Ref. Ares(2024)833604 - 04/02/2024

DigiBUILD

D3.2: 'Second wave' of DigiBUILD AI-based data-driven services for the built environment

WP3 – DigiBUILD AI-based data-driven services for the built environment

January 2024

Funded by the European Union

High-Quality Data-Driven Services for a Digital Built Environment towards a Climate-Neutral Building Stock

Disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor CINEA can be held responsible for them.

Grant Agreement Number	101000158	Acronym	DIGIBUILD
Full Title	High-Quality Data-Driven Services for a Digital Built Environment towards a Climate-Neutral Building Stock		
Topic	HORIZON-CL5-2021-D4-01-03 | Efficient, sustainable and inclusive energy use		
Funding scheme	HORIZON-IA: Innovation Action		
Start Date	June 1st, 2022	Duration	36 months
Project URL	https://digibuild-project.eu/		
Project Coordinator	ENG		
Deliverable	D3.2 - 'Second wave' of DigiBUILD AI-based data-driven services for the built environment		
Work Package	WP3 – DigiBUILD AI-based data-driven services for the built environment		
Delivery Month (DoA)	M20	Version	1.0
Actual Delivery Date	04/02/2024		
Nature	R	Dissemination Level	PU
Lead Beneficiary	NTUA		
Authors	Elissaios Sarmas, Stathis Stamatopoulos, Vaggelis Spiliotis, Daniela Stoian [NTUA], Marilena Lazzaro, Andrea Natalini, Tancredi Testasecca [ENG], Sofía Mulero Palencia, José L. Hernández García [CARTIF], Vittoria Cipollone [UNIVPM], Dimitra Tzani [IEECP]		
Quality Reviewer(s):	Nicole Morresi [UNIVPM], Petteri Rekoma [FVH]		

Keywords	Data-driven Services, Built Environment, Artificial Intelligence, Energy Management, Thermal Comfort, Energy Efficiency, Climate Resilience

Document history

Ver.	Date	Description	Author(s) name
0.1	01/11/2023	Draft ToC	Elissaios Sarmas [NTUA]
0.2	06/11/2023	Final ToC	Elissaios Sarmas [NTUA]
0.31	10/01/2024	1st Round of Contributions	Elissaios Sarmas, Stathis Stamatopoulos, Vaggelis Spiliotis, Daniela Stoian [NTUA], Marilena Lazzaro, Andrea Natalini, Tancredi Testasecca [ENG], Sofia Mulero [CARTIF], Vittoria Cipollone [UNIVPM], Dimitra Tzani [IEECP]
0.32	16/01/2024	2nd Round of Contributions	Elissaios Sarmas, Stathis Stamatopoulos, Vaggelis Spiliotis, Daniela Stoian [NTUA], Marilena Lazzaro, Andrea Natalini, Tancredi Testasecca [ENG], Sofia Mulero [CARTIF], Vittoria Cipollone [UNIVPM], Dimitra Tzani [IEECP]
0.4	22/01/2024	Internal peer review	Nicole Morresi [UNIVPM], Petteri Rekoma [FVH]

0.5	24/01/2024	Minor changes	Stathis Stamatopoulos, Daniela Stoian [NTUA], Tancredi Testasecca [ENG], Sofia Mulero [CARTIF], Vittoria Cipollone [UNIVPM], Dimitra Tzani [IEECP]
0.9	01/02/2024	Quality check	Diego Arnone [ENG]
1.0	04/02/2024	Final version	Diego Arnone [ENG]

Preface

Traditional silo approaches, where stakeholders manage their own data, must be replaced by digital and smart buildings, merging heterogeneous data sources, and placing the stakeholders at the core of these buildings. DigiBUILD will catalyse this much-needed transformation by making use of high-quality data and next generation digital building services, supporting the deployment of EU-wide Framework for a Digital Building Logbook.

An inclusive environment for multi-stakeholder knowledge exchange (based on European Bauhaus initiative) will be applied to co-design end-user-oriented services. DigiBUILD will provide an open, interoperable and cloud-based toolbox to transform current 'silo' buildings into digital, interoperable and smarter ones, based on consistent and reliable data, and support better-informed decision-making for performance monitoring & assessment, planning of building infrastructure, policy making and de-risking of investments. It will be built on top of existing platforms and common EU initiatives, aiming at an Energy Efficient Building Data Space, based on standard cloud-data platform frameworks (FIWARE) and Data Space initiatives (GAIA-X and IDSA). On top of this advanced data governance framework, we will create AI-based data analytics and Digital Building Twins based on high-quality data, aiming at facilitating transparency, trust, informed decision-making and information sharing within the built environment and construction sector, which will be deployed across 10 real-world conditions (TRL 8). DigiBUILD will contribute to the uptake of digital technologies in the build sector to better align the EU Member States' long-term renovation strategies with the EPBD requirements on decarbonisation, and on a path towards a climate-neutral building stock by 2050.

Funded by
the European Union

Who We Are

	Participant Name	Short Name	Country Code	Logo
1	ENGINEERING – INGEGNERIA INFORMATICA SPA	ENG	IT	
2	NATIONAL TECHNICAL UNIVERSITY OF ATHENS	NTUA	EL	
3	FUNDACION CARTIF	CARTIF	ES	
4	UNIVERSITA POLITECNICA DELLE MARCHE	UNIVPM	IT	
5	CNET CENTRE FOR NEW ENERGY TECHNOLOGIES SA	EDP	PT	
6	VEOLIA SERVICIOS LECAM SOCIEDAD ANONIMA UNIPERSONAL	VEOLIA	ES	
7	HERON SINGLE MEMBER S.A. ENERGY SERVICES - HERON ENERGY S.A.	HERON	EL	
8	FOCCHI SPA	FOCCHI	IT	
9	FORUM VIRIUM HELSINKI OY	FVH	FI	
10	MUNICIPIUL IASI	IASI	RO	
11	SITTA RESEARCH SRL	SITTA	RO	
12	ELECTRICITE DE FRANCE	EDF	FR	
13	EMOTION SRL	EMOT	IT	
14	INSTITUTE FOR EUROPEAN ENERGY AND CLIMATE POLICY STICHTING	IEECP	NL	
15	CWARE APS	CWARE	DK	
16	EUROHEAT & POWER	EHP	BE	
17	UNIVERSITY COLLEGE LONDON	UCL	UK	

Table of Contents

Funded by
the European Union

Figures

Funded by
the European Union

Funded by
the European Union

Tables

Funded by the European Union

List of Acronyms

AC	Air Conditioning
EC	European Commission
AHP	Analytical Hierarchy Process
AHU	Air Handling Unit
AI	Artificial Intelligence
API	Application Programming Interface
ARIMA	AutoRegressive Integrated Moving Average
BEV	Battery Electric Vehicles
BM	Building Manager
BMS	Building Management System
BPI	Building Performance Indicator
CPC	Comfort Performance Contract
DHN	District Heating Network
DREEM	Dynamic high-Resolution dEmand-sidE Management
DT	Digital Twin
EC	European Commission
EEM	Energy Efficiency Measures
EU	European Union
EV	Electric Vehicle
GA	General Assembly
GHG	Greenhouse Gas
GUI	Graphical User Interface
HI	Heat Index
HVAC	Heating, Ventilation, and Air Conditioning
IAQ	Indoor Air Quality

IEQ	Indoor Environmental Quality
LCSE	Levelised Cost of Saved Energy
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MCDA	Multi-Criteria Decision Analysis
ML	Machine Learning
MLP	Multi-Layer Perceptron
MLR	Multiple Linear Regression
MSE	Mean Squared Error
MSE	Mean Square Error
NN	Neural Network
NPV	Net Present Value
NSGA	Non-dominated Sorting Genetic Algorithm
OLS	Ordinary Least Squares
PMV	Predicted Mean Vote
PMV	Predicted Mean Vote
PP	Payback Period
PV	Photovoltaic
ReLU	Rectified Linear Unit
RES	Renewable Energy Source
RES	Renewable Energy Source
REST	Representational State Transfer
RMSE	Root Mean Square Error
SCOP	Seasonal Coefficient of Performance
SEER	Seasonal Energy Efficiency Ratio
SoC	State of Charge
SPI	Standard Precipitation Index

sPMV	Simplified Predicted Mean Vote
SuDS	Sustainable Drainage Systems
TSO	Transmission System Operator
TSV	Thermal Sensation Vote
VE SoC	Vehicle System on Chip
WCI	Wind Chill Index
WP	Work Package
XGBoost	eXtreme Gradient Boosting

Funded by
the European Union

Executive Summary

Deliverable 3.2 aims at providing an update on the developmental activities undertaken within Work Package 3 (WP3) of the DigiBUILD project, titled 'DigiBUILD AI-based Data-driven Services for the Built Environment'. It represents the second instalment in a tripartite series of DigiBUILD services, which collectively aim at chronicling the research and development progress of AI-infused and data-driven methodologies and implementations throughout the project's timeline (Months 12, 20, and 28). These services are categorised into five principal domains, addressing (1) AI-based data analytics for high-calibre, (2) data-driven energy management, (3) the construction of energy-efficient and comfortable buildings, (4) data-driven strategies for renovation roadmaps and energy-efficient financing, as well as (5) decision-making processes under uncertainty for the creation of efficient and climate-resilient buildings.

The report offers a detailed technical description of the services under development and their role within the built environment, presenting innovative solutions and approaches. The focal point of the document is to relay detailed information on the developmental progress achieved up to the 20th month of the project, including their implementation in the DigiBUILD pilot schemes. Lastly, the document outlines forthcoming steps for the subsequent months of the project, culminating in a key milestone at month 28, where the developmental activities are expected to conclude, leading to the unveiling of the 'Third Wave' of DigiBUILD AI-based Data-driven Services for the Built Environment.

1 Introduction

1.1 Purpose and audience of the document

Deliverable 3.2 is an essential update on the progress of AI-based and data-driven services within the DigiBUILD project from month 12 to month 20. It extends the foundation laid in the initial report, delving into the advanced stages of integrating these innovative services into the built environment. The report explores the evolved architecture of the AI-based services, highlighting their enhanced data integration with the DigiBUILD Data Lake and the improved mechanisms for data exchange and service integration. A significant focus is placed on the methodological progress, detailing the novelty of the services, the development progress, the practical application of these methodologies in the DigiBUILD pilots, and the future trajectory.

1.2 Relation to other activities

The report on AI-based data-driven services within the DigiBUILD project exhibits significant interdependencies with various other work packages (WPs), highlighting the integrated and collaborative nature of the project. These interrelations are crucial for the cohesive development and implementation of the services and are detailed as follows:

- **WP1 - DigiBUILD user’s stories, data requirements & specifications through co-creation with stakeholders:** The AI-based and data-driven services are intrinsically linked to Task 1.3, which focuses on the "Co-design of the DigiBUILD use-cases and smart energy services for high-quality building stock performance." This task is foundational in defining the user requirements and use cases that the AI services aim at addressing and solving. The development of these services is directly influenced by the needs and scenarios outlined in this task, ensuring that they are user-centric and tailored to real-world applications. Furthermore, an additional critical aspect of the AI service development is their alignment with the principles and guidelines outlined in D1.4 - "DigiBUILD Data Protection and Cyber Security." This document is integral to ensuring that all AI-based services adhere to the highest standards of data protection and cyber security, a paramount concern in today's digital landscape. Additionally, the design and architecture of the AI services play a crucial role in Task 1.5, which deals with the "Architecture definition from existing building data reference platforms." Deliverable D1.5 provides a comprehensive description of how these services are integrated into the overall DigiBUILD architecture. This integration is vital as it aligns the AI services with existing data platforms, enhancing their capabilities and leading to a more robust and efficient building data management system.
- **WP2 - DigiBUILD Data Interoperability & High-Quality Data:** A critical aspect of the AI services is the acquisition of data, which is primarily sourced from the project's Data Lake. This aspect aligns with WP2, focusing on "DigiBUILD Data Interoperability & High-Quality Data." The synergy between WP2 and the AI services is essential, as the quality and interoperability of data directly impact the effectiveness and accuracy of the services. Furthermore, the procurement of data from the DigiBUILD Data Lake is accomplished utilising the data sharing tool that was developed in the context of Task 2.6 and based on the DigiBUILD ontologies and data model which is serving as a federation and data-integration-enabling component.
- **WP4 - Digital Twins and Cloud-Based DigiBUILD Data Toolbox**: The services have a strong correlation with WP4, which encompasses "Digital Twins and cloud-based DigiBUILD data toolbox."

Funded by the European Union

Most of the developed services are intended to be utilised by the digital twins created for each pilot within the project. This integration allows for a more comprehensive use of the AI services, enhancing the capabilities of the digital twins and providing more nuanced insights into building performance.

- **WP5 - Demonstration of DigiBUILD in buildings across Europe:** The primary objective of the services developed, as detailed in this document, is their application within the respective Pilot Clusters outlined in Work Package 5 (WP5). This endeavour aims at bolstering the arsenal of the pilots, facilitating the attainment of not only their energy targets but also of other specified goals. This is to be achieved through the utilisation of smart algorithms, designed to assist in meeting the objectives established by the Key Performance Indicators (KPIs) set forth in WP5.

In summary, the AI-based data-driven services described in this deliverable are designed to utilise data from the project's pilots (WP2 relation) to develop methodologies and algorithms that can more effectively exploit building systems through Digital Twins (WP4 relation). This is done to meet and exceed the user requirements identified in WP1, demonstrating a holistic and integrated approach to enhancing building performance within the DigiBUILD framework.

1.3 Structure of the document

The structure of the document detailing AI-based data-driven services in the DigiBUILD project is designed to provide a comprehensive and systematic exploration of various services and their applications. It is organised into several key sections, each focusing on a distinct aspect of the AI-based services and their implementation.

Section 2 delves into the final version of the architecture of AI-based data-driven services, providing insights into the foundational framework and technical underpinnings of these services.

Section 3 is dedicated to AI-based services for finer-grained energy profiling and forecasting. This section is further subdivided into four subsections, each detailing a specific service, namely: (1) Performance monitoring and benchmarking, (2) Energy performance prediction, (3) District network production forecasting, and (4) Energy forecasting. Each of these subsections includes a detailed description of the service, its novelty, current development progress, application on DigiBUILD pilots, and next steps for further development.

Section 4 explores data-driven services for energy resource management. This section encompasses several services, including (1) Pro-active maintenance and facility management, (2) District network production economic optimization, (3) Power recharging management, (4) Energy vs e-mobility package, (5) Carbon-free buildings, and (6) Optimal electric or thermal load management. Similarly, to Section 3, each service in this section is elaborated upon in terms of its description, novelty, development status, application in DigiBUILD pilots, and future directions.

Section 5 focuses on data-driven energy and non-energy services for enhanced comfort and people well-being. It consists of two services: Enhanced comfort and well-being and Comfort Performance Contract. Each service is described in detail, followed by its novelty, development progress, application in DigiBUILD pilots, and upcoming steps.

Section 6 addresses data-driven services for renovation roadmaps and energy efficiency financing. It includes services related to Energy efficiency financing and policymaking and One-stop-shop energy efficiency hub, with each service section following the same detailed structure as previous sections.

Section 7 presents a decision-making under uncertainty tool for efficient and climate-resilient buildings. This section provides an in-depth look at the service aimed at enhancing building efficiency and resilience in the face of climate change.

Finally, **Section 8** concludes the document and outlines the next steps. This section summarises the findings and progress detailed in the previous sections and provides a forward-looking perspective on the future developments and applications of these AI-based services in the DigiBUILD project.

Throughout the document, each section meticulously details the service development and application, ensuring a thorough understanding of their impact and potential within the DigiBUILD framework.

2 Architecture of AI-based data-driven services (Final Version)

One of the fundamental objectives of DigiBUILD is the development of data-driven services, aimed at empowering the built environment and its stakeholders. These services are designed to facilitate informed decision-making and efficient management of energy and non-energy assets. The delivery of service outputs to users is accomplished through two primary means: Digital Twins and specialized graphical user interfaces (GUIs) that are custom-developed to meet the specific needs of the services and their users. For effective and consistent communication between users and services, or through the Digital Twins, it is imperative to establish a uniform approach to development. This requirement is fulfilled by the AI service architecture delineated in this chapter, which outlines the communication protocols both with the individual technical work packages of the project and internally, to ensure the correct flow of information.

In response to user and technical requirements identified in the WP1 work packages, the WP3 architecture has been segmented into four principal entities/levels, each defined by its functionality and purpose. These are:

- **Artifacts Layer:** This layer offers developers a suite of specialized tools for the appropriate training, storage, and reuse of models developed within the WP3 workflows.
- **Intelligence Layer:** Constituting the core of the data-driven services, this layer encompasses the development of model building algorithms, optimization algorithms, and the necessary algorithms and applications for WP3's simulation services and MCDA (Multi-Criteria Decision Analysis) services.
- **Interface Layer:** Here, the GUIs necessary for the direct pipelining of services and information to the user are developed.
- **Serving Layer:** This layer involves the creation of appropriate APIs to facilitate the utilization of algorithms and models by the Digital Twins.

It is also pertinent to highlight the Business Intelligence Layer and the Digital Twin Suite, which serve as interfaces to WP2 and WP4, respectively. The **Business Intelligence Layer** is instrumental in collating the data necessary for the services from the DigiBUILD Data Warehouse, which is then employed by the data-driven algorithms. Conversely, the **Digital Twin Suite** functions as the interface to WP4 and the corresponding Digital Twins of the pilots, for which intelligent services are developed.

Funded by the European Union

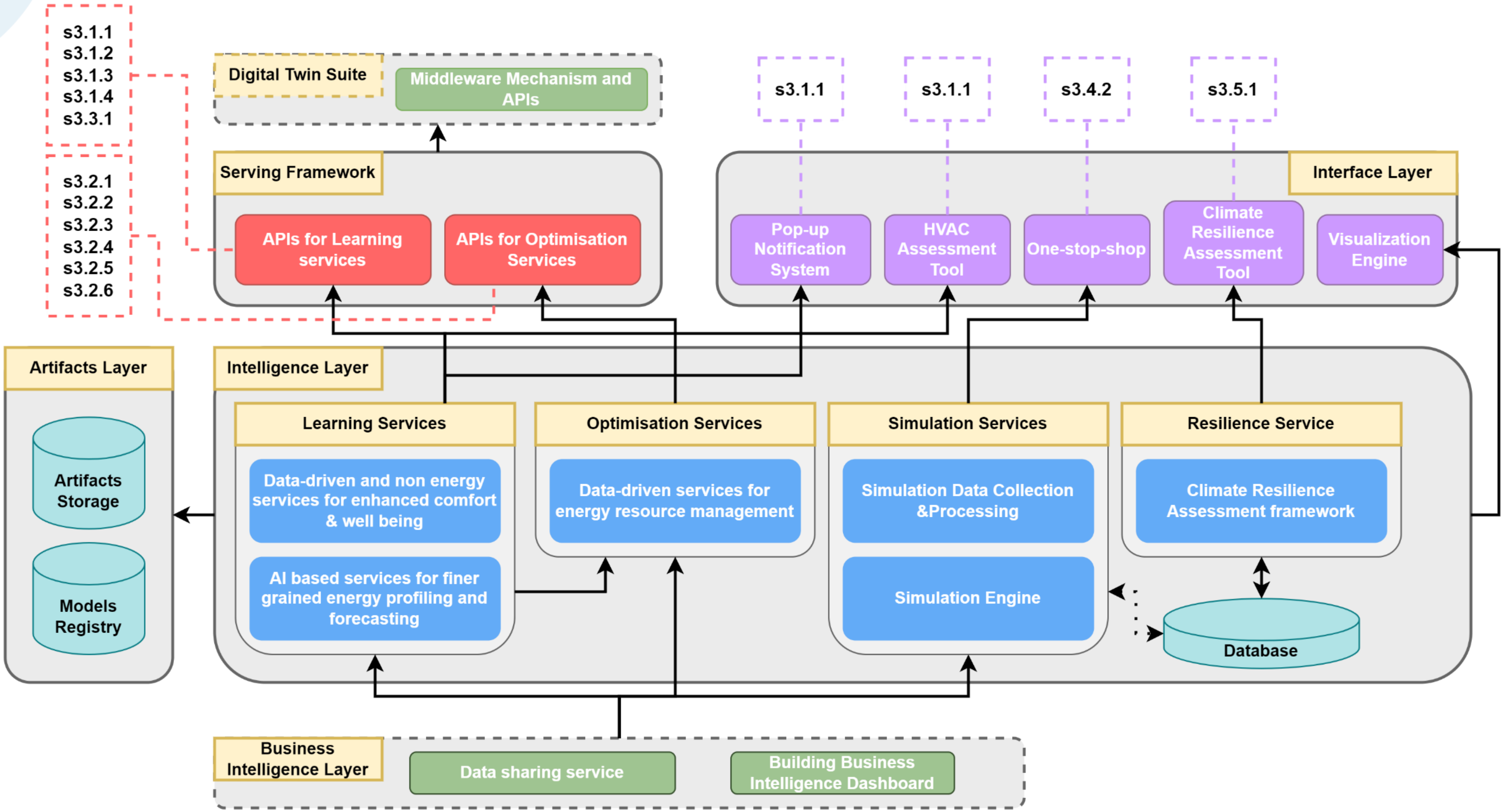


Figure 1: WP3 final architecture

Figure 1 delineates the comprehensive and final architecture of WP3, with which service developers are required to align. This ensures the delivery of homogenous solutions, facilitating their future utility and exploitation. Each layer within this architecture incorporates specific components, aligned with their respective tasks and functionalities:

- **Intelligence Layer:**
  - Learning Services: These encompass trained models developed under T3.1 - "AI-based services for more detailed energy profiling and forecasting" and T3.3 - "Data-driven energy and non-energy services for greater comfort and well-being".
  - Optimization Services: Services and modules emerging from T3.2 - 'Data-driven services for energy resource management'.
  - Simulation Services: Services formulated under T3.4 - 'Data-based services for renovation roadmaps and energy efficiency financing'.
  - MCDA Services: Multi-criteria decision support services for climate resilient buildings, resulting from T3.5 -"Decision-making tools under uncertainty for efficient and climate-resilient buildings".
- **Interface Layer:**
  - Pop-up Notification System: This is an additional graphical user interface (GUI) designed to enhance data-driven energy and non-energy services, aimed at improving comfort and well-being by collecting thermal comfort data from users.
  - HVAC Assessment Tool: Provides insights to end users, particularly for buildings not specified for Digital Twin representation, notably in Pilot 8 (IEECP) and Pilot 9 (NTUA).
  - Climate Resilient Buildings Assessment Tool: A frontend application offering MCDA services to end-users.
  - One-Stop Service User Interface: A front-end application supporting the one-stop service.
  - Visualization Engine: A front-end application utilized by DigiBUILD end-users for visualizing data stored in the DigiBUILD Data Lake.
- **Serving Layer:** This layer includes appropriate API applications, channelling learning results and optimization services to the respective Digital Twins.

This structured architecture exemplifies a well-organized approach to delivering AI-based and data-driven services in DigiBUILD, ensuring seamless integration and functionality across various components and layers.

3 AI-based services for finer-grained energy profiling and forecasting

Energy usage optimization and performance enhancement in buildings is essential, and Machine Learning can be an efficient and powerful weapon towards that. In this context, the services presented in the following section are specifically designed to address various facets of building energy management, including performance monitoring, benchmarking, energy performance prediction, and district network production forecasting. The aim of these services is to develop models that produce accurate forecasts and contribute to sustainable energy management. By using data from each pilot, each service will demonstrate its capabilities in monitoring and improving energy efficiency, while also offering insights into user behaviour and system optimization. The contribution of this research lies in its innovative application of ML algorithms to predict and manage energy usage in buildings, providing a roadmap for smart and energy-efficient building operations.

3.1 Performance monitoring and benchmarking (s3.1.1)

3.1.1 Description of the Service

The high amount of energy used in buildings has raised concerns regarding whether it aligns with their actual requirements. To enable this assessment, it is necessary to establish benchmarks, so that energy performance can be monitored, and an energy-efficient consumption strategy can be provided.

The main objective of this service is to provide energy consumption predictions and forecasts to monitor and compare the actual energy performance of the building under study with that of similar buildings or areas. The literature on this topic is extensive, especially with the use of deep learning and neural networks (1) (2) (3), and different benchmarking approaches will be adopted in parallel depending on the pilot under study.

For the NTUA and IEECP pilots, s3.1.1 involves monitoring the performance of the installed AC systems and supporting decisions related to the replacement of inefficient equipment with new devices of a higher energy class. Specifically, the service processes and analyses electricity consumption data collected from sensors installed at the AC electricity supply lines and correlates them with the respective indoor and outdoor temperature measurements. To that end, the heating and cooling hours are precisely identified for each room (considering the occupancy of the rooms and the behaviour of their users) and the annual energy used for heating/cooling the building is accurately measured. Based on the above, indicative AC replacement scenarios can be evaluated, i.e. actions where the existing AC systems are being replaced by new ones of higher SCOP (Seasonal Energy Efficiency Ratio – cooling efficiency) and SEER (Seasonal Coefficient of Performance – heating efficiency) values. By assuming a reasonable investment cost per scenario and estimating the energy that could potentially be saved after upgrading the equipment based on current or predicted electricity prices, the payback period of the investment is computed and cost-efficient retrofitting actions (e.g. payback period is less than 3 or 5 years) are identified. Note that, due to its nature, the engine of the service can also be exploited to provide further useful insights about the thermal comfort of the users, benchmarking their behaviour, and evaluating if the AC systems are rationally used. An overview of the service architecture is provided in Figure 2.

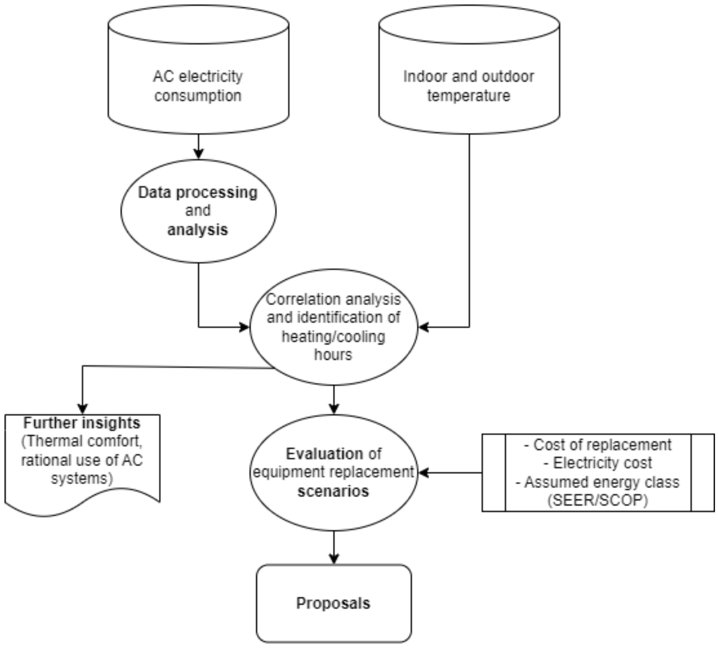


Figure 2: Overview of s.3.1.1 architecture (AC systems monitoring and replacement analysis).

Note that s3.1.1 is currently expanded to also serve the needs of the FVH pilot. To that end, the pilot data have been collected and pre-processed, resulting in a high-quality dataset that represents the total electricity consumed in the pilot building per sector, floor, and room. Data pre-processing and cleaning has taken place to ensure the quality of the data and identify correlations between loads, trends and seasonal patterns. Additionally, suitable time series forecasting methods like SVM, Linear Regression, Exponential Smoothing, Moving Averages, AutoRegressive Integrated methods (ARIMA) and neural networks have been examined with the objective to generate reconciled forecasts for the complete hierarchy of loads, i.e. to ensure both accuracy and coherence. After examining the stated methods and considering the characteristics of the dataset, a neural network, tasked to generate reconciled forecasts for the complete hierarchy, is being developed to allow the inclusion of external variables (e.g. weather) and identifiers (e.g. room, section, floor, and load type), while supporting cross- and transfer-learning. When complete, the service will be capable of effectively supporting hierarchical forecasting tasks, in addition to the individual time series forecasting capabilities provided by s3.1.4, and, thus, extend the novelty and the range of application of the DigiBUILD services.

The neural network's architecture is designed for time series forecasting, while also integrating additional information about the floor, section and room of each individual time series. The architecture is depicted in Figure 3. The model initiates with an input layer of 168 x 97, since 97 different time series are being used and for each of them a window of 1 week (168 hours) is being utilised to predict the output. The time series' length is 5 months which is then split into testing (3 weeks of data) and the rest for training (75%) and validation (25%). After the data has been scaled, it goes through three hidden layers, each consisting of 252 neurons and employing a ReLU activation function. These layers are used to capture the complex, non-linear characteristics of the time series. The second hidden layer is also accompanied by a dropout layer to prevent overfitting during training. In parallel to this, there is an additional feature layer with two neutrons that represent the floor and the section of each time series. These neurons then pass through an embedding layer, which is flattened to make the data suitable for processing in dense layers. These two parallel outputs are then concatenated to achieve the fusion of the categorical features. Finally, after another dense layer, the predicted energy consumption of 24 hours for each window is computed. To further improve the prediction accuracy and take advantage of the hierarchical structure of the time series, the predicted output is then fed into a reconciler.

Funded by the European Union

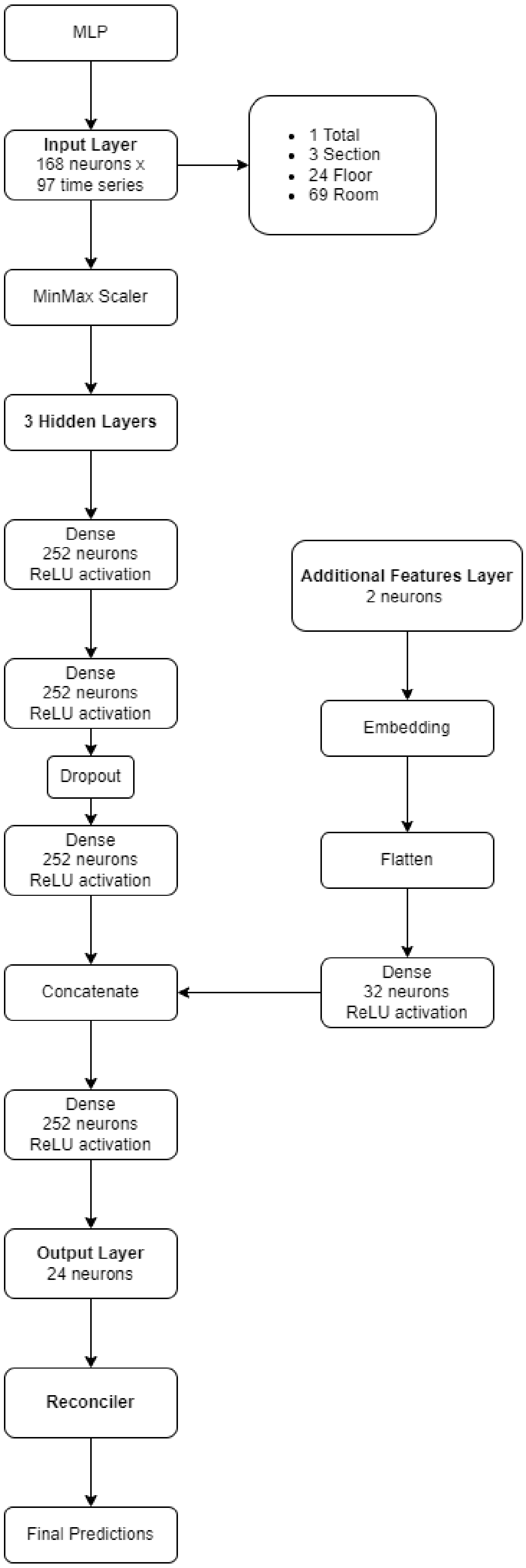


Figure 3: Neural network architecture for FVH pilot.

The performance monitoring and benchmarking service will provide support to reduce the consumption of buildings following a data-driven approach and using ML-based algorithms. Based on the available information and the connection with other identified WP3/4 services, which differ depending on the pilot (s3.2.1 Pro-active maintenance/facility management, s4.1.1 Digital Twin for designing future buildings and s4.3.1 District Digital Twin infrastructure), a tailored approach for the specific pilots will follow, considering different prediction horizons. In terms of architecture, the following aspects are mentioned:

- The tasks of model creation and training will be covered by the back-end environment. Models will be re-trained according to specific needs, stored in a repository and applied to obtain new predictions. The models will be accessed via REST APIs.
- The new calculations will be offered to users in a transparent way, whereby the user will only select the specific parameters to be used for the predictions, and explicit aspects if any (prediction horizon, preferred model). The models will be re-trained based on historical data from the corresponding Data Marts in the final implementation, whereas in this phase they will be based on the historical data samples.
- Performance metrics connected to the available models and other specific calculations connected to the established benchmarks will also be accessible via an endpoint.

The outputs of the models will be used in the aforementioned services with the specifications provided via an API, and the final results will also be integrated in the corresponding versions of the Digital Twins (front environment). The service is developed using Python, taking advantage of several libraries for the development of ML algorithms (NumPy, Pandas, SciPy, Scikit-learn) and for the visualisation of data during the exploration phases The building blocks of the architecture can be seen in Figure 4.

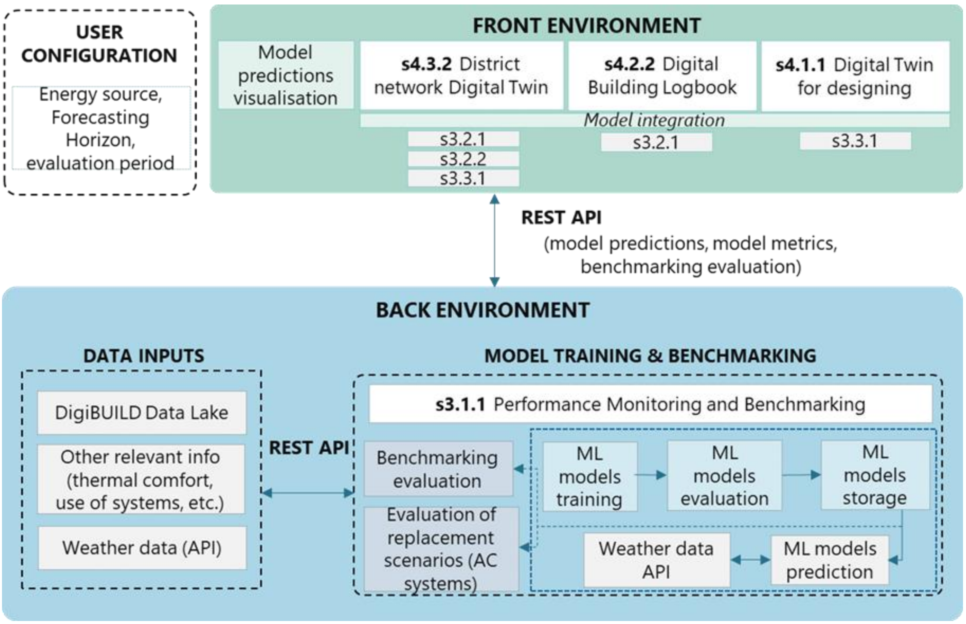


Figure 4: Overall architecture of s3.1.1 Performance monitoring and benchmarking

3.1.2 Novelty

This service is applied within a wide set of pilots, and the approach taken in each of them varies, firstly due to the characteristics under analysis, and secondly due to the main expectations of each pilot. In general terms, the aim is to model energy-related parameters at different scales, and to compare these predictions with real data from the building under study and other similar buildings, to draw conclusions. The way in which this assessment is carried out will differ between pilots, as has been done for the NTUA and IEECP pilots, which are the most advanced in terms of development. In these instances, the possible replacement of existing systems with more efficient ones is assessed, to mitigate the risks associated with energy efficiency investments.

Although the literature involves many methodological approaches for evaluating the impact of renovation and retrofitting actions in buildings, most of them are based on special simulation tools or building energy modelling software, such as Energy Plus, OpenStudio, eQuest and DesignBuilder (4) (5). These approaches can provide accurate results, but they depend on detailed input data, including information about the structure of the building, the materials used for its construction and insulation, as well as the installed mechanical and electrical equipment, among others. In this regard, evaluating the potential of said actions can become particularly time-consuming or even impossible for cases where the required data cannot be collected accurately. On the other hand, approaches that estimate energy savings based on survey and literature review data may provide insights that cannot be safely generalised for all buildings, often overlooking building particularities such as building occupancy and user behaviour. In this regard, the novelty of the developed service lies on its wide applicability, ease-of-use, and data-driven approach. Given data solely collected from smart meters and sensors, the service can accurately estimate the utilisation of AC systems at heating and cooling periods, approximate the efficiency of the infrastructure for each room separately, and compute potential energy savings. Moreover, it can serve as a basis for further analysis and monitoring focusing on the rational use of the systems and the thermal comfort of the building users. In addition, the service can be used to validate the results of building energy modelling software, providing a more data-driven assessment over their results.

As far as the hierarchical forecasting model developed for the FVH pilot is concerned, the novelty of the service lies on the construction of a global forecasting approach that can accurately predict multiple series which represent different types of loads at different spaces of the building, while ensuring coherent results. Conventional forecasting approaches involve the prediction of each load individually. Since said forecasts are not necessarily coherent, coherency is typically ensured through a 2-step process that employs some hierarchical reconciliation method, such as the bottom-up, top-down, or middle-out methods. On the contrary, s3.1.1 opts to achieve coherency through a 1-step process where hierarchical information (e.g. type of load, room, floor) is shared within the model to enhance its performance and simplify the overall forecasting process.

3.1.3 Development Progress

Some preliminary results of the exploratory data analysis and first modelling activities for the NTUA pilot were presented in D3.1. In general terms, work during the current phase has focused on data exploration and correlation of datasets, definition of the approach and modelling.

Activities during the second release for the UCL pilot have focused on conducting a thorough analysis of available information, gathering requirements from other services and initiating preliminary modelling activities. The information considered in this phase corresponds to a subset of zones of the One Pool Street building.

Due to an internal communication limitation in the IT infrastructure in one of the buildings, the monitoring infrastructure remains non-operational, even though all sensors are installed. In another building, there are still some minor modifications pending for the deployed infrastructure that will provide real-time data. This lack of data has delayed the implementation of the service which is scheduled for the next release.

The FOCCHI pilot activities have been developed in parallel, in terms of data exploration and first model development for PV production. The results have been validated using the most common metrics, and the models are available via API. The lack of detailed information for the other datasets, such as building consumption at the expected granularity, has delayed other activities, which will be addressed in the next version of the technology, as the approach to be taken is still in the definition phase.

In order to put s3.1.1 in production for the NTUA and IEECP pilot cases, a dedicated API has been developed. It retrieves, processes, and analyses the AC consumption data, by also identifying heating/cooling hours, and examining the efficiency of the systems based on correlation analysis. Additional APIs have also been developed to execute the evaluation of the equipment replacement scenarios, monitor the thermal comfort of the users, and assess the use of the systems in terms of rationality.

Currently, a GUI has also been developed to allow the users of the service to interact with its results, to adjust its parameters and inputs (details of installed AC systems, electricity cost, efficiency and cost of new AC systems are reported), and to facilitate monitoring through visualisations and detailed reports.

Regarding s3.1.1 for the FVH pilot case, the development progress is still ongoing and is currently focused on fine-tuning a NN architecture that will allow to accurately predict multiple hierarchical time series with a single global model. This includes the expansion of the model so that exogenous variables and identifiers of the building structure are considered when producing the forecasts. At the same time, different modelling approaches are being tested to allow the model to result in reconciled forecasts without deterioration in accuracy. So far, experimental results support the potential of the suggested approach over conventional forecasting approaches.

The source code can be found on GitHub: GitHub s3.1.1

3.1.4 Application on DigiBUILD Pilots

This service will be implemented and validated in the following pilots: UCL (Pilot 1), IASI&SITTA (Pilot 3), FOCCHI (Pilot 5B), FVH (Pilot 7), IEECP (Pilot 8) and NTUA (Pilot 9). Different levels of development run in parallel due to the maturity levels in terms of equipment and data access, and more specific information per pilot is provided below:

- **Pilot 1 (UCL):** the monitoring information available at different scales (building, area and space level) for the pilot buildings will be considered. In this phase, activities have started for the research building with student residential space in Pool Street West (electricity, cooling, and heating consumption). The results of the service will be fed into services s3.2.1 "Pro-active maintenance/ facility management" and s3.2.2 "Enhanced comfort and well-being". In addition, the service will be integrated into the District Digital Twin (s4.3.1). A set of benchmarks will be defined to provide support to reduce energy consumption. UC_3 contains specific details of the implementation of s3.1.1 in the UCL pilot (the reader may consult deliverable D1.2).
- **Pilot 3 (IASI&SITTA):** the aim is to apply the service in administrative buildings where new sensors are being installed. Previously, only static data from invoices were considered to assess building performance. The corresponding application of s3.1.1 defined in D1.2 for IASI&SITTA can be found in UC_11, and the service will be exploited in the Digital Building Logbook.
- **Pilot 5B (FOCCHI):** the data from the PV production, the heating and cooling system and Indoor Environmental Quality will be exploited to balance energy consumption with RES production and reduce energy production and CO2 emissions. At this stage, the activities have exploited the PV production data, while other data sources are still being explored. The service will work in conjunction with the service s3.3.1 "Enhanced comfort and well-being" and the results will be integrated in the pilots' Digital Twin (s4.1.1 "designing future buildings"). The corresponding use case defined in D1.2 for FOCCHI pilot is UC_11.

- **Pilot 7 (FVH):** data from an office building in Helsinki classified per sector, floor and room will be used within s3.1.1. The objective is to predict office usage from available energy consumption data and weather information. UC_3 (reported in D1.2 and D1.6) contains specific details of the implementation of s3.1.1 in FVH.
- **Pilot 8 (IEECP):** the service will be implemented in monitored office buildings in Netherlands, with the main goal of predicting energy consumption of these offices one day ahead. Details of the corresponding use case can be found in UC_3 for s3.1.1 application in IEECP.
- **Pilot 9 (NTUA):** the service will be implemented in an area of an office building which has a BMS. The focus is mainly put on the analysis of the electricity consumption collected by the sensors installed in the AC systems. By proposing AC replacement scenarios and estimating the potential energy that could be saved, a set of benchmarks are established. UC_·3 (D2.1) contains the details of the implementation of s3.1.1 for the NTUA pilot.

Some details of the application of this service during this phase in the different pilots is included below.

UCL pilot

Data from the Pool Street West building have started to be analysed for a subset of spaces. Initially, the aim is to establish groups of rooms according to their nature, to understand the similarities in terms of information to be exploited and, at a later stage, to define a set of benchmarks to monitor performance.

Initial exploration activities of the provided data and preliminary conclusions are presented below. The studied variables are related to comfort, namely temperature (in ºC), occupancy (0-1 values) and CO2 concentration (PPMs), and comprehend several rooms across four of the nine floors that the data sample is composed by. This study has served as an initial and general procedure to classify the different spaces and rooms which will be used later to develop and validate models.

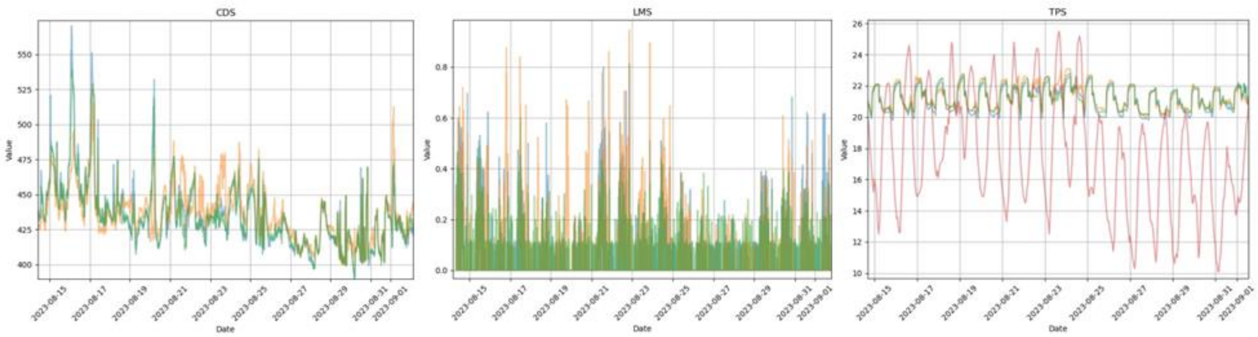


Figure 5: Preview of the studied variables in the UCL pilot (CO2, occupancy and temperature).

After the aforementioned study, several conclusions can be extracted: One of the most relevant issues with the provided data is that data are missing for every variable and sensor over a time period. Specifically, the observed data gaps take place on the 1st of August and from the 9th to the 14th of August. An example of this data gaps is shown below (values in PPMs).

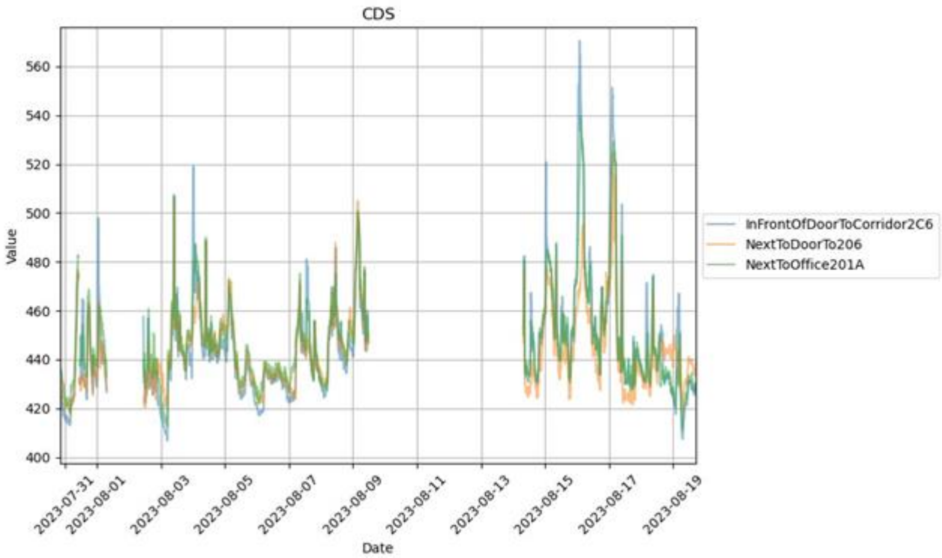


Figure 6: Data gaps in the UCL sample.

Finally, a first and general classification of rooms has been made according to their behaviour and potential use. This enables us to identify which rooms have installed and used AC and which have not over the period of time for which data is available. This classification has been developed analysing daily patterns and comparing the behaviour of the indoor temperature to the outdoor one acquired from an external and trustworthy historical weather source (values in ºC). The following figures showcase in a graphical way the difference of the patterns encountered in this analysis. At a later stage, when the energy data is available, it will be studied together with the comfort parameters already considered, to extract further insights and start modelling activities.

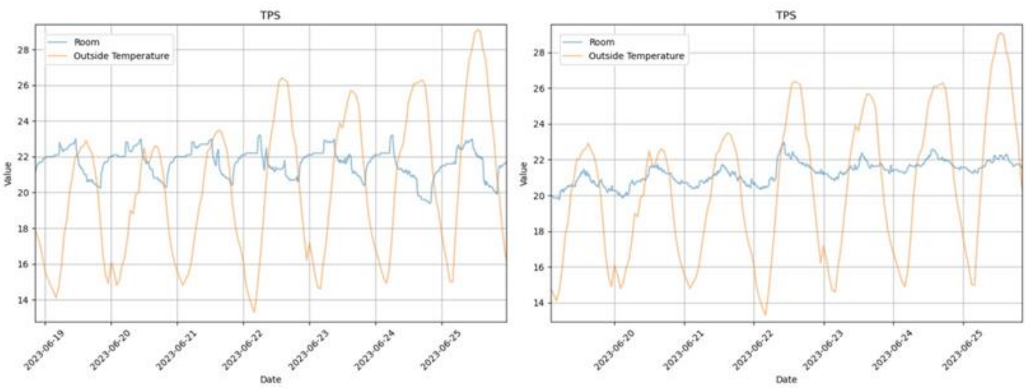


Figure 7: Room temperatures with and without AC in UCL pilot.

FOCCHI pilot

During this phase, PV production data has been explored and the first model development activities have been carried out. The most common metrics have been used for validation, and the initial version includes Random Forest regression and Decision Tree regression. The results show high accuracy, as seen in Table 1. This performance information and access to predictions is available through an API. The information on BMS (HVAC system and gas boilers) and IEQ is still being analysed, as the data is of different granularity and a final approach needs to be consolidated. In addition, synergies with comfort services are being studied to select the most appropriate information. During the final release, model training activities for this subset of information will be finalised and exposed through the API as has been done for other features. The following figures show the results obtained for the PV production.

Funded by
the European Union

Table 1: Forecasting model results for PV production in FOCCHI.

Model	R2	MAPE	MAE	RMSE	MSE
Decision tree	0.87	135.32	2.88	6.09	37.18
Random Forest	0.93	96.71	2.02	4.27	18.27

NTUA pilot

Up to this point, s3.1.1 for AC monitoring and replacement analysis has been successfully applied on the NTUA pilot. Having analysed the data collected from the smart meters and sensors installed in the building, which cover a period of approximately 1.5 years, the service identified potential (payback period of investment is equal or less than 5 years) in the replacement of three AC systems in rooms 25, 30B and 30D of the pilot, as shown in Table 2.

Table 2: Results and proposals of s3.1.1 for the NTUA Pilot

Room	Heating Hours	Cooling Hours	AC Capacity (BTU)	Current Consumption (Wh)	Expected Consumption (Wh)	Energy savings (annual)	Cost savings (annual)	Cost of Investment (€)	Payback Period (years)
24	467	275	9,000	476	327	149	37	350	9
25	509	434	9,000	697	398	299	75	350	5
26	553	609	9,000	513	475	38	9	350	37
27	466	868	9,000	504	510	-	-	350	-
28	301	169	9,000	231	208	23	6	350	61
29	141	284	9,000	125	161	-	-	350	-
30A	203	255	24,000	749	491	258	65	450	7
30B	1,066	591	9,000	1,142	735	408	102	350	3
30D	795	745	9,000	1,188	642	546	136	350	3

As seen, based on the processed data, the service has identified the actual heating and cooling hours of the AC systems per room, measured the electricity currently consumed for heating/cooling purposes, and computed the payback period of the replacement investment assuming the installation of an AC system of the same capacity but better energy efficiency (SEER = 8.5 and SCOP = 5.1 for a modern A+++ AC system). For the cost-benefit analysis, an installation cost of around 400€ was assumed and an electricity price of 0.25€ per kWh (after taxes). It is evident that equipment replacement is mainly encouraged in rooms of relatively high occupancy where AC systems are used more frequently (i.e. for longer heating/cooling hours).

Figure 5 provides further insights on the efficiency of each AC system, depicting the hourly electricity consumed on average for heating/cooling the rooms based on the outdoor temperature. Green lines represent the average energy consumed per hour, while the red and blue lines the average energy consumed during the heating and cooling period. It is evident that some systems are more efficient during the heating or cooling season, a factor that affects the decision of the equipment replacement based on its utilisation per case. It can also be observed that the three systems proposed for replacement, apart from being utilised more heavily, are also less energy efficient, reporting an average consumption of about 0.75kWh/h compared to the rest of the systems (0.5kWh/h or less). These results further justify the proposals of the service.

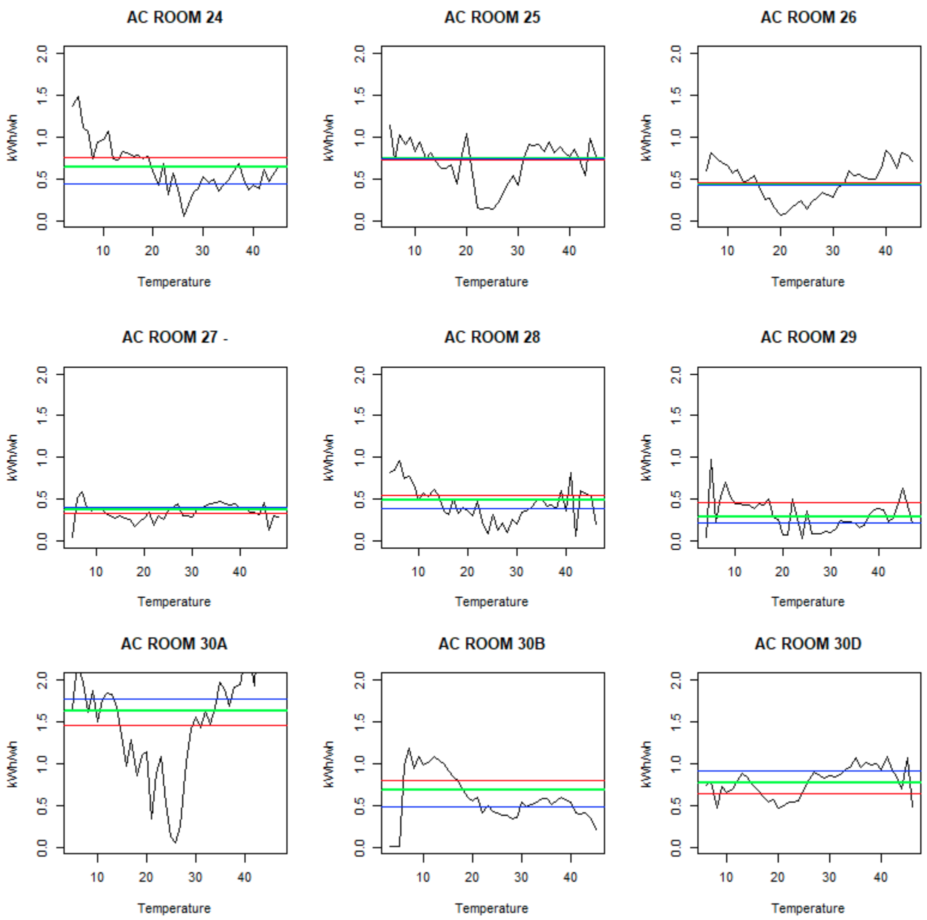


Figure 5: Electricity consumed (average) by the AC systems in the NTUA pilot based on the outdoor temperature.

Finally, as previously noted, the service can provide useful insights regarding the thermal comfort of the users in each room. Figure 6 visualises such results for three indicative rooms. In most cases PMV is within the acceptable range of [-0.5,0.5], meaning that the users of the rooms use the AC systems rationally. Specifically, the analysis indicated that in 80% of the occupancy hours the thermal sensation was acceptable, while in the remaining 20% of the cases the users were mostly cold. Given that these cases can be identified also during the cooling season, we can conclude that the users may sometimes cool their spaces more than required. As a result, behavioural changes on this aspect could result in further energy savings.

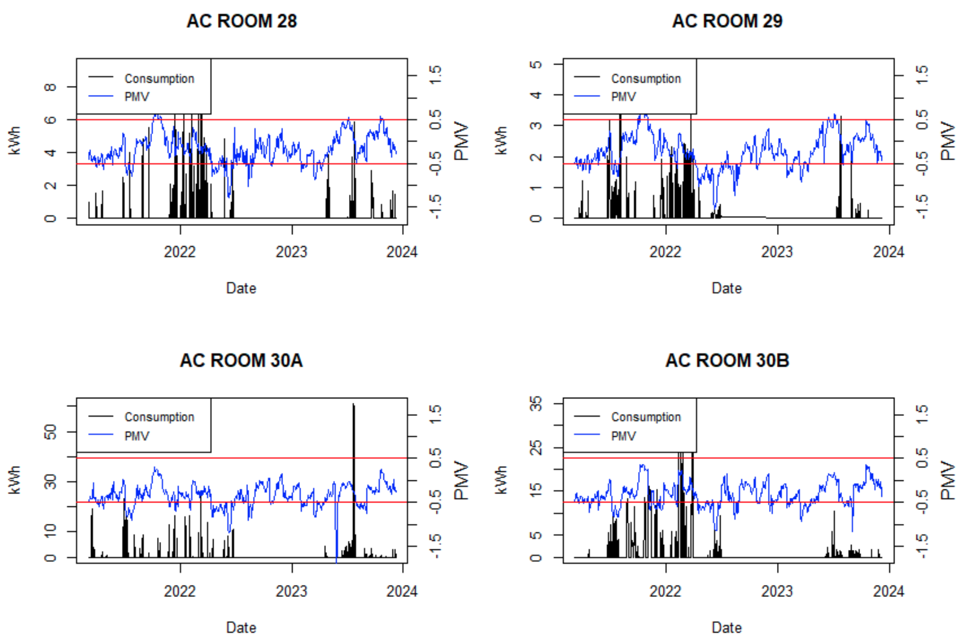


Figure 6: Thermal comfort in three rooms of the NTUA pilot.

The following figures provide visual insights from the analyses conducted using the GUI developed to support the NTUA and IEECP pilots, which do not have a Digital Twin within the DigiBUILD framework.

- Figure 7: This figure displays the result of an analysis focusing on the air conditioning (AC) operation, based on historical data from a specific room in the pilot study. It offers a detailed view of the AC's usage patterns and efficiency.
- Figure 8: Here, we see an economic analysis related to the replacement of the AC unit. This analysis is grounded in the historical operational data of the AC, providing a cost-benefit perspective on the potential replacement.
- Figure 9: This figure illustrates the evaluation of the AC unit replacement in terms of estimated indoor conditions.
- Figure 10: By interacting with the GUI and selecting the 'simulate' button, users can visualise how the new AC unit would perform in maintaining the Predicted Mean Vote (PMV) index, a measure of thermal comfort. The simulation is based on the specific usage profile of the AC unit in the room under analysis.

Funded by the European Union

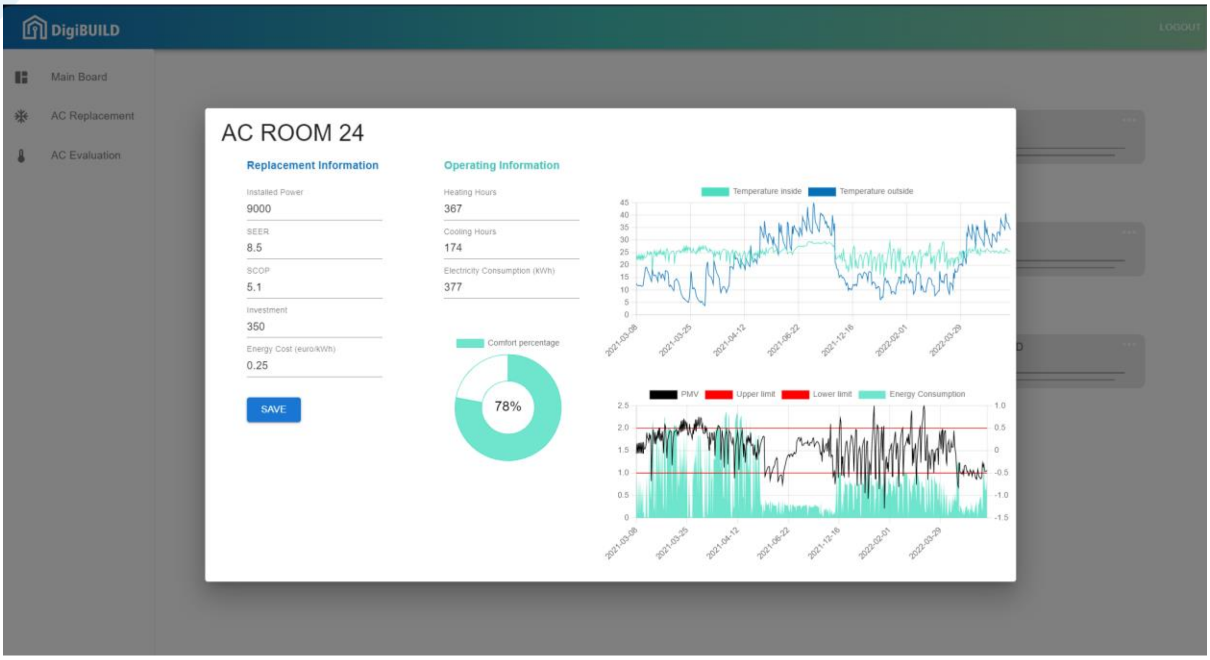


Figure 7: AC usage analysis and replacement information.

Name	Installed Power	Operating Hours Heating	Operating Hours Cooling	Energy Consumption Before	Energy Consumption After	Energy Consumption Savings	Savings (€) Annual	Cost (€) Investment	Payback Period (years)
AC ROOM 24	9000	367	174	377	243	134	33	350	10
AC ROOM 25	9000	487	290	616	341	274	68	350	5
AC ROOM 26	9000	389	280	379	288	91	22	350	15
AC ROOM 27 -	9000	427	546	337	390	0	0	350	Worse than before
AC ROOM 28	9000	287	67	190	169	21	5	350	65
AC ROOM 29	9000	140	192	108	131	0	0	350	Worse than before
AC ROOM 30A	9000	192	112	478	134	344	86	350	4
AC ROOM 30B	9000	1050	395	1086	665	421	105	350	3
AC ROOM 30D	9000	934	432	829	617	212	53	350	6

Figure 8: AC replacement economic analysis per AC unit.

AC Evaluation Results

Name	Installed Power	PMV Mean	PMV STD	Comfortable Hours Real	Comfortable Hours Optimal	Energy Savings (kWh)	Cost Reduction (€) Annual	Payback Period (years)	
AC ROOM 24	9000	-0.17	0.27	77.94	91.77	41	10	10	SIMULATE
AC ROOM 25	9000	-0.23	0.22	82.23	96.54	9	2	5	SIMULATE
AC ROOM 26	9000	-0.23	0.22	82.23	95.24	31	7	15	SIMULATE
AC ROOM 27 -	9000	-0.22	0.22	82.23	96.75	39	9	Inf	SIMULATE
AC ROOM 28	9000	-0.18	0.21	81.89	98.48	4	1	65	SIMULATE
AC ROOM 29	9000	-0.15	0.24	72.22	95.89	11	2	Inf	SIMULATE
AC ROOM 30A	9000	-0.29	0.19	76.27	99.35	50	12	4	SIMULATE
AC ROOM 30B	9000	-0.31	0.21	75.23	86.80	31	7	3	SIMULATE
AC ROOM 30D	9000	-0.31	0.21	75.23	87.66	74	18	6	SIMULATE

Figure 9: AC performance evaluation.

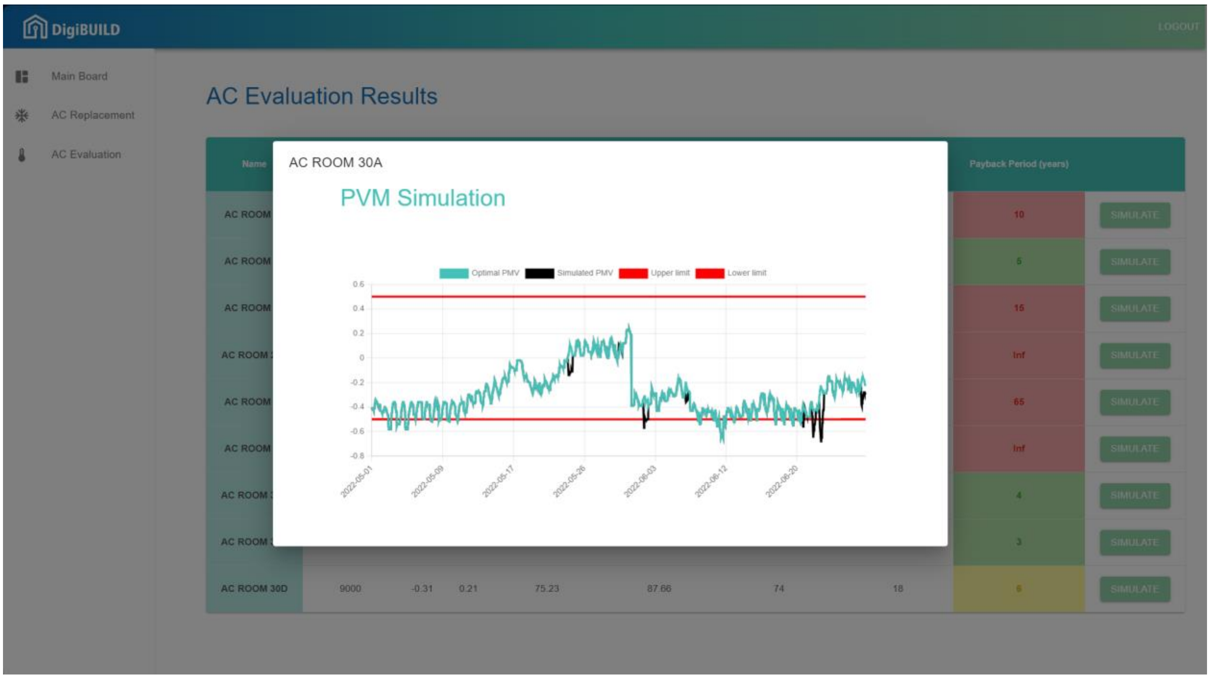


Figure 10: Simulation example of AC replacement and PMV affection.

These figures and the underlying analyses exemplify the GUI's capability to not only assess current operational efficiencies but also to forecast the impact of potential changes in equipment, both in economic terms and in improving indoor environmental quality.

FVH pilot

Simple reconcilers such as Bottom Up, Top Down and MinTrace have been tested and a custom one is currently being developed. The current average mean squared error percentage for all the time series is 2.4%, while the average mean absolute error is 17.3% and with the use of the custom reconciler it is expected to further improve. Figure 11 illustrates some indicative results, particularly focusing on the distribution of error observed when applying the methodology to the FVH pilot. This visual representation provides a crucial insight into the accuracy and reliability of the methodology in a real-world application scenario.

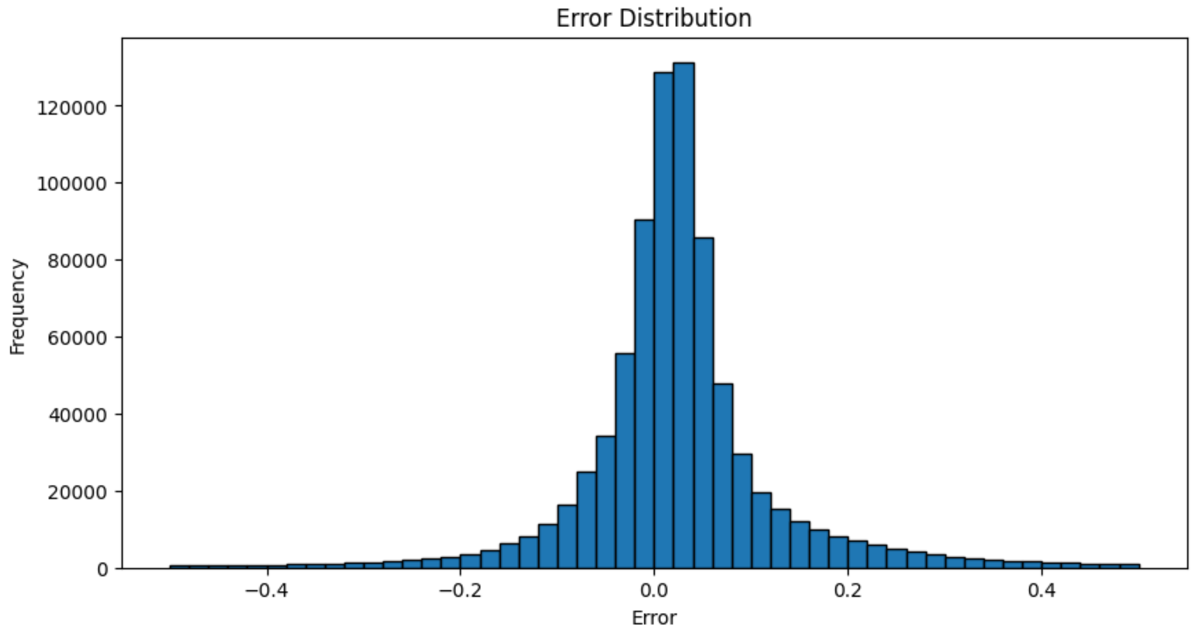


Figure 11: Error distribution of MAE after applying methodology to FVH pilot.

In the following figure, Figure 12, two indicative examples of the day-ahead predictions for two of the rooms can be visualised to showcase the model's performance.

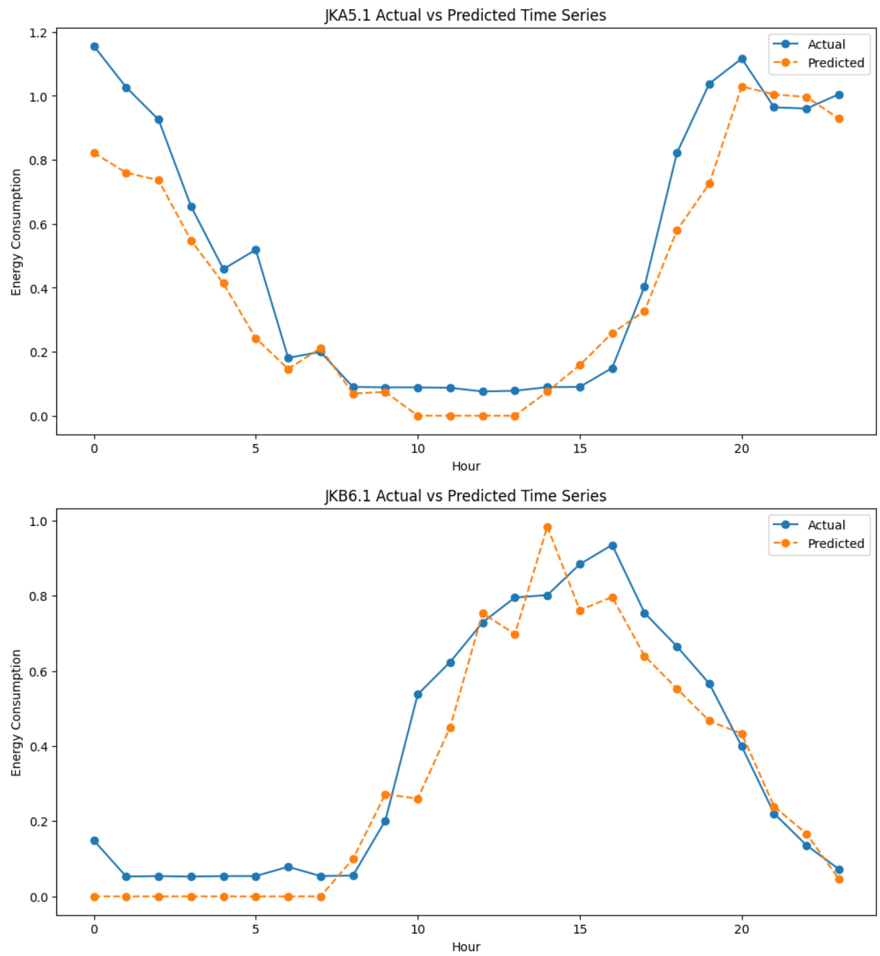


Figure 12: Indicative one day-ahead predictions for rooms JKA5.1 and JKB6.1 of FVH pilot.

3.1.5 Next Steps

During the third technology release, the steps foreseen for this service in the UCL pilot include further data analysis, modelling, benchmark definitions and API exposure. Based on available information, further iterations may also be necessary to redefine the model requirements to meet the needs of other services and the expectations of the pilots. In addition, the already trained algorithms could be refined to improve and ensure the accuracy of the models considering larger amounts of data, as the model training activities have been carried out based on approximately 3 months of data. Finally, new functions could be added to the API interface to expose relevant information not identified at this stage.

Similarly, for the FOCCHI pilot, the data exploration, modelling and API development phases will be carried out for the datasets still under study (building level heating related parameters). Finally, the algorithms already available may also be refined when more historical data become available. In the case of IASI&SITTA, it is expected that new data from the monitoring infrastructure can be exploited in the coming period.

For the FVH pilot, there will be ongoing experimenting and fine tuning with the model in order to increase the accuracy. Additionally, the custom reconciler requires further development and testing so that the desired results can be achieved. For the NTUA pilot further testing of the application will take place, and for IEECP the appropriate configurations must be done on the application, so that it will be able to incorporate the pilot data.

3.2 Energy performance prediction (s3.1.2)

3.2.1 Description of the Service

The objective of the service is to provide predictions to monitor the energy performance of a building while preserving the comfort of its occupants. For this purpose, time series of energy consumption and comfort parameters of the building will be exploited by ML algorithms. This service will be implemented in the EDF pilot and it will provide results for a service of the comfort category. The s3.1.2 service will provide results through a graphical user interface as a stand-alone service, although model results will also be accessible through the REST API. In terms of architecture, the following aspects should be highlighted:

- Model creation and training are covered by the back-end environment. Models will be stored in a repository and applied to obtain new calculations. These models will be accessible via REST APIs, and further calculations can be integrated directly into the stand-alone application.
- The results of the models will be provided in a transparent way, requesting details of the desired models to collect predictions and their performance. These models will be re-trained on the most convenient basis (hourly to weekly horizon), and constantly fed by the corresponding Data Marts.
- Performance metrics from trained prediction models and other specific calculations from the building performance calculation will be accessible via and endpoint. For the building performance calculation, data from IEQ sensors and electricity consumption inside the building will be considered.

The model results will be accessed through the API and the final service will offer a visualisation tool (front environment). The service is being developed using the Python programming language for the ML algorithm training and data exploration phases. The building blocks of the foreseen architecture can be seen in Figure 16.

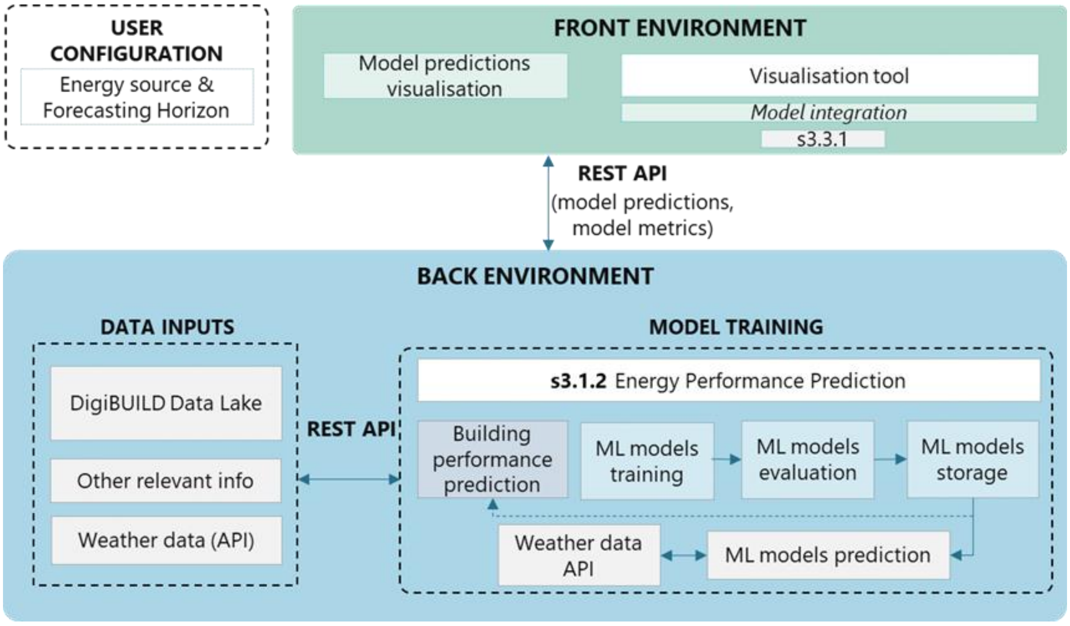


Figure 16: Overall architecture of s3.1.2 Energy performance prediction.

3.2.2 Novelty

Since the final approach to be adopted is still to be refined, the novelty of the service cannot be underlined at this stage. AI-based models will be provided to estimate energy consumption and other considerations will be taken to provide performance prediction, based on literature.

3.2.3 Development Progress

The activities during the second release of this service include the consolidation of the final list of datasets to be used, data exploration and analysis. In addition, the final approach to support building performance monitoring is still being defined. During the final release, AI-based models will be offered to predict the most significant features and as input for comfort services, and finally exposed via REST API to ensure seamless integration.

3.2.4 Application on DigiBUILD Pilots

This service will be implemented and validated in Pilot 2 (EDF). As agreed during this phase, the energy performance prediction service will be used as a stand-alone application and will provide inputs for service s3.3.1 Enhanced comfort and well-being through an endpoint. Integration into the digital twin is not foreseen. The main focus of the pilot is on the exploitation of energy and comfort information, and the key information to be used is related to electricity consumption parameters (i.e. heating, lighting, plugs) and indoor temperatures (IEQ sensors). Therefore, energy performance prediction will be connected to energy consumption and temperature information, as thermal comfort is a priority. Details of the implementation of s3.1.2 in the EDF pilot can be found in UC_28 (D1.2).

During the second release, activities have focused on data exploration and analysis to understand the possibilities offered by the available information to estimate building performance. While the Ellona and Ethera devices (IEQ parameters) cover five different zones, the electrical information comes from three main sources. The correlations between spaces related to the datasets need to be established at an early stage, before modelling activities can start. Furthermore, the final approach has not been consolidated, so an alpha version is not available at this stage, as the purpose of the service is still under analysis.

3.2.5 Next Steps

During the next and final release, the planned steps for the service include further data exploration, modelling, API consolidation and development, and GUI implementation. Although a first identification of datasets to meet the expectations of the service has already been done, the final deliverables and, thus, all dependent activities are still being redefined. The API integration will follow an approach similar to that adopted for other services, while the first mock-up of the interface will be provided with real functionalities and refined so as to be adapted to the final version of the service.

3.3 District network production forecasting (s3.1.3)

3.3.1 Description of the Service

District heating network operators must ensure heat supply under all circumstances and considering economic and technical constraints. Heat production needs to be planned and, for this purpose, it is necessary to know the future demand. The main goal of the District Network production forecasting service is to provide energy predictions to control the energy performance of a district network and help reduce CO_2 emissions by reducing energy consumption.

Heating demand forecasting has been explored extensively in the literature. Some of the most critical challenges are related to the size of the heating network and climatic conditions of socio-economic aspects, among others. The approach adopted in DigiBUILD is based on the exploitation of time series data from the resources available in the network, as well as other relevant characteristics of the buildings under study.

This data-driven approach provides predictions using ML-based algorithms. Specifically, the service is applied at two different locations, namely FASA DHN and Río Vena DHN. Based on the available information and the connection with other WP3/4 services (s3.2.2 District network production economic optimisation and s4.3.2 District network Digital Twin), two different approaches are followed, considering hourly and daily prediction horizons. In terms of architecture, the following aspects are considered:

Model creation and training are covered by the back-end environment. Models can be retrained according to the expected usage, stored in a repository and applied to new calculations. The models can be accessed via REST APIs.

The new calculations are offered to users in a transparent way, so that only the details of the required models, such as the district network under study, the type of parameters to be predicted, the available pre-trained models and the forecasting horizon, in case several options are possible, need to be indicated in the API query. So far, only historical data in the form of files have been used for model training, while at a later stage the models will be fed with the corresponding Data Marts.

Finally, the most relevant metrics of the selected models can also be obtained via the endpoint.

The results of the service will be displayed together with the aforementioned services, although they can also be accessed via the API. Therefore, no visualisation tool is foreseen, while it will be integrated into the front-end environment provided by the corresponding Digital Twin category (s4.3.2). The service is being developed using the Python programming language, and several libraries for the development of ML algorithms (NumPy, Pandas, SciPy, Scikit-learn, etc.) are being considered for the data exploration and model training phases. The high-level updated architecture of the service is shown in Figure 14, while Figure 13 provides some details on the algorithm process.

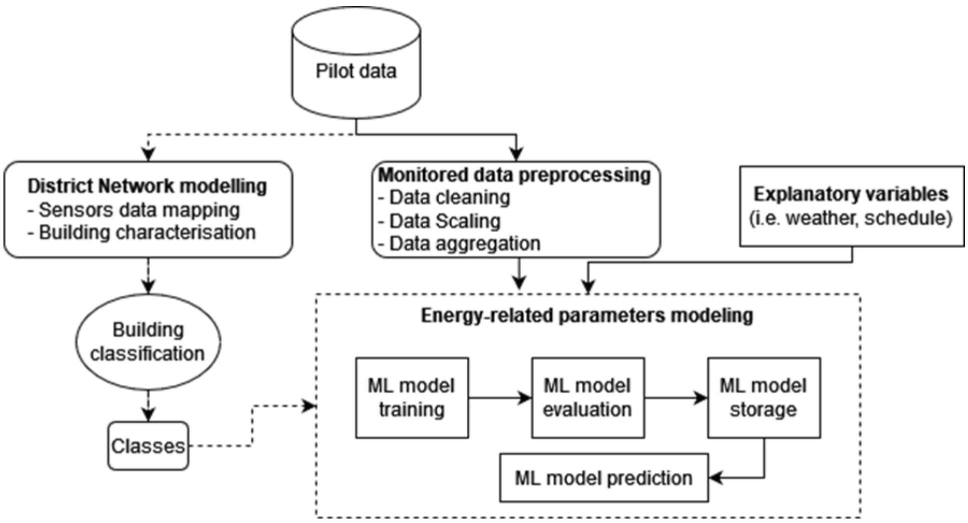


Figure 13: Overview of District Network Production Forecast Service (s3.1.3) approach.

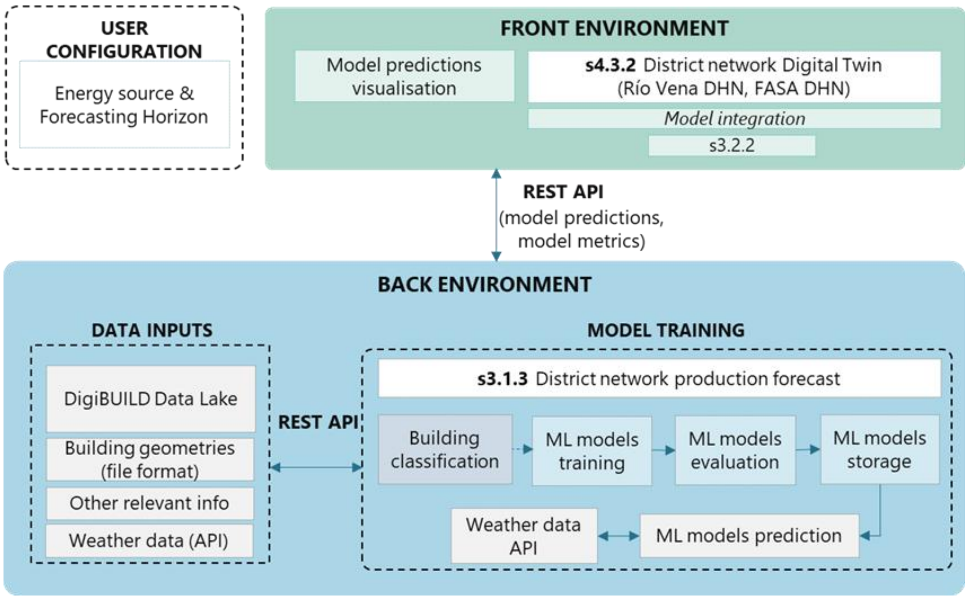


Figure 13: Overall architecture of s3.1.3 District network production forecast.

3.3.2 Novelty

The innovation offered by this service lies within the methodological approach adopted for the two district heating networks under study. Prediction models in this context can be either physical or data-driven. While for the former a detailed insight into the thermal dynamics and energy behaviour of buildings is needed, for the latter little information is needed. Data-driven models can efficiently find relationships, which is the approach selected for this service. In both pilot sites, a regression algorithm is applied to predict the energy demand in buildings belonging to the district heating network. This value can be aggregated at network level and disaggregated at more specific level of detail. In the case of Río Vena, the information from the heating consumption of the 23 buildings in the network is used in an aggregated way in the forecasting models (gas consumption), and the results feed s3.2.2 District network production economic optimisation every hour, according to the specified needs.

On the other hand, the optimisation algorithm based on Particle Swarm Optimisation applied at the second site, (s3.2.2 District network production economic optimisation, details in section 4.2), FASA DHN, requires a higher level of detail, and substation predictions are needed to obtain the total energy that must be provided by the generation systems to cover the buildings demand as well as the heat losses in the distribution system. The integration of clustering with regression has become popular due to its excellent performance in building energy prediction tasks, however, studies on clustering and regression models to achieve optimal performance are lacking (6). Among the advantages of this option, the reduced computational burden can be underlined. In order to reduce complexity and demonstrate its potential, a hybrid method combining building clustering and regression for energy prediction is applied in this context. In this way, a set of representative buildings is selected, and the number of models considered is smaller, thus reducing the complexity in terms of calculation.

3.3.3 Development Progress

During the first release, the effort was devoted to data exploration, to determine and detect patterns and trends within the datasets under study. At this point, some visualisations of the first analyses were obtained and included in D3.1.

During the second release, more comprehensive analyses were carried out and new models were estimated. During this period, data from Río Vena DHN have been explored and pre-processed to determine and detect patterns and trends within the dataset under study. This work has been carried out using Python and through libraries such as matplotlib, or seaborn, some visualisations have been obtained. The models have been validated using the most common metrics, and an API has been created to access information on the performance of the models and the new predicted values.

In parallel, FASA DHN development activities have continued. A two-step methodology has been defined to address the DHN consumption prediction problem, including classification and regression algorithms. Some information related to the geometrical aspects of the buildings is being collected and refined to fit the input format foreseen for the classification phase, while the main features for energy modelling have been identified from the DHN dataset.

3.3.4 Application on DigiBUILD Pilots

The service will be tested and validated in two demonstration sites of Pilot 4 (VEOLIA), corresponding to two differently configured district heating networks. Thus, a different approach has been followed for each location.

- **Pilot 4 - District Heating Network Río Vena:** for this specific application, the heating energy related to gas consumption is predicted, although 15 variables are measured and available, as jointly defined with other services. This result will be used in the optimisation algorithm to reduce resource use.
- **Pilot 4 – FASA DHN:** In this demonstration site, the district heating network has 20 distribution substations, 2 biomass boilers and 1 gas boiler, and the pilot also contains photovoltaic panels. The predictions will include the energy generation of the PV panels and the energy consumption per building. This result will be used in the optimisation algorithm to reduce and balance the use of resources and also reduce CO_2 emissions.

UC_13 (D1.2, D1.6) contains specific details of the implementation of s3.1.3 in the VEOLIA pilot.

Funded by the European Union

Río VENA DHN

During this period, the DHN data have been explored and pre-processed in order to determine and detect patterns and trends within the dataset under study. This work has been carried out using Python and through libraries such as matplotlib, or seaborn, some visualisations have been obtained.

The historical data of this demonstration site includes 14 variables from April 2023 and 15 from August 2023 (previously only 4 variables were available). The most interesting parameter according to the needs of s3.2.2 is stored from April 2023, so the final version of the model may be modified to maintain good accuracy when a full period is available. This model has been validated using the most common metrics, and initial versions include linear regression, Random Forest regression, Decision Tree regression, OLS, Support Vector Machine regression and ARIMA. On average, the results show R2 values of 0.97, MAE: 72.49, RMSE: 285 and MSE: 82644.45. Finally, the model performance information and the new predicted values are accessible through the API. The following figure shows some of the parameters that the user must indicate to retrieve this information.

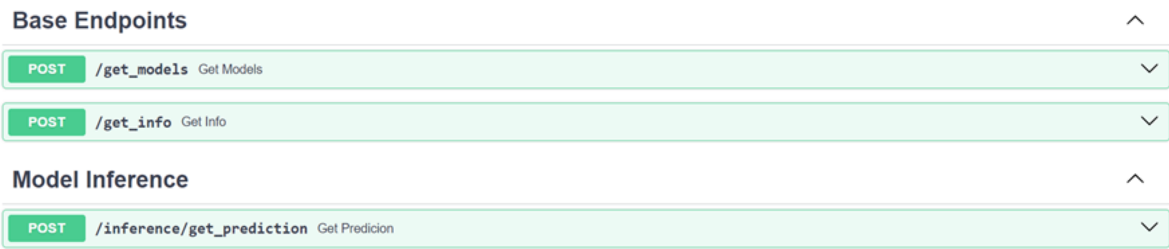
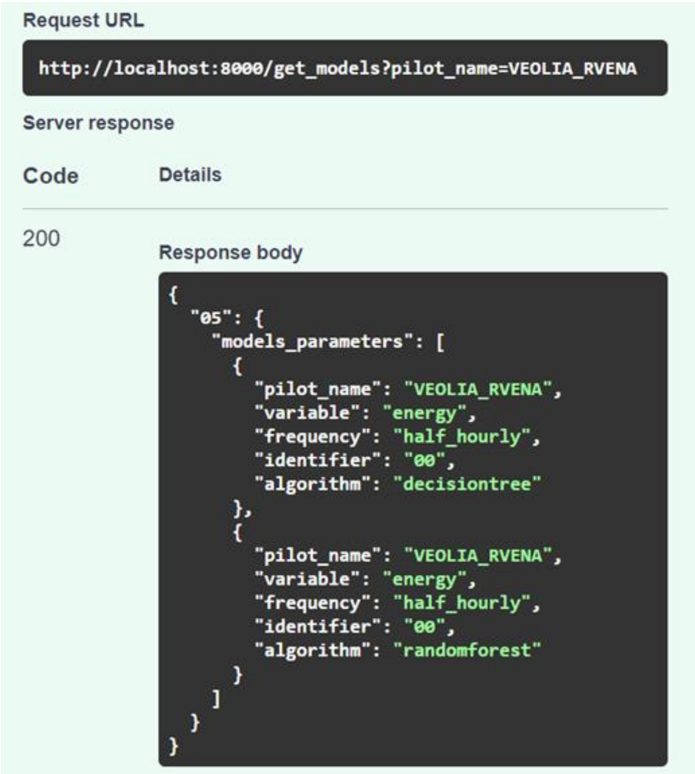
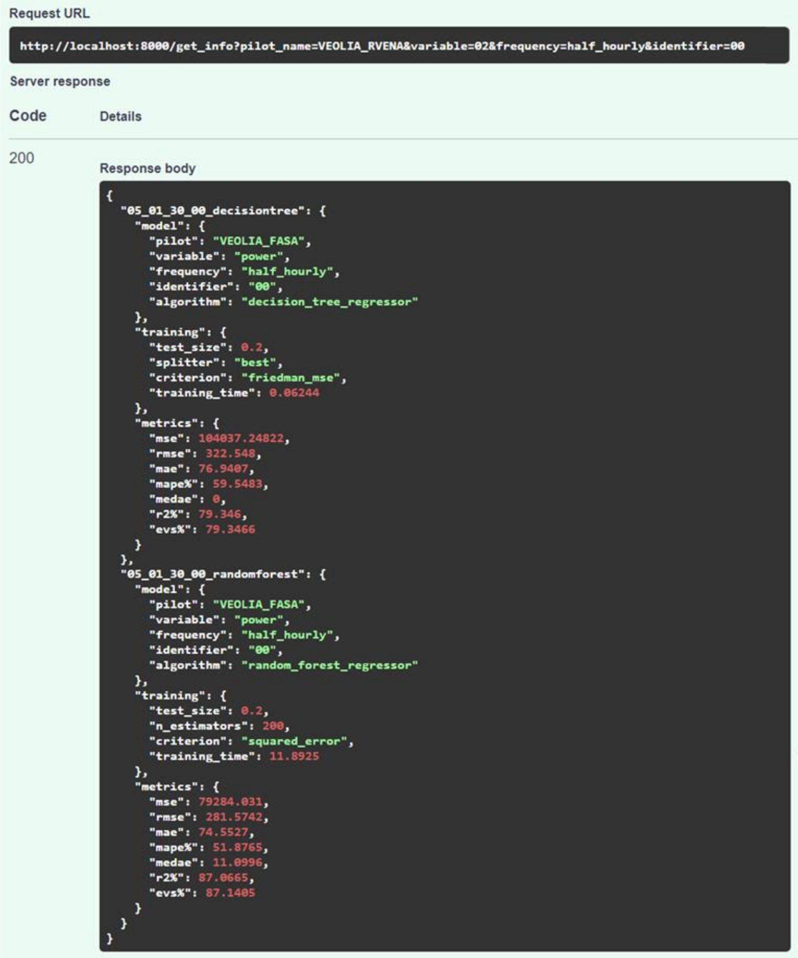


Figure 14: API definition.





Request URL

`http://localhost:8000/get_info?pilot_name=VEOLIA_RVENA&variable=02&frequency=half_hourly&identifier=00`

Server response

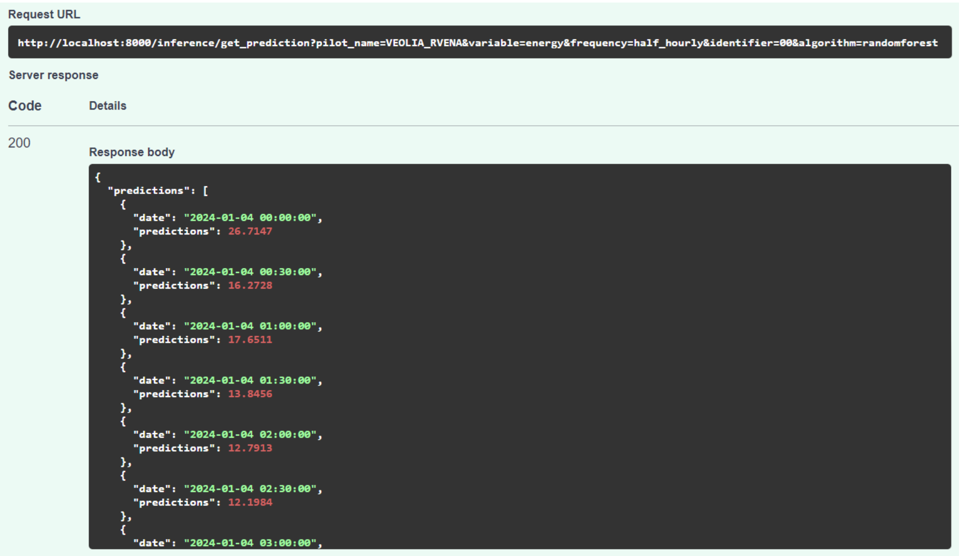
Code Details

200

Response body


```
{
  "05_01_30_00_decisiontree": {
    "model": {
      "pilot": "VEOLIA_FASA",
      "variable": "power",
      "frequency": "half_hourly",
      "identifier": "00",
      "algorithm": "decision_tree_regressor"
    },
    "training": {
      "test_size": 0.2,
      "splitter": "best",
      "criterion": "friedman_mse",
      "training_time": 0.06244
    },
    "metrics": {
      "mse": 104037.24822,
      "rmse": 322.548,
      "mae": 76.9407,
      "mape%": 59.5483,
      "medae": 0,
      "r2%": 79.346,
      "evs%": 79.3466
    }
  },
  "05_01_30_00_randomforest": {
    "model": {
      "pilot": "VEOLIA_FASA",
      "variable": "power",
      "frequency": "half_hourly",
      "identifier": "00",
      "algorithm": "random_forest_regressor"
    },
    "training": {
      "test_size": 0.2,
      "n_estimators": 200,
      "criterion": "squared_error",
      "training_time": 11.8925
    },
    "metrics": {
      "mse": 79284.031,
      "rmse": 281.5742,
      "mae": 74.5527,
      "mape%": 51.8765,
      "medae": 11.0996,
      "r2%": 87.0665,
      "evs%": 87.1405
    }
  }
}
```


Figure 15: API query results for available models in Río Vena (s3.1.3 District Network Production Forecast Service).



Request URL

`http://localhost:8000/inference/get_prediction?pilot_name=VEOLIA_RVENA&variable=energy&frequency=half_hourly&identifier=00&algorithm=randomforest`

Server response

Code Details

200

Response body


```
{
  "predictions": [
    {
      "date": "2024-01-04 00:00:00",
      "predictions": 26.7147
    },
    {
      "date": "2024-01-04 00:30:00",
      "predictions": 16.2728
    },
    {
      "date": "2024-01-04 01:00:00",
      "predictions": 17.6511
    },
    {
      "date": "2024-01-04 01:30:00",
      "predictions": 13.8456
    },
    {
      "date": "2024-01-04 02:00:00",
      "predictions": 12.7913
    },
    {
      "date": "2024-01-04 02:30:00",
      "predictions": 12.1984
    },
    {
      "date": "2024-01-04 03:00:00",
```


Figure 16: API query results for new calculations using the most accurate model for energy consumption prediction model in Río Vena (s3.1.3).

FASA DHN

During the second release for FASA DHN, work has focused on the definition of a methodology to provide DHN consumption predictions in an efficient way. In a preliminary phase, a first analysis of the dataset was carried out to identify the most relevant characteristics of the network to be considered in the modelling activities. Subsequently, the two steps to be performed were defined, involving, firstly, classifying the buildings into multiple sets according to their behaviour, and, secondly, exploring the most convenient regression techniques to apply in order to provide accurate predictions.

Buildings are grouped according to similarity of construction characteristics in order to minimise the computational burden required to estimate the energy consumption of a large set of elements (in this particular case, 20 buildings). One representative element per group is selected and the results are extrapolated to the rest of the buildings in the district considering the corresponding heating surface. The parameters considered for this classification include, the orientation, the size of the building, the existence of an adjoining block, and the presence of shading or windows. In a second step, regression algorithms are explored to model the heating consumption needs at building level. Predictions will be made using hourly or daily consumption and this information will be exploited in the DHN optimisation service and displayed via the digital district twin.

3.3.5 Next Steps

For this service, different levels of development are being managed in parallel, as two different approaches are adopted. Thus, in the case of the Rio Vena DHN, a first preliminary version of the service is ready and fulfils the needs of the pilots and the requirements of other services. On the other hand, for the FASA DHN, although the approach is designed, it is necessary to finalise the training and improvement phases of the models for the different energy assets. Similarly, access to model predictions and performance metrics will be provided through APIs to ensure seamless integration.

3.4 Energy forecasting (s3.1.4)

3.4.1 Description of the Service

Energy forecasting is essential for the smartification of buildings, as it helps optimise energy consumption, enhance efficiency, and improve sustainability. It enables energy managers to plan and allocate resources efficiently, as well as manage loads and schedule tasks and operations to ensure that energy is available when needed, prevent energy wastage, and maximise the use of clean energy. Moreover, it assists them in anticipating high-energy consumption times and manage production and demand accordingly to reduce costs and shave peaks. Finally, energy forecasting can minimise downtime, extend the lifespan of equipment, and improve grid interaction and demand response. As such, s3.1.4 can effectively serve as a basis for supporting data-driven solutions aiming at energy profiling, performance monitoring and benchmarking, energy resources management, load management, and predictive maintenance.

Energy forecasting is distinguished into load and power generation forecasting. In load forecasting, the objective is to predict the energy consumed in a building (or a certain space of it, also possibly per load type) over a time horizon, usually from a few hours to a few days. Energy consumption is characterised by strong seasonal patterns, observed on a daily, weekly and annual basis, but it is also affected by calendar factors (e.g. holidays and special events) and, to a lesser extent, weather conditions (e.g. outdoor air temperature and humidity) and energy prices. In power generation forecasting the objective is to estimate the expected energy production which, when it comes to buildings, typically refers to PV plants. Solar power is stochastic in nature and mostly driven by weather conditions. Therefore, in contrast to load forecasting, weather conditions (e.g. solar radiation and cloud coverage) and time of day are the most critical factors for accurately predicting power production.

Based on the above, in the context of s3.1.4 development, energy forecasting models were designed according to the following axes:

Time series data: Historical load and power generation data are typically provided as time series, where observations are collected over consecutive, constant periods of time. Time is a critical factor in analysing trends and seasonality and can, therefore, provide a strong basis for constructing generalised prediction models.

Explanatory variables: Identifying features that influence energy demand and production can help improve forecast accuracy. In this respect, if available, weather (e.g. air temperature, humidity, solar radiation) and calendar variables are tested in terms of utility and included in the forecasting models as explanatory variables. In addition, past observations of the target series are considered as features to facilitate pattern recognition (e.g. capture seasonality) and assist the specification of the running level of the series. These features are engineered and selected using cross-validation techniques.

Data pre-processing: In most of the cases, raw historical data that are collected by sensors and smart meters will involve missing and unusual values. Therefore, preparing and cleaning the data used for training the model can significantly improve data quality and, consequently, forecast accuracy. The developed service supports data pre-processing by handling missing values, removing outliers, scaling, and transforming variables. Note that data pre-processing may also involve temporal aggregation, i.e. summing/averaging time series data across time so that the frequency of the raw data is transformed into the desired frequency (e.g. transform 5-min data into hourly observations).

Model selection: The rise of ML has created a rich arsenal of forecasting models. As a result, depending on the application and the data available, different models may perform best. In this regard, the service conducts a "forecasting competition" using cross-validation to determine which model is most appropriate for producing forecasts for the forecasting task at hand. For the sake of brevity, the pool of models tested includes conventional time series models (ARIMA and Exponential Smoothing - ETS), multiple linear regression - MLR, and ML algorithms, such as tree-based models (Random Forest- RF - and Gradient Boosting Trees - GBT) and neural networks (Feed-Forward – FF - and Long Short-Term Memory -LSTM- networks).

Hyperparameter tuning: When it comes to ML-based forecasting models, which involve a high number of parameters and hyperparameters controlling their training process, defining the optimal values of the latter is critical to enhance forecast accuracy. To that end, after selecting the most appropriate forecasting model, its hyperparameters are tuned. However, in order to reduce computational cost, for each model the optimisation search space covers a certain set of hyperparameters that are regarded as the most influential ones (e.g. number of trees and learning rate for Gradient Boosting Trees).

Ensembles: It is widely accepted that combining forecasts of multiple, diverse models can help tackle model, parameter, and data uncertainty, thus improving overall forecast accuracy. In this regard, the service supports ensembles of two or more forecasting models from the pool of models defined previously.

Note that, by default, cross-validation assumes an 80-20 split analogy of the historical data into train-test sets so that the results are sufficiently generalised. Moreover, given that the most recent data are more relevant for making new forecasts, the test set covers the last part of the series instead of being created by randomly selected observations. In terms of accuracy measures, the mean squared error of the forecasts is used for feature selection, model selection, and hyperparameter tuning, although other measures can be supported.

Finally, in all the evaluations performed, the forecasting horizon is set equal to the horizon to be considered for the actual forecasts so that simulation results closely represent reality.

An overview of s3.1.4 development process is provided in Figure 17, while the building blocks of the foreseen architecture can be seen in Figure 18.

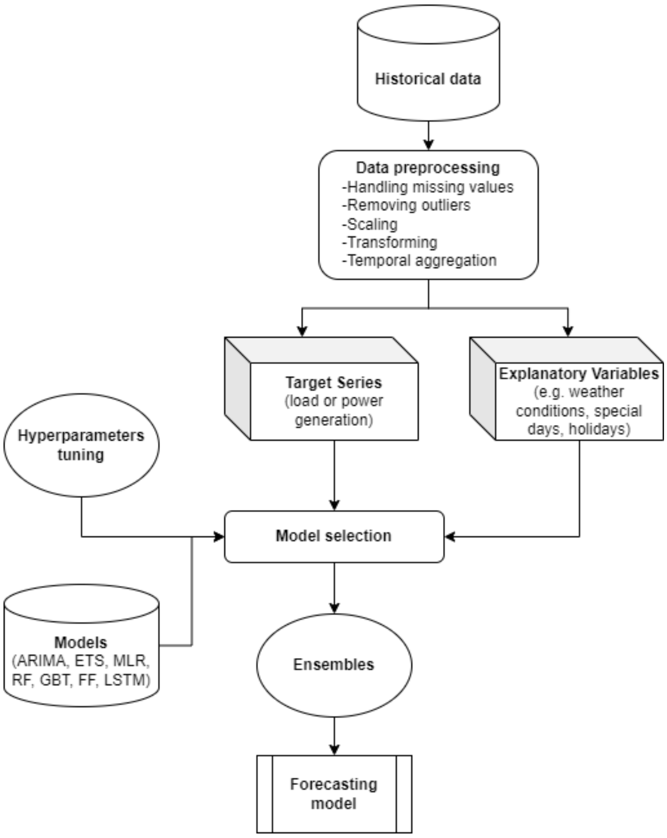


Figure 17: Overview of s3.1.4 approach.

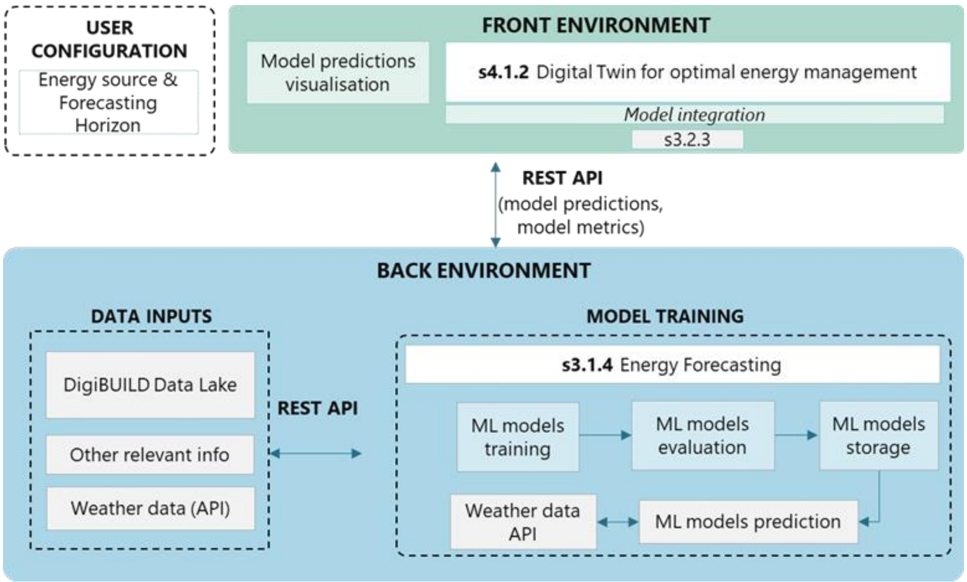


Figure 18: Overall architecture of s3.1.4 Energy Forecasting.

3.4.2 Novelty

A significant number of forecasting methods have been developed over the years to forecast power generation and load (Sarmas et al., 2022). In the field of solar power forecasting, ML regression methods like neural networks, decision-tree-based models, and support vector machines have been identified as the most appropriate ones due to their ability to effectively consider multiple factors influencing PV production and account for nonlinear relationships. Nevertheless, the quantitative comparison among different forecasting methods is challenging in practice as their accuracy depends heavily on the quality and resolution of the historical data available, the forecasting horizon examined, and the precision of the weather forecasts provided. Similar conclusions stand true for load forecasting. The literature suggests that time series methods, although intuitive and fast to compute, are limited in terms of adaptability and capability to model non-linear relationships. ML forecasting methods can effectively deal with these limitations, being also capable to efficiently incorporate information related with weather conditions, calendar effects, special days and events, and other factors that may affect demand. The most popular ML models used in the field are NNs, but regression-tree-based ML models, like GBT, have also become popular, showing promising results in various load forecasting applications, while being relatively faster to compute and easier to parameterise than NNs.

Given these challenges, the novelty of s3.1.4 lies exactly on its ability to generate accurate load and power generation forecasts for buildings of different location, type, technical and behavioural characteristics, also supported by different sets of explanatory variables and historical data of different size and quality. By integrating data pre-processing, model selection, feature engineering, and hyperparameter tuning processes, the service is effectively generalised, allowing the production of forecasts in an automated, unsupervised fashion for a wide range of forecasting applications. Moreover, by introducing ensembles of models, the forecasts become far more robust and trustworthy, thus supporting a variety of decisions related to energy management.

The developed service follows the principles of AutoML solutions that have been recently introduced to automate the end-to-end process of applying ML to real-world problems. As such, s3.1.4 automates the forecasting workflow, significantly reducing the need for manual intervention and expertise at each step of the process. Consequently, it makes forecasting more accessible to individuals with limited expertise, allowing domain experts, data analysts, and business users to leverage ML without requiring an in-depth understanding of the underlying algorithms. Benefits are also present in terms of efficiency and speed as the service iterates through multiple models and configurations, finding the best-performing solution in a fraction of the time it would take through manual trial and error. The service is also more scalable, allowing it to handle large datasets and complex forecasting problems efficiently, it can quickly adapt to changing data, and can be integrated easier with cloud services. Overall, the design principles and elements of s3.1.4 contribute towards the democratisation of energy forecasting, allowing a broader range of users to harness the power of state-of-the-art forecasting algorithms for solving problems in various domains.

3.4.3 Development Progress

The activities during the first release for EMOT focused on the identification of relevant datasets, the connection to other services and the collection of first requirements from the point of view of other services. No significant development activities were reported, while this work progressed over the following months. During the second release, data exploration and analysis activities were carried out to understand consumer partners through visualisations of the available datasets, namely information on electric vehicles and charging station usage. Additionally, the set of requirements and connections between the results of this activity and the implementations of other WP3/4 services were refined. The initial models were developed considering an hourly forecast horizon and validated using the most common metrics, and in a subsequent step the results have been exposed through an API for a seamless integration. The lack of detailed information for the other datasets, such as PV production and energy consumption in buildings, has delayed the related activities, which will be addressed in the next technology release following a similar procedure.

In order to put the energy forecasting models in use according to the principles of the DigiBUILD architecture, two separate APIs are currently developed, the first being responsible for implementing the load forecasting model, while the second for implementing the PV production forecasting model.

The APIs are being developed in Python using the *pandas*, *numpy*, *StringIO*, and *BytesIO* libraries for handling the data and making basic numeric operations, the *fastapi* library for enabling efficient API calls, as well as the *lightgbm* and *sklearn* libraries for implementing the forecasting models and supporting their tuning and training process.

Specifically, each API is tasked to retrieve the historical data required for training the model, create the respective time series, pre-process the data (handle missing values and outliers, apply transformations, scaling, and temporal aggregation), engineer features, select the most suitable forecasting model, and tune its hyperparameters. Once the above workflow has been complete, the trained model is stored at MLflow so that it can be easily accessed and used for inference using a secondary API. As a result, the forecasting models can be retrained either periodically (e.g. on a weekly basis) or on demand to take into account the most recent data and improve performance. At the same time, the inference phase of the model is sped up and simplified.

The source code can be found on GitHub: GitHub s3.1.4

3.4.4 Application on DigiBUILD Pilots

The service will be tested and validated in pilot 5a (EMOT) and pilot 6 (HERON).

- **Pilot 5a (EMOT):** this pilot aims at improving energy management processes. To do so, energy production and consumption values are processed by ML models in order to help obtain optimal energy management, taking advantage of information coming from photovoltaic production, building and electric vehicle fleet consumption and charging station usage. These predictions will be used for optimisation (s3.2.3 Power recharging management) and integrated in a digital twin (s4.1.2 Digital Twin for optimal energy management). Details of the implementation of s3.4.1 for the EMOT pilot are described in UC16 (D1.2).
- **Pilot 6 (HERON):** the objective of HERON is to predict the electricity consumed by EV chargers available at various locations, including residential and business buildings, but also the electricity provided by a PV plant. Thanks to this, it will be possible to cover UC_3. Moreover, through services s3.2.3, s3.2.4, and s3.2.5 of WP3, which will be connected, UC_29, UC_26, and UC_30 will be fulfilled.

During the second release, the available datasets in EMOT have been considered (electric vehicles and charging stations usage) for data exploration and analysis, identification of connection between the results of this activity and other WP3/4 services and the first model development. The models have been trained according to the most appropriate forecast horizon for the needs of other services and validated using the most common metrics. Different state-of-the-art ML models have been trained for the estimation of the VE SoC, such as the Random Forest regressor, the Gradient Boosting regressor, the Lasso regressor or the MLP, although further iterations and refinements might be necessary when more historical data become available. Preliminary results show on average results in the following range: R2 values of 0.98, MAE: 0.36 RMSE: 0.01 and MSE: 4.26. In the case of the charging station data, the average results show an accuracy of 0.89 for R2, MAE: 1.701, RMSE: 2.91 and MSE: 8.57. Finally, the model performance metrics and predictions are accessible through an API.

Up to this point, the service has been successfully applied on one DigiBUILD pilot, HERON, where energy forecasts were required both for predicting load and power generation. The service is provided as an API so that its results can serve as input to other services, such as s3.2.3, s3.2.4, and s3.2.5.

HERON is interested in forecasting the electricity consumed by EV chargers installed in various locations, including residential and business buildings, as well as the electricity produced by a PV plant. The objective is to exploit such forecasts to improve overall energy management and sustainability. Specifically, the forecasts are primarily required on a daily basis (forecasting total consumption for the following day), but hourly forecasts are also relevant (forecasting total consumption for the following day, disaggregated per hour). As such, although the forecasting horizon is constant, the frequency of the forecasts may vary.

Given that EV chargers data were not accompanied by any explanatory variables, in this case feature engineering was limited to the creation of lag and calendar variables. In contrast, PV production forecasts were supported by historical weather forecast data (e.g. solar irradiation) retrieved by Power Data Access Viewer. Following the service’s workflow, a set of time series and ML methods were evaluated and a special case of GBM, namely LightGBM, was identified in both cases as the most suitable model for producing day-ahead forecasts.

Figure 19 presents some indicative results of the application of the load forecasting service at the HERON pilot. The bars summarise the forecast accuracy of the assessed models in terms of Root Mean Squared Error (RMSE), while the Naive model (marked in yellow) effectively serves as a benchmark since it bases its predictions solely on the previous week’s historical data. As seen, the time series forecasting models (marked in red) improve the accuracy significantly over the benchmark, but LightGBM achieves best results. The rest of the ML models tested (marked in blue) perform relatively worse. Some ensembles of the time series and ML models (marked in green) are also tested with moderate results. Based on these results, the API would store the trained LightGBM model at MLflow to be latter used for inference, as indicatively presented in Figure 20. Moreover, once the daily forecasts have been computed, hourly forecasts can be estimated using the seasonality indices depicted in Figure 21. As seen, reasonably enough, most EV users opt to charge their EV vehicles during night hours.

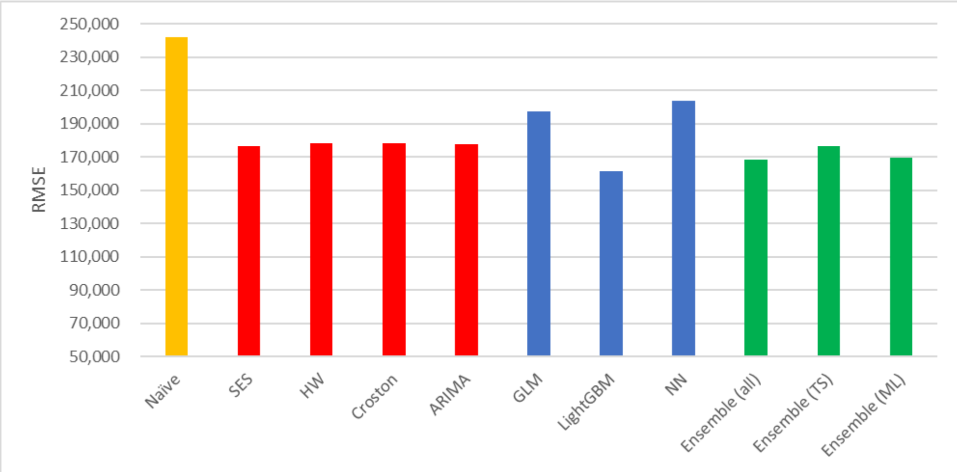


Figure 19: Forecast accuracy (RMSE) of the models tested by the service for predicting (day-ahead).

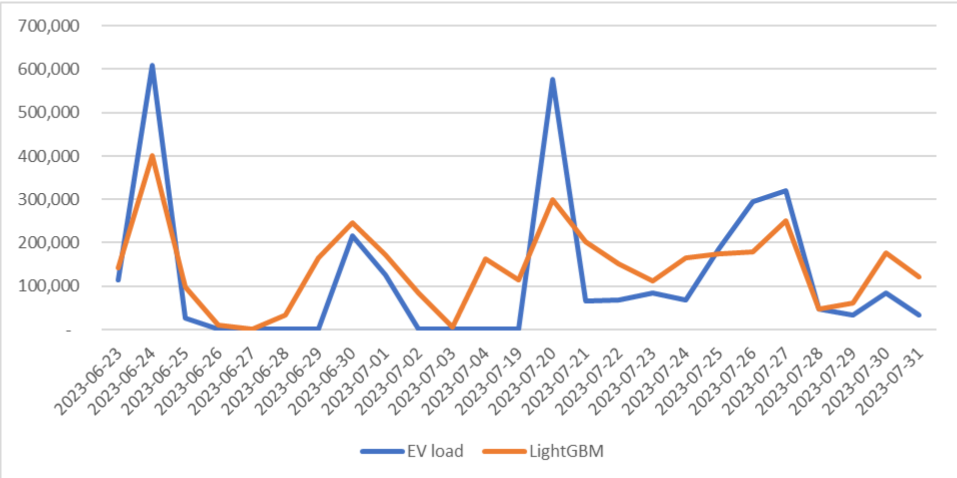


Figure 20: Example of day-ahead forecasts of the LightGBM model used to predict EV load at the HERON pilot.

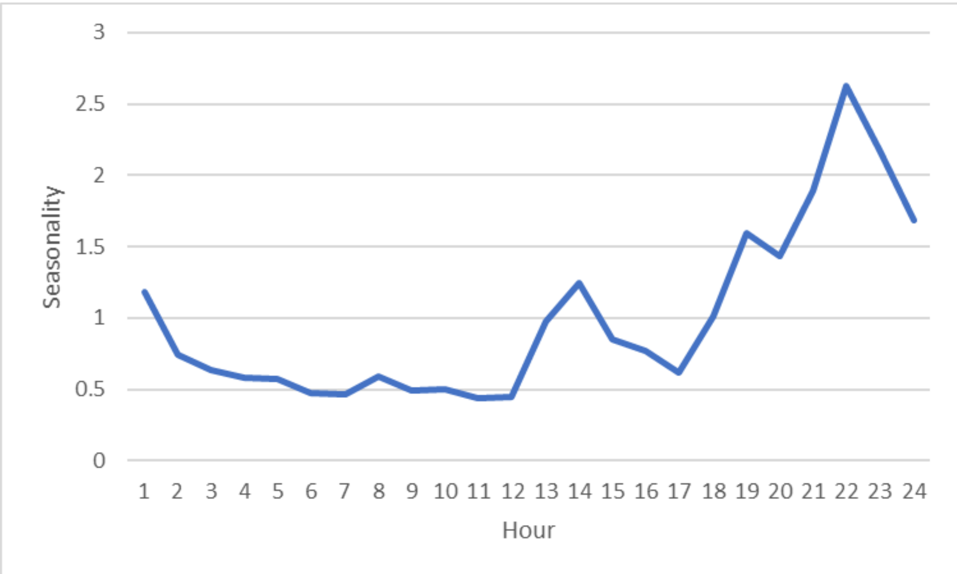


Figure 21: Seasonal indices used for disaggregating daily EV load forecasts into hourly ones at the HERON pilot.

3.4.5 Next Steps

The next steps for s3.1.4 in the HERON pilot are outlined as follows: (1) complete the development of the APIs used for training the forecasting models and their integration in MLflow, (2) enhance the complete forecasting workflow (this will be a continuous process that will be carried out throughout the implementation in order to further improve accuracy), (3) incorporate new monitoring devices and data to retrain the models and improve performance, (4) finally, complete the extension of the service for EMOT pilot and validate its performance.

The next steps foreseen for this service in the EMOT pilot include data exploration, modelling and API development for access and integration of the service. The first, second and third activities are mainly planned for PV production and consumption at building level, as the lack of high-resolution data has delayed these activities. Moreover, in the upcoming months of the project, the algorithms already trained for other datasets available at the demonstration site, such as electric vehicle charging consumption and charging stations, can be refined to improve the accuracy of the model, taking advantage of the use of a larger dataset. Finally, in case other features are identified that need to be provided by existing endpoints, they could be adapted.

4 Data-driven services for energy resources management

DigiBUILD project focusses on reducing carbon footprints and enhancing system resilience through data-driven services, contributing to the energy sector's decarbonisation. In this context, Task 3.2 centralises on optimising different systems through various approaches in all the involved pilots. In s3.2.1 the main focus is the implementation of novel algorithms at the University College of London and IASI municipality for real-time building management, detecting sensor faults and consumption anomalies. District Heating Network Optimisation (s3.2.2) is exploited in two distinct networks (CP Rio Vena and CP Fasa), where different optimisation strategies are employed. Rio Vena aims at minimising gas boiler consumption for cost and greenhouse gas emission reduction. CP Fasa's strategy integrates multiple heat sources (PV, gas, biomass boilers) for a greener optimisation. In the Heron Pilot, services enable knowledge of the hourly carbon footprint to minimise environmental impact. Novel approaches estimate 'greenness' of hours (s3.2.3), optimise the electrical grid (s3.2.5) and provide real-time information (s3.2.4). The Emotion application focusses on maximising PV self-consumption and minimising greenhouse gases. The provided services calculate optimal EV charging strategies (s3.2.3), considering user preferences converted into energy requirements (s3.2.4), and offer real-time control of charging stations and heat pumps (s3.2.6). All these services will be detailed in the next sections of this chapter focussing on their description, novelty and main achievements.

4.1 Pro-active maintenance and facility management (s3.2.1)

The increasing availability of smart sensors and IoT devices leads to a constant availability of data, in most cases real-time, providing users with greater possibilities to manage and control buildings. By exploiting this amount of data, algorithms can be developed to monitor real-time conditions of the building and, in this way, optimize its management, reduce costs, and plan maintenance. This added value also has an important effect on reducing consumption, and therefore reducing the building's environmental impact as well. In addition to other innovative services, this service also allows to provide the best living experience for the occupants of the building by combining real-time monitoring with comfort and well-being monitoring, assessing the real-time condition of the occupants.

4.1.1 Description of the Service

In the context of DigiBUILD, this service will implement a rule-based algorithm that aims at detecting faults at sensor level. Different conditions are monitored, starting from various situations that could arise during the daily operation of the building. For instance, if a room monitored from the sensors installed in the Pilot's premises register a temperature higher than a certain value, the algorithm will identify the issue and notify the Building Manager (BM), providing a relevant solution.

In addition, the algorithm is being developed with the objective of being used in conjunction with the prediction services implemented in s.3.1.1 and s3.3.1. In this regard, the service will be useful in a twofold manner. In the case of the predictions provided by 3.1.1, the service will detect potential discrepancies between the predicted trend of a certain value and the real one and will notify the BM accordingly. On the other hand, the service will allow the monitoring of comfort, exploiting the comfort estimation provided by s3.3.1 and the data coming from the sensors.

Given this comprehensive input, the output of the service will be threefold:

- An alert notification, shown on the Digital Twin Fault Detection page. The type of notification will be different according to the type of Digital Twin the pilot opts to utilise.
- A report on the detected anomaly, indicating the sensor that produced the anomalous value, along with the associated timestamp.
- A strategy or solution to compensate for the detected fault.

A detailed schema of the input/output pattern of s3.2.1 is provided in Figure 22.

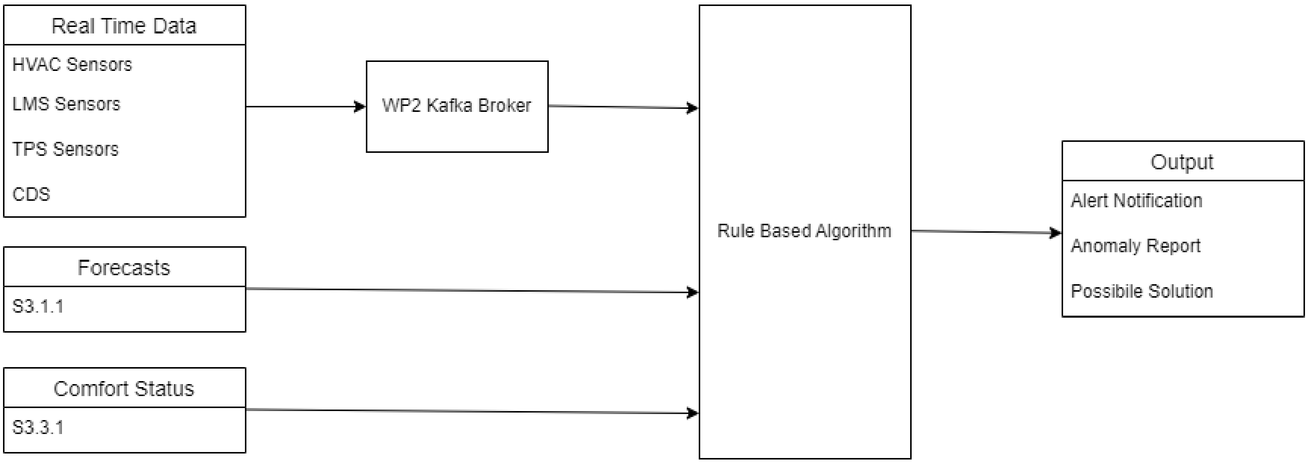


Figure 22: s3.2.1 input/output schema

4.1.2 Novelty

The growing number of available real-time data has made the creation of real-time fault detection algorithms possible. Each of these, with their various particularities, allows the state of the building to be monitored and inspected in real-time. In this case, the lack of historical data did not allow the training of data-driven models. However, the availability of real-time data guided the development towards a rule-based algorithm. Examples of rule-based fault detection algorithms can be found in (7) and in (8) also focusing on HVAC systems AHUs.

There are several aspects of novelty that we can identify in the presented algorithm, starting with its composition: the algorithm has a generalised set of rules that makes the software easily scalable. As a result, the algorithm can be utilised by different Pilots. Moreover, the algorithm is accompanied by a Pilot-Specific set of rules that allows the real-time and specialised monitoring of buildings. Those rules will focus not only on the ventilation system, but it will have a broader scope which also considers temperature and lighting variables. Moreover, the possibility of receiving a real-time notification with a solution suggested to the identified issue will simplify building management operations. The exploitation of the predictions provided by s3.1.1 and the comfort status provided by s3.3.1 will also provide added value to this service, adding the possibility of defining a set of innovative monitoring conditions that will take into account possible future faults and monitor the comfort level of the occupant and expanding the scope of the algorithm beyond the simple device monitoring.

4.1.3 Development Progress

The main core of the algorithm will be represented by a series of "if-then" rules, connected to the real-time data broker provided by the Pilot, resulting to a real-time fault detection model. At the time of writing this deliverable the connection with the data source is implemented using a simple Python MQTT client, that can connect to all the available room-related topics streamed by the Pilot. An example of such rules is reported in

Figure 23.

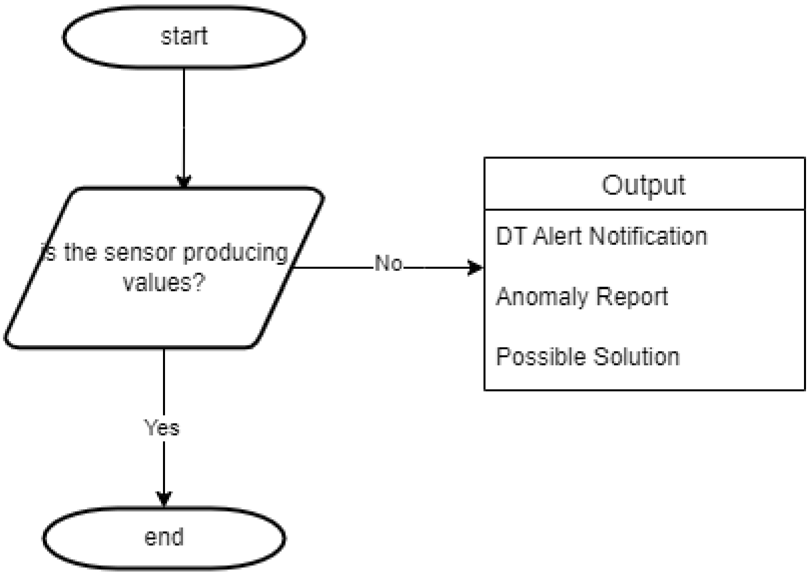


Figure 23: Standard rule diagram of s3.2.1.

At this stage of the development, only five basic and arbitrary rules have been defined to create a baseline for the service and provide control over the most typical fault situations. Rule 1, for instance, aims at controlling the functioning of the sensors: If a sensor is not providing data for a specific time interval, it will be assumed to be problematic in some way, and the system will notify the BM about this fault.

A brief description of the rules initially developed is provided in Table 3. As explained, apart from a set of more generic rules, that can be generalised, a set of personalised rules will also be defined in a pilot-specific manner, to adapt the service to Pilots' specific requirements.

The source code can be found on GitHub: GitHub s3.2.1

Table 3: s3.2.1 first six defined rules

Rule Name	Scope	Applied to
Rule 1	Detect if a sensor is not providing data for a certain time interval	All relevant sensors
Rule 2	Detect if a sensor provides anomalous data (independently of the other specific conditions)	All relevant sensors
Rule 3	Detect if a room temperature is too low/high	Temperature sensors
Rule 4	Monitor the functioning of the Lighting System (e.g. control if the lights are on in an incorrect interval of time)	Lighting management system
Rule 5	Detect if CO2 concentration is higher than a certain threshold	CO2 monitoring sensors

4.1.4 Application on DigiBUILD Pilots

- **Pilot 1 UCL:** Currently, the algorithm has been developed following the structure of the MQTT Broker that registers the data coming from Pool Street West building of the UCL Pilot. When fully operational, the service will be continuously connected to the broker, receiving values regularly. For each room, the system can monitor LMS values, CDS values, and TPS values, which respectively measure the lights (on/off values), the CO_2 levels and the indoor temperature. Using these values, it is then possible to set a range of operativity within which the conditions of the rules are not triggered. For instance, if the temperature drops below 18.5 °C the model will identify an error. An example of this rule is depicted in Figure 24.

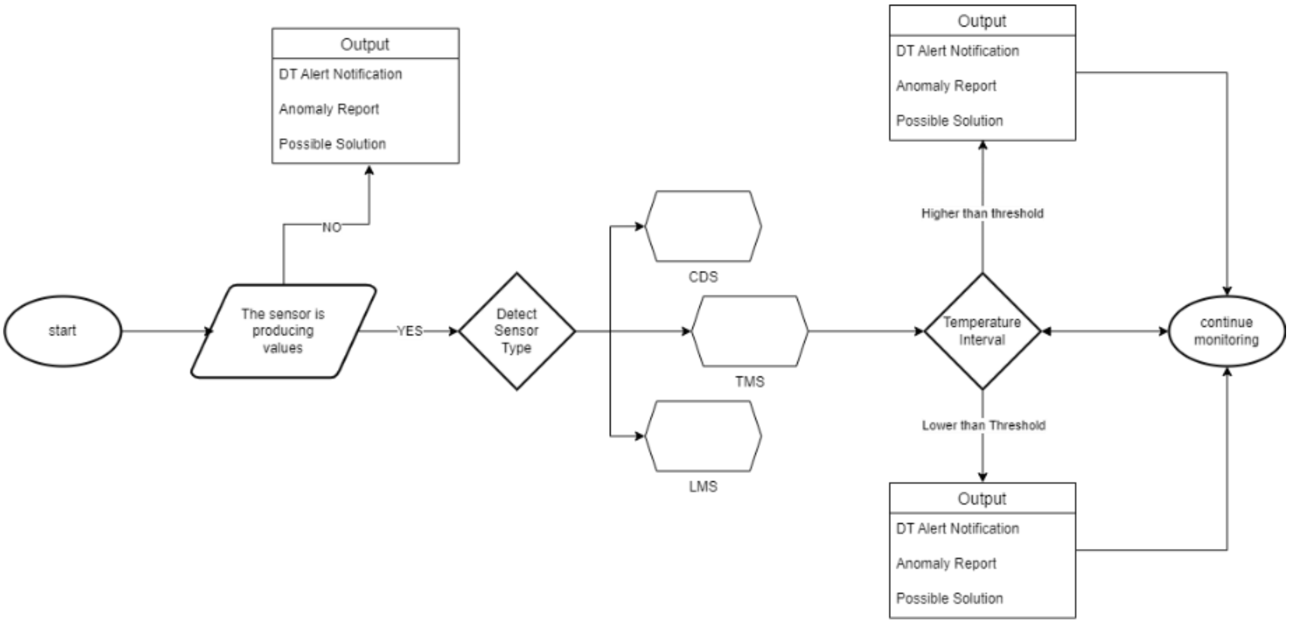


Figure 24: Flow diagram of s3.2.1 for the UCL Pilot.

When an anomaly is identified, the service will continue to work, but will notify the manager through a comprehensive report. If there are no anomalies, the model will periodically return a report on the status of the connection, indicating how many messages it has received so far, how many null values it has recorded, and the total number of anomalies. At this stage of the development, this report is generated every 60 seconds, but due to its exposure in the twin, this time horizon will be extended, and the connection status of the service will be displayed in a different way.

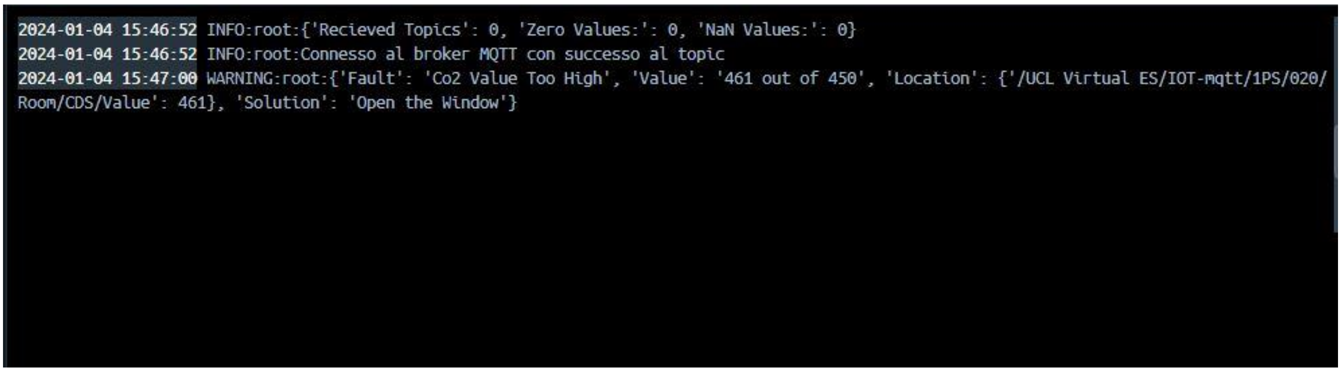


Figure 25: Example of a warning detected by s3.2.1.

In Figure 25, it is shown that each time a rule is triggered, the algorithm returns a complete and explicative response, so that the manager can act to prevent future faults. For this example, the rules have been forcibly triggered to achieve a fault, but clearly, they will be active in real contexts with real and representative thresholds.

- **IASI & SITTA:** Until this point, the service has not been applied to Pilot 3. Therefore, in the following months, the main general rules will be extended to cover the pilot's needs. Moreover, a proper set of dedicated fault detection rules will be defined. Nevertheless, the operation of the service is meant to be similar to Pilot 1, with the required pilot specific adaptation.

4.1.5 Next Steps

The following schedule is envisioned for the service:

- A comprehensive set of conditions will be defined, starting from the pilot's needs, and proceeding to a more general application of the rules. Proper communication between the developers and the Pilot leaders will be planned to finalise the service.
- Application of the service to Pilot 3.
- Connection with interdependent services. S3.2.1 takes as input results from s3.1.1 and s3.3.1. They both communicate through REST APIs, so proper connections need to be developed.
- Exposure of the service in the DT. Currently, the service is neither exposed, nor exploited by any DT interface. In the following months, proper APIs to expose the service to the respective DTs will be developed.

4.2 District network production economic optimisation (s3.2.2)

In earlier deliverables, it has been established that District Heating Networks (DHNs) are crucial to achieve the decarbonisation of the energy sector. The latest networks, known as the fourth- and fifth-generation networks, combine different energy sources like electric heating, solar power, and heat from waste to efficiently boost heat production. In this context, it is important to investigate new approaches to make these networks operate in a greener way. The DigiBUILD project, specifically its service s3.2.2, introduces novel algorithms capable of improving the operation of two different DHNs. This includes one belonging to the widely used third-generation DHN, CP Rio Vena, and another for the more advanced fourth-generation network (CP Fasa), which combine different heat sources (natural gas, biomass and power-to-heat).

4.2.1 Description of the Service

Service s3.2.2 is designed to guide efficient management of District Heating Networks (DHNs) to lower operational costs for the network manager, Veolia, benefiting consumers with cost reductions. Importantly, cutting costs in fossil-fuel-dependent networks also translates to decreased CO2 emissions. Given the declining gas prices, the service at CP Fasa, which utilises multiple heat sources, incorporates carbon footprint considerations, leading to a multi-objective optimisation approach. Distinct optimisation algorithms cater to the unique characteristics of CP Fasa and CP Rio Vena, as detailed in deliverable D5.3.

For CP Rio Vena, the service acts as a decision-support tool, helping the manager choose optimal setpoint temperatures to reduce gas usage, costs, and emissions. It employs data-driven models, trained on the available data since week 17 of 2023 (details available in D5.3), to estimate gas boiler efficiency based on energy demand and supply temperature. These models are integrated into an NSGA-2 algorithm, using the pymoo library, to recommend daily optimal supply temperature setpoints, in timestep up to half hour, using as input the forecasted heating energy demand of the district (from service s3.1.3).

The CP-Fasa service utilises extensive data sets, more than CP-Rio Vena one, offering diverse pathways for the optimisation algorithm. This service, applied to a fourth-generation DHN, extends beyond merely improving gas boiler efficiency. It aids the DHN operator in determining the most appropriate heat source between the three available ones. Initially, the service examines the relationship between Heating Degree Days and the demand for heating, employing linear regression analysis, but being extended to other models with higher accuracy. Following this, Particle Swarm Optimisation is applied to enhance the real DHN's performance, as simulated via the DHNx python library (WP4). This process also accounts for operational limits, such as maintaining a district minimum temperature of 65°C to prevent the growth of Legionella bacteria and adhering to the comfort temperature settings required at user substations.

Both services will be integrated into a digital twin of the respective DHNs, empowering facility managers with actionable insights to smartly and effectively control the network for optimal performance.

4.2.2 Novelty

District Heating Networks (DHNs) are globally growing in complexity and usage, as noted in report (9) by the International Energy Agency (IEA), which also highlights the need to enhance their management to reduce environmental footprints. Currently, many DHNs rely on fossil fuels due to their dependability and cost-effectiveness. However, there are instances, like CP Fasa, where networks are augmented with renewable energy sources (RES), including biomass boilers and photovoltaic-powered heat technology, to assist gas boilers.

As detailed by Lund et al. in (10), researchers are exploring various methods to design or manage new-generation DHNs effectively. Yet, these efforts often lack real-time application or remain theoretical when based on simulations from software like TRNSYS or Modelica. Even when validated with real-case data, these studies seldom explore tangible improvements on actual sites. Service s3.2.2 addresses these shortcomings in three ways. Firstly, it utilises actual data from pilot sites for a tailored, data-driven approach that is adaptable to other locations. Secondly, the implementation in CP Rio Vena serves as a blueprint for future research and DHN companies, presenting a straightforward method, in line with the pilot leader's operational controls, to enhance the main power plant's efficiency, thus reducing costs and emissions. Lastly, this innovative optimisation strategy will be applied in actual sites using a Digital Twin, demonstrating the real-world advantages of these methods.

4.2.3 Development Progress

In the Rio Vena DHN, the primary objective is to optimise boilers for cost reduction and enhanced heat production efficiency. As explained in D5.3, an initial version of this service for Rio Vena has been introduced, featuring preliminary ML models for boilers. These models were designed to predict gas consumption for Boilers 1 and 2, and efficiency for Boiler 3, based on six distinct inputs. A MLP was selected for its proven reliability and accuracy, a common choice in scientific research. The preliminary results, gauged by metrics like R2 and MAE, were promising based on the data collected. Additionally, this early version included an NSGA-2 optimisation algorithm, which iteratively determined optimal supply temperatures and operational hours. Details of this early implementation can be found in D5.3.

As this service is intended for actual deployment and integration into a functional DT, numerous discussions with the pilot leader were conducted to understand operational constraints and controllable variables on site. It was determined that boiler supply temperatures were the only modifiable parameters. Consequently, these temperatures, along with the DHN's energy demand, were incorporated as inputs in the ML model.

Given the new constraints, boiler supply temperature emerged as the sole optimisation parameter. With only

Funded by
the European Union

one variable to optimise, the optimisation problem simplified to an unconstrained one. The NSGA-2 algorithm was chosen for its efficiency and widespread use in swiftly resolving such problems by the use of pymoo library. In the latest version of the algorithm, the algorithm aims at recommending boiler supply temperature setpoints that maximise heat production efficiency, as shown for example in Figure 26.

The source code can be found on GitHub: GitHub s3.2.2

Day/Time	06:00:00	06:30:00	09:00:00	14:00:00	17:30:00	20:30:00	21:00:00
07/06/2023	81.5	-	89.8	69.6	84.7	88.5	88.9
08/06/2023	84.8	-	82.0	77.2	84.0	88.2	84.1
09/06/2023	86.5	84.0	88.9	82.9	88.0	86.1	71.3
10/06/2023	76.8	-	81.6	85.3	87.0	87.9	85.9
	Boilers Setpoint [°C]						

Figure 26: Example of boilers' setpoint suggestion exported by s3.2.2 based on optimisation results.

Second case of the district network operation optimisation is VEOLIA FASA DHN with a similar objective than Río Vena in terms of reduction of cost, but, with the constraint about the prioritisation of the biomass boilers. It should be highlighted that the gas prices have lowered a lot in the latest months, which could derive in a result of the optimisation algorithm that sets the gas boiler as primarily source. Note that the VEOLIA FASA DH network is composed by 2 biomass boilers and a gas boiler used as support for peaks. Then, to avoid the biased result for the price, the carbon footprint is included as objective function.

The concept of the service is depicted in Figure 27, where the optimisation service, which is based on Particle Swarm Optimisation (PSO) technique, is conceptualised. The service is interconnected with the digital twin and service s3.1.3. In this line, the workflow is as follows:

1. The simulation engine (part of the digital twin development and implemented by means of the DHNx python library) calculates the heat and pressure losses that are part of the network due to distribution systems (pipes, forks…).
2. Buildings energy demand should be summed to the losses. The result of this calculation provides the total energy that must be provided by the generation systems to cover both the buildings demand plus the heat losses in the distribution.
3. The case of FASA contains a PV generation system, used as Power-To-Heat. Then, it is positively contributing to the generation (i.e., decreasing the boilers needs as it reduces the demand). Having said that, the input for the optimiser is the boiler needs obtained as the total demand of the district minus the PV contribution.
4. Another input of the optimisation module is the user parameters, such as the comfort set-points that should be established in the demand side (buildings), as well as the energy prices (both gas and biomass).

The source code can be found on GitHub: GitHub s3.2.2

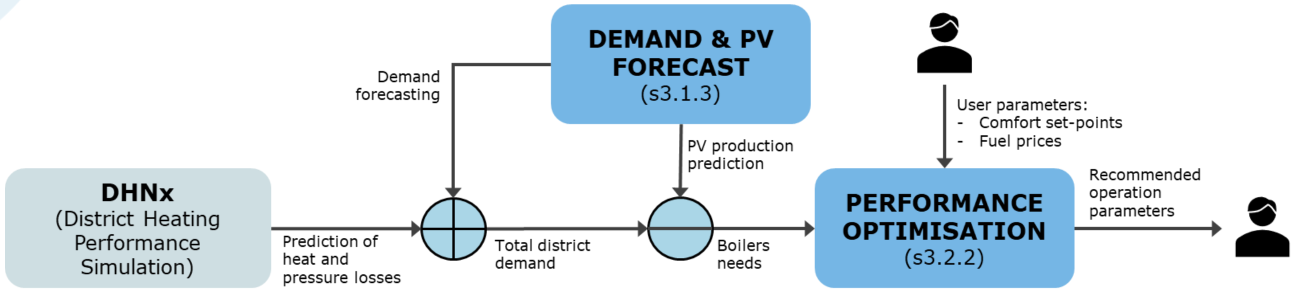


Figure 27: Optimisation service conceptualisation for VEOLIA FASA DHN

4.2.4 Application on DigiBUILD Pilots

VEOLIA - [RIO VENA - ENG] Unlike FASA, RIO VENA had incomplete data storage; relevant data for this service were only gathered starting from week 17 of this year. Prior to this, only the DHN heating demand, gas consumption, and boiler operational hours were collected. This issue led to a delay in RIO VENA's application, and the alpha version featured a ML model tested on just a few weeks of data.

As explained in the previous section, the algorithm's updated version is tailored to the DHN manager's interests and feasible interventions, aiming at maximising service utility. Discussions between ENG and VEOLIA revealed the practical difficulty in modulating individual boilers. In winter, Boilers 1 and 2 operate on alternating weekly schedules, but both are activated when demand exceeds a single boiler's capacity. Therefore, it was decided to modulate and incorporate the overall boiler supply temperature into the ML model, serving as the optimisation algorithm decision variable.

The ML algorithm continues to rely on a multi-layer perceptron, now focusing solely on the boilers' supply temperature and DHN's heating demand. The ML algorithm was trained on data from weeks 17 to 50 of 2023. The correlation matrix was examined to assess boiler efficiency, concentrating on Boilers 1 and 2 for winter (due to variable scheduling) and Boiler 3 for summer. Regarding the winter case, based on the available data, the R2 reported was 0.72 and the MAE was 0.1, while in summer the first metric reached 0.81 and the second improved to 0.05. Some indicative results are graphically shown in Figure 28.

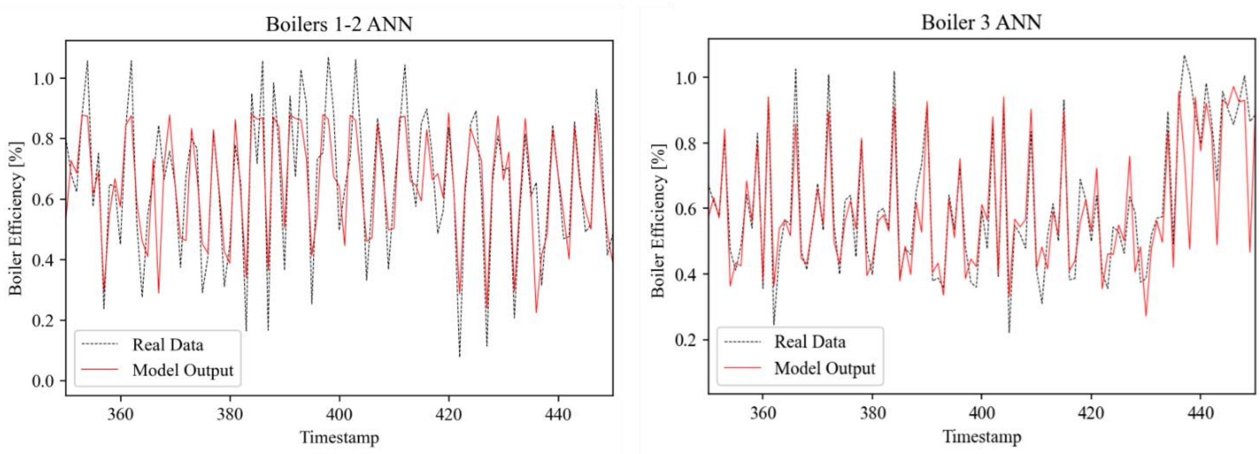


Figure 28: ML model results of s3.2.2. on indicative time periods of October (left) and August (right) for the RIO VENA pilot.

The optimisation process, now simplified, aims to determine the setpoint temperature that maximises boiler efficiency (with a limit of 102% based on boiler datasheets). An enhancement from the previous version is the consideration of optimisation over an entire period (the subsequent day in the DT) rather than at a single timestep. Post-training, the optimisation algorithm excludes periods with no energy demand (maintaining the previous setpoint) and suggests optimal temperatures for remaining hours to maximise efficiency. Simulations on filtered historical data (excluding instances of over 102% efficiency and no energy demand) indicated potential savings of about 32 MWh of NG for Boilers 1-2 in winter and 128 MWh in summer, an energy saving of 6.7% and 31.8%, respectively. Figure 29 visually presents the natural gas usage, both optimal and actual, alongside energy savings for both summer and winter periods.

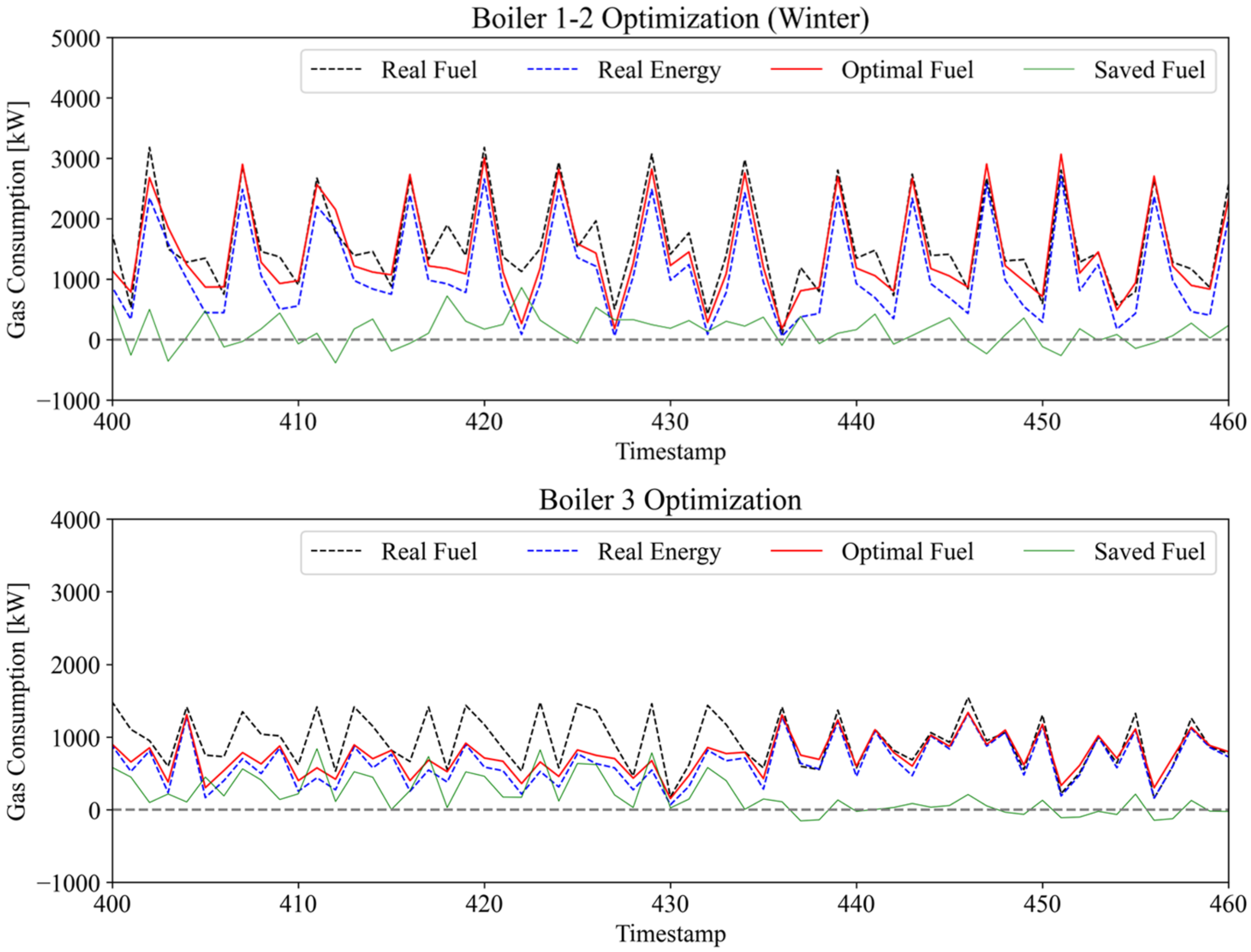


Figure 29: Optimisation results of s3.2.2 on indicative time periods of October (above) and August (below) for the RIO VENA pilot.

VEOLIA - [FASA – CARTIF]: FASA DHN offers a wide set of historical data that is very helpful to analyse the information and provide initial optimisation criteria. In this sense, as also documented in D5.3, years 2018-2019 have been taken as input to analyse the behaviour of the boilers. As observed in Figure 30, the consumed energy (boilers side) is correlated to the external temperature. When climate conditions are improving; then, energy demand is reduced; therefore, consumed energy is decreased, which is mostly happening in the Spring-Summer periods (i.e., summer demand is basically Domestic Hot Water).

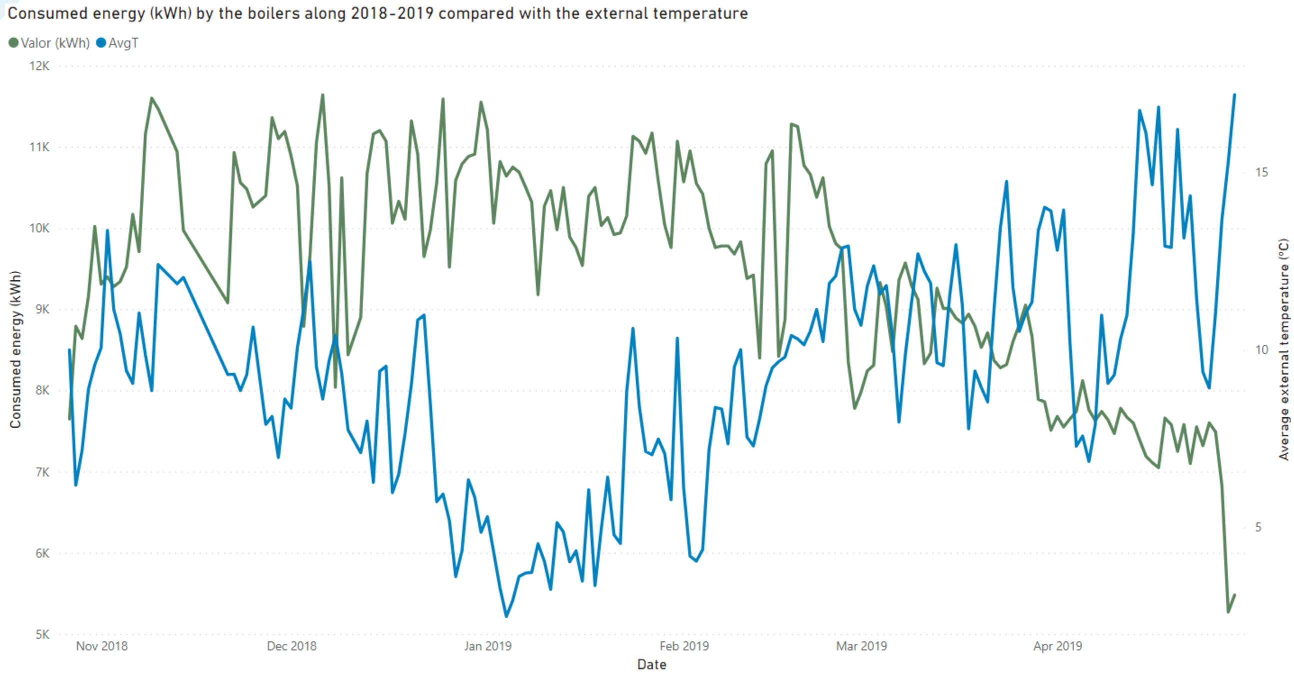


Figure 30: Consumed energy for VEOLIA FASA DHN (s3.2.2)

To better overview the correlation, Figure 31 and Figure 32 show the consumed energy in contrast to the Heating Degree Days (HDD) aggregated per day of the month and month of the year, respectively. It is possible to observe that, even though a relationship exists, there are cases where the HDD (i.e., heating demand) increases, whereas the energy consumption decreases, which requires further investigation.

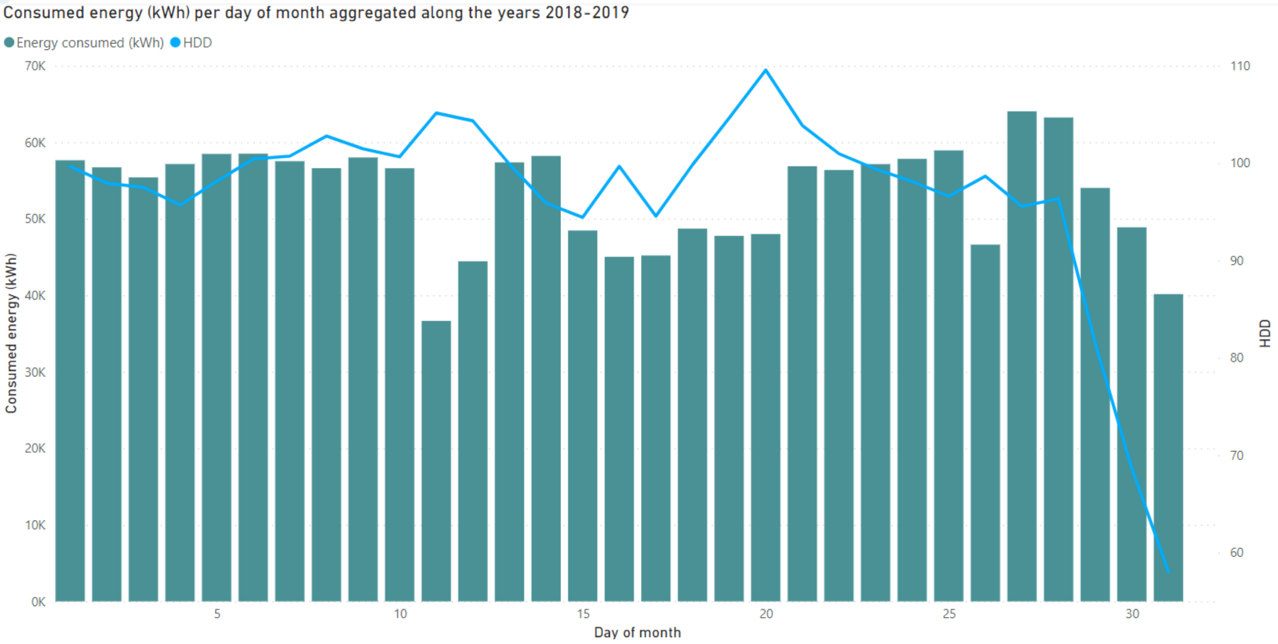


Figure 31: Consumed energy vs HDD in FASA DHN per day of month (s3.2.2)

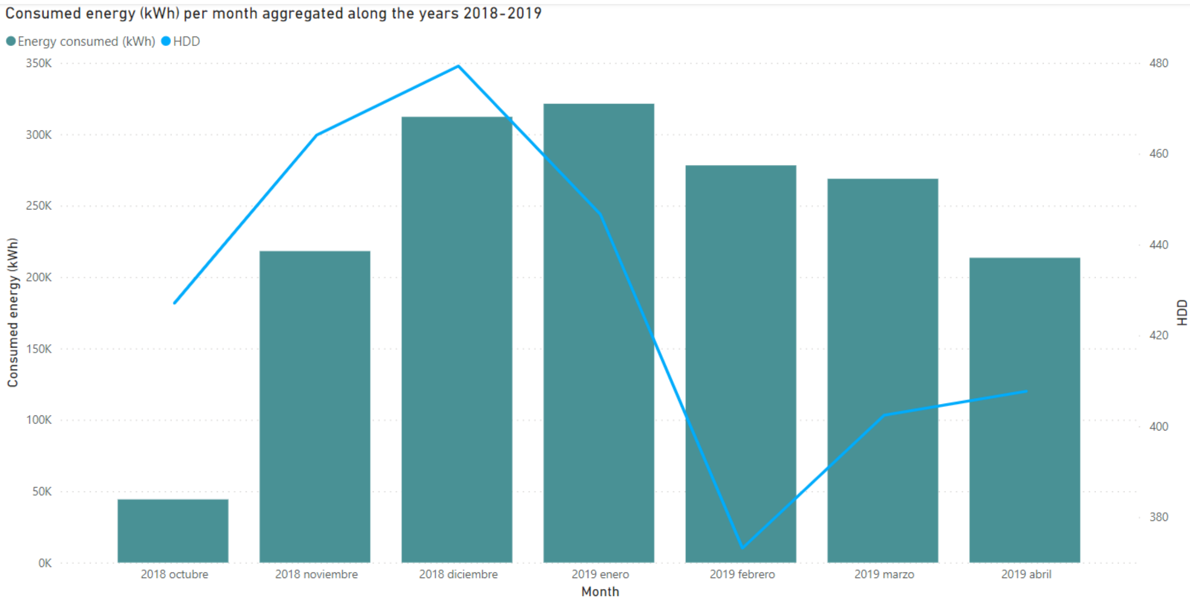


Figure 32: Consumed energy vs HDD in FASA DHN per month (s3.2.2)

According to these values, a correlation analysis has been carried out through linear regression, as depicted in Figure 33. The result of this activity corresponds to equation *(1)*. Statistically speaking, results provide $p < 0.001$ and t-test 93.081, meaning that the 95% of the population has the chance to be within the confidence interval. Hence, independently of the scattering of the points, the linear regression shows a promising value, which needs to be refined (in line with the services models from T3.1).

$$Energy = 0.00036737xHDD + 10.7763 \quad (1)$$

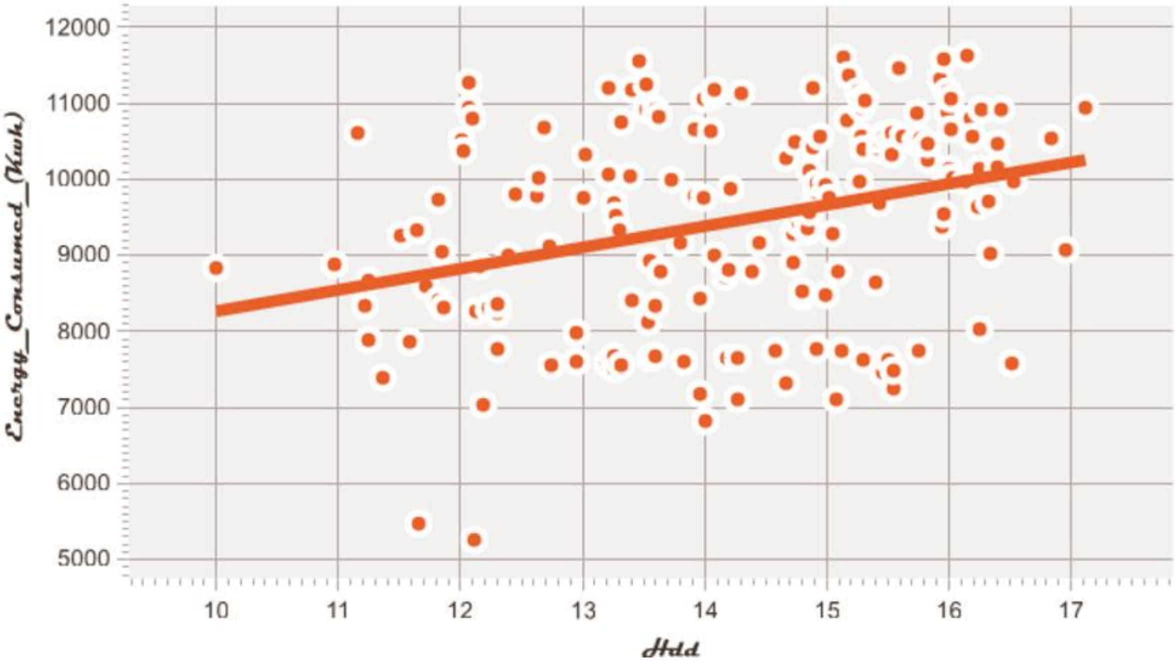


Figure 33: Regression model for the consumed energy in FASA DHN (s3.2.2)

Having the relationship between energy and climate conditions, it can be established the equation *(2)*, where the energy linear regression is equal to the energy calculated through the use of flow and inlet/return temperatures. The inlet temperature set-point is around 87ºC and the constraint of minimum temperature should be 65ºC, thus, the influence of the "constant" inlet temperature is studied.

$$0.00036737\ x\ HDD + 10.7763\ = Flow\ x\ Cp\ x\ (Tin - Tret) \qquad (2)$$

Assuming a constant return temperature (to keep same losses according to the demand) of 65ºC (minimum value, although not rigorous, but it is the one used for the beta version), the linear dependency of the inlet temperature can be obtained as the Figure 34. The preliminary result demonstrates the capability to reduce the constant set-point, even up to 81.43ºC.

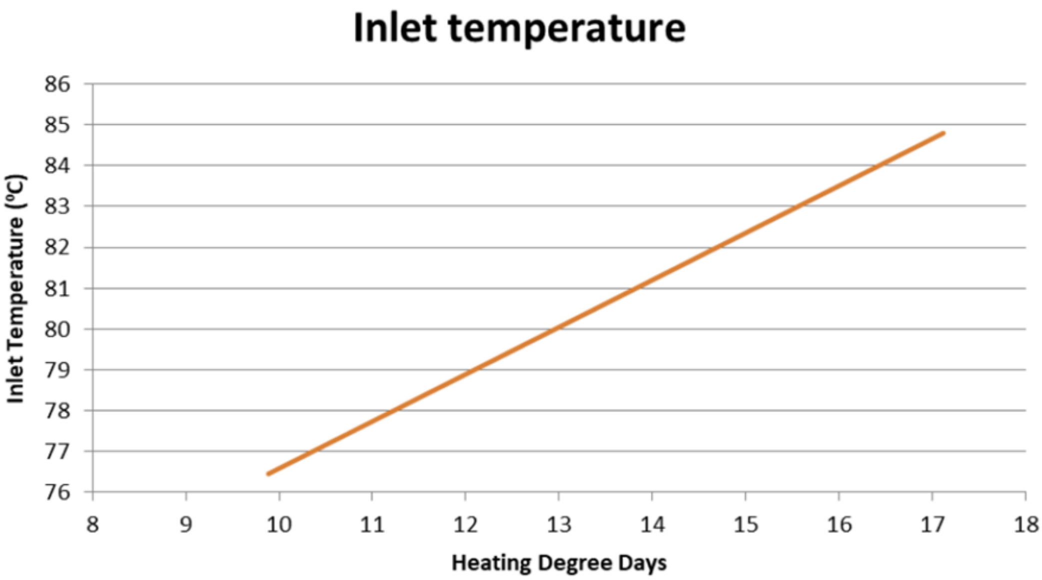


Figure 34: Inlet temperature dependency on HDD for VEOLIA FASA DHN (s3.2.2)

4.2.5 Next Steps

For RIO VENA, the subsequent version of the algorithm will explore new ML models to further enhance the accuracy of the model and reduce potential estimation errors. The precision of the algorithm is anticipated to improve as more data becomes available; at the time of this deliverable's presentation, less than a year's worth of data has been gathered. Furthermore, future research will delve into current best practices to examine the potential for incorporating analyses on how setpoint temperature adjustments impact thermal losses within the DHN. This aspect remains unexplored as substation data are not currently being collected in RIO VENA.

For FASA, the preliminary beta version release looks for optimal set-point of the inlet temperature based on the correlation models. However, further investigation about the correlation between HDD and the consumed energy should be realised to minimise the errors in the model. Moreover, the beta version is focused on the inlet temperature, but the balance among the three boilers would extend the optimisation problem that is being solved within this service and will be part of the final release.

4.3 Power recharging management (s3.2.3)

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, the environmental benefits of EVs depend on the sources of electricity used to charge them. If the electricity is generated from RES, such as solar and wind power, then EVs can help to decarbonise the transport sector. However, if the electricity is generated from fossil fuels, such as coal and natural gas, then EVs can still contribute to GHG emissions.

Moreover, the electrification of mobility poses some challenges for the electric system, as it increases the demand for electricity and requires adequate infrastructure for charging. The integration of EVs with RES can also create some opportunities for balancing the grid, as EVs can act as flexible loads or storage devices that can respond to the fluctuations in supply and demand of electricity. For example, EVs can charge when there is excess production of RES, and discharge when there is a shortage of RES. This can help to smooth out the variability and intermittency of RES and improve their integration into the electric system.

4.3.1 Description of the Service

In the context of DigiBUILD, Service s3.2.3, titled "Power Recharging Management," is designed to assist in the effective management of electric grids and micro-grids through the optimal use of EVs. Specifically, within the Heron Pilot application, this service will facilitate the calculation of a 'green footprint' score for a predetermined future hour within a day. This score will provide critical insights to both the Transmission System Operator (TSO) and end-users, underlining the most environmentally friendly time for EV charging, thereby aiding in emission reduction efforts. Concurrently, in the EMOT Pilot, this service will introduce an advanced optimisation algorithm for end-users. This algorithm will equip facility managers with various charging strategies for the available EVs, tailored to align with user preferences.

4.3.2 Novelty

In the rapidly evolving field of EV technology, a significant focus has been placed on developing sophisticated algorithms for optimizing EV charging. This surge in innovation is driven by the need to integrate EVs into the existing power grids efficiently while minimising environmental impacts. A notable example of this trend is (11) where researchers devised a model to optimise the charging schedules of BEVs. This model, tailored to account for the varying environmental effects of electricity generation, demonstrates that intelligent charging strategies can substantially lower greenhouse gas emissions. It considers factors such as the specific time and location of charging, thereby addressing the dynamic nature of environmental impact. Furthermore, this model also delves into the complexities of balancing various environmental impact categories to minimise the overall ecological footprint of BEVs.

In a different approach, Bao et al. introduced a mixed fleet scheduling methodology specifically for airport ground service vehicles, encompassing both fuel-based and electric tractors (12). Their innovative mixed integer model, complemented by an enhanced adaptive large neighbourhood search algorithm, aims to minimise the cumulative costs associated with time, energy consumption, and emissions. Another significant contribution to this field comes from Yin et al., who presented a strategy to ensure the safe and effective operation of electrical grids through optimal scheduling of EVs (13). Their method utilises advanced machine learning models, including LSTM and XGBoost, combined with the CPLEX solver to address the complexities of optimisation challenges.

The highlighted research emphasises the increasing importance of ML and optimisation algorithms in balancing EV charging demands with the stability of electrical grids. However, there is a gap in applying these algorithms to also address user needs. In the context of this project, combining this service with others, like forecasting models and DT technology, offers a unique approach to solving EV charging issues. Specifically, in the EMOTION pilot, end-users are given a more active role in choosing the optimisation strategy. By inputting their preferences into the algorithm, it's possible to not only improve self-consumption but also cater to the preferences of the facility manager. This method not only makes energy use more efficient but also tailors the experience to better suit individual requirements.

In the field of assessing the carbon intensity of energy used by buildings, extensive research has been undertaken. Notably, an econometric model for evaluating the carbon intensity of buildings was developed by Anshany et al. (14). Additionally, static spatial econometrics and panel co-integration models were employed by Dong et al. (15) to investigate the relationship between regional carbon emission intensity (CEI), urbanisation levels, and energy mix (EM). In DigiBUILD, it was necessary to construct an efficient method for evaluating carbon intensity. This method considers not only house consumption forecasts, but also predicted values of the energy mix in the grid. The aim was to develop a methodology that contributes to optimisation techniques while providing easily understandable results to users, aiding them in informed decision-making to reduce their carbon footprint. A multi-criteria decision analysis approach was adopted, enabling the efficient evaluation of each hour of the upcoming day on a scale from 0 to 1, where 0 is the best and 1 the worst. This scale can be expressed both as a percentage of carbon intensity for user comprehension and incorporated as an indicator in objective functions for further applications, such as optimisation problems, as exemplified in section s3.2.5. Thus, this method allows for the generation of actionable insights from data typically available and often open access in most countries. It enables the provision of information about users' carbon footprints and the integration of these insights into broader optimisation and decision-making frameworks. The dual application of this method underscores its versatility and practical significance in the context of sustainable energy management.

4.3.3 Development Progress

In the development of the Emotion Pilot service, D3.1: 'First wave' of DigiBUILD AI-based data-driven services for the built environment, presented only the general service architecture. This was due to the lack of dynamic data from the pilot until October 23. Consequently, no progress was shown in the alpha release. However, post-release, the development of the algorithm commenced. Initially, after reviewing relevant literature, various algorithms and Python libraries were selected to enhance the service. To kickstart development, synthetic data were created to test these chosen algorithms. The initial strategy involved treating the State of Charge (SoC) of EVs as decision variables, employing Python libraries like pymoo and genetic algorithm for genetic optimisation.

Early versions of the algorithm incorporated basic constraints such as energy balance, SoC limits, and user needs. However, numerous tests revealed that the solutions generated by these algorithms were suboptimal. This issue was primarily due to the decision variable matrix being largely filled with zeros, causing library genetic algorithms (like ES, GA, NSGA-II) to struggle in finding correct solutions, despite significant alterations to evolutionary parameters like crossover or mutation. Alternative optimisation algorithms available in pymoo, such as PSO, Nelder Mead, and Pattern Search, were also explored showing that the last achieved some improvement, but still failed to identify the optimal solution.

A breakthrough was achieved using the PyGAD library for genetic algorithm and finely tuning the evolutionary parameters. This approach produced results closest to the real optimum. However, using SoCs as decision variables imposed severe limitations, as the SoC could only increase due to the pilot's grid-to-vehicle interaction constraint. This restriction significantly narrowed the solution search space.

The final version of the service overcame this challenge by selecting the energy used to charge the EV as the decision variable. This new approach allowed the algorithm to find the optimal solution for each case efficiently. This success confirmed the validity of the approach and opened opportunities to include more complex constraints, such as the need to increase self-consumption and considerations regarding charging station plug power and availability.

In the current phase of development for assessing the carbon footprint of buildings, the methodology has been thoroughly developed and is accessible via a specially designed API. This API facilitates the sharing and distribution of the methodology's outputs. For the specific task of assessing the carbon intensity of buildings, the algorithm was constructed using the VIKOR method. Additionally, a Prefect orchestrator, established at the NTUA premises, enables the scheduled execution and generation of the analysis on a daily basis. This scheduling is based on the predicted values for the upcoming day. To ensure the results are efficiently leveraged and disseminated to subsequent segments of the project, a FastAPI application has been developed. This application is responsible for handling requests and plays a crucial role in the secure sharing of information. It achieves this by integrating with the Identity Access Management framework, Keycloak, which is employed in DigiBUILD project. This approach ensures that the critical data regarding carbon intensity and footprint assessments are not only accurately processed but also securely and efficiently shared within the project's ecosystem, adhering to the necessary standards of data security and accessibility.

The source code can be found on GitHub: GitHub s3.2.3

4.3.4 Application on DigiBUILD Pilots

Following the information available in the deliverable D3.1, since the two pilots where this service must be implemented are strongly different, the results and the details of each application will be clarified better in this section.

- **EMOTION:** Figure 35 presents the updated architecture of the service s3.2.3 for the EMOT Pilot.

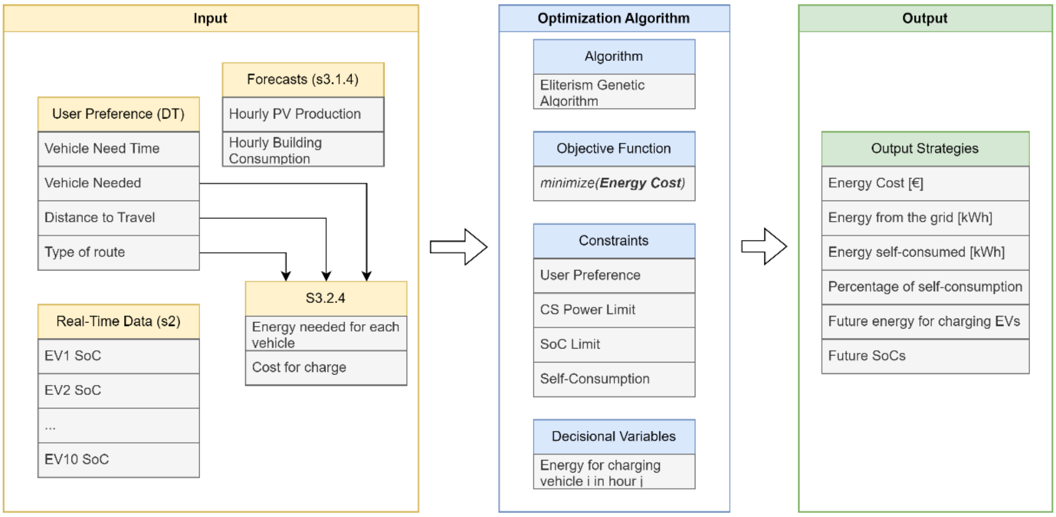


Figure 35: S3.2.3 input/output architecture for the EMOTION pilot.

Since the 70 kW PV plant was installed, the company has faced challenges in synchronizing EV charging with PV production to enhance self-consumption and reduce reliance on grid energy. This service aims to exploit various strategies for effectively charging 10 EVs in the Emotion Pilot fleet. It involves the forecasts of the next day's PV and building energy consumption, estimating the grid energy purchase (E_{buy}) and potential energy sales (E_{sell}). The decisional variable is the energy demanded for charging vehicle j in hour i ($E_{i,j}$). The optimisation algorithm's constraints will be refined once real-time vehicle data and user energy requirements (from s3.2.4) are available. These constraints will consider each EV's SoC, defining the maximum energy capacity per vehicle. Additionally, SoC and user preferences will guide decisions on which vehicle to charge and when *(3)*, while also factoring in each vehicle's battery capacity (C_j) and the charging station's operational limits set for 11 kW for each plug for a total of 44 kW *(4)* considering a limit of two plugs for the available two stations *(5)*.

For each vehicle j:

$$\sum_{i=1}^{h_{desired}} E_{i,j} \geq C_j * (SoC_{1,j} - SoC_{desired}) \qquad (3)$$

Regarding the operational limitations, for each hour i:

$$\sum_{j=1}^{10} E_{i,j} < 44 \tag{4}$$

$$Count.if\ \left(E_{i,j} < 0\right) < 4 \tag{5}$$

Each strategy output of the service focuses on different objectives, such as maximizing self-consumption or ensuring at least 75% usage of potential exported energy ($Perc_{PV}$). The strategy differentiation is exploited by modifying the self-consumption constraint in the problem *(6)*.

$$\sum_{i=1}^{24} \left(\sum_{j=1}^{10} E_{i,j} - E_{sell,i} \right) > (100 - Perc_{PV}) \tag{6}$$

Although equation *(6)* presents an inequality constraint, the objective function in equation *(7)* 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:

$$F(X) = min \left[\sum_{i=1}^{24} \left(\sum_{j=1}^{10} E_{i,j} + E_{buy} - E_{sell} \right) \cdot E_{cost} \right] \tag{7}$$

where E_{cost} is the cost of electric energy which differs if the energy is sold or bought, calculated based on mean cost associated to kWh using available pilot information.

Nevertheless, all strategies share a common goal: minimizing energy costs. A genetic algorithm will be employed to suggest future energy requirements for EV charging, including SoC details. It will also provide key strategy outcomes, such as energy cost savings, free EV charging benefits, PV energy usage percentage, and CO_2 emissions from electricity consumption.

During the 11-day test period from 06/12/23 to 16/12/23, the 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 is considered in the simulation. The algorithm suggested strategies to charge EVs with the available PV-generated energy, as illustrated in Figure 36. Here, the red line represents imported energy, the blue line potential energy export, and the green line the energy used for EV charging. The results confirm that the algorithm effectively guided users to charge EVs using their self-generated power each day.

Funded by
the European Union

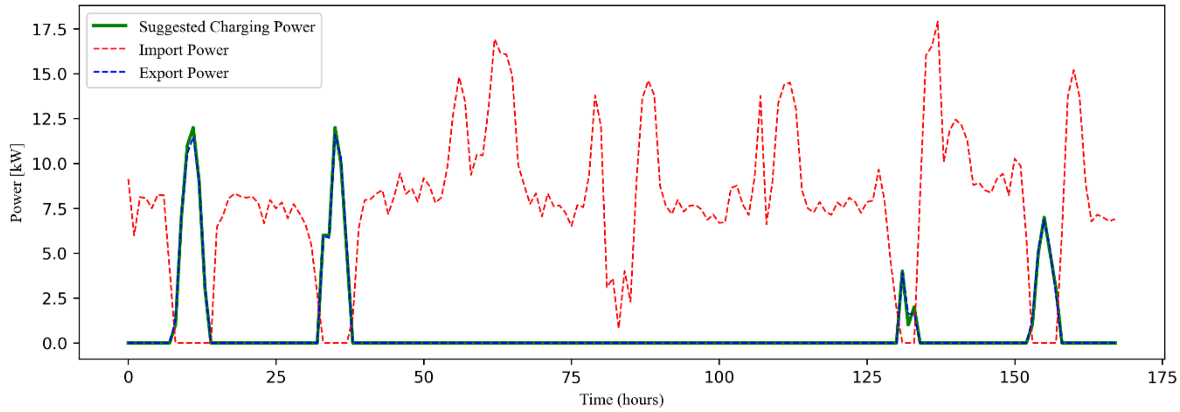


Figure 36: Power trends for charging EVs, as well as power import and export, as computed by s3.2.3 for the EMOTION pilot.

When examining costs over the 11-day testing period, it was observed that costs with optimisation were 1.7% higher compared to scenarios without optimisation, based on the average electricity buying and selling rates in Italy. However, this cost analysis does not include the energy used for EV charging. If we consider also this energy, following optimal strategy, the savings increase to 3.5%. Note also that during the test period, solar radiation was at its lowest. Therefore, there is potential for greater cost savings in future periods with higher solar radiation.

Figure 37 serves as an illustrative example of the dynamic data from a single operational day, showcasing a scenario with overproduction. On this particular day, there is 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, as available energy permits. Consequently, the final SoC for EV4 reaches 66%, exceeding the initial target.

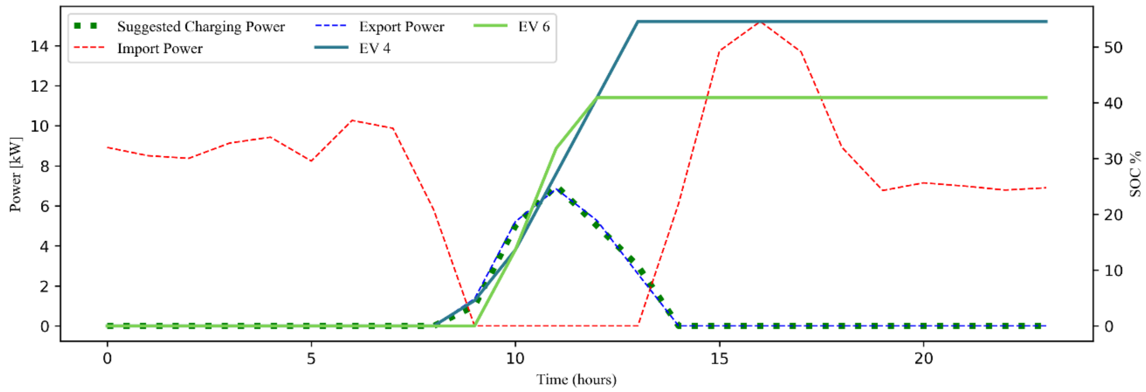


Figure 37: Power trends and SoCs computed by s3.2.3 on 15/12/23 considering EV user preferences, as computed by s3.2.3 for the EMOTION pilot.

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 Figure 38. As expected, in this case only the minimum SoC requirement is satisfied.

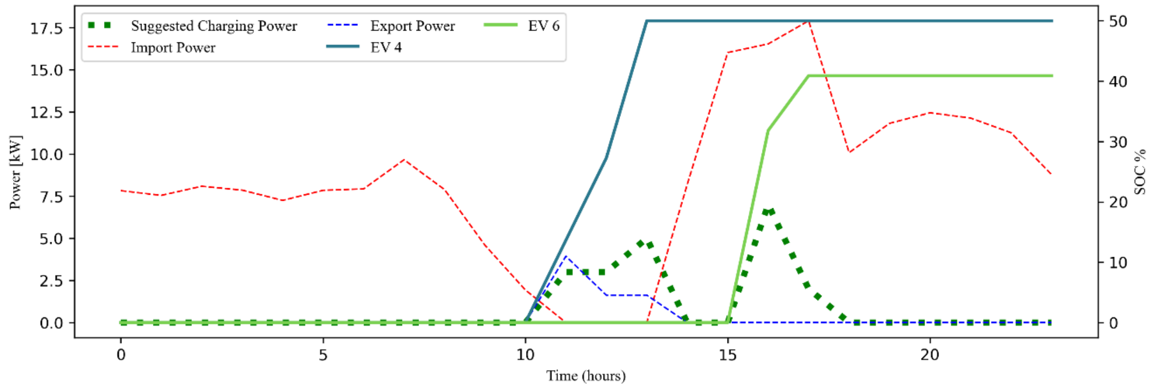


Figure 38: Power trends and SoCs on 14/12/23 considering EV user preferences, as computed by s3.2.3 for the EMOTION pilot.

- **HERON:** In the implementation of the methodology within the HERON pilot for assessing the carbon intensity of buildings, it was essential to gather data from the TSO. In Greece, where the pilot is based, the TSO openly provides this data through an API. The data set includes forecasts of grid consumption, grid losses, and contributions from renewable sources (e.g. PV, Wind, Hydro) for each half-hour of the next day. This data set is then amalgamated with predicted building consumption values, generated by section s3.1.4, for each hour of the following day.

  The process of determining the weights for the multi-criteria analysis involved collaboration with energy experts in the pilot. Utilizing the Analytical Hierarchy Process, the bivariate significance relationships of the criteria were first established (as shown in Table 4), followed by the generation of the weights (depicted in Table 5). These weights are subsequently input into the VIKOR method, which is applied to evaluate the carbon intensity for each hour of the next day. The organisation and scheduling of this entire process are facilitated by the Prefect orchestrator, which provides an efficient and user-friendly environment for orchestration and task execution. Consequently, the results are generated daily for each building and stored in a PostgreSQL database, set up within the task to meet storage requirements.

Table 4: The Analytical Hierarchy Process table presenting the bilateral significance relationships between the criteria considered by s3.2.3 for the HERON pilot.

	RES Production	Hydro Production	Grid Load	Building Consumption	Building tariff	System Losses
RES Production	**1**	4	1	1/9	6	1
Hydro Production	1/4	**1**	1/7	1/9	1/5	1/7
Grid Load	1	7	**1**	1/9	2	1
Building Consumption	9	9	9	**1**	3	9
Building tariff	1/6	5	1/2	1/3	**1**	1/2
System Losses	1	7	1	1/9	2	**1**

Funded by
the European Union

Table 5: Weights calculated via the Analytical Hierarchy Process method of s3.2.3 for the HERON pilot.

RES Production	Hydro Production	Grid Load	Building Consumption	Building tariff	System Losses
0.1318	0.0233	0.1007	0.5718	0.0718	0.1007

To expose the results of this methodology, a FastAPI application is employed. This enables stakeholders, particularly the DT developers, to receive the methodology's results easily, efficiently, and securely for presentation within the pilot's DT. An example of such a request's outcome is illustrated in Figure 39. Through the API, the DT developer is offered two options: to receive results for all buildings in the pilot or to obtain data for each building individually by providing the building's identifier, as defined in DigiBUILD. The output data includes the date (key: 'date') in YYYY-MM-DD format and the payload (key: 'payload'), which contains a list of objects where each object pertains to a separate building. Within the payload, data for each hour of the day include the carbon intensity classification (key: 'co2_class') with classes "Low", "Moderate", and "High", the hour identifier (key: 'index'), the ranking of each hour in terms of carbon intensity compared to other hours (key: 'rank'), the inverted rate generated by the methodology (key: 'inverted rate'), the score for each hour in terms of carbon intensity (key: 'rate'), where 0 is best and 1 is worst, and the reference time in HH:MM format (key: 'time').

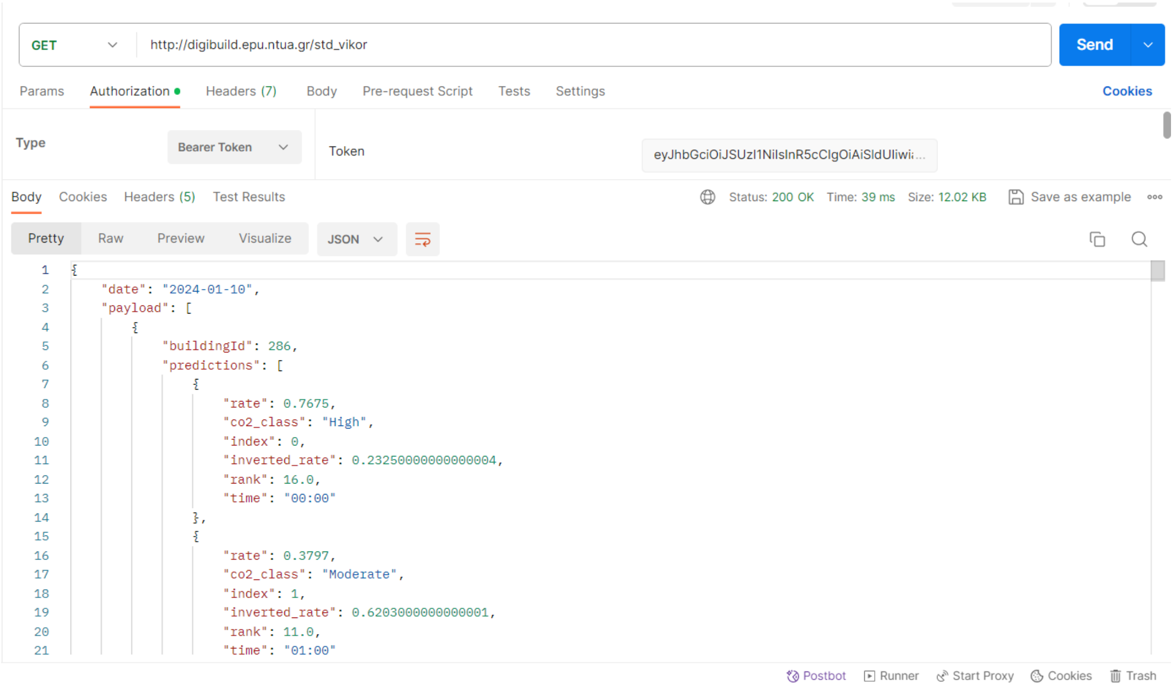


Figure 39: API s3.2.3 response example for the HERON pilot.

Furthermore, Figure 40 displays the API output for a building in a bar chart format. The value of the bars corresponds to the rate achieved by each hour based on the predicted values. Bars are color-coded to reflect their carbon intensity class: "High" in red, "Moderate" in orange, and "Low" in green. This visual representation allows users to easily discern the most favourable times of the day for energy usage, thereby facilitating informed adjustments to their energy consumption patterns.

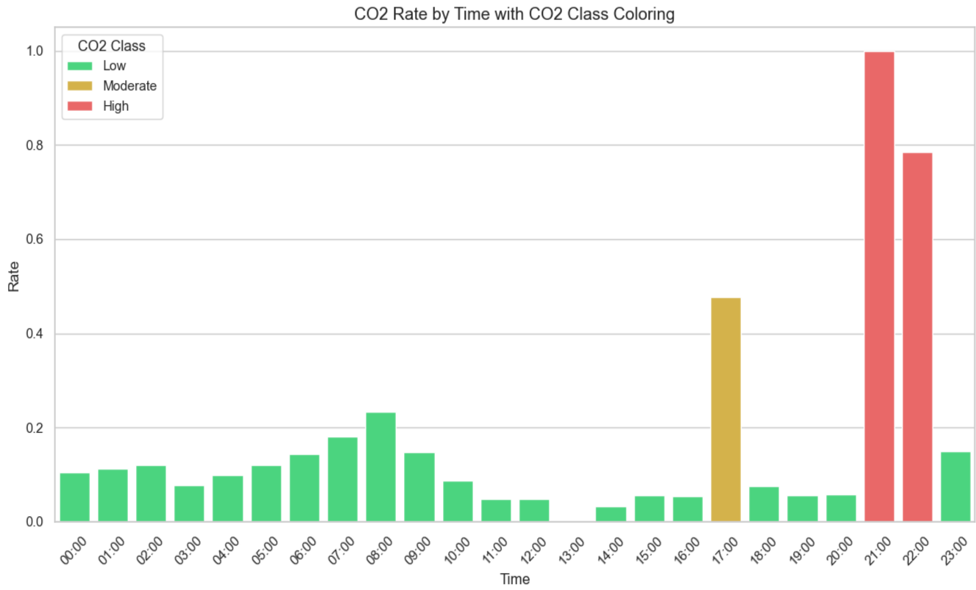


Figure 40: Hourly CO_2 emission rates of building 263, as computed by s3.2.3 for the HERON pilot. The graph illustrates a comparative analysis across time intervals that highlight the environmental impact using CO_2 classifications.

4.3.5 Next Steps

In the forthcoming months of the project, the algorithm is set to undergo refinements for heightened accuracy, leveraging new data collected from the pilots. Currently, the algorithm is based only on data from October '23, but it would benefit from being tested using a more extensive dataset. Furthermore, as the project progresses, there might be changes to the script and algorithms, particularly in their transformation into an API for seamless integration with the DT developed in WP4.

In terms of the EMOTION application, the algorithm will be enhanced to include more precise estimates of charging efficiency. This is particularly crucial as the current data on EVs and charging stations are not synchronised, a status that was evidenced by the varying and mismatched timestamps in the historical data of the vehicles and charging stations.

The application for the HERON pilot has been successfully developed and fully integrated into the pilot's DT for its buildings. An additional feature currently under consideration is the integration of a new methodology capable of analysing historical data from the previous month. This feature would provide valuable insights into past carbon intensity patterns, enhancing the depth and utility of the analysis. A crucial aspect still in development is the establishment of a connection to the semantic ontology of the pilot. This connection is essential for dynamically obtaining building meta-information. The integration of this functionality is not only pivotal for the comprehensive operation of the pilot but also opens avenues for scalable extension. It enables the possibility for the application to be dynamically reused by multiple pilots, significantly broadening the scope and applicability of the service. Once completed, this enhancement will allow for a more holistic and adaptable approach to building energy management, aligning with the evolving needs of various pilot projects. It underscores the commitment to continuous improvement and adaptation in the face of changing requirements and new opportunities in sustainable building management.

4.4 Energy vs e-mobility package (s3.2.4)

4.4.1 Description of the Service

Electric cars are becoming more and more widespread, and their increasing availability is inextricably linked to the possibility of analysing the growing amount of data produced by these vehicles. In this context, this service developed in the DigiBUILD ecosystem is aimed at understanding how to properly estimate EV consumption by starting from analysing different styles of driving through clustering algorithms, with the objective to build a more reliable consumption estimation algorithm that can support s3.2.3 optimisation.

4.4.2 Novelty

As the EU Green Deal and UN sustainability goals prioritise decarbonisation, it is crucial to focus on reducing emissions in the vehicle sector. According to the (16) report, the vehicle sector contributes over 15% to global emissions, and the spread of EVs is vital for reducing GHG emissions. This expansion necessitates dual-sided scientific research; on the first hand to lower emissions during EV charging, and on another to investigate EVs' actual consumption. In the study (17), the authors presented a method for predicting the remaining driving range using XGBoost and LightGBM algorithms based on extensive EV database named NDANEV (18). Although they attained high accuracy, the models included numerous features (e.g. braking ratio, acceleration ratio, motor, and battery output energy) which complicate their replication in simpler scenarios. Alternatively, H.A. Yavasoglu et al. in (19) introduced a range estimation method for EVs using collected data and experiments. They utilised a Decision Tree algorithm to determine road type and a neural network to predict the range based on factors like HVAC usage, vehicle weight, traffic, and road type, achieving over 94% accuracy. While these models are effective and well-crafted, their complexity limits their replication by non-experts and those with less sophisticated equipment.

To overcome these limitations, in the case of the EMOTION pilot, the algorithm was trained using up to two years of EV data to estimate the electric consumption per km travelled. The pilot involves ten EVs, although data from only seven are available. This results in custom and more precise data collection that reflect the actual usage by drivers. A clustering technique, currently K-means, one of the most commonly used methods, is employed to identify clusters based on vehicle speed, which is the sole feature used to estimate consumption. This approach simplifies the model, enhancing its replicability and user-friendliness, even for those without expert knowledge and detailed data. Currently, this parameter can be set by the user as an expected average vehicle speed for future travels, but there is potential for it to be automatically determined in the future.

In the HERON pilot case, a comparatively simpler yet notably effective approach for the user is employed. There is a notable absence of applications capable of providing real-time feedback to users regarding their load usage and carbon footprint. For instance, while applications like Electricity Maps (20) visually represent the carbon intensity for each European Union country, they fail to simultaneously inform the user about their specific load usage. Furthermore, these reports are typically issued daily. Contrastingly, the DigiBUILD methodology analyses the Transmission System Operator (TSO) predictive data hourly, as per section 3.2.3, and furnishes the user with real-time feedback.

4.4.3 Development Progress

Regarding the development that covered the EMOTION pilot, the first step was to analyse the consumption of the EVs by constructing a simple neural network, specifically a MLP that receives input data from the EVs (e.g., odometer, consumption, km travelled, speed). After analysing the different types of data, through an analysis reported in D3.1: "First wave of DigiBUILD AI-based data-driven services for the built environment", an MLP Regressor was constructed with the aim of estimating energy consumption. The pilot specific results were reported in D5.3: "Pilots' execution documentation – Pre-pilot phase". After this, to enhance the performance of the model, clustering algorithms were considered to identify consumption classes based on the estimation of the type of drive. In a first attempt, the covered distance was considered, trying to classify based on distance, with the objective to identify longer/shorter journeys and the respective consumption.

In a most evolved stage of the development, it was decided to continue employing k-means as the clustering algorithm while examining the interplay between various available data. A notable pattern was observed in the clustering of consumption based on vehicle speed, leading to the selection of this factor to influence vehicle consumption predictions. As mentioned earlier, this parameter can be readily estimated based on typical travel speeds or road speed limits, though users also have the option to set it. For the sake of simplicity in this initial phase, it was treated as a constant.

Additionally, data on electricity costs were sourced from records of charging stations spanning 2020 to 2022. Invoice details were also factored in to determine an average electricity cost, which will be incorporated into the service for more accurate cost estimations.

The source code can be found on GitHub: GitHub s3.2.4

4.4.4 Application on DigiBUILD Pilots

Similar to service s3.2.3, this service will be deployed in two pilots: HERON and EMOTION. For the HERON pilot, the service will enable end-users to understand the carbon footprint associated with their real-time charging activities. Meanwhile, in the EMOTION pilot the service is designed to estimate EV consumption, aiding in determining the energy required for traveling specific distances. Notably, this service will play a crucial role in the optimisation process of s3.2.3 by incorporating user preferences. For the EMOTION pilot, this service will not be exposed in the DT, but will act as a component providing information for s3.2.3.

EMOTION: Historical data of EVs provided by the pilot 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 consider this information about vehicle positions, first it was decided to investigate the possibility of using clustering techniques in the latest data. As explained in data analysis of D3.1, 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).

At a latter stage it was investigated the possibility of clustering based on available measurements, speed, consumption, or autonomy variation. However, most of the clustering results were not particularly useful. For instance, studying the relationship between consumption, latitude, longitude, and velocity it emerged that the main clustering reason was the vehicle velocity, as can be seen in Figure 41 which considers the relation between consumption and speed of a Nissan Zoe.

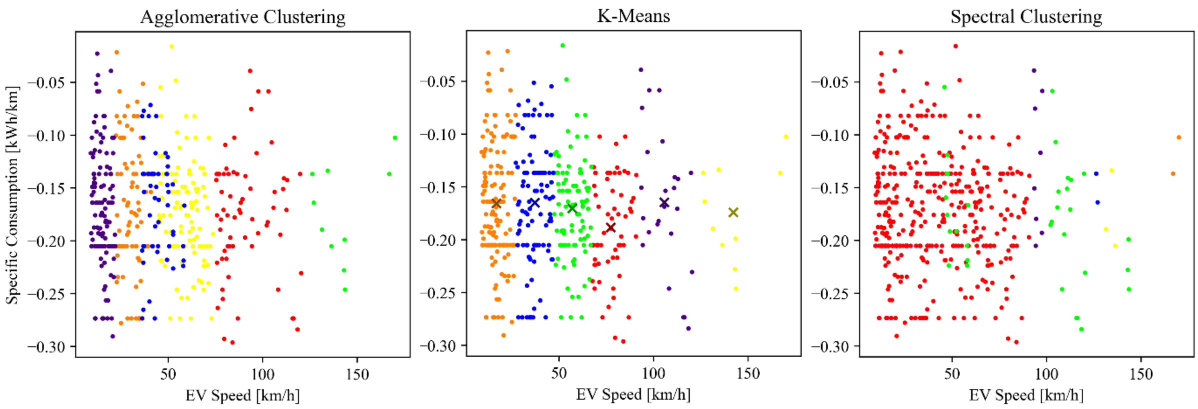


Figure 41: Analysis of different clustering methods

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.

Figure 42 showcases the difficulties present in differentiating clusters for latitude and longitude using the K-means algorithm, rendering the consideration of these two variables for the first evaluation useless.

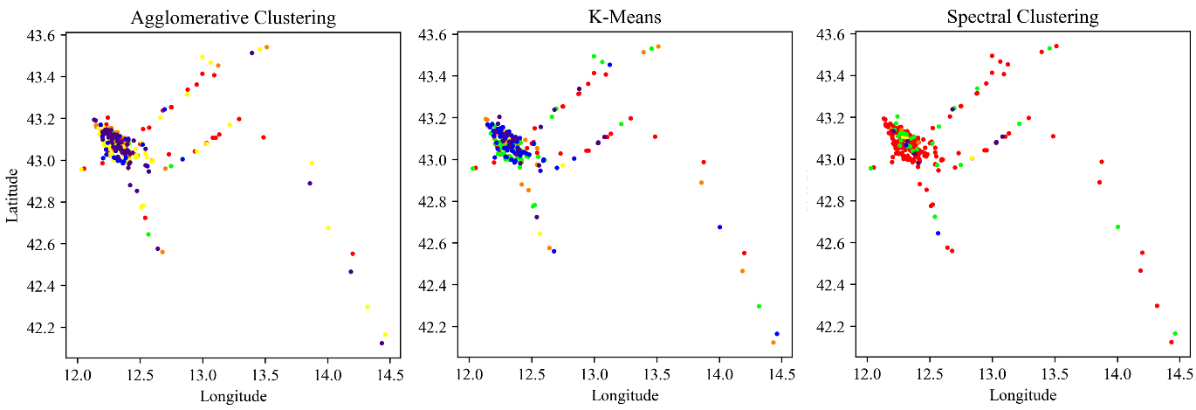


Figure 42: Analysis of different clustering methods considering latitude and longitude information

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 1372 entries post-filtering. Utilizing this data, k-means clustering was selected, and the elbow method was employed to determine the optimal cluster count.

Figure 43 illustrates the elbow method's outcome, including a test with four clusters. To assess the clustering quality, the Silhouette Score (SS) and Davies-Bouldin Index (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.

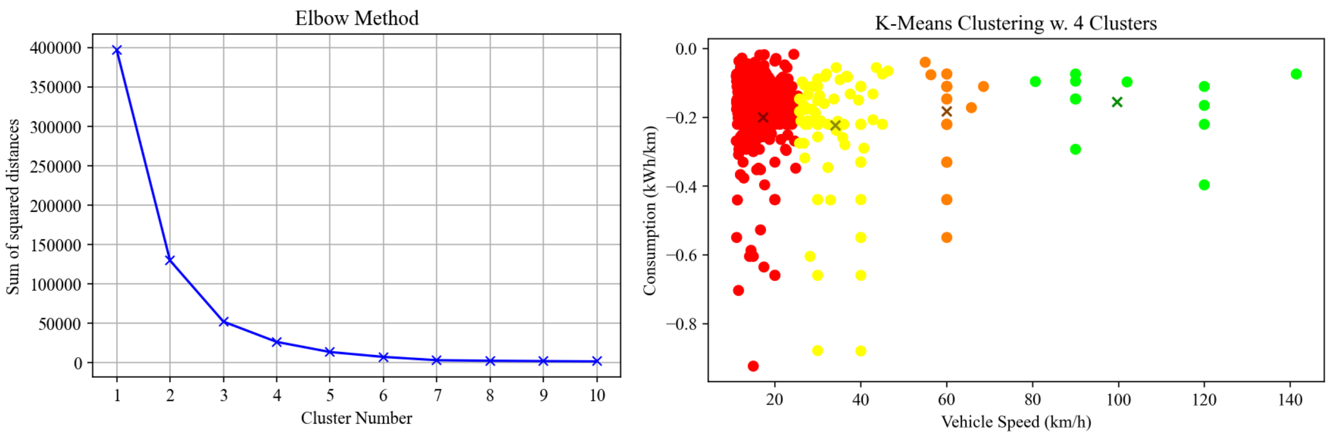


Figure 43: Elbow method application and k-means result considering only 4 clusters

4.4.5 Next Steps

For this service, regarding pilot 5a the following plan is envisioned:

- Refine the clustering method and consequently apply the new findings to the Neural Network model, to enhance consumption estimation.
- Since 1372 values are not enough to develop strong models, the improved algorithm must be tested also on new EV data being collected by the pilot.
- Find new algorithms to improve the accuracy of the models presented.

4.5 Carbon-Free Buildings (s3.2.5)

Minimizing the utilisation of carbon-intensive energy sources remains a paramount objective for the EU, as delineated in the European Green Deal initiative (21). This initiative predominantly aims at achieving carbon neutrality and substantially reducing the carbon footprint associated with energy consumption. In this context, the Carbon-Free Buildings service plays a pivotal role by employing advanced optimisation methodologies to significantly decrease the carbon footprint of entire buildings. This service uniquely addresses both the individual building occupant and the collective user group within an energy community, herein referred to as the 'building cluster'. It aims to refine the energy consumption profile of individual users while concurrently diminishing the overall carbon footprint at the cluster level. This dual approach not only promotes sustainable energy use within individual buildings, but also fosters a more environmentally conscious energy community, aligning with broader European sustainability goals.

4.5.1 Description of the Service

The Carbon-Free Buildings service is dedicated to reducing the carbon intensity of energy consumed by a building cluster through the utilisation of diverse input data. A 'building cluster' refers to a group of buildings selected for the application of this methodology. The cornerstone of this approach is the accurate schedule of flexible loads within these buildings, thereby facilitating a reduction in their carbon footprint.

The assessment of energy consumption is twofold. First, it involves evaluating the potential energy production within the cluster. This evaluation considers both on-site RES, such as rooftop PV systems, and off-site sources, like PV or wind farms, which are virtually mapped to the building cluster. Second, the service assesses energy drawn from the grid, employing a methodology outlined in service s3.2.3 (4.3). This assessment leverages data from the TSO and predictions about the energy mix expected to be present in the grid at any given time.

Subsequently, based on the previously mentioned assessments, as well as predicted demand values of flexible loads and the total predicted energy consumption of each building in the cluster, an optimised schedule for the use of these flexible loads is generated. This schedule is specifically tailored to the user-defined parameters.

In the ensuing sections, we delve into the methodology and framework setup that underpin the utilisation and presentation of the results in the corresponding DT.

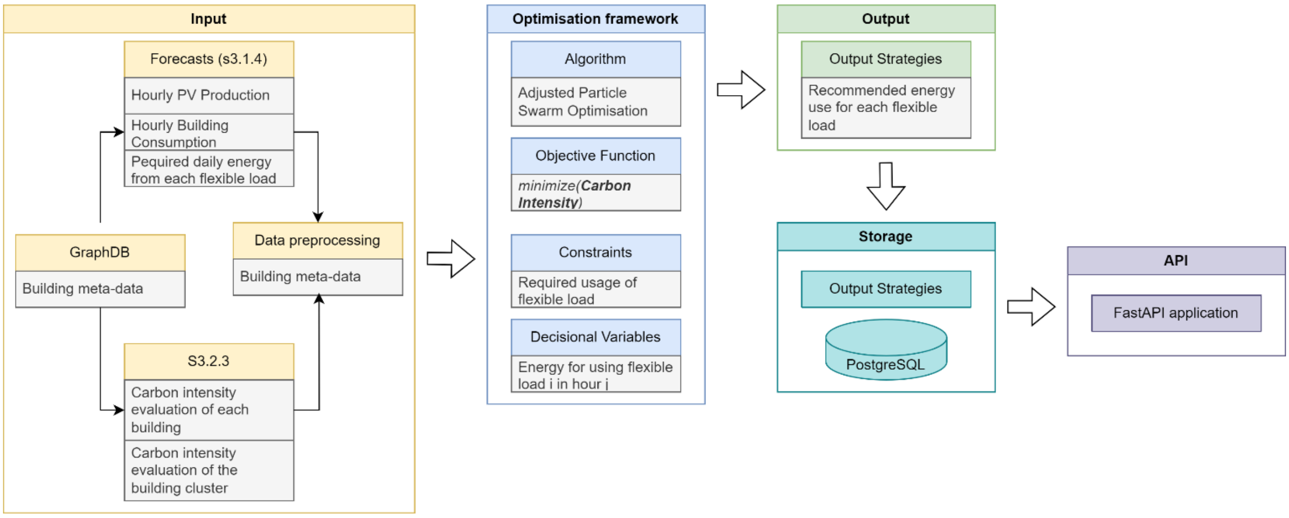


Figure 44: Overview of s3.2.5 architecture.

Figure 44Figure delineates the architecture and information flow across the main modules of the developed service. Data collection is conducted via three primary sources: 1) Metadata acquisition from DigiBUILD's graph database, 2) Forecasting generation and consumption data from section 3.1.4, and 3) Evaluations of the grid's energy mix, as per section 3.2.3. These data undergo a pre-processing pipeline to ensure their compatibility with the optimisation module. Post-optimisation, the data are stored in an intermediate database. To facilitate data retrieval by the DT, a FastAPI application, fully integrated with DigiBUILD's Keycloak security framework, has been structured. This ensures that the results can be securely shared with the respective DT.

The forthcoming section delves into the specific purposes and workflow of each module.

Input Module

The Input Module is the foundational component of the service, tasked with aggregating all necessary data for running the algorithm. The data collection process initiates with GraphQL queries targeting the semantic data of the building cluster. The primary objective here is to delineate the buildings comprising the cluster, along with identifying the building elements and their upgradeable sources. This step is critical as it enables subsequent requests to be issued to requisite services for data re-collection. These services include:

- s3.1.4: This service provides vital energy predictions in three categories: the hourly day-ahead predicted electricity generation from the cluster's renewable energy sources, the day-ahead predicted hourly consumption of the cluster's buildings, and the predicted total daily energy demand of the cluster's flexible loads.
- s3.2.3: It offers data on the carbon intensity of each individual building and the overall cluster.

Post-collection, the data undergoes a process of structuring into necessary data formats, making it suitable for input into the optimisation framework.

Optimisation Framework Module

This module is crucial for determining the optimal use of the flexible loads within the clusters, aiming to diminish the overall carbon intensity. Several optimisation algorithms were initially explored to address this challenge, including the genetic algorithm, NSGA II, and particle swarm optimisation. Due to the high requirements of these algorithms in terms of processing power and runtime, a strategic decision was made to develop an optimised version of the particle swarm optimisation algorithm. This custom algorithm demonstrates a notable reduction in runtime, by up to 30%, compared to standard models. It achieves this efficiency by minimizing the extensive migration of potential solutions, thereby converging on the optimal solution more swiftly.

For the specific problem at hand, we formulated an optimisation problem that leverages the collected data to derive the optimal solution, defined as follows:

$$F(X) = \min[\sum_{j=0}^{23}\sum_{i=0}^{n}\left(\left(C_{j,i} + x_{j,i}\right) * \left(1 + ar_{j,i}\right) + p_{j,i}\right) * \left(1 + bR_j\right) + P_j] \qquad (8)$$

Where:

$x_{j,i}$: The decision variable, denoting the proposed energy consumption from the flexible load of building i at hour j of the following day.

$C_{j,i}$: The total predicted consumption of building i at hour j of the following day.

$r_{j,i}$: The predicted carbon intensity rate of building i at hour j.

$p_{j,i}$: The projected on-site renewable energy production of building i at hour j.

R_j: The predicted carbon intensity rate of the entire building cluster at hour j.

P_j: The predicted off-site renewable energy production at hour j.

a, b: Sensitivity adjustment parameters. They belong to the interval [0, 1] and are used to modulate the impact of the carbon intensity rates ($r_{j,i}$ and R_j) on the overall calculation.

This equation *(8)* aims to minimise the function F(X) which represents the total carbon intensity across all buildings in the cluster for each hour of the next day. The decision variables $x_{j,i}$ are optimised to achieve the lowest possible carbon footprint, taking into account the building's energy consumption, on-site and off-site renewable energy production, and the carbon intensity rates of both the individual buildings and the entire cluster.

With the optimisation framework in place, we generate the decision variables $x_{j,i}$ for each building i at each time j. This results in the optimal recommended energy usage for each flexible load i, tailored to minimise the overall carbon footprint of the building cluster.

Output Module

The Output Module plays a crucial role in the architecture of our service. Its primary responsibility is to efficiently collect the output data from the optimisation framework module. This data, combined with relevant preprocessed information, is then stored in an intermediate PostgreSQL database. The rationale behind this approach is to generate the energy usage schedule on a daily basis, store it, and then make it accessible upon request via the API module. To manage and schedule these processes effectively, we employ a Prefect orchestrator. This tool is instrumental in providing a comprehensive overview of the scheduled tasks and ensuring an efficient and user-friendly interface for both developers and users. Its implementation not only streamlines the workflow but also enhances the overall reliability and accessibility of the service.

API Module

This module fulfils two primary functions. First, it guarantees the secure sharing of the service's output data, achieved through integration with DigiBUILD's Keycloak security framework. Second, it efficiently serves and disseminates the output of section 3.2.5 to stakeholders in a scalable manner, catering specifically to the DTs in the case of DigiBUILD.

4.5.2 Novelty

Load scheduling techniques, such as load shifting and valley filling, along with the utilisation of population and genetic algorithms, have been extensively explored in prior research. Nevertheless, the methodology adopted within the DigiBUILD framework exhibits distinct characteristics, rendering it innovative. Initially, we define the term 'building cluster', a concept that encompasses a variety of buildings, their respective flexible loads, and both on-site and off-site RES. Moreover, while there have been intermittent efforts to integrate energy mix data into building-level energy resource management, DigiBUILD's approach uniquely incorporates a novel method for evaluating the carbon intensity of consumed energy, as delineated in section s3.2.3. This innovation marks the first instance of an optimisation method that employs multi-criteria analysis for the management of flexible loads in buildings. Furthermore, the development of an optimised algorithm, tailored to the specific problem at hand, not only enhances the robustness of our approach but also maximises computational efficiency. Coupled with an API setup that ensures secure information sharing, this service emerges as an innovative asset in the arsenal of facility managers. It significantly aids in informed decision-making for energy management and plays a pivotal role in the reduction of the carbon footprint of facilities.

4.5.3 Development Progress

At this juncture, all modules of the service have been meticulously developed, with the notable exception of integrating metadata through GraphQL queries. Currently, this aspect is provisionally addressed through the use of static information stored locally. The forthcoming incorporation of dynamic metadata querying will significantly enhance the service’s versatility, enabling rapid deployment to any interested stakeholder. While all other modules have been successfully developed to completion, they are currently undergoing a rigorous testing phase. This is to ensure their flawless functionality and to iron out any potential operational issues. The completion of this testing phase is pivotal to the service’s overall efficacy and readiness for practical application.

The source code can be found on GitHub: GitHub s3.2.5

4.5.4 Application on DigiBUILD Pilots

Currently, the service has been applied to the HERON pilot, which encompasses five buildings. In this context, the 'building cluster' is defined as the aggregate of these five buildings, each quantified in terms of total energy consumption. Additionally, the pilot includes an off-site photovoltaic (PV) installation, which is utilised for virtual mapping to supply renewable energy to the building cluster. A notable feature of each building within the pilot is the presence of an electric vehicle (EV) charger. EV chargers, considered significant loads, are integral to the pilot's strategy for minimizing the carbon footprint of the cluster. Observations of the operating patterns of these chargers indicate a randomised usage, primarily driven by the energy needs of the users. Consequently, the tool developed herein is designed to assist users in making informed decisions regarding the charging of their electric vehicles, ultimately aiming to reduce the cluster’s carbon intensity. For the purpose of service testing, two buildings, each equipped with an 11 kW EV charger, were selected. These buildings are also incorporated into the methodologies outlined in sections s3.1.4 and s3.2.3, allowing access to predicted data on their consumption, PV output, and carbon intensity. Application of the service to these buildings demonstrated that the algorithm effectively manages to align charging times with periods of increased energy production, where available. Additionally, when the PV production does not fully meet the cluster's energy needs, the algorithm opts to draw power from the grid during times of lower carbon intensity. Figure illustrates a sample response from the API call. To access some of the API functionalities, users must provide a bearer token received upon signing in to DigiBUILD's Identity Access Management. This token is verified by the API, which then returns the requested data to users with the appropriate permissions. The developed API allows for the retrieval of data both at the overall building level and for specific buildings, given the building_id as a parameter (schema). In the output, users receive information such as the date to which the time program refers (key: 'date'), along with details for each building. These include the building id, the predicted daily energy requirement from the EV charger (key: 'predictedEnergy'), and the charging schedule (key: 'plan'). The 'plan' specifically outlines the index of the reference time (key: 'index'), ranging from 0 to 23, the reference time in 'HH-MM' format (key: 'hour'), the carbon intensity class of the building's energy (key: 'co2_class'), the proposed usage energy from the charger in kW (key: 'energy'), and the percentage of charger usage based on the charger's rated power (key: 'percUsage').

Figure 46 presents an illustrative example of the API's output on a day when the predicted energy requirement for charging was notably high, specifically at 33kWh. The displayed bar chart in the figure graphically represents the service's recommended usage of the charger, quantified in terms of energy in kWh. From the visualisation, we can discern the algorithm's strategic preference for scheduling charger usage. This is particularly evident in the barplot, which aligns with periods of increased energy production. Moreover, the chart highlights that the energy class utilised during these periods is of a low carbon intensity. This aspect is crucial, as it reflects the service's effectiveness, not only in optimising energy usage, but also in ensuring that this usage aligns with environmentally sustainable practices. The bar chart thus serves as a clear demonstration of how the service can manage and optimise energy consumption in real-world scenarios, catering to high-demand situations while maintaining a focus on reducing the carbon footprint.

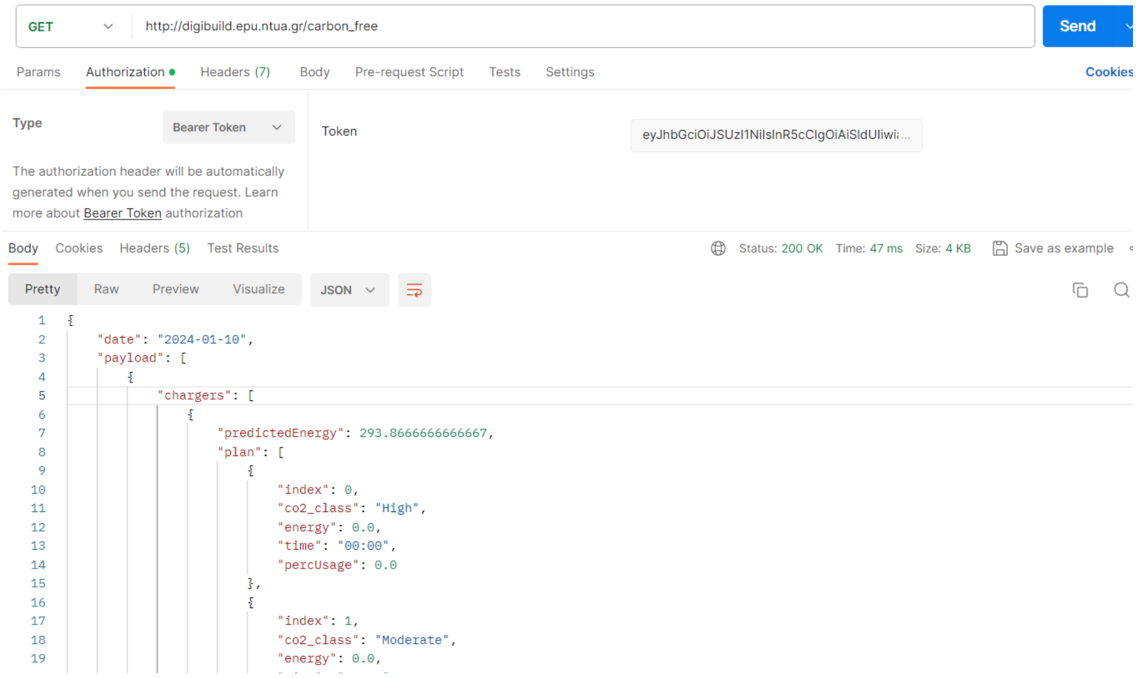


Figure 45: Carbon-free buildings API response example.

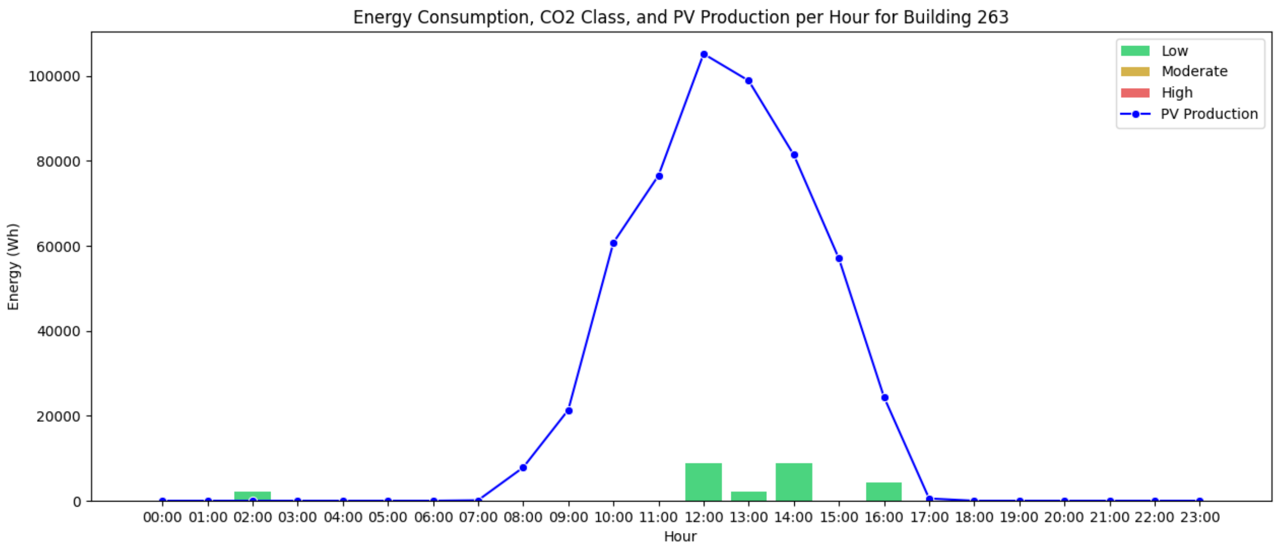


Figure 46: Example of the recommended energy usage for building with id 263, as computed by s3.2.5.

4.5.5 Next Steps

As the subsequent phases of development progress, our primary focus will be on the extensive testing of the service. This testing phase is pivotal for the optimal calibration of the optimisation method, ensuring that it operates with the highest level of efficiency and accuracy. Following this rigorous testing, the implementation of the service across all the buildings within the cluster follows. Currently, the two buildings that have been the subject of our initial tests are already integrated within the DT of the pilot project. The next significant step involves the full integration of the service with the complete building cluster in the DT. This integration will not only expand the scope of our service but also enhance its operational relevance and effectiveness. Throughout this development process, actively seeking and incorporation of feedback from stakeholders is mandatory. This collaborative approach is fundamental to refining the service, as it allows us to understand the practical needs and challenges faced by users. Stakeholder feedback will provide invaluable insights into fine-tuning the service to better meet the requirements of facility managers and other end-users. In summary, these next steps represent a crucial phase in the evolution of the service, though the transition from testing and refinement to broader application.

4.6 Optimal electric or thermal load management (s3.2.6)

The efficient management of energy systems is crucial for reducing consumption and emissions without necessitating significant investments in new equipment or advanced HVAC systems. Given that building operations account for a substantial portion of global energy use and emissions, implementing smart energy management strategies becomes imperative. DigiBUILD, through its s3.2.6-optimal electric or thermal load management, addresses this need by empowering facility managers with advanced tools and insights. This service facilitates the optimal management of electric and thermal loads, enabling a more efficient use of energy resources. By its application in the EMOTION pilot, managers can not only reduce operational costs but also contribute to a significant reduction in environmental impact. This service aligns with contemporary needs for sustainable energy solutions, offering a practical and impactful approach to energy management in buildings.

4.6.1 Description of the Service

The electric or thermal load management service effectively integrates real-time data analysis to manage energy more efficiently at Emotion facilities. This facility has a fleet of 10 EVs and a major energy-consuming device, a 40-kW heat pump. This service aims to address the limitations highlighted in section s3.2.3 by generating an optimal EV charging schedule derived from the predictions in s3.1.4. It is important to note, however, that these predictions might not always be accurate. Therefore, the heat pump's energy use is not considered in the service s3.2.3 which only focuses on optimizing EVs charge. The primary objective of this new service is to offer real-time, adaptable recommendations for EV charging that take into account live data and actual usage patterns.

In this context, the service processes real-time data, including the power output of charging columns, the SoC of the EVs, as well as the real-time energy consumption of the PV system and the building. This approach allows for a comprehensive understanding of the current energy dynamics. Based on the balance between energy production and consumption, and in cases where predictions diverge from actual usage, the service dynamically adjusts the power allocated for EV charging.

Moreover, when the optimal goal of maximizing self-consumption is constrained by plug limitations, this service functions as a decision support system within the DT framework. 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 intelligent approach to energy management is supported by scientific literature, which emphasises the importance of real-time data and adaptive control in optimizing energy use in buildings and facilities. Studies have shown that such systems can significantly reduce energy costs and carbon emissions by optimizing energy consumption patterns, particularly in facilities with renewable energy sources and electric vehicle fleets. This adaptability not only ensures energy efficiency but also contributes to the grid's stability, making it a valuable addition to modern energy management strategies.

4.6.2 Novelty

The primary innovation of this service lies in its ability to manage various energy flows (including electric vehicle charging, grid interaction, and heat pump operations) in real-time, with the aim of reducing energy drawn from the grid. There are numerous instances in scholarly literature where similar approaches have been explored. For example, the work of Mathur et al. (22) and Uzair et al. (23) are relevant examples of real-time operational optimisation. This service plays a crucial role in the DT framework, particularly in determining optimal strategies for real-time EV charging and providing intelligent recommendations to managers. Applying this service in a practical case study allows for the exploration and validation of these algorithms, which are used in simulations or rely on historical data in existing literature. In addition, its innovation lies in its dynamic adaptability to actual usage patterns and live data, offering real-time, adaptable EV charging modulation. Therefore, the service's ability to adjust the power allocated for EV charging based on energy production-consumption balance and divergences from predictions showcases an advanced level of automation. This approach not only ensures energy efficiency but also contributes to grid stability, resonating with the focus on real-time data and adaptive control in optimizing energy use in facilities with renewable energy sources and electric vehicle fleets, as highlighted in current scientific literature (24).

4.6.3 Development Progress

Owing to the technical challenges highlighted in earlier deliverables, such as D3.1 and D5.3, progress towards real-world application was hindered. Specifically, the lack of historical data posed a significant barrier, leading to the creation of a theoretical algorithm that could not be empirically tested in a real-world context.

However, with the eventual availability of data, a comprehensive development phase commenced. The previously conceptualised algorithm was programmed in Python, with clear definitions for inputs and outputs established. Currently, a more streamlined version of this algorithm is under development. This version, by analysing energy exchanges with the grid, is designed to calculate new modulation power levels for each connected electric vehicle and to uncover additional strategic approaches.

4.6.4 Application on DigiBUILD Pilots

The service under discussion, slated for implementation exclusively in the EMOTION pilot, will be explored here with a focus on its practical application in this pilot context. The central equations and operational principles of the algorithm are detailed to illustrate the capabilities of this preliminary version of the service.

The algorithm functions consider the SoC 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 capacity of the plugs (Cap_{plug}), defined as *(9)*, the algorithm recalculates the power assigning the energy in each plug.

$$Cap_{plug} = \sum_{i=1}^{EVs\ Connected} (11 - E_{i,real-time}) \tag{9}$$

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 *(10)*. Finally, the service suggests adjustments to the heat pump's setpoint temperature, depending on the time period.

$$Cap_{EV} = (100 - SoC_{real-time}) \cdot Cap_{battery} \quad (10)$$

In situations of excessive power import, the algorithm initially verifies whether the vehicles connected have reached the SoC demanded by the user, as calculated by s3.2.4 via the DT. Subsequently, the new plug capacity is determined using equation *(11)* which is specifically tailored to reduce power usage, contrasting with equation *(9)*, in which E_{min} is the minimum charging power that a charging station could reach.

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

If the plug capacity is inadequate, particularly during high heating or cooling demands that cannot be covered by PV generation, the system suggests adjusting the setpoint temperature of the heat pump. Additionally, it recommends ensuring that such a modification does not result in discomfort.

The source code can be found on GitHub: GitHub s3.2.6

4.6.5 Next Steps

In this preliminary form, the algorithm successfully meets the pilot's goal of real-time management of thermal and electrical loads at the charging stations. Future enhancements could refine its optimisation capabilities. For example, instead of modulating electric vehicle charging uniquely based on the achievement of the desired SoC, the algorithm could also consider whether the SoC can be attained within the required timeframe even with adjusted modulation levels. Additionally, a more precise selection of setpoint temperatures for the heat pump could be explored as an improvement once more data is gathered and analysed. This approach offers a promising avenue for advancing the algorithm's functionality and effectiveness in managing energy loads.

5 Data-driven energy and non-energy services for enhanced comfort and people well-being

Quantifying comfort has become increasingly significant given the substantial time people spend in indoor spaces during their daily routines, such as in offices, homes, and schools. Furthermore, the impact of indoor environmental conditions on human well-being and work productivity emphasises the importance of assessing and ensuring comfort. The need for comfort maintenance and monitoring is elevated by the potential advancements in sustainability and energy efficiency. Creating a comfortable indoor environment involves aligning energy consumption with user needs, thereby reducing energy costs. As a result, buildings with effective insulation and efficient energy systems, including lighting and HVAC systems, can not only minimise energy usage but also foster a conducive environment for occupant well-being.

Within this context, the DigiBUILD project introduces two services: Enhanced Comfort and Wellbeing (s3.3.1) and the Comfort Performance Contract (s3.3.2). Based on the outcomes of Task 1.3 activities outlined in deliverable D1.2, the end-users poised to benefit from these services include facility managers, sustainability stakeholders, and policy makers. This is because the maintenance and monitoring of comfort contribute to improving buildings in terms of energy consumption and overall costs. Subsequent paragraphs will provide a detailed and technical description of the methodology that has been idealised and employed these services implementation.

5.1 Enhanced comfort and well-being (s3.3.1)

5.1.1 Description of the Service

The objective of the service is to ensure comfort in indoor environments by harmonizing factors related to Indoor Environmental Quality (IEQ), such as indoor air quality (IAQ), thermal comfort, visual, acoustic comfort with considerations of energy consumption and costs. This is achieved through the analysis of data collected from sensors installed in various buildings. The evaluation of comfort and well-being is facilitated by the development and application of AI-based algorithms designed to quantify and predict comfort-related elements, such as the Thermal Sensation Vote (TSV), utilizing building data. The accuracy of these predictions relies on both historical and real-time data acquired from the sensor network in the pilot buildings. This AI-driven approach aids various stakeholders in efficiently and sustainably managing buildings while ensuring a specified level of comfort for occupants. Beyond maintaining acceptable comfort levels, the service also targets the reduction of energy consumption and associated costs.

The service will deploy a coaching solution, empowering end-users to effectively manage indoor building conditions. The comfort predictions generated by the AI model serve as feedback for end-users, allowing them to adjust indoor environmental conditions and restore optimal comfort. Additionally, within this service framework, direct feedback on indoor environmental and comfort conditions is gathered from end-users using survey-based approaches. This information is utilised in implementing the AI model to generate comfort predictions aligned with occupants' preferences.

As described in deliverable D3.1 "First wave of DigiBUILD AI-based data-driven services for the built environment" (M12), the s3.3.1 service is implemented by considering the following phases (the reader can find their detailed description in D3.1):

- **Phase 0**: *Assessment model for Enhanced Comfort and Wellbeing.*

Funded by the European Union

- **Phase 1**: *Integration of the data inside the service Data Lake—Interoperability and Quality.*
- **Phase 2**: *Definition of the technical specifications of the service.*
- **Phase 3**: *Baseline monitoring. Track energy consumption (T3.1) and track comfort parameters.*
- **Phase 4**: *Profiling and individuation of anomalies*.
- **Phase 5**: *Delivery of coaching solution for enhanced comfort and wellbeing.*

Following the submission of D3.1 (M12), a more detailed and technical definition of the comfort service was undertaken. This involved providing a technical overview of the components essential for the development of the service. UNIVPM specifically contributed by defining the software components and visualisation tools required for the practical application of s3.3.1 in pilot scenarios. In terms of software components, two comfort models were taken into consideration: the simplified Predicted Mean Vote (sPMV) and the adaptive comfort model. The reader will appreciate the developments for the sPMV model in the next paragraphs, while the adaptive model will be computed in the upcoming period, as reported in the paragraph 5.1.5.

The sPMV and adaptive models utilise environmental parameters such as indoor temperatures, indoor relative humidity, CO_2 levels, and user feedback as inputs. UNIVPM opted to implement both machine learning (ML) and artificial intelligence (AI) algorithms. Baseline ML models, including Random Forest, Naïve Bayes, K-Nearest Neighbour, Adaboost, and Bagging algorithms, were incorporated for comfort classification. Additionally, the Long Short-Term Memory (LSTM) Network was chosen to predict comfort by leveraging historical data from pilots considering different forecasting horizons. Moreover, within this context, the ML and AI approaches will be also exploited to predict the feedback of the users in terms of Thermal Sensation Vote (TSV). Another software component within the s3.3.1 architecture focuses on providing a model to forecast energy consumption based on comfort maintenance.

Parallel to these software blocks, the s3.3.1 implementation involves the creation of two types of visualisation tools: a pop-up notification system and a dashboard. The pop-up notification system serves a dual purpose, collecting user feedback on their thermal sensation and offering recommendations to end-users for personalised actions to restore or maintain the identified comfort condition. The dashboard, on the other hand, is integrated into the s3.3.1 structure to provide an overview of indoor environmental conditions and energy consumption, considering the outputs of the implemented AI/ML-based models.

Within the context of s3.3.1 deployment, the UNIVPM team has conceptualised the service architecture by highlighting each connection between the included elements mentioned before. The reader can appreciate three coloured boxes in Figure 47:

- The pink box, denoted as 'A' in Figure 48, embodies the two comfort models integrated into the framework of s3.3.1. Specifically, Figure 48 illustrates that these two comfort models draw input from the DigiBUILD repository, identified as "Storage" in the figure. The input data requested encompass environmental parameters (e.g., indoor temperature, CO2, etc.) gathered through sensors installed at the pilots' sites. Additionally, users' feedback will also be sourced from this repository. For a more detailed view of this block, the reader can refer to Figure 38, where a zoomed-in image elucidates the implementation details of the two comfort models, namely the simplified PMV (sPMV) and adaptive models. They have been found in literature (25), (26) and are based on the comfort standard ASHRAE 55. The sPMV model (25) has been selected since it simplifies the quantification of comfort with respect to the standard PMV, making it more applicable in real-life contexts by relying on only two environmental parameters (such as indoor temperature and relative humidity). Furthermore, it has been already described in deliverable D5.3. As regards the adaptive model, it is based on the computation of an upper ($\theta_i\ max$) and a lower ($\theta_i\ min$)

comfort limit expressed as temperatures (26). Equation *(12)* illustrates the calculation of these limits, emphasizing their dependence on the mean outdoor temperature (θ_{rm}). The adaptive model has been selected since it provides a range of temperatures (Equation 12) within which optimal thermal comfort for occupants is guaranteed. Furthermore, as reported in Figure 47, the implementation of this model will include also the user feedback in order to incorporate the occupant's experience in the building and observe its influence on the comfort measurement. As previously mentioned, the adaptive model is planned to be developed in the next months.

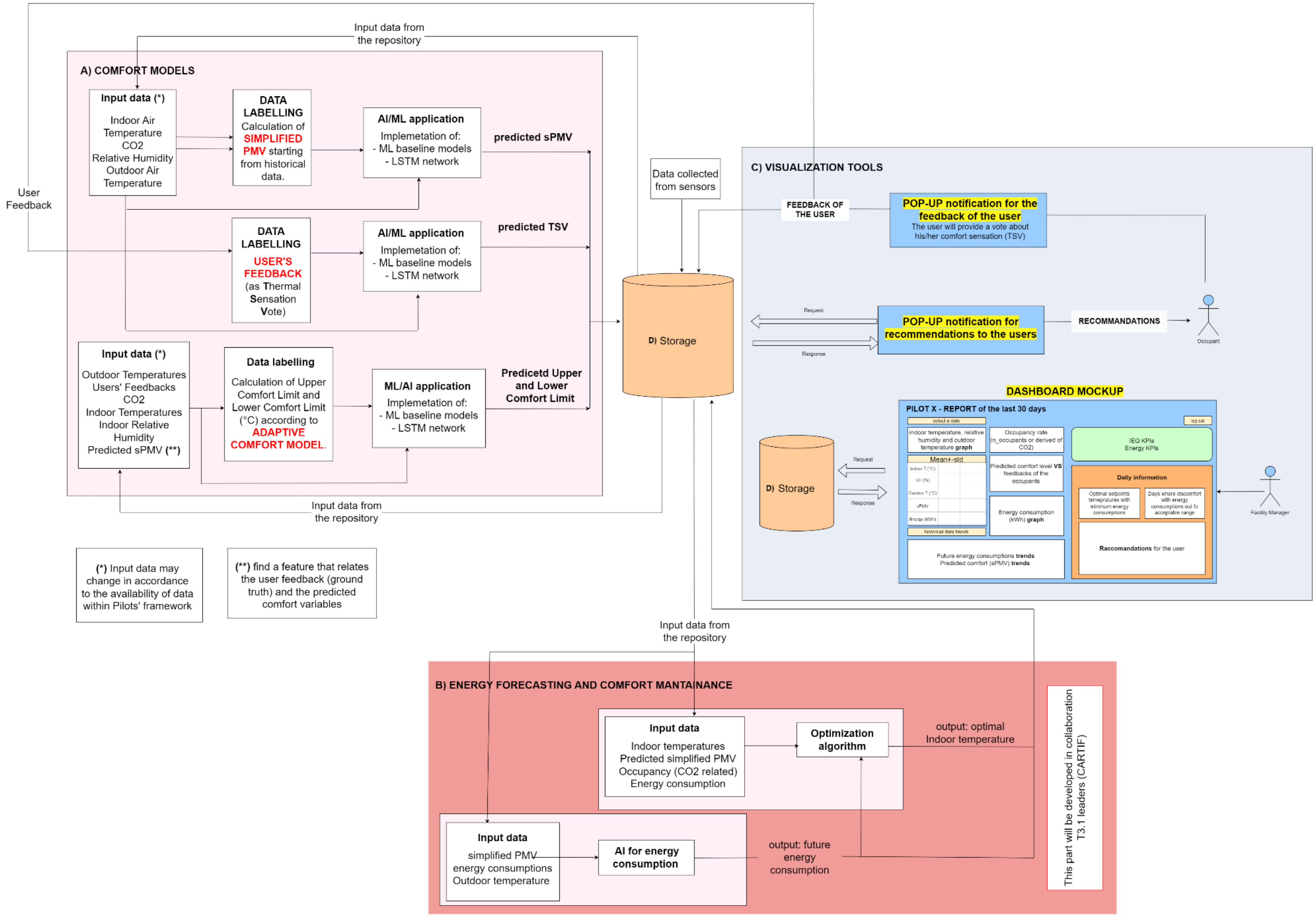


Figure 47: Block scheme of the whole service architecture.

$$\theta_i \max = 0.33 * \theta_{rm} + 18.8 + 4$$
$$\theta_i \min = 0.33 * \theta_{rm} + 18.8 - 4 \qquad (12)$$

The last software component included in this block of the service architecture (Figure 47) aims at predicting the users' feedback, considering the feedback collected with the pop-up notification system () mentioned in the introductory part. The objective of this element is to forecast the subjective thermal sensation of occupants and examine its potential correlation with the indoor comfort conditions. This aspect will be technically developed in the next months.

Figure 48 emphasises that the implementation of both models involves two distinct steps: data labelling and the application of ML baseline models and an LSTM network. In data labelling, environmental parameters serve as inputs. For the sPMV model, the label is the simplified PMV, while for the adaptive model, the temperatures outlined in Equation *(12)* serve as the labels. The results of this operation and the other parameters retrieved from the repository (e.g., CO_2, indoor Temperature, Outdoor temperature, users' feedback, etc.) contribute to the computation of ML and AI models illustrated in Figure 48. The generated outcomes are stored in the project repository, as indicated in Figure 48.

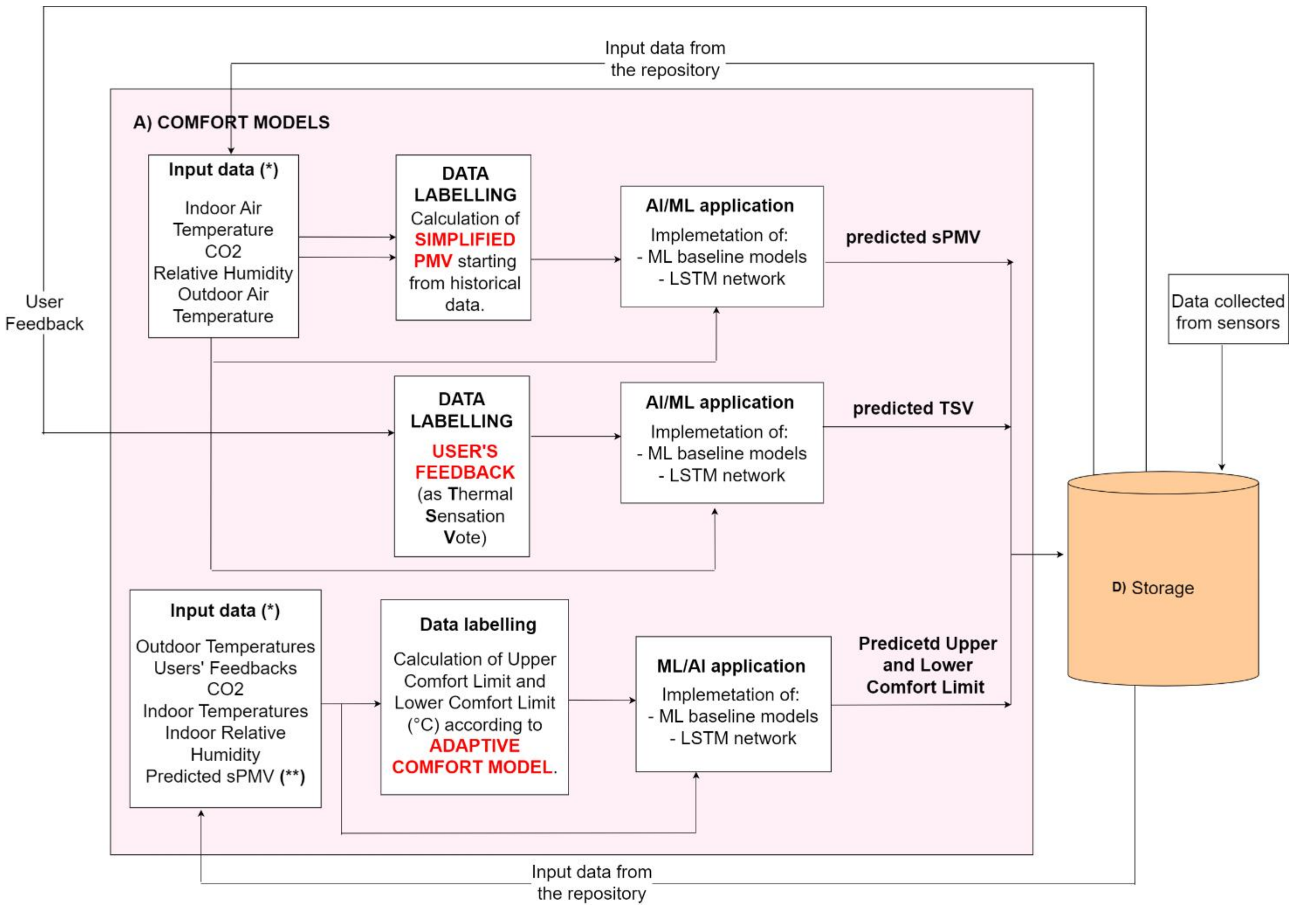


Figure 48: Section of s3.3.1 architecture dedicated to comfort models.

- In Figure 47, it is possible to note also a blue box. It includes the visualisation tools included in the technical architecture of the service. The first visualisation tool is a pop-up notification system which retrieves the users' feedback about their thermal sensation but also provides recommendations to maintain or improve the indoor comfort condition. The reader can appreciate from Figure 47 that the users' feedback will be stored in the "Storage", while the recommendations will be provided in accordance with the data collected in the project repository. Furthermore, as it is possible to note from Figure 37, the user feedback collected with this methodology will be exploited as label in the comfort models software block (Figure 48). Concerning the second visualisation tool, Figure 49Figure 51 presents a mock-up of the dashboard where end-users can consult graphs, charts, and tables depicting predicted comfort levels and the environmental conditions of the building. Data for populating the dashboard will be sourced from the repository. The reader can also appreciate from Figure 49 that recommendations are also intended to be visualised.

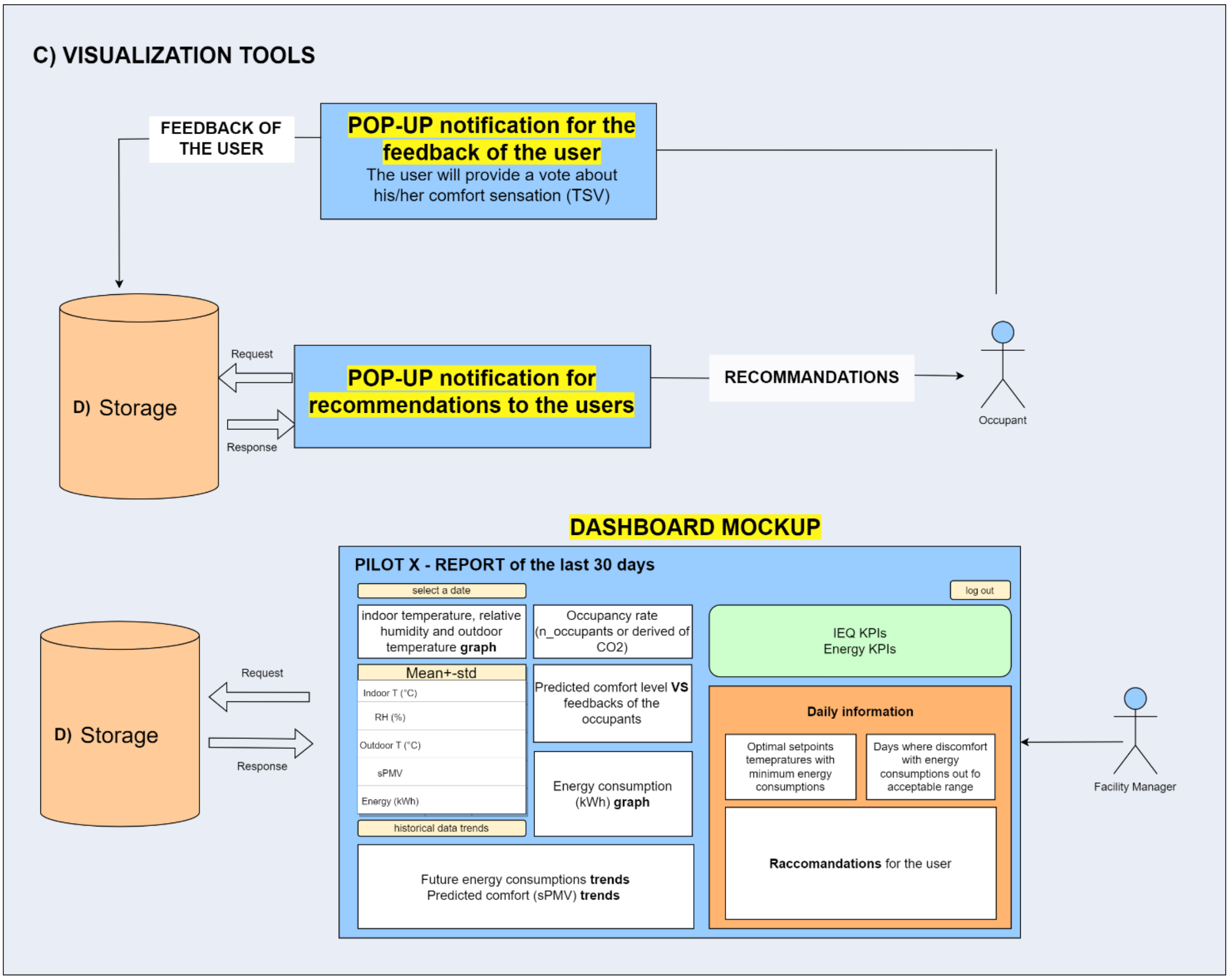


Figure 49: Section of s3.3.1 architecture dedicated to the visualisation tools.

- The last box shown in Figure 47 is the red one. This section aims at developing an optimisation algorithm considering forecasted energy consumption and comfort. Specifically, this approach aims to determine an optimal temperature ensuring minimal energy consumption and proper comfort level. For this reason, Figure 50 presents a part dedicated to energy forecasting (through dedicated AI algorithms, such as the LSTM network) and a part dedicated to the optimisation activity. For the computation of both algorithms, input data will be retrieved from the project repository (as shown in Figure 47) and the outcomes of the algorithms will be stored in it. However, given the pivotal role that energy forecasting component has, a collaboration with CARTIF (leader of Task 3.1 "AI-based services for finer grained energy profiling & forecasting") has been established and future technical developments will follow.

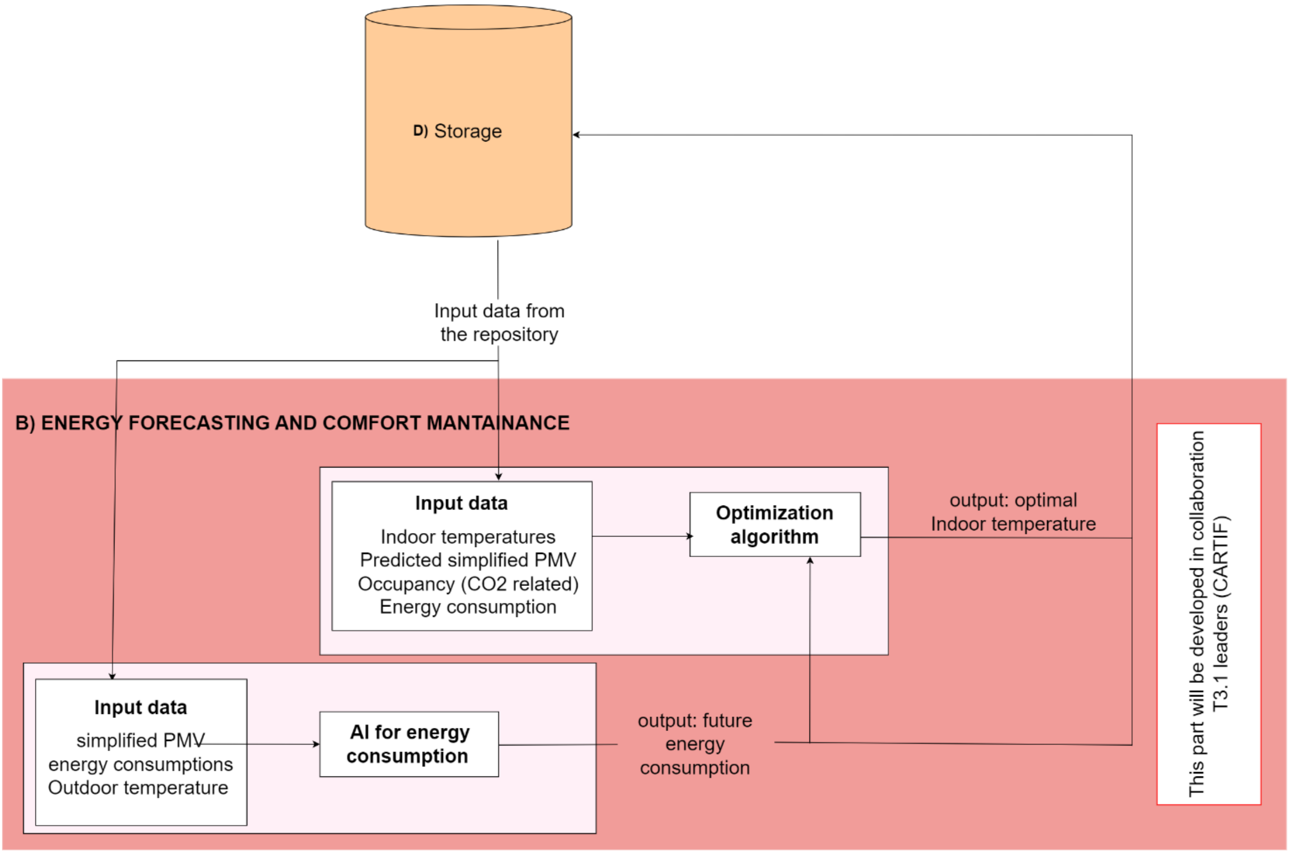


Figure 50: Section of s3.3.1 architecture dedicated to energy models.

The reader can find the original file of the described block scheme (Figure 47) on GitHub in order to provide a deeper and more detailed analysis of the whole architecture.

Additional technical aspects were already reported into deliverable D5.3 titled "Pilots' Execution Documentation – Pre-pilot Phase" (M17). Together with D3.1, this document shows the initial deployment of the service during the pre-pilot phase, presenting associated results and limitations. It serves as a foundational reference for the initiation and improvement of the first concrete implementation of the specified software components. Consequently, readers seeking deeper insights into the models and analysed data are advised to refer also to D5.3 for comprehensive technical details.

All technologies enclosed in s3.3.1 architecture will be implemented through the Python language as already outlined in D3.1 and all the Python codes developed within this framework have been uploaded in DigiBUILD repositories (i.e., Jupyer, GitHub, MLflow). The links to access them follows:

- Jupyter: http://digibuild.epu.ntua.gr:8888/lab/tree/s3.3.1 .
- MLflow: http://digibuild.epu.ntua.gr:5000/ .
- Github: https://github.com/digibuild-technology-release/s3.3.1_s3.3.2_UNIVPM .

5.1.2 Novelty

The innovative aspects enclosed in s3.3.1 touches different dimensions of digital service development. From a technical point of view, the exploitation of ML and AI techniques to classify and predict comfort in terms of sPMV detaches s3.3.1 from the state of the art (25), (26).Moreover, the pivotal role of the user in s3.3.1 development and application represent also an important novelty. In fact, in most cases of comfort assessment, user feedback is often sought for a direct but subjective evaluation of thermal sensation, yet no coaching solution is usually considered. In this context, users' feedback is pivotal for the implementation of the ML and AI algorithms. Furthermore, the important role of the end-user is based on the recommendations that should be delivered to actively adjust the environmental surroundings and maintain or guarantee a proper comfort level (coaching solution).

5.1.3 Development Progress

This paragraph presents the progress brought by the activities conducted during these months in accordance with the s3.3.1 outcomes reported in D3.1 and D5.3. Indeed, deliverable D3.1 presents the description of the general architecture designed for s3.3.1 together with the preliminary results of an early baseline monitoring (Phase 3) developed considering the available historical data from Pilot 5b and Pilot 2. These outcomes have led to the MS4 (M13) with the "'First wave' of DigiBUILD solutions (technology micro-services components/enablers)". In deliverable D5.3, instead, the technical progress of the previous results has been presented. In particular, the development and application of ML baseline models and of a preliminary version of the LSTM network has been described in the document together with the results in the framework of the pre-pilot phase.

The upcoming months have been focused on the improvement of the service from data and models point of view, in accordance to the outcomes reported in previous deliverables (D3.1, D5.3). Thus, the development progress has been focused on:

- Data Analysis and data labelling,
- Improvement of the baseline models and LSTM network,
- Development of the pop-up notification system.

These steps will be described in the following paragraphs.

Data analysis and data labelling

The examination of the results outlined in D5.3, particularly the validation metrics for both baseline models and LSTM, has revealed the necessity for improvements at dataset level. Indeed, the accuracies obtained during the pre-pilot phase highlight the influence of an unbalance among the input data and of the unavailability of data in some pilots. Consequently, an additional round of data analysis has been conducted to improve the service. For instance, with respect to what has been presented in both D3.1 and D5.3, for pilot 5b and Pilot 2 new environmental data have been included in the input dataset of the ML and AI algorithms. This has triggered also a new data labelling session which has been computed with the python function already available on GitHub from M13. This activity has been implemented with Python and the codes are on Jupyter repository.

Improvement of baseline models and LSTM network

In accordance with the progress described in the above lines, baseline models and LSTM network have been improved with respect to the results shown in previous deliverables (D3.1, D5.3). As regards the baseline models, they have been implemented to classify comfort in terms of sPMV. The dedicated Python codes have been applied to different pilots, and their validation metrics will be discussed in the following paragraphs. As regards the LSTM network, it has been implemented as regression model to predict comfort in terms of sPMV. Specifically, two trials have been computed to predict comfort considering a forecasting horizon of 1 hour and 24 hours. In both cases, the LSTM model has been built with two layers characterised by 64 neurons and a dense layer characterised by a number of units equal to the forecasting horizon. Python was used as the programming language, and the dedicated codes are available on DigiBUILD Jupyter repository and on Github.

Pop-up notification system development

Unlike the alpha version of s3.3.1, the current stage of the service development, the pop-up notification system has been implemented and tested internally in some of the UNIVPM offices. In particular, the developed system is dedicated only to collecting user feedback on their thermal sensation, expressed as Thermal Sensation Vote (TSV). From a technical point of view, this system consists of a webpage (Figure 51), where the user should provide the following information: the pilot, the room, the building, the floor, and his/her TSV. This web page is accessible through the following link: https://digibuild-demo.eu/digibuild/. All the requested info, once inserted, is stored in a dedicated database which is reachable at the following link: https://digibuild-demo.eu/phpmyadmin/ . The system has been already tested internally in UNIVPM offices. For pilot-level testing, scheduled for January, a detailed protocol has already been shared with the pilots' leaders to facilitate feedback collection.

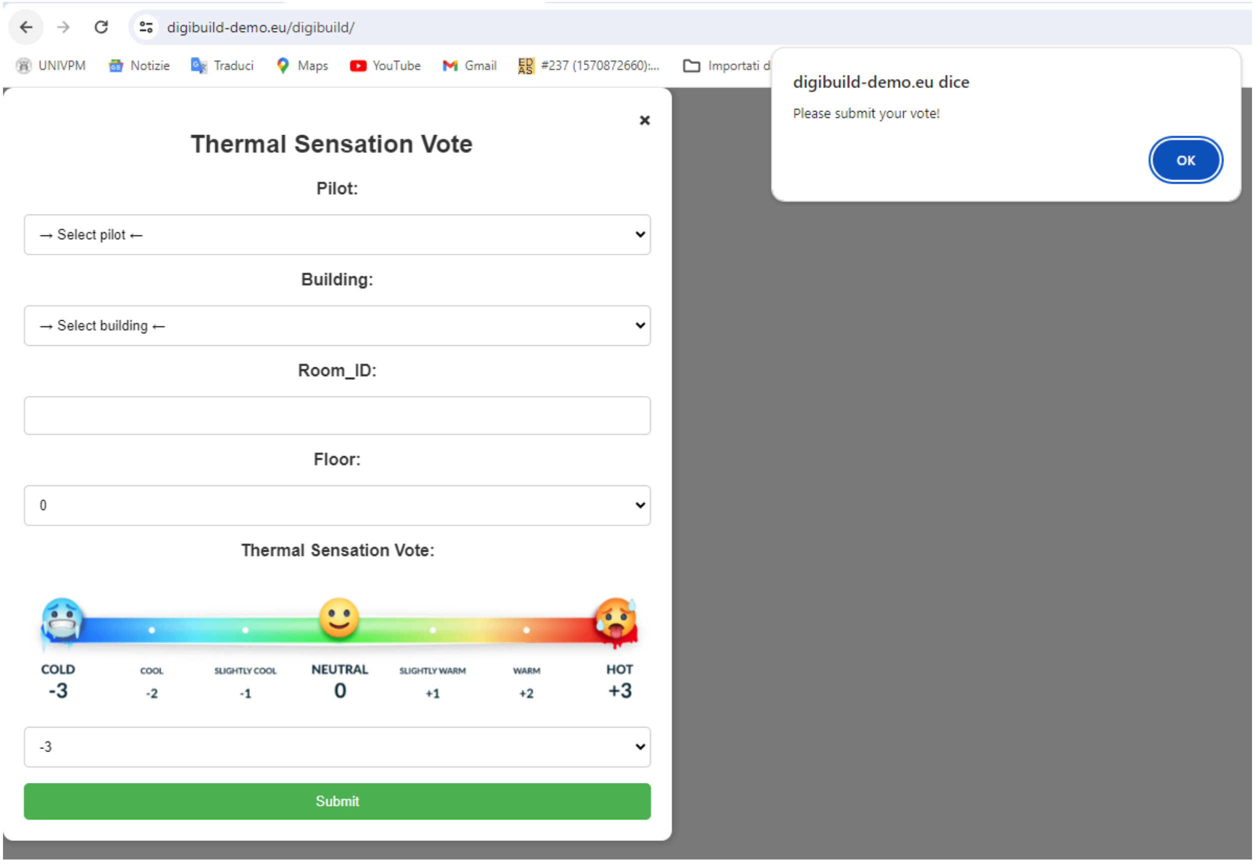


Figure 51: Web page developed to collect users' feedback.

The pop-up notification system has been developed through PHP programming language, JavaScript and HTML in order to build up a dynamic web page that stores data on the mentioned database. The codes are available on the GitHub repository and on Jupyter.

5.1.4 Application on DigiBUILD Pilots

〉 Pilot 2 – EDF

In this paragraph it will be possible to appreciate some results of the application of the baseline models and LSTM network in Pilot 2 buildings. As already described in D3.1, Pilot 2 develops in the EDF offices. Environmental data are collected from the ground floor and first floor of the building from two kind of sensors called respectively Ellona sensors and Ethera sensors. In this pilot, energy meters are also installed to collect energy (kWh) and power data (kW). For more details about the pilot sites and the exploited data, the reader can consult D3.1. Moreover, the reader is also invited to analyse deliverable D5.3, for the outcomes of the first application of the algorithms during the pre-pilot phase. Specifically, in the document the early application of the service on the EDF pilot scenario can be appreciated. Starting from the results of D5.3 and the issues raised during the pre-pilot phase, s3.3.1 has been applied to EDF pilot in order to get better performances of the algorithms.

The first action point has been implemented by the EDF team and it regarded data analysis and the subsequent data labelling. Indeed, during these months some data formats have changed due to data refreshment. Consequently, data have been analysed and resampled gaining three datasets regarding the ground floor, the first floor north and the first floor south of the pilot building. The codes applied for this activity are available in the DigiBUILD Jupyter Repository. At this stage, feature selection has been applied on the data collected from the ground floor and first floor. In particular, three correlation matrices have been computed in python and they are reported in Figure 52, Figure 53, Figure 54. According to them, the features that have been selected as inputs data for the ML and AI algorithms have been the following: indoor temperature (°C), indoor relative humidity (%), CO_2 (ppm) and outdoor temperature (°C). Specifically, from the correlation matrices, the service developers have selected the parameters with high accuracy value with the label (sPMV). Then, the input datasets have been enhanced with those parameters that do not correlate each other. In other words, if some variables show high correlation values (Figure 52) within each other, only one of them is kept in order to avoid inter-correlation among input data and reduce the possible overfitting issues that may emerge during the algorithms' implementation. The reader can appreciate the exploited input datasets in Table 6. The technical partners have decided to include data collected from the ground floor and first floor north in accordance to the quantity of available historical data.

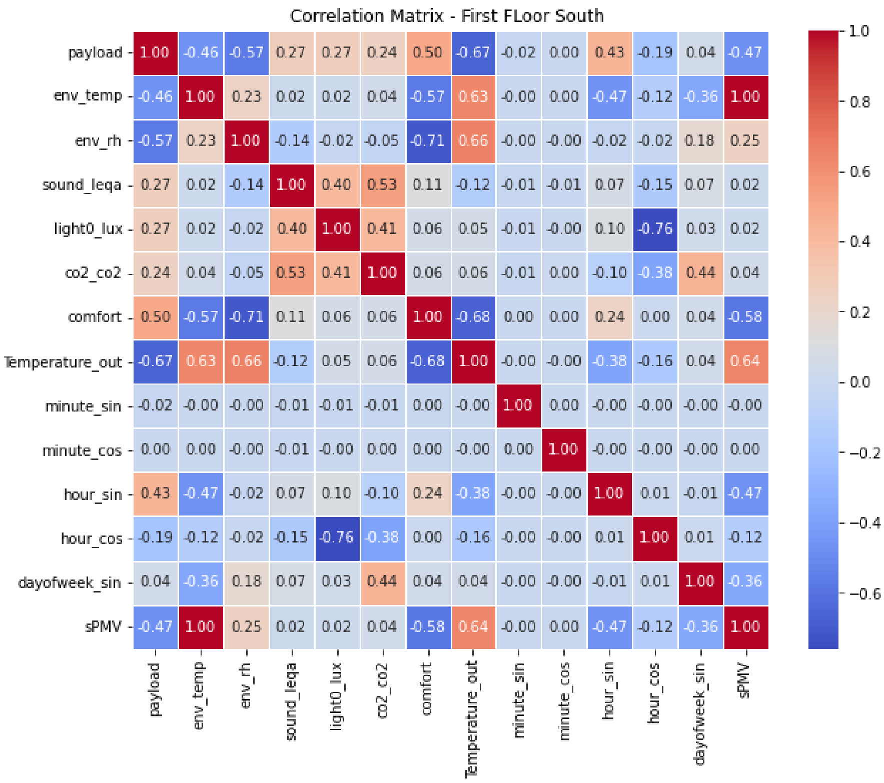


Figure 52: Correlation matrix - First Floor South case.

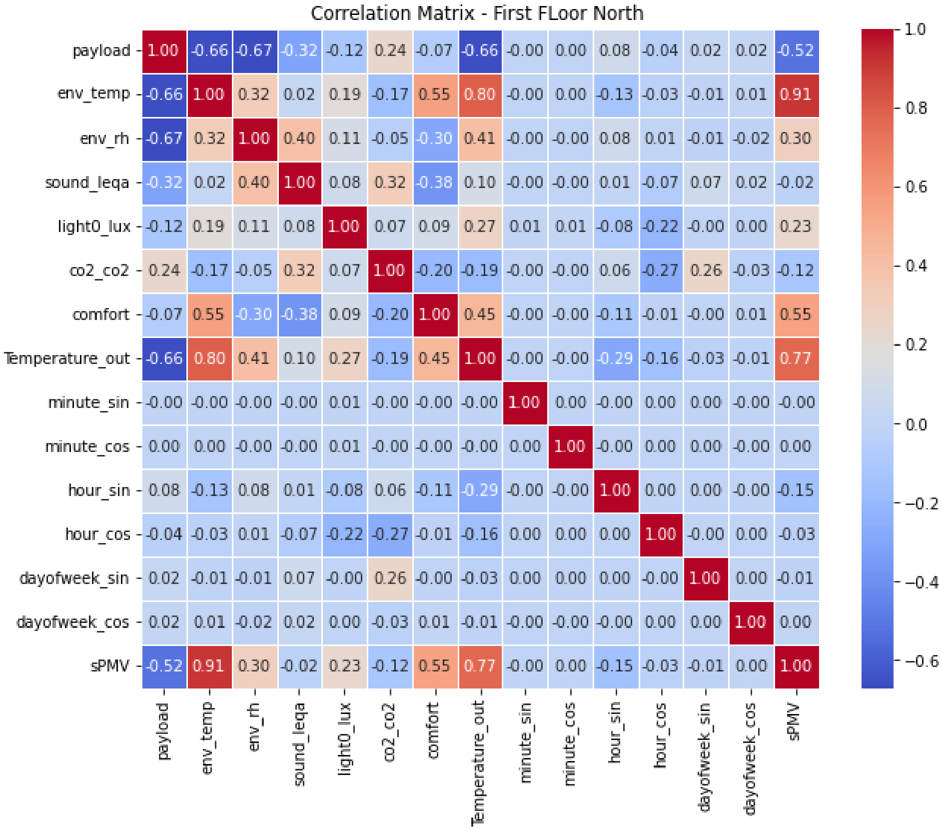


Figure 53: Correlation matrix - First Floor North case.

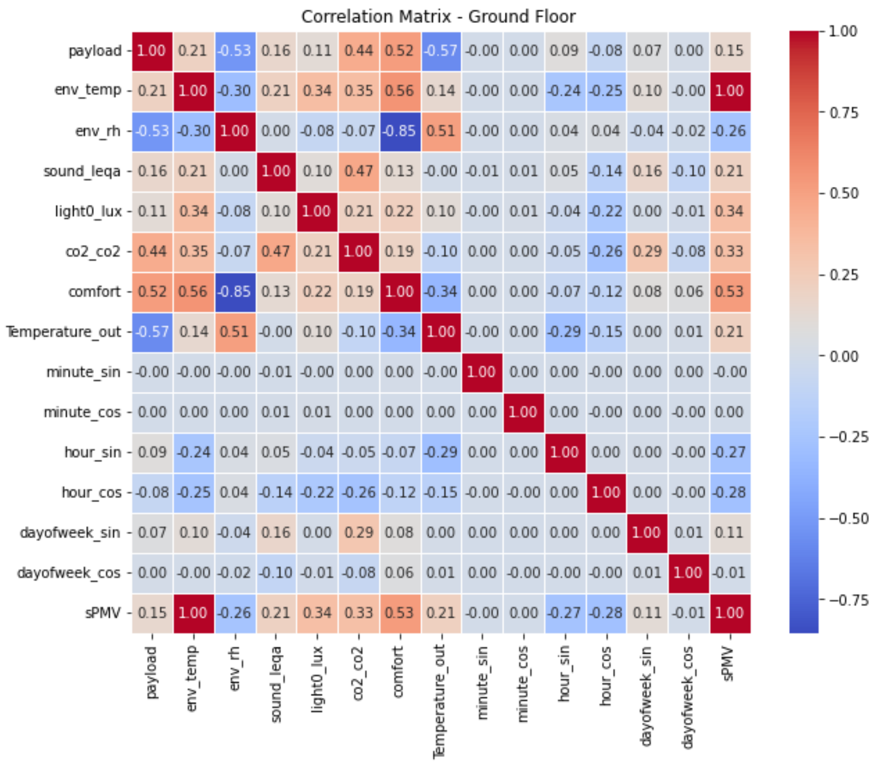


Figure 54: Correlation matrix - Ground Floor case.

Data labelling has been applied exploiting the python function implemented for the computation of sPMV (link to Jupyter repository).

At this point, the implementation of baseline models (Random Forest, Naïve Bayes, Decision Tree, K-nearest Neighbours, Adaboost, Bagging) have started. The algorithms have been trained and tested as classificators and their validation metrics are reported in Table 7 and Table 8, respectively for the first floor north and ground floor. The outcomes reported in Table 7 and Table 8 highlight the high performance of the algorithms and the improvements of the models with respect to past results (D5.3). For instance, in the First Floor North case (Table 7), the mean accuracy is 97.1%, underlining the proper capability of the chosen algorithms in classifying sPMV starting from the available historical data.

Table 6: Datasets exploited for the baseline models and LSTM algorithm implementation – Pilot 2.

	Input data	Output
Dataset 1 – First Floor North	• Indoor Temperature (°C) • Indoor Relative Humidity (%) • CO_2 (ppm) • Outdoor temperature (°C)	sPMV
Dataset 2 – Ground Floor	• Indoor Temperature (°C) • Indoor Relative Humidity (%) • CO_2 (ppm) • Outdoor temperature (°C)	sPMV

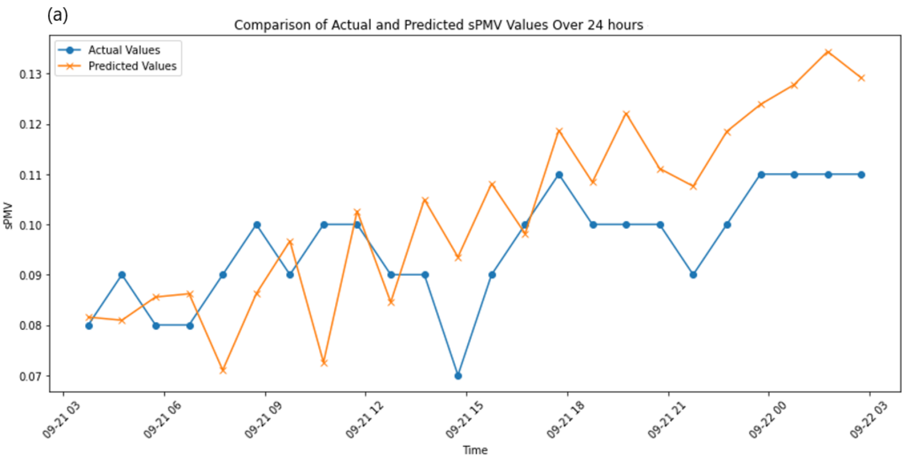
Table 7: Validation metrics of the baseline models implemented with data collected from the First Floor North.

Baseline model	Accuracy	F1_score	Precision	Recall
Random Forest	99.9%	99.7%	99.8%	99.7%
Decision Tree	99.8%	99.5%	99.7%	99.3%
Naive Bayes	93.9%	82.6%	82.8%	82.8%
KNN	98.6%	96.1%	96.9%	95.3%
Adaboost	90.5%	55.2%	66.4%	75.5%
Bagging	99.8%	99.5%	99.8%	99.4%

Table 8: Validation metrics of the baseline models implemented with data collected from the Ground Floor.

Baseline model	Accuracy	F1_score	Precision	Recall
Random Forest	99.9%	98.9%	97.9%	99.9%
Decision Tree	99.8%	98.9%	97.9%	99.9%
Naive Bayes	99.1%	81.5%	73.2%	99.6%
KNN	99.6%	83.2%	86.7%	80.4%
Adaboost	99.9%	98.9%	97.9%	99.9%
Bagging	99.9%	99.5%	98.9%	99.9%

As regards the application of the LSTM network (described in section s1.1.3), some example of its outcomes can be appreciated in Figure 55 and Figure 56. In particular, the LSTM has been implemented considering data collected from both ground floor and first floor north to forecast comfort (as sPMV) for the next hour (Figure 55b, Figure 56b) and for the next 24 hours (Figure 55a, Figure 56a). The performance of the algorithm has been evaluated through the metrics reported in Table 9 and Table 10. The values of MAE and MSE points out the high performance of the algorithm in predicting comfort (as sPMV) for future time periods. For example, in the case of the LSTM applied in the First Floor North of the building (Table 8), the validation metrics (e.g., MAE= 0.09, MSE=0.01) reveal accurate predictions of the sPMV for the chosen forecasting horizons.



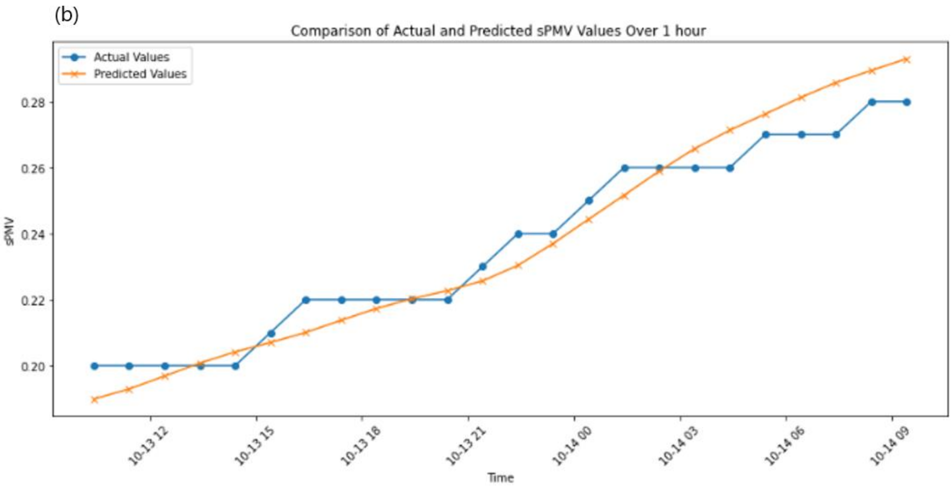
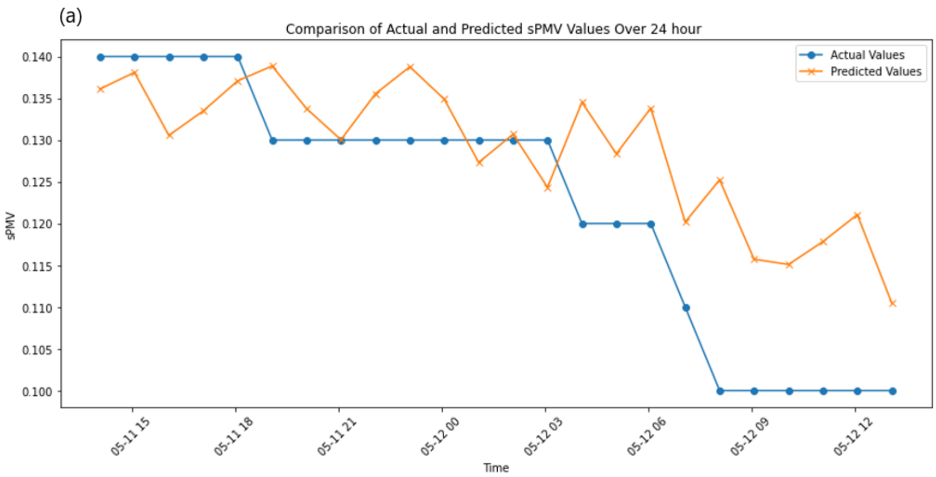


Figure 55: Comparison of Predicted sPMV and the actual sPMV for the four different seasons. In this case, the sPMV prediction has been done for the next 24 hours (a) and the next hour (b). - First floor north case.



Funded by
the European Union

(b)

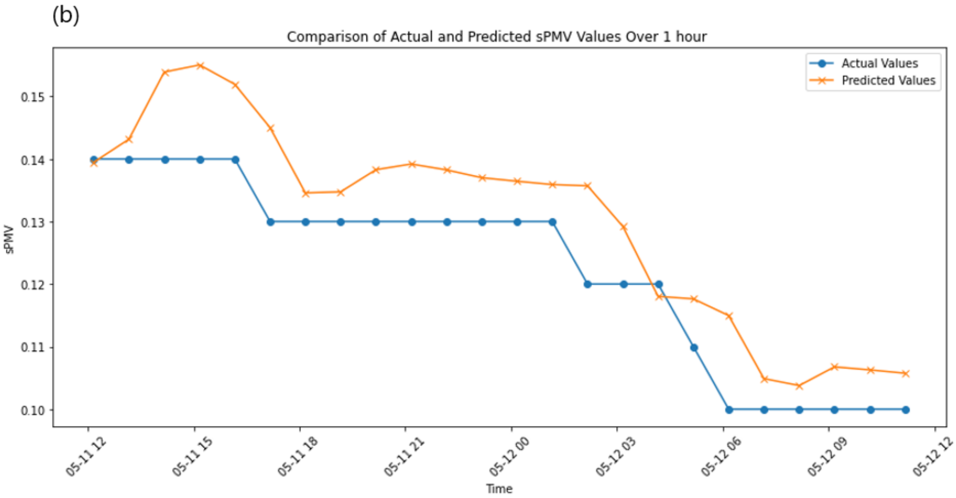


Figure 56: Comparison of Predicted sPMV and the actual sPMV for the four different seasons. In this case, the sPMV prediction has been done for the next 24 hours (a) and the next hour (b). - Ground floor case.

Table 9: Validation metrics of the LSTM network for the prediction of sPMV of the next hour and next 24 hours (First Floor North case) – Pilot 2.

	MAE	MSE
LSTM for predicting the next hour	0.09	0.01
LSTM for predicting the next 24 hours	0.08	0.02

Table 10: Validation metrics of the LSTM network for the prediction of sPMV of the next hour and next 24 hours (Ground Floor case) – Pilot 2.

	MAE	MSE
LSTM for predicting the next hour	0.007	6.18×10^{-5}
LSTM for predicting the next 24 hours	0.01	0.003

〉 Pilot 5b – FOCCHI

In deliverable D5.3, there are the last updated results of s3.3.1 application on Pilot 5b. From M15 up to M19, s3.3.1 development in Pilot 5b has been focused on the analysis of Focchi historical data collected from Demoroom 2 and their reorganisation to create a balanced input dataset for the ML and AI algorithms, overcoming the data issues highlighted in D5.3 (the reader is invited to consult the mentioned document). In particular, the features selected for the implementation of the models in Pilot 5b are: indoor temperature (°C), indoor relative humidity (%), indoor CO_2 (ppm), outdoor temperature (°C) and external light intensity (lux). This feature selection step has been guided by analysing the correlation matrix reported in Figure 57. Firstly, the service developers have observed which parameter had a high correlation value with the label (sPMV). Then, other parameters have been chosen with the following criteria: to avoid inter-correlation among input variables and the subsequent overfitting issue that may emerge during the models computation, the parameters that exhibited high correlation within each other have been excluded. This step has produced the input dataset reported in Table 11.

The next step has been the implementation of the baseline ML models (uploaded on MLflow), whose validation metrics are reported in Table 12. These results reveal the improvements of these algorithms, given the high performance reached in terms of accuracy, f1_score, recall and precision. For example, from Table 10, it is possible to infer that the chosen ML algorithms successfully classify comfort (as sPMV) given the mean accuracy of 91.6%.

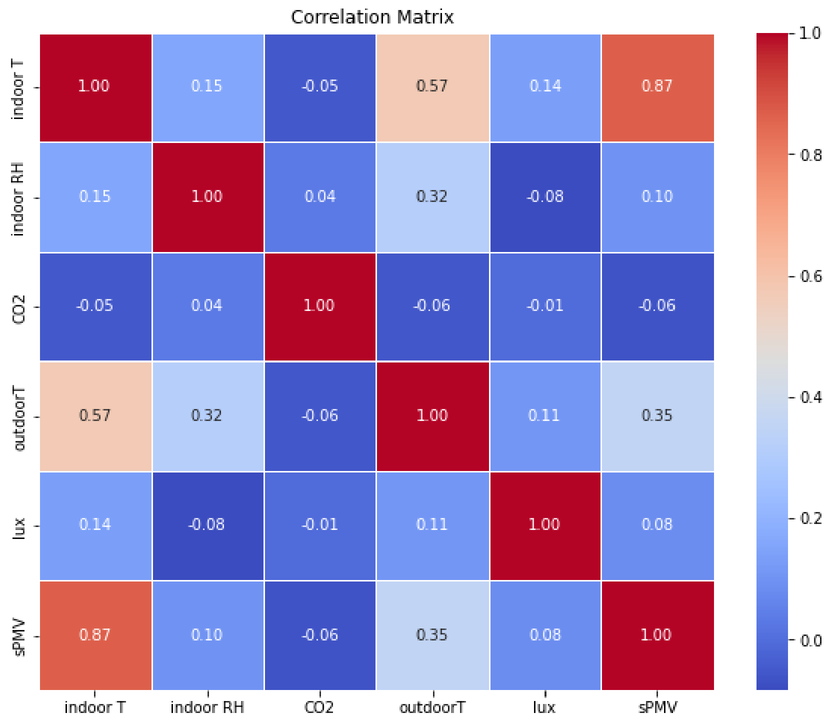


Figure 57: Correlation matrix to select the input features.

Table 11: Dataset exploited for the baseline models and LSTM algorithm implementation – Pilot 5b.

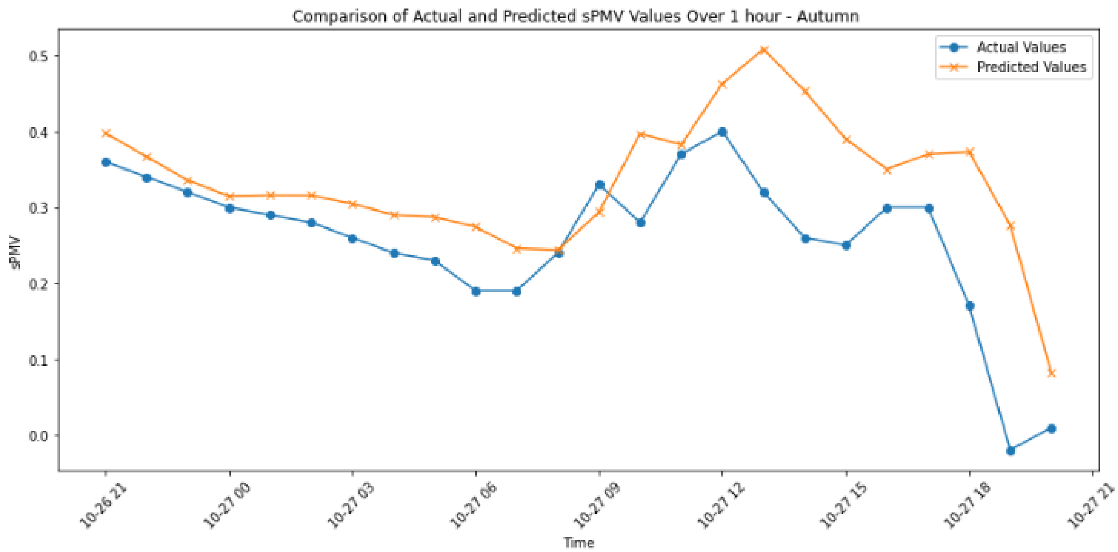
Input data	Output

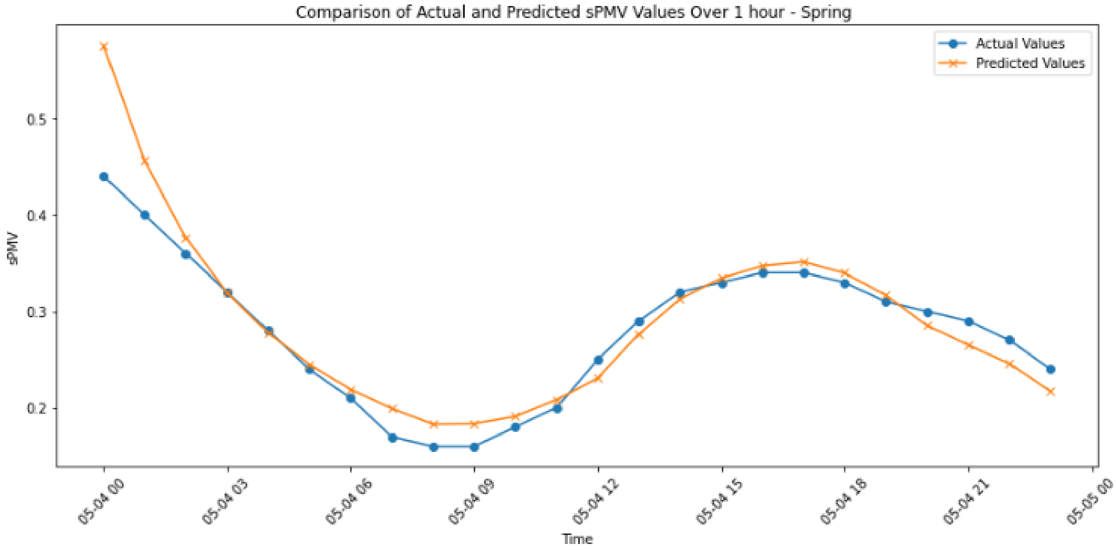
Dataset – Demoroom 2	• Indoor Temperature (°C) • Indoor Relative Humidity (%) • CO_2 (ppm) • Outdoor temperature (°C) • External light intensity (lux)	sPMV

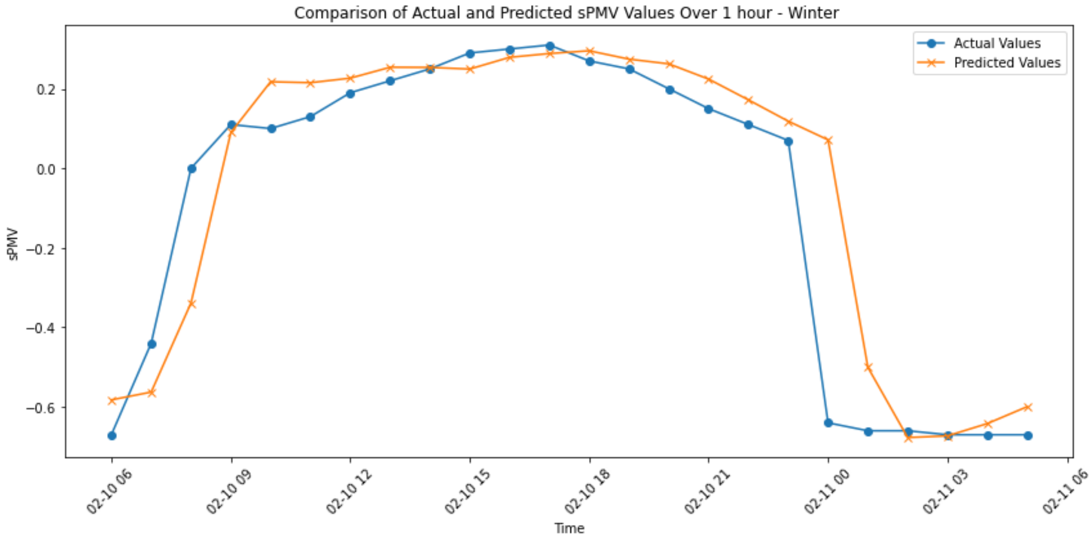
Table 12: Validation metrics of the baseline models implemented for Pilot 5b.

Baseline model	Accuracy	F1_score	Precision	Recall
Random Forest	96.8%	93.3%	95.5%	91.6%
Decision Tree	95.6%	91.3%	90.8%	91.9%
Naive Bayes	90.3%	76.1%	76.8%	77.7%
KNN	86.2%	51.8%	68.2%	52.4%
Adaboost	84.2%	72.5%	72.7%	72.9%
Bagging	96.7%	93.2%	95.3%	91.4%

As regards the LSTM network, the model has been trained by considering the above mentioned input data (Table 11) collected from 6 out of 4 rooms of the Demoroom 2. The outcomes of LSTM model application can be appreciated in Figure 58 and Figure 59. In particular, the reported graphs compare the sPMV calculated with Focchi historical data (label) with the sPMV predicted for the next hour (Figure 58) and the next 24 hours (Figure 59). The high performance of the algorithm is underlined by the validation metrics reported in Table 13. Indeed, MAE and MSE values demonstrate the accurate capability of the developed network in predicting sPMV for the two different forecasting horizons.




Comparison of Actual and Predicted sPMV Values Over 1 hour - Spring
Actual Values
Predicted Values
sPMV
0.5
0.4
0.3
0.2
05-04 00
05-04 03
05-04 06
05-04 09
05-04 12
05-04 15
05-04 18
05-04 21
05-05 00
Time


Comparison of Actual and Predicted sPMV Values Over 1 hour - Winter
Actual Values
Predicted Values
sPMV
0.2
0.0
−0.2
−0.4
−0.6
02-10 06
02-10 09
02-10 12
02-10 15
02-10 18
02-10 21
02-11 00
02-11 03
02-11 06
Time

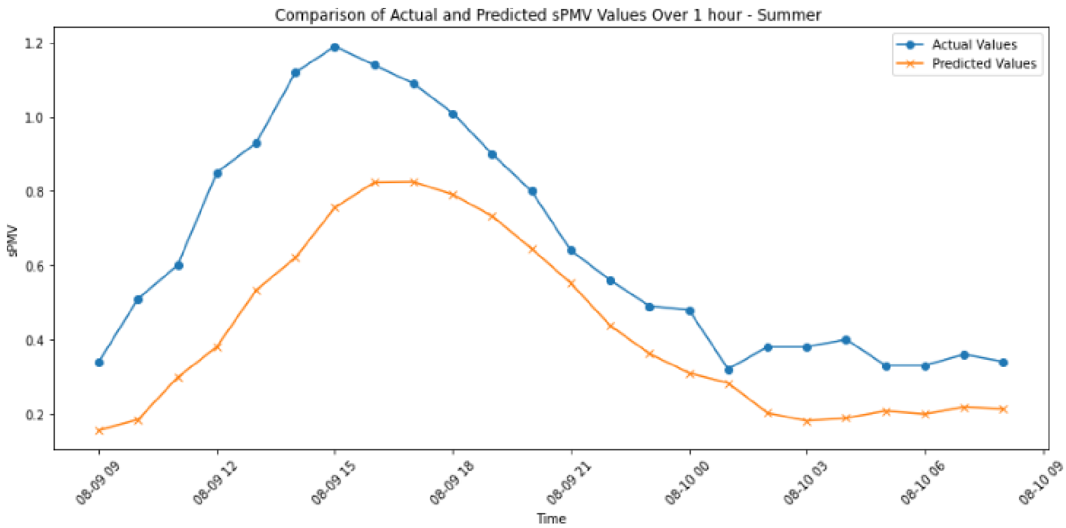
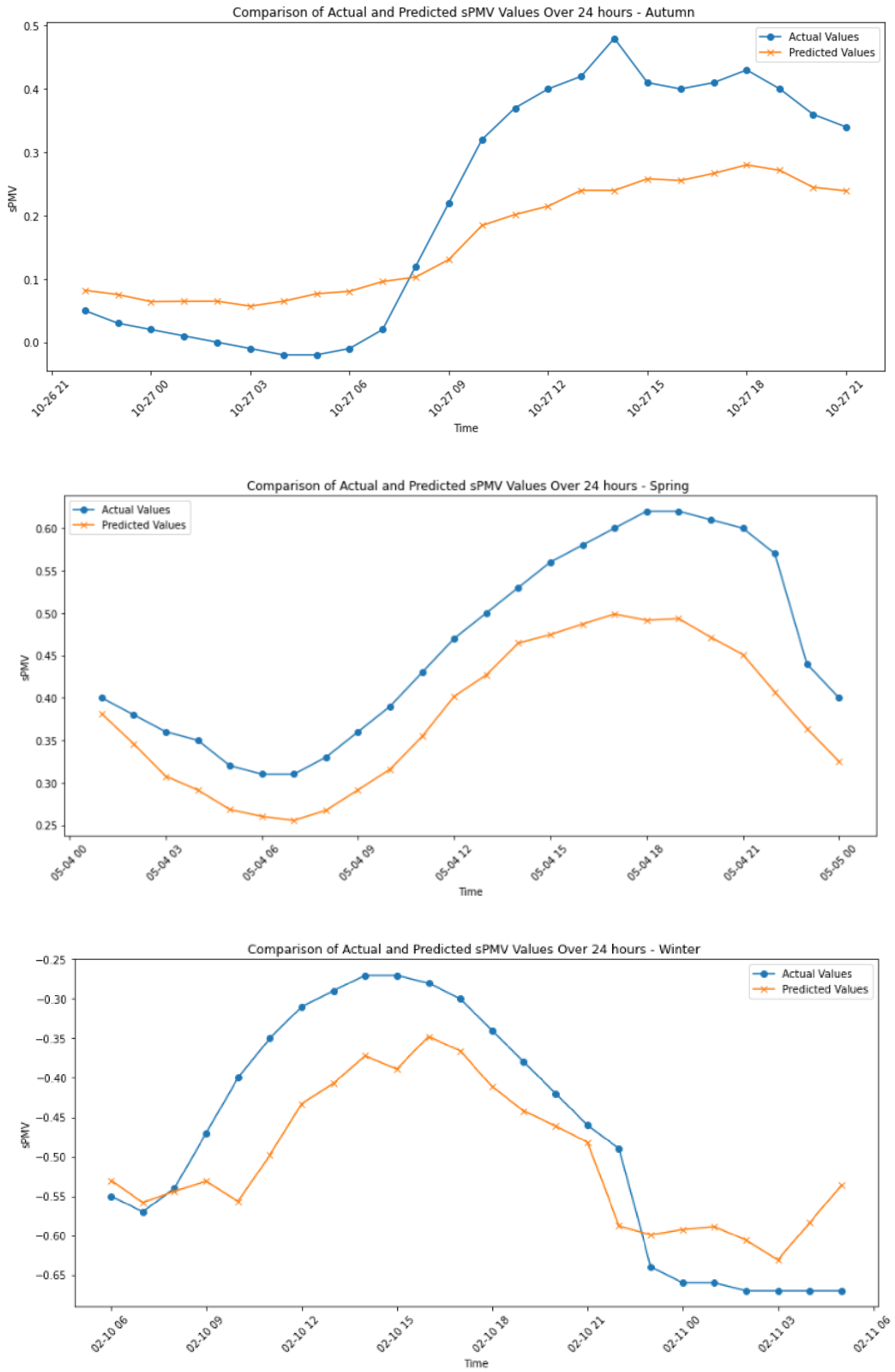

Comparison of Actual and Predicted sPMV Values Over 1 hour - Summer
Actual Values
Predicted Values
sPMV
1.2
1.0
0.8
0.6
0.4
0.2
08-09 09
08-09 12
08-09 15
08-09 18
08-09 21
08-10 00
08-10 03
08-10 06
08-10 09
Time

Figure 58: Comparison of Predicted sPMV and the actual sPMV for the four different seasons. In this case, the sPMV prediction has been done for the next hour.



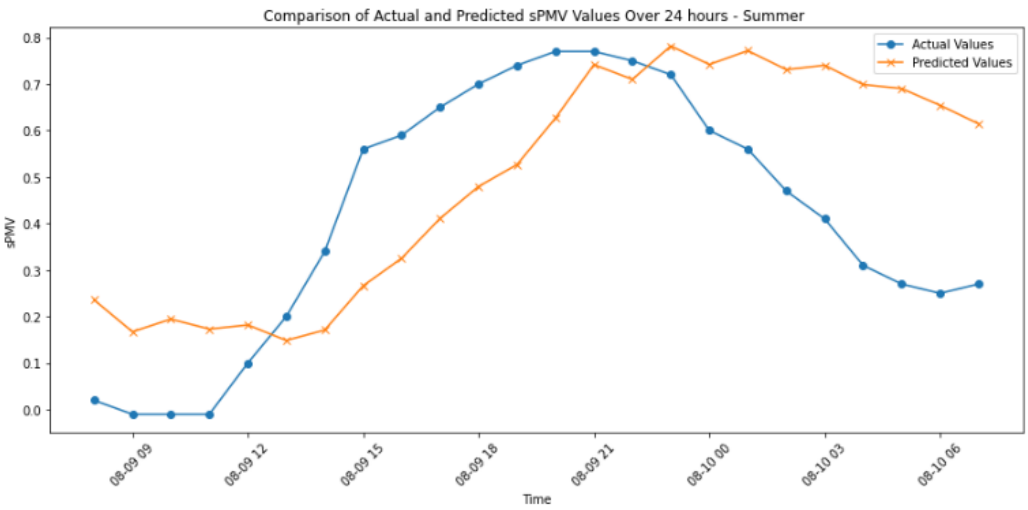


Figure 59: Comparison of Predicted sPMV and the actual sPMV for the four different seasons. In this case, the sPMV prediction has been done for the next 24 hours.

Table 13: Validation metrics of the LSTM network for the prediction of sPMV of the next hour and next 24 hours - Pilot 5b.

	MAE	MSE
LSTM for predicting the next hour	0.1	0.019
LSTM for predicting the next 24 hours	0.085	0.015

The python codes used at this stage of the project are online in the Jupyter repository and on the Github repository.

5.1.5 Next Steps

In order to enter the full pilot operation phase at M21 with s3.3.1 ready to be tested and used in the pilots' sites, the next steps should be taken into account:

1. Apply the previously described software components on the remaining pilots.
2. Start the real-time implementation of the service.
3. Compute ML and AI models (Figure 48) to predict the TSV starting from the feedback collected with the pop-up notification system.
4. Start the implementation of the adaptive comfort model (Figure 48).
5. Exploring additional comfort dimensions (e.g., visual comfort, IAQ, etc.) could be a valuable next

step in achieving a comprehensive assessment of comfort enhancing its overall evaluation.

6. Finalise the integration of the s3.3.1 with WP4 tasks, specifically Task 4.1 for Pilot 5b Digital Twin.
7. Test the pop-up notification system in the pilot sites.
8. Start the implementation and development of the dashboard.
9. Start the implementation of the software component dedicated to the energy forecasting related to comfort maintenance. Within this framework, UNIVPM has started a collaboration with CARTIF (leader of Task 3.1) in the past months.

5.2 Comfort Performance Contract (s3.3.2)

5.2.1 Description of the Service

The primary objective of the Comfort Performance Contract (CPC) is to deliver a service ensuring the satisfaction of the end-user by maintaining optimal indoor conditions decided by the end user at a pre-established energy cost. This service assumes responsibility for the comfort and well-being of individuals within the built environment through the establishment of a contractual agreement linking its energy costs (and consumption) directly to the indoor comfort condition. Users have the flexibility to specify their desired comfort levels based on the building type and historical energy usage. Additionally, the option exists to incorporate sensor network installations if necessary to ensure the agreed level of comfort.

This service will be deployed considering the phases listed below which have been already described in deliverable D3.1:

- **Phase 0**: *Assessment model for Enhanced Comfort and Wellbeing.*
- **Phase 1**: *Integration of the data inside the service Data Lake—Interoperability and Quality*
- **Phase 2:** *Definition of the technical specifications of the service*.
- **Phase 3**: *Baseline monitoring. Track energy consumptions (T3.1) and track comfort parameters.*
- **Phase 4**: *Profiling and individuation of anomalies.*
- **Phase 5**: *Service Level Agreement (SLA) with stakeholder.*
- **Phase 6**: *Budget simulation and refinement of technical specification and SLAs.*
- **Phase 7**: *CPC management.*

It is possible to note that, at this stage, s3.3.2 shares the first 4 phases of its development with s3.3.1. This implies that the software components and visualisation tools described in the previous chapter regards also the CPC. Hence, s3.3.1 and s3.3.2 developed parallel from a technical point of view. However, the peculiarity of s3.3.2 is the development of the financial and legal framework that links the comfort assessment and maintenance with the energy consumption and the related costs. In other words, what differentiates s3.3.1 and s3.3.2 is the economic bond that s3.3.2 aims to establish between indoor comfort maintenance and energy consumption in order to enhance people's well-being at proper energy costs.

5.2.2 Novelty

The novelty introduced by this service is the possibility to provide an economic plan which guarantees comfort at established energy costs in specific indoor environment; in opposition with the current situation in the energy market, now at the centre of the contract there is the comfort of the occupant, rather than the energy consumption. This financial link that s3.3.2 will establish aims at improving people experience in a certain building and the efficiency of this last from an economic point of view. Furthermore, given that s3.3.2 shares the same software blocks with s3.3.1, the usage of AI and ML approaches for comfort prediction and classification represents another innovative aspect that CPC service will bring in the built environment scenario.

5.2.3 Development Progress

Starting from what has been reported in D3.1 and in D5.3, s3.3.2 progress has been characterised by the establishment of a collaboration with the UNIVPM faculty of Economy. Indeed, the main focus during this period has been producing the document that formalises the financial bond between comfort maintenance and energy consumption to guarantee indoor comfort. The technical knowledge about finance and management provided by the UNIVPM colleagues of economy has been pivotal for the creation of the first draft of the contract. In accordance with their expertise, the document has been systematically organised through the employment of a modular approach, delineating every aspect of IEQ, including the norms and standards that regulates IEQ. This methodological framework has been necessary, given that each dimension of IEQ can impact on a building in terms of energy consumption and efficiency. Furthermore, adopting this approach provides the potential end-user with the freedom to invest in a comfort aspect that aligns more conveniently with the specific needs of the case study. Consequently, the document encompasses the following key sections:

1. **Definition of the Parties:** in this part, the roles of client and operator are clarified. In particular, this part should be filled with the personal data of both client and economic operator.
2. **Premises:** the document points out two primary premises where the fundamental aspects and concepts at the base of the IEQ are explained with the support of the reference standards (such as UNI EN ISO 9920, ASHRAE 55, UNI EN ISO 7726, etc.). The contractual objective is to furnish a comprehensive plan that encompasses and addresses various facets of comfort (e.g., thermal comfort, visual comfort, and indoor air quality) within an indoor environment. It is fundamental that the client is aware of all of these aspects.
3. **Object of the Contract:** This section encapsulates three distinct articles. The initial article delineates the terms and conditions governing the relationship between the involved parties, specifying these terms for each aspect of IEQ. The focus of the second article revolves around privacy concerns and the administration of personal data. The concluding section delineates the financial bond between the parties, defining the fees and payment modalities that will be activated upon the execution of the contract.

5.2.4 Application on DigiBUILD Pilots

The CPC document, designed in Italian language, is currently undergoing translation into English (M20). It will be applied on two pilots of the project, which are Pilot 5b (FOCCHI) and pilot 7 (IEECP). While it has not yet been tested on the Pilots in this stage of the project, it is noteworthy that all the technical applications concerning comfort prediction and classification, as detailed in section 1.1.4, extend to s3.3.2. Indeed, the contract emphasises the assurance of comfort at specified energy costs, achieved through a comprehensive comfort evaluation utilizing techniques developed within the s3.3.1 context. The next action point is focused on retrieving feedback from Pilot leaders to gather insights from the end user perspective about the contract, and also enhance the quality and completeness of the document. This collaborative effort aims also to engage potential users effectively. In fact, in the upcoming period the involvement of end-users will start to provide valuable feedback within a real-life context (Pilot sites). This approach aims at refining the document for optimal applicability in the next full-pilot operation phase.

5.2.5 Next Steps

With the aim of reaching the full pilot operation phase with a complete version of the CPC in the framework of s3.3.2, the following activities will be developed in the upcoming period:

1. Provide the CPC document to the pilots' leaders to retrieve feedback.
2. Engage end-users of the related pilots to collect feedback about the contract.
3. Finalise the document.

All the next steps listed and described in section 1.1.5 apply also in this case given the common path that s3.3.2 and s3.3.1 shares from a software point of view.

6 Data-driven services for renovation roadmaps and energy efficiency financing

The DigiBUILD project continues to advance in its mission to foster renovation and heighten energy efficiency in the building sector. This next phase of the project will leverage advanced data-driven solutions to chart renovation paths and provide innovative financing options for energy efficiency. Our focus in this section is to outline the newly developed services aimed at realizing these goals.

One of the key targets is the execution of sophisticated modelling strategies to enhance demand-side management and pinpoint effective energy efficiency upgrades in a benchmark building. In the subsection titled "3.4.1 – Financing and Policy Making for Energy Efficiency," we will introduce an evolved model: the Enhanced Dynamic high-Resolution Demand-side Management (e-DREEM). This upgraded model is designed to evaluate building performance more comprehensively, offering tailored solutions for optimal energy usage. Its objectives are not only to facilitate long-term energy savings and tackle energy poverty but also to incorporate cutting-edge intelligent algorithms and advanced analytical techniques for a more effective building energy management system.

Further, under "3.4.2 – Comprehensive One-stop-shop for Energy Efficiency," we are developing an integrated platform to streamline energy efficiency retrofitting processes in buildings. This platform stands out for its holistic evaluation approach, examining the impact of retrofitting actions not just in energy terms but also considering economic, environmental, and social benefits. The platform's primary aim is to heighten the awareness of building owners and managers about energy efficiency measures, equip them with the necessary knowledge and tools for informed decision-making, and provide access to various financing options and incentives at both national and local levels for energy efficiency upgrades.

In essence, the expanded services under the DigiBUILD project are set to significantly lower energy consumption and greenhouse gas emissions in the building sector. This initiative promotes sustainable and cost-effective energy solutions. The ongoing development and implementation of these services are anticipated to foster a more efficient, transparent energy market, yielding benefits for building owners, users, and the broader environment and society.

6.1 Energy efficiency financing and policy making (s3.4.1)

6.1.1 Description of the Service

The objective of s.3.4.1 is to facilitate the decision-making process of building owners/ users with regards to their building's renovation considering the energy consumption and the related costs of interventions. In this respect, the ultimate premise of this service is to develop integrated renovation roadmaps including different portfolios of measures and financing options using the DREEM model (27). The DREEM model is described in detail in deliverable D3.1 "First wave of DigiBUILD AI-based data-driven services for the built environment" (M12).

Furthermore, as described in D3.1, the s3.4.1 service is implemented following 7 main steps (the reader can find their detailed description in D3.1):

- Step 1: Collection of data related to building characteristics and energy consumption related data.
- Step 2: Creating the dynamic energy simulation model of the building before interventions (baseline)
- Step 3: Calibrating the model using the data collected for the building
- Step 4: Collecting the preferences of building interventions (a checklist of possible renovation actions has been deployed)
- Step 5: Simulating the energy efficient scenarios
- Step 6: Conducting the technoeconomic assessment
- Step 7: Wrapping up the possible renovation actions with potential policy/ financing solutions into concrete renovation roadmaps.

The service is currently being implemented for the IEECP pilot buildings. By developing and calibrating the models for these buildings has helped us understand the data needs and requirements, as well as potential limitations for the users and comments after the submission of D3.2 and the 1st Technical workshop have given us constructive feedback to start working on a user interface for s.3.4.1.

6.1.2 Novelty

Estimating heating and cooling demands is a major issue related to evaluating potential energy-saving actions for buildings. Existing building models are usually either too complex and computationally expensive or too simple to adequately predict a precise load profile. With developments in building energy tools such as TRNSYS and Energy+, detailed building models have been directly used for energy simulations. However, the simulations of such models are often time-consuming, the number of buildings to simulate is limited and modelling multiple energy-saving actions further increases the demand, time, and complexity of the simulations (28). Additionally, deep energy retrofit measures in buildings require very high initial investment costs and their benefits accrue only slowly over time. It is therefore crucial to identify retrofit measures which are not only beneficial for the environment but will also incentivise the owner of the buildings and will ensure effective private and public budget spending (29). The s.3.4.1 aims to tackle both these obstacles in order to facilitate building owners and users to choose their renovation pathway considering not only the energy savings potential of different renovations measures but also their cost-effectiveness and other potential financing options. By providing their building data characteristics (as displayed in detail in 6.1.4) and filling in a checklist with their actions of preferences building owners can get back a detailed renovation roadmap facilitating them with the decision-making process and reassuring them about the effectiveness of the measures.

6.1.3 Development Progress

The service has been successfully applied to two pilot buildings, with two additional applications currently in progress. Specifically, it has been implemented in two IEECP pilot buildings, one located in Attica, Greece, and another in the Netherlands. Data have been currently collected for two more IEECP buildings. Furthermore, have contacted the FOCCHI pilot to gather some missing information regarding the building characteristics (e.g., the total aera of the walls of the buildings) in order to proceed with the simulations. Finally, discussions have been initiated with the WP3 leader (NTUA) in order to start working also on the interface that will allow the users to use the service.

6.1.4 Application on DigiBUILD Pilots

〉 **IEECP pilot – Attica, Greece**

6.1.4.1 Pilot characteristics

The first application of s.3.4.1 for the case of the IEECP pilot concerns an apartment located in Attica, Greece. The building specifications were collected through a template (presented in detail in the DigiBUILD Deliverable 3.1 *First wave' of DigiBUILD AI-based data-driven services for the built environment*) and are displayed below. In addition, data related to the energy consumption of the apartment (e.g., data from an electricity bill) were shared with the modelling team to facilitate establishing the baseline consumption of the building/apartment and for calibration purposes.

Table 14: Characteristics of the IEECP pilot in Attica.

Pilot:	**IEECP - Attica Pilot**
Location and climate characteristics	
Country:	Greece
Region (e.g., town, municipality, etc.):	Attica
Climate/ Climate Zone:	Climate Zone B
Building Characteristics	
Type of building/ usage:	Apartment
Year of Construction or Renovation:	1985
Building size:	1 basement level and 4 ground levels
Total Floor area of the building	125 m^2
Total area of external walls of the buildings:	52.5 m^2
Total area of Wall1:	24 m^2
Total area of Wall2:	24 m^2
Total area of Wall3:	4.5 m^2
Total Roof area of the building:	0 m^2
Windows system:	
Total windows area:	24 m^2
Total area of Window1:	6.6 m^2
Total area of Window2:	6.6 m^2
Total area of Window3:	2.75 m^2
Total area of Window4:	2.75 m^2
Total area of Window5:	2.75 m^2
Total area of Window6:	1.3 m^2
Total area of Window7:	1.3 m^2
Building envelope features/ Construction features (U-values)	

Uwall:	1 W/m²/K
Ufloor:	2 W/m²/K
Ufloor:	2 W/m²/K
Uwindow:	3 W/m²/K
Building system	
Heating system:	Central Oil Boiler
Nominal capacity:	13kW
Cooling system:	AC Split
Nominal capacity:	8.9 kW
EER (if available):	5
Lighting equipment:	46 lightbulbs
Lighting equipment capacity:	30 kW
Other Parameters	
Occupancy:	2 people
Operating schedule:	24hrs

〉

We have started the application by using the data presented above to create a simulation model of a residential building in Attica, Greece as presented in Figure 60.

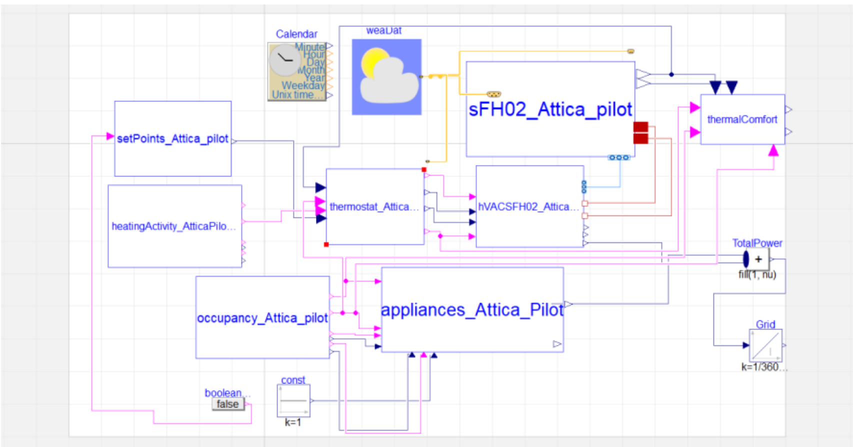


Figure 60: Simulation environment in Dymola showing the model of the IEECP pilot in Attica, Greece.

6.1.4.2 Analysis and results

After calibrating the model and deriving the baseline energy consumption of the apartment we analyse the effect of the implementation of different energy efficiency measures (EEM). These include:

- **EEM1**: building envelop refurbishment (this includes windows replacement with more energy-efficient glazing, and exterior wall insulation)
- **EEM2**: replacement of the oil boiler with a gas boiler

Funded by
the European Union

- **EEM3**: replacement of the oil boiler with a heat pump

The total annual energy consumption of the apartment for the baseline scenario as well as after the implementation of the measures are presented in the following graphs. In addition, the energy consumption is disaggregated to the thermal energy consumption and the energy used for cooling and operation of other appliances.

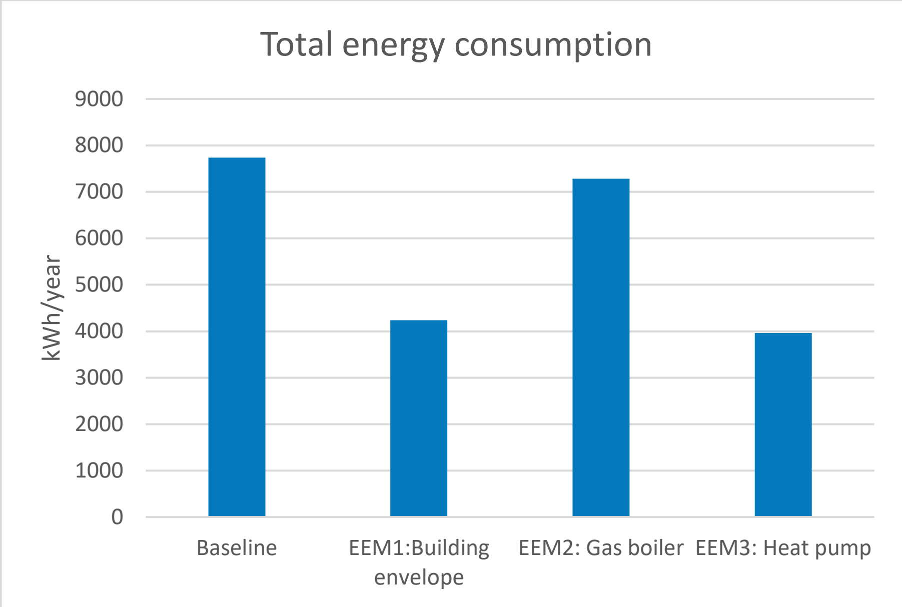


Figure 61: Total energy consumption for the baseline scenario and after the implementation of EEM1, EEM2, and EEM3

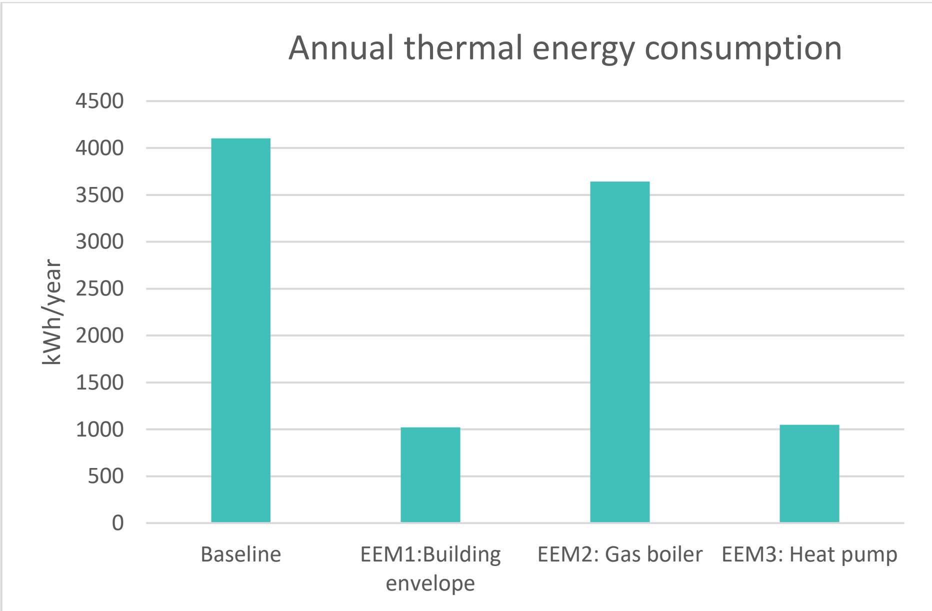


Figure 62: Annual thermal energy consumption for the baseline scenario and after the implementation of EEM1, EEM2, and EEM3

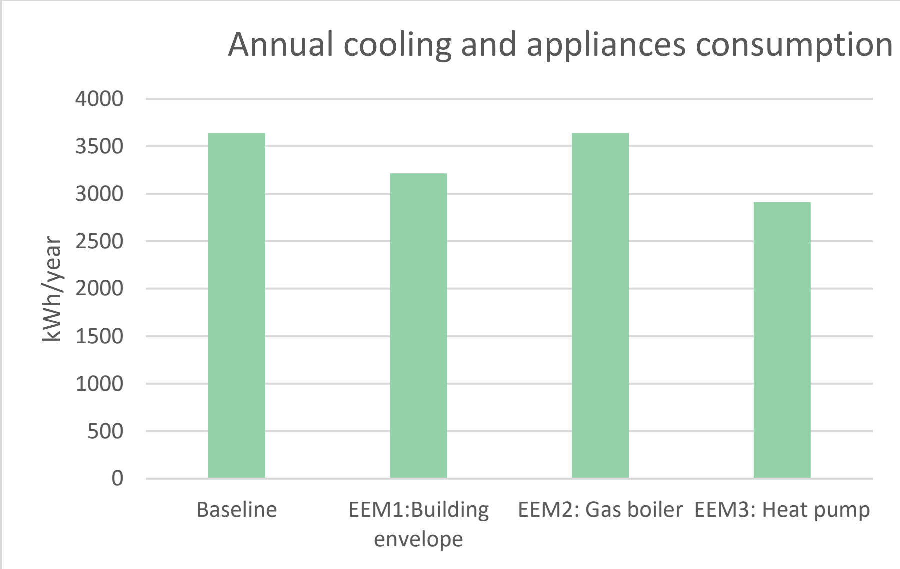


Figure 63: Annual energy used for total cooling purposes and operation of other appliances for the baseline scenario and after the implementation of EEM1, EEM2, and EEM3

Furthermore, the following graphs present the average daily total, heating and cooling and other appliances energy consumption respectively.

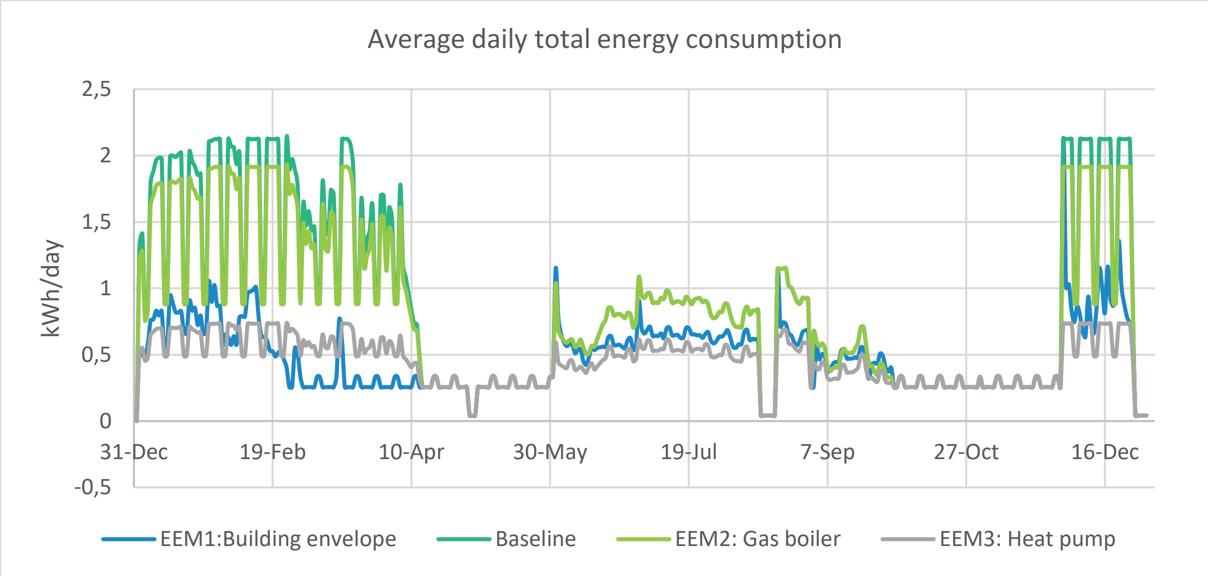


Figure 64: Average daily energy consumption for the baseline scenario and after the implementation of EEM1, EEM2, and EEM3

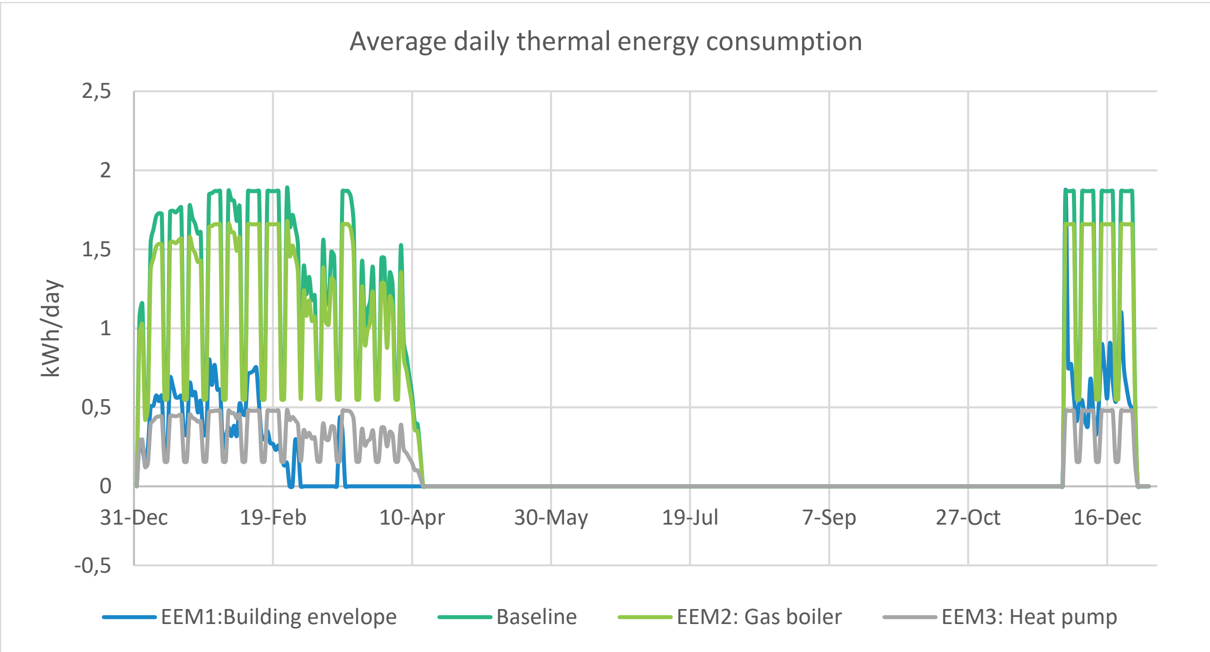


Figure 65: Average daily thermal energy consumption for the baseline scenario and after the implementation of EEM1, EEM2, and EEM3

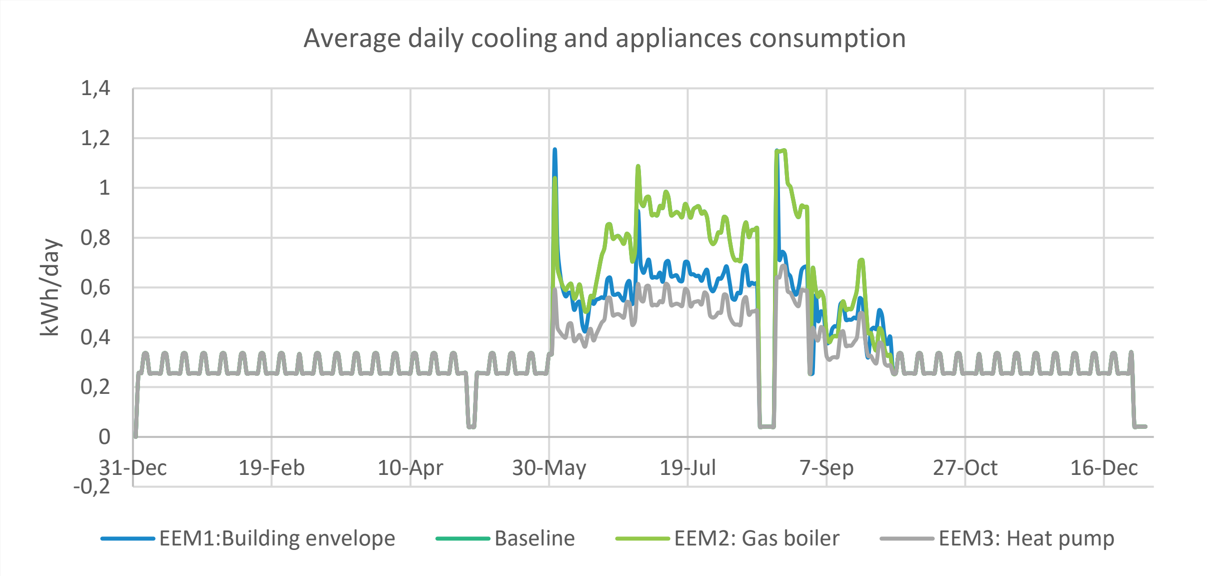


Figure 66: Average daily energy use for cooling and operation of other appliances for the baseline scenario and after the implementation of EEM1, EEM2, and EEM3.

In addition to the energy savings derived from the different energy efficiency measures, we perform a technoeconomic assessment analysis to evaluate the economic viability of the measures. For this analysis we have chosen three indicators, the net present value (NPV), the simple payback period (PP) and the levelized cost of saved energy (LCSE).

The NPV is calculated using the following formula:

$$NPV = \sum_{i=0}^{\tau} \left(\frac{CF_i}{(1+d)^i} \right) \tag{13}$$

Where:

- τ is the calculation period or the lifetime of the measure
- d is the discount rate
- CF_i is the annual cash flow in year i; $CF_i = \Delta cost_{energy,i} + \Delta cost_{om,i} - I_i$,

  Where:

  - $\Delta cost_{energy,i}$ is the energy cost savings in year i
  - $\Delta cost_{om,i}$ is the change of annual operation and maintenance cost in year i
  - I_i is the investment cost in year i

The payback period (PP) is calculated using the following formula:

$$PP = \frac{Initial\ investment}{Annual\ energy\ savings} \tag{14}$$

The LCSE is calculated using the following formula:

Funded by the European Union

$$LCSE = \frac{Discounted\ Cash\ Flow}{Discounted\ Energy\ Savings} = \frac{I_0 * CRF - \Delta cost_{energy} - \Delta cost_{om}}{\Delta E} = \frac{-NPV}{\sum_{i=0}^{\tau}\left(\frac{\Delta E}{(1+d)^i}\right)} \quad (15)$$

Where:

- I_0 is the initial investment
- $\Delta cost_{energy}$ is the annual energy cost savings
- $\Delta cost_{om}$ is the change in annual operation and maintenance cost
- CRF is the cost recovery factor; $CRF = \frac{d*(1+d)^\tau}{(1+d)^\tau - 1}$,

  Where

  - τ is the calculation period or the lifetime of the measure
  - d is the discount rate
- ΔE is the annual energy savings

For the evaluation of the three alternative measures, we use two different discount rates, a 7% discount rate as used in the Greek long-term renovation strategy 2020[1], and a 4% discount rate as suggested by the Commissions Impact Assessment guidelines[2].

The results are presented in Table 15.

Table 15: Technoeconomic analysis results for the three energy efficiency measures with 7% and 4% discount rates

	Lifetime (years)	Investment cost (€)	Net Present Value (Disc. Rate `7%) (€)	Net Present Value (Disc. Rate 4%) (€)	Payback period (Years)	Levelized cost of saved energy (Disc. Rate 7%) (€)	Levelized cost of saved energy (Disc. Rate 4%)(€)
EEM1: Building envelope	30	5103[3]	-62.59	1920.82	12.56	0.001	-0.032
EEM2: Gas boiler	15	3910[4]	-5,064.82	-5,319.73	>20	1.206	1.038
EEM3: Heat pump	20	6700	-2,123.89	-829.62	18.45	0.065	0.020

We observe that the NPV is negative, thus the investment is not cost-effective, for a discount rate of 7% for all three measures. For the refurbishment of the building envelope and for changing the oil boiler

[1] https://energy.ec.europa.eu/system/files/2021-08/el_2020_ltrs_en_version_0.pdf
[2] https://energy.ec.europa.eu/system/files/2020-07/fr_ltrs_2020_en_0.pdf
[3] The investment costs for the measures included in EEM1 have been derived from the Commission's Comprehensive study of building energy renovation activities and the uptake of nearly zero-energy buildings in the EU https://op.europa.eu/en/publication-detail/-/publication/97d6a4ca-5847-11ea-8b81-01aa75ed71a1/language-en/format-PDF/source-119528141
[4] The investment costs for EEM2 and EEM3 are derived from the Comparison study of heating costs from different technologies from the Laboratory of Steam Boilers & Thermal Plants, NTUA and the Thermal Processes Laboratory, NTUA

with heat pump the negative result is mostly explained due to the high initial investment costs required. On the other hand, when it comes to changing the oil boiler with natural gas boiler, our results suggest that the low energy savings and the increased costs of natural gas compared to oil affect greatly the investment. The corresponding results for the 4% discount rate indicate a positive NPV for the refurbishment of the building envelope and a negative NPV for EEM2 and EEM3.

It is evident from our analysis that high investment costs often make energy efficiency interventions cost-inefficient especially when the renovation includes measures such as heating system replacement and envelope interventions. This highlights the need for public subsidies and also the need for new tools and mechanisms to mobilise private financing, in order to allow the uptake and implementation of all these proposed energy efficiency interventions in buildings.

In Greece, the programme "Exoikonomo 2023[5]" was running from June 2023 until September 2023 and is the continuation of the successful programmes "Exoikonomo autonomo", "Energy Saving at Home II" and "Energy Saving at Home I". The fund of the "Exoikonomo 2023" programme will contribute to the energy saving of at least 213 killotonnes of oil equivalent and the energy renovation of at least 105,000 residencies by 2025. Special support is foreseen for energy poor and vulnerable households in the form of an increased rate of grants and a separate budget of 60 million €. The subsidy can cover up to 75% of the total investment for the renovation depending on the income of the individual or the family and the typology of the building (e.g., older buildings, or buildings with low energy performance based on their EPC indication are prioritised). While the lowest contribution is at the level of 40% for individuals or families belonging to the highest income category.

In the following tables (Table 16 and Table 17) the results of the technoeconomic analysis taking into consideration the highest level of subsidation (75%) and the lowest level of subsidation (40%) are presented.

Table 16: Technoeconomic analysis results taking into account 75% subsidisation of the investment cost

	Capital cost required after subsidy (75%) (€)	Net Present Value (Disc. Rate `7%) (€)	Net Present Value (Disc. Rate 4%)(€)	Payback period (Years)	Levelized cost of saved energy (Disc. Rate 7%) (€)	Levelized cost of saved energy (Disc. Rate 4%)(€)
EEM1: Building envelop	1275.75	3,764.66	5,748.07	3.14	-0.086	-0.095
EEM2: Gas boiler	977.5	-2,132.32	-2,387.23	17.97	0.508	0.466
EEM3: Heat pump	1675	2,901.11	4,195.38	4.61	-0.089	-0.100

Table 17: Technoeconomic analysis results taking into account 40% subsidisation of the investment cost.

	Capital cost required after subsidy (40%) (€)	Net Present Value (Disc. Rate `7%) (€)	Net Present Value (Disc. Rate 4%) (€)	Payback period (Years)	Levelized cost of saved energy (Disc. Rate 7%) (€)	Levelized cost of saved energy (Disc. Rate 4%) (€)

[5] https://exoikonomo2023.gov.gr/to-programma

EEM1: Building envelop	3061.8	1,978.61	3,962.02	7.54	-0.045	-0.065
EEM2: Gas boiler	2346	-3,500.82	-3,755.73	>20	0.834	0.733
EEM3: Heat pump	4020	556.11	1,850.38	11	-0.017	-0.044

As concluded from the results of the analysis for both of the subsidation cases the investment with own capital for EEM1 and EEM3 is cost-effective, verifying the importance of public subsidies. The only measure that is not cost-effective is EEM2 (the replacement of the oil heating system with a gas boiler).

Considering our results these investments prove to be not viable due to the current high gas prices, making them less attractive also for consumers. However, building owners and/or users often consider only the initial/ capital investment needed from their side when it comes to financing renovation measures since they may have limited access to technoeconomic assessments or whole renovation plans for their buildings to facilitate their decision-making process. This highlights the need for the DigiBUILD services and specifically s3.4.1.

IEECP pilot – Hague, Netherlands

6.1.4.3 Pilot characteristics

The second application of s.3.4.1 for the case of the IEECP pilot concerns an apartment located in Netherlands. The building specifications are displayed below. Again, in this case data related to the energy consumption of the apartment (e.g., data from an electricity bill) were shared with the modelling team to facilitate establishing the baseline consumption of the building/apartment and for calibration purposes.

Table 18: Characteristics of the IEECP pilot in Netherlands.

Pilot:	IEECP - Netherlands Pilot
Location and climate characteristics	
Country:	Netherlands
Region (e.g., town, municipality, etc.):	Village
Climate/ Climate Zone:	Moderate maritime (or oceanic) climate
Building Characteristics	
Type of building/ usage:	House/ Office
Year of Construction or Renovation:	1936
Building size:	3 floors
Total Floor area of the building	125 m^2
Total area of external walls of the buildings :	110 m^2
Total area of Wall1:	30 m^2
Total area of Wall2:	60 m^2
Total area of Wall3:	20 m^2

Total Roof area of the building:	180 m^2
Windows system:	
Windows system	Combination of Plastic and Wooden Framed, Double glass windows (10mm thick)
Total windows area:	33.8 m^2
Total area of Window1:	3 m^2
Total area of Window2:	8 m^2
Total area of Window3:	5.2 m^2
Total area of Window4:	0.4 m^2
Total area of Window5:	2.4 m^2
Total area of Window6:	4 m^2
Total area of Window7:	6 m^2
Total area of Window8:	4.8 m^2
Building envelope features/ Construction features (U-values)	
Uwall:	1.5 $W/m^2/K$
Ufloor:	1.7 $W/m^2/K$
Ufloor:	1.6 $W/m^2/K$
Uwindow:	1.2 $W/m^2/K$
Building system	
Heating system:	Gas boiler
Nominal capacity:	8,5 - 35,7 kW
Lighting equipment:	45 LED bulbs
Lighting equipment capacity:	350 W
Other Parameters	
Occupancy:	3 people
Operating schedule:	24hrs

The simulation model of the building in the Netherlands is presented below.

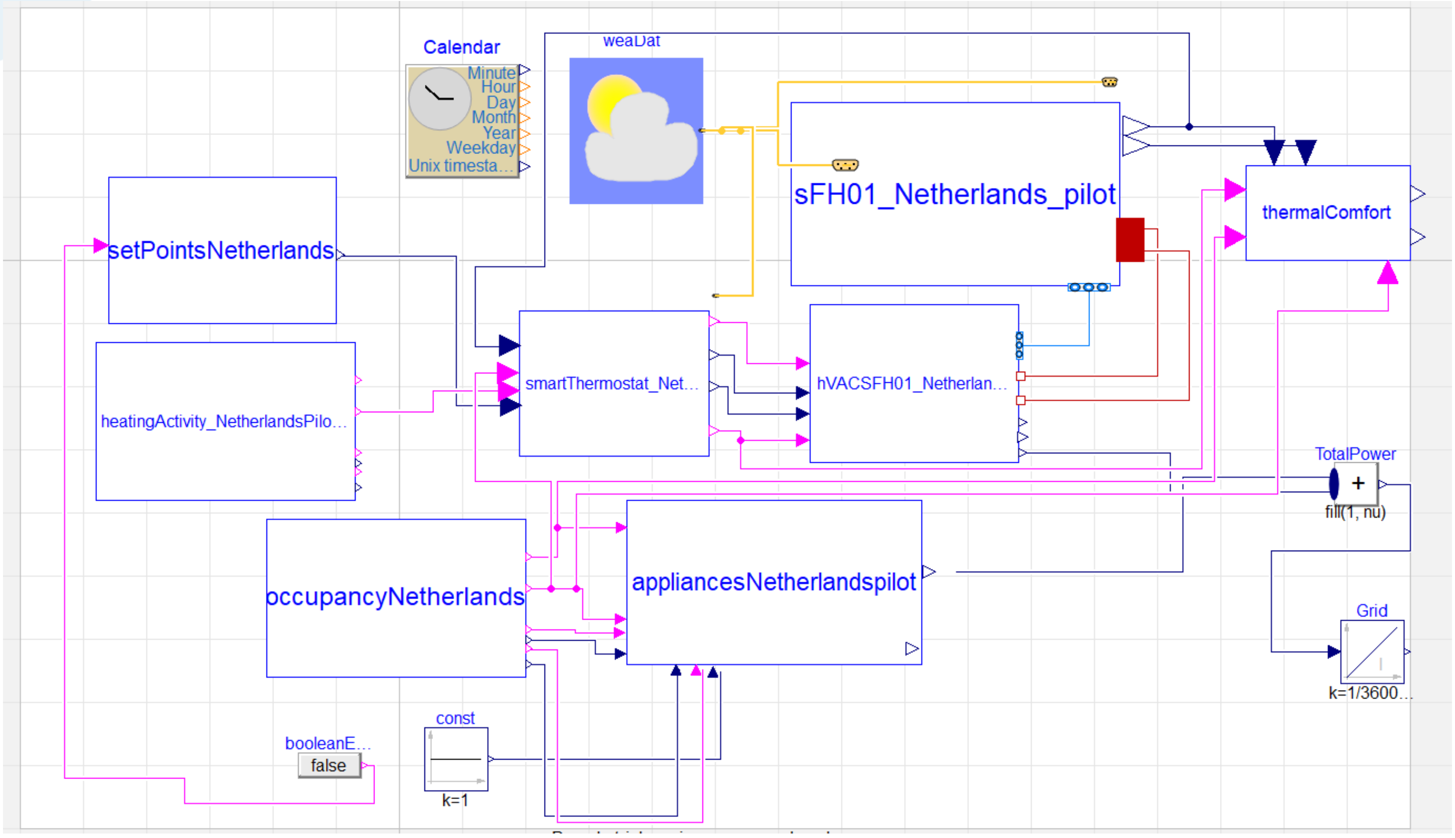


Figure 67: Simulation environment in Dymola showing the model of the IEECP pilot in Netherlands.

6.1.4.4 Analysis and results

After calibrating the model and deriving the baseline energy consumption of the apartment we analyse the effect of the implementation of different energy efficiency measures. After consultation with the building owner, these include:

- **EEM1**: Refurbishment of external walls
- **EEM2**: Replacement of double-glazed windows with triple-glazed
- **EEM3**: Refurbishment of the roof
- **EEM4**: Replacement of the gas boiler with a heat pump
- **EEM5**: Total building envelope refurbishment

The total annual energy consumption of the apartment for the baseline scenario as well as after the implementation of the measures are presented in the following graphs. In addition, the energy consumption is disaggregated to the thermal energy consumption and the energy used for cooling and operation of other appliances.

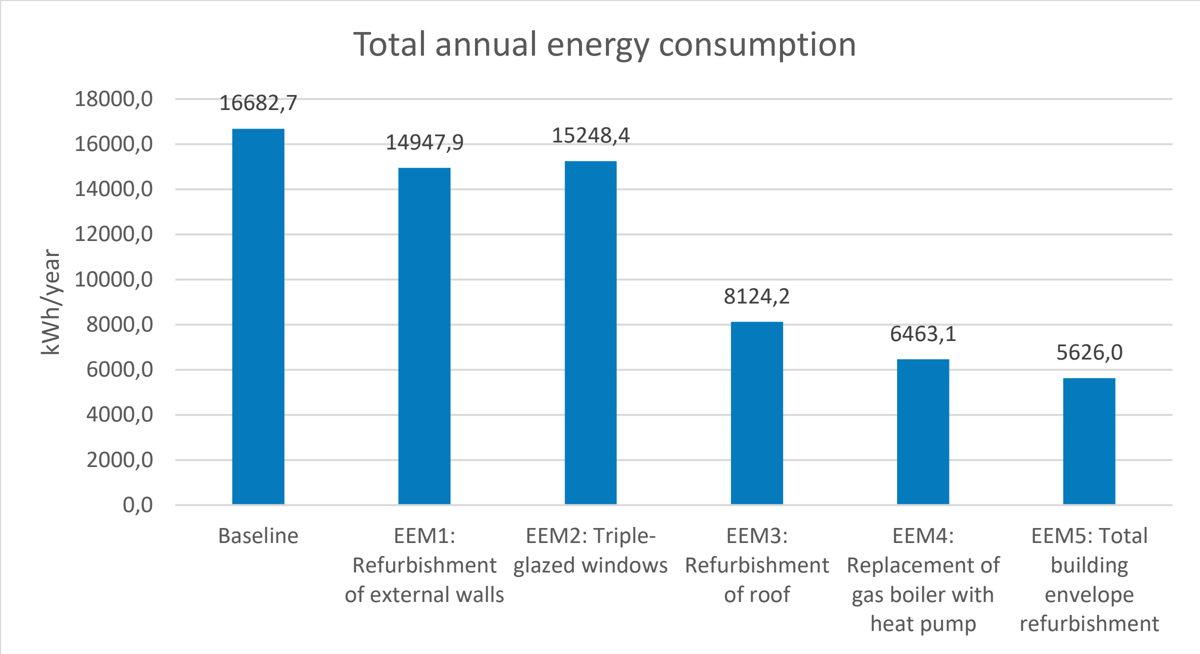


Figure 68: Total annual energy consumption for the baseline scenario and after the implementation of the energy efficiency measures.

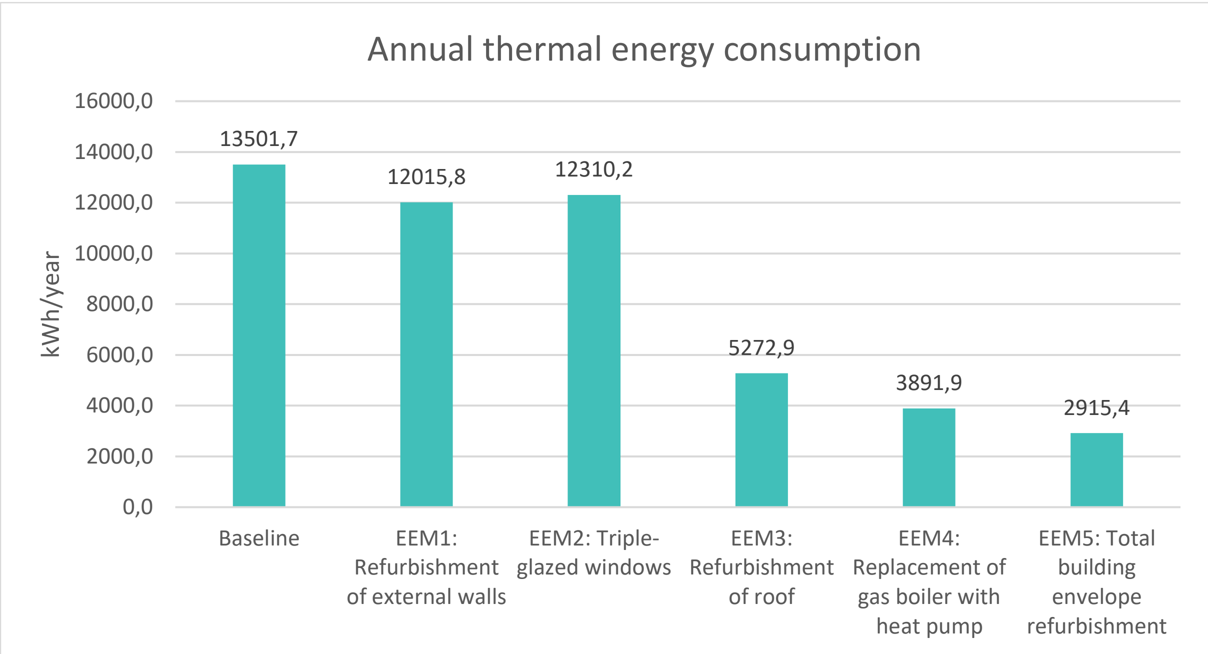


Figure 69: Annual thermal energy consumption for the baseline scenario and after the implementation of the energy efficiency measures.

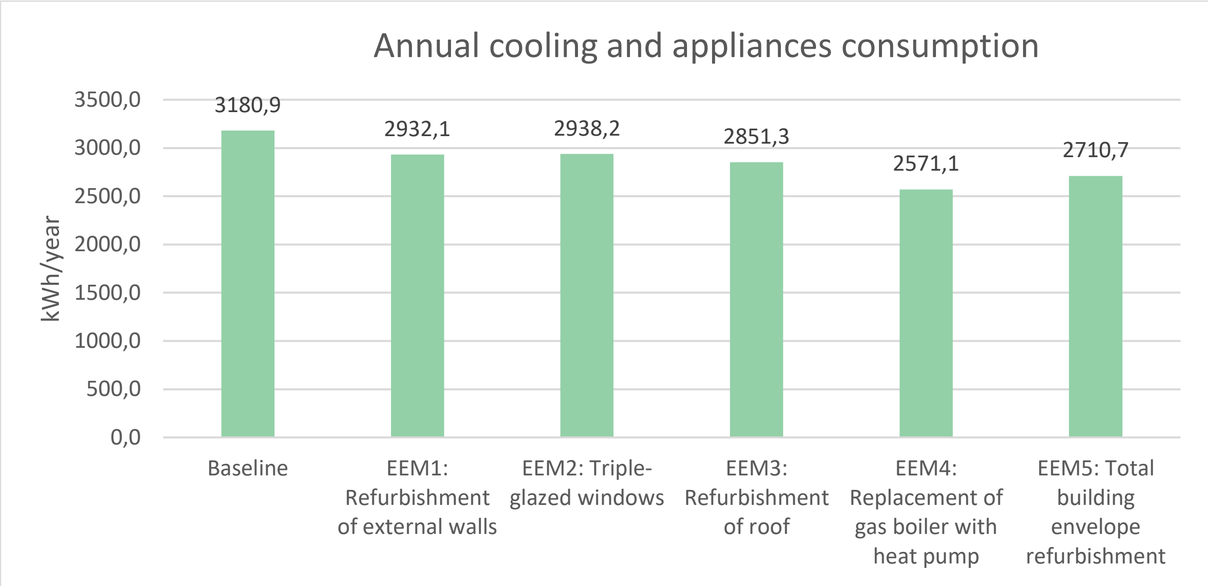


Figure 70: Annual energy used for total cooling purposes and operation of other appliances for the baseline scenario and after the implementation of the energy efficiency measures.

Furthermore, the following graphs present the average daily total, heating and cooling and other appliances energy consumption respectively.

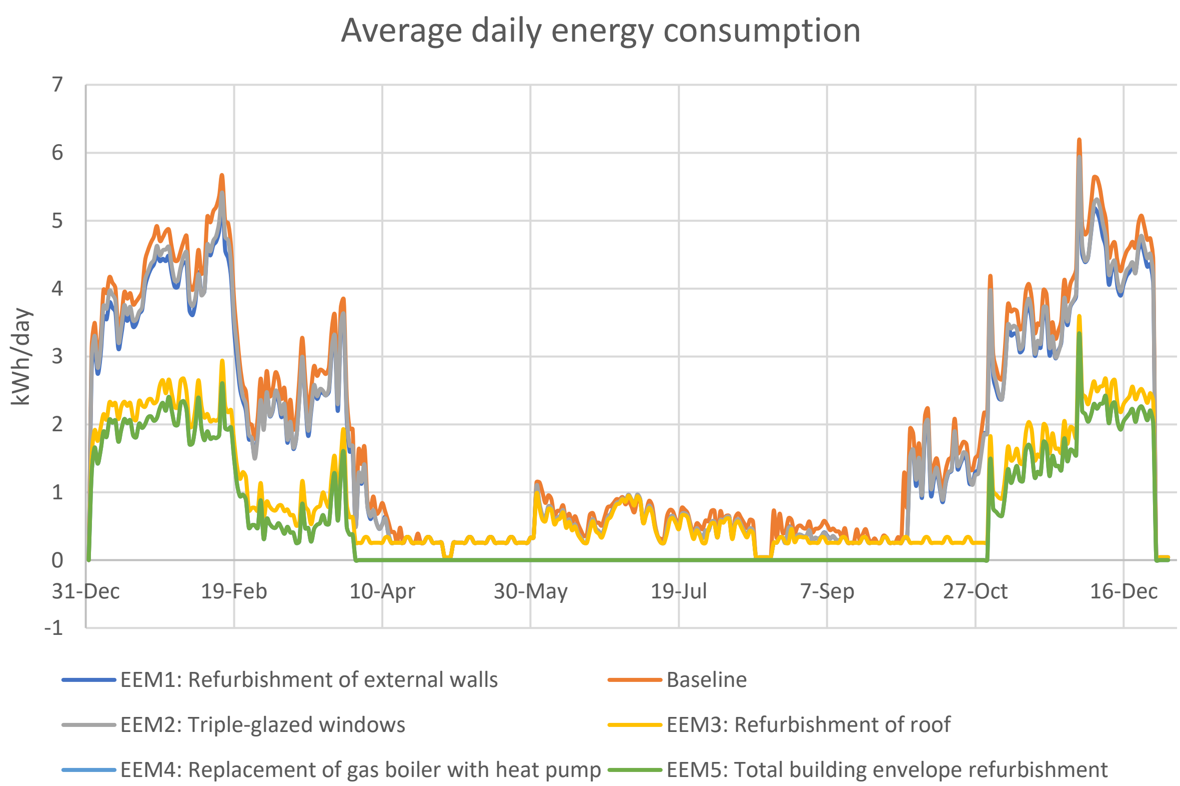


Figure 71: Average daily energy consumption for the baseline scenario and after the implementation of the energy efficiency measures.

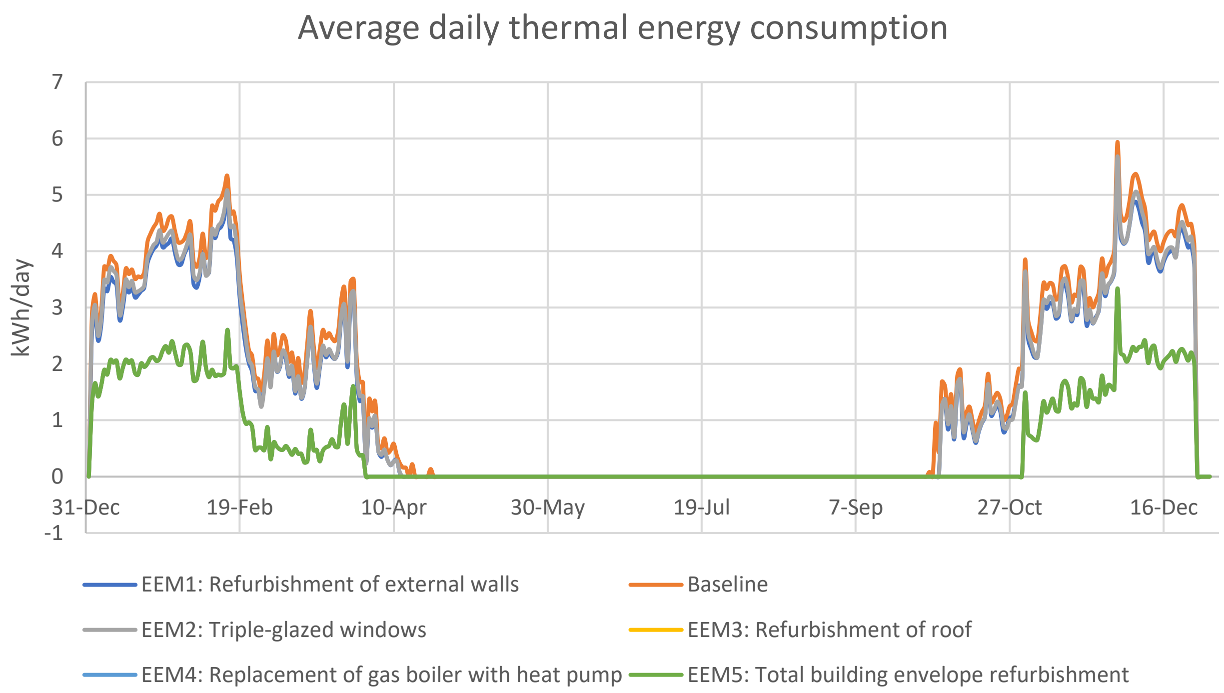


Figure 72: Average thermal energy consumption for the baseline scenario and after the implementation of the energy efficiency measures.

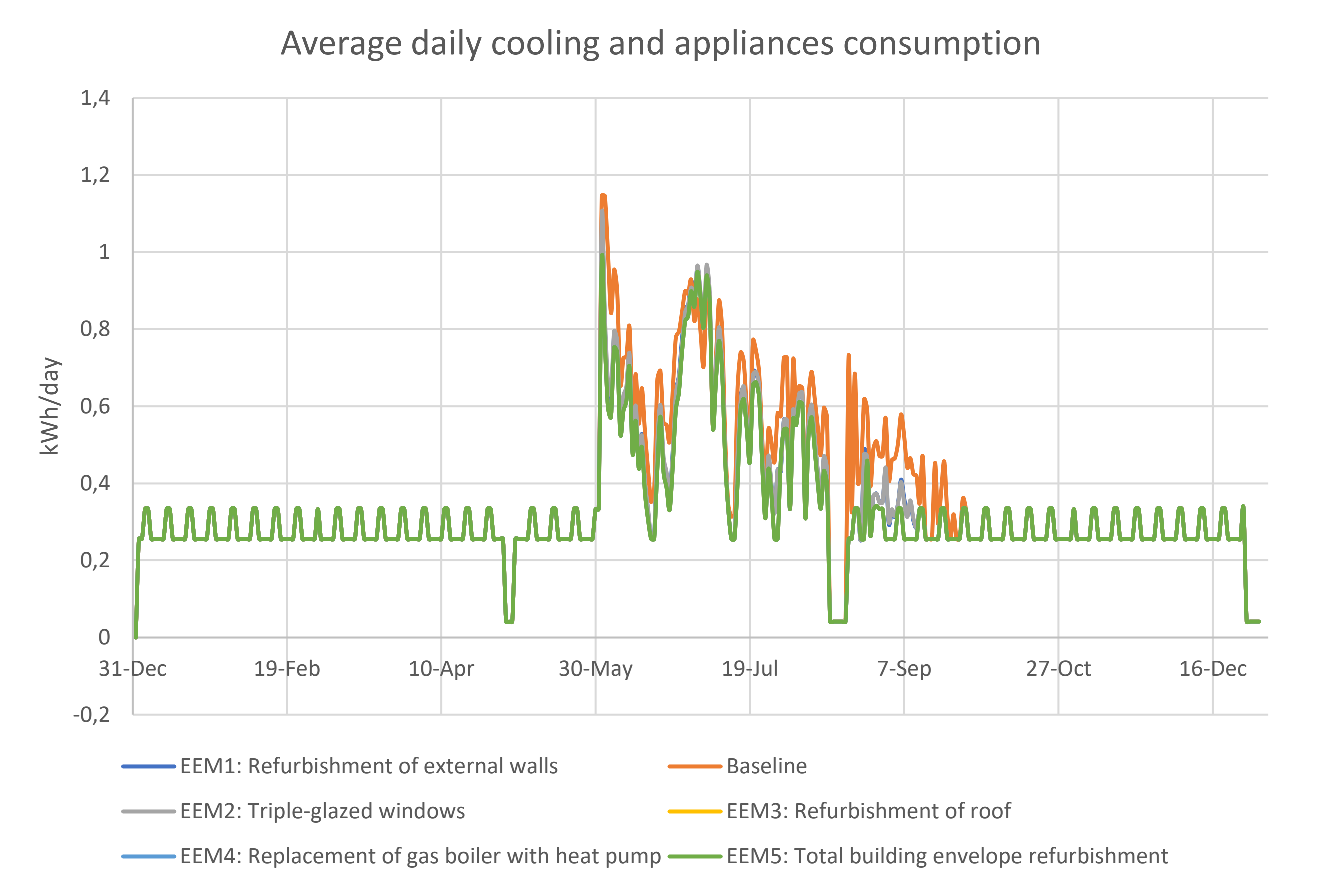


Figure 73:Average cooling and appliances energy consumption for the baseline scenario and after the implementation of the energy efficiency measures.

The results of the technoeconomic analysis are presented in the following table:

Table 19: Technoeconomic analysis results for the three energy efficiency measures with 7% and 4% discount rate

	Lifetime (years)	Investment cost[6] (€)	Net Present Value (Disc. Rate `7%) (€)	Net Present Value (Disc. Rate 4%) (€)	Payback period (Years)	Levelized cost of saved energy (Disc. Rate 7%) (€)	Levelized cost of saved energy (Disc. Rate 4%)(€)
EEM1	30	3750	-1,265.5	-287.78	18	0.174	0.125
EEM2	30	3380	-1,334	-529.00	20	0.189	0.136
EEM3	30	3657	8800	13,702.98	3.65	0.034	0.024
EEM4	20	6875	7,784	9,364.55	5.75	0.063	0.049
EEM5	30	8874	-7210.8	13,540.20	6.80	0.064	0.046

6 The investment costs for the measures included in Table X have been derived from the following study: https://www.sciencedirect.com/science/article/pii/S0306261917317816?via%3Dihub

We observe that the NPV is negative, thus the investment is not cost-effective, for the first two measures (refurbishment of external walls and replacement of windows) in both cases of 7% and 4% discount rate. On the other hand, when it comes to refurbishing the roof, replacing the gas boiler with heat pump or going for a total envelope renovation NPV values are positive in both cases, and we also notice very short payback periods.

In Netherlands there are two Incentive schemes for energy savings in the rental housing sector:

- The Energy Performance Incentive Scheme for the Rental Sector (STEP) enables landlords to improve the energy performance of their rental properties.
- The Energy Savings Fund for the Rental Sector (FEH) offers low-interest loans for landlords to make their rental properties more energy-efficient.

The ISDE is an investment subsidy with a one-off payment that runs until 2030. Applications can be made from 3 January to 31 December 2024. The amount of subsidy for which one can be eligible depends on type of intervention and the energy-savings potential and is generally about 20% of the investment amount.

6.1.5 Next Steps

In order to enter the full pilot operation phase at M21, the next steps should be taken into account for s.3.4.1:

1. Apply the service on the remaining pilot building of IEECP and FOCCHI.
2. Start developing the interface to collect the building owners/ users data in an automatic way.
3. Test the interface system in the pilot sites.
4. Setting up a standardized renovation roadmap that will be fed back to the users based on the feedback received from the applications.

6.2 One-stop-shop energy efficiency hub (s3.4.2)

6.2.1 Description of the Service

The primary objective of the one-stop shop energy efficiency hub is to generate a framework designed to automate the generation of optimal building renovation roadmaps, complete with cost analyses. This aids in decision-making for optimal financing of building renovations. The framework takes into account not only energy and economic factors but also environmental and social considerations. The platform aims to enable users, through straightforward steps, to receive a list of energy efficiency actions. These actions are intended to enhance both the energy performance and the economic return of their buildings while also making them more beneficial from an environmental and social point of view. A prerequisite for utilising the tool's functionality is the possession of a correct and valid Building Energy Performance (BEP) model of the building in the .idf format, along with a corresponding .epw file containing local weather data. Additionally, knowledge of the cost of the actions to be examined through the application is required.

As detailed in the previous Deliverable D3.1, the service is divided into four fundamental pillars: data collection, iterative scenario simulation, evaluation, and ranking. Each of these pillars, or phases, will be analysed further.

Data Collection

The goal of this pillar is to gather the appropriate files from the user, analyse the baseline BEP file provided by the user, and collect pertinent information from it for subsequent use. Thus, in Figure 74, the data flow in this component is outlined. In this phase, the user submits their building's BEP file (.idf) along with a suitable weather file (.epw). These are collected and examined for their correctness and validity. Initially, the file is checked for its geometry and surfaces. The BEP file's ability to run with the simulation engine used (EnergyPlus), along with the provided weather data, is then verified. Once these files pass the checks, data related to the building's construction (constructions, materials, orientation, etc.) and its energy requirements are collected. The BEP file is then analysed using suitable Python scripts to collect data on the surface and construction of external walls, roofing, and windows. This will assist later in evaluating the cost of replacing these with more efficient constructions. Concurrently, the user is asked about the aspects of the building they wish to modify, presenting them with a catalogue of available actions and their estimated construction costs. These data are collected, thus completing the data collection phase. The next step is the iterative scenarios simulation.

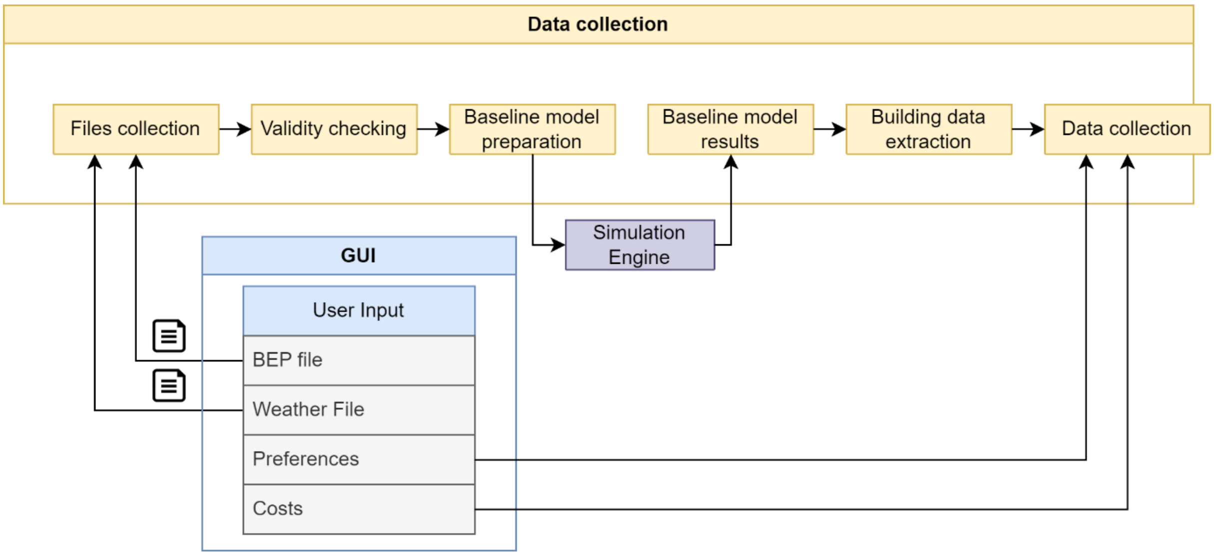


Figure 74: s3.4.2 - Data collection and validation

Iterative simulation

In this phase, various alternative Building Energy Performance (BEP) models are generated, based on the user's selections from the available energy efficiency actions. These models are essential for running simulations and gathering energy data of each alternative BEP model. Consequently, it is apparent that this methodology contemplates multiple distinct scenarios per building. Naturally, this list is subject to further research and expansion. illustrates the workflow of the iterative scenario simulation.

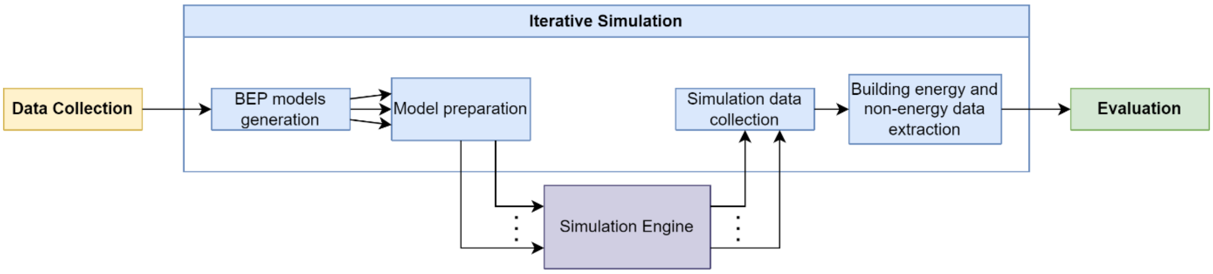


Figure 75: s3.4.2 - Iterative simulation flow

The initial step involves the collection of user preferences and the construction of different BEP models. These models are designed to encapsulate all feasible combinations of energy upgrades in the building. Once constructed, they are submitted to the simulation engine to execute the simulations. Upon the completion of the simulations, the building's energy requirements are collected, mirroring the baseline methodology. Specifically, the data collated includes the total energy demand for the simulation period, the energy demand for heating and cooling, the energy sources utilised for these purposes, and the PMV (Predicted Mean Vote) and PPD (Predicted Percentage of Dissatisfied) indices, which pertain to comfort levels during the simulation period. After the collection of this information, a second phase of calculations follows to determine the costs of meeting energy needs and the carbon dioxide emissions associated with them. This marks the completion of the iterative simulation phase, leading us to the evaluation phase to scrutinise the results.

Evaluation

At this stage, the evaluation of the Building Performance Indicators (BPIs), as termed in this service for ease of reference, is conducted. The Building Performance Indicators (BPIs) have been streamlined from what was presented in Deliverable D3.1. This revision aims to represent the value of each criterion more accurately and to eliminate any overlaps. Final BPIs are depicted in Table 20, categorised accordingly. Some indicators necessitate additional calculations. The methodologies for these calculation formulas, are presented below.

Table 20: Building Performance Indicators to be evaluated.

BPI	ID	Type	Formula	Units
Total Energy Consumption annually	BPI_1	energy	$\sum(Energy_{per\ source})$	J/y
Heating/Cooling Total Energy Consumption annually	BPI_2	energy	$Energy_{heating} + Energy_{cooling}$	j/y
Energy savings annually	BPI_3	energy	$Energy_{before} - Energy_{after}$	J/y

Energy used from RES	BPI_4	energy	$\frac{Energy_{produced}}{Energy_{consumed}}$	%
Energy used from HC	BPI_5	energy	$\frac{\sum(Energy_{per\ HC\ source})}{\sum(Energy_{per\ source})}$	%
Net Present Value	BPI_6	economic	$\sum_{t=1}^{5} \frac{\text{Net Cash Flow at Time t}}{(1+r)^t}$, $where\ r\ the\ discount\ rate$	€
Payback Period	BPI_7	economic	$\frac{Initial\ Investment}{Annual\ Cash\ flow\ from\ energy\ savings}$	years
Discounted Payback Period	BPI_8	economic	$Initial\ Investment = \sum_{t=1}^{5} \frac{\text{Net Cash Flow at Time t}}{(1+r)^t}$	years
Internal Rate of Return (IRR)	BPI_9	economic	$0 = \sum \frac{\text{Net Cash Flow at Time t}}{(1+IRR)^t}$	%
Cost of Effectiveness	BPI_10	economic	$\frac{Initial\ Investment}{Annual\ energy\ savings}$	€/J
Energy source diversity	BPI_11	social	$\frac{1}{2N^2\bar{y}} \sum_{i=1}^{N} \sum_{j=1}^{N} \lvert y_i - y_j \rvert$ $Where$: n: $number\ or\ energy\ sources$ $\bar{y}$: the average proportion of total energy consumption per source y_i, y_j: the proportions of total energy consumption for sources $\boldsymbol{i}$ and $\boldsymbol{j}$, respectively	0 to 1
Predicted Mean Vote	BPI_12		-	-3 to +3
Therman Comfort Satisfaction Rate	BPI_13	social	-	%
Mean UA Value	BPI_14	social	$\frac{\sum U - Value \times A_{floor}}{Total\ A_{floor}}$ $where\ A\ the\ area.$	W/°C
Non-HC Energy Percentage	BPI_15	environmental	$\frac{Energy_{total} - Energy_{H/C}}{Energy_{total}} \times 100\%$	%

Renewable Energy Capacity	BPI_16	environmental	-	J
Average annual CO_2 emissions	BPI_17	environmental	$\sum f_{emission} \times Energy_{per\ source}$ $where\ f_{emission}:\ CO_2\ emmission\ factor$	tn

Once all BPIs have been computed, the process culminates with the optimisation phase.

Ranking

This is the final phase where the optimal scenarios are generated through their evaluation based on the BPIs. To make this possible, a multi-criteria decision analysis system was devised to evaluate the different alternative scenarios. As is obvious, alternatives are the different configurations of possible interventions and criteria are the BPIs. The weighting of the criteria was done with the help of the pilot as it has experts in the field of energy management and utilisation. Bilateral significance relationships between the criteria were completed by the pilot energy experts and are shown in Table 31. Based on this table the weights for each criterion were calculated by utilizing the Analytical Hierarchy Processes method. Consequently, the weights of the criteria are shown in Table 32 and Table 33.

The TOPSIS method was chosen for the evaluation of the alternatives for its computational efficiency, Comprehensive and Logical Approach. This method ranks the different alternatives for us. The top ten resulting from it are presented to the user as the best scenarios for energy upgrades in their building. The flow of ranking along with evaluation phase is presented in Figure 76.

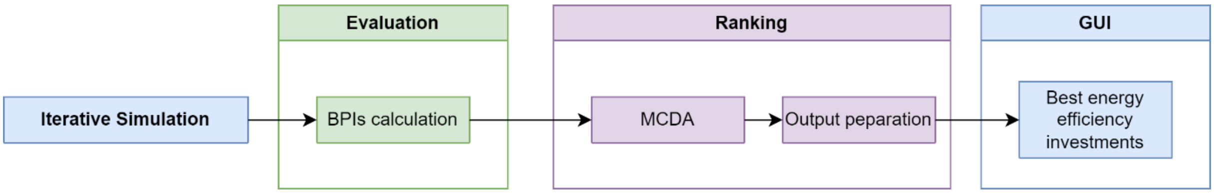


Figure 76: Evaluation and ranking flow.

6.2.2 Novelty

Numerous successful attempts have been made in the past to develop models and frameworks that utilise simulation data to identify optimal energy upgrade measures for buildings. Despite this, the concept is not widely adopted within the building industry. A significant reason for this is that most buildings lack up-to-date Building Information Models and BEP models. Nevertheless, in the current era of digital transformation in the building sector, a promising foundation appears to be emerging for the advancement of such methodologies. Furthermore, the strategy employed in DigiBUILD seeks to incorporate considerations beyond the traditional focus on energy and economics. This is achieved through the introduction of appropriate indicators capable of capturing the significance of factors such as comfort, energy poverty, and the diversity of energy sources. Consequently, DigiBUILD presents a renewed endeavour and an innovative approach to the computation of energy upgrades in buildings.

6.2.3 Development Progress

Regarding the progress of the service, the data collection component is partially complete. It efficiently

Funded by
the European Union

gathers essential data from users, validates this information, and acquires baseline simulation data. The most significant segment yet to be finalised pertains to the iterative simulation pillar. This phase is particularly challenging from a programming perspective due to its reliance on comprehensive modelling techniques, which has consequently protracted the overall implementation of the service. Specifically, the manipulation of baseline files and the generation of all possible simulations are ongoing tasks. Regarding the evaluation segment, as delineated in the service description, it is thoroughly developed. However, its integration with the iterative simulation component is pending. The fundamental alternative evaluation methodology, constituting the evaluation segment, has been meticulously constructed. This is poised to be amalgamated with the preceding elements, thereby completing the entire value chain. Ultimately, the Graphical User Interface (GUI) for the one-stop-shop is anticipated to amalgamate the iterative simulation, evaluation, and ranking segments. This will occur once the developing activities for the components are finalised, ensuring the complete integration of the service's full value chain.

The source code can be found on GitHub: GitHub s3.4.2

6.2.4 Application on DigiBUILD Pilots

According to the Description of Action (DOA), the service is being developed for implementation in Pilot 9 (NTUA), specifically for Use Case 31 (UC_31). Although the entire value chain of the tool has not yet been completed, and its completion is essential for the pilot's usability, the pilot has significantly contributed to its definition during the development phase. In particular, the pilot includes experts specialising in energy management and policy formulation. These experts have engaged in co-creation activities at various stages of the methodology development and data collection. Specifically, the technical developers of the service, in collaboration with the pilot's energy management experts, have jointly developed the criteria that will be used to calculate energy upgrades. Furthermore, individuals involved in the pilot have determined the significance of these criteria to ensure they accurately and optimally reflect the impact of each criterion based on its specific needs. Therefore, while the service has not yet been implemented in this pilot, it has been directly involved in shaping the specifications to meet its requirements and achieve its objectives.

6.2.5 Next Steps

The forthcoming stages in the service's development are primarily concerned with the integration of modelling techniques, which represent the most substantial portion of the workload. Following the completion of this process, it is imperative that the entire value chain of the methodology – encompassing data collection, iterative scenario simulation, evaluation, and optimisation – is thoroughly comprehended. Additionally, this methodology must undergo comprehensive testing. This testing, to be conducted by both the service developers and the Pilot, is crucial for ensuring the service's effective functioning in real-world conditions. This will be facilitated through the Graphical User Interface (GUI) which is, in parallel, under development.

Funded by
the European Union

7 Decision-making under uncertainty tools for efficient and climate resilient buildings

This chapter delves into the increasingly important topic of building resilience in the context of climate change. It presents a comprehensive approach to assessing and enhancing the resilience of buildings against a variety of climate-related risks.

The chapter introduces a novel service designed to evaluate the climate resilience of buildings. This service is critical in today's context, where global climates are constantly changing, posing new challenges to the built environment. The approach outlined in this chapter aims to provide a clear understanding of a building's vulnerabilities to specific climate hazards and suggests measures to bolster its resilience.

7.1 Efficient and climate resilient buildings (s3.5.1)

7.1.1 Description of the Service

As global climates undergo continuous evolution, the imperative to fortify built environments against environmental challenges is heightened. Therefore, it becomes imperative to thoroughly assess buildings' resilience to climate-related risks and acquire the essential foundation for understanding its vulnerabilities to specific climate hazards, hence enabling the implementation of appropriate measures to enhance its resilience. The proposed service introduces a comprehensive framework able to assess and evaluate the climate resilience of buildings. The initial stage of the climate resilience assessment for a building involves preparing a climate vulnerability and risk assessment for the reference building. To achieve this, we identify related climate hazards and evaluate the building's exposure to each of the outlined hazards, both current and future, and assess the vulnerability of the building to the identified hazards. Subsequently, this information is utilised to assign appropriate weights to the technical features of the building, which are considered in the overall climate resilience assessment of the reference building. The following step involves identifying the characteristics of the building infrastructure of the reference building. These features can be categorised into the building's structural characteristics, energy systems, water management systems, sustainability practices, and crisis management systems (in the case of workplaces and large buildings). The objective is to ensure that non-expert personnel can easily complete all the above tasks and draw conclusions regarding the climate resilience of a building. To achieve this, thorough research is conducted on the building characteristics that need to be considered in calculating the climate resilience indicator. This is to ensure that the tool is user-friendly and accurate at the same time.

7.1.2 Novelty

Numerous studies have delved into the realm of building resilience, each with its unique focus on specific aspects and hazards. For instance, Hung et al. (30) have pioneered a methodology aimed at assessing building resilience within the context of metropolitan land use planning, albeit primarily at a broader urban scale rather than the individual building level. Similarly, Cere et al. (31) have contributed by qualitatively characterizing urban-scale building resilience through delphi-based expert consultations. However, their approach may not readily apply to individual building owners and entails reliance on expert input.

On a more specialised front, Himoto & Suzuki (32) have devised a computational model to measure fire-resilience, although this method is confined exclusively to addressing fire-related risks. Lopez-Garcia et al. (33) have put forth diagnostic techniques for assessing heat resilience using temperature and CO2 data, but their approach may not comprehensively cover all dimensions of resilience. Sun et al. (34) have conducted a case study focusing on thermal resilience and energy efficiency within buildings, but it might not encompass the full spectrum of resilience factors. Meanwhile, Menna et al. (35) have undertaken a study concentrating on methodologies for assessing seismic resilience, particularly geared toward earthquake hazards.

Burroughs (36) has innovated by developing the ARMS tool to gauge building resilience from the building owner's perspective, albeit taking a broader approach that doesn't specifically address climate-related hazards. Finally, Duarte et al. (37) have proposed a resilience classification system founded on five dimensions, facilitating the classification and comparison of building performance and vulnerability assessment.

While these studies have made valuable contributions to the understanding of building resilience, they often concentrate on specific hazards or dimensions. Hence, there remains a pressing need for a more comprehensive approach to assess building resilience, particularly in light of the concerns stemming from the challenges posed by climate change. Here, we introduce a comprehensive framework able to assess and evaluate the climate resilience of buildings, with the aim to, on the one hand provide a detailed assessment of the buildings' climate resilience, and, on the other hand support easy access to the broad audience. The analysis covers the main climate hazards and integrates weather data for exposure analysis. The proposed framework also introduces an innovative building climate resilience rating based on a detailed building climate resilience scoring system which supports diverse building typologies. The scoring is based on a structured and quantifiable assessment approach achieved by drawing from extensive research on building infrastructure to determine the appropriate weights in the assessment tool. The results provide a robust understanding of the environmental challenges faced by buildings and the framework's adaptability is validated with application in two DigiBUILD pilot sites in Finland (Pilot 7 – FVH) and Greece (Pilot 9 – NTUA). In parallel, a web application is currently in development, aiming to improve the tool's user-friendliness and expand its accessibility, ensuring its utilisation by a diverse audience.

7.1.3 Development Progress

7.1.3.1 Climate exposure analysis

The resilience of buildings to climate-related challenges is intricately tied to their exposure and vulnerability to severe weather events. An illustrative statement from the Intergovernmental Panel on Climate Change (IPCC) underscores the importance of embedding information on climate risks into the architectural design, construction, and retrofitting of housing." These climate risks, which buildings must confront and withstand, encompass a wide range of direct and indirect threats. In this section, we will provide a general overview of the climate risks that buildings must address and cope with, along with an explanation of the methodology employed to utilise Numerical Weather Data for quantifying these risks on a scale.

7.1.3.1.1 Climate hazards

In the European Union, substantial endeavours have been undertaken to meticulously delineate climate hazards for the purpose of providing stakeholders with comprehensive information. According to the EU Taxonomy Classification, climate hazards are categorised as chronic and acute, further sub-divided into temperature, wind, water, and soil mass-related phenomena, as illustrated in Table 21. This study concentrated on specific climate hazards, namely, heat waves, cold waves, heavy precipitation, storms, flooding, and drought. The selection of these hazards is based on their direct correlation with weather data, enabling the calculation of the building's exposure intensity to these climate phenomena.

1.1.1.1.1 Weather data and exposure analysis

This subsection establishes the correlation between climate hazards and climate as well as non-climate data, elucidating how their analysis culminates in the quantification of climate hazard exposure. For the purpose of analysis, indicators were selected to assess exposure to climate hazards outlined in this section. The methodology employed involves establishing thresholds for the metrics associated with each climate hazard. This allows the quantification of exposure on a scale ranging from 0 to 5, where 0 signifies negligible exposure and 5 denotes very high exposure to the specified hazard.

Table 21: Classification of Climate Hazards.

	Temperature	Wind	Water	Solid Mass
Chronic	Changing temperature	Changing wind pattern	Changing precipitation patterns	Coastal erosion
	Heat stress	Precipitation variability	Soil degradation	
	Temperature variability	Ocean acidification	Soil erosion	
	Permafrost thawing	Saline intrusion	Solifluction	
	Sea-level rise	Water stress		
Acute	Heat wave	Cyclone / Hurricane	Drought	Avalanche
	Cold wave	Storm	Heavy precipitation	Landslide
	Wildfire	Tornado	Flood	Subsidence
			Glacial lake outburst	

- **Heat waves:** For heat wave hazard, the indicator utilised for assessing building locations is the heat index, also recognised as apparent temperature. The development of apparent temperature aimed to gauge thermal comfort rather than investigate human health Steadman **(38)**. Nevertheless, it has gained popularity as an exposure metric in environmental health, particularly in its approximate "heat index" form.

  The formula employed to calculate the heat index is as follows:

$$\mathrm{HI} = c_1 + c_2 \times T + c_3 \times R + c_4 \times T \times R + c_5 \times T^2 + c_6 \times R^2 + c_7 \times T^2 \times R + c_8 \times T \times R^2 + c_9 \times T^2 \times R^2$$

$$where: \begin{array}{l} HI \text{ is the heat index (in degrees Fahrenheit)}, \\ T \text{ is the air temperature (in degrees Fahrenheit)}, \\ R \text{ is the relative humidity (as a percentage)}, \\ c_1 = -42.379, \\ c_2 = 2.04901523, \\ c_3 = 10.14333127, \\ c_4 = -0.22475541, \\ c_5 = -6.83783 \times 10^{-3}, \\ c_6 = -5.481717 \times 10^{-2}, \\ c_7 = 1.22874 \times 10^{-3}, \\ c_8 = 8.5282 \times 10^{-4}, \\ c_9 = -1.99 \times 10^{-6}. \end{array} \qquad (16)$$

The mapping of the thresholds for the maximum values of the heat index observed in the weather data analysis of the last decade for the area where each reference building reference building is located is depicted in Table 22 , both with the qualitative exposure scale and the quantitative one.

Table 22: Heat Waves exposure levels

Risk Level	No Risk	Very Low	Low	Moderate	High	Very High
Max HI	**<26.7**	**32.2**	**39.4**	**46.1**	**54.4**	**>54.4**

Cold waves: In the context of the cold wave climate hazard, the index employed for computing the exposure fraction of the reference building is the Wind Chill Index (WCI). This index, beyond considering ambient temperature, incorporates wind speed, acknowledging its impact on heat transfer. The formula for the WCI is as follows:

$$WCI = 35.74 + 0.6215T - 35.75V^{0.16} + 0.4275TV^{0.16}$$

$$where: \begin{array}{l} WCI \text{ is the Wind Chill Index}, \\ T \text{ is the air temperature in Fahrenheit}, \\ V \text{ is the wind speed in miles per hour}. \end{array} \qquad (17)$$

The mapping of the thresholds for the maximum values of the WCI observed in the weather data of the last decade for the area where the reference building is located is depicted in Table 23 both with the qualitative exposure scale and the quantitative one.

Funded by
the European Union

Table 23: Cold Waves exposure levels

Risk Level	No Risk	Very Low	Low	Moderate	High	Very High
Max WCI	<10	4.4	-6.7	-17.8	28.9	>28.9

Heavy Precipitation: Regarding the hazard of heavy precipitation, an analysis was conducted on the hourly rainfall amount (mm/h). The mapping of the thresholds for the maximum values of the rainfall observed of the last decade for the area where the reference building is located is depicted in Table 24.

Table 24: Heavy Precipitation exposure levels

Risk Level	No Risk	Very Low	Low	Moderate	High	Very High
Max Rainfall (mm/h)	<12.7	38.1	76.2	127	192	>192

Heavy Storm: For the exposure analysis in storms, an evaluation of both the hourly rainfall amount (as already calculated in Table 24) and the wind speed (m/s) are considered. The categorisation in the different exposure levels is thus formulated by considering the site’s rating in both factors. The mapping of thresholds for the exposure assessment is depicted in Table 25. It should be noted at this juncture that both conditions must simultaneously be met to categorise the exposure of the building in the worst-case category.

Table 25: Storm exposure levels

Risk Level	No Risk	Very Low	Low	Moderate	High	Very High
Max Rainfall (mm/h)	<12.7	38.1	76.2	127	192	>192
Max Wind speed (m/s)	<7	14	28	35	42	>42

Flooding: In relation to the climate hazard of flooding, three key parameters were identified that predominantly influence exposure to this hazard. Apart from the evident factor of hourly rainfall amount (mm/h), consideration is given to the distance from bodies of water (sea, river, lake) and the elevation from sea level of the site where the reference building is situated. Consequently, the mapping of thresholds for the assessment of exposure in the flooding phenomenon is delineated in Table 26. Similarly, to storm analysis, all conditions must simultaneously be met to categorise the exposure of the building in the worst-case category.

Table 26: Flooding exposure levels

Risk Level	No Risk	Very Low	Low	Moderate	High	Very High
Max Rainfall (mm/h)	<12.7	38.1	76.2	127	192	>192

Funded by
the European Union

Distance (km)	**>5**	**2**	**1**	**0.5**	**0.1**	**<0.1**
Elevation (m)	**>300**	**100**	**50**	**30**	**10**	**<10**

Drought: The exposure of the building to the drought hazard was estimated using the Standardised Precipitation Index (SPI). This index, proposed by experts in 2009 at the Interregional Workshop on Indices and Early Warning Systems for Drought, was recommended for adoption by all National Meteorological and Hydrological Services (NMHSs) worldwide Zargar et al. (39). Its purpose is to characterise meteorological droughts, supplementing other existing drought indices utilised within their respective services. As per the official guide of the Standardised Precipitation Index, the thresholds for drought exposure classification are defined, as presented in Table 27.

Table 27: Drought exposure levels

Drought level	near normal	moderately dry	severely dry	extremely dry
SPI values	-0.99 to 0.99	-1.0 to -1.49	-1.5 to -1.99	-2 and less

However, in order to uphold scaling within the range of 0 to 5, the thresholds for drought exposure classification listed in Table 28 were ultimately employed.

Table 28: Drought exposure levels (0-5 range)

Risk Level	No Risk	Very Low	Low	Moderate	High	Very High
Min SPI	**>0**	**-0.5**	**-1**	**-1.5**	**-2**	**<-2**

In conclusion, this subsection has established the correlation between climate hazards and relevant data, providing a comprehensive understanding of the factors influencing the quantification of climate hazard exposure. Indicators, meticulously selected for their relevance to each climate hazard, facilitated the analysis, with thresholds set to enable a nuanced assessment.

7.1.3.2 Building climate resilience

Resilience to climate change can be defined as the ability of interconnected social, economic, and ecological systems to effectively manage and adapt to hazardous events, trends, or disturbances. This involves responding to and reorganizing in ways that ensure the preservation of their essential functions, identities, and structures. In the context of building infrastructure, the concept of resilient buildings encompasses a comprehensive approach. Resilient buildings should be strategically planned, meticulously designed, skilfully constructed, and efficiently operated in a manner that proactively anticipates, prepares for, and adapts to evolving climate conditions IPCC (40). Furthermore, resilient structures must have the capacity to withstand, respond to, and swiftly recover from disruptions resulting from these changing climate conditions. In this regard, buildings should play a significant role in mitigating or preventing the adverse impacts caused by the current climate situation or the expected challenges of the future EC (41). This pertains not only to safeguarding the building itself but also to ensuring the well-being of the individuals who inhabit these structures, preserving the surrounding natural environment, and protecting the valuable assets contained within these buildings FEMA (42). The

European Commission's report EC (41). outlines best strategies to enhance building resilience and adaptation solutions for each priority hazard and underscores the importance of strategic design and infrastructure choices in bolstering building resilience to diverse climate hazards. In the followings we explore how specific design and infrastructure considerations can contribute to building resilience in the context of the identified climate hazards and define a set of criteria based on which the climate resilience of the building is assessed. The criteria correspond to specific building characteristics, which are classified into six categories, these are: the building envelope, the energy and water systems, adaptation solutions throughout the building's infrastructure, sustainability practices and actions related to the buildings functional and environmental performance throughout its lifespan.

7.1.3.2.1 Building climate resilient rating

Building envelope: The role of building techniques and materials in enhancing the resilience of structures, either actively or passively, is underscored by the building envelope, while their influence on degradation levels is also significant Xiong et al. (43) In terms of the climate resilience they offer, various components of the building were examined, including wall construction material, interlayer insulation, external wall insulation, roof insulation, window-to-wall ratio, glazing systems, and shutters.

For the materials and wall constructions techniques the assessment is based on their contribution to enhancing the climate resilience of the building. Considering the diverse and non-uniform regulatory framework and its direct influence in the construction techniques across different European countries Papadopoulos (44) in this paper, the is conducted on some of the indicative techno types used in Europe since 1960. The criterion by which the framework is established is the thermal transmittance of the construction (U-value); however, other factors, such as air and humidity tightness and sustainability aspects, are also being factored European Parliament, Council of the European Union (45). Older technologies, aside from having lower U-values, thereby rendering buildings vulnerable to phenomena such as heat waves, cold waves, and drought, are constructed from less sustainable materials that have undergone temporal degradation, impacting the tightness of the building, especially in regions with high rainfall, storms, and potential flooding United Nations Environment Programme (46). With these considerations in mind, the lower tier is occupied by single-layer reinforced masonry walls, followed by lightweight timber and gypsum board walls. These represent older techniques with very low thermal transmittance and limited resistance to water-related phenomena. In the subsequent tiers, basic wood framing and brick/concrete walls or similar options were explored, showing improved thermal resistance results in comparison to their predecessors. Moving to the upper tiers, we find reinforced concrete/steel framing with masonry/concrete block cladding cement and specialised structural insulated panels or similar engineered resilient systems. The latter techniques are regarded as the most modern, aligning with the most up-to-date regulatory frameworks in European countries, indicating increased levels of thermal transmittance and durability.

Another component assessed for its contribution to the building envelope's resilience against severe climatic phenomena is interlayer insulation. It involves the application of insulation between the studs of walls or roofs. Similar to walls, it introduces an additional level of thermal transmittance, aiding in the reduction of the likelihood of air leakage and the risk of water penetration during heavy precipitation and flooding IEA (47). Once again, the regulatory framework varies across different European countries and evolves over time Papadopoulos (44). Moreover, the methods of implementation can differ based on the chosen material, manufacturer, specification, and installer. The styles considered in this paper, ranked in order from those offering little to no resilience to those significantly enhancing building strength, include: missing interlayer insulation, an air gap between studs, fibreglass batt insulation between studs, cellulose insulation between studs, open-cell spray foam insulation between studs, closed-cell spray foam insulation between studs, interior insulating sheathing, and aerogel insulation Arivumani et al. (48); Dong et al. (49). It is important to note that these techniques are indicative.

External wall insulation is a major factor in fortifying buildings against extreme weather events. This form of insulation is applied externally to walls and roofs, bolstering the overall building envelope. Beyond its energy-efficiency benefits, it provides an added layer of defence against severe weather conditions Lee et al. (50)Its primary function is to enhance the masonry's overall U-value, enabling the maintenance of indoor conditions in the face of prolonged heat waves or cold waves Berger et al. (51). As interlayer insulation, it also mitigates the risks of air leakage and water ingress Al-Homoud (52). The degree of reinforcement is contingent on various factors, including insulation type, installation quality, and layer thickness Ekici et al. (53). In this study, we examined and classified several indicative styles and materials of external insulation, assessing their resilience against climate-related hazards. Detailed ranking table can be found in the supplementary material, air leakage, water ingress and damage from extreme weather events.

A roof without insulation will allow heat to escape in the winter and enter in the summer, making buildings uncomfortable, expensive to heat and cool and vulnerable to climate hazards. A poorly insulated roof or a standard insulated roof with asphalt shingles is only slightly better than a roof with no insulation. Asphalt shingles are a common roofing material, but they are not the most climate-resilient option. Asphalt shingles can absorb heat from the sun, which can make your home hotter in the summer. They are also more susceptible to damage from high winds and hail. Metal roofs reflect sunlight, which can help to keep your buildings cooler in the summer. They are also more durable and less susceptible to damage from extreme weather events. A solar reflective or cool roof is a type of roof that is designed to reflect sunlight and reduce heat absorption. A green roof is a type of roof that is partially or completely covered with vegetation. Green roofs can help to insulate buildings, reduce stormwater runoff, and improve air quality. Hurricane-resistant roofing is a type of roofing that is designed to withstand high winds and hail. It is important to note that not all hurricane-resistant roofing is created equal, as is the case with earlier roofing styles.

Another noteworthy building characteristic scrutinised for its climate resilience is the window-to-wall ratio. This factor significantly influences a building's response to heatwaves and cold snaps by determining its capacity to harness or repel solar radiation, thereby influencing internal conditions. Research by Goia (54)has identified the optimal window-to-wall ratio for versatile performance across various climates to fall within the 35% to 45% range. Deviations from this range result in diminishing benefits in leveraging the local climate and solar exposure. When the ratio falls below the 35% to 25% range, and more significantly below the 25% to 15% range, a building's ability to maintain optimal indoor temperatures, especially in colder climates, is compromised. Conversely, when the ratio exceeds the 45% to 55% range, and further exceeds the 55% to 65% range, the building's ability to maintain optimal indoor temperatures diminishes, especially in warmer climates with extensive solar radiation exposure. Ratios outside of these mentioned ranges are generally regarded as impractical and do not confer climate resilience attributes to the building.

The final characteristic, though not intrinsically part of the building envelope, holds an indirect yet significant connection. It pertains to the presence, or absence, as well as the type of shutters adorning the building's openings. These shutters wield immense importance in bolstering the building's resilience against the array of climate hazards outlined in this manuscript. During heatwaves and in the context of xeriscaping, they serve as effective shading mechanisms, effectively preserving the internal temperature of structures Alawadhi (55). In scenarios involving flooding, heavy precipitation, and storms, shutters assume a dual role, acting as an additional layer of waterproofing and as a means of redirecting water away from the building's apertures Vutukuru et al. (56). Particularly in regions frequently subjected to severe storms, specialised shutters are deployed to enhance overall resilience. Consequently, the hierarchy, ranging from the least resilient to the most, spans from temporary plywood panels or cardboard to accordion-style shutters, roll-down or roll-up shutters, storm panels (metal or polycarbonate), Bahama shutters, and impact-resistant windows and doors.

Energy Systems: Energy systems and services play a key role in enhancing building resilience against a range of climate related hazards. In this context, we propose a rating approach for evaluating the resilience of buildings based on a set of energy systems, namely buildings cooling equipment, heating systems, ventilation systems and the use of energy efficient and smart appliances.

The assessment of cooling strategies' resilience is based on the analysis outlined in Zhang et al. (57). This study undertakes an in-depth examination of the performance of state-of-the-art cooling strategies, with a particular focus on their effectiveness during heatwaves and power outages. The resilient characteristics of the examined cooling strategies are summarised across four criteria – absorptive capacity, adaptive capacity, restorative capacity and recovery speed. Furthermore, aside from resilience capabilities during extreme events, the suitability of the cooling strategies in terms of climate zone and technology readiness level is also assessed. To gain a more comprehensive insight into the characteristics of each cooling system, we have drawn upon information from the U.S. Department of Energy (58) Results are summarised below, while the resulting scoring is presented in the supplementary material. Compression refrigeration, relying on vapor compression technology, demonstrates high adaptive capacity as it can retain sufficient cooling capacity even during heatwaves. However, it's not very robust to power outages, and integrating it with local electricity production or energy storage can enhance its resilience. Absorption refrigeration, including desiccant cooling, offers energy-efficient and eco-friendly cooling, making it resilient during heatwaves if backed by alternative heat sources. Ground source cooling, benefiting from stable ground temperatures, exhibits high resilience to heatwaves. Still, its cooling capacity may be affected by climate change, and the recovery speed depends on environmental conditions. Sky radiative cooling, while renewable and resilient to blackouts, relies on favourable climate conditions and material properties. Finally, high-temperature cooling systems like radiant cooling show high adaptability and resilience under heatwaves and power outages, particularly when properly designed and controlled. These cooling strategies vary in their suitability depending on climate, building type, and specific needs, underscoring the importance of thoughtful planning and system selection for optimal performance and resilience. Ventilative cooling harnesses the potential of outdoor air for cooling and can encompass both natural and mechanical methods. While ventilative cooling systems' effectiveness depends on factors like ventilation type, building characteristics, and local climate, well-designed natural ventilation can offer substantial relief during heatwaves. Adiabatic/evaporative cooling offers an alternative cooling approach based on the adiabatic process of reducing air temperature through water evaporation. Climate change impacts reveal that evaporative cooling, exhibits more resilience compared to ventilative cooling.

Insights on the assessment of different heating technologies are retrieved by the papers of Vakiloroaya et al. (59) and Bac et al. (60). For residential usage, the U.S. Department of Energy website was also advised. The heating systems reviewed are gravity air furnaces, portable or plug-in space heaters, hot water baseboard heater systems, electric resistance, hot water baseboard heater systems, in-floor radiant, traditional boilers and radiator systems, heat pump, forced air distribution systems, furnace and hybrid heating home systems (heat pump systems and power of gas furnace). Gravity air furnaces, which rely on natural convection, are robust but often less energy efficient. Portable or plug-in space heaters offer quick, localised heating but are not suitable for large surfaces. Hot water baseboard heater systems are effective, distributing heat evenly through radiators or baseboards, but again their capacity is limited. Electric resistance systems are easy to install but can be costly to operate. In-floor radiant heating provides a comfortable, energy efficient solution. Traditional boilers and radiator systems offer reliable, widespread warmth, but they can be less efficient. Heat pumps are energy-efficient options that work well in moderate climates. Forced air distribution systems and furnaces are common choices for quick heating, while hybrid heating systems, combining a heat pump with a gas furnace, offer a balance between efficiency and power U.S. Department of Energy (DOE) (58). It is important to note that the site's climatic conditions, expected thermal comfort, costs, the availability of energy sources and the application of the building are key factors that affect the performance and scoring of each heating technology Vakiloroaya et al. (61). Also, ergonomic, environmental, reliability, technical, and economical aspects should be considered Bac et al. (62).

For ventilation systems, the scoring is based on the articles of Ahmed et al. (63) and Zaniboni & Albatici (64). Ahmed et al. (63) focus on natural ventilation and examine the capacity of different natural ventilation strategies to deliver heatwave resilience and sufficient thermal comfort and maintain high indoor air quality within warm climatic conditions. The study of Zaniboni & Albatici (64) expands the results of Ahmed et al. (63) and offers a systematic review to compare both natural and mechanical ventilation systems under a sustainable design perspective. Among the various natural ventilation techniques, single sided and cross ventilation leverage prevailing wind patterns to cool indoor spaces, but their effectiveness can be limited by wind availability. Windcatchers, architectural structures designed to capture and direct breezes into buildings, are effective in harnessing wind for cooling, particularly in arid regions. Solar chimneys, on the other hand, exploit solar-induced buoyancy to create airflow, making them reliable in sunny climates. Natural ventilation with evaporative cooling is especially effective in extremely hot environments, utilizing water evaporation to reduce indoor temperatures. While air handling units represent mechanical ventilation, they are energy-intensive and often less sustainable. Heat recovery ventilation strikes a balance by recovering heat from exhaust air, enhancing energy efficiency. Hybrid or mixed-mode ventilation systems combine the strengths of natural and mechanical ventilation, adapting to changing conditions for optimal comfort and air quality, offering a comprehensive suite of solutions for warm-climate buildings. Yet, it is important to highlight that selecting the best ventilation solution must take into account the context, type, and condition of energy efficient buildings.

Last, we recognise the resilience benefits of the strategic deployment of energy efficiency and automated building monitoring and management systems Frankoni et al. (65), and evaluate the use of energy efficient and smart appliances. We propose a rating comprising of 11 evaluation levels. The initial level corresponds to the absence of energy efficient or smart appliances, while the subsequent levels represent an incremental increase of 10% in the adoption of such appliances relative to the total number of appliances within the buildings.

Adaptation solutions on systems and services: The list below comprises a compilation of potential solutions for adapting buildings to the priority hazards outlined in Section 1.1.1.1.1. The alignment or interaction between each adaptation solution and the associated hazards is determined using the methodology presented in the EC report (41). Most of these solutions are assessed using a two-tiered approach, which corresponds to whether they are present or not in the examined building. However, the evaluation of the installation of storm protection devices utilises a three-tiered approach, assigning varying scores to different percentages of critical building devices with installed protection against storms.

1. Disconnecting surface water from sewage system: The separation of surface water and sewage systems is essential for flood resilience, as flooding has the potential to harm a building's infrastructure services in case overflow or backflow water infiltrates the building through sewage pipes EC (41). Ensuring the separation of these systems minimises the potential for sewage overflow during storms and flooding, safeguarding public health Houghton & Castillo-Salgado (66).
2. Placement of electrical and mechanical systems above flood level: The report of Low13 underscores the importance of elevating critical building components such as electrical and mechanical systems above flood levels. This strategic placement offers robust resistance to a spectrum of hazards, including, apart from floods, heavy precipitation, and storms. By safeguarding these vital systems ensures the continuity of building operations but also yields long-term cost savings by mitigating potential damage and maintenance expenses.
3. Dimensioning drainage networks to future runoff projections: Designing drainage networks to accommodate future runoff projections is vital for resilience against increasing precipitation and storm events. Research suggests that incorporating climate projections into drainage system design can prevent urban flooding Arnbjerg-Nielsen et al. (67). Proper dimensioning ensures efficient water management and minimises flood risks.
4. Placement of sinks/toilets at a minimum height above flood level: In line with the floodplain hazard management regulations in Montana (68), plumping systems such as toilets, stools, sinks, urinals, vaults, and drains must be placed at a minimum height above flood levels to reduce water damage and contamination during flooding. This also ensures essential sanitary facilities remain functional in emergency situations.
5. Water-efficient fixtures and fittings: Indoor water efficiency, installation of water-efficient fixtures and fittings, contributes to building resilience during drought and water scarcity episodes. Fixtures like delayed inlet valves, flow restrictors and low-flush toilets are examples of technologies that not only reduce the environmental impact but also ensure a reliable water supply during dry periods. It is important to note that these solutions require regular checking and monitoring for potential leakages EC (69).
6. Greywater recycling: Greywater recycling systems enhance buildings' resilience by conserving freshwater resources, promoting water sustainability, and reducing the dependency on central - municipal water supplies. Being effective against water scarcity and supply disruptions, such solutions thereby contribute on mitigating drought-related challenges Gikas & Tchobanoglous (70).
7. Onside water sources – water storage: Water storage solutions provide a secure and readily available water source, particularly in regions prone to droughts and water shortages. These solutions, which may include rainwater tanks for greywater provision or watering plants during dry periods, not only offer a reliable local water source but also support conservation efforts. Local governments and municipalities play a critical role in this context by enforcing the installation of water retention systems in new buildings, which should be designed to withstand longer flood return periods, ensuring a proactive response to extreme weather events EC (69).

Such onsite water sources, like storage tanks or wells, address both high-demand periods and the preservation of water resources, thus effectively guaranteeing a continuous water supply even during times of scarcity. Water storage solutions thus play a crucial role in building resilience, reinforcing water security, and adapting to changing rainfall patterns.

8. Burial of distribution lines: Burial of distribution lines reduces vulnerability to power outages, enhancing resilience in the face of high winds or storms EC (69).
9. Installation of backup generators: Similar to the burial of distribution lines, the installation of backup generators enables building to continue operating in case of grid failures and electrical disruptions caused by high winds and storm events EC (69).
10. Storm protection devices: Thunderstorms and lightning strikes, in addition to causing a power surge, can also cause damage to electronic devices. Storm protection devices enhance building resilience by safeguarding against extreme weather, reducing damage and ensuring occupant safety. As mentioned in the beginning of this section, the scoring related to the installation of storm protection devices utilises a three-tiered system, assigning different scores based on the percentages of critical building devices with installed protection against storms.

Sustainability practices: Sustainable Drainage Systems (SuDS) are natural solutions designed to retain a specific percentage of rainfall, effectively accommodating water for short durations and mitigating heavy precipitation events. There are a variety of different SuDS options, each serving different purposes and their development is dependent on the different locations or typical rainfall landing area. Under the 2019 version of the UK National Planning Policy Framework Department for Levelling Up & Communities (71), SuDS form one of five key considerations to the provision of development, when considering how to manage flood risk. Also, the use of SuDS as a central tool to reduce flooding was outlined in the Pitt Review (72). Depending on their design, SuDS can enhance biodiversity and reduce the risk of sewage overflow and associated health hazards, and vegetation, as part of SuDs, contributes to heat reduction. However, they require regular and potentially costly maintenance. The types of SuDS that are examined here are: permeable surfaces and filter drains (gravelled area, solid pavins blocks, porous paviors), infiltration devices (soakaways, infiltration trenches and basins), filter strips and swales, basins and ponds (constructed wetlands, balancing ponds, detention basins, retention ponds) and living roofs. The hierarchy of the different SuDS types, as detailed in the supplementary material, is based on a sustainability scale and incorporates associated amenity and environmental benefits as in Ambiental (73).

Water systems: In examining water systems, two critical indicators for a comprehensive life cycle evaluation were assessed. Firstly, Anti-return valves for toilets and sinks/sewage pumps are scrutinised for presence or absence, addressing contamination risks. Secondly, the evaluation includes Rainwater tanks, evaluating their integration or exclusion for sustainable water management. This systematic approach ensures a thorough understanding of water system resilience and sustainability, aligning with principles established in building assessments.

Site & Location: In addition to the building itself, the development and construction of the broader area in which it is situated and the available infrastructure to address extreme weather conditions are of critical importance for the building's resilience. Therefore, two factors related to the site sustainability of the building and the surrounding area were incorporated into this research: land use and zoning, and utilities and infrastructure to address severe climate events. Land use and zoning consider urban development and land-use planning as part of climate hazard mitigation efforts Burby et al. (74). The following criterion was divided into five levels, listed below from the one that contributes least to the most resilient building location:

1. Urban sprawl with limited green spaces
2. Mixed-use zoning with green corridors
3. Smart growth and compact development
4. Mixed-use zoning with planned green infrastructure
5. Development with climate-adaptive zoning

In terms of utilities and infrastructure, this pertains to the availability of emergency services and plans, hospitals, and evacuation plans for the area. These factors collectively play a crucial role in a dynamic climate scenario to ensure that residents receive immediate assistance, have effective evacuation plans in place, and receive proper care during potential climate crises et al. Lewis & Aghababian (75). Therefore, the five defined levels, from the least to the most resilient, are as follows:

1. Limited or underfunded emergency services
2. Basic emergency services and hospital facilities
3. Well-equipped hospitals and trained emergency personnel
4. Advanced healthcare facilities and comprehensive emergency plans
5. Integrated emergency services, hospitals, and evacuation routes.

7.1.3.3 Building Climate Resilience Scoring System

Following the definition of building's exposure to the considered climate hazards and determining the various levels of each direct and indirect building element, and calibrating their resilience to specific hazards, the subsequent phase involves devising a scoring system. This system aims to encapsulate both the overall climate resilience of the building and the individual resilience of its diverse systems, as well as its vulnerability to the considered climate hazards. This section will elucidate the methodology of the developed scoring system and its utilisation of the climate exposure analysis and resilience ratings for each building component. These elements are instrumental in calculating the appropriate weights and, ultimately, deriving the building's final resilience score.

7.1.3.3.1 Weight Calculation

To calculate the weights of each climate hazard, the first step involves defining a table that associates each climate hazard with a value represented as j.

Let e_j denote the result of the exposure analysis, as derived from the methodology detailed in 1.1.1.1.1, for each hazard. Subsequently, these results are normalised using the equation provided in *(18)*.

$$\overline{e}_j = \frac{e_j}{\max(e_j)} \tag{18}$$

At this stage, the injection of external importance to each hazard, tailored to the assessor's requirements, becomes feasible. Thus, i_j ' is defined for each climate hazard 'j' and must adhere to *(19)*.

$$0 \leq i_j \leq 1 \quad \text{and} \quad \sum_j i_j = 1 \tag{19}$$

Therefore, the non-normalised intermediate weights, w_j, for each climate hazard j are computed using the equation *(20)*.

$$w_j = i_j \cdot \widetilde{e}_j \tag{20}$$

Finally, the weights are normalised using the equation ***(21)***.

$$\overline{w}_j = \frac{w_j}{\sum_{i=1}^{6} w_i} \tag{21}$$

Table 29: Building Categories and components indices.

	Envelope (1)	Energy Systems (2)	Resilience (3)	Sustainability Practices (4)	Water Systems (5)	Site & Location (6)
1	Walls	Cooling Equipment	Connection between surface water & sewage	Sustainable urban drainage systems	Anti-return valves	Land use and zoning
2	In. insulation	Heating Equipment	Placement of electrical & mechanical systems		Rainwater tanks	Emergency services, hospitals, and evacuation route setups
3	Roof Insulation	Ventilation	Drainage network dimensioning			
4	Window/ Wall ratio	EE appliances	Placement of toilets & sinks			
5	Glazing system	Smart appliances	Water efficient fixtures and fittings			
6	Ex. Insulation		Greywater recycling			
7	Shutters		Water storage			
8			Burial of distribution lines			
9			Installation of backup generators			
10			Storm protection devices			

Funded by the European Union

7.1.3.3.2 Climate Resilience Score Calculation

The calculation of the climate resilience score of the building is based on the current level of the different components of the building, as defined in 7.1.3.2.1. For easier indexing in the equations that follow, Table 29 is defined which maps the different components of the different categories of the building to y and z indices respectively. For easier indexing in the equations that follow, in Table 29 the different building components (y) are mapped to the corresponding category (z).

Let $C_{z,y,j}$ be the resilience score achieved by the reference building for component y of category z against hazard j. Each score $C_{z,y,j}$ is normalised by the maximum resilience score that the reference building achieves among all components y of category z against hazard j.

$$\overline{C}_{z,y,j} = \frac{C_{z,y,j}}{C^{\max}_{z,y,j}} \quad (22)$$

By summing up the normalised values of the score in each component y of each category z for each hazard j as, the overall climate resilience score is generated in a scale from 0 to 1.

$$S_j = \frac{\sum_z \sum_y \overline{C}_{z,y,j}}{N} \quad (23)$$

Then, the normalised weights calculated in *(21)* are incorporated through *(24)* to account for the intensity of the building's exposure to the considered climate hazards.

$$\mathrm{WS}_j = \mathrm{NWS}_j \cdot \overline{w}_j \quad (24)$$

Finally, by summing the resilience score for each climate hazard j, the total climate resilience score of the reference building on a scale of 0 to 1 is obtained, which is ultimately multiplied by 100 to yield the percentage resilience score of the reference building against the optimal scenario.

$$S_{total} = \sum_{j=1}^{7} WS_j \quad (25)$$

7.1.3.4 Development of Web Application

Currently a web application has been developed for the pilots, where they can assess their buildings and compare them afterwards. The application features a user-friendly interface where the user can enter information about multiple buildings and then proceed with the assessment. The methodology mentioned has been applied in the app and the user can answer a questionnaire to determine and weight the climate hazards that affect each building, and then, after choosing the appropriate information for each building category, the assessments results can be seen. The web application can be used for all pilots without the need for custom modifications and is currently in the stage of being tested and fine-tuned.

The source code can be found on GitHub: GitHub s3.5.1

7.1.4 Application on DigiBUILD Pilots

- **Pilot 7 (FVH):** The climate of Finland exhibits characteristics of both maritime and continental climates, depending on the direction of air flow. Despite its northern location, the average temperature in Finland is several degrees higher than that of most other regions at similar latitudes. The Finnish climate is characterised by irregular precipitation, and it often experiences swift changes in weather conditions World Bank (76). Additionally, Finland is expected to witness an increase in both humidity and temperature, with the most significant warming occurring during the winter months. Winters are projected to become more humid, overcast, and with reduced snowfall, as highlighted in the study by Ruosteenoja et al. (77). The rise in sea level, estimated to be 29cm in the Gulf of Finland, is partially compensated by the residual land mass uplifting still following the loss of ice cover from the last ice age Johansson et al. (78). The main climate-related hazards in Finland are storms, floods and rapid flooding in or near the population centers Pilli-Sihvola et al. (79). The finish pilot takes place in a large office building, used and operated by the City of Helsinki. The City of Helsinki is actively focusing on the enhancement of data-driven management for its portfolio of city-owned buildings. This particular building was finalised in 2020, covering an area of 35,261m2 across seven floors and a basement and can accommodate approximately 2,000 people. The building is part of district heating and district cooling networks. These networks are currently powered by coal, but there are plans to transition to gas and biomass as alternative fuel sources.

- Pilot 9 (NTUA): To further broaden the analysis, the second case study is conducted in Greece, a country situated in Southeast Europe. Greece exhibits a diverse climate owing to its unique geographical location and varied topography, resulting in distinct climatic zones within relatively short distances Angra & Sapountzaki (80). Climate hazards in Greece are primarily influenced by its Mediterranean climate, characterised by mild and wet winters in the southern lowland and island regions, contrasted with cold winters with heavy snowfalls in the mountainous areas in the central and northern regions, alongside hot and dry summers World Bank (76). These climate related hazards include heatwaves, droughts, wildfires and occasional winter storms and flooding. Our methodology is implemented in a ground floor office building of 150 m^2 located on the campus of National Technical University of Athens.

7.1.4.1 Exposure and Weight Results

According to the framework outlined in 7.1.3.1, the exposure of the considered buildings, along with the resulting weights derived from equations *(18)*, *(19)* and *(21)(19)*, are presented in Table 30. It is evident from these tables that the exposure of each site to climate hazards is accurately depicted. On one hand, a site located in a Nordic country, Helsinki, Finland, is characterised by intensely cold winters with normal humidity levels, moderate rainfall, and limited storm events. On the other hand, Athens, Greece, experiences hot and dry summers, thereby increasing exposure to drought phenomena. However, its winters are milder. Additionally, occasional storms and high rainfall are observed, elevating the risk of flooding. The meteorological data for both pilot cases were sourced from the open-access repository Weather2023 (81).

Table 30: Exposure and Weight Results for Helsinki and Athens pilot sites.

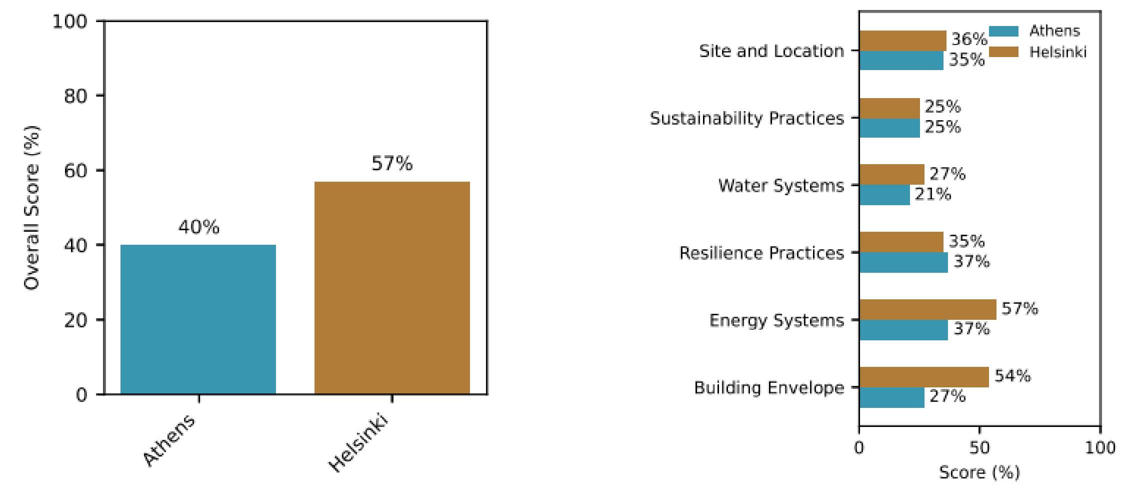
	Heat waves	Cold Waves	Heavy Precipitation	Storms	Flooding	Drought
			Helsinki			

Exposure	1	4	3	2	2	1
Weight	0.08	0.31	0.23	0.15	0.15	0.08
			Athens			
Exposure	4	2	2	1	3	2
Weight	0.29	0.14	0.14	0.07	0.21	0.14

7.1.4.2 Climate Resilience Score Results

The criteria outlined in 10 and 11 have been fulfilled for both buildings. Detailed tables for each building are available in the supplementary material. In Figure 77(a), the total score attained by each building within scores. the climate resilience scoring framework is presented. It is evident that the building located in Finland demonstrates superior reinforcement against the severe weather events to which it is exposed, compared to the building situated in Greece. Shifting focus to Figure 77(b), the discernible difference is primarily observed in the energy systems and building envelope components. This discrepancy can be readily explained by the Finnish building being a more modern construction than its Greek counterpart. By examining Figure 77, insights into the resilience of the buildings against their respective hazards are obtained. Specifically, in 1a, it is observed that the building in Finland achieves a comparable resilience score against all climate hazards, nearing 50%, and notably excels in drought, heavy precipitation, and cold waves, where the resilience score approaches 60%. Conversely, the Greek pilot case demonstrates good resilience against flood, drought, and heavy precipitation, achieving a score close to 50%. However, it appears to lag in resilience against storms, cold waves, and heat waves, where the resilience score is just above 30%. Lastly, Figure 76a, the weighted resilience of each building reference to each hazard contributes to the overall score. By aggregating the percentages for each hazard, the total score is obtained, as described by the equation in *(25)*.

〉



(a) Comparison of climate resilience overall score per building.

(b) Comparison of non-weighted climate resilience scores per climate hazard.

Figure 77: Comparison of climate resilience scores

7.1.4.1 Web Application

As already mentioned, the web application can be used for both pilots without any further modifications. The first version of the application is complete and is currently being tested by the pilots. In the following figures the interface is presented with an NTUA building as an example.

- Figure 78: The main page of the application, with which the user is greeted. The user can add a new building or click on any previously added ones to review them or update them. The user can also view the overall score of the assessment for each building and compare them.
- Figure 79: A popup where the user can add the necessary information to create a new building.
- Figure 80: The information entered for a building.
- Figure 81: A portion for the climate hazards questionnaire, that will be used to calculate the weight for each hazard.
- Figure 82: The results presented for the climate exposure assessment. The user can see what hazards are more likely to inflict the building.
- Figure 83: A portion for the building characteristics selection. The user needs to select the characteristics for the building for each category, so that the assessment can proceed.
- Figure 84: The results of the assessment for each category and per climate hazard. The results are shown weighted and non-weighted.

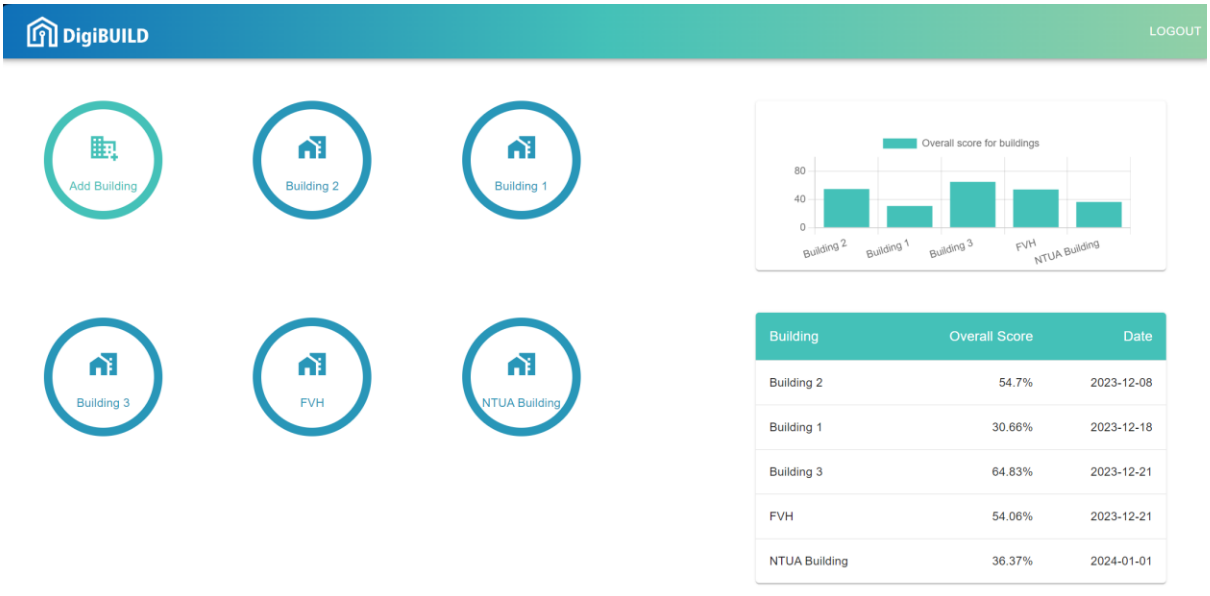


Figure 78: Dashboard of the application

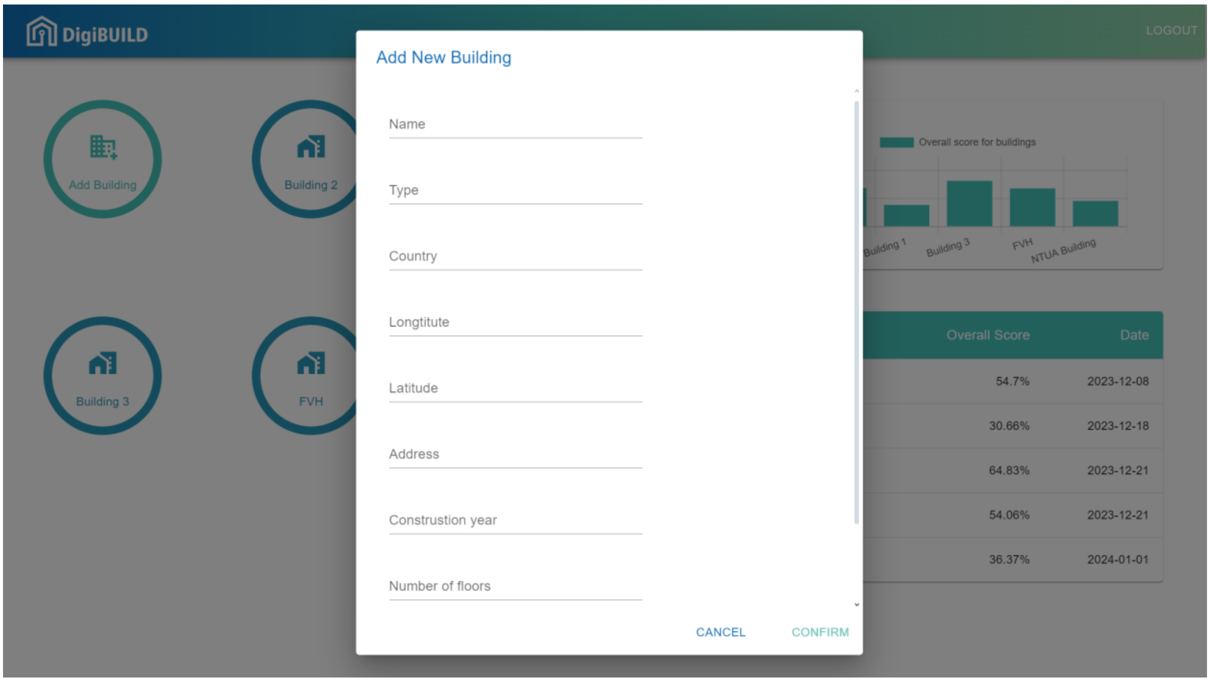


Figure 79: Add new building pop-up

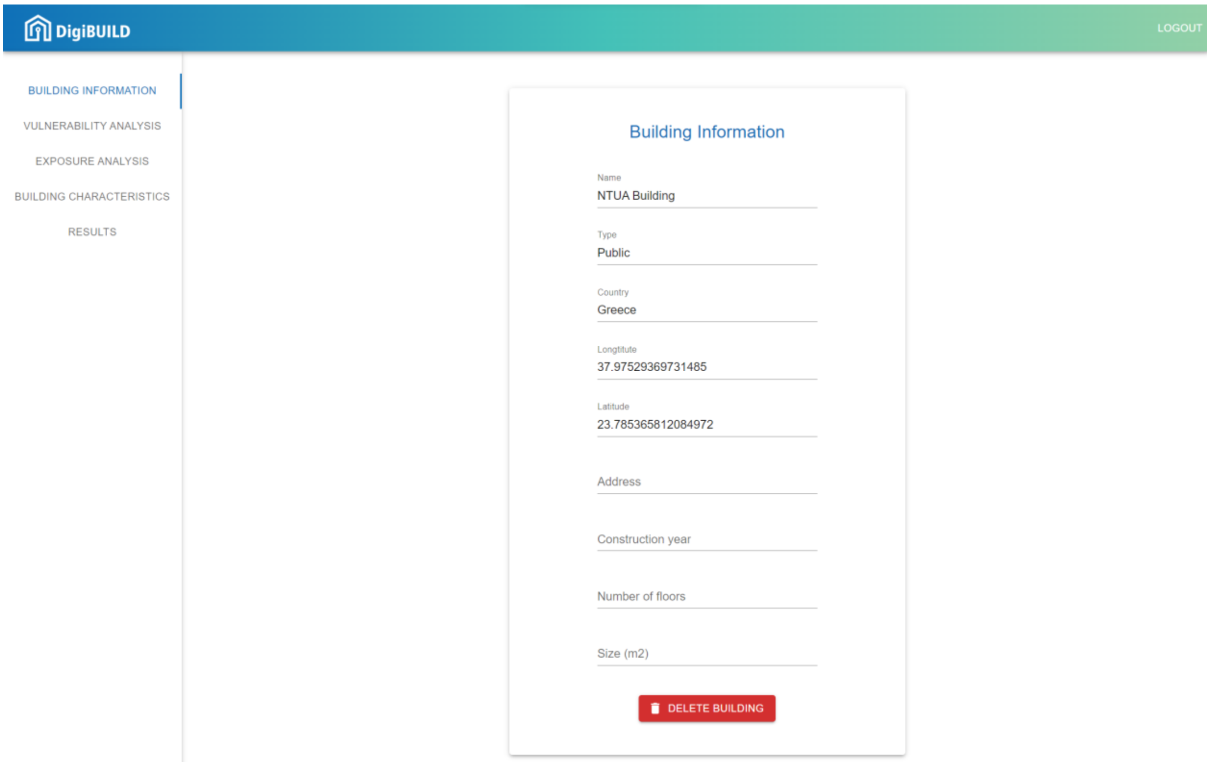


Figure 80: Basic information of selected building

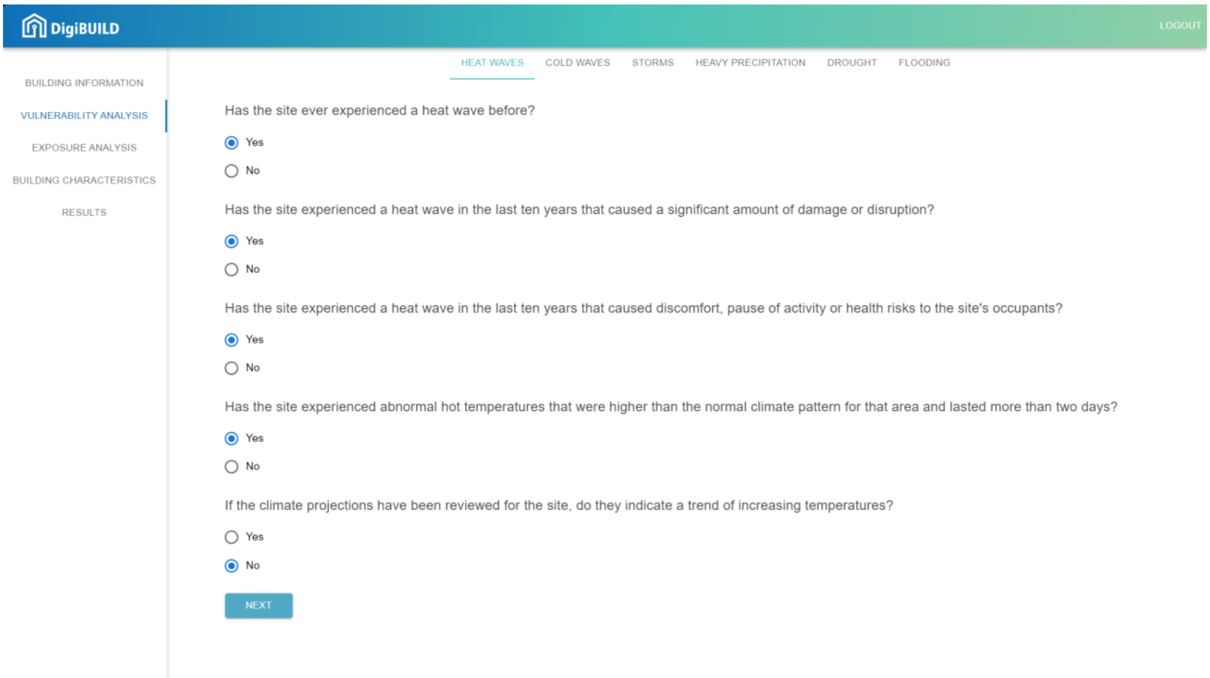


Figure 81: Part of the climate hazards questionnaire

Figure 82: Results of exposure analysis for the building

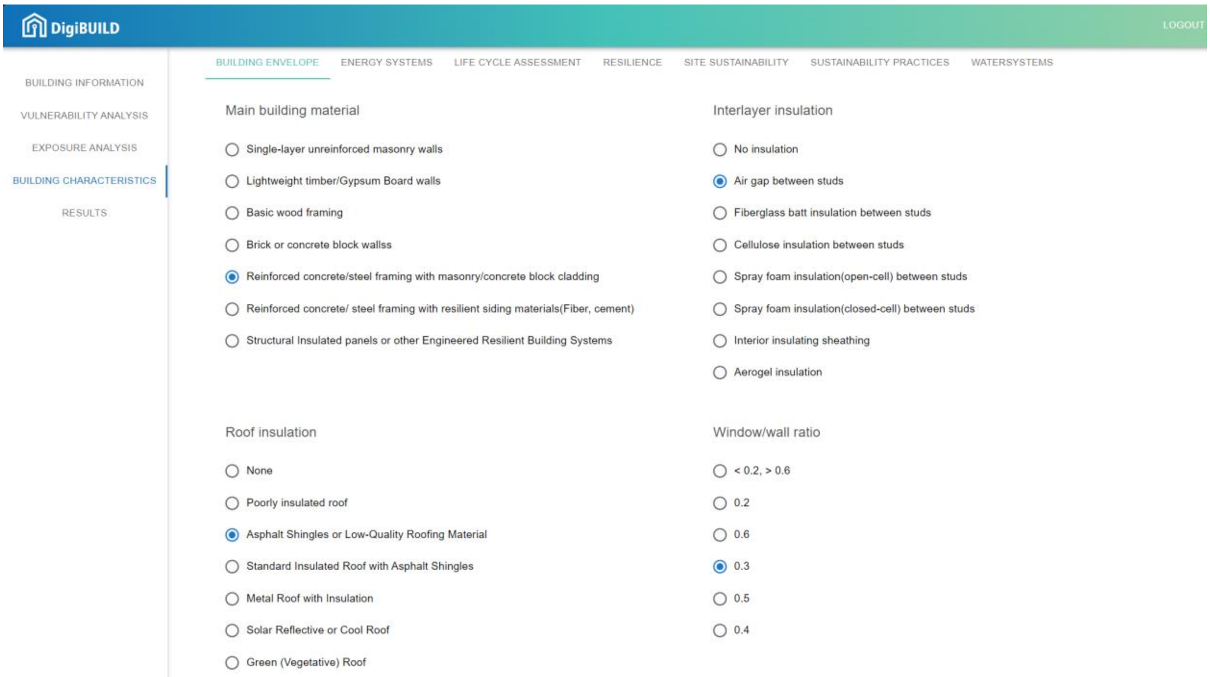


Figure 83: Part of the building characteristics selection

Figure 84: Building assessment final results

7.1.5 Next Steps

The proposed tool builds upon three key elements i) the examination and evaluation of essential structural and infrastructural building elements, ii) the comprehensive assessment of the building's capacity to withstand climate-related hazards, ii) the contribution to the improvement of the building's overall climate resilience.

Extending prior research, the primary aim of the proposed tool is to simplify building assessment, making it accessible to individuals without specialised technical knowledge. Additionally, it covers all the main climate hazards and places a strong emphasis on precision. This is achieved by drawing from extensive research on building infrastructure to determine the appropriate weights in the assessment tool. As next step, the goal is to create a user-friendly and accurate tool that can be used by a broad audience. Towards that scope, there is an ongoing development of a web application designed to enhance the user-friendliness of this tool and broaden its accessibility, making it available to anyone interested in its utilisation.

In parallel, among the proposed next steps to enhance the tools precision and efficacy in the evaluation of a building's resilience to climate hazards, is the fine-tuning of the assessment tool's weights. At the same time, it is essential to pursue new information regarding building structures that may wield significant influence over a building's vulnerability to climate-related threats. Notably, the future incorporation of established building evaluation methodologies into the existing framework, such as the Energy Performance Certificate (EPC), the Smart Readiness Indicator (SRI), and the Post Occupancy Evaluation (POE), holds particular significance. By continually enriching the tool with fresh information and insights, the aspiration is to provide a more comprehensive and robust resource for assessing climate resilience in the built environment.

Funded by
the European Union

8 Conclusions and Next Steps

As we conclude our thorough reporting of the 'Second wave' of DigiBUILD AI-based data-driven services, it is crucial to consider the significant advancements in integrating advanced AI into the built environment sector. This journey, outlined in the preceding sections, marks a considerable advancement in utilising data-driven technologies to enhance energy efficiency, building management, and occupant well-being, while in parallel addressing climate resilience. However, there were also cases where services were ahead of some others in terms of development. Nevertheless, there have been instances where certain services have surpassed others in deployment. This underscores the clear significance of immediate availability and enhanced data quality. Simultaneously, it highlights the pressing need for modernisation and improvement of the built environment.

Nevertheless, all these efforts contribute towards realising the primary objectives outlined in the Grant Agreement. These include delivering at least 20 novel pre-trained machine learning and deep learning models for buildings. Additionally, it is anticipated that the degree of uncertainty associated with the monitored and existing data will be reduced to below 10%. Furthermore, it aims to ensure comfort within the Predicted Mean Vote (PMV) range of -0.5 to 0.5 during the occupancy period of the buildings.

This report also places significant emphasis on the identification of novel and innovative aspects of services where they exist. In this domain, which is currently evolving, certain services have been identified that either exhibit a robust and innovative methodology or present a fresh, novel approach to the well-established issues of energy management in buildings. However, it is noteworthy that some services demonstrate less pronounced innovation.

In conclusion, the following objectives have been established for the future development and successful forthcoming full Pilot operation of DigiBUILD:

- Completing the methodologies by adding final refinements where necessary.
- Developing synchronous applications related to the Work Package (WP) services host, especially in areas where this has not been implemented yet.
- Operating the services within an interoperable environment, particularly ensuring communication with the relevant Work Packages (WP2, WP4) as required.
- Providing comprehensive documentation of scenarios for the effective training and guidance of the pilots involved. This is crucial to maximise the benefits of the developed technologies and support them in achieving their objectives.
- Ensuring the security of pilot and application data at every stage of development.

Appendix

Table 31: s3.4.2 - AHP matrix containing pairwise importance of each BPI.

	BPI_1	BPI_2	BPI_3	BPI_4	BPI_5	BPI_6	BPI_7	BPI_8	BPI_9	BPI_10	BPI_11	BPI_12	BPI_13	BPI_14	BPI_15	BPI_16	BPI_17
BPI_1	1	2	4	3	2	5	6	5	4	4	3	2	3	4	3	4	5
BPI_2	1/2	1	3	2	3	4	5	4	3	3	2	2	3	3	2	3	4
BPI_3	1/4	1/3	1	2	1/2	6	7	6	5	5	4	3	4	5	4	5	6
BPI_4	1/3	1/2	1/2	1	1/3	4	5	4	3	4	5	3	4	4	6	7	5
BPI_5	1/2	1/3	2	3	1	4	5	4	3	3	4	2	3	3	2	3	4
BPI_6	1/5	1/4	1/6	1/4	1/4	1	2	3	7	6	2	1	2	5	1	4	3
BPI_7	1/6	1/5	1/7	1/5	1/5	1/2	1	2	6	5	1	1/2	1	4	1/2	3	2
BPI_8	1/5	1/4	1/6	1/4	1/4	1/3	1/2	1	5	4	1	1/2	1	3	1/2	2	1
BPI_9	1/4	1/3	1/5	1/3	1/3	1/7	1/6	1/5	1	7	1/2	1/3	1/2	6	1/3	4	3
BPI_10	1/4	1/3	1/5	1/4	1/3	1/6	1/5	1/4	1/7	1	1/2	1/3	1/2	5	1/3	3	2
BPI_11	1/3	1/2	1/4	1/5	1/4	1/2	1	1	2	2	1	1/2	1	4	5	6	4
BPI_12	1/2	1/2	1/3	1/3	1/2	1	2	2	3	3	2	1	2	3	2	3	4
BPI_13	1/3	1/3	1/4	1/4	1/3	1/2	1	1	2	2	1	1/2	1	3	2	4	3

BPI_14	1/4	1/3	1/5	1/4	1/3	1/5	1/4	1/3	1/6	1/5	1/4	1/3	1/3	1	1/2	2	2
BPI_15	1/3	1/2	1/4	1/6	1/2	1	2	2	3	3	1/5	1/2	1/2	2	1	4	3
BPI_16	1/4	1/3	1/5	1/7	1/3	1/4	1/3	1/2	1/4	1/3	1/6	1/3	1/4	1/2	1/4	1	2
BPI_17	1/5	1/4	1/6	1/5	1/4	1/3	1/2	1	1/3	1/2	1/4	1/4	1/3	1/2	1/3	1/2	1

Table 32: s3.4.2 - Calculated weights per BPI. (a)

	BPI_1	BPI_2	BPI_3	BPI_4	BPI_5	BPI_6	BPI_7	BPI_8	BPI_9
weights	0.14123193	0.11120199	0.11583311	0.10242537	0.09943315	0.05823796	0.04052578	0.03365667	0.03632785

Table 33: s3.4.2 - Calculated weights per BPI. (b)

	BPI_10	BPI_11	BPI_12	BPI_13	BPI_14	BPI_15	BPI_16	BPI_17
weights	0.02529783	0.04907615	0.05558917	0.03867374	0.01894081	0.04113634	0.01686202	0.01555013

Funded by
the European Union

References

1. *Energy prediction techniques for large-scale buildings towards a sustainable built environment: A review. Energy and Buildings, 2020, 224: 110238.* GASSAR, Abdo Abdullah Ahmed and CHA, Seung Hyun.

2. *Building energy prediction using artificial neural networks: A literature survey. Energy and Buildings, 2022, 262: 111718.* LU, Chujie, LI, Sihui and LU, Zhengjun.

3. *Prediction of building energy consumption by using artificial neural networks. Advances in Engineering Software, 2009, 40.5: 356-362.* EKICI, Betul Bektas and AKSOY, U. Teoman.

4. *Energy baseline prediction for buildings: A review. Results in Control and Optimization, 2022, 7: 100129.* QAISAR, Irfan and ZHAO, Qianchuan.

5. *Methodology to estimate building energy consumption using EnergyPlus Benchmark Models. Energy and Buildings, 2010, 42.12: 2331-2337.* FUMO, Nelson, MAGO, Pedro and LUCK, Rogelio.

6. Ding, Z., Wang, Z., Hu, T., & Wang, H. A comprehensive study on integrating clustering with regression for short-term forecasting of building energy consumption: case study of a green building. *Buildings.* 12(10), 1701, 2022.

7. *Digital Twin framework for automated fault source detection and prediction for comfort performance evaluation of existing.* Haidar, Hosamo Hosamo, et al. 2923, Vol. 281.

8. *A Digital Twin predictive maintenance framework of air handling units based on automatic fault detection and diagnostics.* Hosamo, Haidar Hosamo, et al. 2022.

9. Agency, International Energy. Heating. *iea.org.* [Online] https://www.iea.org/energy-system/buildings/heating.

10. *Perspectives on fourth and fifth generation district heating.* Henrik Lund, Poul Alberg Østergaard, Tore Bach Nielsen, Sven Werner, Jan Eric Thorsen, Oddgeir Gudmundsson, Ahmad Arabkoohsar, Brian Vad Mathiesen. 2021.

11. *Environmental optimization of the charge of battery electric vehicles.* Leon Zacharopoulos, Nils Thonemann, Marcel Dumeier, Jutta Geldermann. 2023.

12. *Mixed fleet scheduling method for airport ground service vehicles under the trend of electrification.* Dan-Wen Bao, Jia-Yi Zhou, Zi-Qian Zhang, Zhuo Chen, Di Kang. 2023.

13. *Study on orderly charging strategy of EV with load forecasting.* Wanjun Yin, Jianbo Ji, Tao Wen, Chao Zhang.

14. El Anshasy, A. A., & Katsaiti, M. S. Energy intensity and the energy mix: What works for the environment? *Journal of environmental management.* 136, 85-93, 2014.

15. Dong, F., Long, R., Li, Z., & Dai, Y. Analysis of carbon emission intensity, urbanization and energy mix: evidence from China. *Natural Hazards.* 2016, Vols. 82, 1375-1391.

16. Electric Vehicles. *iea.org.* [Online] 2023. https://www.iea.org/energy-system/transport/electric-vehicles.

17. *Machine Learning-Based Method for Remaining Range Prediction of Electric Vehicles.* Zhao, Liang, et al. 2020.

18. [Online] http://www.ndanev.com/.

Funded by
the European Union

19. *Implementation of machine learning based real time range estimation method without destination knowledge for BEVs.* H.A. Yavasoglu, Y.E. Tetik, K. Gokce. 2019.

20. Electricity Maps. Electricity Maps. [Online] [Cited: Jan 15, 2024.] https://app.electricitymaps.com/map.

21. European Commision. European Commision. *The European Green Deal.* [Online] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en.

22. *A Cost-Efficient Energy Management of EV Integrated Community Microgrid.* Divya Mathur, Neeraj Kanwar & Sunil Kumar Goyal. 2023.

23. *A multi-criteria decision model to support sustainable building energy management system with intelligent automation.* Muhammad Uzair, Syed Ali Abbas Kazmi. 2023.

24. *Joint scheduling optimization of a microgrid with integration of renewable energy sources and electric vehicles considering energy and reserve minimization.* Tao Hai, Jincheng Zhou, Jasni Mohamad Zain, Farah Jamali. 2023.

25. *HVAC systems testing and checking: A simplified model to predict thermal comfort in conditions with moderate environment.* Buratti C., Ricciardi P., Vergoni M. s.l. : Applied Energy, 2012.

26. *Towards the quantification of energy demand and consumption through the adaptive comfort approach in mixed mode office buildings considering climate change.* Sànchez-Garcìa D., Rubio-Bellido C., Martìn del Rìo J.J. s.l. : Energy & Builidngs, 2019.

27. Stavrakas, V., & Flamos, A. A modular high-resolution demand-side management model to quantify benefits of demand-flexibility in the residential sector. *Energy Conversion and Management.* 205, 112339, 2020.

28. Li, Y., O'Neill, Z., Zhang, L., Chen, J., Im, P., & DeGraw, J. Grey-box modeling and application for building energy simulations-A critical review. *Renewable and Sustainable Energy Reviews.* 146, 111174, 2021.

29. Lai, Y., Papadopoulos, S., Fuerst, F., Pivo, G., Sagi, J., & Kontokosta, C. E. Building retrofit hurdle rates and risk aversion in energy efficiency investments. *Applied Energy.* 306, 118048, 2022.

30. Hung, H. C., Yang, C. Y., Chien, C. Y., & Liu, Y. C. Building resilience: Mainstreaming community participation into integrated assessment of resilience to climatic hazards in metropolitan land use management. *Land use policy.* 2016, Vols. 50, 48-58.

31. Cerè, G., Rezgui, Y., & Zhao, W. Urban-scale framework for assessing the resilience of buildings informed by a delphi expert consultation. *International journal of disaster risk reduction.* 36, 101079., 2019.

32. Himoto, K., & Suzuki, K. Computational framework for assessing the fire resilience of buildings using the multi-layer zone model. . *Reliability Engineering & System Safety.* 2021, Vols. 216, 108023.

33. Lopez-Garcia, E., Lizana, J., Serrano-Jimenez, A., Diaz-Lopez, C., & Barrios-Padura, A. Monitoring and analytics ´ to measure heat resilience of buildings and support retrofitting by passive cooling. *Journal of Building Engineering.* 2022, Vols. 57, 104985.

34. Sun, K., Specian, M., & Hong, T. Nexus of thermal resilience and energy efficiency in buildings: A case study of a nursing home. *Building and Environment.* 177, 106842, 2020.

35. Menna, C., Felicioni, L., Negro, P., Lupisek, A., Romano, E., Prota, A., & Hajek, P. Review of methods for the combined assessment of seismic resilience and energy efficiency towards sustainable retrofitting of existing European buildings. *Sustainable Cities and Society.* 77, 103556, 2022.

36. Burroughs, S. Development of a tool for assessing commercial building resilience. *Procedia engineering.* 180, 1034–1043, 2017.

37. Duarte, M., Almeida, N., Falcao, M. J., & Rezvani, S. M. Resilience rating system for buildings against natural hazards. *In World Congress on Engineering Asset Management .* (pp. 57–68)., 2021, Vol. Springer.

38. Steadman, R. G. Norms of apparent temperature in Australia. *Aust. Met. Mag.* 43, 1–16., 1994.

39. Zargar, A., Sadiq, R., Naser, B., & Khan, F. I. A review of drought indices. *Environmental Reviews.* 19, 333–349., 2011.

40. Climate Change 2014: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. . *IPCC.* [Online] IPCC, 2014. https://www.ipcc.ch/site/assets/uploads/2018/02/WGIIAR5-FrontMatter.

41. EC. EU-level technical guidance on adapting buildings to climate change. 2023. Vol. doi:10.2834/558395.

42. FEMA. Building Codes Save: A Nationwide Study. Losses Avoided as a Result of Adopting Hazard-Resistant Building Codes. [Online] 2020. https://www.fema.gov/sites/default/files/2020-11/fema_building-codes-save_study.pdf.

43. Xiong, D., Cheng, K., & Chen, J. Climate-adaptive design strategies of sports stadia in a hot summer and cold winter zone: A case study of Nanjing. *Buildings.* 13, 2238., 2023.

44. Papadopoulos, A. M. Forty years of regulations on the thermal performance of the building envelope in Europe: Achievements, perspectives and challenges. *Energy and Buildings.* 127, 942–952., 2016.

45. Parliament, European. Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings (recast). [Online] 2010. http://data.europa.eu/eli/dir/2010/31/oj.

46. Programme, United Nations Environment. *2022 Global Status Report for Buildings and Construction: Towards a Zero-emission, Efficient and Resilient Buildings and Construction Sector. .* [Online] 2022. https://globalabc.org/sites/default/files/2022-11/FULL%20REPORT_2022.

47. IEA. Technology and Innovation Pathways for Zero-carbon-ready Buildings by 2030. . [Online] 2022. https://www.iea.org/reports/technology-and-innovation-pathways-for-zero-carbon-ready-buildings-by-2030.

48. Arivumani, V., Iyer, V., & Velraj, R. Evaluating the thermal transmittance of various building materials with insulating materials. 2019.

49. Dong, Y., Kong, J., Mousavi, S., Rismanchi, B., & Yap, P.-S.. Wall insulation materials in different climate zones: A review on challenges and opportunities of available alternatives. *Thermo.* 3, 38–65., 2023.

50. Lee, J., Kim, J., Song, D., Kim, J., & Jang, C. Impact of external insulation and internal thermal density upon energy consumption of buildings in a temperate climate with four distinct seasons. *Renewable and Sustainable Energy Reviews.* 75, 1081–1, 2017.

51. Berger, T., Amann, C., Formayer, H., Korjenic, A., Pospichal, B., Neururer, C., & Smutny, R. Impacts of external insulation and reduced internal heat loads upon energy demand of offices in the context of climate change in Vienna, Austria. *Journal of Building Engineering.* 2016.

52. Al-Homoud, M. S. Performance characteristics and practical applications of common building thermal insulation materials . *Building and environment.* 40, 353–366., 2005.

53. Ekici, B. B., Gulten, A. A., & Aksoy, U. T. A study on the optimum insulation thicknesses of various types of external walls with respect to different materials, fuels and climate zones in turkey. . *Applied Energy.* 92, 211–217., 2012.

54. Goia, F. Search for the optimal window-to-wall ratio in office buildings in different European climates and the implications on total energy saving potential. *Solar Energy.* 132, 467–492., 2016.

55. Alawadhi, E. M. Using phase change materials in window shutter to reduce the solar heat gain. *Energy and Buildings.* 47, 421–429., 2012.

56. Vutukuru, K. S., Moravej, M., Elawady, A., & Chowdhury, A. G. Holistic testing to determine quantitative wind-driven rain intrusion for shuttered and impact resistant windows. *Journal of Wind Engineering and Industrial Aerodynamics.* 206, 104359., 2020.

57. Zhang, C., Kazanci, O. B., Levinson, R., Heiselberg, P., Olesen, B. W., Chiesa, G., Sodagar, B., Ai, Z., Selkowitz, S., Zinzi, M. Resilient cooling strategies–a critical review and qualitative assessment. . *Energy and Buildings.* 251, 111312., 2021.

58. Energy, US Department of. Home Heating Systems. [Online] https://www.energy.gov/energysaver/home-heating-systems.

59. Vakiloroaya, V., Samali, B., Fakhar, A., & Pishghadam, K. A review of different strategies for HVAC energy saving. *Energy conversion and management.* 77, 738–754., 2014.

60. Bac, U., Alaloosi, K. A. M. S., & Turhan, C. A comprehensive evaluation of the most suitable HVAC system for an industrial building by using a hybrid building energy simulation and multi criteria decision making framework. . *Journal of Building Engineering.* 37, 102153, 2021.

61. Vakiloroaya, V., Samali, B., Fakhar, A., & Pishghadam, K. A review of different strategies for HVAC energy saving. *Energy conversion and management.* 77, 738–754., 2014.

62. Bac, U., Alaloosi, K. A. M. S., & Turhan, C. A comprehensive evaluation of the most suitable HVAC system for an industrial building by using a hybrid building energy simulation and multi criteria decision making framework. *Journal of Building Engineering.* 37, 102153, 2021.

63. Ahmed, T., Kumar, P., & Mottet, L. Natural ventilation in warm climates: The challenges of thermal comfort, heatwave resilience and indoor air quality. *Renewable and sustainable energy reviews.* 138, 110669., 2021.

64. Zaniboni, L., & Albatici, R. Natural and mechanical ventilation concepts for indoor comfort and well-being with a sustainable design perspective: a systematic review. *Buildings.* 12, 1983, 2022.

65. Frankoni, E., Hotchkiss, E., Hong, T., & Reiner, M. Enhancing Resilience in Buildings Through Energy Efficiency. [Online] 2023. https://www.energycodes.gov/sites/default/files/2023-07/Efficiency_for_Building_Resilience_PNNL-32727_Rev1.pdf.

66. Houghton, A., & Castillo-Salgado, C. Health co-benefits of green building design strategies and community resilience to urban flooding: A systematic review of the evidence. *International journal of environmental research and public health.* 14, 151, 2017.

67. Arnbjerg-Nielsen, K., Willems, P., Olsson, J., Beecham, S., Pathirana, A., Bulow Gregersen, I., Madsen, H., & Nguyen, V.-T.-V. Impacts of climate change on rainfall extremes and urban drainage systems: a review . *Water science and technology.* 68, 1, 2013.

68. Floodplain Hazard Management Regulations. [Online] 2021. https://www.fema.gov/sites/default/files/2020-07/fema_p-936_floodproofing_non-residential_buiildings_110618pdf.pdf..

Funded by
the European Union

69. EC. EU-level technical guidance on adapting buildings to climate change. 2023. Vol. doi:10.2834/558395.

70. Gikas, P., & Tchobanoglous, G. The role of satellite and decentralized strategies in water resources management . *Journal of environmental management.* 90, 144–152., 2009.

71. UK Department for Levelling Up & Communities. National Planning Policy Framework. [Online] 2023. https://assets.publishing. service.gov.uk/government/uploads/system/uploads/attachment_data/file/1182995/NPPF_Sept_23.pdf .

72. Pitt, M. Lessons learned from the 2007 floods. [Online] 2008. https://webarchive.nationalarchives.gov.uk/ukgwa/20100812084907/http:/archive.cabinetoffice.gov.uk/ pittreview/_/media/assets/www.cabinetofice.gov.uk/flooding_review/pitt_review_full%20pdf.pdf.

73. Ambiental. Sustainable Drainage Strategies, the pros and cons, what they do, and when you should be using them. . [Online] 2023. https://www.ambiental.co.uk/what-are-suds-sustainable-drainage-systems-guide/.

74. Burby, R. J., Deyle, R. E., Godschalk, D. R., & Olshansky, R. B. Creating hazard resilient communities through land-use planning. *Natural hazards review.* 1, 99–106., 2000.

75. Lewis, C. P., & Aghababian, R. V. Disaster planning, part i: Overview of hospital and emergency department planning for internal and external disasters. *Emergency Medicine Clinics.* 14, 439–452., 1996.

76. Bank, World. Finland - Climatology, Climate Change Knowledge Portal. *World Bank.* [Online] 2023. https://climateknowledgeportal.worldbank.org/country/finland/climate-data-historical.

77. Ruosteenoja, K., Raisanen, J., Jylha, K., Makela, H., Lehtonen, I., Simola, H., Luomaranta, A., & Weiher. Climate change estimates for Finland on the basis of global cmip3 climate models. *FMI reports.* 2013: 4., 2013.

78. Johansson, M. M., Pellikka, H., Kahma, K. K., & Ruosteenoja, K. Global sea level rise scenarios adapted to the Finnish coast. . *Journal of Marine Systems.* 129, 35–46, 2014.

79. Pilli-Sihvola, K., Harjanne, A., & Haavisto, R. Adaptation by the least vulnerable: Managing climate and disaster risks in finland. *International journal of disaster risk reduction.* 31, 1266–1275., 2018.

80. Angra, D., & Sapountzaki, K. Climate change affecting forest fire and flood risk—facts, predictions, and perceptions in central and south Greece . *Sustainability.* 14, 13395, 2022.

81. Weather2023. [Online] [Cited: November 8, 2023.] https://rp5.ru/.

82. Tolkien, John Ronald Reuel. *The Lord of Rings.* Crows Nest, New South Wales : George Allen & Unwin, 1954. pp. 1-5.

83. Nigel Bevan, Jim Carter, Jonathan Earthy, Thomas Geis & Susan Harker. What Are User Requirements? Developing an ISO Standard. *Human-Computer Interaction. Theories, Methods, and Human Issues .* 2018.

84. Guiney, Daryl Kulak Eamonn. *Use Cases: Requirements in Context.*

85. Home Cooling Systems. [Online] 2023. https://www.energy.gov/energysaver/home-cooling-systems.

86. Zaniboni, L., & Albatici, R. Natural and mechanical ventilation concepts for indoor comfort and well-being with a sustainable design perspective: a systematic review . *Buildings.* 12, 1983., 2022.

Funded by
the European Union