Energy Conversion and Management: X 21 (2024) 100512

Contents lists available at ScienceDirect

Energy Conversion and Management: X

journal homepage: www.sciencedirect.com/journal/energy-conversion-and-management-x

Building energy performance metamodels for district energy management optimisation platforms

Víctor F. Sánchez-Zabala [a,*], Tomás Gómez-Acebo [b]

[a] TECNALIA, Basque Research and Technology Alliance (BRTA), Bizkaia Science and Technology Park, Astondo bidea 700, Derio, Spain
[b] University of Navarra, Tecnun School of Engineering, Manuel de Lardizábal 13, 20018 San Sebastián, Spain

ARTICLE INFO

Keywords:
District modelling
Model predictive control
Metamodels
Ensemble learning
Stochastic optimization algorithms
Co-simulation

ABSTRACT

Reactive control strategies lack the flexibility necessary to optimize the operational costs of buildings and district systems. To overcome this limitation and to enable the transition to model predictive control strategies (MPC), the development of dedicated control platforms and models is required. Predictive models for district systems management should provide supply and demand side integrated modelling, high accuracy, generalization capacity and reduced computational times. However, traditionally available MPC solutions do not meet these requirements as simplified models offer short computational times but lack the required accuracy; detailed physics-based models provide satisfactory generalization but at the expense of high computational costs; and the generalization capacity of data models is constrained by the quality and availability of data. In contrast, metamodels developed through the combined use of physics-based models and machine learning techniques offer a powerful alternative at reduced computational cost. This paper describes an upgraded Integrated District Model concept developed through co-simulation coupling metamodels of buildings with a district heating infrastructure Modelica model. Furthermore, the process to produce the metamodels and optimization engine required to generate demand flexibility optimization functionalities for the buildings of the Stepa Stepanovic subnetwork (Belgrade) is depicted. Starting from the development of metamodels of instances of specific buildings (residential and educational use) the process was expanded to provide additional generalization to define, (1) a generic metamodel with the capacity to reproduce the behaviour of any instance of building of the residential typology, and (2) metamodels with generalization capacity in relation to operational settings. As part of this process the potential of several machine learning algorithms (e.g Support Vector Machines, etc) was evaluated including the latest ensemble boosting methods (e.g. Adaboost, Gradient Boosting and Extreme Gradient Boosting) with comparatively low use in the building simulation community. Finally, a virtual test bed consisting in metamodels coupled to an optimization engine based on genetic algorithms, was implemented, and compared to a traditional Physics-based model MPC solution (EnergyPlus-GENOPT), to evaluate the potential of the developed building level optimization functionalities. The metamodels and optimization engine were able to reproduce the optimized settings identified by the EnergyPlus-GENOPT MPC solution with cost savings potentials of 5–10%.

1. Introduction

The operation of most existing buildings and thermal networks is managed using reactive control strategies implemented through static rules. These strategies rely on fixed setpoints and schedules based on historical data to minimize service deficiency risks. While these strategies can ensure reliable and reasonably efficient system operation when control rules and setpoint values are properly defined, they lack the flexibility to adapt system settings in real-time without human intervention, to changing user behavior and external conditions (e.g. weather and energy prices, etc) [1]. Therefore, a transition to Model Predictive Control (MPC) solutions, is necessary [2,3]. MPC solutions involve dedicated control platforms and models that allow real-time evaluation of alternative operational criteria to dynamically maximize system performance and exploit distributed energy resources, while considering building and tariff flexibility and continuous commissioning functionalities.

In the context of building and district operation, MPC refers to an advanced approach to constrained control that utilizes a model of the

* Corresponding author.
E-mail addresses: vfrancisco.sanchez@tecnalia.com (V.F. Sánchez-Zabala), tgacebo@unav.es (T. Gómez-Acebo).

https://doi.org/10.1016/j.ecmx.2023.100512
Received 28 September 2023; Received in revised form 22 November 2023; Accepted 8 December 2023
Available online 10 December 2023

Nomenclature

ADA	Adaboost
AHU	Air Handling Unit
ANFIS	Adaptive Neuro-Fuzzy Interference System
ANNs	Artificial Neural Networks
BT	Boosted Tree
BTGP	Bayesian trees Gaussian process
CBECS	Commercial Buildings Energy Consumption Survey
CSI	Cubic Spline Interpolation
CPU	Central Processing Unit
DT	Decision Tree
DDM	Data Driven Model
DH	District Heating
DHW	Domestic Hot Water
DL	Deep Learning
DEAPS	Dwelling Energy Assessment Procedure software
EBT	Ensemble Bagging Trees
ELM	Extreme learning machines
ET	Extremely Randomized Trees
FFNN	Feed Forward Backpropagation Neural Network
FMI	Functional Mock-up Interface Standard
FMU	Functional Mock-up Unit
GA	Genetic Algorithm
GB	Gradient Boosting
GPR	Gaussian Process Regression
GP	Gaussian processes
GPE	Gaussian process emulator
GPU	Graphics Processing Unit
GRFM	Kriging
HEE	Heterogeneous Ensemble
HMPC	Hybrid Model Predictive Control
HOE	Homogeneous Ensemble
HVAC	Heating, Ventilation, and Air Conditioning
IDM	Integrated District Model
IBPN	Improved Back-Propagation Network
KNN	K-Nearest Neighbors
L	Metamodel for L – Shaped floor plan buildings
LOLS	Linear Ordinary Least Squares
LM	Linear models
LR	Lasso regression
LSTM	Long-Short-Term-Memory Neural Network
MARS	Multivariate Adaptive Regression Splines
ML	Machine Learning
MLP	Multi-Layer Perceptron
MLRM	Multiple Linear Regression Model
MPC	Model Predictive Control
MSA	Multi Step Ahead
NSGA-II	Non-Dominated Shorting Genetic Algorithm II
OLS	linear regression with ordinary least squares
PCA-ANN	Principal component Analysis Artificial Neural Networks
PSO	Particle Swarm Optimization
PR	Polynomial Regression
Q2	Second quartile
Q3	Third quartile
R	Metamodel for rectangular floor plan buildings
R + L	Metamodel for rectangular or L shape floor plan buildings
R_PLUS	Upgraded metamodel for rectangular floor plan buildings
R + L_PLUS	Upgraded metamodel for rectangular or L shape floor plan buildings
RF	Random Forest
RBFN	Radial-basis function networks
RC	Resistance Capacitance
ResNet	Residual network
RNNs	Recurrent Artificial Neural Networks
RVFL	Random Vector Functional Link
R^2	Coefficient of determination
SGD	Stochastic Gradient Descent
SRC	Standardized Regression Coefficient
SSA	Single Step Ahead
SVM	Support Vector Machine
SVR	Support Vector Regression
XGBoost, XGB	Extreme Gradient Boosting

supervised systems to optimize control inputs while minimizing a cost function (e.g., energy cost, etc) subject to certain constraints (e.g., thermal comfort, etc) [4]. During each sampling interval, an optimal control problem over a finite prediction horizon is solved based on future disturbances (e.g., energy prices, weather forecast, etc) to generate a trajectory of inputs and states that satisfies system dynamics and constraints while optimizing the cost function [5].

MPC solutions have demonstrated significant potential for improving the efficiency of building and district systems, including cost and peak demand reduction, increased renewable energy penetration and improved control precision and stability [6,7]. However, traditional deterministic MPC approaches assume perfect disturbance forecasts, disregarding uncertainties that can impact the optimization process when inaccurate input data is used. To address these uncertainties, alternative MPC formulations such as Stochastic MPC have been developed [6]. While this approach increases the complexity and computational cost of the optimization process, it enables the incorporation of uncertainties into the identification of control signals.

The second major challenge for MPC solutions, particularly in district-scale applications, is the cost associated with developing models while ensuring compatibility with real-time applications in terms of computational time. Ideally, predictive models for MPC solutions in district systems should integrate supply and demand side modeling, offer high accuracy and generalization capacity, and have reduced computational times. Two main types of models, physics-based models and data-driven models (DDMs), have been traditionally used in building and district domains, each with its own advantages and limitations.

Due to their detailed nature, physics-based models provide satisfactory accuracy and high generalization capacity for evaluating new operational strategies. However, this comes at the expense of high computational costs. Additionally, deep knowledge of the physical domain and the availability of information are required for model definition. There is abundant literature available on MPC implementations based on this type of model. In [8,9], a MPC solution combining EnergyPlus [10] and a Genetic Algorithm (GA) was proposed to enable multi-objective optimization of hourly set point temperatures with a day-ahead planning horizon. This solution was successfully applied to a multi-zone residential building in Naples. In [11], an MPC system based on an EnergyPlus energy prediction engine was utilized to optimize the control rules of an administrative building and minimize energy consumption. Additionally, [12] presented an MPC system based on a control-oriented dynamic thermal model. This system was tested in a TRNSYS [13]-MATLAB [14] co-simulation testbed and demonstrated that it could effectively utilize building thermal mass to optimize energy consumption during low-price periods.

When it comes to DDMs, a deep understanding of the physical domain and detailed information about the dynamic system's physical parameters become less important. These models are highly suitable for real-time applications due to their short computational times. However, their generalization capacity is constrained by the quality and availability of data. The MPC domain offers a diverse literature on the use of DDMs. For instance, in [15], an MPC system with Machine Learning (ML) based building models was proposed to control the Heating Ventilation and Air Conditioning (HVAC) system of a building located in Singapore. The system incorporated a multi-objective function to optimize energy efficiency and indoor thermal comfort. In [16] an approximate MPC system that emulated the dynamic behaviors of a building in Singapore equipped with an MPC system using a Recurrent Artificial Neural Network (RNNs) was described. Another study [17] proposed a system based on Support Vector Machines (SVM) classification to approximate the optimal behavior of an MPC controller for blind operation. In [18], Several ML algorithms were trained to model the HVAC system of office buildings using data collected at the Energy Resource Station of the Iowa Energy Center. The system incorporated a Particle Swarm Optimization (PSO) algorithm to optimize HVAC system settings. In [19], an MPC solution was defined to optimize the comfort settings of several rooms in the Saint George Hospital in Chania, Greece. The system incorporated dedicated Artificial Neural Networks (ANNs) to predict outdoor and indoor temperatures. Furthermore, [20] described an MPC solution based on ANNs models developed for the HVAC system of a residential house in Ontario. Nonetheless, owing to their limited ability to generalize, DDMs find their primary application in forecasting tasks, such as predicting building heating and cooling demands [21–23], estimating building energy consumption [24,25], and making predictions regarding district demands [26–28].

As an approach to maintain the generalization capacity of detailed physics-based models while addressing the computational time limitations, researchers have explored the potential of parallelization techniques and simplified physics-based models. Parallel computing enables optimization algorithms to distribute simulation tasks among multiple process threads simultaneously, thereby reducing the overall computation time. However, it requires advanced programming skills [29]. In [30], a high-performance cluster for parallel computing integrating the Non-Dominated Shorting Genetic Algorithm II (NSGA-II) with EnergyPlus was presented. The effectiveness of this method was demonstrated through its application to the optimal design of a typical single-family house. Moreover, the optimization time was reduced from almost 12 days to 4.4 h.

The use of simplified physics-based models is often considered as an alternative approach. However, it carries the risk of oversimplification, potentially leading to inaccurate modelling [29,31–33]. Despite these limitations, several successful applications can be found in the literature. In [4], a MPC solution was developed based on a Resistance Capacitance (RC) model to identify optimal comfort settings in a multizone building with the goal of minimizing energy consumption. The solution was tested in a two-month experiment conducted in an educational building in Prague with energy savings ranging from 15 % to 28 %. Another study [34] focused on the development of an ANNs model to predict the evolution of indoor temperatures in multi-zone buildings. The temperature predictions were utilized by a MPC solution based on a RC model to optimize HVAC control. In [35], a Hybrid Model Predictive Control (HMPC) scheme was developed to minimize operating costs of HVAC systems in commercial buildings. The HMPC model was developed using a RC model, while the nonlinearity associated with the HVAC process was handled by an ANNs model. Furthermore, [36] focused on an MPC energy management system based on a RC model for optimizing the operation of a residential building equipped with air-sourced heat pumps for heating and Domestic Hot Water (DHW) production.

Contrasting the limitations of simplified models and parallelization, the combination of detailed physics-based models with ML techniques offers a powerful alternative in the form of metamodels. These metamodels approximate the behavior of complex physics-based models at a reduced computational cost [33]. By utilizing a set of simulation input and output data from the original model, metamodels are generated through ML techniques. Once trained and validated, they can reproduce the behavior of the physics-based models and generate predictions, such as energy demand and internal temperatures, with reduced computational requirements [37–39]. Past research has successfully employed metamodels for predicting energy behavior in specific instances of buildings. In [40], ANNs, SVR, and Long-Short-Term-Memory Neural Network (LSTM) were proposed for simultaneous prediction of heating, cooling, and lighting loads. Another study [41] presented an ANNs model trained with synthetic data produced through a calibrated model for predicting energy consumption and thermal comfort levels in an indoor swimming pool. [42] focused on ML models to predict energy consumption. A virtual test case of a French residential building simulated in TRNSYS was evaluated, resulting in heating energy load predictions with values of up to 0.98 for the Coefficient of determination (R^2).

However, these surrogate models, specifically developed for individual buildings, are limited to the specific problem they were trained for. As a result, they lack the necessary generalization capacity to be reused for similar problems, such as alternative locations or different building geometries. Nonetheless, researchers are actively exploring ways to enhance the generalization capacity of metamodels to broaden their applicability. These efforts aim to extend their usage in areas such as building design optimization, building model calibration, and prediction of building stock energy behavior. The approach employed in these studies involves expanding the solution spaces used to train the metamodels by incorporating additional features, such as thermal envelope characteristics, orientation and other relevant factors specific to the desired application [29], in order to increase their generalization capacity. According to the existing literature, most metamodels developed for these applications rely on simplified geometric and architectural inputs, providing aggregated outputs with low temporal resolution, such as annual heating consumption. Additionally, there is a noticeable scarcity of studies focusing on building energy performance forecasting at the urban scale [43].

In [44], an approach combining metamodels and metaheuristic algorithms was developed to optimize the indoor thermal comfort and energy performance of residential buildings in Morocco. The solution was successfully tested on a residential building using a dataset generated with various variants of a TRNSYS model to train an ANNs metamodel. In another study [45], a solution for multi-use building performance optimization based on a Principal Component Analysis Artificial Neural Networks (PCA-ANN) metamodel coupled with a GA was presented. In [46], a combined physical and data-driven modeling approach was applied to predict the annual thermal energy use intensity of buildings at the stock level. The approach was successfully tested in a case study in Chongqing, China. In another study [31], a surrogate model based on a deep temporal convolutional neural network was developed with the generalization capacity required to predict heating and cooling demands for buildings exposed to different climatic conditions.

Similarly, metamodels with an appropriate level of generalization capacity offer an ideal alternative for MPC solutions in the building and district domain. However, research focusing on MPC solutions based on metamodels remains relatively scarce. In one study [47], an MPC solution based on RNNs was presented to enhance the operation of thermal substations. In another study [18] an ensemble of Multi-Layer Perceptrons (MLPs) was integrated into an energy optimization system including a PSO algorithm to optimize Air Handling Unit (AHU) settings. In [48], a MPC system for multizone HVAC systems was developed for a non-residential building in France. An ANNs model was constructed using a dataset generated from an EnergyPlus model. The ANNs model was coupled with a GA to enable real-time optimization of the operational schedules. Table 1 and Table 2 provide a summary of the key features of the analysed literature on the use of metamodels for the aforementioned applications.

In summary, the development of MPC solutions, especially at the district level, remains a complex technical field, and various limitations remain unresolved, including:

Table 1
References of metamodels use for building design/retrofitting optimization and model calibration.

Ref	Algorithm	Prediction horizon	Time step	Target	Domain	Application	Data Source	Testbed
[49]	MLP, ANNs, SVM, MARS, GPE	Monthly and annual	Month	Monthly and annual energy use intensity	Office buildings	Calibration	Model. EnergyPlus. medium office building Archetype	Virtual. EnergyPlus. medium office building Archetype
[50]	GPE	Annual	Annual	Annual normalized electricity and natural gas demand	Office buildings	Calibration	Model	Retrofitted Building
[51]	FFNN, LR	From yearly to 15 min	15 min	Building energy consumption	Residential buildings	Calibration	Model. EnergyPlus. Residential building	Virtual. EnergyPlus. Residential building
[52]	CSI	Annual	Annual	Annual energy consumption	Buildings	Design optimization	Model. EnergyPlus. Residential building	Virtual. EnergyPlus. Residential building
[53]	ANNs	Annual	Hourly	Annual energy load and summer comfort index	Residential buildings	Design optimization	Model. French residential building	Virtual. TRNSYS-GENOPT. French residential building
[54]	MARS	Hourly	Hourly	Annual indoor environmental indexes	Residential buildings	Design optimization	Model. EnergyPlus. Residential building in Hong Kong	Virtual. EnergyPlus. Residential building in Hong Kong
[55]	PR	Annual	Hourly	Annual Heating/cooling energy needs	Single family houses	Design optimization	Model. TRNSYS. Single family houses in six Moroccan climatic zones	Virtual. TRNSYS. Single family houses in six Moroccan climatic zones
[56]	ANNs	Annual	Hourly	Yearly Total Energy Consumption	Institutional buildings	Design optimization	Model. EnergyPlus. Institutional building in Montreal	Virtual. EnergyPlus. Institutional building in Montreal
[57]	SVR	Annual	Annual	Annual energy consumption	Multi use buildings	Design optimization	Model. EnergyPlus. Multi use buildings in Great Lakes (USA)	Virtual. EnergyPlus. Multi use buildings in Great Lakes (USA)
[58]	PR, GRFM, RBFN, MARS, SVM	Annual	Annual	Annual Primary Energy for Heating	Residential buildings	Renovation optimization	Model. TRNSYS. 3 residential building typologies in Milan	Virtual. TRNSYS. 3 residential building typologies in Milan
[59]	ANNs	Annual	Annual	Annual energy consumption intensity	Buildings	Retrofitting optimization	Model. TRNSYS school building in Coimbra (Portugal)	Virtual. TRNSYS school building in Coimbra (Portugal)
[60]	ANNs	Annual	Annual	Annual energy consumption for heating and cooling	Single-family buildings	Design optimization	Model. EnergyPlus. Single-family house in Argentina	Virtual. EnergyPlus. Single-family house in Argentina
[61]	LM, MARS, GP, SVM, BTGP	Monthly	Monthly	Monthly heating and cooling demand	Educational buildings	Design optimization	Model. ISO13790. 114 buildings in the University of Pennsylvania and 30 buildings in the Georgia Institute of Technology	Virtual. ISO13790. 114 buildings in the University of Pennsylvania and 30 buildings in the Georgia Institute of Technology
[62]	ANNs	Annual	Annual	Annual CO2 emission	Residential buildings	Retrofitting optimization	Real data (baseline) and simulated data (refurbishment scenarios). EnergyPlus. Group of buildings of the city of Zurich	Real data (baseline) and simulated data (refurbishment scenarios). EnergyPlus. Group of buildings of the city of Zurich
[63]	RBFN	Annual	Annual	Annual energy consumption	Buildings	Design optimization	Model. EnergyPlus. Buildings in Birmingham (UK), Chicago (USA) and San Francisco (USA)	Virtual. EnergyPlus. Buildings in Birmingham (UK), Chicago (USA) and San Francisco (USA)
[64]	IBPN	Annual	Hourly	Annual energy consumption	Residential buildings	Design optimization	Model. EnergyPlus. Residential building in China	Virtual. EnergyPlus. Residential building in China
[65]	DL	Annual	Monthly	Annual Heating and cooling energy consumption	Office buildings	Design optimization	Model. EnergyPlus. Office building in Brussels	Virtual. EnergyPlus. Office building in Brussels

- The high development cost and long computational times associated with detailed Physics-based models.
- The challenge of efficiently providing detailed physics-based demand-side modeling.
- The limited generalization capacity of DDMs.
- The lack of accuracy in simplified models.
- The cost and complexity of parallel computation solutions.
- The open challenges still existing in metamodel-based optimization methods.

In their previous work [90], the authors attempted to address some of these limitations, such as demand-side modeling and computational time reduction, through the concept of an Integrated District Model (IDM). The IDM utilized co-simulation techniques and a combination of data models and physics-based models (Modelica [91] and EnergyPlus) to maximize generalization capacity and reduce computational times simultaneously. The developed solution was tested in the Stepa Stepanovic subnetwork of the city of Belgrade, and in spite of an 85 % reduction of the physics-based building (EnergyPlus) models used in real-time the defined solution did not fully exploit the potential for computational time reduction.

Therefore, the objective of the present work is to develop an enhanced version of the IDM concept by coupling metamodels of buildings with a DH infrastructure Modelica model for use as an energy prediction engine in district MPC solutions. This upgraded model offers integrated supply/demand side modeling, improved generalization capacity, and reduced computational times, enabling the evaluation of numerous alternative operational strategies in real-time. To test this proposed solution, metamodels and optimization engines were developed to generate demand flexibility optimization functionalities for the

Table 2
References of metamodels use for energy demand forecasting at building, district and building stock level.

Ref	ML algorithm	Prediction horizon	Time step	Target	Domain	Application	Data Source	Testbed
[66]	SVR, ANNs	Hourly	Hourly	Heating/cooling loads	Buildings	Heating/cooling loads forecasting	Model. 768 variations of 12 buildings	Virtual. 768 variations of 12 buildings
[67]	ANNs	Annual	Annual	Annual cooling loads	Office buildings	Cooling load forecasting in different climates	Model. EnergyPlus. Office building in Brazil	Virtual. EnergyPlus. Office building in Brazil
[68]	MARS, SVM, GP, RF. ANNs	Annual	Annual	Annual cooling loads	Commercial building	Cooling load forecasting (generic buildings)	Model. EnergyPlus. Commercial building in Brazil	Virtual. EnergyPlus. Commercial building in Brazil
[69]	Clustering	Annual	Annual	Annual energy use intensity	Building groups	Energy use intensity forecasting	Model. EnergyPlus. Building	Real. 2646 buildings in Geneva
[70]	RF, KNN, ANNs	Annual	Annual	Urban energy use (operational and transportation)	Urban energy	Urban energy use forecasting	Real. The city of Chicago	Real. The city of Chicago
[71]	ANNs	Annual	Annual	Annual heating and cooling consumption	Residential buildings	Residential building labelling	Simulated. Brazilian regulation labelling dataset	Virtual. Brazilian regulation labelling dataset
[72]	ResNet	Hourly, daily, monthly, annual	15 min	Electricity use (single building, block, urban)	Buildings and building stock	Electricity use forecasting (single building, block, urban)	Simulated and real. University campus in California.	Real. University campus in California.
[73]	MLP, RF, SVM, GB	Annual	Annual	Annual building energy consumption	Commercial buildings	Annual building energy consumption forecasting	Real. National survey data from CBECS	Real. New York City Local Law 84 energy consumption dataset
[74]	LR, RF, SVR	Annual	Annual	Building annual electricity and gas use intensity	Buildings, districts and city	Annual electricity and gas use intensity forecast (building, district, city)	Real. Subset from 23.000 buildings of the city of New York	Real. Subset from 23.000 buildings of the city of New York
[75]	MLRM	Annual	Annual	Annual dwelling natural gas and electricity consumption	Residential buildings	Dwelling annual gas and electricity consumption forecasting	Real. GIS database of the city of Rotterdam	Real. GIS database of the city of Rotterdam
[76]	ANNs	Hourly and daily	Hourly	Electric load forecasting (individual and cluster level)	Heterogeneous buildings districts	Load forecasting (individual and cluster level)	Real. Urban educational district in Montréal	Real. Urban educational district in Montréal
[77]	RF, KNN RNNs, LSTM, MLP	15 min, 2 h, 24 h	15 min	Load forecast (individual building and cluster level)	Heterogeneous buildings districts	Load forecasting (individual and cluster level)	Real. Urban educational district in Montréal	Real. Urban educational district in Montréal
[78]	SVR, ANNs, RF, BT, GP	Daily	Hourly	Daily electricity forecast (individual and clusters)	Commercial buildings	Electricity demand forecasting (individual buildigns and clusters)	Real. 47 commercial buildings in Nederlands (2 years of data)	Real. 47 commercial buildings in Nederlands (1 year of data)
[79]	EBT	Hourly	Hourly	Hourly electricity demand	Institutional building	Forecast of electricity demand	Real. Institutional building in Florida	Real. Institutional building in Florida
[80]	PR	Annual	Daily	Annual heating and cooling intensity	Office buildings in Korea	Heating and cooling demand forecasting	Model. TRNSYS. Office building in Korea	Virtual. TRNSYS. Office buildings in Korea
[81]	OLS, RF, SVR, MARS, GPR, ANNs	Annual	Annual	Annual energy demand intensity	Tertiary buildings	Energy forecasting	Office Model (BSim) and educational building model (ISO 13790)	Virtual. Office Model (BSim) and educational building model (ISO 13790)
[82]	SRC, MARS	Annual	Annual	Annual gas consumption intensity	Secondary school buildings	Demand forecasting for Building stock	Model. EnergyPlus. School buildings in London	Virtual. EnergyPlus. school buildings in London
[83]	KNN, RF, GB, ANNs, DT	Annual	Annual	Annual primary energy use intensity	Residential buildings	Building stock refurbishment	Model. DEAPS. Irish building stock	Virtual. DEAPS. Irish building stock
[84]	MLRM	Annual	Annual	Annual gas and electricity consumption	Residential buildings	Dwelling gas and electricity consumption forecasting	Real data. GIS database of the City of Rotterdam	Real data. GIS database of the City of Rotterdam
[85]	ANNs	Annual	Annual	Annual heating energy use	Residential buildings	Heating energy use forecasting	Model. ESP-R. Different typologies of residential buildings	Virtual. ESP-R. Different typologies of residential buildings
[86]	SVR, ANNs, Ensemble	Hourly	Hourly	Cooling load forecasting	Office buildings	Cooling load forecasting	Model. TRACE 700. Dataset of 243 office buildings in Taiwan	Virtual. TRACE 700. Dataset of 243 office buildings in Taiwan
[87]	FFNN, RBFN, ANFIS, HEE	Daily	Daily	Daily heating consumption of the campus	University campus	Campus Heating forecasting	Actual data. Campus building in Sweden	Actual data. Campus building in Sweden
[88]	ANN, SVR, LOLS, KNN, HEE	Hourly	Hourly	Building level Hourly electricity demand	Buildings	Electricity demand forecasting (single generic building)	Real. 24 residential buildings and 8	Real. 24 residential buildings in California

(*continued on next page*)

Table 2 (*continued*)

Ref	ML algorithm	Prediction horizon	Time step	Target	Domain	Application	Data Source	Testbed
[89]	HOE of ANNs	Hourly	Hourly	Hourly cooling loads	Buildings	Cooling load forecasting (specific building instances)	university buildings (California) Model. Trnsys. Real commercial skyscraper in Hong Kong	Virtual. Trnsys. Real commercial skyscraper in Hong Kong

buildings in the Stepa Stepanovic subnetwork. Initially, metamodels were created for specific instances of residential and educational buildings. The process was then expanded to provide additional generalization, resulting in (1) generic metamodels capable of reproducing the behavior of any residential building instance, and (2) metamodels with generalization capacity for operational settings. To maximize the solution space covered in the dataset used to train the metamodels of residential building typology, an incremental process was implemented. Despite the limited availability of EnergyPlus models for the buildings in the Stepa Stepanovic neighborhood (Belgrade, Serbia), the incremental approach allowed the utilization of building simulation models from previous projects, thereby improving the accuracy of the metamodels without additional effort in defining physics-based models.

The development process of the metamodels involved evaluating the potential of various ML algorithms, including SVM, Random Forest (RF), k-nearest neighbors (KNN), Extremely Randomized Trees (ET), and different architectures of ANNs. Additionally, the applicability of ensemble boosting methods, such as Adaboost, Gradient Boosting (GB), and Extreme Gradient Boosting (XGBoost), which are relatively less commonly used in the building simulation community, was assessed. Boosting algorithms are gaining research interest in the building domain. However, further investigation is still required to assess the strengths and limitations of boosting methods in this domain, and the present work contributes to narrowing this knowledge gap.

Lastly, in order to facilitate the assessment of metamodels for optimizing building operations, a virtual testing environment was established. This environment incorporated metamodels coupled with a GA-based optimization engine. The control commands generated through this setup were then compared against those generated by a conventional physics-based model MPC solution.

The remaining sections of this paper are structured as follows: Section 2 describes the methods used to develop the metamodel-based IDM, while Section 3 describes the implementation in the buildings of the Stepa Stepanovic neighbourhood of the process defined to evaluate the performance of different ML algorithms in predicting building energy behavior, considering various levels of generalization capacity (specific building instances, residential typology, and demand flexibility optimization). The results of the testing processes for the defined models are discussed in Section 4. Finally, Section 5 concludes the paper.

2. Methods

2.1. Metamodels definition procedure

As previously described, metamodels serve as surrogates, approximating the behavior of detailed physics-based models at a reduced computational cost. Contrasting the limitations of simplified models, parallelization, detailed physics-based models and data models they offer a powerful alternative for MPC functionality development in the building and district domains. The workflow to produce metamodels through supervised ML techniques can be summarized in the following sequence:

- Identification of the alternative operational strategies relevant to the systems to be supervised.
- Definition of the feature space considering all the inputs that characterize buildings and systems dynamics.
- Feature engineering to optimize the feature space (e.g. dimensionality reduction through unsupervised methods, etc). An adequate selection of inputs is extremely important to define accurate and efficient metamodels. The inputs should be highly correlated with the predicted parameters, while having the least possible correlation with each other [43].
- Definition of representative samples of inputs to train and validate the metamodel, and calculation of the corresponding outputs through dedicated simulation sets of the detailed physics-based models.
- Training and validation of the metamodel selecting the most suitable supervised regression ML algorithm according to appropriate metamodel performance evaluation metrics.

Fig. 1 displays the workflow to develop building heating demand and interior temperature metamodels to be used in MPC systems.

Despite the potential displayed by metamodels, the strength and weakness of the metamodels based optimization methods is still a great research field within the building simulation community with several open challenges such as the number of building model evaluations required to construct and validate metamodels and the comprehensive evaluation of the accuracy and sensitivity of the optimization results [37,39,60]. For the sake of brevity, a comprehensive description of the ML algorithms available for metamodeling is beyond the scope of this paper. Therefore, the provided overview will be limited to ensemble methods, which are successfully applied in many domains, but are relatively scarce in use in the building/district domain. Additional information related to metamodeling techniques can be found in [92–96].

2.2. Ensemble methods

ML models can be classified as single or ensemble models based on their structure and the number of prediction algorithms used. The single prediction method relies on a single prediction algorithm, while the ensemble prediction method consists of multiple models with an integration process to achieve improved predictive performance. The main distinction between these two model types lies in the process of selecting and training the learning algorithms. While the single prediction method offers several advantages such as ease of implementation and fast computation, when compared to ensembles their main disadvantage is limited accuracy. The ensemble prediction method, on the other hand, offers improved prediction accuracy and stability. However, ensembles require more computation time and a higher level of knowledge.

In the case of ensembles, the base models are developed through resampling data [95]. The most common resampling techniques include Bagging, Stacking, and Boosting [97]. The bagging process starts by randomly selecting instances with replacement (Bootstrap) from the original training dataset. Each bootstrap sample is used to train a separate model, typically using the same algorithm. The predictions from all the individual models are then combined to obtain the final prediction. Bagging reduces the impact of individual model errors, increases the overall accuracy and stability of the ensemble and helps to reduce overfitting.

On the other hand, Boosting is an in-series ensemble that trains weak

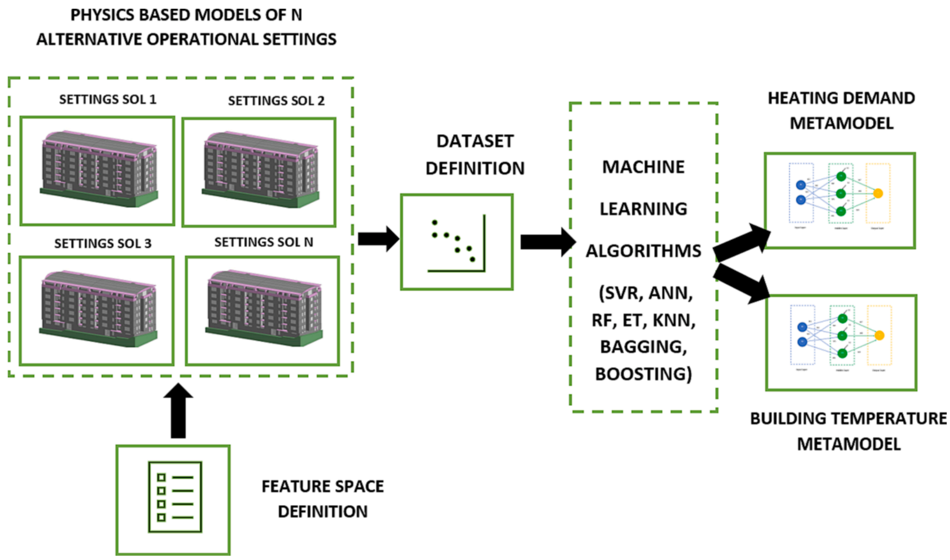


Fig. 1. Development procedure for the residential building flexibility optimization metamodels.

learners sequentially, where each subsequent learner focuses on the samples that the previous learners struggled to predict correctly. Initially, all instances are given equal weights, and a weak learner is trained on the weighted data. The subsequent weak learners are then trained on modified versions of the training data, with the weights adjusted to emphasize the poorly predicted instances from previous iterations.

Stacking involves training several base models on a subset of the dataset and then using their predictions as input for a final model, which learns how to combine these predictions. Base models are trained independently with specific subsets of the complete dataset and then used to make predictions on the remaining data. The final model is trained on these intermediate predictions along with the actual target values improving prediction accuracy by reducing bias and variance.

Ensemble models can be further classified into two types: homogeneous and heterogeneous, according to the selection of the base models. In the case of homogeneous ensembles, the same base learner is applied to different distributions of the training set. On the other hand, a heterogeneous ensemble consists of different learning models that are trained using the same data set. The heterogeneous ensemble improves prediction performance by leveraging the complementarity among different types of learning algorithms. In contrast, the homogeneous ensemble can be seen as an optimization process that enhances the performance of a specific learning algorithm by training it multiple times with varied datasets and combining their predictions [79].

According to [95], 91 % of ML-based building energy use predictions rely on single algorithms, while studies on ensemble prediction methods are limited to only 9 %. Therefore, further research to validate model superiority and computational efficiency is necessary before ensemble models can be widely adopted in the building/district energy domain.

Boosting ensemble algorithms such as AdaBoost, GB and XGBoost have been successfully applied to real-world problems with high predictive accuracy. However, again according to [95], there are only a few studies focused on the application of this type of homogenous ensemble models for building energy use prediction. For instance, in [98], twelve DDMs were developed to predict building thermal load. XGBoost was recommended for long-term predictions based on the obtained results. In [99] a data-driven workflow based in the XGBoost algorithm for office building performance assessment was constructed and successfully applied to an office building. In [100], an integrated optimization system combining a GA and GB surrogate model was used to achieve energy-optimal thermal designs for residential buildings in Turkey. This study highlighted the performance of the rarely used GB-based surrogate model in predicting building thermal loads. However, further investigation is still required to assess the strengths and limitations of boosting methods in this domain, and the present work aims to contribute to narrowing this knowledge gap. Table 3 provides a short overview of the most popular boosting methods. Further details related to boosting methods can be found in [101–104].

2.3. Metamodel-based integrated district model

Ideally, to develop MPC solutions for district systems, it is crucial to have an integrated model that can accurately predict all the phenomena affecting its behavior across various scales, including buildings, district infrastructure, and interfaces. This model should demonstrate a strong

Table 3
Overview of boosting methods.

Boosting methods	
Adaboost	Gradient Boosting and Extreme Gradient Boosting
• Operates by iteratively training a series of weak regression models, typically decision trees, on weighted subsets of the training data. Each weak model focuses on instances with high prediction errors from the previous models. During each iteration, the weak regression model is fitted to the weighted training data, and the weights are updated based on the model's errors, assigning more importance to poorly predicted instances. The final ensemble model is constructed by combining the predictions of all the weak models, weighted by their individual performance. Can deliver robust predictions, but it may be prone to overfitting. Therefore, it is crucial to employ proper regularization techniques.	• GB combines multiple weak prediction models, typically decision trees, to create a robust model. Models are trained sequentially, with each subsequent model aiming to correct the mistakes of the previous ones fitting them to the residuals of the previous models minimizing a loss function. By combining these weak models, GB can effectively capture complex patterns in the data, resulting in enhanced accuracy. GB offers several advantages, including regularization parameters to prevent overfitting and demonstrates high robustness and accuracy in various domains. Extreme Gradient Boosting, commonly referred to as XGBoost is a particularly powerful implementation of GB specifically designed to optimize performance. With its robustness, scalability, and high performance, XGBoost has proven to be very effective in diverse applications of ML.

capacity for generalization and enable real-time evaluation of a wide range of alternative operating scenarios, while maintaining efficient computation times.

However, the definition of a model with these characteristics remains an unresolved challenge. In their previous work [90], the authors introduced the concept of an IDM, capitalizing on the potential offered by co-simulation techniques and data models. The IDM allowed for the detailed physical modeling of both the demand side and district infrastructures by combining EnergyPlus models for representative buildings with a Modelica model of the thermal network infrastructure. This model offered high generalization and a significant reduction in the number of real-time simulated EnergyPlus models compared to conventional physics-based MPC solutions. However, this solution did not fully tap into the existing potential for reducing computation times.

To address this limitation, the current work has evolved the concept of an IDM by replacing EnergyPlus models with metamodels. This advancement allowed for the concurrent reduction of computation times while maintaining a significant portion of the model's generalization capability for both the demand and supply sides. The co-simulation solution was implemented using the capabilities of the Functional Mock-up Interface (FMI) standard [105]. The designed co-simulation scheme comprises metamodels that depict the energy behavior of each building, encapsulated as Functional Mock-up Units (FMUs) (slaves), and coupled with the Modelica model (master). Dynamic coupling is accomplished by exchanging relevant variable values at the interface (the substation of each building) between the master and slaves during time integration.

During each simulation time step, the energy demand from each building to the network is incorporated by considering the instantaneous setpoint temperature of the heating system, building temperature (B_TEMP), as well as the flow rate (RET_FLOW) and return temperature (RET_TEMP) of the heating circuit in the secondary side of building thermal substation. Using these values, the IDM enables the computation of the required flow rate through the primary side of each building's substation and the return temperature to the network. Precisely predicting the network's return temperature is crucial for optimizing its operation.

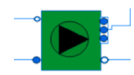
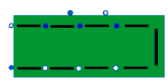
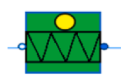
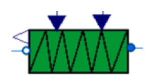
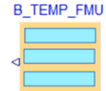
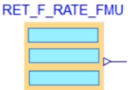
Likewise, for each calculation time step, the impact of the energy supply quality from the network on the comfort conditions of each building can be evaluated. This integration is accomplished by utilizing the temperature value at the secondary side outlet of each building's substation (SUP TEMP) as an input parameter for the metamodels.

Hence, in accordance with the designed co-simulation scheme, the energy performance of each building is characterized by three distinct metamodels. These metamodels facilitate the forecasting of the heating system's flow rate (RET_F_RATE_FMU), return temperature (RET_TEMP_FMU), and building indoor temperature (B_TEMP_FMU), considering predicted disturbances. Table 4 provides an overview of the scope and connectivity of the components of the Integrated District Model and Fig. 2 displays a simplified version of the architecture of the co-simulation scheme of the IDM, where the infrastructures (distribution network, heating plant, etc) are modelled using the DH Modelica library described in the previous work of the authors [90].

The purpose of the IDM is to serve as a predictive engine for MPC solutions for district energy networks. Based on predictions of weather conditions and energy prices, the MPC system will exploit the predictions (hourly resolution) provided, over a 24-hour prediction horizon, by the IDM to assess the energy consumption (thermal production and pumping) associated with different operational strategies in real-time. Through an optimization engine, the MPC system will be able to identify operational strategies that minimize the cost associated with the energy consumption of the DH system. Typically, the complete optimization process will be carried out on an hourly basis to minimize the impact of uncertainties associated with disturbances forecasting (e.g. weather conditions, etc).

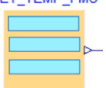
The conversion of the metamodels into FMUs is achieved through the

Table 4
Integrated district model components scope and connectivity overview.

Component	Component scope	Connectivity	Outputs
 HEATING_PLANT	District heating plant sub-system model including hot water generators, storage tanks, control valves and hydraulic pumps	DH heating plant sub-system model hydraulically connected to the distribution pumping station sub-system model	Energy consumption of DH heating plant hot water generators. Electricity consumption of DH heating plant pumps.
 PUMPING_STATION	DH heating distribution pumping station sub-system model including all the distribution hydraulic pumps	Pumping sub-system model hydraulically connected to the DH plant and distribution network sub-system models	Electricity consumption of distribution pumps.
 MAIN_LOOP	Distribution network sub-system model including pipes and the connection nodes of the buildings as necessary to define network topology	Distribution network sub-system model hydraulically connected to the pumping station and building thermal substation models	Water temperature distribution over the distribution network. Distribution thermal and pressure losses.
 SOLAR_PLANT	Solar plant sub-system model including solar collectors, storage tanks, control valves and hydraulic pumps	Solar plant sub-system model hydraulically connected to the distribution network sub-system model	Solar collector production. Electricity consumption of solar plant pumps.
 SUBSTATION_B1	Building thermal substation model including heat exchangers and control valves	Building thermal substation model hydraulically connected to the distribution network sub-system model and to building metamodels (encapsulated as FMUs)	Primary side (network) inlet/outlet temperatures and water flow rate. Secondary side (building) outlet temperature.
B_TEMP_FMU 	Building interior temperature metamodel (encapsulated as an FMU)	In order to receive the needed input, building interior temperature metamodel connected to the substation model (hot water supply temperature)	Interior temperature of the building
RET_F_RATE_FMU 	Building thermal substation secondary side flow rate metamodel (encapsulated as an FMU)	In order to receive the needed input, building substation secondary side flow rate metamodel connected to the substation model (hot water supply temperature) and to the building temperature metamodel (building interior temperature)	Water flow rate of the heating system of the building (water flow rate on the secondary side of the building thermal substation)

(continued on next page)

Table 4 (*continued*)

Component	Component scope	Connectivity	Outputs
RET_TEMP_FMU 	Building substation secondary side inlet temperature metamodel (encapsulated as an FMU)	In order to receive the needed input, building substation secondary side flow rate metamodel connected to the substation model (hot water supply temperature) and to the building temperature metamodel (building interior temperature)	Return temperature of the heating system of the building (inlet temperature to the secondary side of the building thermal substation)

use of the SimulatorToFMU software package [106]. Written in Python, this package allows users to export a Python-driven simulation program or script as a FMU that adheres to the FMI.

Due to space constraints, the comprehensive evaluation of the IDM's potential at the district scale will not be covered in this article. Instead, it will be the subject of a separate paper. As a consequence, this paper focuses on the development and experimental evaluation of the necessary components (metamodels and optimization algorithms) for implementing and exploiting the IDM. In order to assess the potential of the enhanced IDM, sections 3 and 4 will quantify the metamodels' capability to predict the energy behavior of buildings within the Stepa Stepanovic subnetwork. This evaluation focused on using metamodels to predict heat demand and indoor temperature in buildings connected to DH systems. The choice to utilize metamodels for heat demand, instead of the previously mentioned metamodels for flow rate and return temperature to the substation, was due to the limitation of available historical data, which only included heat demand information without corresponding profiles for flow rate and return temperature.

2.4. Optimisation techniques

Optimization algorithms can be categorized into various groups, including but not limited to deterministic or stochastic methods, derivative-based or derivative-free methods, trajectory or population-based methods, and single-objective or multi-objective algorithms [37]. Selecting the appropriate algorithm for a specific problem is crucial to achieve the highest accuracy and performance. The selection of optimization algorithms should consider both the nature of the problem and the features of the optimization method [33,37]. Generally, three criteria should be considered when selecting an optimization method: robustness, efficiency in terms of computational time and required memory, and accuracy. Additionally, the nature of design variables and objective functions, the presence of constraints, and the availability of derivatives of the objective functions should also be taken into account. In general, optimization algorithms can be classified into Enumerative, Calculus and Stochastic methods [33]. Their main characteristics are presented in Table 5.

From the perspective of their applicability to optimization problems in the building and district domain, the most relevant classification of optimization algorithms consists of gradient-based and gradient- free methods, as gradient-based methods can only be applied to smooth and continuous functions [109]. Since building energy behaviour is primarily nonlinear and discontinuous, MPC solution development should be based on gradient-free optimization methods. Consequently, and according to the features depicted in Table 5, evolutionary algorithms are applied to overcome the shortages of derivative-based methods.

According to the literature, the stochastic population-based algorithms are the most frequently used methods in building performance optimization [37]. Among these, the highest use increase was experienced by GAs, followed by PSO. The analysed literature was in agreement with the prevalence of GA and PSO methods. Therefore, GA and PSO were selected as the optimization methods to be used in this work.

Genetic algorithms mimic the process of biological evolution to find optimal or near-optimal solutions to complex problems. It is a population-based algorithm that falls under the category of global optimization algorithms, incorporating both global search operations for the optimal solution and local improvement operations. Further details related to GAs can be found in [117,118].

Similarly, PSO is inspired by the collective behavior of social organisms, such as bird flocking, fish schooling or swarming. It aims to find the optimal solution by simulating the cooperation of particles in a multi-dimensional search space. The process involves both individual intelligence and the communication between individuals, therefore PSO involves both local and global optimum finding processes. Further details related to PSO can be found in [119].

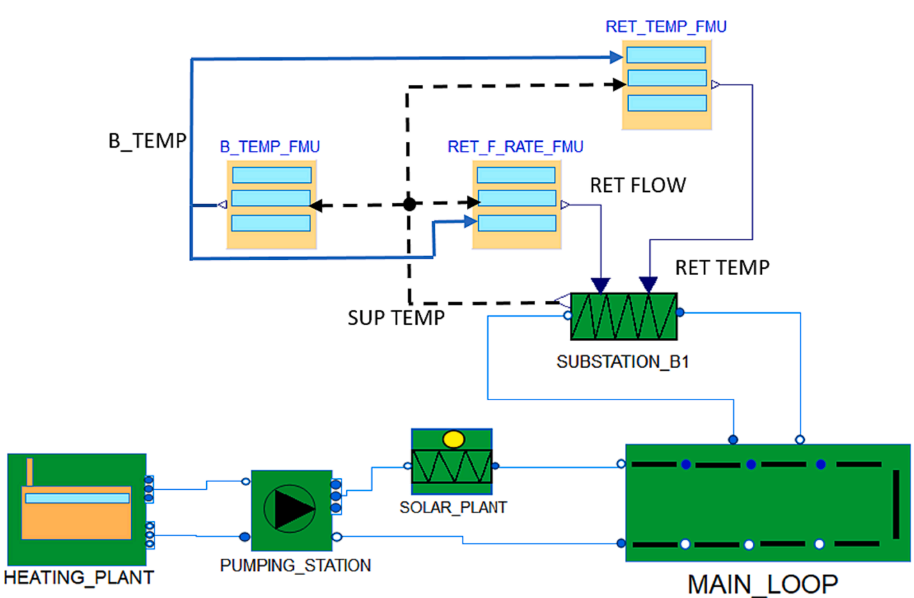


Fig. 2. Simplified architecture of the co-simulation scheme of the upgraded IDM.

Table 5
Overview of optimization methods.

Algorithms	Approach	Advantages	Limitations	References
Enumerative (Hooke and Jeeves, Brute Force, Derivative Free method, Simplex Nelder–Mead, Non-Linear programming, Hill Climbing Method, Pattern Search, Adaptive Search, etc)	Based on the analysis of all possible points in the search space	Finding the global optimum, given enough time and computational resources guaranteed. Simplicity and ease of implementation. Differentiability, convexity, or smoothness of the objective function not required. Applicable to nonlinear problems.	Can become computationally intractable or impractical when the search space is vast. Can get trapped in local optima.	Brute Force:[58] Derivative free method: [57] Simplex Nelder–Mead: [55] Non-Linear programming: [52]
Calculus or Gradient based (Newton's Method, Steepest Descent Method, Interior Point, etc)	Based on mathematical expressions or gradients	Fast convergence to the optimal solution exploiting information provided by the derivatives. Convergence to the global optimum guaranteed, for convex problems and to local optima for non-convex problems.	Inapplicable to most building studies due to the nonlinear and discontinuous nature of the problem	Interior Point: [57] and [107]
Stochastic (GA, PSO, Harmony Search, Ant Colony Optimisation, Simulated Annealing etc)	Randomness is included in the optimization process.	Offer a high probability of finding a near- optimal solution. Efficient even with non-differentiable and non-convex functions. Properly used can avoid getting stuck in local optima. Can make use of parallel or distributed computing techniques.	If not properly used risk of becoming trapped into local optimum instead of finding the global optimum. Their reliance on randomness can lead to different results on different runs.	GA: [108,109,48,8,60,58,56,110,111,112,59,44,45,64,53,10,63,113,29,19,114,100,99] PSO: [108,47,115,44,45,18,116] Harmony Search: [115]

3. Stepa Stepanovic use case

In this section, the process designed to evaluate the performance of the developed metamodels in predicting building energy behavior within the Stepa Stepanovic neighborhood will be described. The evaluation took into account different levels of generalization capacity, encompassing metamodels for specific building instances, metamodels for the residential typology, and metamodels for building demand flexibility optimization.

3.1. Vozdovac DH system and Stepa Stepanovic subnetwork

The Vozdovac system is one of the several heating networks that form the DH system of the city of Belgrade. It comprises of a DH plant and a two-pipe distribution network based on a branched topology and meets the energy demand of different areas of the city, including the Stepa Stepanovic neighbourhood. The latter is covered by the Stepa Stepanovic subnetwork which supplies the energy required to cover the heat demand of 52 residential buildings and two educational buildings (kindergarten and primary school) of the neighbourhood. A detailed description of the technical features of the Vozdovac system and of the buildings of the Stepa Stepanovic neighbourhood can be found in [120].

As described in the following sections, the evaluation process of the predictive capabilities of building surrogate models was completed taking advantage of the EnergyPlus models defined in the previous work of the authors [90] for each of the 6 residential representative buildings (1D, 2A, 3G, 4E, 6D and 6L) and the 2 educational buildings of the Stepa Stepanovic neighbourhood. Detailed information on the modelling specification followed to develop these models can be found in [120].

3.2. Testing scenarios

In this section, the scenarios defined to assess the metamodels' ability to predict the energy performance of buildings in the Stepa Stepanovic neighbourhood are described, including all the details related to defining the feature space, the process of generating datasets for training, testing, and validation, as well as the packages, algorithms, and metrics used for metamodel definition.

3.2.1. Metamodels for residential and educational building instances

The evaluation of metamodel potential involved developing specific metamodels for six residential buildings and two educational buildings within the demonstration district. Taking into account that the ultimate purpose of these metamodels is their integration into MPC systems, a time resolution (time step) of one hour and a predictive horizon of 24 hours were established, values typically used in MPC solutions.The analysis initially focused on a 1-hour prediction horizon and was later extended to a 24-hour ahead prediction horizon. The primary objectives of these metamodels were twofold: predicting the energy supplied from the DH to the buildings and forecasting the interior temperature of each building.

To assess the feasibility of metamodel development, diverse ML algorithms (SVR, RF, ET, KNN, ANNs) were applied to each representative building in the district. Additionally, special attention was given to boosting methods such as GB, XGBoost, and Adaboost during the evaluation. As a result, building-specific synthetic data for an entire heating season was used to generate the metamodels for each representative building. This synthetic data was derived from the EnergyPlus models of the representative buildings, which had been previously calibrated with real data in the authors' prior work [90]. The feature space of the dataset was carefully defined, drawing from the authors' experiences in their previous work [90]. Table 6 presents the parameters that comprised the initially employed feature space.

Given the high number of zones in the different buildings (ranging from 40 to 175), several approaches were explored to optimize the employed feature space. Finally, thermal zone setpoints and temperatures at the beginning of each time step were replaced with an equivalent setpoint and interior temperature for the entire building. The equivalent setpoint temperature was calculated as the volume-weighted mean value of the heating setpoint values of the heated zones of each building. The equivalent building temperature at the beginning of each time step was defined similarly. This resulted in a much more compact feature space without significantly impacting the model's accuracy, as depicted in section 4. Fig. 3 describes the complete process of generating the metamodels for residential building 1D, including the definition of the feature space, the generation of synthetic data for training, testing, and validation, and ultimately, the generation and evaluation of the metamodel's performance for each of the considered ML algorithms. This procedure was replicated for each of the residential and educational buildings.

To train the metamodels for each building, 75 % of the generated synthetic data was utilized, while the remaining 25 % of the data was reserved for the testing phase. The generation of synthetic data for training and testing was accomplished by leveraging climatic data from the EnergyPlus weather file for the city of Belgrade. Furthermore, a specific synthetic dataset was generated for validating the performance of the metamodels for each building. This validation dataset was created using the EnergyPlus models of the representative residential and educational buildings. In this case, the models were fed with weather data corresponding to another location in Serbia with weather conditions similar to those of Belgrade (Banja Luka). The produced datasets underwent a standardization process, where the values of different features were transformed to a range between 0 and 1 using scaling.

The hyperparameter tuning for different algorithms was carried out using a grid search procedure with cross-validation, typically employing 5 K-folds. For metamodels based on ANNs, the network architecture was defined through an iterative manual process, including determining the number of hidden layers and neurons per layer. Given the substantial number of buildings and ML algorithms considered, a large number of metamodels were developed. Therefore, providing detailed hyperparameter values for all the developed metamodels would be impractical for the sake of brevity. The development of metamodels was accomplished through dedicated ML platforms, including Scikit-learn [121], KERAS [122], Tensorflow [123], and XGBoost [124]. As the ultimate purpose of these metamodels was to be integrated into an MPC solution, their potential was evaluated based on accuracy and latency. To assess the performance of different ML algorithms R^2 [125] was selected.

Table 6
Feature space for metamodels development. Specific building instances.

Initial Feature Set		
Climatic	User behaviour and activity patterns	Ambient conditions
• Air dry bulb temperature Solar irradiation Wind speed and direction Absolute humidity	• Zone heating setpoint temperature Ventilation rates Occupancy patterns Day of the week, month, hour of the day	• Zone temperatures at the beginning of the prediction time step

Expanding on the metamodels developed for one-step-ahead predictions (hourly time resolution), Fig. 4 illustrates the recursive procedure established to generate multi-step-ahead forecasting capabilities necessary for providing 24-hour ahead predictions. As depicted in Fig. 4, in order to generate heating demand and temperature predictions for a specific timestep (N), it is necessary to have the corresponding input vector for that timestep, which includes the building's temperature at the beginning of the timestep, predicted by the temperature metamodel in the previous timestep (N-1).

3.2.2. *Metamodels for residential building typology*

As previously mentioned, developing metamodels for building typologies could offer improved efficiency by creating valid models that could be applied to any building within a specific typology. This evolution could allow a reduction of the required workload without significantly compromising the accuracy of the models. As a consequence, after evaluating the potential of metamodels in predicting the energy behavior of specific building instances, the possibility of expanding their predictive capacity was explored by defining dedicated metamodels for predicting the heating demand and interior temperature of any residential building equipped with hot water radiators and having either a rectangular or L-shaped floor plan (R + L). To achieve this, a procedure was developed to upgrade the generalization capacity of these metamodels as necessary to predict the energy behavior of buildings with diverse characteristics. In the initial step, additional parameters were incorporated into the feature space described in Table 6 to enhance the metamodels' generalization capacity concerning the architectural design of the buildings, including the following:.

- Heated area.
- Orientation and height.
- Long and short façades lengths and widths.
- Percentage of glazing area.
- Façade, roof, and slab thermal transmittances.
- Glazing Thermal and solar transmittances.

The metamodels for predicting heating demand and building temperature were trained, tested, and validated using synthetic data generated from 6 representative residential buildings in the district (1D, 6D and 6L with rectangular floor plans and 2A, 3G and 4E with L-shaped floor plans), using the EnergyPlus weather data file for Belgrade. To configure the training and testing datasets, synthetic data from 2 buildings with rectangular floor plans and 2 buildings with L-shaped floor plans were employed. The validation dataset, on the other hand, was generated using data from de two remaining buildings. Fig. 5 provides a simplified overview of the metamodels generation process for the residential building typology.

The analysis was expanded by defining 2 additional sets of metamodels for the residential typology, considering, on one hand, only buildings with a rectangular floor plan (R), and on the other hand, only buildings with an L-shaped floor plan (L). In both cases, a total of 3 buildings were available for training, testing and validation, with 2 buildings allocated for training and testing and 1 building for validation. The distribution of synthetic data for training and testing followed the same criterion as mentioned earlier.

To evaluate the impact on the model quality resulting from the inclusion of extra buildings in the datasets for model generation, the metamodel generation process was repeated for both the metamodel set for rectangular and L-shaped floor plan building typology (R + L_PLUS), and for the metamodel set for rectangular floor plan building typology (R_PLUS). In this iteration, synthetic data from 3 additional rectangular floor plan buildings (NSA1, NSA5, NSA7) obtained from a previous

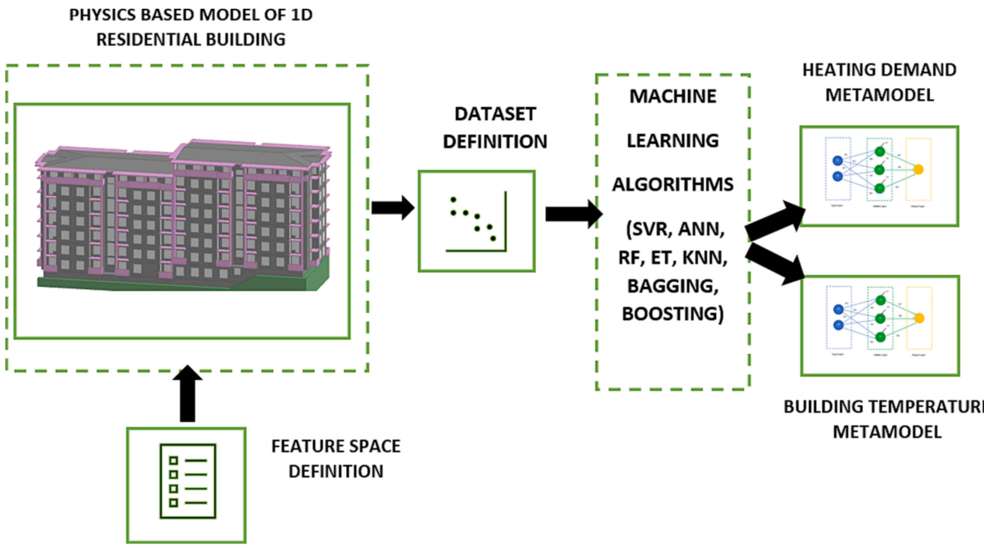


Fig. 3. Metamodels development procedure for specific building instances (residential and educational buildings).

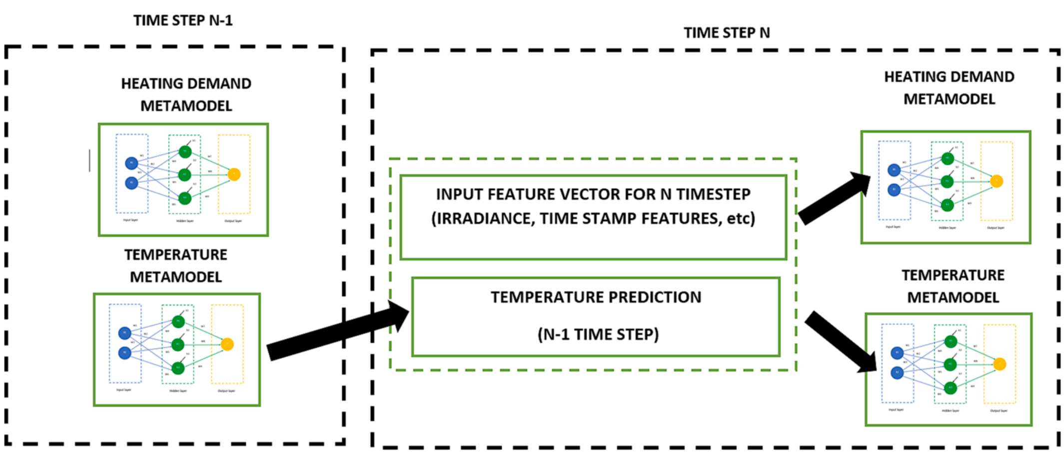


Fig. 4. Recursive procedure for multi-step ahead predictions.

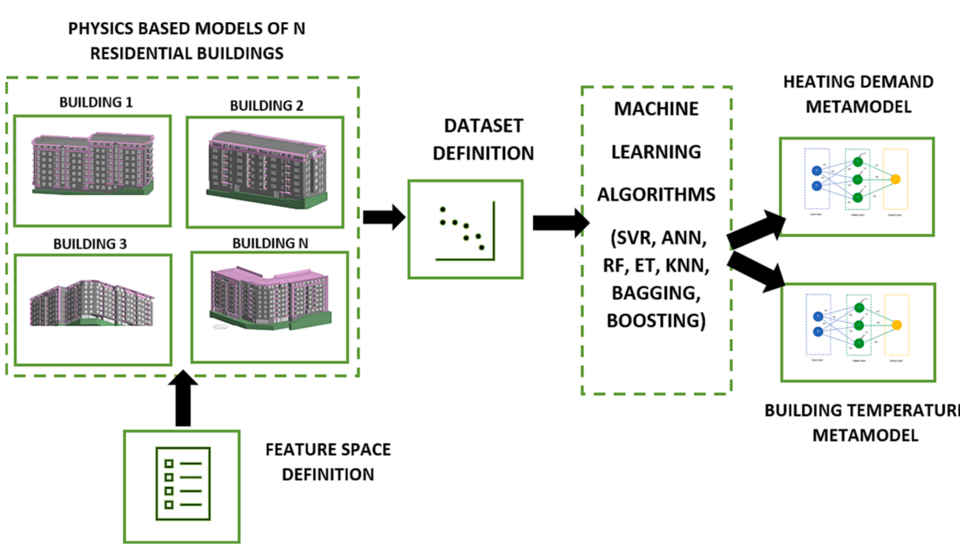


Fig. 5. Development procedure for the residential building typology metamodels.

project were added to the dataset after making slight modifications to align them with the architectural criteria of the Stepa Stepanovic buildings. Table 7 provides a summary of the evaluated metamodels, identifying the buildings used for training, testing, and validation.

To generate multi-step ahead predictions, the procedure described in the preceding section was followed. Similarly, the evaluation of metamodel performance considered both latency and accuracy. The process of normalizing the datasets, the ML algorithms employed, hyperparameter fine tuning, and the ML packages used to develop these metamodels followed the same criteria mentioned in the preceding section.

3.2.3. Metamodels for MPC functionalities

The ultimate purpose of the work described in this paper, and specifically of the IDM, is to develop MPC functionalities that can be applied at the building/district scale to optimize the operation of building/district systems. To achieve this, it is necessary to develop metamodels with sufficient generalization capability concerning the strategies and settings to be optimized in each case. In contrast to the metamodels analysed in the two preceding sections, which can predict the behavior of buildings operated based on predefined control logic and setpoints, the metamodels intended for MPC solutions must be capable of predicting the behavior of buildings operated under a wide range of alternative strategies or settings.

As an initial step to assess the feasibility of developing metamodels for MPC applications in the buildings/district domain, a building-scale proof of concept was conducted. Specific metamodels were developed for residential building 6L with the necessary generalization capability to predict the impact on the building's energy behavior resulting from optimal utilization of its own flexibility. Therefore, the developed metamodels provided the capacity to predict the energy behaviour of the building depending on the existing heating setpoint temperature. Since the ventilation of the residential buildings, including Building 6L, was solved through natural ventilation and considering the characteristics of their thermal envelopes, this proof of concept was carried out for a modified design of building 6L, adding an improved thermal envelope and a mechanical ventilation system with heat recovery to increase the capacity of the building to store energy. Furthermore, to moderately expand the limited optimization potential dictated by the prevailing meteorological conditions during the heating season (marked by minimal outdoor temperature fluctuations throughout the day), a flexible heat tariff was introduced. This tariff encompassed a 50 % reduction in the heat supply cost during nights and the central hours of the day. While flexible tariffs are not commonly adopted by district heating operators, the growing integration of renewables and the adoption of decarbonized heat production technologies, along with the transition toward low distribution temperatures, anticipate the increasing viability of such tariff structures in the future.

The dataset required to develop the aforementioned metamodels (training and testing) was generated by utilizing the results provided by the EnergyPlus model of building 6L adapted to its upgraded design. This involved a specific set of simulations carefully designed to provide the most accurate representation of the solution space while minimizing the number of required simulations. The designed set consisted of a total of 320 simulations. Each simulation was used to generate synthetic data corresponding to the building's energy behavior for a complete heating season, with a specific profile for the heating setpoint.

The setpoint temperature profiles were defined in a manner that enabled the generation of necessary synthetic data to describe the building's dynamics with the minimum number of profiles. Due to space limitations, a detailed description of these profiles is omitted here. Fig. 1 provides a simplified depiction of the process used to define the metamodels for optimizing the flexibility of building 6L, including dataset generation, evaluation of different ML algorithms, and the final definition of the metamodels.

The data from a series of preselected days was excluded from the training and testing datasets for later use in the optimization phase described below. The procedures for dataset normalization, data split to define training and testing datasets, generation of multi-step ahead predictions, hyperparameter fine tuning, used ML packages, and evaluation of metamodel performance were all carried out according to the approaches described in the two preceding sections.

In addition, optimization algorithms based on NSGA-II and PSO were defined, and the necessary testbed was implemented to integrate the metamodels with the optimization algorithms. The developed testbed enabled the identification of the heating setpoint hourly profile with a 24-hour prediction horizon, facilitating the optimization of the daily energy cost of the building. Thus, optimizing building 6L's flexibility became an optimization problem of the 24-value vector defining the heating setpoint hourly profile, using the daily energy consumption of the building as the objective function. To validate the developed optimization capabilities, days not used in the generation of training and testing datasets were employed. The testbed implementation was conducted in Python, utilizing the Pymoo [126] optimization package. Due to space limitations, specific details concerning the parameter values used to fine-tune the employed stochastic algorithms (e.g., number of generations, population size, etc.) are omitted. Fig. 6 illustrates the conceptual high-level architecture of the developed testbed, showing the coupling between the metamodels and optimization algorithms, as well as the main information flows involved in the optimization process.

For validation purposes and to establish the ground truth of the evaluated optimization problem, a second testbed was designed by coupling EnergyPlus models to the GenOpt [127] general-purpose optimization platform, thus emulating a physics-based model MPC solution. Fig. 7 illustrates the conceptual architecture of this testbed, depicting the coupling of the model and optimization algorithm (PSO), as well as the existing information flows.

Table 7
Summary of developed residential building typology metamodels.

METAMODELS FOR THE RESIDENTIAL BUILDING TYPOLOGY				
Metamodel	Metamodel code	Training and testing buildings	Validation	
			buildings	Code
Rectangular and L shaped floor plan	R + L	6D, 6L, 2A, 4E	1D, 3G	1D, 3G
Rectangular floor plan	R	6D, 6L	1D	1D_SPEC
L-shaped floor plan	L	2A, 4E	3G	3G_SPEC
Rectangular floor plan Plus	R_PLUS	6D, 6L, NSA1, NSA5, NSA7	1D	1D_SP_PLUS
Rectangular and L shaped floor plan Plus	R + L_PLUS	6D, 6L, 2A, 4E, NSA1, NSA5, NSA7	1D, 3G	1D_PLUS, 3G_PLUS

4. Results and discussion

In this section, the results derived from the application of the process outlined in Section 3.2 will be presented and analysed. This process aims to assess the potential of the developments required for the implementation of the IDM as presented in Section 2.2.

4.1. Building instance metamodels

As described in Section 3.2.1, firstly, based on the feature space outlined in Table 6, the feasibility of reducing the number of parameters considered in the feature space was evaluated. This reduction involved substituting the parameter set related to setpoint temperature and internal temperature of each thermal zone with an equivalent set of setpoint temperature and interior temperature for the entire building. This modification of the feature space enabled a significant reduction in the number of required parameters (depending on the building, from 78 to

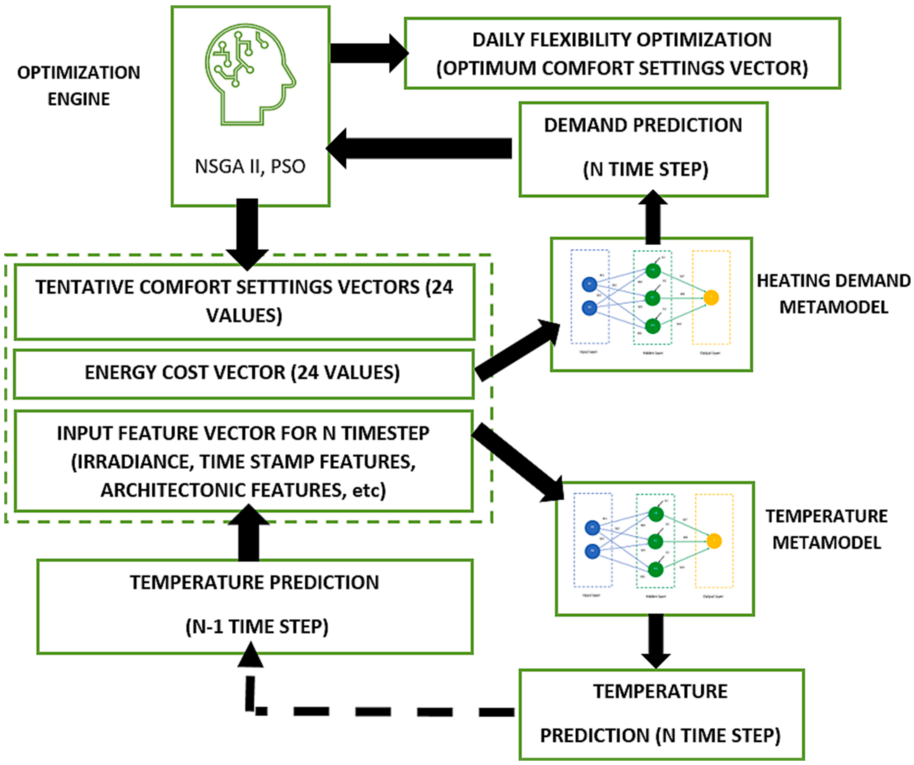


Fig. 6. Residential building flexibility optimization test bed.

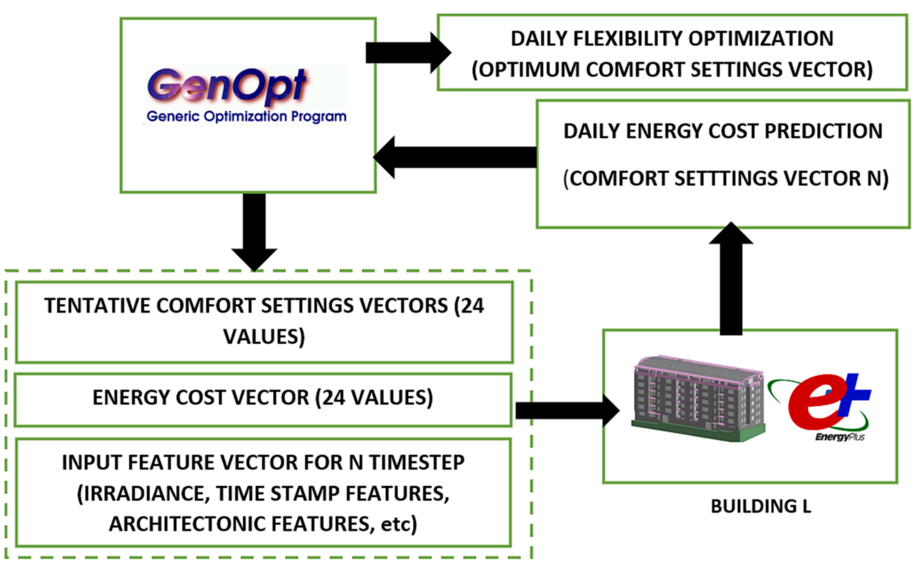


Fig. 7. Genopt-EnergyPlus building residential flexibility optimization test bed.

348). To evaluate this potential reduction, separate heating demand prediction metamodels were developed for the 1D residential building using different ML algorithms, employing both the initial feature space and the reduced feature space. Fig. 8 illustrates the comparison of R^2 values obtained from these metamodels for both the testing and validation datasets. With the exception of the KNN metamodel, it can be observed that the simplification introduced in the feature space produced a minor model prediction (SSA) accuracy degradation.

In terms of model fitting durations and latency for the testing dataset, in the case of 1D residential building, diminishing the feature space led to a reduction of model training time of between 42.81 % and 87,01 % and a decrease of latency of between 27,57 % and 95,22 %, depending on the used ML algorithm. Given the substantial frequency of re-fittings required for each model throughout the grid search with cross-validation process, the adoption of the reduced feature space had the potential to considerably decrease the necessary computation times. The exercise was replicated for the rest of the residential buildings, yielding analogous outcomes. Therefore, taking into account the modest reduction in algorithm precision and the potential cost savings in computation, the reduced feature space was adopted for the completion of the rest of the work.

Fig. 9 illustrates the distribution of R^2 values obtained for the heating metamodels (SSA prediction) of the residential buildings in relation to the applied algorithms, encompassing both the testing dataset (left) and

Fig. 8. Residential building 1D. Heating metamodels R^2. Testing dataset (left) and validation dataset (right). SSA prediction.

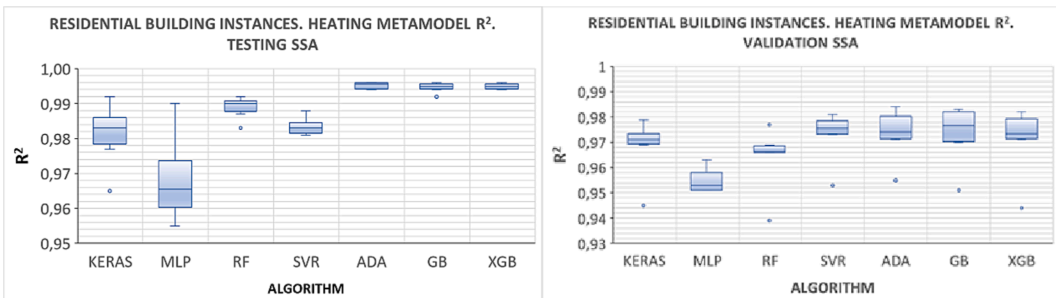


Fig. 9. Residential building instances. Heating metamodels R^2. Testing dataset (left) and validation dataset (right). SSA prediction.

the validation dataset (right). For both the testing dataset and the validation dataset, KNN yielded significantly inferior results compared to the rest of the considered algorithms. Specifically, in the case of the testing dataset, the values for the second (Q2) and third (Q3) quartiles were 0.964 and 0.968, respectively. For the validation dataset, the Q2 and Q3 values were 0.878 and 0.89, respectively. These values signify a significant decrease in comparison to the figures from the testing dataset. This is thought to be a result of an incomplete similarity between the climatic conditions of Belgrade and those in Banja Luka, potentially giving rise to an overfitting concern for a straightforward model like KNN. In any case, to facilitate the relative comparison of the rest of the considered algorithms, the results pertaining to KNN have been excluded from Fig. 9. As displayed there the rest of the methods provided high R^2 values for the testing and validation datasets. In the case of the testing dataset, boosting methods clearly display the highest performance with Q2 values of 0.996, 0.995 and 0.995 for Adaboost (ADA), GB and XGBoost (XGB) respectively. Among classic ML algorithms RF (0.99) and SVR (0.983) display the best performance, followed by ANNs methods (KERAS and MLP) with Q2 values of 0.983 and 0.966 respectively. If Q3 is adopted as comparison parameter almost the same algorithm relative performance is obtained with values of 0.997, 0.996, 0.996, 0.991, 0.986, 0.985 and 0.974 for ADA, GB, XGB, RF, KERAS, SVR and MLP respectively.

As already indicated for KNN, the results obtained for the validation dataset present a reduction of performance in relation to the testing dataset. However, in this case and owing to the much higher generalization capacity of the considered algorithms the observed reduction is moderate. However, this effect seems to affect the performance of the boosting methods with a slightly higher intensity. As a consequence, the prevalence of boosting methods in terms of performance, is less evident with SVR presenting almost comparable R^2 values. More specifically, in terms of Q2 values, GB (0.9765) presents the highest performance closely followed by SVR (0.9755), ADA (0.974) and XGB (0.973). The rest of the algorithms also display high Q2 values with KERAS (0.971), RF (0.966) and MLP (0.953) completing the performance evaluation. If Q3 is adopted to evaluate the relative performance of the considered algorithms, Boosting (GB, ADA, XGB) methods prevail followed by SVR with Q3 values of 0.982, 0.98, 0.979 and 0.978 respectively.

Fig. 10 illustrates the distribution of the MSA prediction R^2 values obtained for the heating (left) and indoor temperature (right) metamodels of the residential buildings. For both, KNN yielded significantly inferior results. Specifically, in the case of the heating metamodel, the values for Q2 and Q3 were 0.8755 and 0.883, respectively. For the interior temperature metamodel, the Q2 and Q3 values were 0.934 and 0.939, respectively. Therefore, in the case of the heating metamodel the existing performance decrease in relation to SSA prediction was moderate. In any case, to facilitate the relative comparison of the performance of the rest of the considered algorithms, the results pertaining to KNN have been excluded from Fig. 10.

As displayed in Fig. 10 the rest of the methods provide high R^2 values for the heating and interior temperature metamodels. In the case of the heating metamodel, in terms of Q2 values, the prevalence of boosting methods observed for SSA prediction disappears, and instead, the best performance is displayed by SVR (0.974), closely followed by ADA (0.969), GB (0.968), XGB (0.966) and MLP (0.966), with RF (0.952) offering the lowest performance. In relation to the results obtained for SSA prediction, these values represent a moderate, performance degradation. If Q3 is adopted as comparison criteria almost the same algorithm relative performance is obtained with values of 0.975, 0.972, 0.969, 0.969, 0.967, and 0.955 for SVR, ADA, GB, MLP, XGB and RF respectively.

In the case of the interior temperature metamodel, as displayed by Fig. 10 (right), in terms of Q2 values, XGB (0.9845) slightly outperforms SVR (0.984), ADA (0.983) and GB (0.983), with RF (0.98) and MLP (0.973) displaying the lowest performance. If, instead, Q3 is adopted as comparison parameter boosting methods prevail with values of 0.986, 0.986, 0.985, 0.985, 0.984, and 0.974 for XGB, ADA, GB, SVR, RF and MLP respectively.

Fig. 11 illustrates the distribution of R^2 values obtained for the heating metamodels (SSA prediction) of the educational buildings in relation to the applied algorithms, encompassing both the testing dataset (left) and the validation dataset (right). For both datasets, KNN yielded significantly inferior results. Specifically, in the case of the testing dataset, the values for Q2 and Q3 were 0.936 and 0.944,

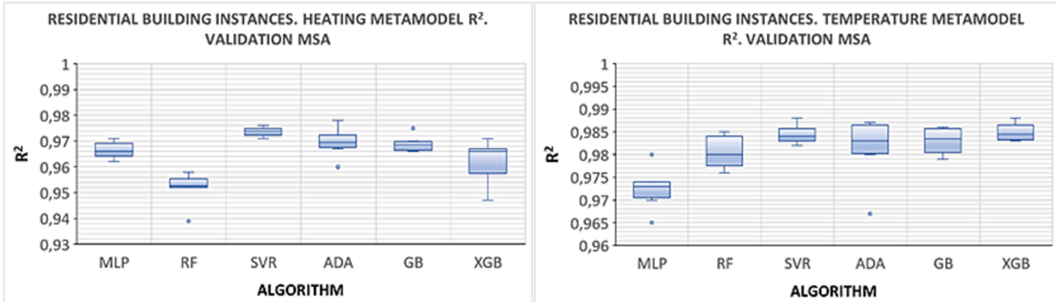


Fig. 10. Residential building instances. Heating metamodels R^2 (left) and indoor temperature metamodels R^2 (right). MSA prediction.

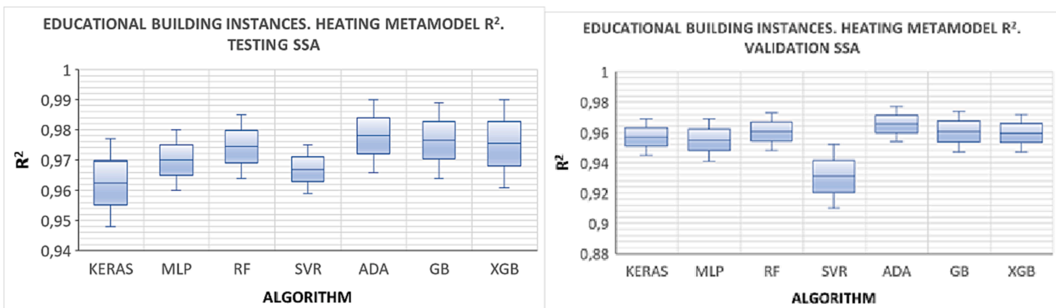


Fig. 11. Educational building instances. Heating metamodels R^2. Testing dataset (left) and validation dataset (right). SSA prediction.

respectively. For the validation dataset, the Q2 and Q3 values were 0.899 and 0.944, respectively. As already mentioned in the case of residential buildings the performance degradation displayed for the validation dataset is believed to be caused by some level of disparity between the climatic conditions of Belgrade and those of Banja Luka. For the reasons already stated KNN results have been excluded from Fig. 11. The rest of the methods provide high R^2 values for the testing and validation datasets. In the case of the testing dataset, as displayed by Fig. 11 (left), boosting methods clearly display the highest performance with Q2 values of 0.984, 0.982 and 0.982 for ADA, GB and XGB respectively. Among classic ML algorithms RF (0.974) and MLP (0.97) display the best performance, followed by SVR and KERAS with Q2 values of 0.967 and 0.962 respectively. If Q3 is adopted as comparison parameter the same algorithm relative performance is obtained with values of 0.984, 0.982, 0.982, 0.979, 0.975, 0.971 and 0.969 for ADA, GB, XGB, RF, MLP, SVR and KERAS respectively.

As already mentioned for the residential buildings, the results obtained for the validation dataset presented a moderate reduction of performance in relation to the testing dataset that affected the performance of the boosting methods with a slightly higher intensity. As a consequence, as displayed in Fig. 11 (right), the prevalence of boosting methods in terms of performance, is less evident with RF presenting comparable R^2 values. More specifically, in terms of Q2 values, ADA (0.965) presents the highest performance closely followed by GB (0.96), RF (0.96) and XGB (0.959). The rest of the algorithms also display high Q2 values with KERAS (0.957), MLP (0.955) and SVR (0.951) completing the performance evaluation. If Q3 is adopted to evaluate the relative performance of the considered algorithms, no modifications can be observed.

It is worth mentioning that in relation to the performance displayed by metamodels for residential buildings, in this case a moderate but relevant and consistent performance decrease was observed. This is consequence of the higher impact of using the reduced feature space on the educational buildings, derived from the existing significantly more complex and diverse activity patterns in relation to those of residential buildings.

Fig. 12 illustrates the distribution of the MSA prediction R^2 values obtained for the heating (left) and indoor temperature (right) metamodels of the educational buildings. For both KNN yielded significantly inferior results. Specifically, in the case of the heating metamodel, the values for Q2 and Q3 were 0,745 and 0,799, respectively. For the interior temperature metamodel, the Q2 and Q3 values were 0,853 and 0.883, respectively. For the reasons already mentioned the results for KNN have been excluded from Fig. 12. As displayed in Fig. 12 the rest of the methods provide relatively high R^2 values for the heating and interior temperature metamodels, but with a considerable performance loss in relation to the performance observed for SSA prediction. This effect is much more intense than in the case of the residential buildings. This is consequence of the higher impact of using the reduced feature space on the educational buildings derived from the existing significantly more complex and diverse activity patterns.

In the case of the heating metamodel, in terms of Q2 values, the prevalence of boosting methods observed for SSA prediction disappears, with MLP and SVR presenting comparable levels of performance. More specifically, the best performance was displayed by ADA (0.93), followed by MLP (0.928), GB (0.926), SVR (0.924) and RF (0.922), with XGB (0.921) offering the lowest performance. Although suboptimal these performance levels remained acceptable in relation to the existing goals. If Q3 is adopted as comparison parameter almost the same algorithm relative performance is obtained with values of 0.94, 0.939, 0.935, 0.933, 0.928 and 0.927 for ADA, MLP, GB, SVR, XGB and RF respectively.

In the case of the interior temperature metamodel, in terms of Q2 values, ADA (0.962) slightly outperforms XGB (0.960), SVR (0.960)and GB (0.953), with MLP (0.942) and RF (0.92) displaying the lowest performance. If, instead, Q3 is adopted as comparison parameter the relative performance of the considered algorithms remained the same with values of 0.972, 0.972, 0.969, 0.963, 0.952 and 0.935 for ADA, XGB, SVR, GB, MLP and RF respectively.

Fig. 13 depict the distribution of the R^2 value of the heating metamodels developed for the educational buildings according to the building type. The obtained results display the impact in terms of metamodels performance degradation associated to the complexity of the activities performed in the buildings. There, it can be observed that the

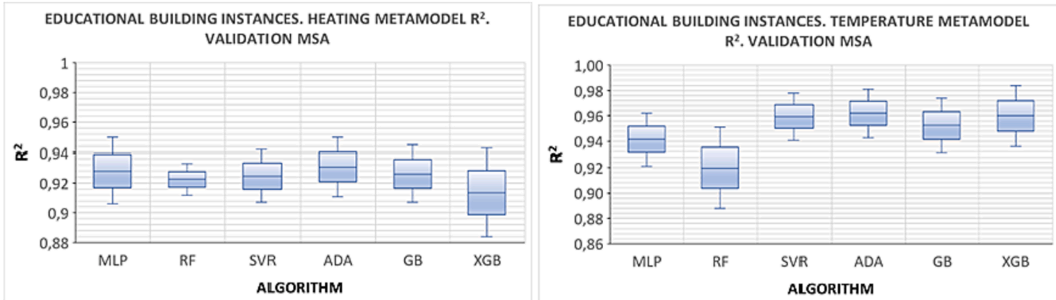


Fig. 12. Educational building instances. Heating metamodels R^2 (left) and indoor temperature metamodels R^2 (right). MSA prediction.

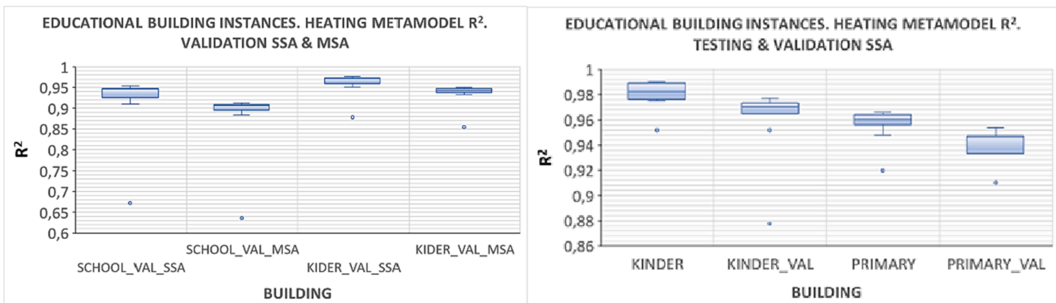


Fig. 13. Educational building instances. Heating metamodels R^2. Validation dataset SSA and MSA prediction (left). Testing and validation datasets SSA prediction (right).

performance results for the primary school, which displays significantly more complex activity patterns, are consistently below those obtained for the kindergarten.

Regarding the latency associated with the developed metamodels, it is important to note that the computation times presented by all of them are entirely suitable for real-time usage. For example, in the case of the residential building with the highest number of thermal zones (building 3G), except for ADA, the latency values corresponding to the testing dataset (881 instances) fall within the range of 0.018 (XGB) to 0.13 (SVR) seconds. ADA was clearly the slowest algorithm with a latency of 3.6 s.

According to the obtained results and performed analysis the following conclusions can be delineated:

- The proposed simplified feature space effectively captured the dynamics of energy behavior for specific instances of residential and educational buildings, reducing computation times with a moderate impact on model accuracy. However, as the considered building usage patterns become more complex, the impact of using the reduced feature space increased, though without compromising its validity.
- The developed metamodels were able to provide energy behavior predictions for specific instances of residential and educational buildings, with suitable levels of accuracy concerning the established objectives, both for 1-hour and 24-hour prediction horizons, with a moderate degradation in performance quality when extending the prediction horizon.
- With the exception of KNN, all evaluated ML algorithms exhibited high R^2 values. Particularly, boosting methods demonstrated a significant potential generally outperforming the rest of the considered algorithms.

4.2. Residential typology metamodels

Fig. 14 (left) displays the distribution of the R^2 value of the metamodels for heating (SSA prediction and testing dataset) developed for the residential typology with rectangular and L-shaped floor plan (R + L), the residential typologies with rectangular floor plans (R and R_PLUS), and the residential typology with L-shaped floor plan (L), respectively. It can be observed that in all cases, the values obtained for both Q2 (from 0.988 to 0.991) and Q3 (from 0.991 to 0.996) are very high, indicating a high level of precision for the developed metamodels. Regarding the Q2 value, all four considered models exhibited very high levels of accuracy without significant differences among them. On the other hand, when Q3 is used as the evaluation parameter, the metamodel for the residential typology with L-shaped floor plan (0.996) showed slightly higher levels of accuracy, closely followed by the rest of the metamodels with values of 0.992, 0.992, and 0.991 for the R_PLUS, R + L, and R metamodels, respectively.

In contrast, Fig. 14 (right) illustrates the distribution of the R^2 value obtained for the heating metamodels (SSA prediction and testing dataset) for the different considered residential building typologies, based on the algorithms employed for the development of these metamodels. It can be observed that for all considered algorithms, high accuracy values were achieved, with boosting methods prevailing over the others. In terms of Q2, ADA (0.996) exhibited the best performance, closely followed by GB (0.996), RF (0.991), ET (0.989), SVR (0.987), KERAS (0.979), MLP (0.972), KNN (0.973), and XGB (0.964). Using Q3 as the evaluation parameter, the relative performance levels of the considered algorithms was very similar, with the prevalence of boosting methods being somewhat more evident. In this case, GB (0.997) displayed the best performance, followed by ADA (0.996), XGB (0.996), RF (0.992), ET (0.99), KERAS (0.989), SVR (0.988), MLP (0.987), and KNN (0.976).

As displayed in Fig. 14 (right), in the case of R and R_PLUS building typologies, and in relation to the rest of boosting methods, XGBoost algorithm demonstrated a higher susceptibility to overfitting issues. This could cause an increased variance in the model and a notable loss of precision for the validation datasets (SSA and MSA), when compared to the rest of boosting algorithms. In order to achieve a better balance between bias and variance in XGBoost models for building typologies R and R_PLUS, a decision was made during the training process to adjust these models with slightly less flexibility. This adjustment resulted in

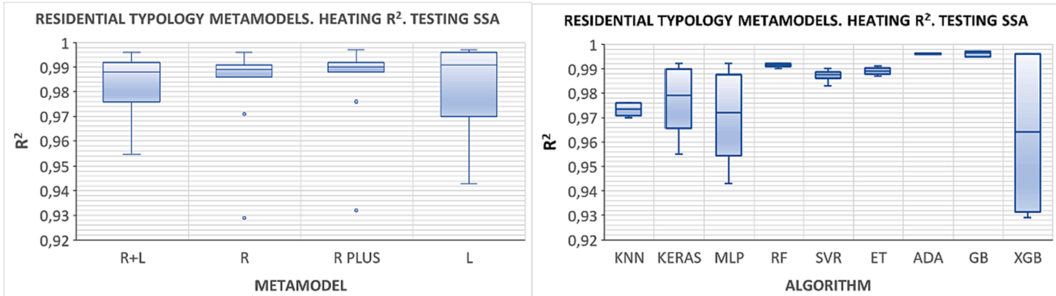


Fig. 14. Residential typology. Heating metamodels R2 vs building typology (left). Heating metamodels R2 vs ML algorithms (right).

higher bias in relation to the testing dataset (lower R^2 values). However, it enhanced the models' generalization capability for the validation datasets (lower variance). This phenomenon occurred to a lesser extent for the rest of building typologies, producing a greater dispersion in the R^2 values for XGBoost method compared to other boosting algorithms.

Fig. 15 (left) depicts the distribution of the R^2 value obtained for the heating metamodels (SSA prediction and validation dataset) for the different considered residential building typologies, based on the algorithms employed for the development of these metamodels. It can be observed that reasonably high performance values were achieved for all considered algorithms, but they were significantly lower than those observed for the testing dataset, with Q2 values ranging from 0.9 to 0.951 and Q3 values ranging from 0.919 to 0.964. It is believed that this effect was due to a suboptimal definition of the solution spaces associated with the different metamodels for the residential typology, resulting from the availability of a suboptimal number of buildings for training. Additionally, this effect was slightly more pronounced in the case of boosting methods, causing their prevalence relative to other methods to be reduced. In terms of Q2, XGB (0.951) exhibited the best performance, closely followed by ET (0.945), ADA (0.938), SVR (0.935), GB (0.931), RF (0.925), MLP (0.917), KERAS (0.907), and KNN (0.9). Using Q3 as the evaluation parameter, the relative performance levels of the considered algorithms underwent some modifications, further aligning the performance of the boosting methods with that of the rest. In this case, KERAS (0.964) presented the best performance, followed by ET (0.959), ADA (0.952), XGB (0.952), MLP (0.952), GB (0.951), SVR (0.945), RF (0.9251), and KNN (0.919).

On the other hand, Fig. 15 (right) displays the distribution of the R^2 value obtained for the heating metamodels (MSA prediction and validation dataset) for the considered residential building typologies, based on the algorithms employed for the development of these metamodels. It can be observed that, except for MLP and RF, reasonably high precision levels were obtained for all the considered algorithms, with Q2 values of up to 0.948 and Q3 values of up to 0.959. Additionally, it can be noted that in relation to the SSA prediction in Fig. 15 (left), the loss of metamodel performance was very moderate. In terms of Q2, XGB (0.948) presented the best performance, followed by ADA (0.943), SVR (0.931), GB (0.93), ET (0.918), KNN (0.897), MLP (0.871), and RF (0.841), which yielded the worst results. Using Q3 as the evaluation parameter, the relative performance levels of the considered algorithms underwent minimal changes, with a slight increase in the prevalence of boosting methods. In this case, XGB (0.959) presented the best performance, followed by ADA (0.959), GB (0.943), SVR (0.940), ET (0.932), RF (0.931), MLP (0.922), and KNN (0.913).

Fig. 16 (left) displays the distribution of the R^2 value for the validation buildings summarized in Table 7 (1D and 3G, 1D_SPEC, and 1D_SP_PLUS and 3G_SPEC) used in the development of the heating metamodels (SSA prediction) for the residential typology (R + L, R and R_PLUS and L). It can be observed that, in all cases the values obtained for both Q2 (from 0.915 to 0.959) and Q3 (from 0.938 to 0.964) were acceptably high, but displayed a significant performance degradation in relation to the values observed for the testing dataset. As already mentioned this degradation was caused by the limitations on the definition of the solution spaces of these metamodels, consequence of the suboptimal number of buildings available for training. The sensitivity of the metamodel performance to the number of building models available for training can be deduced comparing the performance provided by the metamodels for building 1D_SP_PLUS (R_PLUS metamodel) and building 1D_SPEC (R metamodel). The addition of 3 extra models to train the R_PLUS metamodel enabled an increase in Q2 from a 0.919 value for building 1D_SPEC to a value of 0.959 in the case of building 1D_SP_PLUS. If Q3 is used as evaluation parameter, it grows from 0.951 for building 1D_SPEC to a value of 0.964 for building 1D_SP_PLUS. On the other hand, according to the Q2 and Q3 values observed for 1D (0.931 and 0.938) and 3G (0.915 and 0.945) buildings the R + L metamodel displayed a similar performance for both rectangular plan buildings and L shape plan buildings. Finally, according to the Q2 and Q3 values displayed by buildings 1D (0.931 and 0.938) and 1D_SPEC (0.919 and 0.951) and buildings 3G (0.915 and 0.945) and 3G_SPEC (0.925 and 0.952), the performance of the R metamodel and of the L metamodel seemed to be slightly higher than that provided by the R + L metamodel.

Similarly, Fig. 16 (right) displays the distribution of the R^2 value for the validation buildings summarized in Table 7 (1D and 3G, 1D_PLUS and 3G_PLUS, 1D_SPEC and 1D_SP_PLUS and 3G_SPEC) used in the development of the heating metamodels (MSA prediction) for the

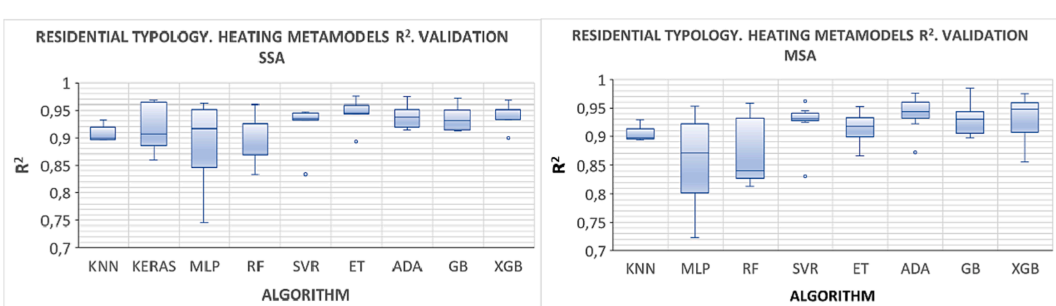


Fig. 15. Residential typology. Heating metamodels R^2. SSA prediction (left). MSA prediction (right). Validation dataset.

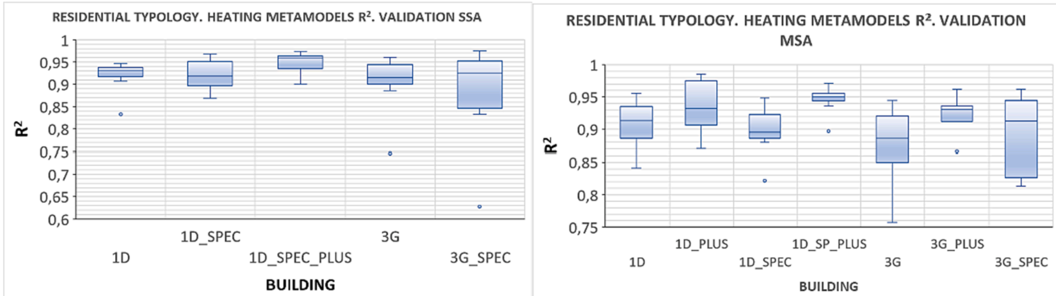


Fig. 16. Residential typology. Heating metamodels R^2. SSA Prediction (left) and MSA prediction (right). Validation dataset.

residential typology (R + L and R + L_PLUS, R and R_PLUS and L). It can be observed that, excluding buildings 3G and 1D_SPEC, in all cases the values obtained for both Q2 (from 0.886 to 0.95) and Q3 (from 0.92 to 0.975) were acceptably high and the performance degradation in relation to the SSA prediction moderate. The impact associated with the availability of 3 additional models for the training of the R + L_PLUS and R_PLUS metamodels can be observed by comparing the Q2 and Q3 values of the 1D (0.914 and 0.935) and 3G (0.886 and 0.920) buildings with those of the 1D_PLUS (0.932 and 0.975) and 3G_PLUS (0.930 and 0.936) buildings, and the values of the 1D_SPEC building (0.895 and 0.922) with those of the 1D_SP_PLUS building (0.950 and 0.955). In all cases, the availability of even a limited number of additional models allowed for a significant increase in the metamodels' performance. On the other hand, according to the Q2 and Q3 values obtained for 1D (0.914 and 0.935) and 3G (0.886 and 0.920) buildings in the case of MSA predictions the R + L metamodel seemed to display a higher performance predicting the behaviour of rectangular plan buildings.

Regarding the latency associated with the developed metamodels, it is important to note that the computation times presented by all of them were entirely suitable for real-time usage. For example, in the case of the R + L residential building typology, except for ADA, the latency values corresponding to the testing dataset (3524 instances) fall within the range of 0.016 (ET) to 4.16 (SVR) seconds. ADA was clearly the slowest algorithm with a latency of 23.4 s. These computation times were obtained using a portable workstation equipped with an Intel(R) Core(TM) i7-9750H Central Processing Unit (CPU) and 32 GB of RAM, exclusively utilizing the computing capabilities of its CPU without resorting to Graphics Processing Unit (GPU) capabilities. According to these preliminary results, with a suitable computing infrastructure and model implementations tailored to exploit GPU computing capabilities, computation times will be compatible with the integration of such models into predictive engines for district MPC solutions, even with an arbitrarily high number of connected buildings. In any case, as mentioned earlier, the utilization of the IDM in MPC solutions for DH systems will be the subject of a comprehensive district scale study in a subsequent specific paper.

According to the obtained results and performed analysis, as a summary, the following conclusions can be compiled:

- The degradation in model accuracy between the testing datasets and the validation datasets indicated that the number of buildings available for model generation (6) was suboptimal to achieve a sufficiently precise solution space definition to fully exploit the potential of metamodels in terms of precision.
- Despite this, for all the evaluated residential building typologies, several of the considered algorithms allowed for the generation of metamodels with the capacity to provide 24-hour ahead predictions, with acceptable levels of precision relative to the set objectives. Particularly, boosting methods demonstrated a significant potential in general outperforming the rest of the considered algorithms.
- For all the considered building typologies, the availability of even a small additional number of models (3) for metamodel training led to significant increases in their performance. Therefore, the proper definition of the solution space was the critical aspect for harnessing its potential. As a consequence, this process could potentially serve to progressively generate metamodels for residential typologies with increasing accuracy and generalization capability over time.

4.3. MPC metamodels

As described in the previous sections, boosting methods demonstrated significant potential for the development of metamodels. Therefore, for the development of metamodels aimed at optimizing the flexibility of specific instances of residential buildings, only boosting methods, as well as some stacking-based solutions, were investigated. However, stacking-based solutions were ultimately discarded since the marginal improvements in accuracy did not outweigh the lengthy training times resulting from the size of the dataset used (324.406 instances). Furthermore, the latency of these solutions was significantly higher than that of the rest of the methods, slowing down the real-time identification of optimal setpoints.

As described in Section 3.2.1, the generation of building heating demand predictions for time horizons greater than 1 hour is based on a recursive procedure involving the coupling of a metamodel to predict the indoor temperature of the building and another metamodel to predict building heating demand. From a purely predictive perspective the best performance, among all the considered algorithm combinations, was delivered by a combination of XGBoost heating demand metamodel ($R^2 = 0.978$. SSA prediction) and a nested Adaboost ($R^2 = 0.995$. SSA prediction) temperature metamodel. This combination displayed the best balance of high predictive precision levels over the complete solution space relevant for the addressed optimization problem. The nested AdaBoost model was created by employing a standard AdaBoost algorithm, using an AdaBoost model as the base regressor instead of the more commonly used decision tree-based regressor. While the implementation of the nested AdaBoost model led to longer computation times compared to other algorithms discussed in earlier sections, it became particularly important in optimizing building flexibility. Accurately predicting the indoor temperature evolution during the central hours of the day, when the heating system is inactive, proved essential for precise forecast of the building's heating demand over prediction horizons exceeding 1 hour. In any case, as already mentioned, with suitable computing infrastructures and model implementations tailored to harness the computing capabilities of GPUs, the computation times associated with these models will be compatible with their integration into MPC solutions for large-scale DH systems.

The evaluation of the simultaneous capacity of metamodels to predict building behavior and of optimization algorithms to identify optimum heating setpoint temperature profiles, was carried out by replicating the system's operation for a series of validation days not included in the training and testing datasets. To identify the optimal setpoint temperature profile, the optimization engine calculates, for each of the alternative operational settings, the cost associated with heating energy consumption based on the heating demand predictions provided by the metamodels, applying the district heating supply tariff as outlined in Section 3.2.3.

As an illustrative example, the results corresponding to January 10th are included, which can be considered representative of the performance provided by the developed solution for the entire set of evaluated validation days. Fig. 17 (left) displays the base hourly temperature setpoint profile for January 10th. This profile consists of a setback temperature of 16 °C maintained during the night and the central part of the day characterized by the lack of occupancy in the building, and a setpoint temperature of 20 °C maintained during the hours when the apartments remained occupied (from 7:00 to 10:00 h and from 17:00 to 24:00 h). On the other hand, Fig. 17 (right) shows the hourly heating demand profile of building 6L for January 10th (E +) corresponding to the base setpoint temperature profile, as well as the SSA and MSA predictions provided by the heating metamodel. It can be observed that the SSA predictions accurately replicated the demand profile throughout the day, including the peak load occurring at 17:00 h. The MSA predictions also reasonably captured most of the profile but did not achieve sufficiently precise prediction of the peak occurring at 17:00 hours.

Fig. 18 (left) depicts the comparison between the optimized hourly setpoint temperature profile for January 10th (E_PLUS_GENOPT) and the approximation provided by the system developed based on metamodels and stochastic optimization algorithms (GA). The optimized profile was characterized by a shift of two hours in the activation times of the heating system. As can be observed, the approximation provided by the developed system was able to closely replicate the optimized profile. On the other hand, Fig. 18 (right) shows the hourly heating demand profile of building 6L for January 10th (E +), corresponding to the optimized profile, as well as the SSA and MSA predictions provided by the heating metamodel. It can be seen that the SSA predictions accurately reproduced the demand profile throughout the day, including the load peaks occurring during the two heating system startups (at 7:00 and 15:00 h). Despite a certain degradation in performance, the MSA predictions also reasonably captured most of the profile, but from 17:00 hours onwards, significantly underestimated the heating demand value.

The combined effect of energy savings associated with smoothing heating peaks, as well as shifting a portion of energy consumption to times with lower heat purchase prices, resulted in daily cost savings ranging from approximately 5 to 10 % in relation to the cost associated to the baseline settings, depending on the validation day. The assessment of cost reduction related to heating energy throughout the entire heating season will undergo a thorough analysis in a subsequent, dedicated paper. This upcoming publication will delve into the evaluation of the impact at district scale of the integration of the IDM into district network MPC solutions.

Based on the results obtained, as a summary conclusion, it is worth noting that although further efforts are still necessary to improve the accuracy of the metamodels for specific regions within the solution space, the developed metamodels and optimization engine successfully enabled an effective determination of optimized control commands with a dramatic reduction regarding the required computational times compared to traditional physics- based model MPC solutions (from hours to minutes).

5. Conclusions

This paper introduced an enhanced version of the IDM concept, achieved through co-simulation. The approach involved coupling metamodels of buildings with a DH infrastructure Modelica model, resulting in an energy prediction engine for district MPC solutions with integrated supply/demand side modeling, improved generalization capacity, and reduced computational costs.

The developed metamodels provided energy behavior predictions for individual cases of residential and educational buildings, yielding satisfactory levels of accuracy for 24-hour prediction horizons. This achievement was accomplished by effectively leveraging the devised recursive predictive process. All evaluated ML algorithms, except for KNN, exhibited notable performance with R^2 values of up to 0.975 and 0.94 for residential and educational buildings respectively. Boosting methods displayed remarkable potential, generally outperforming other algorithms. Based on the acquired results, the suggested feature space, formulated by leveraging the concepts of building equivalent temperature and heating setpoint temperature, led to a reduction in computation times while exerting only a moderate influence on model accuracy.

The process was expanded to encompass further generalization, resulting in three sets of generic metamodels capable of replicating the behavior of any residential building instance with rectangular or L shape floor plans. Furthermore, an incremental procedure was defined to enable the utilization of building simulation models from previous projects for metamodel training, enhancing their accuracy potential. Based on the attained results, the number of buildings available for the development of metamodels (up to 9 after implementing the incremental procedure) proved suboptimal in achieving a sufficiently precise definition of the solution space, hindering the complete realization of metamodels' potential in terms of precision. Nevertheless, the developed metamodels managed to generate reasonably accurate 24 hour ahead heating demand predictions. Notably, all algorithms considered, except for KNNs, showcased more than acceptable performance levels. Boosting methods, in particular, exhibited substantial potential, generally surpassing the performance of other algorithms with R^2 values of up to 0.959. The inclusion of even a small number of additional models for training yielded notable enhancements in metamodels performance.

Finally, dedicated metamodels were created for residential building 6L, with the necessary generalization capability to predict the effects on the building's energy behavior resulting from optimal utilization of its own flexibility. The best performance was delivered by a combination of XGBoost heating demand metamodel ($R^2 = 0.978$) and a nested Adaboost ($R^2 = 0.995$) temperature metamodel. The implemented testbed enabled the seamless integration of the metamodels with optimization algorithms based on NSGA-II, and the successful identification of the optimum daily heating setpoint hourly profiles with a 24-hour prediction horizon, leading to daily cost savings ranging from 5 % to 10 %. While additional efforts are still required to enhance the accuracy of the metamodels for certain specific regions within the solution space, they

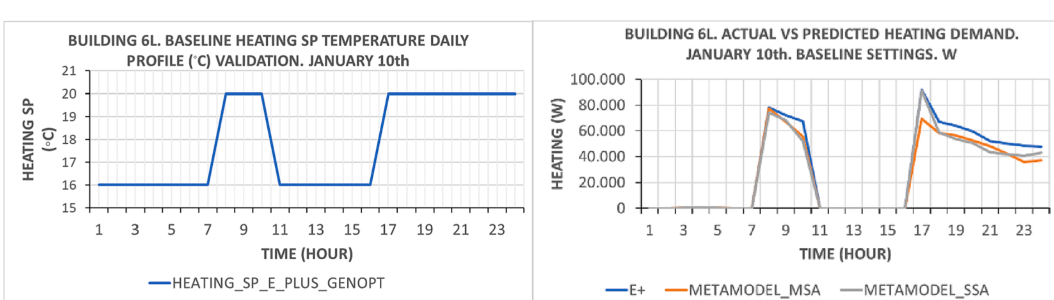


Fig. 17. Building 6L. Baseline heating SP temperature daily profile (left). Actual vs predicted heating demand (right). January 10th.

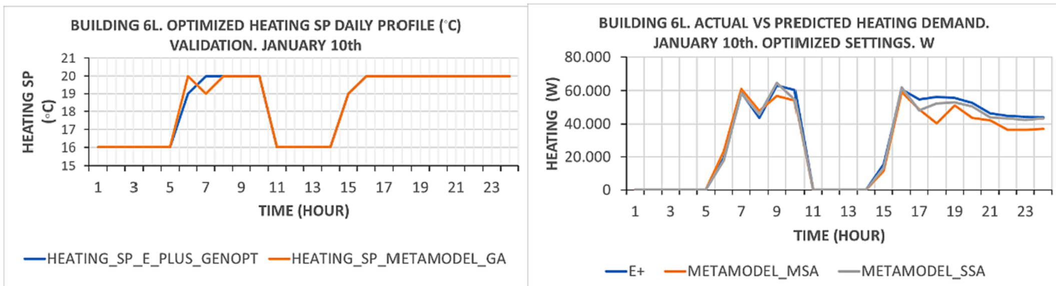


Fig. 18. Building 6L. Optimized heating SP temperature daily profile (left). Actual vs predicted heating demand for optimized settings (right). January 10th.

successfully facilitated the effective determination of optimized control commands.

To further enhance the described developments, the authors are working on refining the metamodels for the considered residential typologies. To accomplish this, an adequate number of building models will be utilized to facilitate an optimal delineation of the solution spaces essential for metamodel training. Additionally, as previously mentioned, the authors are actively involved in assessing at a district-scale the potential of the IDM as a predictive engine for MPC solutions in district systems.

CRediT authorship contribution statement

Víctor F. Sánchez-Zabala: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Tomás Gómez-Acebo:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

Authors would like to express their gratitude to the operator of the Vozdovac district heating system (Beogradske elektrane) for the specifications used to develop and calibrate the EnergyPlus models, and to Solintel M&P, SL for developing the initial versions of the EnergyPlus models (including only the geometrical and constructive definition of the buildings) used in the previous work of the authors [90].

References

[1] Gholamzadehmir M, Del Pero C, Buffa S, Fedrizzi R, Aste N. Adaptive-predictive control strategy for HVAC systems in smart buildings – A review. Sustain. Cities Soc. 2015;63:102480. https://doi.org/10.1016/j.scs.2020.102480.
[2] Afram A, Janabi-Sharifi F. Theory and applications of HVAC control systems – A review of model predictive control (MPC). Build. Environ. 2014;72:343–55. https://doi.org/10.1016/j.buildenv.2013.11.016.
[3] D. Mariano-Hernández, L. Hernández-Callejo, A. Zorita-Lamadrid, O. Duque-Pérez, F. Santos García. A review of strategies for building energy management system: Model predictive control, demand side management, optimization, and fault detect & diagnosis. Journal of Building Engineering. 2021;33:101692. 10.1016/j.jobe.2020.101692.
[4] Široký J, Oldewurtel F, Cigler J, Prívara S. Experimental analysis of model predictive control for an energy efficient building heating system. Appl. Energy 2011;88:3079–87. https://doi.org/10.1016/j.apenergy.2011.03.009.
[5] Oldewurtel F, Parisio A, Jones CN, Gyalistras D, Gwerder M, Stauch V, et al. Use of model predictive control and weather forecasts for energy efficient building climate control. Energ. Buildings 2012;45:15–27. https://doi.org/10.1016/j.enbuild.2011.09.022.
[6] Thieblemont H, Haghighat F, Ooka R, Moreau A. Predictive control strategies based on weather forecast in buildings with energy storage system: A review of the state-of-the art. Energ. Buildings 2017;153:485–500. https://doi.org/10.1016/j.enbuild.2017.08.010.
[7] Yao Ye, Shekhar DK. State of the art review on model predictive control (MPC) in Heating Ventilation and Air-conditioning (HVAC) field. Build. Environ. 2021; 200:107952. https://doi.org/10.1016/j.buildenv.2021.107952.
[8] Fabrizio Ascione, Nicola Bianco, Claudio De Stasio, Gerardo Maria Mauro, Giuseppe Peter Vanoli. Simulation-based model predictive control by the multi-objective optimization of building energy performance and thermal comfort. Energy and Buildings. 2016;37:131–144. doi: 10.1016/j.enbuild.2015.11.03+3.
[9] Fabrizio Ascione, Nicola Bianco, Claudio De Stasio, Gerardo Maria Mauro, Giuseppe Peter Vanoli. A new comprehensive approach for cost-optimal building design integrated with the multi-objective model predictive control of HVAC systems. Sustainable Cities and Society. 2017. 31: 136-150. doi: 10.1016/j.scs.2017.02.010.
[10] EnergyPlus building energy simulation program. https://energyplus.net/.
[11] Kontes GD, Giannakis GI, Sánchez V, Agustin-camacho PD, Romero-amorrortu A, Panagiotidou N. Simulation-based evaluation and optimization of control strategies in buildings. Energies 2018;11:1–23. https://doi.org/10.3390/en11123376.
[12] Maomao Hu, Fu Xiao, John Bagterp Jørgensen, Rongling Li. Price-responsive model predictive control of floor heating systems for demand response using building thermal mass . Applied Thermal Engineering. 2019;153:316-329. doi: 10.1016/j.applthermaleng.2019.02.107.
[13] TRNSYS, Transient System Simulation Tool. http://www.trnsys.com/index.html.
[14] MATLAB. https://www.mathworks.com/products/matlab.html.
[15] Yang S. Man Pun Wan, Wanyu Chen, Bing Feng Ng, Swapnil Dubey Model predictive control with adaptive machine-learning-based model for building energy efficiency and comfort optimization. Appl. Energy 2020;271:115147. https://doi.org/10.1016/j.apenergy.2020.115147.
[16] Shiyu Yang, Man Pun Wan, Wanyu Chen, Bing Feng Ng, Swapnil Dubey. Experiment study of machine-learning-based approximate model predictive control for energy-efficient building control. Applied Energy 2021;288: 116648. doi: 10.1016/j.apenergy.2021.116648.
[17] Le K, Bourdais R, Guéguen H. From hybrid model predictive control to logical control for shading system: A support vector machine approach. Energ. Buildings 2014;84:352–9. https://doi.org/10.1016/j.enbuild.2014.07.084.
[18] Kusiak A, Li M, Tang F. Modeling and optimization of HVAC energy consumption. Appl. Energy 2010;87:3092–102. https://doi.org/10.1016/j.apenergy.2010.04.008.
[19] Papantoniou S, Kolokotsa D, Kalaitzakis K. Building optimization and control algorithms implemented in existing BEMS using a web based energy management and control system. Energ. Buildings 2015;98:45–55. https://doi.org/10.1016/j.enbuild.2014.10.083.
[20] Afram A, Janabi-Sharifi F, Fung AS, Raahemifar K. Artificial neural network (ANN) based model predictive control (MPC) and optimization of HVAC systems: A state of the art review and case study of a residential HVAC system. Energ. Buildings 2017;141:96–113. https://doi.org/10.1016/j.enbuild.2017.02.012.
[21] Jain RK, Smith KM, Culligan PJ, Taylor JE. Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy. Appl. Energy 2014;123:168–78. https://doi.org/10.1016/j.apenergy.2014.02.057.
[22] Henrik Lund, Sven Werner, Robin Wiltshire, Svend Svendsen, Jan Eric Thorsen, Frede Hvelplund Brian Vad Mathiesen. 4th Generation District Heating (4GDH) Integrating smart thermal grids into future sustainable energy systems. Energy. 2014;68:1–11. doi: 10.1016/j.energy.2014.02.089.
[23] Cheng Fan Fu, Xiao YZ. A short-term building cooling load prediction method using deep learning algorithms. Appl. Energy 2017;195:222–33. https://doi.org/10.1016/j.apenergy.2017.03.064.
[24] Fan Zhang, Chirag Deb, Siew Eang Lee, Junjing Yang, Kwok Wei Shah. Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique. Energy and Buildings. 2016. 126: 94-103. doi: 10.1016/j.enbuild.2016.05.028.

[25] Lei L, Chen W, Bing Wu, Chen C, Liu W. A building energy consumption prediction model based on rough set theory and deep learning algorithms. Energ. Buildings 2021;240:110886. https://doi.org/10.1016/j.enbuild.2021.110886.
[26] Potočnik P, Škerl P, Govekar E. Machine-learning-based multi-step heat demand forecasting in a district heating system. Energ. Buildings 2021;233:110673. https://doi.org/10.1016/j.enbuild.2020.110673.
[27] Laakkonen L, Korpela T, Kaivosoja J, Vilkko M, Majanne Y, Nurmoranta M. Predictive Supply Temperature Optimization of District Heating Networks Using Delay Distributions. Energy Procedia 2017;116:297–309. https://doi.org/10.1016/j.egypro.2017.05.076.
[28] Chen S, Ren Y, Friedrich D, Yu Z, Yu J. Sensitivity analysis to reduce duplicated features in ANN training for district heat demand prediction. Energy and AI 2020;2:100028. https://doi.org/10.1016/j.egyai.2020.100028.
[29] Weili Xu, Chong A, Karaguzel OT. Khee Poh Lam. Improving evolutionary algorithm performance for integer type multi-objective building system design optimization. Energ. Buildings 2016;127:714–29. https://doi.org/10.1016/j.enbuild.2016.06.043.
[30] Ascione F, Bianco N, De Stasio C, Mauro GM, Vanoli GP. CASA, Cost-optimal analysis by multi-objective optimisation and artificial neural networks: a new framework for the robust assessment of cost-optimal energy retrofit, feasible for any building. Energ. Buildings 2017;146:200–19. https://doi.org/10.1016/j.enbuild.2017.04.069.
[31] Westermann P, Welzel M, Evins R. Using a deep temporal convolutional network as a building energy surrogate model that spans multiple climate zones. Appl. Energy 2020;278:115563. https://doi.org/10.1016/j.apenergy.2020.115563.
[32] Papadopoulos S, Azar E. Integrating building performance simulation in agent-based modeling using regression surrogate models: A novel human-in-the-loop energy modeling approach. Energ. Buildings 2016;128:214–23. https://doi.org/10.1016/j.enbuild.2016.06.079.
[33] Kheiri F. A review on optimization methods applied in energy-efficient building geometry and envelope design. Renew. Sustain. Energy Rev. 2018;92:897–920. https://doi.org/10.1016/j.rser.2018.04.080.
[34] Huang H, Chen L, Eric Hu. A neural network-based multi-zone modelling approach for predictive control system design in commercial buildings. Energ. Buildings 2015;97:86–97. https://doi.org/10.1016/j.enbuild.2015.03.045.
[35] Huang H, Chen L, Eric Hu. A new model predictive control scheme for energy and cost savings in commercial buildings: An airport terminal building case study. Build. Environ. 2015;89:203–16. https://doi.org/10.1016/j.buildenv.2015.01.037.
[36] Gasser J, Cai H, Karagiannopoulos S, Heer P, Hug G. Predictive energy management of residential buildings while self-reporting flexibility envelope. Appl. Energy 2021;288:116653. https://doi.org/10.1016/j.apenergy.2021.116653.
[37] Nguyen A-T, Reiter S, Rigo P. A review on simulation-based optimization methods applied to building performance analysis. Appl. Energy 2014;113:1043–58. https://doi.org/10.1016/j.apenergy.2013.08.061.
[38] Westermann P, Evins R. Surrogate modelling for sustainable building design –A review. Energ. Buildings 2019;198:170–86. https://doi.org/10.1016/j.enbuild.2019.05.057.
[39] Roman ND, Bre F, Fachinotti VD, Lamberts R. Application and characterization of metamodels based on artificial neural networks for building performance simulation: A systematic review. Energ. Buildings 2020;217:109972. https://doi.org/10.1016/j.enbuild.2020.109972.
[40] Luo XJ, Oyedele LO, Ajayi AO, Akinade OO. Comparative study of machine learning-based multi-objective prediction framework for multiple building energy loads. Sustain. Cities Soc. 2020;61:102283. https://doi.org/10.1016/j.scs.2020.102283.
[41] Yuce B, Li H, Rezgui Y, Petri I, Jayan B, Yang C. Utilizing artificial neural network to predict energy consumption and thermal comfort level: An indoor swimming pool case study. Energ. Buildings 2014;80:45–56. https://doi.org/10.1016/j.enbuild.2014.04.052.
[42] Paudel S, Elmitri M, Couturier S, Nguyen PH, Kamphuis R, Lacarrière B, et al. A relevant data selection method for energy consumption prediction of low energy building based on support vector machine. Energ. Buildings 2017;138:240–56. https://doi.org/10.1016/j.enbuild.2016.11.009.
[43] Soheil Fathi, Ravi Srinivasan, Andriel Fenner, Sahand Fathi. Machine learning applications in urban building energy performance forecasting: A systematic review. Renewable and Sustainable Energy Reviews. 2020. 133: 1364-0321. doi: 10.1016/j.rser.2020.110287.
[44] Chegari B, Tabaa M, Simeu E, Moutaouakkil F, Medromi H. Multi-objective optimization of building energy performance and indoor thermal comfort by combining artificial neural networks and metaheuristic algorithms. Energ. Buildings 2021;239:110839. https://doi.org/10.1016/j.enbuild.2021.110839.
[45] Razmi A, Rahbar M, Bemanian M. PCA-ANN integrated NSGA-III framework for dormitory building design optimization: Energy efficiency, daylight, and thermal comfort. Appl. Energy 2021;305:117828. https://doi.org/10.1016/j.apenergy.2021.117828.
[46] Li X, Yao R. Modelling heating and cooling energy demand for building stock using a hybrid approach. Energ. Buildings 2021;235:110740. https://doi.org/10.1016/j.enbuild.2021.110740.
[47] Buffa S, Soppelsa A, Pipiciello M, Henze G, Fedrizzi R. Fifth-Generation District Heating and Cooling Substations: Demand Response with Artificial Neural Network-Based Model Predictive Control. Energies 2020;13, no. 17:4339. https://doi.org/10.3390/en13174339.
[48] Garnier A, Eynard J, Caussanel M, Grieu S. Predictive control of multizone heating, ventilation and air-conditioning systems in non-residential buildings. Appl. Soft Comput. 2015;37:847–62. https://doi.org/10.1016/j.asoc.2015.09.022.
[49] Lim H, Zhai ZJ. Comprehensive evaluation of the influence of meta-models on Bayesian calibration. Energ. Buildings 2017;155:66–75. https://doi.org/10.1016/j.enbuild.2017.09.009.
[50] Manfren M, Aste N, Moshksar R. Calibration and uncertainty analysis for computer models – A meta-model based approach for integrated building energy simulation. Appl. Energy 2017;103:627–41. https://doi.org/10.1016/j.apenergy.2012.10.031.
[51] Edwards RE, New J, Parker LE, Cui B, Dong J. Constructing large scale surrogate models from big data and artificial intelligence. Appl. Energy 2017;202:685–99. https://doi.org/10.1016/j.apenergy.2017.05.155.
[52] Carreras J, Pozo C, Boer D, Guillén-Gosálbez G, Caballero JA, Ruiz-Femenia R, et al. Systematic approach for the life cycle multi-objective optimization of buildings combining objective reduction and surrogate modelling. Energ. Buildings 2016;130:506–18. https://doi.org/10.1016/j.enbuild.2016.07.062.
[53] Gossard D, Lartigue B, Thellier F. Multi-objective optimization of a building envelope for thermal performance using genetic algorithms and artificial neural network. Energ. Buildings 2013;67:253–60. https://doi.org/10.1016/j.enbuild.2013.08.026.
[54] Chen Xi, Yang H, Sun Ke. Developing a meta-model for sensitivity analyses and prediction of building performance for passively designed high-rise residential buildings. Appl. Energy 2017;194:422–39. https://doi.org/10.1016/j.apenergy.2016.08.180.
[55] Romani Z, Draoui A, Allard F. Metamodeling the heating and cooling energy needs and simultaneous building envelope optimization for low energy building design in Morocco. Energ. Buildings 2015;102:139–48. https://doi.org/10.1016/j.enbuild.2015.04.014.
[56] Seyed Amirhosain Sharif, Amin Hammad. Developing surrogate ANN for selecting near-optimal building energy renovation methods considering energy consumption, LCC and LCA. Journal of Building Engineering. 2019;25:100790. doi: 10.1016/j.jobe.2019.100790.
[57] Eisenhower B, O'Neill Z, Narayanan S, Fonoberov VA, Mezić I. A methodology for meta-model based optimization in building energy models. Energ. Buildings 2012;47:292–301. https://doi.org/10.1016/j.enbuild.2011.12.001.
[58] Prada A, Gasparella A, Baggio P. On the performance of meta-models in building design optimization. Appl. Energy 2018;225:814–26. https://doi.org/10.1016/j.apenergy.2018.04.129.
[59] Ehsan Asadi, Manuel Gameiro da Silva, Carlos Henggeler Antunes, Luís Dias, Leon Glicksman. Multi-objective optimization for building retrofit: A model using genetic algorithm and artificial neural network and an application. Energy and Buildings. 2014;81: 444-456. doi: 10.1016/j.enbuild.2014.06.009.
[60] Bre F, Roman N, Fachinotti VD. An efficient metamodel-based method to carry out multi-objective building performance optimizations. Energ. Buildings 2020;206:109576. https://doi.org/10.1016/j.enbuild.2019.109576.
[61] Tian W, Choudhary R, Augenbroe G, Lee SH. Importance analysis and meta-model construction with correlated variables in evaluation of thermal performance of campus buildings. Build. Environ. 2015;92:61–74. https://doi.org/10.1016/j.buildenv.2015.04.021.
[62] Thrampoulidis E, Mavromatidis G, Lucchi A, Orehounig K. A machine learning-based surrogate model to approximate optimal building retrofit solutions. Appl. Energy 2021;281:116024. https://doi.org/10.1016/j.apenergy.2020.116024.
[63] Brownlee AEI, Wright JA. Constrained, mixed-integer and multi-objective optimisation of building designs by NSGA-II with fitness approximation. Appl. Soft Comput. 2015;33:114–26. https://doi.org/10.1016/j.asoc.2015.04.010.
[64] Wei Yu, Li B, Jia H, Zhang M, Wang Di. Application of multi-objective genetic algorithm to optimize energy efficiency and thermal comfort in building design. Energ. Buildings 2015;88:135–43. https://doi.org/10.1016/j.enbuild.2014.11.063.
[65] Singaravel S, Suykens J, Geyer P. Deep-learning neural-network architectures and methods: Using component-based models in building-design energy prediction. Adv. Eng. Inf. 2018;38:81–90. https://doi.org/10.1016/j.aei.2018.06.004.
[66] Chou J-S, Bui D-K. Modeling heating and cooling loads by artificial intelligence for energy-efficient building design. Energ. Buildings 2014;82:437–46. https://doi.org/10.1016/j.enbuild.2014.07.036.
[67] Lopes MN, Lamberts R. Development of a Metamodel to Predict Cooling Energy Consumption of HVAC Systems in Office Buildings in Different Climates. Sustainability 2018;10:4718. https://doi.org/10.3390/su10124718.
[68] Melo AP, Versage RS, Sawaya G, Lamberts R. A novel surrogate model to support building energy labelling system: A new approach to assess cooling energy demand in commercial buildings. Energ. Buildings 2016;131:233–47. https://doi.org/10.1016/j.enbuild.2016.09.033.
[69] Tardioli G, Narayan A, Kerrigan R, Oates M, O'Donnell J, Finn DP. A methodology for calibration of building energy models at district scale using clustering and surrogate techniques. Energy and Buildings, Volume 226 2020; 226:110309. https://doi.org/10.1016/j.enbuild.2020.110309.
[70] Narjes Abbasabadi, Mehdi Ashayeri, Rahman Azari, Brent Stephens, Mohammad Heidarinejad . An integrated data-driven framework for urban energy use modelling (UEUM). Applied Energy. 253;2019: 113550. doi: 10.1016/j.apenergy.2019.113550.
[71] Melo AP, Fossati M, Versage RS, Sorgato MJ, Scalco VA, Lamberts R. Development and analysis of a metamodel to represent the thermal behavior of naturally ventilated and artificially air conditioned residential buildings. Energ. Buildings 2016;112:209–21. https://doi.org/10.1016/j.enbuild.2015.12.021.
[72] Nutkiewicz A, Yang Z, Jain RK. Data-driven Urban Energy Simulation (DUE-S): A framework for integrating engineering simulation and machine learning methods

in a multi-scale urban energy modeling workflow. Appl. Energy 2018;225: 1176–89. https://doi.org/10.1016/j.apenergy.2018.05.023.
[73] Robinson C, Dilkina B, Hubbs J, Zhang W, Guhathakurta S, Brown MA, et al. Machine learning approaches for estimating commercial building energy consumption. Appl. Energy 2017;208:889–904. https://doi.org/10.1016/j.apenergy.2017.09.060.
[74] Kontokosta CE, Tull C. A data-driven predictive model of city-scale energy use in buildings. Appl. Energy 2017;197:303–17. https://doi.org/10.1016/j.apenergy.2017.04.005.
[75] Mastrucci A, Baume O, Stazi F, Leopold U. Estimating energy savings for the residential building stock of an entire city: A GIS-based statistical downscaling approach applied to Rotterdam. Energ. Buildings 2014;75:358–67. https://doi.org/10.1016/j.enbuild.2014.02.032.
[76] Dagdougui H, Bagheri F, Le H, Dessaint L. Neural network model for short-term and very-short-term load forecasting in district buildings. Energ. Buildings 2019; 203:109408. https://doi.org/10.1016/j.enbuild.2019.109408.
[77] Bellahsen A, Dagdougui H. Aggregated short-term load forecasting for heterogeneous buildings using machine learning with peak estimation. Energ. Buildings 2021;237:110742. https://doi.org/10.1016/j.enbuild.2021.110742.
[78] Walker S, Khan W, Katic K, Maassen W, Zeiler W. Accuracy of different machine learning algorithms and added-value of predicting aggregated-level energy performance of commercial buildings. Energ. Buildings 2020;209:109705. https://doi.org/10.1016/j.enbuild.2019.109705.
[79] Wang Z, Wang Y, Srinivasan RS. A novel ensemble learning approach to support building energy use prediction. Energ. Buildings 2018;159:109–22. https://doi.org/10.1016/j.enbuild.2017.10.085.
[80] Yong S-G, Kim J, Cho J, Koo J. Meta-models for building energy loads at an arbitrary location. Journal of Building Engineering 2019;25:100823. https://doi.org/10.1016/j.jobe.2019.100823.
[81] Østergård T, Jensen RL, Maagaard SE. A comparison of six metamodeling techniques applied to building performance simulations. Appl. Energy 2020;211: 89–103. https://doi.org/10.1016/j.apenergy.2017.10.102.
[82] Wei Tian R, Choudhary.. A probabilistic energy model for non-domestic building sectors applied to analysis of school buildings in greater London. Energ. Buildings 2012;54:1–11. https://doi.org/10.1016/j.enbuild.2012.06.031.
[83] Usman Ali, Mohammad Haris Shamsi, Mark Bohacek, Cathal Hoare, Karl Purcell, Eleni Mangina, James O'Donnell. A data-driven approach to optimize urban scale energy retrofit decisions for residential buildings. Applied Energy. 2020. 267: 114861. doi: 10.1016/j.apenergy.2020.114861.
[84] A. Mastrucci, O. Baume, F. Stazi, U. Leopold, Estimating energy savings for the residential building stock of an entire city: A GIS-based statistical downscaling approach applied to Rotterdam, Energy and Buildings 75 (2014) 358–367, URL: http://www.sciencedirect.com/science/article/pii/S0378778814001467.
[85] Magalhães SMC, Leal VMS, Horta IM. Modelling the relationship between heating energy use and indoor temperatures in residential buildings through Artificial Neural Networks considering occupant behavior. Energ. Buildings 2017;151: 332–43. https://doi.org/10.1016/j.enbuild.2017.06.076.
[86] Ngo N-T. Early predicting cooling loads for energy-efficient design in office buildings by machine learning. Energ. Buildings 2019;182:264–73. https://doi.org/10.1016/j.enbuild.2018.10.004.
[87] Jovanović RŽ, Sretenović AA, Živković BD. Ensemble of various neural networks for prediction of heating energy consumption. Energ. Buildings 2015;94:189–99. https://doi.org/10.1016/j.enbuild.2015.02.052.
[88] Burger EM, Moura SJ. Gated ensemble learning method for demand-side electricity load forecasting. Energ. Buildings 2015;109:23–34. https://doi.org/10.1016/j.enbuild.2015.10.019.
[89] Wang L, Lee EWM, Yuen RKK. Novel dynamic forecasting model for building cooling loads combining an artificial neural network and an ensemble approach. Appl. Energy 2018;228:1740–53. https://doi.org/10.1016/j.apenergy.2018.07.085.
[90] Sánchez VF, Marijuan AG. Integrated model concept for district energy management optimisation platforms. Appl. Therm. Eng. 2021;196:117233. https://www.sciencedirect.com/science/article/pii/S1359431121006712.
[91] The Modelica Association. https://www.modelica.org/.
[92] Zeyu Wang, Ravi S. Srinivasan, J. Yu. A Review of Artificial Intelligence Based Building Energy Prediction with a Focus on Ensemble Prediction Models. In: Proceedings of the 2015 Winter Simulation Conference, California, USA. December, 2015. p.3438–3448.
[93] Zhao H, Magoulès F. A Review on the Prediction of Building Energy Consumption Renewable and Sustainable Energy Reviews 2012;2012(16):3586–92. https://doi.org/10.1016/j.rser.2012.02.049.
[94] Aurélie Foucquier, Sylvain Robert, Frédéric Suard, Louis Stephan, Arnaud Jay. State of the art in building modelling and energy performances prediction: A review. Renewable and Sustainable Energy Reviews, Elsevier, 2013, 23, pp.272 - 288. 2017;116:297–309. https://hal-cea.archives-ouvertes.fr/cea-01792021.
[95] Wang Z, Srinivasan RS. A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models. Renew. Sustain. Energy Rev. 2017;75:796–808. https://doi.org/10.1016/j.rser.2016.10.079.
[96] Seyedzadeh S, Rahimian FP, Glesk I, Roper M. Machine learning for estimation of building energy consumption and performance: a review. Visualization. Engineering 2018;6:5.. https://doi.org/10.1186/s40327-018-0064-7.
[97] Alobaidi MH, Chebana F, Meguid MA. Robust ensemble learning framework for day-ahead forecasting of household based energy consumption. Appl. Energy 2018;212:997–1012. https://doi.org/10.1016/j.apenergy.2017.12.054.
[98] Wang Z. Tianzhen Hong, Mary Ann Piette. Building thermal load prediction through shallow machine learning and deep learning. Appl. Energy 2020;263: 114683. https://doi.org/10.1016/j.apenergy.2020.114683.
[99] Yan H, Yan Ke, Ji G. Optimization and Prediction in the Early Design Stage of Office Buildings Using Genetic and XGBoost Algorithms Building and Environment 2022;218:109081. https://doi.org/10.1016/j.buildenv.2022.109081.
[100] Sadik Yigit. A machine-learning-based method for thermal design optimization of residential buildings in highly urbanized areas of Turkey. Journal of Building Engineering. 189;2019: 116324. doi: 10.1016/j.jobe.2021.102225.
[101] Freund Y, Schapire RE. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. J. Comput. Syst. Sci. 1997;55(1):119–39. https://doi.org/10.1006/jcss.1997.1504.
[102] Drucker H. Improving regressors using boosting techniques. En *Icml* 1997: 107–15.
[103] Friedman JH. Stochastic gradient boosting. Comput. Stat. Data Anal. 2002;38: 367–78. https://jerryfriedman.su.domains/ftp/stobst.pdf.
[104] CHEN, Tianqi; GUESTRIN, Carlos. Xgboost: A scalable tree boosting system. En Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. 2016. p. 785-794. doi: 10.48550/arXiv.1603.02754.
[105] The Functional Mock-up Interface (FMI) standard. https://fmi-standard.org/.
[106] SimulatorToFMU. https://github.com/LBNL-ETA/SimulatorToFMU.
[107] Kusiak A, Guanglin Xu, Zhang Z. Minimization of energy consumption in HVAC systems with data-driven models and an interior-point method. Energ. Conver. Manage. 2014;85:146–53. https://doi.org/10.1016/j.enconman.2014.05.053.
[108] Hamdy M, Nguyen A-T, Hensen JLM. A performance comparison of multi-objective optimization algorithms for solving nearly-zero-energy-building design problems. Energ. Buildings 2016;121:57–71. https://doi.org/10.1016/j.enbuild.2016.03.035.
[109] Magnier L, Haghighat F. Multiobjective optimization of building design using TRNSYS simulations, genetic algorithm, and Artificial Neural Network. Build. Environ. 2010;22:739–46. https://doi.org/10.1016/j.buildenv.2009.08.016.
[110] Rosso F, Ciancio V, Dell'Olmo J, Salata F. Multi-objective optimization of building retrofit in the Mediterranean climate by means of genetic algorithm application. Energ. Buildings 2020;216:109945. https://doi.org/10.1016/j.enbuild.2020.109945.
[111] Carlucci S, Cattarin G, Causone F, Pagliano L. Multi-objective optimization of a nearly zero-energy building based on thermal and visual discomfort minimization using a non-dominated sorting genetic algorithm (NSGA-II). Energ. Buildings 2015;104:378–94. https://doi.org/10.1016/j.enbuild.2015.06.064.
[112] Perera ATD, Wickramasinghe PU, Nik VM, Scartezzini J-L. Machine learning methods to assist energy system optimization. Appl. Energy 2019;243:191–205. https://doi.org/10.1016/j.apenergy.2019.03.202.
[113] Bre F, Fachinotti VD. A computational multi-objective optimization method to improve energy efficiency and thermal comfort in dwellings. Energ. Buildings 2017;154:283–94. https://doi.org/10.1016/j.enbuild.2017.08.002.
[114] Ascione F, Masi RFD, de Rossi F, Ruggiero S, Vanoli GP. Optimization of building envelope design for nZEBs in Mediterranean climate: Performance analysis of residential case study. Appl. Energy 2016;183:938–57. https://doi.org/10.1016/j.apenergy.2016.09.027.
[115] He X, Zhang Z, Kusiak A. Performance optimization of HVAC systems with computational intelligence algorithms. Energ. Buildings 2014;81:371–80. https://doi.org/10.1016/j.enbuild.2014.06.021.
[116] Yang T, Pan Y, Mao J, Wang Y, Huang Z. An automated optimization method for calibrating building energy simulation models with measured data: Orientation and a case study. Appl. Energy 2016;179:1220–31. https://doi.org/10.1016/j.apenergy.2016.07.084.
[117] Katoch S, Chauhan SS, Kumar V. A review on genetic algorithm: past, present, and future. Multimed. Tools Appl. 2021;80:8091–126. https://doi.org/10.1007/s11042-020-10139-6.
[118] McCall J. Genetic algorithms for modelling and optimisation. J. Comput. Appl. Math. 2005;184(1):205–22. https://doi.org/10.1016/j.cam.2004.07.034.
[119] OZCAN, Ender; MOHAN, Chilukuri K. Analysis of a simple particle swarm optimization system. Intelligent engineering systems through artificial neural networks, 1998, vol. 8, p. 253-258.
[120] Sánchez V, et.. D3.6 Local and Global Energy performance models. MOEEBIUS Project 31/03/2017.. https://www.moeebius.eu/images/MOEEBIUS_D3.6_Local_and_Global_Energy_performance_models_V1.0.pdf.
[121] Scikit-Learn. https://scikit-learn.org/.
[122] Keras. https://keras.io/.
[123] Tensorflow. https://www.tensorflow.org.
[124] XGBoost. https://xgboost.readthedocs.io.
[125] Investigation of performance metrics in regression analysis and machine learning-based prediction models. The 8th European Congress on Computational Methods in Applied Sciences and Engineering ECCOMAS Congress 2022 5 – 9 June 2022, Oslo, Norway. DOI: 10.23967/eccomas.2022.155.
[126] Blank J, Deb K. pymoo: multi-objective optimization in python. *IEEE*. Access 2020;8:89497–509. https://doi.org/10.1109/ACCESS.2020.2990567.
[127] GenOpt Generic Optimization Program. https://simulationresearch.lbl.gov/GO/index.html.