate the inner conditions of SOFC components through digital modelling. Hwang et al. [15,16] used DL-assisted semantic segmentation to gain insight into cathode and anode microstructure (surface and volume fractions, data on two and three-phase boundaries, etc.) starting from focused-ion beam scanning electron microscope images. A DL process was also adopted in combination with FEM to foresee the elastic properties of cathode materials. Regarding system operation, a widely investigated assessment is related to fault detection and diagnosis [17,18], by using different ML approaches like Random Forest, or ANN. These processes can effectively determine fault conditions like component degradation or gas leakage.

The prediction of system efficiency by modelling the polarisation curve was conducted by Subotic et al. [19] through the application of ANN. The authors trained the model with both numerical and experimental data, while for validation, only experimental values were used. Performance optimisation was based on operating temperature, fuel mix, and current density. Output parameters, in terms of voltage and thermoelectric efficiency, were also investigated in [6,20,21], including the prediction of voltage level concerning system degradation.

6.2 Data-Driven Models for Solid Oxide Fuel Cell Digital Twin

The studies reviewed in this section provide a comprehensive overview of modelling techniques in the literature, focusing on fault detection or efficiency improvement. However, from a high-level perspective, these techniques offered a level of detail that does not

significantly impact the practical management of an SOFC, particularly in commercial applications. This insight motivated the current research to estimate fuel efficiency based solely on parameters that can be controlled and adjusted during operation. In this work, the DT of the SOFC is employed to maintain an accurate ML model capable of monitoring ageing, identifying potential faults, and replicating electric efficiency as operational conditions change. The efficiency estimation, which considers electric power and includes ramp-up and ramp-down phases for the SOFC, is based on a novel approach integrating both RF and DL algorithms. Moreover, the DT will consistently provide the most accurate models, utilising the best available data. This SOFC DT will be part of a broader framework that integrates with other energy systems, optimising both the cell and the entire energy hub, which is based on hydrogen and sustainable energy resources.

In the remaining sections of this paper, Section 2 reports the DT architecture, defining the algorithms used to model the SOFC, including both DL and ML approaches, together with all the validation metrics adopted. Section 3 first describes the experimental setup of the case study and the results from data analysis, including a description of obstacles in data collection. Finally, the models are compared, and validation scores are presented. Conclusions are then reported in Section 4 to summarise the outcomes of the paper and highlight the next steps.

6.2.1 Methodology

In this work, the presented models are used as the model layer of the DT of a real SOFC powered by NG. In particular, the DT being developed is composed of different blocks to be used to produce electric energy efficiently. Focusing on the schema available in Fig. 68, which represents the architecture of the proposed twin, it can be noted how the ML models constitute one of the cores of the twin. Specifically, the DT comprises three primary elements where the first focuses on displaying key parameters through a dashboard. Additionally, the back-end of this interface collects data and initiates Python scripts to update the ML models and explore optimization strategies.

The ML models, the focus of this work, include regressors such as Random Forest, MLP and GBoost, which are updated daily. This frequent updating facilitates three beneficial outcomes: firstly, the models evolve similarly to the actual SOFC, mirroring its ageing process: a distinct advantage of having a DT of a real system. Secondly, the models enable assessments of how the fuel cell ages, which is crucial for identifying potential malfunctions or quality degradation. Lastly, daily updates ensure that optimization algorithms are simulated using the most current model, which may vary as more data is accumulated. Indeed, performance metrics such as MAE, MSE, RMSE, and others are calculated to identify the algorithms that best represent the

actual behaviour of the fuel cell. In the future, this process could be extended by using all data available in BlueGen Database in order to increase the prediction accuracy and develop fault detection algorithms.

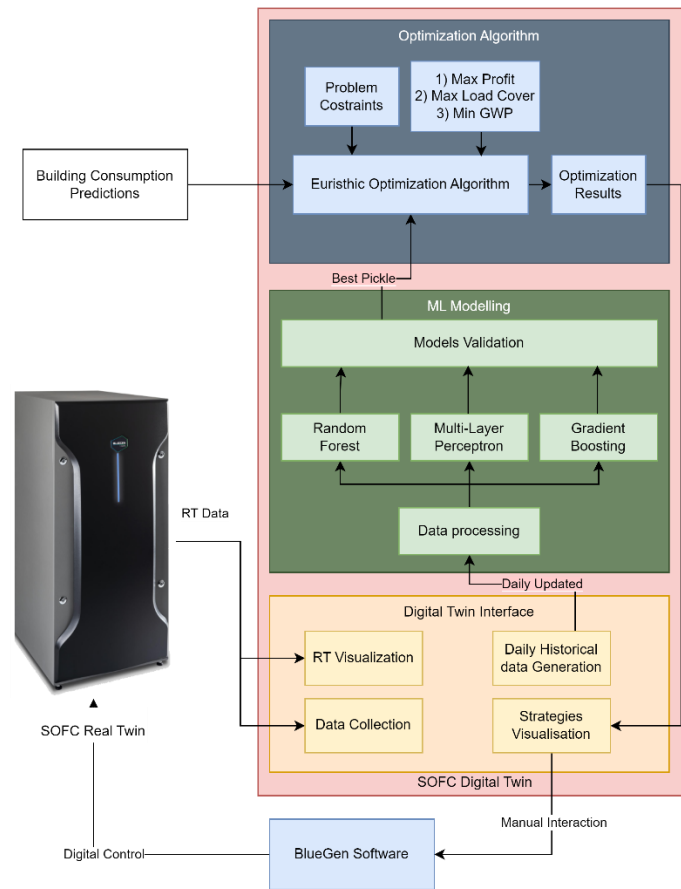


Fig. 68. SOFC DT Architecture

The third block of the SOFC DT includes the optimization algorithm which is used to calculate different operating strategies to modulate the Fuel Cell in the actual day. Precisely, the algorithm objective functions include three different objectives: maximizing profit to assess economic benefits, maximizing the load cover factor to evaluate the building's autonomy from the grid, and minimizing the Global Warming Potential to reduce the carbon footprint of the cell. Based on these results, for each strategy, the operator can choose one of these strategies, that will be visualized on the DT interface, and select using BlueGen Software, the correct power modulation to operate the fuel cell.

Focusing on the RF model, in this work, 280 decision trees were utilized, each with a depth of 20 and minimum sample values for split and leaf set to 3 and 1, respectively. It is notable that the number of trees and their depth, particularly the latter, are higher than those used in GBoost and XGBoost. This was chosen because, in boosting processes, new trees are incrementally added, and it is preferable to have simpler individual trees to avoid a sure overfitting. Conversely, in the RF algorithm, a single forest composed of diverse individual

trees is used, allowing each tree to possess a more complex structure. The XGBoost Python library was then utilized [22], with 100 estimators selected, a learning rate of 0.1, and a maximum tree depth of 5. Specifically, the number of estimators corresponds to the number of trees iteratively added to the model. The learning rate acts as a shrinkage factor applied to new trees to prevent overfitting. Moreover, although increasing tree depth could enhance model precision by adding complexity to each tree, it was determined that a maximum depth of 5 was optimal for this model. This choice finally aligns with the recommendation in [22] to use values between 4 and 8. Finally, the Gboost model developed in this study includes: a number of estimators (the decision trees) equal to 200, each with a depth of 5 and minimum sample values for split and leaf set to 3 and 1, respectively. Finally, including a learning rate set to 0.1, the combination of these hyperparameters led to the most robust model.

Concerning DL algorithms, an MLP NN was structured with six hidden layers, respectively composed of 10, 52, 50, 100, and 10 neurons. An iteration process of finding different layers was carried out to find the correct values of hyperparameters, which in ANN represent the number of hidden layers and neurons for layers. This process was necessary to finetune the network to maximize the accuracy of the model while reducing the time spent in the iterative process. Additionally, since using GAs and the internal optimization process of DL algorithms, is affected by the randomness of parameters, different processes were carried out for different random states.

In this study, despite, to the knowledge of authors, no recurrent dependencies as expected, this model was even selected for investigating this possibility. The model architecture, developed using Keras, included an LSTM layer followed by a dense layer with one unit. The LSTM layer includes 100 units (or neurons), iteratively selected in order to achieve the most accurate model. Secondly, the dense layer is set equal to 1 as required for the regression process of time-series value.

6.2.2 Case Study and Framework

The SOFC generator comprises a hot box containing an SOFC stack, a reformer, a burner, a heat recuperator, and a condenser to recover the water for the reforming process. A control and diagnosis board, thanks to numerous sensors that constantly provide readouts of all operational parameters necessary for operations and safety, manages the functions of ancillaries (Balance of Plant) [23]. The experimental setup of the SOFC installed at the CNR-ITAE facilities in Messina, Italy can be seen in Fig. 69. It is observable that the SOFC is directly connected to the gas network via a yellow valve, enabling a direct connection to the national NG grid and the potential for integrating a blended mixture of NG and hydrogen. In addition to the electrical

connection, this SOFC also can produce heat with a nominal thermal output of 0.75 kW. Accordingly, connections for both inlet and outlet water flow rates are installed, ensuring the proper operation of the cell.



Fig. 69. SOFC Experimental Setup.

For the sake of completeness, the main operational parameters, according to the producer datasheet, are reported in Table 25.

Table 25. SOFC Operational Parameters.

Parameter	Value
Nominal electric power in AC [kW]	1.3
Minimum power output in AC [kW]	0.5
Maximum power output in AC [kW]	1.5
Electric efficiency in nominal condition on NG LHV	57%
Cogeneration efficiency in nominal condition on NG LHV	88%
Nominal thermal power output [kW]	0.75
Nominal NG consumption [m ³ /h]	0.24
NG supply pressure [mbar]	17–25
Start-up time [h]	24

The SOFC in this study allowed for the collection of cumulative and instantaneous data on

an hourly basis. The cumulative outputs included hours of operation, electric power produced, and CO₂ avoided, while the instantaneous data included the SOFC status, electric power, inlet gas flow rate, inlet gas power, electric efficiency, and CO₂ emissions. The status specifically indicated whether the cell was in “Heat Up”, “Power Export”, “Cool Down”, or “Off”. An initial data overview indicated that efficiency depended on whether the FC was increasing or decreasing its power, or operating in a steady state. Therefore, the difference between power at two consecutive timesteps was calculated to distinguish moments of ramp-up, ramp-down, or constant power. The status was updated with this information, generating three distinct time series for power ramping up, ramping down, and remaining constant.

6.2.3 Results and Findings

6.2.3.1 Data Analysis

The experimental phases extended over six months, accumulating 3935 h of data, which included periods of shutdowns. In Fig. 70, the power data of electrical efficiency, NG, and electric power during the periods when the SOFC was operational, which were used for training the model, are shown. Initially, the SOFC was operated at nominal power (1.5 kW) and as observed in the graph, up to hour 750, the efficiency of the FC did not vary significantly. After experimenting with a systematic ramp-up and ramp-down step-like procedure, Fig. 70 shows that the FC efficiency decreases as the power production decreases.

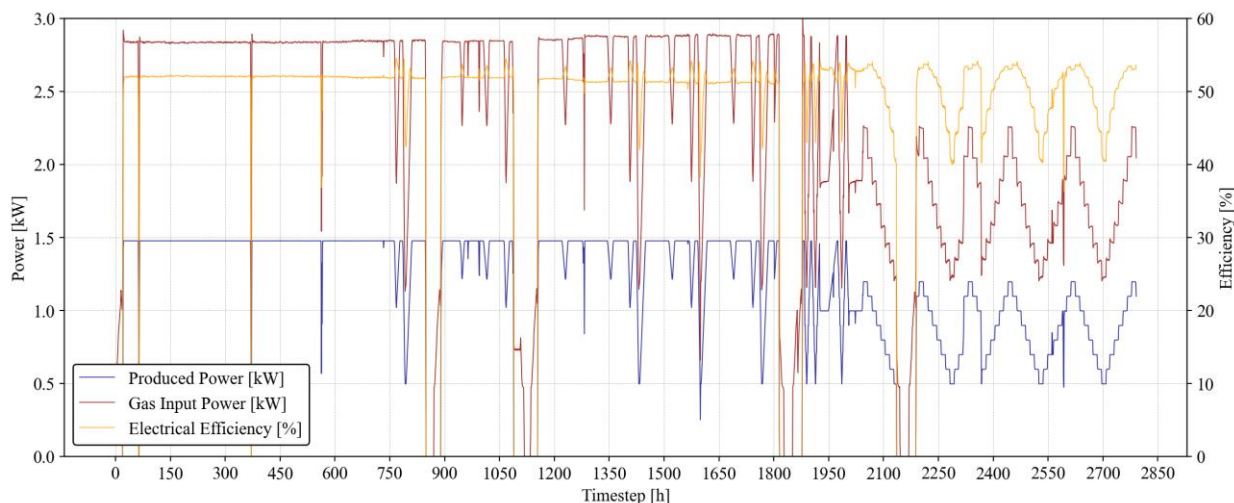


Fig. 70. Behaviour of Power and Efficiency in the Operational Experimental Period.

To develop models capable of predicting energy efficiency, correlations between efficiency and other variables were investigated, as illustrated in Fig. 71. The second-order polynomial trendline, which fits the variation in electrical efficiency concerning electric power, demonstrates that power increments significantly influence the SOFC’s efficiency. As seen on the left, during ramp-up phases the cell always has a lower efficiency compared to the case in

steady functioning, while considering ramp-down, at lower cell power, the efficiency in steady state is lower than when it operates in ramp-down. Conversely, when the SOFC is operating in power higher than 1000 W, it appears that ramp-down phases achieve lower efficiency. The correlation matrix on the right of Fig. 71a reveals how the electric efficiency, as expected, could be easily correlated with produced power, and obviously from gas-related quantities such as gas input and CO₂ emissions, which strictly depend on the quantity of gas used. This revealed that increasing the power should lead to higher efficiency, as also highlighted in the right graph (Fig. 71b). However, it seems that the power increment is not strictly correlated to electrical efficiency, conversely to what is encountered in reading the graph on the right. These reasons led to the choice of ML and DL models, which can account for non-linear relations such as XGboost, RF, or ANN. In all of these models, to calculate FC electric efficiency, the input features selected were the produced power and the power increment.

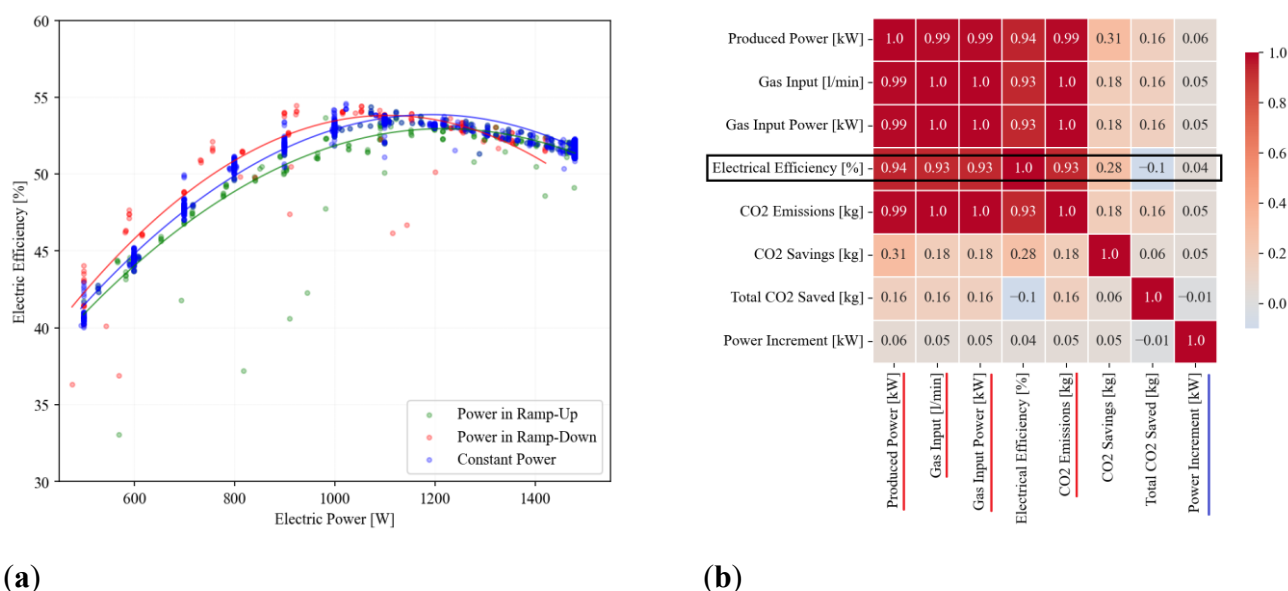


Fig. 71. Polynomial Fits on Efficiency–Power Graph (a) and Correlations on Experimental Data (b).

6.2.3.2 Model Validation

In this work, various algorithms were examined to identify the most accurate model for estimating cell efficiency. Fig. 72 depicts the model prediction curves alongside the actual data for the entire dataset. It is evident that all the models can predict efficiency behaviour smoothly, even during efficiency drops due to partialisation. These drops shown in the graph correspond to timesteps when the SOFC was turned off, resulting in very low-efficiency readings. Despite these peaks, the models generally followed the curve accurately, with some exceptions for models such as the ANN, LSTM, and polynomial regression, which displayed more “constant” behaviour compared to other models that better tracked the real data curve. Additionally, the graph shows that when the SOFC operated initially at constant power, the efficiency remained

constant, and all the models replicated this behaviour.

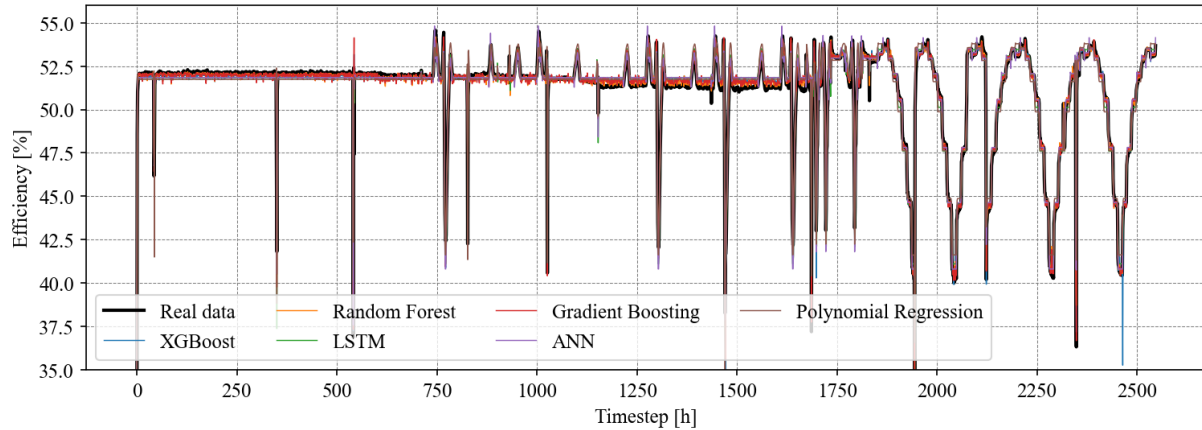


Fig. 72. Models and Real Data Behaviour During the Entire Experimental Period.

However, for the sake of clearness, Fig. 73 illustrates the models' performance in three different sections, from left to right: a nearly constant zone, a ramp-down section, and a highly variable situation. In the first section, it is evident that LSTM and polynomial regression failed to accurately follow the real efficiency behaviour, despite its limited variation. In contrast, the RF and GBoost models successfully replicated the efficiency trend across the 100 h considered in this time window. In the second section, from hour 1600 to 1700, all the models closely aligned with the real data; however, when the cell operated at an almost constant efficiency of about 51%, an error of 0.7% was observed for almost all the models. Nonetheless, RF, GBoost, and XGBoost remained the most accurate. Finally, in the ramp-up and ramp-down behaviours shown in the right section of Fig. 73, all the models, except for polynomial regression, accurately followed the efficiency curve, with decision-tree-based models significantly outperforming the other algorithms.

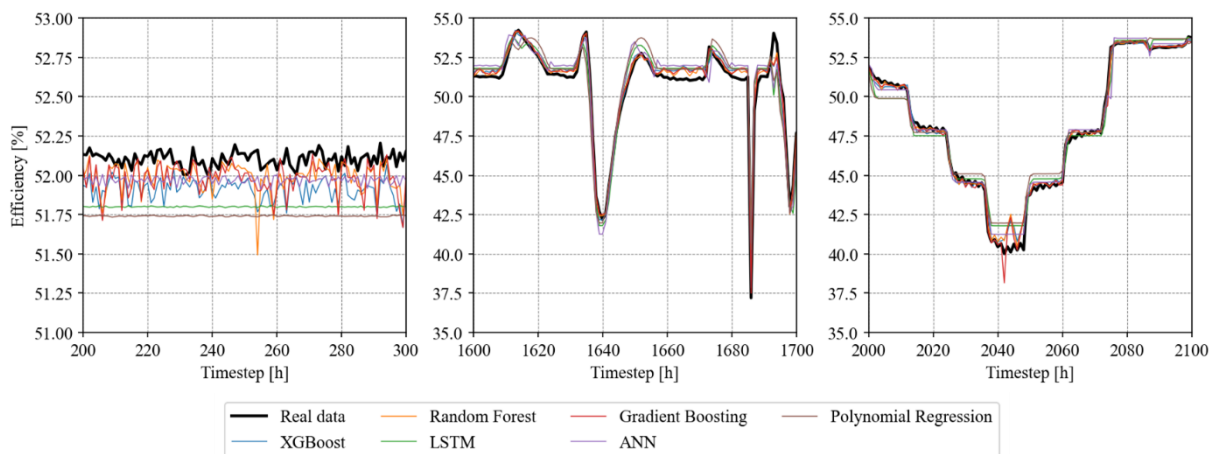


Fig. 73. Models and Real Data Behaviour in Different Time Horizons.

Table 26 summarises the scores for all the models under both test and overall conditions. Test conditions represent the validation scores during the training process, while overall scores

pertain to the complete dataset. Confirming observations from the graphical comparison between the models and the actual data, XGBoost, GBoost, and RF demonstrated superior accuracy compared to the other models. Specifically, these three models achieved a minimum R^2 of 0.98 for the entire dataset, and even when considering only test data, the scores reached 0.98 for RF and 0.97 for GBoost. Similar MSEs were observed for these models, with the lowest values recorded at 0.10 for the RF for the overall dataset and 0.14 during the validation phase. Regarding the MAPE, the ANN reported a value of 0.04, identical to that of GBoost, XGBoost, and RF, where the worst value was recorded for the polynomial regressor. However, in terms of the RMSE, which can be compared to the MAE due to their similar measurement units, its value was 1.7 times higher than that in the RF algorithm and 1.5 times higher than in XGBoost, which recorded an RMSE of 0.36. As anticipated, polynomial regression and the LSTM algorithm were the least accurate, achieving R^2 values of 0.94 and 0.96, respectively, in the total validation. In this study, even considering the WS, tree-based models such as RF, GBoost, and XGBoost outperformed DL methods, likely due to the tree-based models' superior handling of non-linear relationships and interactions within a small dataset. DL models, while generally powerful as highlighted in the current state of the art (SOA), often require larger datasets for accurate training and may experience overfitting, as evidenced by the performance of the LSTM and ANN in this study.

Table 26. Overview of Model Score Results.

Model	MSE		MAE		RMSE		MAPE		R^2		WS	
	Total	Test	Total	Test	Total	Test	Total	Test	Total	Test	Total	Test
XGBoost	0.13	0.27	0.16	0.25	0.36	0.52	0.04	0.04	0.98	0.96	0.03	0.04
RF	0.10	0.14	0.17	0.24	0.31	0.38	0.04	0.04	0.99	0.98	0.03	0.03
LSTM	0.29	0.24	0.36	0.36	0.54	0.49	0.07	0.07	0.96	0.96	0.06	0.05
GBoost	0.11	0.17	0.20	0.25	0.33	0.41	0.04	0.04	0.98	0.97	0.03	0.03
ANN	0.28	0.26	0.34	0.34	0.53	0.51	0.04	0.04	0.96	0.95	0.04	0.04
Polynomial Regression	0.41	0.36	0.45	0.44	0.64	0.60	0.09	0.09	0.94	0.94	0.07	0.07

The primary goal of this paper was to develop a DT capable of self-improvement by utilising the most accurate model based on the current data collected. The accuracy and selection of the best model based on the WS were assessed every 500 h of data collection, simulating an ageing DT. In the final release of the DT, models are expected to be trained almost daily to ensure the best model is available for optimisation purposes. However, for model validation, a time frame of 500 h was deemed more appropriate. For this reason, Table 27 summarises the scores of all

the models, considering WS and R^2 for each time window, revealing the evolution of scores as the dataset increased. Initially, XGBoost was selected as the superior model, achieving a WS of 0.004 with an R^2 of 0.998, surpassing the other models; during this time, the SOFC operated at a constant and maximum power of 1.4 kW. Further analysis indicated that between 1,000 and 2000 h of data collection, the best model was GBoost, with a WS ranging from 0.022 to 0.027 and a peak R^2 of 0.97 encountered after 200 h. Finally, as anticipated from the earlier results, at 2500 h, RF once again proved to be the best model, achieving an R^2 of 0.98 and a WS of 0.031. Regarding ANNs, Table 3 shows that at 1500 h, the R^2 value was 0.6 lower compared to the value at 1000 h. This can be explained by the transition from steady-state operation to a highly variable pattern, where the NN encountered input and output features it could not accurately predict. However, by 2500 h, the network achieved an R^2 value of 0.95. These results demonstrated how the DT framework presented in this paper can be used to develop the most accurate and “ageing” DT capable of precisely replicating SOFC behaviour across different periods. Furthermore, when used for optimisation purposes, this enables the most accurate results based on the current conditions of the FC.

Table 27. Model Scores in Different Periods.

Model	500 h		1000 h		1500 h		2000 h		2500 h	
	R^2	WS	R^2	WS	R^2	WS	R^2	WS	R^2	WS
XGBoost	0.998	0.004	0.876	0.067	0.941	0.038	0.957	0.035	0.975	0.035
RF	0.837	0.085	0.929	0.041	0.943	0.037	0.969	0.028	0.983	0.031
LSTM	0.957	0.030	0.872	0.081	0.883	0.090	0.914	0.080	0.953	0.060
GBoost	0.938	0.034	0.966	0.022	0.958	0.029	0.971	0.027	0.974	0.036
ANN	0.845	0.081	0.961	0.025	0.892	0.062	0.911	0.057	0.950	0.047
Polynomial Regression	0.993	0.009	0.878	0.079	0.859	0.105	0.888	0.099	0.941	0.074

Finally, Table 28 summarises the main findings related to each algorithm presented in this work. It highlights the performance scores of different algorithms and the challenges in fine-tuning the models. For instance, it highlights that ML models achieved better scores and faster training times compared to DL ones.

Table 28. Comparison of ML Models: Advantages, Disadvantages, and Limitations.

Model	Advantages	Disadvantages	Limitations
XGBoost	High accuracy (MAPE = 0.04) Fast training times (0.19 s)	Requires a lot of memory	Possible overfitting if not properly regularised
RF	Highest R^2 (0.99) and WS (0.03) Fast training times (0.54 s)	Poor performance on small dataset ($R^2_{500h} = 0.84$)	Not always suitable for sequential or temporal data
LSTM	Excellent for time series	Very long training times (6.47 s)	Large amounts of data required
GBoost	Robustness as the number of data variates ($R^2 > 0.93$)	Many iterations needed to fine-tune hyperparameters	Can overfit with too many iterations
ANN	Ability to capture complex patterns	High training times (3.72 s)	Large amounts of data for training needed
Polynomial Regression	Very fast training times (0.008s)	Not as accurate as other models	Sensitive to outliers

6.3 Final Discussion and Future Directions

The need to decarbonise the energy sector has never been more pressing, considering that carbon emissions are still increasing globally and the effects of climate change, especially in Europe, are spreading. To meet the goals of the Paris Agreement, the shift toward sustainable energies and technologies must be expedited, and hydrogen technologies such as SOFCs offer remarkable potential. SOFCs allow for high thermoelectric efficiency when supplied by hydrogen and offer the advantages of producing both heat and electricity with zero emissions; additionally, a CO₂ capture system can produce methane by using green hydrogen and playing a very significant role in this transition. Despite SOFC's benefits from the clean-burning properties of hydrogen, as water is the only byproduct, even NG gives SOFCs high efficiency and extremely low emissions compared to conventional ways of power generation. The integration of hydrogen technologies thus supports a dual strategy toward greater energy security and smaller carbon footprints in the energy mix. In this framework, DTs are changing the way the management and optimisation of energy systems—including SOFCs—are being carried out. A DT is a dynamic, virtual synchronised twin of a physical system using real-time data to simulate, predict, and optimise its performance. For SOFCs, a DT can monitor operational conditions in terms of performance prediction or fault diagnosis and enhance efficiency and reliability.

This section was oriented at the development of a DT model for a gas-feed SOFC, using ML techniques to get an accurate efficiency estimation of the system under different conditions.

The approach that was followed during this paper was to first gather four months of data from a real-world setup of an SOFC. Using this data, several deep and ML models, like ANNs and GBoost, were trained to find a reliable model capable of predicting electrical efficiency for the SOFC using operational parameters. The main findings of this study can be summarised as follows:

- The experimental setting of the study considered the measurement data acquisition during 3935 h of SOFC operation, with parameters like electric power output, gas flow rates, and system efficiency, whose correlation revealed strong dependencies used to conceptualise the models' features.
- Validation results, including the MAE, MAPE, and RMSE, revealed that XGBoost, GBoost, and RF algorithms were accurate and fast enough for real-time efficiency prediction within a DT framework.
- RF Regressor was found to be the best regression, with R^2 almost equal to 0.99 and an RMSE equal to 0.31 using the entire dataset. Such high accuracy of the model potentially supports the integration of SOFCs into energy management systems, improving operational efficiency.
- Additionally, the potential of having a DT was demonstrated by observing how the most accurate model evolved over time based on the data collected, transitioning from XGBoost to GBoost, and finally to RF.

In summary, data-driven modelling techniques and RF have numerous advantages in the development of DTs for SOFCs. Accurate and reliable DTs will play a big role in the energy sector, towards a more digitalised and sustainable future. These virtual models optimise operational efficiency but further provide deeper insight into the behaviour of the system, enabling proactive maintenance and reducing downtime. One challenge that could restrict implementation is the use of proprietary software to control the SOFC, which requires an operator to manage the system through closed-source tools. However, the methodology used in this study is not limited to SOFCs and could be applied to other fuel cells, vehicles, or cogeneration systems. This approach would allow for the development of a self-evolving DT capable of calculating system efficiency, supporting optimal management. Furthermore, analysing differences between actual and predicted efficiency could help identify potential system faults. Future works will be aimed at the enlargement of the dataset, covering more operational scenarios and environmental conditions, including the development of fault detection algorithms. In this framework, testing with blends of hydrogen and NG at various concentrations will also be considered in future works to investigate the effects on cell efficiency. Additionally, optimisation will be integrated into the DT to optimise the operation of 4 different SOFCs in an expanded context including PV systems, energy storage solutions, and buildings.

References

- [1] International Energy Agency. CO2 Emissions in 2023. 2024.
- [2] Rabuni MF, Li T, Othman MHD, Adnan FH, Li K. Progress in Solid Oxide Fuel Cells with Hydrocarbon Fuels. *Energies (Basel)* 2023;16:6404. <https://doi.org/10.3390/en16176404>.
- [3] Rahmani M, Maharluie HN. Study of syngas-powered fuel cell, simulation, modeling, and optimization. *Advances in Synthesis Gas : Methods, Technologies and Applications*, Elsevier; 2023, p. 493–531. <https://doi.org/10.1016/B978-0-323-91879-4.00014-X>.
- [4] Ming W, Sun P, Zhang Z, Qiu W, Du J, Li X, et al. A systematic review of machine learning methods applied to fuel cells in performance evaluation, durability prediction, and application monitoring. *Int J Hydrogen Energy* 2023;48:5197–228. <https://doi.org/10.1016/j.ijhydene.2022.10.261>.
- [5] Bozorgmehri S, Hamed M. Modeling and Optimization of Anode-Supported Solid Oxide Fuel Cells on Cell Parameters via Artificial Neural Network and Genetic Algorithm. *Fuel Cells* 2012;12:11–23. <https://doi.org/10.1002/fuce.201100140>.
- [6] Wu Y, Wu X, Xu Y, Cheng Y, Li X. A Novel Adaptive Neural Network-Based Thermoelectric Parameter Prediction Method for Enhancing Solid Oxide Fuel Cell System Efficiency. *Sustainability* 2023;15:14402. <https://doi.org/10.3390/su151914402>.
- [7] Guo Z, Ye Z, Ni P, Cao C, Wei X, Zhao J, et al. Intelligent Digital Twin Modelling for Hybrid PV-SOFC Power Generation System. *Energies (Basel)* 2023;16:2806. <https://doi.org/10.3390/en16062806>.
- [8] Yang J, Li X, Jiang JH, Jian L, Zhao L, Jiang JG, et al. Parameter optimization for tubular solid oxide fuel cell stack based on the dynamic model and an improved genetic algorithm. *Int J Hydrogen Energy* 2011;36:6160–74. <https://doi.org/10.1016/j.ijhydene.2011.02.019>.
- [9] Gong W, Cai Z, Yang J, Li X, Jian L. Parameter identification of an SOFC model with an efficient, adaptive differential evolution algorithm. *Int J Hydrogen Energy* 2014;39:5083–96. <https://doi.org/10.1016/j.ijhydene.2014.01.064>.
- [10] Jiang B, Wang N, Wang L. Parameter identification for solid oxide fuel cells using cooperative barebone particle swarm optimization with hybrid learning. *Int J Hydrogen Energy* 2014;39:532–42. <https://doi.org/10.1016/j.ijhydene.2013.09.072>.
- [11] Wei Y, Stanford RJ. Parameter identification of solid oxide fuel cell by Chaotic Binary Shark Smell Optimization method. *Energy* 2019;188:115770. <https://doi.org/10.1016/j.energy.2019.07.100>.
- [12] Yang B, Chen Y, Guo Z, Wang J, Zeng C, Li D, et al. <scp>Levenberg-Marquardt</scp> backpropagation algorithm for parameter identification of solid oxide fuel cells. *Int J Energy*

Res 2021;45:17903–23. <https://doi.org/10.1002/er.6929>.

- [13] Yang B, Guo Z, Yang Y, Chen Y, Zhang R, Su K, et al. Extreme learning machine based meta-heuristic algorithms for parameter extraction of solid oxide fuel cells. *Appl Energy* 2021;303:117630. <https://doi.org/10.1016/j.apenergy.2021.117630>.
- [14] Wang J, Xu Y-P, She C, Xu P, Bagal HA. Optimal parameter identification of SOFC model using modified gray wolf optimization algorithm. *Energy* 2022;240:122800. <https://doi.org/10.1016/j.energy.2021.122800>.
- [15] Hwang H, Ahn J, Lee H, Oh J, Kim J, Ahn J-P, et al. Deep learning-assisted microstructural analysis of Ni/YSZ anode composites for solid oxide fuel cells. *Mater Charact* 2021;172:110906. <https://doi.org/10.1016/j.matchar.2021.110906>.
- [16] Hwang H, Choi SM, Oh J, Bae S-M, Lee J-H, Ahn J-P, et al. Integrated application of semantic segmentation-assisted deep learning to quantitative multi-phased microstructural analysis in composite materials: Case study of cathode composite materials of solid oxide fuel cells. *J Power Sources* 2020;471:228458. <https://doi.org/10.1016/j.jpowsour.2020.228458>.
- [17] Meng-ting C, Xiao-wei F, Zhong-hua D, Xi L, Xiao-long W, Yuan-wu X, et al. Data-Driven Fault Detection for SOFC system based on Random Forest and SVM. 2019 Chinese Automation Congress (CAC), IEEE; 2019, p. 2829–34. <https://doi.org/10.1109/CAC48633.2019.8997454>.
- [18] Xue T, Wu X, Xu Y, Jing S, Li Z, Jiang J, et al. Fault Diagnosis of SOFC Stack Based on Neural Network Algorithm. *Energy Procedia* 2019;158:1798–803. <https://doi.org/10.1016/j.egypro.2019.01.423>.
- [19] Subotić V, Eibl M, Hochenauer C. Artificial intelligence for time-efficient prediction and optimization of solid oxide fuel cell performances. *Energy Convers Manag* 2021;230:113764. <https://doi.org/10.1016/j.enconman.2020.113764>.
- [20] Kang J-L, Wang C-C, Wong DS-H, Jang S-S, Wang C-H. Digital twin model and dynamic operation for a plant-scale solid oxide fuel cell system. *J Taiwan Inst Chem Eng* 2021;118:60–7. <https://doi.org/10.1016/j.jtice.2021.01.001>.
- [21] Marra D, Sorrentino M, Pianese C, Iwanschitz B. A neural network estimator of Solid Oxide Fuel Cell performance for on-field diagnostics and prognostics applications. *J Power Sources* 2013;241:320–9. <https://doi.org/10.1016/j.jpowsour.2013.04.114>.
- [22] Brownlee Jason. XGBoost with Python: Gradient Boosted Trees with XGBoost and scikit-learn. 2018.
- [23] Palomba V, Ferraro M, Frazzica A, Vasta S, Sergi F, Antonucci V. Dynamic simulation of a multi-generation system, for electric and cooling energy provision, employing a SOFC cogenerator and an adsorption chiller. *Energy Procedia* 2017;143:416–23.

<https://doi.org/10.1016/j.egypro.2017.12.705>.

Chapter 7

Buildings Digital Twins and Models

Nowadays, digitalization is not merely influencing, but rather revolutionizing all sectors, from healthcare to education, from industries to buildings. This digital integration which includes IoT devices, 5G connectivity, and AI applications is pushing the research toward optimal control of indoor equipment to achieve the objective of decarbonisation, cost reduction and well-being improvements. In this context, Eini and Abdelwahed [1] developed a NN-based Model Predictive Control (MPC) approach for building comfort management, focusing on energy consumption and occupants' comfort in smart buildings. Yang et al. [2] proposed a MPC of air-conditioning, lighting, and shading systems, aiming to optimize energy efficiency and indoor thermal and visual comfort reducing up to 20% of energy consumption.

At the same time, individuals are spending an increased amount of time in indoor environments, working both in offices and at home. The use of digital devices such as PCs, smartphones or tablets has become integral to daily life, but while these bring new levels of productivity and efficiency across various applications, prolonged use can lead to discomfort and strain. One of the primary reasons for the growing importance of studying Indoor Environmental Quality (IEQ) parameters is not only its fundamental role in our daily lives but also its potential to enhance the productivity of occupants [3]. IEQ covers a broad spectrum of factors that need to be examined to ensure an optimal comfort state for individuals. It regards the achievement of satisfactory air quality and acoustic conditions, as well as adhering to standards of thermal and visual comfort. The latter, in particular, focuses special attention due to its direct impact on the performance of workers. Enhanced visual comfort can lead to increased focus and productivity, and is indispensable for safeguarding the health and safety of users. In this framework, building DTs, replicating real-world systems [4], will play a key role in this digital and green transition. DTs can increase the efficiency of different energy systems such as DHN or microgrid but they are also used to improve the proactive management of buildings [5].

7.1 Overview of Thermal and Visual Comfort in Buildings

In Chapter 2, the concept and development of DTs for buildings were explored extensively, presenting a thorough review of the current state of the art. For this reason, this section intends to switch attention to the occupants' comfort in indoor spaces, specifically offering an overview of the research related to assessing thermal and visual comfort in buildings. The primary objective is to examine existing works in this domain, while also introducing and discussing the KPIs that influence comfort evaluation. By highlighting the latest findings and methodologies, this section aims to underscore the importance of optimizing indoor conditions for occupant satisfaction and overall building performance.

The recognition of comfort in indoor environments is widely accepted as crucial, directly influencing human well-being, quality of life, and the sustainable functioning of buildings. In particular, the thermal hygrometric is evaluated by considering two important indices: the PMV and the PPD. The first index, proposed by Fanger and adopted by the UNI-EN-ISO 7730 standard [6,7], is a function of the six independent variables on which the terms of the heat balance depend and which is defined on the 7-value scale: metabolic activity, the thermal resistance of clothing, air temperature, relative air velocity, partial vapour pressure and MRT. This latter is the weighted average temperature of the temperatures of the surfaces that delimit the environment including the effect of incident solar radiation. So, the thermal condition is significantly affected by the radiative heat exchange between the human body and the surrounding surfaces [8] (it contributes as far as 30% of the whole thermal exchanges of the subject) [9]. For this reason, it is a very important parameter to be considered when radiant systems are installed in a space [10]. The calculation of radiant heat flux often involves measuring the emissivity and temperature of each surface within the room and utilizing the MRT method, as recommended by ASHRAE [11].

Many researchers investigated on this topic as well by presenting innovative methods [3]. Moreover, many studies presented adaptive control strategies applications based on these indices. Choi et al., [12] introduced a control algorithm that takes into account real-time PMV by utilizing a new model and studied its effects on thermal comfort and energy efficiency. Yu et al. [13] based their control strategy on PMV to determine the best Natural Ventilation and to propose a change-over control strategy. Ren et al. [14] proposed a method by integrating the software TRNSYS with the parametric simulation manager jEPlus to determine the optimal starting and operation mode of a combined radiant floor cooling and ventilation system. To do this, they developed an optimization process constrained by the thermal comfort defined by PMV to minimize energy consumption. Focusing on smart buildings application, Hosamo et

al. developed a DT to better detect faults in Heat Ventilation and Air Conditioning (HVAC) systems [15]. In another study they investigated how to optimal manage HVAC to reduce consumption while satisfying thermal comfort of occupants [16]. Despite DT is a novel technology in the entire energy sector, its applications in indoor comfort evaluation are more limited thus far. For instance, Desogus et al. in [17] presented a framework to develop twins by integrating BIM tools and sensors for building thermal comfort analysis and visualization.

To improve the well-being of building occupants, visual comfort should be guaranteed as well. Khosravi et al. [18] developed a MPC scheme with thermal and visual comfort constraints. This latter was analysed by considering the Daylight Glare Probability (DGP) in a semi-linear support vector regression model. As well Yang et al. developed a MPC system to enable coordinated control of multiple building services for overall optimized energy efficiency, indoor thermal and visual comfort, as well as a hybrid model for predicting indoor visual comfort and lighting power [2]. Shum and Zhong [19] evaluated the impact of various window shade control strategies on the heating and cooling load of cold climate zone test spaces by considering as well the visual comfort conditions. Lee and Song [20] analyzed the visual and thermal environments from various perspectives by taking into account smart windows in residential buildings. in Sejong City, South Korea. In the study of Rabani et al. [21], energy use and thermal comfort and visual comfort were evaluated. They developed an optimization method that integrated building envelope, glazing parameters, HVAC setpoints, shading device, and window opening control methods. The shading and window opening control strategies were implemented using various control methods including the indoor air temperature, the CO₂ level, the daylight level, the wind velocity, and direct solar radiation on the façade. Baghoolizadeh et al. [22] investigated the effect of Venetian blinds on reducing the energy consumption of the building as well as improving the thermal and visual comfort of the building occupants.

7.2 SolarLab Case Study

All the research presented in this chapter were conducted considering the same case study located in the University of Palermo: the SolarLab laboratory (Fig. 74). It is located on the third floor of the building 9 of the Department of Engineering of the University of Palermo. It covers an area of 106 m² and has a height of 4.40 m, inclusive of the false ceiling, and 3.40 m, exclusive of the false ceiling. Along the elongated side, there are four windows (2.40 x 2.90 m) equipped with double-glazing and aluminum frames, lacking any blind system. They are positioned on the southeast side. Opposite them lies a green roof with an average albedo value of 0.25, hosting medium-height plants that obscure two windows. Additionally, a shelter (2.70x19 m) partially

shades the facade.

Within the laboratory, the lighting fixtures include four suspended LED luminaires (each with a power of 49W). The lighting fixtures were chosen to achieve the recommended illuminance values according to the UNI-EN 12464 [23] standard for office environments



Fig. 74. Pictures of the laboratory

Regarding the instrumentation used to measure the values (Fig. 75), Delta Ohm HD have been installed in the laboratory to monitor the illuminance and luminance values during the experimental campaign. consumption has been measured using SIEMENS SENTRON Power Monitoring Device PAC3200, an instrument for measuring electric power distribution variables such as voltage, current, power, electrical work and frequency. Finally, the outdoor illuminance has been measured by a Delta Ohm LP PHOT 02 luxmeter equipped by a “shading band” in order to cover the direct contribution of the daylight.



Fig. 75. Pictures of the instrumentation used to collect the measures

7.2.1 SolarLab Energy Model

In the context of this thesis, an energy model of the laboratory was developed. This allowed for initial investigations on thermal and visual comfort behaviour while waiting for new sensors to be installed including Globe Thermometer (GT), drybulb and wetbulb temperature sensors, air velocity sensor and thermal-flux meter. All of these will be used to increase the robustness of the developed model and to create an advanced system for optimally controlling the HVAC plant.

EnergyPlus was selected as energy simulation software for its widespread use in the current literature [24]. Starting from the envelope, it is composed of elements shown in Table 29. As can be seen, since the building was built in the 1960s, its characteristics are not enough efficient as required by current legislation. However, the solar shelter was designed to reduce opportunely the solar contribution while guaranteeing an appropriate angle for the installation of the hybrid solar PV/th system. This laboratory is normally occupied by students from 8:00 AM to 6:00 PM while the maximum occupancy is normally reached at 11:00 AM and 17:00 PM. Regarding electric equipment, the main contribution is due to the presence of PCs or electric devices inside the room which leads to an increase in internal gains in combination with the presence of people. The lighting system is composed of four pendant luminaires equipped with LED sources.

Table 29. Information of SolarLab thermal envelope

Building Element	Number of Layers	Thickness	Thermal Resistance	Thermal Transmittance	U_{lim}
	[-]	[m]	[m²K/W]	[W/m²K]	[W/m²K]
Exterior Wall	4	0.355	0.711	1.406	0.41
Interior Wall	3	0.12	0.584	1.713	-
Floor Slab	4	0.35	0.472	2.117	-
Roof Slab	5	0.36	0.499	2.004	0.32
Window (double-glazing)	3	0.012	0.256	3.905	2.5

This information was included for modelling the laboratory in EnergyPlus [25], including the schedule for each internal load, the stratigraphy of each element of the building envelope and finally the HVAC and lighting system. Fig. 76 shows a detail of the modelled SolarLab thermal zone in Sketchup.

The room borders two sections that correspond to rooms without air conditioning: a stairwell and a space with elevator machinery. It is important to note that the Southwest façade and the floor are treated as adiabatic, given that they are adjacent to rooms with air conditioning maintaining identical internal conditions. To better simulate temperature cycles in the ceiling

roof due to the HVAC use, an accurate model of the cooling and heating ceiling was needed. A boiler was assumed to distribute hot water for heating at 35°C and a chiller to produce chilled water at 15°C. In addition, an air ventilation system was incorporated in the model to prevent potential issues related to high relative humidity and condensation during the summer months. As a result, during the whole heating season, the temperature of the roof was maintained around 30°C, while during summer this temperature remains around 22°C, even considering that the external surface (the roof) temperature exceeds 45°C for several hours a day.

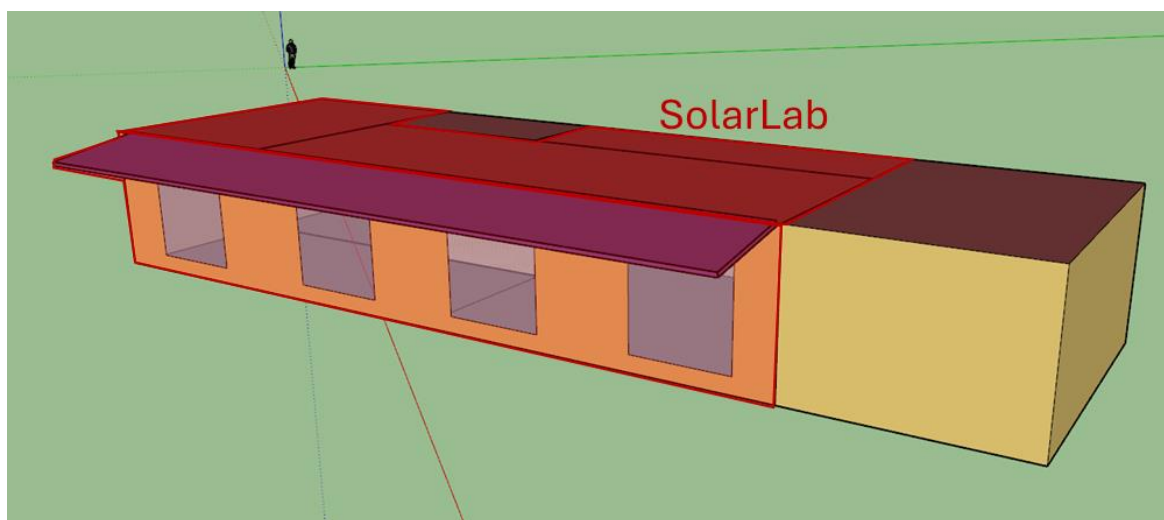


Fig. 76. Three-Dimensional Representation of the EnergyPlus Model

7.3 Digital Twin of SolarLab to Improve Visual Comfort

Numerous studies explored predictive lighting control and thermal comfort conditions, yet the application of DT technology in enhancing visual comfort remains almost underexplored. The utilization of a DT of a facility offers a dynamic understanding of real-time comfort parameters, while available studies focus only on the potential energy-saving [26]. To fill this gap, this section presents a DT of the real laboratory SolarLab aiming to investigate in real-time the most valuable visual comfort parameters inside the lab. More importantly, it functions as an interactive service, potentially guiding end-users through a dashboard with tailored suggestions for optimal comfort. Aligning with the current SoA, illuminance will be monitored through the use of an experimental facility. Crucially, as a DT should accurately mirror and synchronize with its physical counterpart, it holds the potential for automated adjustments in lighting parameters, moving towards a more personalized comfort experience. This concept aligns with the emerging research in Personal Environmental Comfort systems, particularly in the domain of visual comfort. In this context, a DT could be useful to achieve individualized optimal conditions for each user.

7.3.1 Methodology

To optimize visual comfort within the laboratory environment, a DT was constructed to enhance occupants' awareness of their comfort levels. The DT is designed to display a range of parameters, including Daylight Factor (DF) and uniformity of illuminance. Furthermore, it employs rule-based algorithms to proffer suggestions that guide users back to optimal comfort conditions. These recommendations are tailored to individual preferences, thanks to the precise placement of sensors. Notably, the DT possesses the capability for self-updating, refining its functionality based on new user data. The subsequent section provides an in-depth elaboration on the development process and intricate features of the DT.

7.3.1.1 Digital Twin Architecture

The concept of DT, first introduced by Grieves et al. in 2007, has undergone significant refinement and clarification over time. The architectural and framework specifics of a DT have been comprehensively delineated by the DTC whitepaper [4]. This document provides an array of DT architectural examples and enumerates the essential components required for their construction. It surely represents an incisive resource for understanding the evolution and current state of DT technology. In particular, drawing from the conceptual framework outlined in [27], DTs can be categorized into six distinct maturity levels. A model that integrates real-time data to enhance operational efficiency is classified as a Level 3 DT. This aligns with the concept of an Operational Twin, defined as a Level 2 (out of 4) DT in [28].

The DT developed in this study is based on real-time data sourced from illuminance and luminance sensors within the laboratory, feeding this information into models executed via LabView software [29]. LabView serves as both the interface (front-end) and the processing platform (back-end), enabling real-time computation of comfort parameters. Based on these calculations, the DT generates customized recommendations displayed on a user dashboard. The interfacing of sensors with the computational system, where the DT is developed, utilizes a wired connection instead of relying on 5G infrastructure or cloud-based solutions.

This setup ensures continuous user awareness of their environmental conditions and allows manual interaction with the lab infrastructure, such as adjusting lighting or window shading. Furthermore, a simulation model of the laboratory was created in DIALux [30] and in DAYSIM [31] to pre-emptively assess the impact of various solutions on energy efficiency and comfort parameters.

This simulation model serves as an experimental tool to explore the efficacy of different interventions without physically intervening in the laboratory. It can be strategically employed for predictive control, such as determining the optimal movements of automated shading

systems to reduce cooling consumption during summer, while simultaneously considering visual comfort parameters. Additionally, the model capabilities extend beyond the current laboratory sensor infrastructure. E.g., utilizing Palermo Climate Data within these models, it facilitates a more comprehensive investigation that surpasses the limitations of the KPIs measurable uniquely by the existing sensors. The high-level structure of the developed DT is shown in Fig. 77.

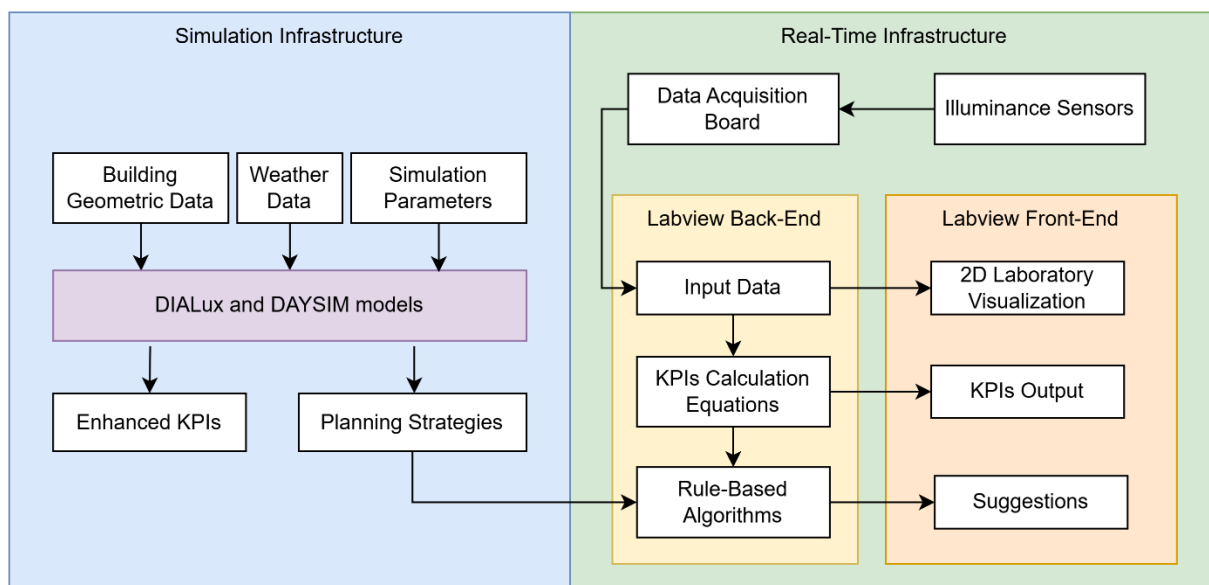


Fig. 77 Architecture of DT of the Laboratory

7.3.1.2 Digital Models of the Twin

In the context of lighting design and analysis, various KPIs play a crucial role in assessing visual comfort and lighting quality. These metrics, described in detail in [32], include DF [33], Daylight Autonomy (DA) [34], Spatial Daylight Autonomy (sDA) [35], Continuous Daylight Autonomy (csDA) [36], uniformity of illuminance, Discomfort Glare Probability (DGP) [37], Color Rendering Index (CRI) [38], etc.. Each KPI serves a unique purpose, contributing to the overall assessment of lighting systems and their impact on visual comfort, occupant well-being, and energy efficiency. In the context of this study, the limitations of the available illuminance and luminance sensors in the case study necessitated a two-pronged approach. Initially, KPIs feasible for real-time computation through the LabView interface were identified (i.e. illuminance on the workplanes, vertical illuminance on the desktop, uniformity, DF, and luminance). Subsequently, a virtual replica of the laboratory was constructed using Dialux and Daysim software to simulate other parameters to be included in DT (better following described).

The illuminance levels are monitored against two threshold values to identify instances of inadequate lighting, in accordance with UNI EN 12464-1 standard [23], as indicated in Equation (36).

$$450 \text{ lux} < E_i < 550 \text{ lux} \quad (36)$$

Furthermore, the legislation does not define a specific illuminance threshold for glare discomfort, except for the Unified Glare Rating (UGR), calculated as shown in Equation (37), with the threshold for the laboratory set at 19.

$$UGR = 8 \log_{10} \left[\frac{0.25}{L_b} \sum_{i=1}^n \frac{L_i^2 \omega_i (P_i)^2}{p_i} \right] \quad (37)$$

Where L_b is the background luminance, L_i denotes the luminance of the i th luminaire, ω_i is the solid angle subtended by the i th luminaire at the observer's eye, P_i represents the Guth position index of the i th luminaire, p_i is the perceived luminance of the i th luminaire, and n signifies the number of luminaires or light sources considered in the calculation.

Moreover, if illuminance and Correlated Colour Temperature (CCT) values fall within the comfort zone delineated by the Kruithof curve [39] (Fig. 78), users could receive a comfort notification on the dashboard. As it is possible to see, this curve, developed by using psychophysical data, collected by Dutch physicist Arie Andries Kruithof, includes illuminance values and CCT values in general considered as comfortable for the users. So, it is useful to understand the comfortable conditions according to the CCT and the illuminance. E.g., in a typical office space with illuminance levels of about 500 lux the suggested CCT to have a pleasing condition should be between 3000 and 6000 K. By considering daylighting, it has a CCT of about 6500 K of 6500 K. If the illuminance values are around 100 lux, the lighting will appear bluish.

In such instances, users could either concur with the assessment or interact with the twin to indicate discomfort, leading to an automatic adjustment of the comfort zone based on user preferences, thereby updating the DT.

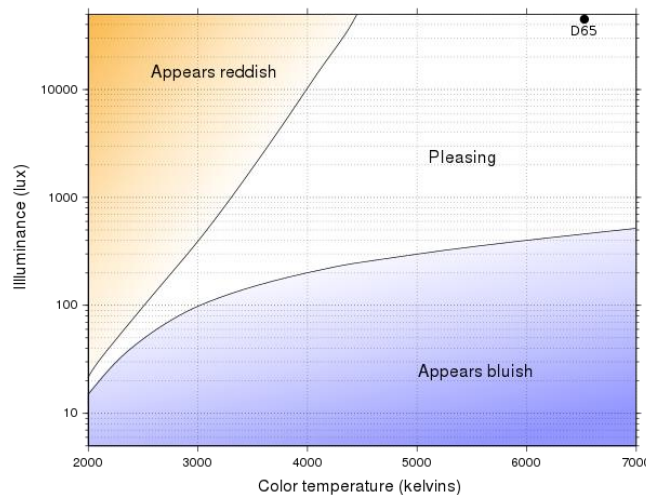


Fig. 78. Kruithof Curve

DF is quantified as the ratio of internal to external illuminance on an overcast day, expressed as a percentage, as per Equation (38):

$$DF = \frac{E_i}{E_o} \times 100\% \quad (38)$$

where E_i denotes the indoor illuminance under diffuse sky conditions, and E_o is the outdoor illuminance under overcast sky conditions

Finally, uniformity of illuminance U evaluates how evenly light is distributed within a space. It is often expressed as a ratio, comparing the minimum illuminance level to the average illuminance level as in (39).

$$U = \frac{E_{min}}{E_{med}} \quad (39)$$

E_{min} denotes the minimum illuminance level within the specified area, and E_{med} is the median illuminance level within that same area.

However, in this first implementation, only DF and U are integrated into LabView for real-time estimation, with individual workstation values displayed on the dashboard. The “recommendations” are generated through rule-based algorithms that utilize calculated KPIs values, according to (36), to propose various strategies for optimizing indoor comfort. For example, in scenarios where desk board illuminance is excessively high, potentially impacting eye health, the algorithm advises a 20% increase in shading system coverage, visible in the user interface. Contrarily, in cases of insufficient illuminance, the laboratory temperature and flux-controllable LED lamps are engaged. The algorithm suggests to interact with lamps in order to address the issue. These instructions are then displayed on the dashboard for user action.

As before mentioned, the Labview input, the *back end* of the twin, incorporated a series of parameters which rely on sensors installed indoors and outdoor. By connecting the measurement data through functions labelled as “Multiply”, “Divide”, and “Add”, it was possible to apply equation (38) for both the workplans. To implement equation (39), the “Min” function was employed to identify the lowest illuminance value, which was then divided, using “Divide” function by the average illuminance calculated as the total of all measurements, achieved using the “Add” function, and subsequently divided by the count of inputs fixed as a “Numeric Constant” function. Moreover, comparative operators including “Greater?” and “Less?” functions were both utilized with Boolean operators and a “Case Structure” for implementing rule-based algorithms. The latter was particularly important for enabling multiple Boolean outputs, such as providing recommendations or altering the colour of a fictional LED to indicate whether the visual comfort condition meets the setpoint criteria presented in equation (1). Ultimately, these outputs were all displayed on the Labview front-end panel, serving as a

control dashboard (Fig. 79). This interface included a specific led indicator for each workplane, describing comfort conditions, and illuminance on workplanes or additional sensor locations (Plane 1, Plane 2 and Plane 3). It also presented data regarding external conditions (i.e. diffuse solar radiation and outdoor diffuse illuminance) which were also used for calculating DF, displayed for both desks.

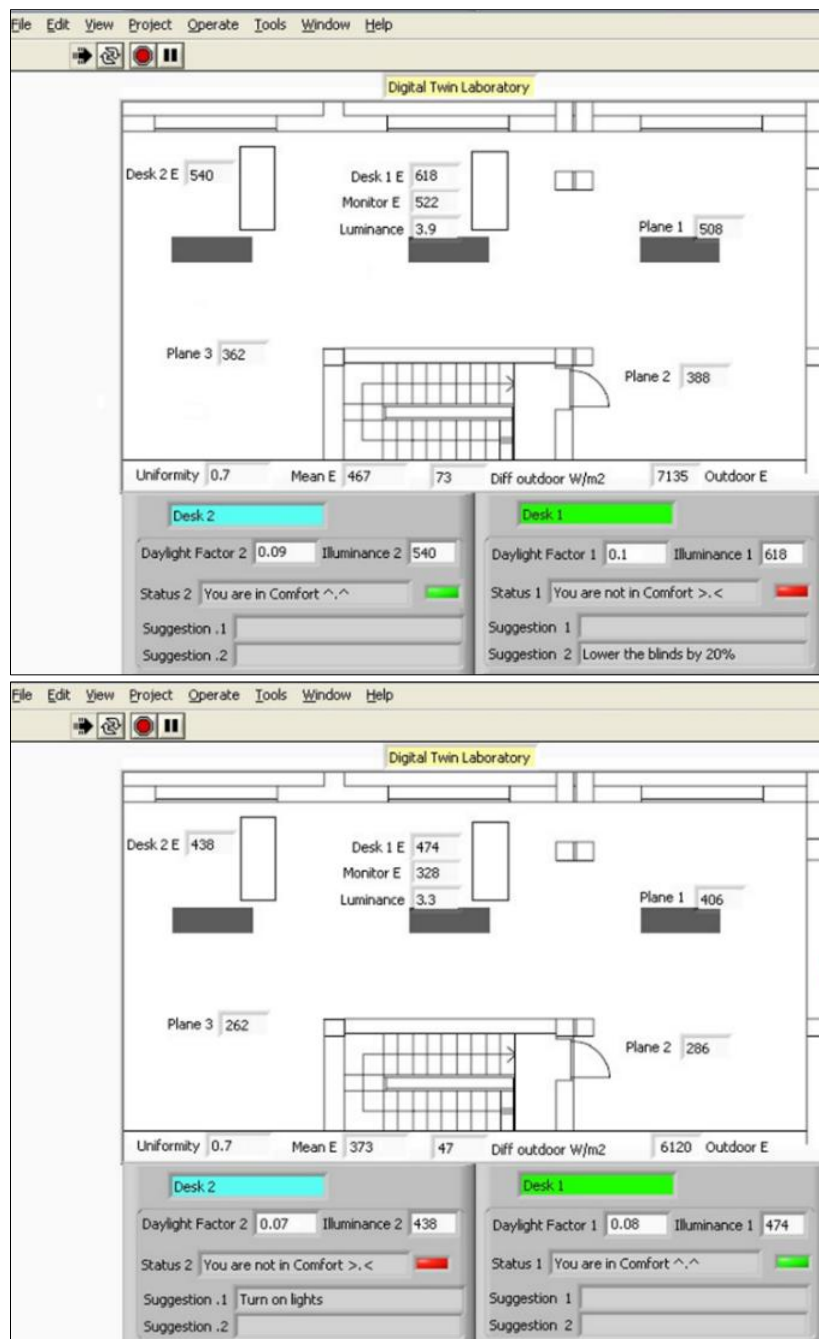


Fig. 79. Examples of LabView dashboards of the Laboratory DT.

In the simulation models created using DIALux and DAYSIM, a wider spectrum KPIs can be examined within the laboratory, encompassing metrics such as DA, sDA, csDA, and DGP. For instance, initial calculations were conducted through the use of the lighting simulation software DIALUX (Fig. 80) to facilitate the design of the lighting system. However, real-time

monitoring of these diverse parameters in a physical laboratory setting would necessitate the installation of numerous sensors, significantly escalating the cost of the experimental setup. Within this context, these tools emerge as an invaluable asset, not only in evaluating the impact of design choices on indoor comfort (such as selecting appropriate lighting fixtures or determining the positioning and angle of roofing) but also in exploring novel predictive strategies. For instance, it can be decisive in assessing how shading should be implemented to minimize cooling loads in summer, without compromising comfort. Following this analysis, the newly derived strategies can be effectively visualized by updating the rule-based algorithm and visualized through the LabView interface.

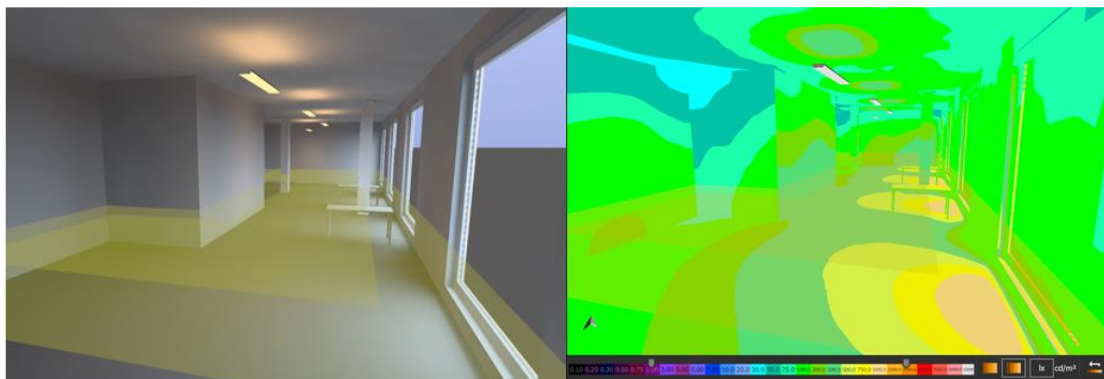


Fig. 80. Pictures of the Dialux model with real colours (left) and false colour (right)

7.3.2 Results and Findings

This initial version of the laboratory DT incorporates basic yet promising functionalities, indicating the significant potential of this technology while also underscoring the need for further enhancements. Displayed in Fig. 79 is the dashboard of the twin real-time component, developed using LabView.

The preliminary focus is on two desks: one used by an individual operating a PC, and the other by a user primarily engaged with books and documents. For the PC user, the dashboard displays all previously mentioned parameters, including also luminance, alongside comfort status information. For the second desk, similar parameters are assessed, excluding luminance and monitor illuminance, with corresponding suggestions provided.

Currently, the DT offers two suggestions related to illuminance. Dialux simulations have shown that a 20% reduction in blind elevation can effectively lower desk illuminance, thereby decreasing glare probability. Consequently, if desk illuminance exceeds 550 lux, a suggestion prompts user interaction to lower the blinds. Conversely, if illuminance falls below 450 lux, users are alerted to activate lights, which are sensor-controlled to adjust flux for desired illuminance levels. The situations under discussion are depicted in Fig. 79, where two distinct

scenarios are presented.

In the right image, Desk 2 experiencing insufficient illuminance levels conversely, the left image illustrates a case where in Desk 1, the illuminance surpasses the threshold of 550 lux. In each instance, the DT provides suggestions, offering users guidance to address their discomfort based on the specific lighting conditions observed at their respective workstations.

Future simulations with the DAYSIM and DIALux-developed model aim to uncover more impactful strategies influencing heating or cooling loads, considering thermal comfort. Additionally, rather than basing individual comfort solely on illuminance, these simulations will incorporate a range of KPIs to provide a holistic assessment of visual comfort enhancements.

7.4 Visual Comfort and Energy Efficiency Evaluation Using Shadings

Transparent envelope plays a key role in the energy balance of the buildings and in the indoor environmental conditions (e.g. visual and thermal comfort). Windows can provide the correct contribution of natural lighting and, during colder seasons, an advantage in terms of solar gain. Many studies presented in the literature investigated the influences of transparent envelopes and the role of shading systems as well as innovative ones. E.g. Rizi and Etlaweel [40] evaluated an adaptive facade towards visual and thermal comfort by using environmental parametric tools in the location of Iran. Shi et al. [41], investigated on the visual-thermal comfort and energy performance of a hydrogel-based thermotropic glazing with different transition temperatures. In particular, they compared them with conventional double-clear glazing and low-emissivity double-glazing under five climatic conditions in China. In this framework, this study aims to propose a methodology that combines different tools to analyse the visual comfort conditions by applying different shading systems by evaluating both visual and thermal comfort, and by considering energy consumption. This novel approach includes the development of a python script for creating custom schedules for controlling shading systems in EnergyPlus, in addition, it also presents a method to implement roller blinds in the model.

7.4.1 Methodology, Models and Scenarios

The analysis reported in this section was performed by using a combination of different tools. In particular, the first model was developed by using Sketchup combined with Euclid extension. Secondly, the base case model was detailed and defined using Energy+ software by using the idf editor to specify the detailed thermal characteristics of the building, the schedules, and HVAC system. Being the aim of this work the assessment of the visual comfort condition and the power absorbed by the lighting system, in relation to the solar gain contribution and the

relative use of the solar cooling system, three different scenarios were analysed (described in the next section).

The analysis was made in three different points of the room where the workplanes were located. The first simulation was aimed at simulating the illuminance values on the workplanes. The tool Python was used to develop a script able to set custom schedules. These detailed the hourly functioning of shading systems on each different day for each scenario. Since using Energy+ Version 8.9 does not enable the possibility of choosing how much window is covered by the roller blinds shading system, these were simulated by using fixed shading objects, which partially cover the window differently.

The methodology presented in this work is schematized in Fig. 81. A Base Case model was developed to study the thermal demand, the power absorbed by the lighting system and the illuminance maps. These were used to develop, by using Python, detailed schedules of the functioning of shading systems to avoid glare issues in the three different workplanes. These schedules were the core of Scenario 1 which was further divided in Scenario 1a, 1b and 1c depending on blinds' slat angles. Finally, Scenario 2a allows, by modifying Base Case with fixed shading objects, the investigation of solar gain and illuminance maps.

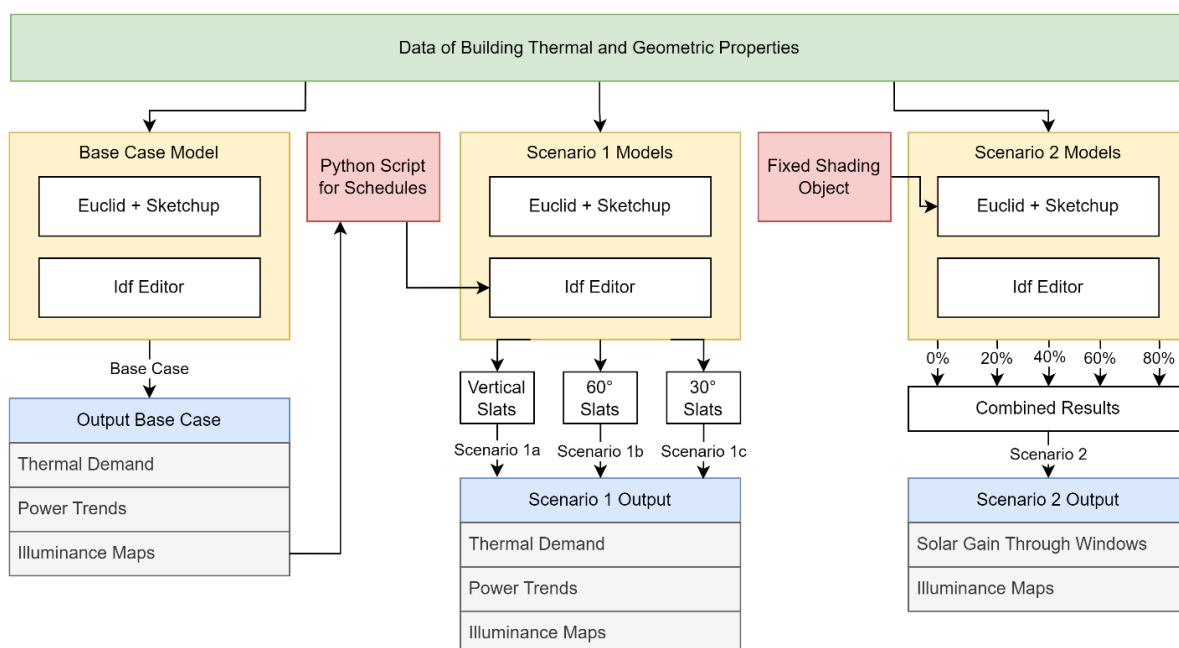


Fig.81. Flow Diagram of Methodology for Investigating Effects of Shading Systems.

As before mentioned, the analyses of different scenarios (Fig. 82) assessed the impact of various shading systems on visual comfort, thermal comfort, and energy efficiency within the laboratory setting. In the Baseline Case, used for the comparison, the laboratory shading systems are considered open. In this scenario no shading system was included, so, it provided

an unobstructed view of daylight contribute on the space without any interventions to mitigate glare or adjust solar gains (offering insight into the laboratory inherent energy and lighting dynamics).

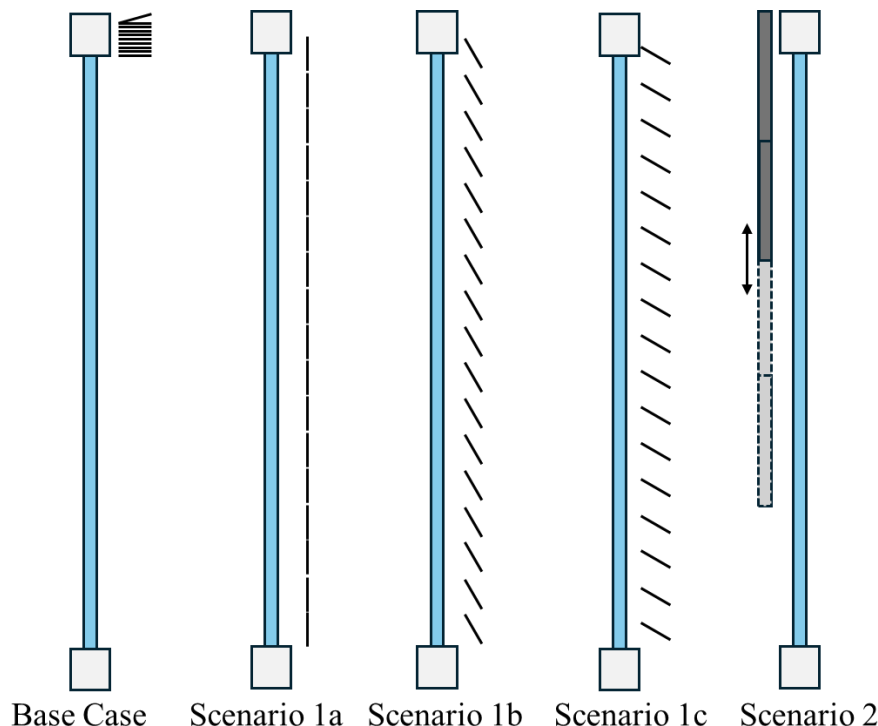


Fig. 82. Schema of shading system layout for each scenario.

Scenario 1 considered a proactive approach to managing glare with the use of slatted blinds that completely cover the windows. The activation of these blinds was determined by a custom schedule activated by the possible presence of glare (according to the illuminance level). Notably, the slat angle undergoes different adjustments, from 90° , which means slat parallel to the glass, to 30° on the vertical, aiming to optimize the balance between natural light utilization and glare reduction. In particular, Scenario 1a considered a slat angle of 90° , Scenario 1b considered a slat angle fixed to 60° and Scenario 1c a slat angle fixed to 30° .

Scenario 2 explored a different shading strategy by employing a fixed shading object to simulate the effect of roller blinds. Contrary to Scenario 1, the blinds can partially cover the windows, with coverage ranging from 0 to 80%. This approach assessed the impact of partial window coverage on glare mitigation and its subsequent effects on energy efficiency and IEQ. Together, these scenarios provide a comprehensive examination of various shading interventions and their role in optimizing energy performance and occupant comfort in the laboratory. Through comparative analysis, the study identifies effective strategies for integrating shading systems into building designs to achieve sustainability goals.

7.4.1.1 Energy Plus Modelling Details

As said, to develop the geometry of the energy model the software Sketchup with the Euclid extension was used. It facilitated the development of a detailed model with thermal and geometrical characteristics. In this step, as well the positioning of windows and doors was specified. The Euclid extension provided capabilities for setting the boundaries of each surface, enabling the classification of surfaces as adiabatic (adjacent to other surfaces within similarly air-conditioned spaces) or as directly exposed to the sun, covered, among others. This level of detail extended to the consideration of shading (e.g., the impact of projecting roofs can be modelled through the "Shading Group Zone" feature available within the Euclid extension).

Upon the development of the model, it can be transformed into an .idf file, which is then subject to further refinement using the EnergyPlus IDF Editor in order to define the analytical phase from geometric and physical modelling. It was useful to evaluate the energy performance of the building under various conditions. Through this process, from defining the building geometry to refining the model in EnergyPlus, a detailed representation of the building energy dynamics was achieved. In particular, utilizing the IDF Editor it was important to outline the building construction elements and the materials constituting each component. The model was implemented by specifying material properties, which vary by country and location, necessitating the input of thermal characteristics like conductivity, density, specific heat and, as well, materials, the structural aspects of the building, including walls, roofs, floors, and windows, are assembled using various material layers. Subsequently, shading systems, such as low-reflectance blinds, were designed and integrated into the EnergyPlus model.

For HVAC systems, a cooling roof was employed as the emission source, complemented by an air-to-water chiller for cooling and a boiler for heating. The operational setpoints were set according to the UNI 11300 standards (temperatures was set at 26°C in summer and 20°C in winter) [42]. After establishing the building components, the operational schedules for each element were determined, covering the HVAC system, blind activation, occupancy, and lighting. As outputs of the simulation, an illuminance map and a grid assessing illuminance levels were chosen.

7.4.1.2 Custom Shading Controls and Lighting Energy Demand Assessment

One of the main contributions of this work is the development of a methodology aimed at optimizing the scheduling of shading devices based on the illuminance at various locations, such as different workplanes within a room. EnergyPlus options give the possibility to sei scheduling and control options. Nevertheless, it relies on average values for reference points. E.g, it permits the closure of shading devices like diffusive shades or slatted blinds based on

predetermined glare indices or solar radiation levels but does not allow for adjusting the slat opening or considering illuminance across multiple workplanes.

To across these limitations, in this study, two distinct strategies were adopted. Initially, an hourly illuminance map was generated using a baseline model with inactive shading devices to assess the impact of solar radiation and illuminance at three specific points. This facilitated the identification of hours during the year when illuminance exceeds 2000 lx, a threshold established to prevent glare for occupants. This step involved creating a Python algorithm to compile a DataFrame marking each hour with the presence or absence of glare over the year.

Precisely, the algorithm in Fig. 83 creates a text file, including a text formatted to meet EnergyPlus requirements for the planned schedules. In this work, as specified, the algorithm checks illuminance values at three points and closes the shading in cases of glare, but potentially it could be used for more complex and flexible control strategies. Unfortunately, by using Energy Plus is possible to write a limited number of rows and characters. For this reason, in the schedule, consecutive hours and days were merged when identical patterns were identified.

```
# Sort the DataFrame by DateTime
df_dates.sort_values('DateTime', inplace=True)

# Initialize the schedule string
schedule_str = "Schedule:Compact,\nMhantz, \nFraction,\n"

# Initialize variables for day control
current_day = None
complex_field_count = 1

# Iterate over the DataFrame
for index, row in df_dates.iterrows():
    day = row['DateTime'].strftime('%m/%d')
    time = row['DateTime'].strftime('%H:%M')
    value = row['Value']

    # If it's a new day, start a new Through section
    if day != current_day:
        if current_day is not None:
            # Complete the previous day until midnight
            schedule_str += f"Until: 24:00, 0.0,\n"
            complex_field_count += 1
        schedule_str += f"Through: {day},\nFor: AllDays ,\n"
        complex_field_count += 2
        current_day = day

    # Add the Until line for the current time
    schedule_str += f"Until: {time}, {value},\n"
    complex_field_count += 1

# Complete the last day until midnight
schedule_str += f"Until: 24:00, 0.0,\n"

# Print the first lines of the schedule for verification
print(schedule_str[:500])

# Write the schedule to a text file
with open('schedule.txt', 'w') as file:
    file.write(schedule_str)
```

Fig. 83. Section of the Script for Creating Custom Schedule.

Once the schedule was generated, this study proposes two different methods to mitigate glare: cover the entire window with slatted blinds partially obscuring the window with a surface

that simulates the employment of roller shutters. While the former (slatted blinds) can be readily implemented in EnergyPlus, allowing for control over blind activation and slat angles, the latter (roller shutter) is not directly supported by the software. To approximate the effect of roller shutters, a static shading device was simulated to partially obstruct the window. Then, the base case was simulated with five levels of shading coverage, from 0 to 80% in 20% increments, to evaluate solar heat gain through the windows and update the illuminance map.



Fig. 84. Energy Plus Model with Shadings Covering the 40% of the Windows

Based on values of illuminance measured in the workplane, a simplified algorithm was employed to calculate the electric energy required to reach the threshold value on the workplane according to the technical standard [23] (e.g. 500 lx). In particular, the luminaires were considered switched on if the illuminance values were lower than 500 lx (as well as in moments where shading devices were closed).

Once the lamps were chosen and the related electric demand was estimated, then a logic ON-OFF was used to calculate hourly the electric demand of the lamps. Finally, since, for heating and cooling needs, different inputs could be required, such as natural gas for boilers or electricity for air-water chiller, it was necessary to convert them into PE to achieve comparable results. In particular, for assessing the PE related to natural gas and electricity were considered the conversion factor available in [43].

7.4.2 Results and Findings

This chapter presents the findings from the analysis of energy efficiency and visual comfort trade-offs. Utilizing the EnergyPlus software, this research delves into the balance required to maintain comfortable indoor environments while optimizing energy use. As said, three different scenarios were considered, each offering insights into different aspects of building performance, particularly focusing on the power absorbed by the lighting system, the contribution of solar

gains, and the utilization of the solar cooling system.

7.4.2.1 Base Case

The temperature and humidity values, both calculated outside and inside the zone, are plotted in Fig. 85. As can be seen, when the HVAC system is switched on, an operative temperature of 20°C in winter and 26°C in summer is maintained. During night, when the laboratory is considered unoccupied, a maximum of 34°C and a minimum of 13 °C were simulated. As can be seen, since the laboratory was considered occupied by a maximum of 3 to 5 people, even without humidity control, most of the time the relative humidity was around 50%.

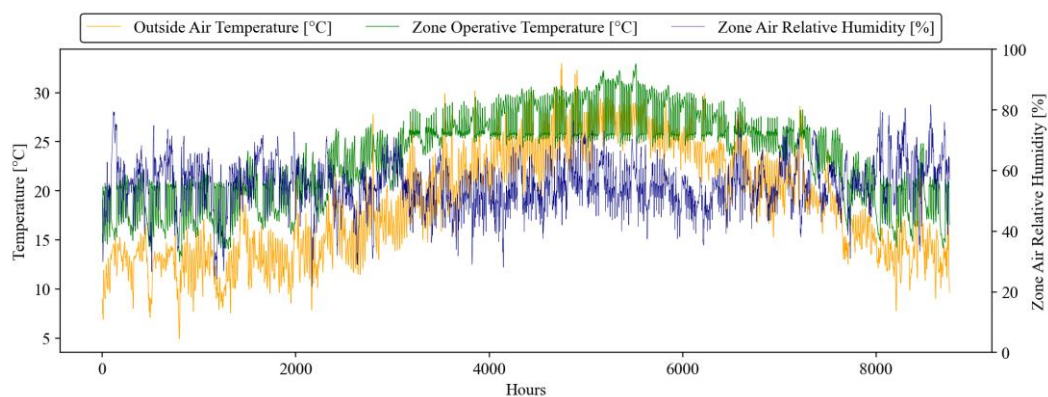


Fig. 85. Indoor and outdoor conditions in Base Case

Regarding thermal loads, the contribution of heating energy was lower than cooling as expected in a hot climate as Palermo (Fig. 86). Additionally, a contribution for solar radiation was significant in winter since the windows are exposed to the south-west direction. Of course, the shelter mitigated the quantity of radiation in summer. Moreover, it can be noted that the reduced heating energy was also related to the contribution of solar gain through the window, which can reach almost 6 kW in winter; while in summer, thanks to the shading roof and the higher sun, this contribution was even lower. The peak of heating demand estimated was 9.9 kW for heating and 11.4 kW for cooling purposes, demanding respectively an amount of energy of 2111 kWh and 7372 kWh during the entire year.

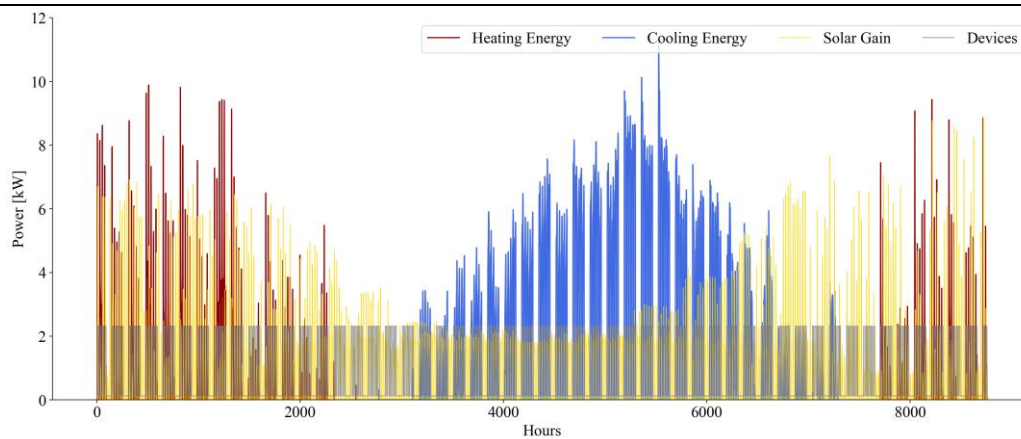


Fig. 86. Energy behaviour in Base Case.

7.4.2.2 Scenario 1

In the Scenario 1 cases, namely Scenario 1a, 1b and 1c, to mitigate glare issues, slatted blinds with different slat angles were implemented. In particular, the blinds were considered closed when illuminance on three selected workplanes in the laboratory exceeded 2000 lux. Based on this criterion, the blinds closed upon higher sunlight intensity, consequently reducing also solar heat gains and increasing and reducing respectively heating and cooling demand, as shown in Fig. 87. Regarding the winter period, the initial heating energy demand in the Base Case was similar to the demands related to the other scenarios (primarily influenced by the overnight heat loss). Throughout the day, in all scenarios, heating energy was slightly higher than the Base Case since lower solar gains due to shading were also observed but contributions from occupants and devices partially compensated the diminished solar input.

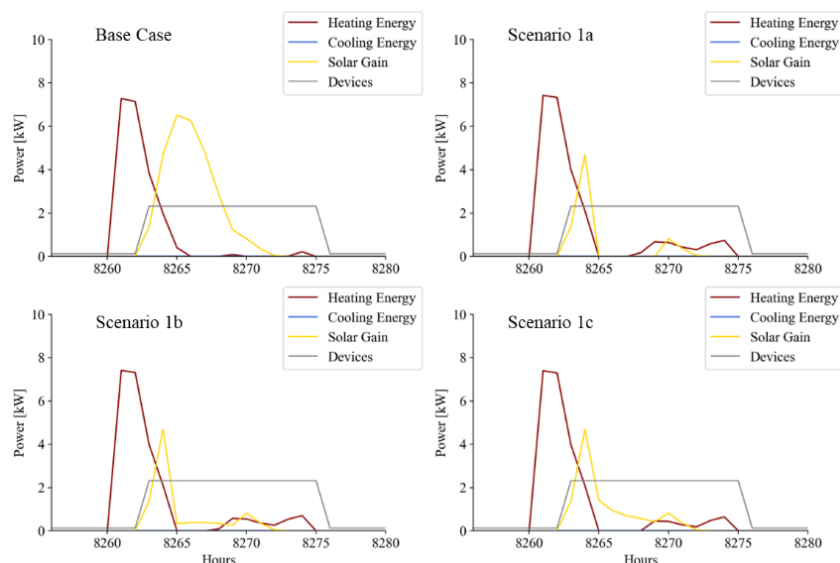


Fig. 87. Comparative Analysis of Hourly Energy Contributions in a Winter Day.

However, the summer day analysis in Fig. 88 revealed a reduction in cooling energy. This is because, after 11 AM, hour 4115, the higher solar gains were mitigated by the external shading

roof, which obstructs direct sunlight. Nevertheless, the effect of shading in the morning reduced cooling energy, achieving an annual saving of approximately 306.6 kWh in Scenario 1a. The reduction of solar heat gain through the window influenced cooling energy consumption.

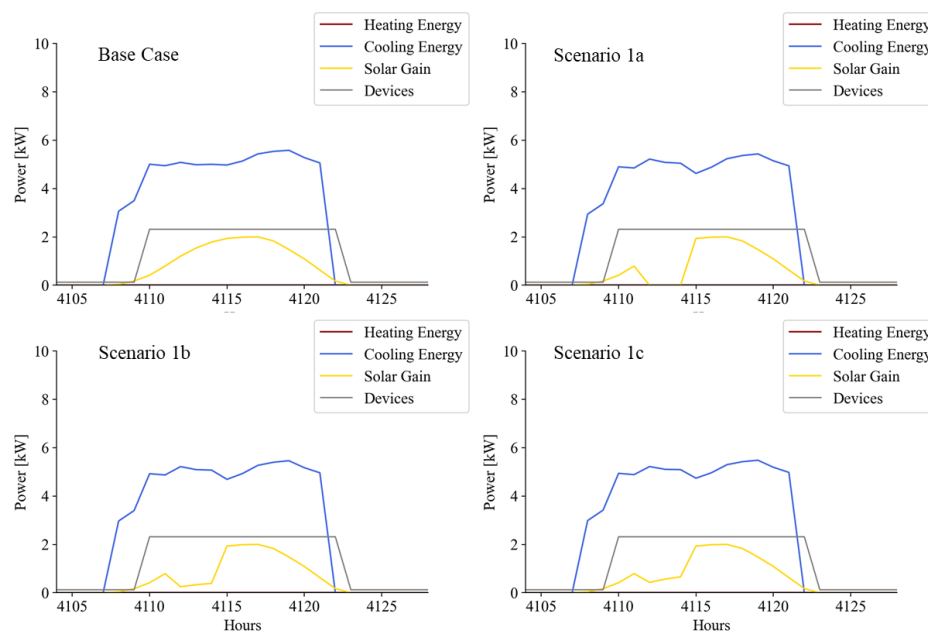


Fig. 88. Comparative Analysis of Hourly Energy Contributions in a Summer Day.

Table 30. Annual Heating and Cooling Energy Demand Across Various Scenarios.

		Base Case	Scenario 1a	Scenario 1b	Scenario 1c
Heating Demand	[kWh]	2110.8	2350.4	2308.6	2266.5
Heating Peak	[kW]	9.9	9.9	9.9	9.9
Cooling Demand	[kWh]	7372.8	7066.2	7130.5	7185.8
Cooling Peak	[kW]	11.4	11.3	11.3	11.3

Table 30 reports the annual outcomes from the proposed shading strategies in this scenario. In particular, their impact on peak and annual thermal loads was examined. During the heating period, Scenario 1a, employing slats parallel to the window, resulted in a 240-kWh increase in heating energy. In contrast, Scenario 1c, with slats angled at 30 degrees on the horizontal, reduced this increase to 155.7 kWh. However, this disadvantage in winter was offset in summer, with Scenario 1a achieving a cooling energy saving of 306.6 kWh compared to the Base Case. Concerning peak loads, shading devices showed a minimal effect since in winter, peak loads occurred early in the day due to the need to heat the air that has cooled overnight. In summer, with the sun at its zenith around noon, the canopy effectively blocked direct solar radiation.

Finally, a comparison of the values simulated at 11:00 AM of the 21st of December was carried out to evaluate the effectiveness of the shading solution for visual comfort (Fig. 89).

The red rectangles in Fig. 89 show that, at 11:00 in the Base Case, the values of illuminance were higher than 10,000 lux. By using blinds, at that time, the illuminance was still sufficient being about 600 lux in Scenario 1b and 1c. In all cases, there were no glare issues.

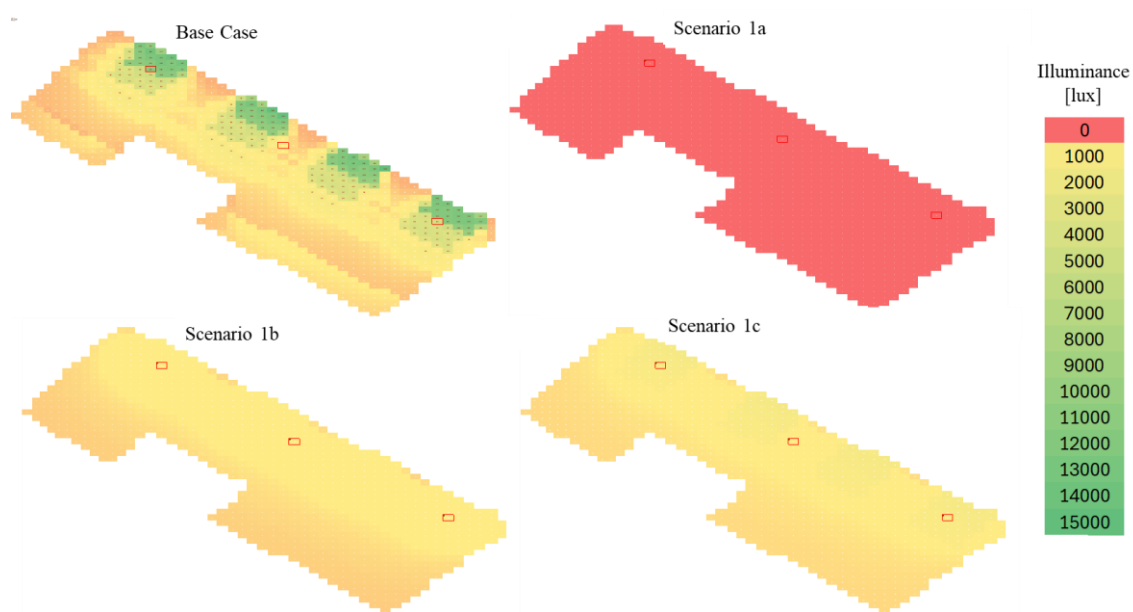


Fig. 89. Illuminance Maps at 11:00 AM in Winter Day: Base Case Versus Scenario 1.

As shown in Fig. 90, similar to the winter day, by considering a typical summer day, given that the slats are vertical and considered closed, the scenario results show zero illuminance in each point. Enhanced summer solar radiation contributed to higher diffuse lighting, identifying areas with illuminance above 2000 lux in the other scenarios, and in areas particularly close to the window in Scenarios 1b and 1c, thereby preventing glare at the workplanes. The results of Scenarios 1b and 1c were similar exhibited similar. This is due to the predominant contribution of diffuse radiation, as direct sunlight was obstructed by the external shading.

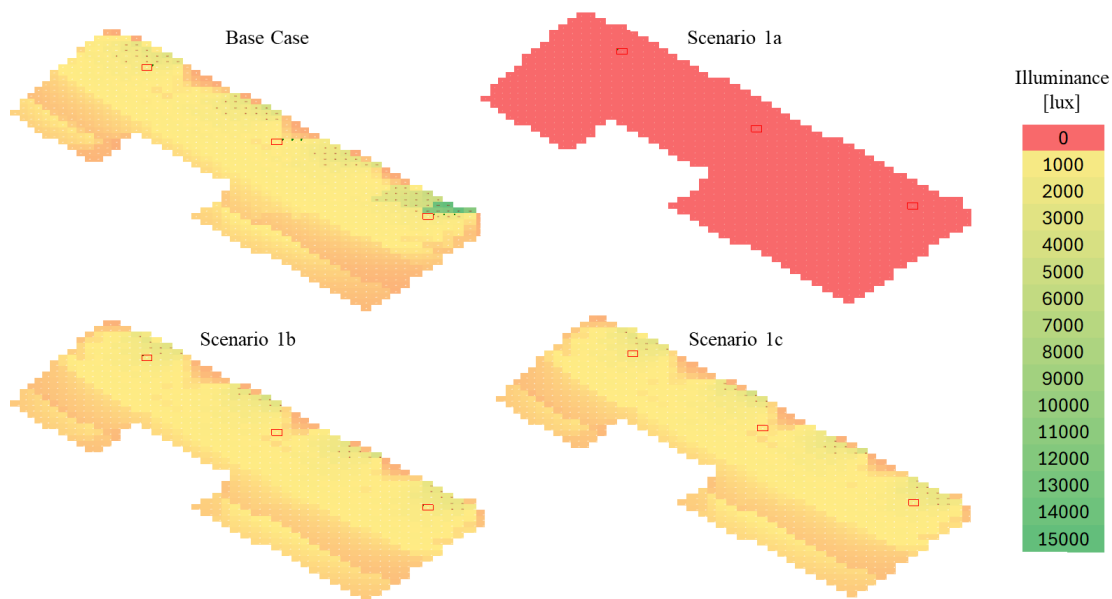


Fig. 90. Illuminance Maps at 11:00 AM in Summer Day: Base Case Versus Scenario 1.

Regarding the thermal comfort evaluation, PMV was studied for all scenarios to examine the impact of using slat blinds in the laboratory [6]. This method was developed by Fanger. It provides a thermal sensation index on a seven-point scale ranging from -3 (cold sensation) to $+3$ (hot sensation) and is determined by six environmental variables (indoor air temperature, MRT, relative humidity, and air velocity), and two factors related to occupants, namely their clothing level and metabolic rate. The HVAC system regulated by an air temperature thermostat was considered. As can be seen in Fig. 91, similar values of PMV were found for Scenario 1a, 1b, and 1c.

In all three scenarios, the control strategy for closing the slatted blinds is the same; the only variation is the angle of the slats, which changes with each scenario. Specifically, in Scenario 1a, where the slats are vertical and cover the entire window, the PMV is constantly lower than in Scenarios 1b and 1c. In Scenario 1c, the slats are tilted at 60 degrees from the vertical, resulting in less coverage of the window and leading to a higher contribution of solar radiation.

However, comparing the Base Case with Scenario 1 results, the values of PMV in the two days are different. In summer, as shown on the right side of Fig. 91, the PMV of the Base Case consistently equals or exceeds the PMVs in Scenario 1. This occurs because the direct solar radiation in the morning warms the internal surface, and the lack of shades in the afternoon promotes radiative heat gain. On the contrary, during winter mornings, the PMV is higher due to the heating effect of solar radiation on the surfaces, which results in a positive gain. During the afternoon, when the sun is sufficiently high to be obscured from the external shading, the slatted blinds reduced heat losses in Scenario 1 leading to comparable PMV values. After 13:00,

the PMV of Base Case is again higher than Scenario 1. It is because, in both cases, the blinds do not cover the windows. Furthermore, thermal inertia of floor, which has reached a higher temperature due to solar radiation, influences the overall MRT. It enhances thermal comfort.

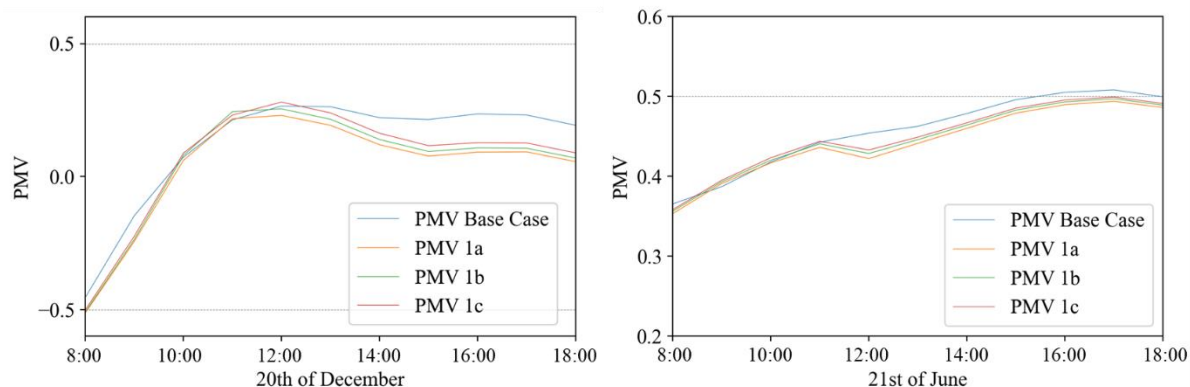


Fig. 91. PMV trend in a typical winter and summer day during work hours.

It is important to underline that in some cases regulating air temperature is not sufficient to ensure indoor thermal comfort. As illustrated in Fig. 92, on distinct winter and summer days, even though the indoor air temperature is maintained at 20°C and 26°C respectively (in compliance with current Italian legislation), the required PMV threshold between -0.5 and 0.5 was not achieved in both instances. However, it was met for up to 64% of the working hours. Despite this, focusing on the trend of PMV in both cases, the behaviour is comparable to the two days previously analysed. This fact confirmed that radiant temperature has to be considered, and only controlling the air temperature, which most thermostats are based on, is insufficient to guarantee comfort. Therefore, in extremely cold or hot conditions, particularly affecting uninsulated envelopes, higher setpoint temperatures (in winter) and lower temperatures (in summer) should be considered in the HVAC system.

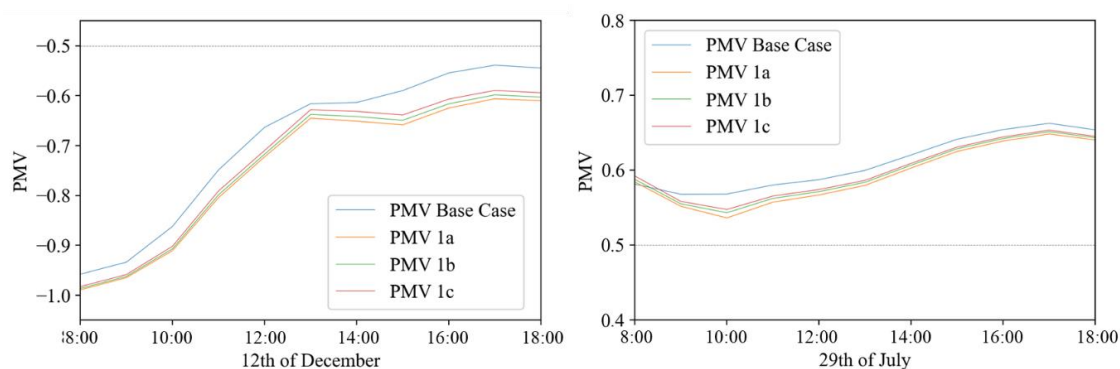


Fig. 92. PMV trend in an especially cold day (left) and hot day (right) during work hours.

7.4.2.3 Scenario 2

In Scenario 2, the influence of the roller blinds was explored through the use of five distinct EnergyPlus model simulations. To do this, 5 simulations were run by considering respectively

the shading with shades at 0%, 20%, 40%, 60% and 80%. Each simulation served to identify for each hour the percentage of closure of the shading system that avoided glare issues. This methodology allows for an examination of factors such as hourly illuminance and solar radiation through the windows, despite not being capable to simulate in detail heating and cooling demands which are influenced by factor such as air or building inertia. For the daily analysis, the same representative days chosen for Scenario 1, covering one winter and one summer day, were considered.

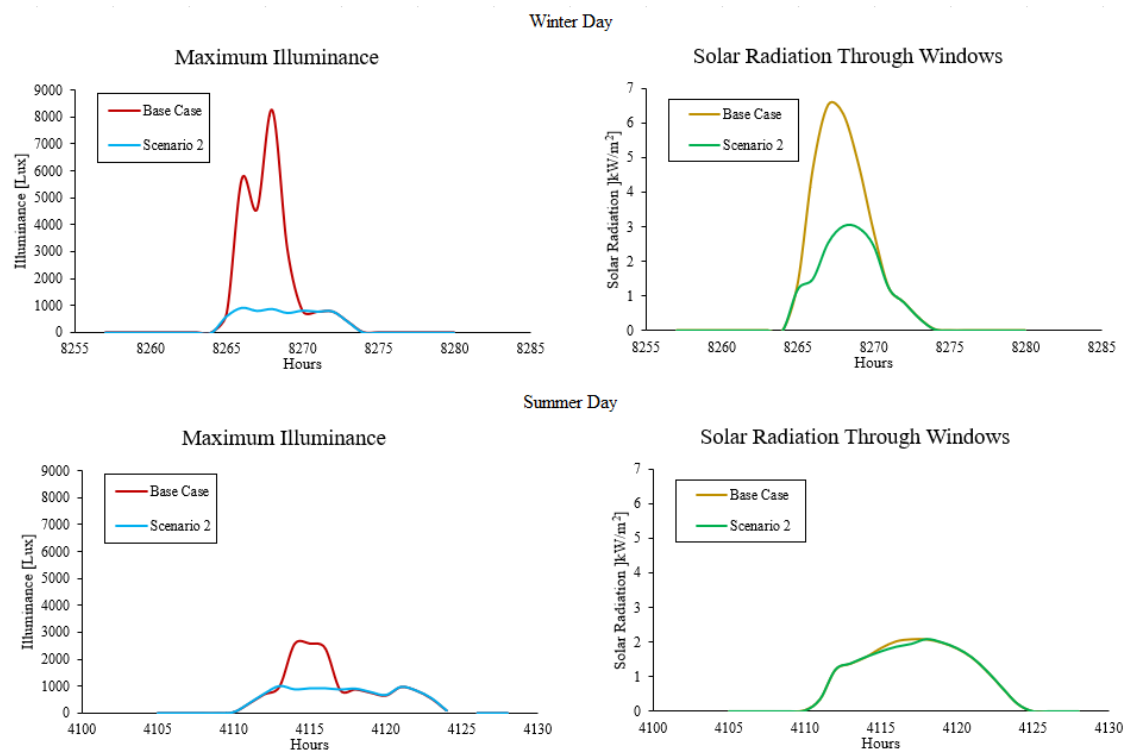


Fig. 93. Analysis of Roller Blind Impact During Typical Winter (Above) and Summer (Below) Days.

In winter, as depicted in Fig. 93, lower sun positions in the morning led to workplanes illuminance levels exceeding 8000 lx, causing visual discomfort. The activation of roller blinds reduced this illuminance, not only reducing glare, but also maintaining an average daylight illuminance of 1000 lux. It ensured visual comfort from natural light and reduced potential reliance on artificial lighting. This reduction in solar radiation resulted in a higher heating demand while, during summer, desk illuminance slightly exceeded 2000 lux from 9:00 AM to 12:00 AM, allowing the roller blinds to provide partial shading and reducing cooling. Also in this case, the maintained illuminance level of about 1000 lux when blinds were active indicated an effective glare control. When considering solar radiation, the reduction impact from Scenario 2 compared to the Base Case was less pronounced in summer, attributed to the elevated position of the sun.

7.4.3 General Discussion

To assess the impact of various shading strategies on energy consumption, all building energy needs were converted into PE. Specifically, natural gas consumption was converted with a factor of 1.05, and electricity from grid consumption utilized a conversion factor of 2.42, following the guidelines outlined in [43]. Furthermore, given that Scenario 2 integrates results from multiple simulations, each representing a different level of shading, the PE calculations for heating and cooling are based on Base Case thermal demand output. Precisely, during the occupancy period, the difference in solar gain was then algebraically aggregated to approximate Scenario 2 energy consumption. In this sense, it is important to underline that the consumptions related to Scenario 2 were calculated as Energy Plus and Euclid are not able to simulate the adjusting of shading objects dynamically.

Table 31. Analysis of Glare Occurrence Across Scenarios

	Base Case	Scenario 1a	Scenario 1b	Scenario 1c	Scenario 2*
Glare Hours	900	0	0	12	0
Yearly	10.27%	0.00%	0.00%	0.14%	0.00%
Work period	20.21%	0.00%	0.00%	0.38%	0.00%

The results reported in Table 31 show the effectiveness of the examined methods in mitigating glare at all three workplanes in the study. Specifically, in the base case scenario, despite the existing shading shelter, illuminance exceeded the threshold for 900 hours annually (10.27% of the year); while, considering work periods, this represented 20% of the time. All scenarios successfully eliminated glare in desk positions, with Scenario 1c being the exception since it only reduced this percentage below 0.4%.

Table 32. Comparison of Primary Energy Consumption Among Scenarios

PE		Base Case	Scenario 1a	Scenario 1b	Scenario 1c	Scenario 2*
Heating	[kWh]	2333.0	2597.8	2551.6	2505.1	2629.9
Cooling	[kWh]	3784.3	3628.2	3660.5	3688.2	3758.9
Lighting	[kWh]	154.3	400.1	392.1	360.5	175.4
Total	[kWh]	6271.6	6626.0	6604.2	6553.7	6564.3

Results presented in Table 32 clarify that by the use of shading systems, despite the presence of glare was notably reduced, the PE consumption could increase. In this case study, the increase in total PE ranged from 4.6% in Scenario 2 to 5.6% in Scenario 1a since the reductions in cooling energy achieved through shading were offset by increases in heating and lighting demands. Specifically, the presence of the shelter in summer led to minimal cooling savings,

and the production of cooling energy, using air-to-water chillers, was more efficient than heating resulting in energy consumption comparable. Moreover, the demand for artificial lighting significantly increased in Scenarios 1a to 1c. It is because the shading effect is achieved by fully closing the blinds, a condition that can be modelled using Energy Plus. Although adjusting the angle of the slats was feasible, the effectiveness of this adjustment heavily relies on the sun position and its hourly scheduling is limited by the character and text limit for EnergyPlus schedules. For this reason, using a fixed slat angle results in a substantial decrease in desk illuminance and, thus, more frequent reliance on artificial lighting systems. Scenario 2 approach of using modularly adjusted roller blinds resulted in an illuminance pattern that maintains electrical demand for lighting akin to the base case, showcasing the efficacy of this adaptive shading method.

7.5 A Methodology to Enhance Thermal Comfort Monitoring in Buildings

One of the most important parameters needed for PMV evaluation is the MRT. In the absence of instrumentation to measure it (i.e. a globethermometer), it needs to be calculated. For instance, La Gennusa et al. [9] presented a simple method to assess the MRT by using the temperature of the internal surfaces of the environment, the intensity of the diffuse and direct solar radiation entering the room through the glazed surfaces, the angle factor between the opaque and glazed surfaces of the environment and the subject, and the projected area factor of the subject in the solar beam direction. Of course, different tools include the calculation of MRT, but right now there are some lacks as underlined by Alfano et al. [44].

The focus of this research activity is on devising novel methodologies and instruments for monitoring thermal comfort parameters across various zones within indoor environments. The objective is to determine the optimal types, quantities, and placements of sensors required to regulate heating and cooling systems in diverse zones, alongside exploring control strategies for radiant systems. Within this context, this section aims to introduce a methodology for evaluating punctual indoor thermal comfort conditions. It aims particularly to compute comfort conditions at specific points within the space (e.g., where occupants are situated), thereby facilitating the generation of MRT maps across different locations and incorporating additional surfaces with varying temperatures (e.g., to simulate radiant conditioning systems). Moreover, ML models were developed to estimate the PMV of two occupants based on MRT evaluated in two different zones of the room. The contribution of this work is to present this methodology that could be potentially applied to any Building Energy Model output to enhance investigation on local PMV.

7.5.1 Methodology

The methodology proposed in this study is illustrated in the architectural schema in Fig. 94. The EnergyPlus model is simulated to generate various outputs, including surface temperatures, outdoor and indoor temperature and humidity, and global solar radiation. Using the surface temperatures, a Python script calculates the view factors between the surfaces of two complex objects (e.g., a person and the room), which are then used to determine the MRT of the individual object or person. Subsequently, the PMV for people is calculated based on MRT and other parameters, such as air temperature, provided by the EnergyPlus model. The output of this script can be utilized for both mapping MRT within the room or estimating MRTs and PMVs for various objects. Finally, ML models are trained using different indoor and outdoor parameters, as well as the simulated temperature from MRT sensors, to estimate PMVs for people. This approach will assist in identifying both the optimal model and the most suitable position for placing the MRT sensor (i.e., GT).

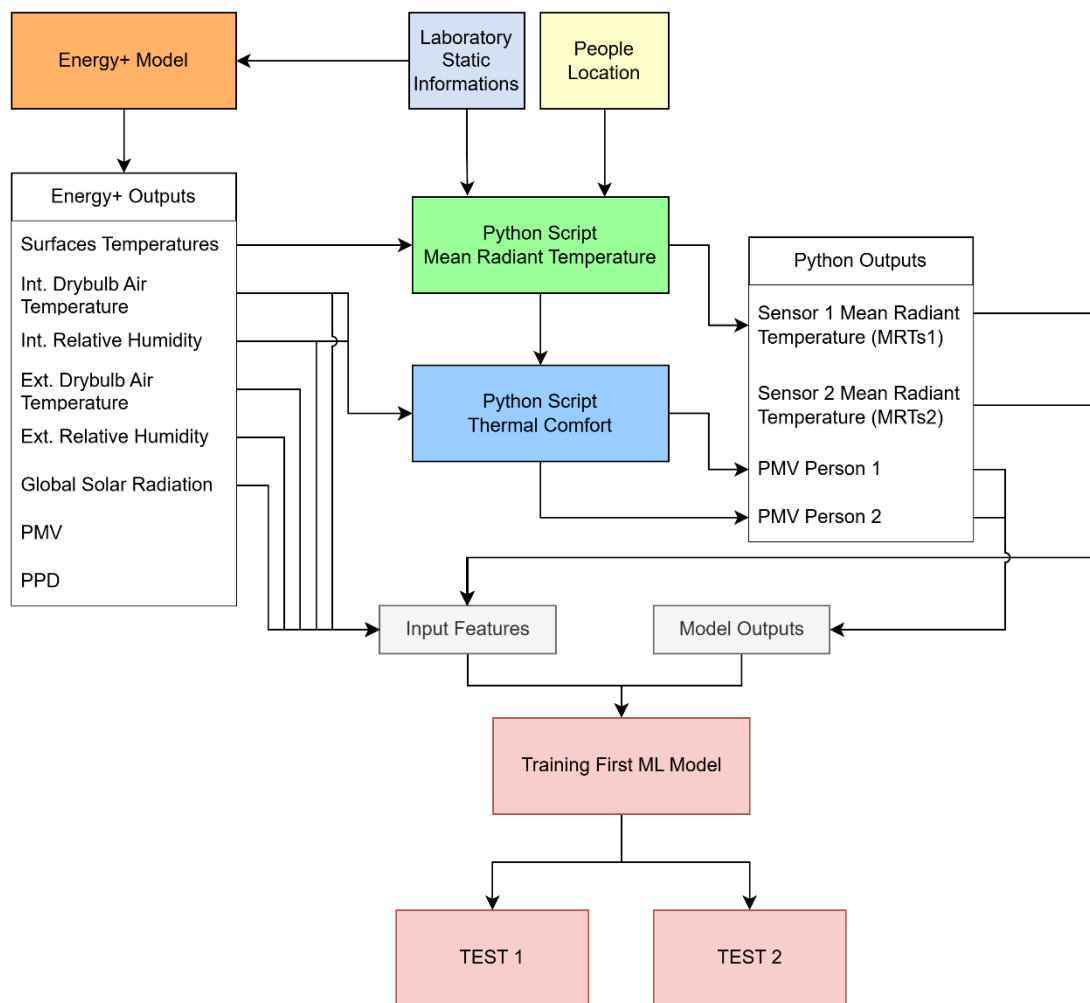


Fig. 94. Block diagram showing the workflow adopted for the development of the software.

7.5.1.1 Simplified Local Mean Radiant Temperature Algorithm

To improve the capability of estimating thermal comfort, a novel methodology is presented in this study which combines three distinct Python libraries: PyVista [45], PyViewFactor [46] and PyThermalComfort [47]. The first is used to build and represent complex 3d meshed object inside Python environments to simplify mesh creation and plotting without losing functionality. PyViewFactor, as the name suggests, is a Python library that computes exact view factors between planar faces, applicable to both indoor and outdoor environments. PyThermalComfort is a comprehensive library that includes various algorithms related to thermal comfort calculation, such as PMV, PPD, adaptive PMV and Heat Index.

Since EnergyPlus calculates the average PMV and PPD values within a room using a simplified calculation for the MRT, the algorithm presented here extends this calculation throughout the three-dimensional space within the laboratory. Specifically, PyVista is employed to define the problem's geometry, which includes the room and the human body, both modelled as parallelepipeds. Subsequently, the view factors are computed between each surface of the individual and all the surfaces comprising the room under study. With the knowledge of all view factors, equation (40) is introduced to calculate the MRT for each surface of the individual's parallelepiped:

$$MRT_i = \sqrt[4]{T_1^4 F_{p-1} + T_2^4 F_{p-2} + \dots + T_n^4 F_{p-n}} \quad (40)$$

Where MRT is the MRT of the surface i of the individual, T_n^4 is the temperature of the room's surface n (an output of the EnergyPlus model), and F_{p-n} the view factor between the individual's surface i and the room's one n.

Once the MRT for each surface of the individual i is calculated, a single MRT for the individual, weighted on the surfaces A, is calculated as per equation (41):

$$MRT = \frac{\sum_{i=1}^{n_{surf}} A_i \cdot MRT_i}{\sum_{i=1}^{n_{surf}} A_i} \quad (41)$$

Finally, the outputs of the EnergyPlus model, including dry bulb air temperature, relative humidity, and MRT, are required to calculate the PMV and PPD using PyThermalComfort. To estimate these parameters, a metabolic rate of 1 met and clothing insulation of 1 clo in winter and 0.5 clo in other seasons are assumed in this case study.

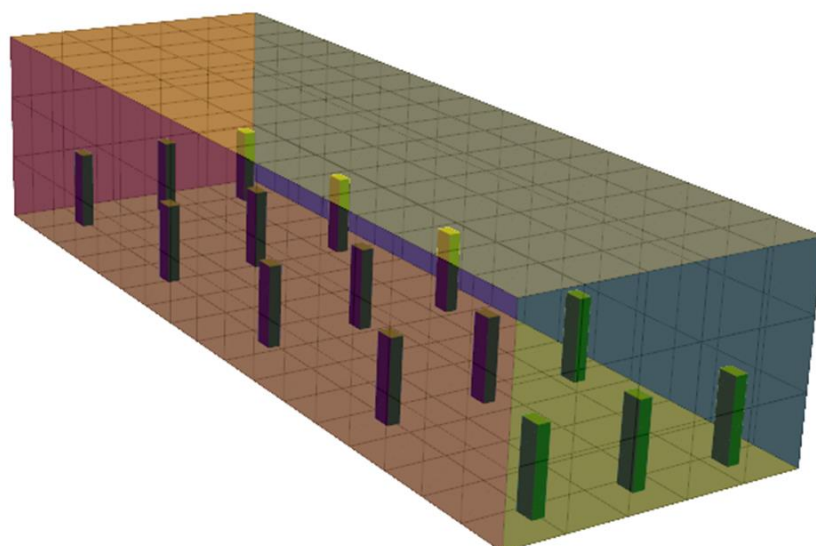


Fig. 95. SolaLab Model for MRT Simplified Algorithm.

In this manner, the presented algorithm is capable of calculating various thermal comfort parameters, considering only rectangular rooms, and expanding the possibilities available in EnergyPlus. It can even consider the effects of different surface temperatures, simulating the effects of zonal radiator roofs or other personal devices that could locally affect the individual's thermal behaviour.

7.5.1.2 Advanced Local Mean Radiant Temperature Algorithm

This advanced parallel method, driven by simulation, has been developed to explore the feasibility of improving control strategies with only one GT in order to allow the evaluation of comfort conditions for different occupants within the space. For this purpose, view factors and MRT were calculated for over 6000 hours for four different objects: two parallelepipeds to represent standing persons and two small cubes to represent the GTs. The simulated data obtained above will be used as one input to different ML algorithms outlined in the following sections for the evaluation of thermal comfort for humans using GT data with other variables. How these variables are precisely generated is shown in the flowchart in Fig. 96.

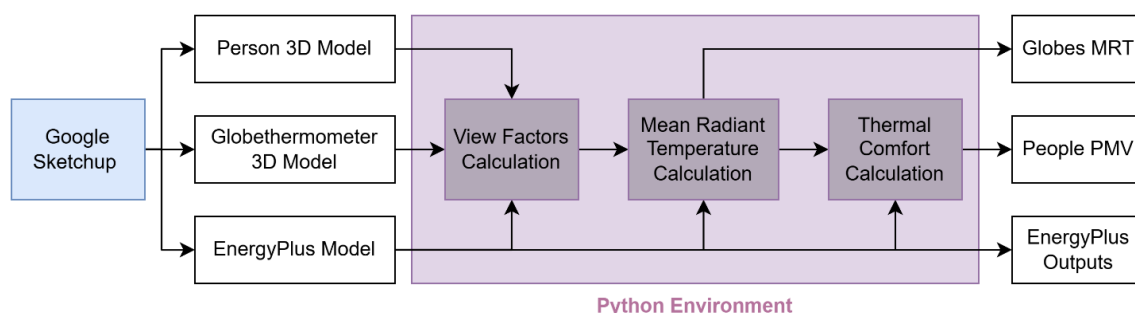


Fig. 96. Flowchart for the estimation of the indoor parameter needed my ML algorithms.

Google SketchUp was utilized to create the 3D models of the SolarLab and for the representations of the individuals and the globes. As the SolarLab energy model was developed by integrating the Euclid extension in Sketchup, the 3D models were exported as .obj files and used by the algorithm for view factor calculations. In this advanced version, the algorithm was enhanced to accept any 3D model and compute the view factors between different surfaces of the models. Next, the view factors and surface temperatures from the EnergyPlus simulations were used to calculate the MRT for all four models (GTs and individuals) in the room, while the PMV for the individuals was calculated using the indoor conditions simulated by EnergyPlus. Finally, the outputs were refined and post-processed to be utilized by the ML algorithms.

7.5.1.3 Machine Learning Algorithms for PMV Estimation and Optimal Sensors Position

After conducting simulations using EnergyPlus software and developing a script for MRT, the outputs from these simulations have been employed to create ML models that can estimate the PMV. The primary objective is to predict the PMV for occupants based solely on external weather conditions, roof surface temperature, internal data from MRT sensors, air temperature, and relative humidity. To accomplish this, a range of commonly used ML algorithms, including DL and decision-tree-based methods, have been evaluated with different hyperparameters and input features.

The testing scenarios have been selected to achieve the following objectives:

1. Identify the best ML model and the optimal combination of input features and hyperparameters.
2. Test the chosen model for estimating user PMVs using the first virtual MRT sensor.
3. Evaluate the model using the second sensor.

Following this process, it will be possible to have an ML model that, once trained using real historical data or simulation data (if historical data are not available), will assess the PMV of different individuals inside the room to enhance the control of the HVAC system. Specifically, since the laboratory is expected to incorporate a radiant heating and cooling system, this methodology relies on the accurate estimation of MRT, which depends on the individual's position as well as the roof temperature.

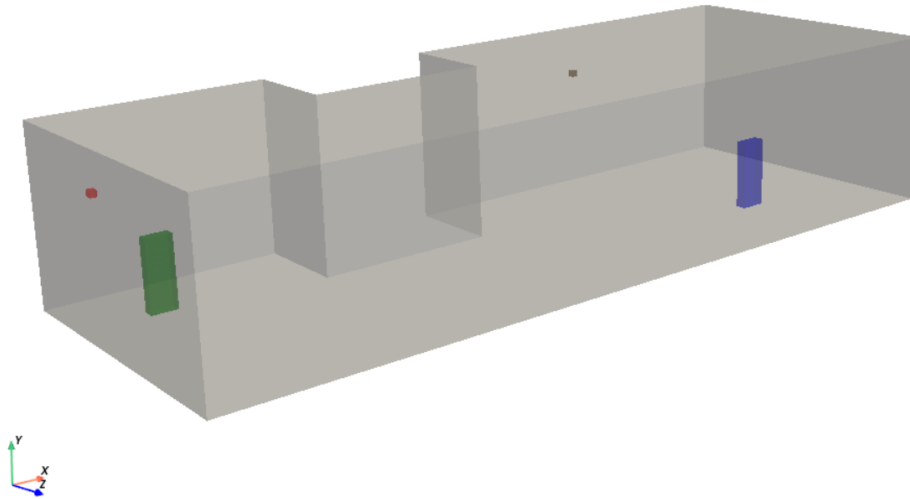


Fig. 97. 3D spatial view of sensors (cubes) and individuals (parallelepipeds)

Additionally, simulations for the last two objectives are necessary to evaluate how the position of the sensor will impact the final accuracy of the model. The positions chosen, depicted in Fig. 97, are deliberately very different from each other to evaluate this phenomenon. Moreover, in the first scenario, the model is based on the mean values of MRT and PMV for the two sensors and two individuals, representing a general model that can be applied under various conditions, once it has been accurately validated.

7.5.2 Results and Findings

7.5.2.1 Thermal Comfort Parameters Mapping in SolarLab in Free Floating

The simplified python-based model enhances the monitoring of thermal parameters within the laboratory. Consequently, it allowed for a detailed visualization of the behaviour of the MRT experienced by an individual in various positions within the room, as depicted in Fig. 98. Notably, the left and right walls of the laboratory, which are adiabatic boundaries with heated rooms, consistently exhibited higher temperatures than those recorded during winter.

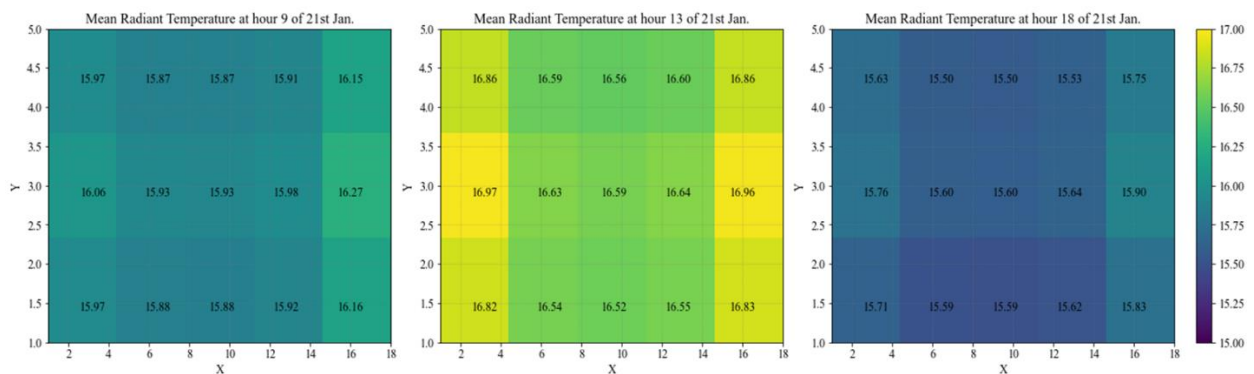


Fig. 98. MRT on January 21st at different times.

On a typical summer day, such as July 31st, the situation is similar. However, due to increased radiation and the south-west facing left wall, higher temperatures are recorded closer to the left wall as shown in Fig.99.

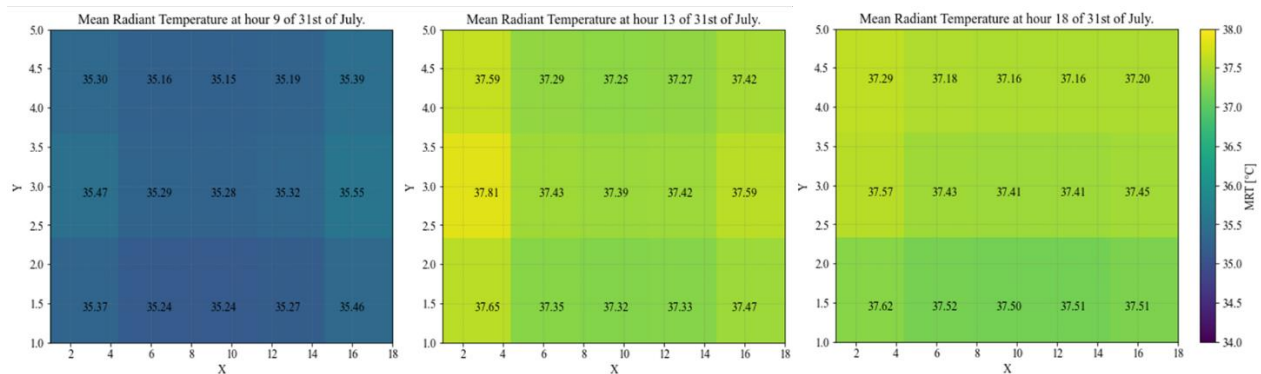


Fig. 99. MRT on the 31st of July at different hours

Free-floating analysis reveals, as expected, that the MRT varies between different points in the zone by 1% to 2%. Since the only difference in calculating PMV and PPD across different zone positions is the MRT, similar variations are observed. This is illustrated in Fig.100, which shows the PMV trends for two different days based on the chosen position. Notably, for both days, the individual is outside the comfort zone. On the winter day, the highest PMV value reached is -1.7, between ‘Cool’ and ‘Slightly Cool’ “sensation”; while on July 31st, the value is consistently around 3, corresponding to an ‘Hot’ sensation.

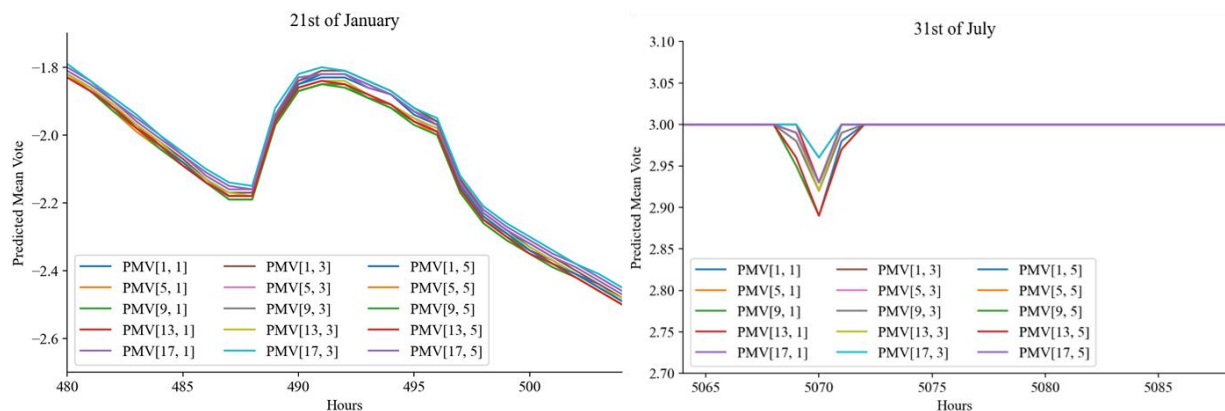


Fig. 100. PMV trend on the 21st of January and 31st of July at different hours

Focusing on the equinoxes, specifically June 21st and September 21st, the situation improves slightly. In spring, a PMV value closer to the legislative limit is achieved. However, in autumn, given that Palermo remains warm in September, the sensation is still ‘warm’ to ‘hot’, as can be seen in Fig.101.

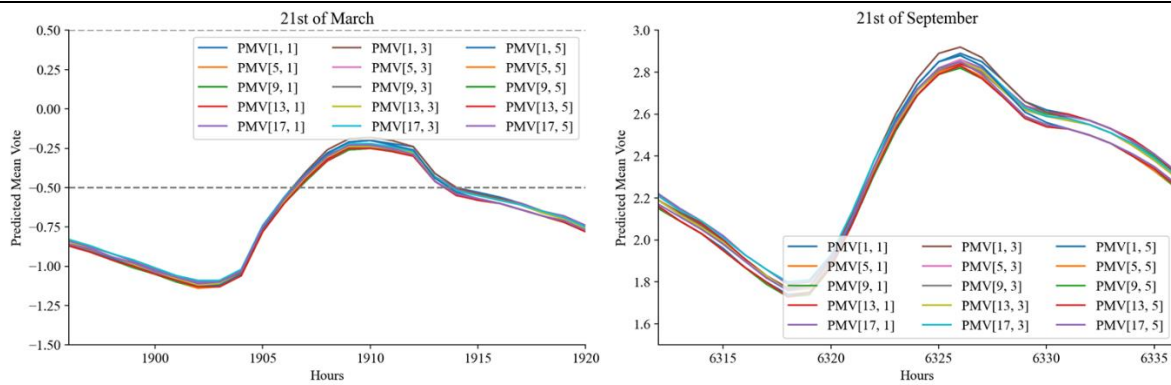


Fig. 101. PMV trend on the equinoxes at different hours.

7.5.2.2 Outcomes of Advanced Local Thermal Comfort Output

To develop the updated version of the local thermal comfort algorithm, a similar methodology was followed, combining the three Python libraries previously presented: PyVista, PyViewFactor, and PyThermalComfort. The simplified version of this algorithm could only be applied to rectangular-shaped rooms because PyViewFactor was unable to effectively calculate view factors between partially covered surfaces. This limitation led to the consideration of each surface of the object for which view factors were to be calculated, taking into account only the three-dimensional section of the room that was fully visible to that surface, effectively segmenting the partially viewed surfaces. While this approach was correct and effective in rectangular zones, it was not applicable in the real context of SolarLab, which has a C-shaped room. Therefore, to minimize the number of partially viewed surfaces and simultaneously allow for buildings and objects with complex geometries, this new approach was adopted.

Once the building and objects models were developed on Sketchup, the model was further divided in the smallest possible surfaces, bearing in mind that the view factor of every surface of the room for all surfaces of the object had to be calculated. This required a trade-off between achieving a highly accurate model and the increased computational effort required by the algorithm. Fig. 102 shows the meshed model of the SolarLab, precisely the mesh was created before extracting the model .obj (left) that was then imported to the developed script using PyVista (right). It was important to set correctly the orientation of each surface while creating the three-dimensional models, in such a way that object surfaces had to be oriented in an opposite direction to room surfaces.

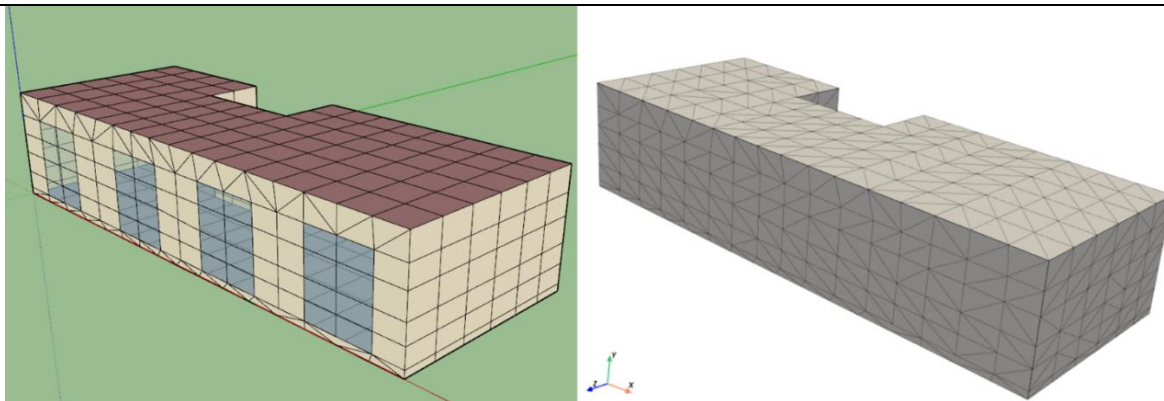


Fig. 102. 3D Model of SolarLab in Sketchup (left) and after being meshed by PyVista (right)

Once the model was imported, the script began calculating the view factors between each surface of the object, 12 surfaces in the case of parallelepipeds representing people and cubes representing GTs, based on their relative positions. For example, considering the developed model with a mesh composed of 928 triangles, calculating the view factors between the entire object and the room required 11.136 calculations. After this, as ray tracing is implemented in PyViewFactor, the view factors, including obstacles, were calculated. Fig. 103 shows an example of the view factor distribution for the lower surface of the person facing the stairwell (represented by the blue triangle inside the person's parallelepiped). As seen in Fig. 103, the surfaces with a view factor higher than a threshold value of 0.00001 are all located opposite the blue surface. Furthermore, the closest surfaces, which are those nearest to the floor of the zone, have higher view factor values up to 0.1. Additionally, the wall on the left is only partially visible, while the walls of the stairwell that are perpendicular to the person's surface are not visible at all. Finally, since user-defined meshes can still lead to inaccuracies in cases of partially visible meshes, all the calculated view factors, whose sums in simulation ranged between 0.98 and 1.01, were scaled so that the final sum equals 1.

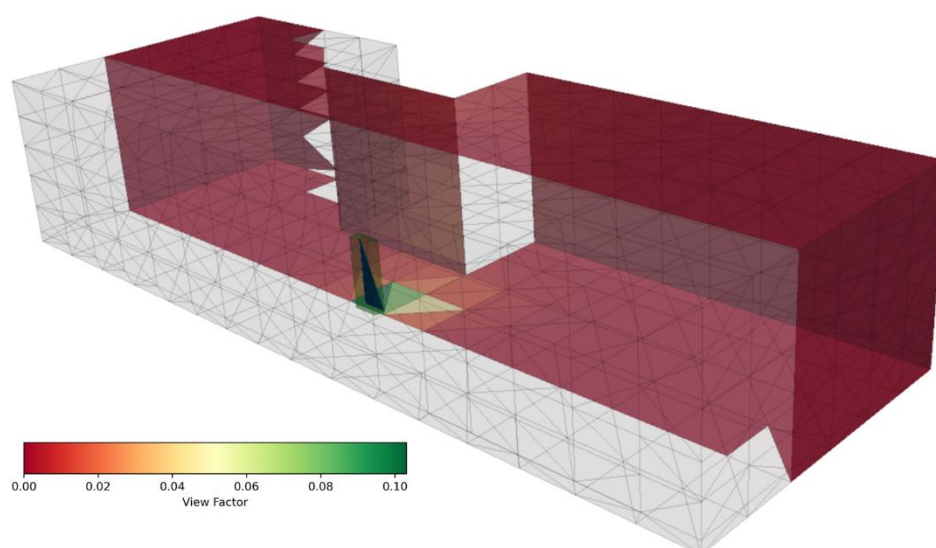


Fig. 103. View Factors Distribution Inside SolarLab

Having determined view factors for all combinations of surfaces, it was necessary to assign the correct temperature to each surface. Although this can be done manually in the case of simple geometrical configurations or limited numbers of surfaces, in a more complex model, an automated methodology was required. To solve this issue, temperature values had to be assigned to each macro-surface (e.g. roof and windows), which were either the output of EnergyPlus simulations or, if available, surface temperature measurements. Then, the function will calculate, for a given face, the nearest surface temperature based on the distances between the selected surface and the macro-surfaces. To avoid increasing computational costs and time by repeating this process for each timestep, the function is applied once, and the information is stored in a dictionary with each configuration as a key. This method allows temperatures to be automatically assigned to each surface at every timestep, enabling the calculation of the MRT, as described in the intermediate report. Finally, the MRT of each object's surface-person or GT is computed for each time step by using a function that takes into account the view factors, the temperatures of room surfaces, and the area of each object's surface.

In order to save computation time, this procedure was conducted for only four different objects, as depicted in Fig. 104 below for working days (from Monday to Friday) for a total of 6241 hours. Specifically, for S1 and S2, representing GTs (named also MRT sensors), the MRT was calculated. For the two individuals, P1 and P2, located at desks within the room, once the MRT was determined, PMVs were also calculated using the PyThermalComfort library, based on indoor conditions (e.g., temperature and humidity) obtained from the EnergyPlus simulation. For the individuals, an activity rate of 1 MET was assumed, with a clothing insulation value of 0.5 for the cooling period and 1 for the heating period, following current legislation.

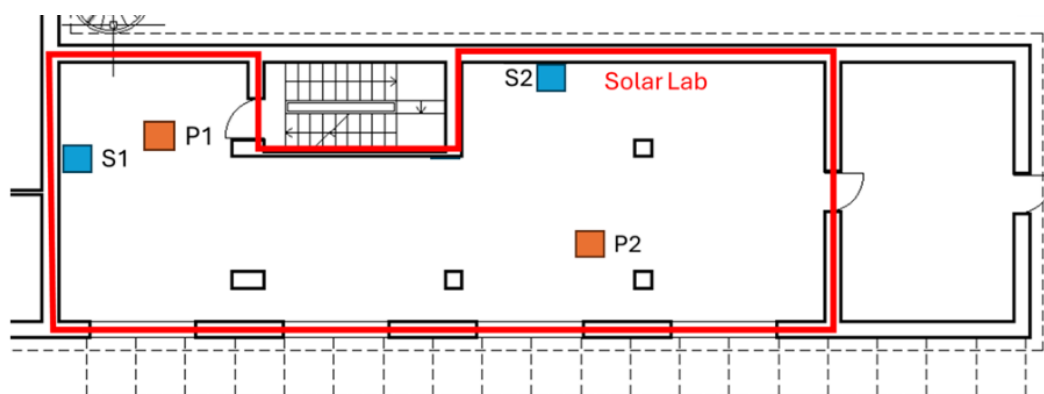


Fig. 104. People (P1 and P2) and Globe Thermometers (S1 and S2) position inside SolarLab

7.5.2.3 Optimal Position and Machine Learning Model for PMV Estimation

In this study, various algorithms in ML and DL have been implemented and tested to estimate PMVs to assess indoor thermal comfort. The tested algorithms include XGBoost Regression,

GBoost Regression, LSTM networks, Feedforward ANNs, RF Regression, Decision Tree Regression and K-Nearest Neighbors (KNN) Regression. Said models have then been assessed on various key performance metrics such as MAE, MSE, RMSE, and Coefficient of Determination (R^2) to present the most reliable and accurate model.

Input features were chosen by first analysing correlations between variables, and secondly by testing different combinations of inputs to calculate the average PMV calculated as the mean between PMVs of the two simulated people. The input features selected include the MRT of the sensor, roof surface temperature, site dry bulb temperature, site total radiation, and SolarLab indoor conditions (i.e. indoor air temperature and relative humidity). To select the optimal ML model for the estimation of the PMVs of the two people, PMV_1 and PMV_2 , models were trained and accuracy for PMV estimations was evaluated as shown in Fig. 105.

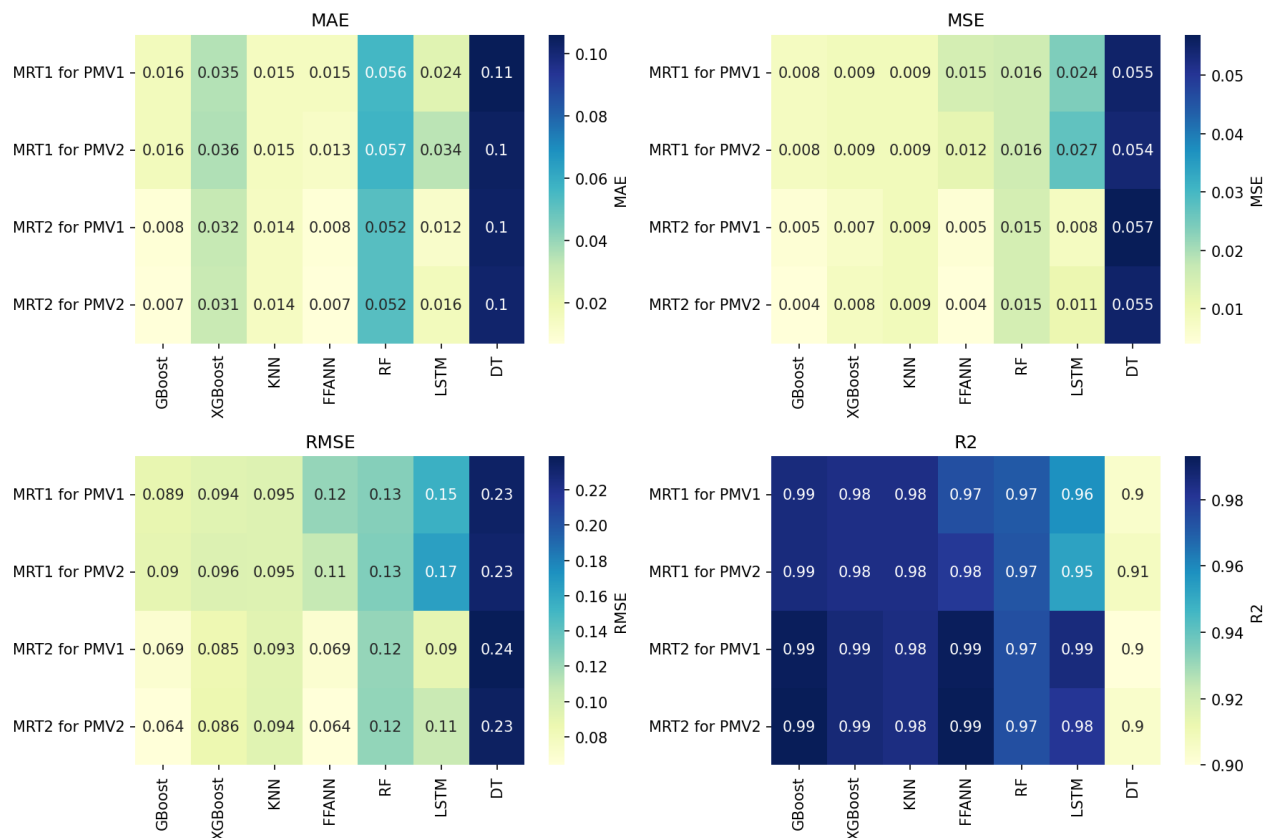


Fig. 105. Accuracy of Models for PMV Estimation Using Total Dataset

Fig. 105 presents the detailed performance of various ML models concerning the estimation of PMV_1 and PMV_2 in Scenario 3, by making use of the total dataset. Exception for the use of two different MRT inputs (MRT_1 and MRT_2), all the models used the same input features for estimating PMVs. Across all combinations of MRT and PMV, Feedforward Artificial Neural Network (FFANN) consistently emerges as the top-performing model. For MRT_2 and PMV_1 , FFANN achieved MAE of 0.008, MSE of 0.005, RMSE of 0.069 and R^2 of 0.99, which clearly indicates good estimations by this model. Similarly, FFANN maintained strong accuracy for

both MRT_2 and PMV_2 with an MAE of 0.007 and an R^2 value of 0.99, which means small prediction errors. Other well-performing techniques are XGBoost, GBoost and LSTM models which obtained comparable accuracy. The worst performance in estimating both PMV_1 and PMV_2 was achieved by Decision Tree, which generates higher error rates in all combinations of MRT and PMV. In general, the models with MRT_2 achieved lower errors with higher R^2 values across all metrics, making this the better input toward PMV estimation from both datasets. This outcome revealed that the most promising position of the MRT sensor concerning the models for both PMV_1 and PMV_2 estimation is MRT_2 . Also, the models provided better accuracy in MRT_2 , thus establishing the same as the best choice for the task of PMV estimation across different datasets.

7.6 Final Discussion and Future Directions

Nowadays, where users predominantly stay in indoor environments, the importance of safeguarding their health and comfort becomes crucial. Visual comfort is fundamental in creating an optimal working environment, synergizing with other comfort parameters like thermal comfort and IAQ. Rigorous investigation of lighting levels is not only essential to comply with technical standards but also creates new possibilities for efficiently managing interior illumination systems to reduce energy consumption.

DTs, as a peak of Industry 4.0 and the digitalization of the building and energy sectors, present novel opportunities to augment the benefits of visually comfortable environments for individuals. As virtual counterparts of physical systems, DTs can aid building managers and users in maintaining awareness of their comfort status and facilitate optimized control of HVAC systems, lighting, and other electrical loads. Although there are numerous examples in literature showcasing the advantages of DTs in thermal comfort monitoring and HVAC system optimization, a notable gap exists in their application to visual comfort assessment.

To bridge this gap and elucidate the benefits of employing a DT for visual comfort monitoring, this chapter introduced the DT application in a real laboratory at the University of Palermo. The DT in question is proficient in calculating key visual comfort KPIs based on illuminance and luminance sensor data within the laboratory. For instance, at two selected workstations, metrics such as DF, illuminance distribution, and luminance are computed. By accessing the DT, developed on LabView, users can monitor their comfort status, communicate discomfort, and receive actionable suggestions for improvement. Furthermore, the model in DiaLux extends sensor capabilities beyond the current experimental setup. It enables the simulation of the real system and the exploration of various strategies to enhance the management of lighting and shading systems. Even if easy, the use of Labview however limits

the scalability of this methodology, as depending by a non open source tool which is Labview. Moreover Labview components, the “vi” are limited so it must be considered that for difficult task it could be needed to use python-based vis. In future, the DT will be improved by including other parameters that are useful to evaluate visual comfort conditions e.g. contrast, and colour rendering index of the luminaires, and more precise uniformity by adding further luxmeter. As mentioned, Future simulations with the DAYSIM and DIALux-developed model aim to uncover more impactful strategies influencing heating or cooling loads, considering thermal comfort. Additionally, rather than basing individual comfort solely on illuminance, these simulations will incorporate a range of KPIs to provide a holistic assessment of visual comfort enhancements.

This chapter even introduced research based on SolarLab energy model to explore the trade-off between thermal and visual comfort, and energy consumption. The analysis reported in this chapter were performed by using a combination of different tools: Sketchup combined with Euclid, Energy+ and python. Considering three different points of the room where the workplanes were located and two different scenarios, results showed that in all scenarios there are not glare issues in desk positions. Furthermore, by the use of shading systems, even if the presence of glare was notably reduced, the PE consumption could increase, ranging from 4.6% in Scenario 2 to 5.6% in Scenario 1a. This is because the reductions in cooling energy achieved through shading were offset by increases in heating and lighting demands. Specifically, the presence of the shelter in summer led to minimal cooling savings, and the production of cooling energy, using air-to-water chillers, was more efficient than heating resulting in energy consumption comparable. Moreover, the demand for artificial lighting significantly increased in Scenarios 1a to 1c. Although adjusting the angle of the slats was feasible, the effectiveness of this adjustment heavily relies on the sun position and its hourly scheduling is limited by the character and text limit for EnergyPlus schedules. For this reason, using a fixed slat angle results in a substantial decrease in desk illuminance and, thus, more frequent reliance on artificial lighting systems. Scenario 2 approach of using modularly adjusted roller blinds resulted in an illuminance pattern that maintains electrical demand for lighting akin to the base case, showcasing the efficacy of this adaptive shading method. The methodology proposed is useful to simulate the effects of two different shading devices and the methodology used to extend the possibility of creating new control strategies extending Energy Plus possibilities (roller shutter is not available in the included shadings systems library). Moreover, the Python script developed was used to create new shading schedule based on the illuminance monitored in three different workplanes. Although simple to use, LabVIEW limits the scalability of this methodology due to its reliance on proprietary, non-open-source software. Additionally,

LabVIEW components, or "VIs", have limitations, and for more complex tasks, Python-based VIs may be required.

The development of a beta version of software designed to estimate local MRT was presented in this chapter. Using simulated building surface temperatures and various geometries, the methodology provided calculates the MRT for four different objects: two representing people and two simulating the GTs in the analysed room. Additionally, an algorithm was developed to estimate MRTs in a rectangular-shaped zone, simulating a simplified version of SolarLab. Using over 6,000 hours of simulation data, these were used to train ML models to estimate people PMV based on MRT values calculated by the two different GTs. Results achieved revealed that estimation of PMVs using MRT_2 performed better compared to MRT_1 indicating that FFANN yields the best performance, having the lowest errors and highest R^2 values. This approach, while based on validated Python libraries, would greatly benefit from validation using real data collected from SolarLab. Additionally, evaluating real-time comfort parameters presents implementation challenges when using Arduino or Raspberry components, as view factor calculations are computationally intensive. These challenges result from the numerous calculations required for the large number of surfaces in the mesh. However, with the recent arrival of real sensors, these issues can be addressed, including testing real-time dashboards using the methods outlined in Chapter 2 of this thesis.

These findings could be valuable for architects, engineers, and environmental designers, offering insights into the complexities of creating energy-efficient, comfortable, and sustainable indoor spaces. Furthermore, this analysis can be used as input to design building automation control systems and to develop advanced control strategies in DTs.

References

- [1] Eini R, Abdelwahed S. A Neural Network-based Model Predictive Control Approach for Buildings Comfort Management. 2020 IEEE International Smart Cities Conference (ISC2), IEEE; 2020, p. 1–7. <https://doi.org/10.1109/ISC251055.2020.9239051>.
- [2] Yang S, Wan MP, Ng BF, Dubey S, Henze GP, Chen W, et al. Model predictive control for integrated control of air-conditioning and mechanical ventilation, lighting and shading systems. *Appl Energy* 2021;297:117112. <https://doi.org/10.1016/j.apenergy.2021.117112>.
- [3] Ribino P, Bonomolo M, Lodato C, Vitale G. A Humanoid Social Robot Based Approach for Indoor Environment Quality Monitoring and Well-Being Improvement. *Int J Soc Robot* 2021;13:277–96. <https://doi.org/10.1007/s12369-020-00638-9>.
- [4] Mckee D. Platform Stack Architectural Framework: An Introductory Guide - A Digital Twin Consortium White Paper. 2023.
- [5] Testasecca T, Lazzaro M, Sirchia A. Towards Digital Twins of buildings and smart energy networks: Current and future trends, 2023. <https://doi.org/10.1109/MetroLivEnv56897.2023.10164035>.
- [6] Fanger PO. Thermal Comfort: Analysis and Applications in Environmental Engineering. Danish Technical Press; 1970.
- [7] UNI EN ISO 7730:2006 Ergonomia degli ambienti termici - Determinazione analitica e interpretazione del benessere termico mediante il calcolo degli indici PMV e PPD e dei criteri di benessere termico locale n.d.
- [8] Huan C, Zhang S, Lin Z. Performance evaluation of mean radiant temperature calculated from inner surface temperatures of envelope with various emissivities. *Build Environ* 2021;206:108334. <https://doi.org/10.1016/j.buildenv.2021.108334>.
- [9] La Gennusa M, Nucara A, Rizzo G, Scaccianoce G. The calculation of the mean radiant temperature of a subject exposed to the solar radiation—a generalised algorithm. *Build Environ* 2005;40:367–75. <https://doi.org/10.1016/j.buildenv.2004.06.019>.
- [10] Ye L, Lv G, Ding Y. Structure optimization of radiant cooling panel: A literature review and assessment. *Energy Build* 2024;310:114097. <https://doi.org/10.1016/j.enbuild.2024.114097>.
- [11] ASHRAE. ASHRAE Handbook – HVAC Systems and Equipment. 2012.
- [12] Choi EJ, Yun JY, Choi YJ, Seo MC, Moon JW. Impact of thermal control by real-time PMV using estimated occupants personal factors of metabolic rate and clothing insulation. *Energy Build* 2024;307:113976. <https://doi.org/10.1016/j.enbuild.2024.113976>.
- [13] Yu Y, Xiang T, Wang D, Yang L. Optimization control strategy for mixed-mode buildings based on thermal comfort model: A case study of office buildings. *Appl Energy* 2024;358:122627.

<https://doi.org/10.1016/j.apenergy.2024.122627>.

- [14] Ren J, Liu J, Zhou S, Kim MK, Miao J. Developing a collaborative control strategy of a combined radiant floor cooling and ventilation system: A PMV-based model. *Journal of Building Engineering* 2022;54:104648. <https://doi.org/10.1016/j.jobbe.2022.104648>.
- [15] Hosamo HH, Nielsen HK, Kraniotis D, Svennevig PR, Svidt K. Improving building occupant comfort through a digital twin approach: A Bayesian network model and predictive maintenance method. *Energy Build* 2023;288. <https://doi.org/10.1016/j.enbuild.2023.112992>.
- [16] Hosamo H, Hosamo MH, Nielsen HK, Svennevig PR, Svidt K. 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* 2023;17:125–71. <https://doi.org/10.1080/17512549.2022.2136240>.
- [17] Desogus G, Frau C, Quaquero E, Rubiu G. From Building Information Model to Digital Twin: A Framework for Building Thermal Comfort Monitoring, Visualizing, and Assessment. *Buildings* 2023;13:1971. <https://doi.org/10.3390/buildings13081971>.
- [18] Khosravi M, Huber B, Decoussemaeker A, Heer P, Smith RS. Model Predictive Control in buildings with thermal and visual comfort constraints. *Energy Build* 2024;306:113831. <https://doi.org/10.1016/j.enbuild.2023.113831>.
- [19] Shum C, Zhong L. Optimizing automated shading systems for enhanced energy performance in cold climate zones: Strategies, savings, and comfort. *Energy Build* 2023;300:113638. <https://doi.org/10.1016/j.enbuild.2023.113638>.
- [20] Lee S-J, Song S-Y. Energy efficiency, visual comfort, and thermal comfort of suspended particle device smart windows in a residential building: A full-scale experimental study. *Energy Build* 2023;298:113514. <https://doi.org/10.1016/j.enbuild.2023.113514>.
- [21] Rabani M, Bayera Madessa H, Nord N. Achieving zero-energy building performance with thermal and visual comfort enhancement through optimization of fenestration, envelope, shading device, and energy supply system. *Sustainable Energy Technologies and Assessments* 2021;44:101020. <https://doi.org/10.1016/j.seta.2021.101020>.
- [22] Baghoolizadeh M, Rostamzadeh-Renani M, Rostamzadeh-Renani R, Toghraie D. Multi-objective optimization of Venetian blinds in office buildings to reduce electricity consumption and improve visual and thermal comfort by NSGA-II. *Energy Build* 2023;278:112639. <https://doi.org/10.1016/j.enbuild.2022.112639>.
- [23] UNI EN 12464-1:2021 Luce e illuminazione - Illuminazione dei posti di lavoro - Parte 1: Posti di lavoro in interni n.d.
- [24] Pan Y, Zhu M, Lv Y, Yang Y, Liang Y, Yin R, et al. Building energy simulation and its application for building performance optimization: A review of methods, tools, and case studies.

Advances in Applied Energy 2023;10:100135. <https://doi.org/10.1016/j.adapen.2023.100135>.

- [25] Energy Plus n.d.
- [26] Seo H, Yun WS. Digital Twin-Based Assessment Framework for Energy Savings in University Classroom Lighting. *Buildings* 2022;12. <https://doi.org/10.3390/buildings12050544>.
- [27] Institution of Engineering and Technology. Digital twins for the built environment - An introduction to the opportunities, benefits, challenges and risks. 2019.
- [28] Cooper JP, Jackson S, Kamojjala S, Owens G, Szana K, Tomic S. Demystifying Digital Twins: Definitions, Applications, and Benefits. *J Am Water Works Assoc* 2022;114:58–65. <https://doi.org/10.1002/awwa.1922>.
- [29] Bitter R, Mohiuddin T, Nawrocki M. LabVIEW: Advanced programming techniques. Crc Press; 2006.
- [30] DIAL GmbH. DIALux n.d.
- [31] Reinhart C, Breton P-F. Experimental Validation of Autodesk® 3ds Max® Design 2009 and Daysim 3.0. *LEUKOS* 2009;6:7–35. <https://doi.org/10.1582/LEUKOS.2009.06.01001>.
- [32] Galatioto A, Beccali M. Aspects and issues of daylighting assessment: A review study. *Renewable and Sustainable Energy Reviews* 2016;66:852–60. <https://doi.org/10.1016/j.rser.2016.08.018>.
- [33] Tregenza PR. The daylight factor and actual illuminance ratios. *Lighting Research & Technology* 1980;12:64–8. <https://doi.org/10.1177/096032718001200202>.
- [34] Reinhart CF, Weissman DA. The daylit area – Correlating architectural student assessments with current and emerging daylight availability metrics. *Build Environ* 2012;50:155–64. <https://doi.org/10.1016/j.buildenv.2011.10.024>.
- [35] IESNA - Lighting measurement e spatial daylight autonomy (draft) Illuminating Engineering Society of North America, New York, NY (2011) n.d.
- [36] Z. Rogers, D. Goldman, D. “Daylighting metric development using daylight autonomy calculations in the sensor placement optimization tool”. Boulder, Colorado, USA: Architectural Energy Corporation: http://www.archenergy.com/SPOT/SPOT_Daylight%20Autonomy%20Report.pdf. 2006 n.d.
- [37] Schiler M. Examples of Glare Remediation Techniques Four buildings. 2009.
- [38] Ahn YN, Kim K Do, Anoop G, Kim GS, Yoo JS. Design of highly efficient phosphor-converted white light-emitting diodes with color rendering indices ($R_1 - R_{15} \geq 95$ for artificial lighting. *Sci Rep* 2019;9:16848. <https://doi.org/10.1038/s41598-019-53269-0>.
- [39] Ashdown I. The Kruithof Curve: A Pleasing Solution. n.d.
- [40] Rizi RA, Eltaweel A. A user detective adaptive facade towards improving visual and thermal comfort. *Journal of Building Engineering* 2021;33:101554.

<https://doi.org/10.1016/j.jobbe.2020.101554>.

- [41] Shi F, You Y, Yang X, Hong X. Annual evaluation of the visual-thermal comfort and energy performance of thermotropic glazing in a reference office room of China. *Build Environ* 2024;254:111378. <https://doi.org/10.1016/j.buildenv.2024.111378>.
- [42] UNI/TS 11300-1:2014 Prestazioni energetiche degli edifici - Parte 1: Determinazione del fabbisogno di energia termica dell'edificio per la climatizzazione estiva ed invernale n.d.
- [43] Decreto Ministeriale 26/06/2015 - Allegato 1 Criteri generali e requisiti delle prestazioni energetiche degli edifici. n.d.
- [44] d'Ambrosio Alfano FR, Pepe D, Riccio G, Vio M, Palella BI. On the effects of the mean radiant temperature evaluation in the assessment of thermal comfort by dynamic energy simulation tools. *Build Environ* 2023;236:110254. <https://doi.org/10.1016/j.buildenv.2023.110254>.
- [45] Sullivan C, Kaszynski A. PyVista: 3D plotting and mesh analysis through a streamlined interface for the Visualization Toolkit (VTK). *J Open Source Softw* 2019;4:1450. <https://doi.org/10.21105/joss.01450>.
- [46] Bogdan M, Walther E, Alecian M, Normale Supérieure Paris-Saclay É. Calcul des facteurs de forme entre polygones-Application à la thermique urbaine et aux études de confort. 2022.
- [47] Tartarini F, Schiavon S. pythermalcomfort: A Python package for thermal comfort research. *SoftwareX* 2020;12:100578. <https://doi.org/10.1016/j.softx.2020.100578>.

Chapter 8

Conclusions

This industrial PhD, conducted in collaboration with Engineering S.p.A., researchers from the CNR, and academics from the University of Palermo, addressed various research objectives from multiple perspectives. The research presented in this thesis primarily focuses on the modelling and optimization of SESs to advance the development of DTs in the energy sector. Between November 2021 and October 2024, interest in DT technology grew exponentially, accompanied by a significant increase in related publications. Numerous European and Italian research projects have set the development of at least one DT as a key research objective. Reflecting this trend, a portion of this thesis is linked to the DigiBUILD European project, coordinated by Engineering, other internal CNR projects and projects followed by the University of Palermo. Participation in these projects facilitated innovative studies, enabled by collaboration with partners from across Europe. This collaboration, which included an international research exchange at the National Technical University of Athens, fostered the continuous exchange of ideas. The interdisciplinary background of the researchers from Athens, comprising both electrical and computer engineers, provided an invaluable opportunity for mutual growth and the development of algorithms and models for smart grids and DTs.

DTs are a key technology rapidly spreading across various fields, including the energy sector. Originally conceived during NASA's Apollo missions in the 1960s as "living models", DTs today enhance the safety, resilience, and efficiency of energy systems. These systems benefit from the models embedded within the DT due to their ability to self-evolve and optimize based on real-world data. Unlike static DMs, DTs are synchronized replicas that use real-time data during system operation to improve model accuracy, detect faults or anomalies, and maximize efficiency. For a system to be classified as a true DT, this synchronization must be automated, ensuring data flows continuously from the real-world system to the digital environment. Advanced applications further automate interactions between the twin and the physical system, enabling real-time control without manual intervention. To provide a comprehensive overview, Chapter 2 of this thesis presented several DT categorization methods, including two based on classification by category and two based on DT maturity levels. After reviewing over 100 works

on DTs applied to energy-related fields, such as smart energy networks, buildings, and individual components, a systematic analysis of 64 papers from 2022 and 2023 was conducted. Despite the promising results shown across all works, only 36% could be classified as true DTs, with the remaining studies representing advanced models or dashboards. This finding highlights a significant gap that must be addressed to meet the expectations set by the EU and the IEA for DT progress. Moreover, given the wide range of potential applications and approaches that can benefit from this technology, this thesis explores multiple case studies to demonstrate both the feasibility and the challenges associated with its implementation.

For this reason, a single chapter was dedicated to outlining the methodology used to develop the DTs in this thesis. This begins by addressing data management, injection, and processing methods, and includes a discussion of both wired and wireless approaches. Then the focus was on the models developed for the various DTs and case studies examined in this thesis. It discussed data-driven models, highlighting different types of ML algorithms, including decision tree-based algorithms, DL methods, and unsupervised algorithms. Then, the physics-based models applied in various case studies using TRNSYS software were presented, with a particular focus on models for absorption chillers, pipes, and other DHN components. Finally, one of the most important aspects of DTs, the methods used to develop the dashboard (e.g., Grafana, LabVIEW, or React), were introduced.

One of the applications that stands to benefit most from DTs is DHNs, as many studies in this review demonstrated how this technology can enhance fault detection in pipes and improve overall system performance. Moreover, as modern DHNs, namely, the fourth and fifth generations, rely on distributed heat sources and prosumers, proactive management is increasingly necessary. In this context, three different studies were conducted in this thesis, focusing on the latest DHN generations. Specifically, these involved the optimization of an existing third-generation DHN through DT technology, the development of a model to analyse the hydraulic and thermal behaviour of a fourth-generation network, and a comparison of solar-driven fourth and fifth-generation networks for both cooling and heating applications. Despite Rio Vena DHN being a third-generation network powered by natural gas, the DT, which integrates GBoost models for boiler efficiency and a genetic optimization algorithm, enabled gas savings of up to 30%, depending on the day. Additionally, the fourth-generation model of a double-loop DHN with prosumers showed that thermal losses could be reduced by 26.4% to 52.8% simply by lowering operational temperatures. The introduction of prosumers led to a 27% reduction in energy consumption for pumping and a 31.3% decrease in thermal energy produced by the CHP. Finally, a study comparing fourth and fifth-generation networks supported by a central solar collector plant revealed that CO₂ emissions could be reduced by

64% to 68%, depending on the scenario, compared to independent heating and cooling production. Notably, thermal losses in the fifth-generation network were reduced by 76% compared to the fourth-generation network. However, while prosumers had a positive impact on energy generation, pumping energy increased by a factor of 1.08 due to inefficiencies in the return loop, which offset the benefits of improved flow rate distribution on the supply side.

The review conducted between the first and second years of this PhD showed that the majority of DT applications are in the domain of electrical engineering, particularly in smart grids or for individual components such as battery DTs. However, as previously highlighted, there remains a gap in research and the implementation of synchronized DTs in real-world case studies. Therefore, one of the studies in this thesis focused on the development of a comprehensive DT for the Emotion microgrid and models for the SOFC DT provided by CNR. In the first case, the DT incorporates clustering models for EVs, enabling the estimation of consumption for all vehicles within the microgrid, as well as algorithms for scheduling and real-time management of EV charging to maximize PV self-consumption. Specifically, using a hybrid GWO-WOA, it was possible to schedule the optimal charging strategy, balancing potential export under both cloudy and sunny conditions in under 25 seconds. Since user interaction is required, inputting preferences through the DT dashboard, response time was a critical factor. For the SOFC DT, various ML models were compared to estimate efficiency based on power modulation. After analysing more than 3,900 hours of SOFC operation, the best-performing models were XGBoost and RF, with the latter achieving a RMSE of 0.31 and an R^2 value of nearly 0.99. This SOFC model will be implemented in future studies within optimization algorithms to manage a larger energy system, including EVs, buildings, and PV systems. This aligns with the future trend toward “DTs of DTs”, where multiple components interact optimally within an integrated energy system.

In addition to smart grid applications, DTs for buildings are emerging as a critical focus, due to the complexity and multiphysics nature of buildings, which include various components. The importance of their optimal management relies not only on the fact that buildings are a significant contributor to global CO₂ emissions but also because people are spending an increasing amount of time in indoor environments. This situation drove the research toward the development of models and DTs aimed at enhancing visual and thermal comfort while considering building energy consumption. All of the studies discussed were conducted in the SolarLab laboratory at the University of Palermo, which features a wide range of sensors and serves as a testbed for various research projects. The first study focused on developing a DT for improving visual comfort through a dashboard implemented in LabVIEW and models developed using Dialux and Daysim. This DT was capable of calculating and displaying key

visual comfort indicators using available illuminance and luminance sensors, while also providing suggestions to users. Another study compared different strategies for controlling various shading systems to reduce visual discomfort. This model-based approach, potentially applicable to building automation systems, successfully reduced glare from 20% in the base case to zero in three out of four scenarios during working hours. Furthermore, advanced control of roller blinds was effective in reducing glare while maintaining optimal natural daylight use, with only a 14% increase in lighting energy, compared to a 154% increase when using slatted blinds. Additionally, an algorithm was developed to identify the distribution of MRT in complex geometries. An analysis was also conducted to determine the optimal placement of MRT sensors for estimating individual PMV using ML algorithms. Results demonstrated that an optimal sensor position was identified, and using a FFANN, the PMV for both individuals in the case study was calculated with an R^2 of up to 0.98.

During the development of this research project, several challenges were encountered. A key issue was data management, which could fail in various ways: malfunctioning sensors, difficulties in installing sensor networks, or challenges in connecting sensors via MQTT to databases. Another challenge was balancing end-user requirements, preferences, and the limitations of research outcomes, as end-users are the primary recipients of this technology. When dealing with highly complex systems, methodologies involving software such as EnergyPlus or TRNSYS may be the only viable approach for creating accurate models. In such cases, end-users must be thoroughly instructed on how the DT can enhance the operation of the real system.

The goal of this thesis was to demonstrate the applicability of DT technology in various contexts, integrating both advanced physics-based models and AI models. For the former, while developing accurate models is challenging for highly complex and articulated systems, they do not require large amounts of data to be built. The latter, by contrast, requires extensive data for training but can be applied to systems with a pre-existing sensor infrastructure. This thesis also demonstrated, through the application of optimization algorithms (e.g., genetic, mathematical, or bio-inspired), that DTs can help achieve decarbonization goals simply by optimizing management, without the need for extensive retrofitting, as seen in the Emotion microgrid and Rio Vena DHN case studies. Furthermore, applications in buildings revealed that it is possible to enhance indoor comfort for occupants through the use of appropriate algorithms implemented in building automation systems.

The findings of this thesis could serve as valuable tools for shaping new policies aimed at the digitalization of the energy sector. The literature review and methodologies presented clarify the concept of DTs and lay the groundwork for the informed adoption of this technology. This

approach can help avoid funding research activities that lack practical benefits, ensuring that DTs achieve their primary goal: enhancing real system performance. Additionally, most methodologies proposed in this work rely on open-source tools, providing a practical foundation for researchers and energy managers seeking to optimize the performance of systems ranging from individual buildings to complex smart energy networks. Finally, findings such as those related to the Rio Vena DHN offer a straightforward methodology for DHN managers to reduce carbon emissions and costs in gas-driven networks. This is particularly important for decarbonization efforts, as most existing networks still rely heavily on fossil fuels.

Still remains room for improvement in the case studies analysed. In some of the applications discussed a real-time implementation is still needed to provide tangible benefits to end-users and contribute to global sustainability. Additionally, focusing on the models presented for actual DTs, their accuracy is expected to improve, as self-evolving DTs based on real-time data will continuously use updated information to train the models.

List of Figures

Chapter 1 Introduction	5
<i>Fig. 1. Summary Schema of the DigiBUILD Project [11]</i>	6
Chapter 2 Digital Twins of Buildings and Energy Systems	12
<i>Fig. 2. Example of DT Reference Architecture according to IBM [5].</i>	13
<i>Fig. 3. Paper Publication Trend.</i>	14
<i>Fig. 4. Kritzing Classification revised schema [6].</i>	15
<i>Fig. 5. Digital Twin of University of Texas Campus [25]</i>	22
Chapter 3 Methodology for Developing Digital Twins	41
<i>Fig. 6. SolarLab Experimental Facility in University of Palermo.</i>	42
<i>Fig. 7. DigiBUILD DT Architecture.</i>	43
<i>Fig. 8. Example of Correlation Matrix Heatmap for SOFC (left) and EV (right).</i>	45
<i>Fig. 9. Example of code for XGBoost model for the SOFC</i>	47
<i>Fig. 10. Example of code for ANN model for the SOFC</i>	49
<i>Fig. 11 Process of Clustering Based on k-means</i>	51
<i>Fig. 12. Macro of the fifth-generation user substation.</i>	59
<i>Fig. 13. Example of Optimization Algorithm Script using PyGAD</i>	64
<i>Fig. 14. Veolia Digital Twin Dashboard in Grafana</i>	66
<i>Fig. 15. A Section of the Emotion Digital Twin Dashboard in React</i>	67
<i>Fig. 16. Dashboard of DT of SolarLab for Visual Comfort Monitoring</i>	68
Chapter 4 Digital Twins and Models of District Heating Networks	74
<i>Fig. 17. I/O Scheme for Service s3.2.2 for CP – Rio Vena (VEOLIA)</i>	77
<i>Fig. 18. CP Rio Vena Optimization Function</i>	80
<i>Fig. 19. Average and Extreme Values of Temperature and Humidity</i>	81
<i>Fig. 20. Sky View of Rio Vena District Heating Network Cluster of Buildings [17]</i>	81
<i>Fig. 21. Daily behaviour of average temperature in Rio Vena network</i>	83
<i>Fig. 22. Hourly behaviour in summer and winter solstice of temperatures and power</i>	84
<i>Fig. 23. TRNSYS Tank Model for Rio Vena.</i>	84
<i>Fig. 24. Correlation Heatmap for Boilers 1-2 (left) and Boiler 3 (right) operation.</i>	85
<i>Fig. 25 Boiler Efficiency ML Model Behaviour for Boilers 1-2 Operation</i>	87
<i>Fig. 26 Boiler Efficiency ML Model Behaviour for Boiler 3 Operation</i>	87
<i>Fig. 27. Optimization Results During a Generic Day of Boiler 3 Operation (right) and Boiler 1-2 (left).</i>	88
<i>Fig. 28 Digital Twin Dashboard for Data Monitoring and KPIs Visualization.</i>	89
<i>Fig. 29 Rio Vena Digital Twin Simulation Page</i>	90
<i>Fig. 30. Case study DHN model developed on TRNSYS</i>	92
<i>Fig. 31. Case Study Cluster</i>	93
<i>Fig. 32. Cumulative user loads and outdoor air temperature in the case study</i>	94

<i>Fig. 33. Overall thermal load for all the buildings in the case study</i>	95
<i>Fig. 34. Trend of pressure for each node of the network in the design phase</i>	98
<i>Fig. 35. Hourly pressures, on the sides of the pump, during a year in the DHN</i>	99
<i>Fig. 36. Hourly pressures during standard weeks in the DHN</i>	100
<i>Fig. 37. Hourly temperatures during standard weeks in the DHN</i>	101
<i>Fig. 38. Pressures comparison during standard hours</i>	102
<i>Fig. 39. Temperatures comparison during standard hours</i>	102
<i>Fig. 40. Temperatures of supply-side nodes operating at 3 different temperatures (100 °C, 80 °C, 60 °C)</i>	103
<i>Fig. 41. Pressures distribution on a winter and on a summer week in Scenario 2</i>	105
<i>Fig. 42. Solar Loop Model Developed on TRNSYS</i>	108
<i>Fig. 43. Solar Loop Operation Modes</i>	109
<i>Fig. 44. District Heating Network model in TRNSYS</i>	111
<i>Fig. 45. Hourly Trend of Ambient Temperature and Relative Humidity of Palermo</i>	112
<i>Fig. 46. Trend of Global Solar Radiation on the Horizontal Plane in Palermo</i>	112
<i>Fig. 47. Representation of buildings cluster of the case study.</i>	112
<i>Fig. 48. Hourly Heating and Cooling demands of the entire building cluster</i>	113
<i>Fig. 49. PE and CO₂ Emissions by User in Scenario 0.</i>	116
<i>Fig. 50. Yearly pressures trend in Scenario 1</i>	117
<i>Fig. 51. Yearly pressures trend in Scenario 2a</i>	119
<i>Fig. 52. Heat losses trend in Scenario 1 and 2a</i>	119
<i>Fig. 53. Pipes heat losses trend in Scenario 1 and 2a.</i>	120
<i>Fig. 54. Comparison between improvement scenarios</i>	121
<i>Fig. 55. PTES average temperature in different scenarios</i>	122
<i>Fig. 56. Comparison of pressure losses between Scenario 2a and Scenario 2b</i>	122
<i>Fig. 57. Comparison of Pressure Trend in the last week of January and July</i>	123
Chapter 5 Digital Twin of a Microgrid	131
<i>Fig. 58. DT Architecture of Emotion Digital Twin</i>	134
<i>Fig. 59. The Emotion Microgrid Case Study [21]</i>	140
<i>Fig. 60. Microgrid Consumption and Production Behaviour in a Winter and Summer Week</i>	141
<i>Fig. 61. Correlation Matrixes for 7 Vehicles from Emotion Microgrid</i>	142
<i>Fig. 62. Analysis of different clustering methods</i>	143
<i>Fig. 63. Analysis of different clustering methods considering position</i>	144
<i>Fig. 64. Elbow method application (left) and k-means result considering only 4 clusters (right)</i>	144
<i>Fig. 65. Weekly Power trends for optimal EV charging compared to power import and export.</i>	147
<i>Fig. 66. Power trends and SoCs on winter days considering EV user preferences.</i>	148
<i>Fig. 67. Sections of the Emotion Digital Twin Dashboard: (a) vehicle selection for optimization; (b) results visualization; (c) DSS page.</i>	149
Chapter 6 Digital Twin of a Solid Oxide Fuel Cell	154
<i>Fig. 68. SOFC DT Architecture</i>	157
<i>Fig. 69. SOFC Experimental Setup.</i>	159

<i>Fig. 70. Behaviour of Power and Efficiency in the Operational Experimental Period.</i>	160
<i>Fig. 71. Polynomial Fits on Efficiency–Power Graph (a) and Correlations on Experimental Data (b).</i>	161
<i>Fig. 72. Models and Real Data Behaviour During the Entire Experimental Period.</i>	162
<i>Fig. 73. Models and Real Data Behaviour in Different Time Horizons.</i>	162
Chapter 7 Buildings Digital Twins and Models	170
<i>Fig. 74. Pictures of the laboratory</i>	173
<i>Fig. 75. Pictures of the instrumentation used to collect the measures</i>	173
<i>Fig. 76. Three-Dimensional Representation of the EnergyPlus Model</i>	175
<i>Fig. 77 Architecture of DT of the Laboratory</i>	177
<i>Fig. 78. Kruithof Curve</i>	178
<i>Fig. 79. Examples of LabView dashboards of the Laboratory DT.</i>	180
<i>Fig. 80. Pictures of the Dialux model with real colours (left) and false colour (right)</i>	181
<i>Fig. 81. Flow Diagram of Methodology for Investigating Effects of Shading Systems.</i>	183
<i>Fig. 82. Schema of shading system layout for each scenario.</i>	184
<i>Fig. 83. Section of the Script for Creating Custom Schedule.</i>	186
<i>Fig. 84. Energy Plus Model with Shadings Covering the 40% of the Windows</i>	187
<i>Fig. 85. Indoor and outdoor conditions in Base Case</i>	188
<i>Fig. 86. Energy behaviour in Base Case.</i>	189
<i>Fig. 87. Comparative Analysis of Hourly Energy Contributions in a Winter Day.</i>	189
<i>Fig. 88. Comparative Analysis of Hourly Energy Contributions in a Summer Day.</i>	190
<i>Fig. 89. Illuminance Maps at 11:00 AM in Winter Day: Base Case Versus Scenario 1.</i>	191
<i>Fig. 90. Illuminance Maps at 11:00 AM in Summer Day: Base Case Versus Scenario 1.</i>	192
<i>Fig. 91. PMV trend in a typical winter and summer day during work hours.</i>	193
<i>Fig. 92. PMV trend in an especially cold day (left) and hot day (right) during work hours.</i>	193
<i>Fig. 93. Analysis of Roller Blind Impact During Typical Winter (Above) and Summer (Below) Days.</i>	194
<i>Fig. 94. Block diagram showing the workflow adopted for the development of the software.</i>	197
<i>Fig. 95. SolaLab Model for MRT Simplified Algorithm.</i>	199
<i>Fig. 96. Flowchart for the estimation of the indoor parameter needed my ML algorithms.</i>	199
<i>Fig. 97. 3D spatial view of sensors (cubes) and individuals (parallelepipeds)</i>	201
<i>Fig. 98. MRT on January 21st at different times.</i>	201
<i>Fig. 99. MRT on the 31st of July at different hours</i>	202
<i>Fig. 100. PMV trend on the 21st of January and 31st of July at different hours</i>	202
<i>Fig. 101. PMV trend on the equinoxes at different hours.</i>	203
<i>Fig. 102. 3D Model of SolarLab in Sketchup (left) and after being meshed by PyVista (right)</i>	204
<i>Fig. 103. View Factors Distribution Inside SolarLab</i>	204
<i>Fig. 104. People (P1 and P2) and Globe Thermometers (S1 and S2) position inside SolarLab</i>	205
<i>Fig. 105. Accuracy of Models for PMV Estimation Using Total Dataset</i>	206
Chapter 8 Conclusions	214

List of Tables

Chapter 1 Introduction	5
Chapter 2 Digital Twins of Buildings and Energy Systems	12
<i>Table 1. DT Classification according to DTC.</i>	<i>16</i>
<i>Table 2. DT Maturity Level according to IET [4].</i>	<i>17</i>
<i>Table 3. DT Maturity level according to Cooper et al. [7].</i>	<i>18</i>
<i>Table 4. Overview on Buildings-related DTs.</i>	<i>19</i>
<i>Table 5. Overview on smart grid-related DTs.</i>	<i>24</i>
<i>Table 6. Overview on energy component-related DTs.</i>	<i>28</i>
<i>Table 7. Overview of classification results.</i>	<i>31</i>
Chapter 3 Methodology for Developing Digital Twins	41
Chapter 4 Digital Twins and Models of District Heating Networks	74
<i>Table 8. Metrics Comparison Between ANN and Gboost Models</i>	<i>86</i>
<i>Table 9. User and cumulative thermal load</i>	<i>96</i>
<i>Table 10. Electric Energy requested by the pumps</i>	<i>99</i>
<i>Table 11. Pressure drops in network branches</i>	<i>100</i>
<i>Table 12. Heat losses in the pipes during winter</i>	<i>103</i>
<i>Table 13. Main results for each scenario</i>	<i>106</i>
<i>Table 14. CO₂ emissions in the main cases of the case study</i>	<i>107</i>
<i>Table 15. Annual energy demand and maximum power of each user in the building cluster</i>	<i>113</i>
<i>Table 16. Overall Results for Scenario 0.</i>	<i>116</i>
<i>Table 17. Overall Results for Scenario 1.</i>	<i>117</i>
<i>Table 18. Overall Results for Scenario 2a.</i>	<i>118</i>
<i>Table 19. Overall Results for Scenario 2b.</i>	<i>120</i>
<i>Table 20. Overall Results for Scenario 2c.</i>	<i>121</i>
Chapter 5 Digital Twin of a Microgrid	131
<i>Table 21. ANN Model Hyperparameters</i>	<i>136</i>
<i>Table 22. Metrics of ANN Model for EVs Energy Consumption Prediction</i>	<i>143</i>
<i>Table 23. Optimization Results in Winter Days</i>	<i>145</i>
<i>Table 24. Optimization Results in Summer Days</i>	<i>146</i>
Chapter 6 Digital Twin of a Solid Oxide Fuel Cell	154
<i>Table 25. SOFC Operational Parameters.</i>	<i>159</i>
<i>Table 26. Overview of Model Score Results.</i>	<i>163</i>
<i>Table 27. Model Scores in Different Periods.</i>	<i>164</i>
<i>Table 28. Comparison of ML Models: Advantages, Disadvantages, and Limitations.</i>	<i>165</i>
Chapter 7 Buildings Digital Twins and Models	170

<i>Table 29. Information of SolarLab thermal envelope</i>	<i>174</i>
<i>Table 30. Annual Heating and Cooling Energy Demand Across Various Scenarios.</i>	<i>190</i>
<i>Table 31. Analysis of Glare Occurrence Across Scenarios</i>	<i>195</i>
<i>Table 32. Comparison of Primary Energy Consumption Among Scenarios</i>	<i>195</i>
Chapter 8 Conclusions	214

List of Research Products

Scientific Articles:

1. Catrini P, Testasecca T, Buscemi A, Piacentino A. Exergoeconomics as a Cost-Accounting Method in Thermal Grids with the Presence of Renewable Energy Producers. *Sustainability (Switzerland)* 2022;14. <https://doi.org/10.3390/su14074004>.
2. Catrini P, Testasecca T, La Villetta M, Morale M, Piacentino A. Thermodynamic-Based Method for Supporting Design and Operation of Thermal Grids in Presence of Distributed Energy Producers. *Journal of Sustainable Development of Energy, Water and Environment Systems* 2023;11. <https://doi.org/10.13044/j.sdewes.d11.0459>.
3. Ciulla G, Testasecca T, Mangione S, Longo S, Tupenaite L. Timber Houses in the Mediterranean Area: A Challenge to Face, 2023, p. 77–86. https://doi.org/10.1007/978-3-031-33906-6_7.
4. Testasecca T, Stamatopoulos S, Lazzaro M, Sarvas E. Recent advances on data-driven services for smart energy systems optimization and pro-active management. 2023 IEEE International Workshop on Metrology for Living Environment (MetroLivEnv), 2023. <https://doi.org/10.1109/MetroLivEnv56897.2023.10164056>.
5. Testasecca T, Lazzaro M, Sirchia A. Towards Digital Twins of buildings and smart energy networks: Current and future trends, 2023. <https://doi.org/10.1109/MetroLivEnv56897.2023.10164035>.
6. Testasecca T, Ferraro M, Arnone D, Lazzaro M, Beccali M, Ciulla G. Classification and maturity of Digital Twins for energy systems: A systematic review 2023. *Under Review in Energy and Buildings (Submitted on May 13th 2024)*
7. Testasecca, T., Catrini, P., Beccali, M., & Piacentino, A. (2023b). Dynamic simulation of a 4th generation district heating network with the presence of prosumers. *Energy Conversion and Management X*, 20, 100480. <https://doi.org/10.1016/j.ecmx.2023.100480>.
8. Testasecca T, Catrini P, Alba GE, La Villetta M, Beccali M, Piacentino A. The Role of Solar Heating and Cooling Systems in Modern District Heating Networks with Prosumers. *SDEWES 2024, Rome (Italy): 2024. (Selected for Special Issue in Renewable Energy)*.
9. Testasecca T, Maniscalco MP, Brunaccini G, Airò Farulla G, Ciulla G, Beccali M, et al. Toward a Digital Twin of a Solid Oxide Fuel Cell Microcogenerator: Data-Driven Modelling. *Energies (Basel)* 2024;17:4140. <https://doi.org/10.3390/en17164140>.
10. Bonomolo M, Testasecca T, Buscemi A, Alberto Munafò FL, Beccali M. Digital Twin-Based Assessment Framework For Monitoring Visual Comfort. 2024 IEEE 22nd Mediterranean Electrotechnical Conference (MELECON), IEEE; 2024, p. 1139–45. <https://doi.org/10.1109/MELECON56669.2024.10608560>.
11. Bonomolo M, Testasecca T, Beccali M. Using shading systems to enhance energy efficiency, visual and thermal comfort: a case study. *SDEWES 2024, Rome (Italy): 2024. (Selected for Special Issue in e-Prime)*.

12. Bonomolo M, Testasecca T, Cirrincione L, Beccali M, Scaccianoce G. Development of a methodology for assessing indoor punctual thermal comfort conditions. *EEEIC 2024*, Rome, Italy, 2024, pp. 1-6, doi: 10.1109/EEEIC/ICPSEurope61470.2024.10750969.
13. Testasecca T, Stamatopoulos S, Natalini A, Lazzaro M, Capizzi CM, Sarma E, et al. Implementing Digital Twins for Enhanced Energy Management in Three Case Studies. *2024 IEEE International Workshop on Metrology for Living Environment (MetroLivEnv)*, IEEE; 2024, p. 343–8. <https://doi.org/10.1109/MetroLivEnv60384.2024.10615496>.
14. Baby BA, Alberto Munafò FL, Testasecca T, Beccali M, Ferraro M, Poma G. Digitalizing Pipeline Network for Hydrogen-Blended Natural Gas Distribution Assessments. *2024 IEEE International Workshop on Metrology for Living Environment (MetroLivEnv)*, IEEE; 2024, p. 349–54. <https://doi.org/10.1109/MetroLivEnv60384.2024.10615772>.
15. Bonomolo M, Guarino S, Leone G, Testasecca T, Coga A, Beccali M. Assessment of performance of Hybrid Heat Pump system during winter season in different climatic conditions. *EEEIC 2024*, Rome, Italy, 2024, pp. 1-6, doi: 10.1109/EEEIC/ICPSEurope61470.2024.10750969.

Scientific Reports and Deliverables:

1. Beccali M, Bonomolo M, Scaccianoce G, Cirrincione L, Testasecca T. INTERMEDIATE REPORT UNIPA - T8.3.5 - A.PA.2 (M12-M18) REPORT ON DEVELOPMENT OF A TOOL FOR LOW-COST PORTABLE DEVICES BASED ON MACHINE LEARNING ALGORITHM 2024. Project NEST: Project code PE0000021, Concession Decree No. 1561 of 11.10.2022 adopted by Ministero dell'Università e della Ricerca
2. Beccali M, Bonomolo M, Scaccianoce G, Cirrincione L, Testasecca T, Sharif B. SPOKE 8- WP3- TASK 8.3.5 – ACTIVITY A.PA.2 (M24) DELIVERABLE D.PA.2– DEVELOPMENT OF A TOOL FOR LOW-COST PORTABLE DEVICES BASED ON A MACHINE LEARNING ALGORITHM 2024. Project NEST: Project code PE0000021, Concession Decree No. 1561 of 11.10.2022 adopted by Ministero dell'Università e della Ricerca
3. Sarma E, Panagiotis K, Stamatopoulos S, Konstantinos Touloumis, Sofía Mulero Palencia, José L. Hernández García, et al. D3.1: 'First wave' of DigiBUILD AI-based data-driven services for the built environment. 2023. EU Project DigiBUILD: grant agreement no. 101069658.
4. Sarma E, Stamatopoulos S, Spiliotis V, Lazzaro M, Natalini A, Testasecca T, et al. D3.2: "Second wave" of DigiBUILD AI-based data-driven services for the built environment 2024. EU Project DigiBUILD: grant agreement no. 101069658.
5. Sarma E, Stamatopoulos E, Spiliotis E, Mulero S, Hernandez J, Natalini A, et al. D3.3: Final release of DigiBUILD AI-based data-driven services for the built environment 2024. EU Project DigiBUILD: grant agreement no. 101069658

-
6. Antonino Sirchia, Milos Crvic, Marilena Lazzaro, Andrea Natalini, Tancredi Testasecca, Mathias Bouquerel, et al. Deliverable 4.1: Digital Twin Suite and cloud-based DigiBUILD data toolbox (Alpha version). 2023. EU Project DigiBUILD: grant agreement no. 101069658.
 7. Lazzaro M, Capizzi C, Natalini A, Testasecca T, Sayegh H, García JLH, et al. Deliverable D4.3: Digital Twin Suite and cloud-based DigiBUILD data toolbox (final version) 2024. EU Project DigiBUILD: grant agreement no. 101069658.
 8. Lazzaro M, Capizzi C, Natalini A, Testasecca T, Sayegh H, García JLH, et al. D4.2: Digital Twin Suite and cloud-based DigiBUILD data toolbox (Beta version) 2024. EU Project DigiBUILD: grant agreement no. 101069658.
 9. Serena Serroni, Gian Marco Revel, Sara Casaccia, Nicole Morresi, Vittoria Cipollone, Edoardo Mancinelli, et al. D5.3: Pilots' execution documentation – Pre-pilot phase. 2023. EU Project DigiBUILD: grant agreement no. 101069658.
 10. Beccali M, Lo Brano V, Biondi A, Buscemi A, Guarino S, Testasecca T. TASK 8.3.4 Development of materials, technologies, innovative systems and algorithms to increase the exploitation of RES, sustainability, resilience and flexibility in energy end uses. ACTIVITY A.PA.1 DYNAMIC MODELS OF RTC AND BTES STATE OF THE ART ASSESSMENT AND PRELIMINARY SURVEY OF THE SOFTWARE PLATFORM FOR THE IMPLEMENTATION OF THE DT. 2024. Project NEST: Project code PE0000021, Concession Decree No. 1561 of 11.10.2022 adopted by Ministero dell'Università e della Ricerca
 11. Beccali M, Brano V Lo, Biondi A, Buscemi A, Guarino S, Testasecca T. SPOKE 8 - M3: Consolidation of NEST EXTENDED - PARTNERSHIP: planning phase, M8.4: Design and development of innovative materials, components and systems, PART 1, D.PA.2: Development and validation of lumped parameter dynamic models (M24) 2024. Project NEST: Project code PE0000021, Concession Decree No. 1561 of 11.10.2022 adopted by Ministero dell'Università e della Ricerca

Appendix

List of Acronyms

4GDH	Fourth Generation District Heating
5GDH	Fifth Generation District Heating
AI	Artificial Intelligence
ANN	Artificial Neural Networks
API	Application Programming Interface
BEF	Building Energy Flexibility
BEV	Battery Electric Vehicle
BIM	Building Information Model
CCT	Correlated Colour Temperature
CFD	Computational Fluid Dynamics
CGO	Chaos Game Optimization
CHP	Combined Heat and Power Plant
CNR	National Research Council
CO ₂	Carbone Dioxide
COP	Coefficient of Performance
DA	Daylight Autonomy
DBI	Davies-Bouldin Index
DES	Distributed Energy Sources
DF	Daylight Factor
DGP	Daylight Glare Probability
DH	District Heating
DHC	District Heating and Cooling
DHN	District Heating Network
DHW	Domestic Hot Water
DL	Deep Learning
DM	Digital Model
DN	Nominal Diameter
DNN	Deep Neural Network
DS	Digital Shadow
DSS	Decision Support System
DT	Digital Twin
DTC	Digital Twin Consortium
DTS	Digital Twin System
EER	Energy Efficiency Ratio
EES	Engineering Equation Solver

EP	EnergyPlus
ES	Evolutionary Strategy
EU	European Union
EVs	Electric Vehicle
FCs	Fuel Cell
FEM	Finite Element Method
FFANN	Feedforward artificial neural network
GA	Genetic Algorithm
GBoost	Gradient Boosting
gCO ₂	grams of carbon dioxide
GHG	Greenhouse Gas
GT	Globe Thermometer
GUI	Graphic User Interface
GWO	Grey Wolf Optimization
GWO-WOA	Grey Wolf Optimizer–Whale Optimization Algorithm
HP	Heat Pump
HVAC	Heating, Ventilation, and Air Conditioning
IAQ	Indoor Air Quality
IEA	International Energy Agency
IET	Institution of Engineering and Technology
IGWO	Improved Grey Wolf Optimization
IoT	Internet of Things
IT	Information -Technology
KPI	Key Performance Indicator
LHV	Lower Heat Value
LSTM	Long Short Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MD	Model Dashboard
ML	Machine Learning
MLP	Multilayer Perceptron
MPC	Model Predictive Control
MRT	Mean Radiant Temperature
MSE	Mean Squared Error
NG	Natural Gas
NH ₃ -H ₂ O	Ammonia-Water
NMG	Networked Microgrid
NN	Neural Network
NSGA	Non-dominated Sorting Genetic Algorithms
PC	Personal Computer

PE	Primary Energy
PMV	Predicted Mean Vote
PPD	Predicted Percentage of Dissatisfied
PS	Pattern Search
PSO	Particle Swarm Optimisation
PTES	Pit Thermal Energy Storage
PV	Photovoltaic
RES	Renewable Energy Sources
RF	Random Forest
RMSE	Root Mean Square Error
RNN	Recurrent Neural Networks
RTDS	Real-Time Digital Simulator
SCADA	Supervisory Control and Data Acquisition
sDA	Spatial Daylight Autonomy
SES	Smart Energy System
SoA	State of the Art
SoC	State of Charge
SOFC	Solid Oxide Fuel Cell
SS	Silhouette Score
TCL	Thermostatically Controlled Loads
TES	Thermal Energy Storage
UGR	Unified Glare Rating
WOA	Whale Optimization Algorithms
WP	Work Package
WS	Weighted Score
XGboost	Extreme Gradient Boosting

Acknowledgements

This research project involved many individuals who contributed significantly to its development. Special thanks go to the European DigiBUILD and NEST projects for the opportunity to explore such innovative and engaging research topics.

First, I sincerely thank my academic supervisors. Prof. Ciulla provided invaluable guidance in introducing me to academia, from writing my first papers to conducting exams and lectures, fostering my personal and professional growth. Prof. Beccali, a constant mentor, provided unwavering support and thoughtful advice. His blind trust in my ability to handle various responsibilities proved to be a foundation for all my achievements during this PhD.

Collaboration with non-academic partners enriched my experience with a versatile education. I am grateful to Eng. Ferraro for instilling a critical and analytical approach to research and to Eng. Arnone for teaching me the qualities of effective leadership and project management through his kindness and the opportunities provided during DigiBUILD. I also thank my close colleagues, Marilena and Andrea, whose support and collaboration strengthened both the project outcomes and our personal connections. I also want to thank Prof. Vangelis, along with Stathis and other NTUA friends, for the opportunity to join them in Athens.

I extend my heartfelt thanks to Prof. Bonomolo for her continuous guidance and dedication throughout this journey (and beyond!), and to Prof. Catrini, for introducing me to the world of research. I am also grateful to Dr. Guarino for being an excellent advisor, as well as to Prof. Lo Brano and Dr. Buscemi for their support as well as their humour.

Lastly, I deeply thank my closest colleagues and friends: Giuseppe, a constant presence since my undergraduate studies; Ben, my first office roommate; Filippo, my first thesis student and now a valued colleague; and Hamid and Anson, whose fellowship has been invaluable.

Heartfelt thanks go to Roberta, my main support and number one fan, who stood by me even in the most challenging moments. I also thank her and my enlarged family for their patience and understanding, especially when my numerous commitments required significant sacrifices.