# Predictive chiller operation: A data-driven loading and scheduling approach

Authors: Enric Sala-Cardoso<sup>a,\*</sup>, Miguel Delgado-Prieto<sup>a</sup>, Konstantinos Kampouropoulos<sup>b</sup>, Luis Romeral<sup>a</sup>

Emails: enric.sala@upc.edu\*, miguel.delgado@upc.edu, luis.romeral@upc.edu

\* Corresponding author.

<sup>a</sup> Department of Electronic Engineering, Universitat Politècnica de Catalunya, MCIA Research Center, Terrassa, Spain.

<sup>b</sup> Fundació CTM Centre Tecnològic, Manresa, Spain.

Abstract— The proper sequencing and optimal loading of chillers is one of the major avenues for energy efficiency improvement in existing heating, ventilating and air conditioning installations. The main enabler for the success of such applications is the access to accurate chiller performance maps that allow to operate the equipment in optimal conditions. However, current solutions are excessively reliant on maps obtained through suboptimal means, such as manufacturer datasheets, extensive instrumentation campaigns or burdensome modelling and simulation methodologies. Furthermore, recent studies show that strategies based on model-predictive control may lead to increased savings by anticipating the future cooling de- mand and scheduling the operation of the chillers, selecting the optimal operation configuration and ex- tending the remaining life by reducing switching. In this regard, this study presents a novel data-driven and multi-criteria chiller orchestration strategy that combines a chiller performance characterization stage for obtaining performance maps based on a neural network-based learning methodology and a state-of- theart hybrid load forecasting scheme for calculating the future load profiles. The effectiveness of the proposed methodology is tested with experimental data from a multi-chiller installation in a tertiary sec- tor building, where nearly a 20% average performance increase is achieved compared to the standard realtime controller of the HVAC installation.

Keywords—chiller scheduling, deep learning, demand-side management, model-predictive control, optimal chiller loading.

## 1. Introduction

Heating, Ventilating and Air Conditioning (HVAC) systems are one of the main energy consumers in buildings, accounting for up to 40% of their total energy consumption [1] . Naturally, HVAC systems have become a target of study that currently represents a great deal of scientific effort f or the reduction of energy consumption and preservation of resources [2]. Two main avenues are possible for tackling this problem, either by implementing strategies for reducing power demand, or by improving the efficiency of the involved energy systems. On the avenue of improving energy efficiency, production equipment such as chillers are one of the areas that present the greatest potential for improvement [3] since these are one of the largest consumers in buildings, especially in the residential and tertiary sector where they can account for up to 40% of the building's energy consumption [4] . This is especially true in multi-chiller plants commonly found in medium to large buildings, where proper chiller loading and coordination is critical for increasing efficiency, and achieving an optimal control solution is of great interest with many studies attempting to tackle this issue with a variety of methods [5]. The potential for improvement is due to the fact the Coefficient of Performance (COP) of production equipment is not uniform throughout their operation range, meaning that differences in the control strategy of multi-chiller systems can lead to significant changes in the energy consumption result [6,7]. Indeed, a recent study concluded that up to 70% of the annual power consumption occurred while the chillers were operating at low Partial Load Ratio (PLR) which is the range were the coefficient of performance of chillers is typically at its lowest [8] , while a study evaluating ideal operation levels concluded there was a potential for improvement of up to 23.4% [9] . Thus, a strategy able to optimally control the chillers' operation has the potential to increase the operational efficiency of the overall system, minimizing the energy consumption by ensuring that each machine operates at its optimal COP while considering the affecting operating state, like weather conditions, inlet temperature from the distribution bus and future load demand.

#### A. Literature review

The chiller sequencing and optimal loading problems refer to the necessity to find a strategy for coordinating the operation of a group of chillers in order to meet the cooling demand while minimizing energy consumption, in terms of the loading ratio of each chiller [10] and the sequence at which they should be turned on or off according to the cooling load requirements [11] . A thorough review of the related state-of-the-art reveals that this is a complex problem having several facets that need to be carefully considered in order to achieve a performant solution. The main key aspects identified are: (i) the need of chiller performance characterization; (ii) the objective function of the controller and (iii) the choice of optimization strategy and implementation. Regarding the chiller performance characterization, chiller sequencing and optimal loading methodologies rely on the capability of mapping control actions to expected performance in or- der to select the chiller's settings that lead to their operation at optimal COP, given a set of operating conditions [12]. The most common approach in recent methodologies is the utilization of manufacturerprovided datasheets that specify the performance ac- cording to specific operating conditions or the simulation and modeling of the equipment using software tools like TRNSYS [13] . For example, performance data obtained from datasheet lookup tables were employed for implementing a predictive HVAC controller based on a mixed-integer approach, which has the down- side of providing only static figures and not considering all affecting parameters [14]. Instead, a study focused on the determination of a probabilistic approach for chiller replacement discussed the utilization of simulation compared to other methods, concluding that these could be viable for energy estimation if calibrated using measured data [15]. Even though the employment of such solutions in control applications is useful for research purposes, in practice, it is essential to have performance maps that accurately reflect the actual behavior of the equipment considering the multiple affecting factors besides the PLR, such as weather conditions, operating state and aging. In contrast, a methodology for optimizing the operation of a chiller plant employed a datadriven approach based on the modeling of the chiller group using a neural network and the implementation of a two-level algorithm, which allowed to achieve energy savings of 14% under simulation [16]. Another key aspect of the problem is the selection of a proper objective function for the optimization process. The de facto standard objective function is based on the aggregate power consumption of the group of chillers, being a useful strategy when the minimization of the used energy is the ultimate goal of the solution [13]. However, the need to tackle other concerns indicate that a multi-objective strategy may be preferred. In particular, some studies make reference to the switching problem, which can be described as the frequency and magnitude of changes to the relevant control settings. An application of real-time HVAC optimization looked into minimizing the disturbances caused by controller actions, comparing rate-limited setpoint reset to controlled step- changes [17]. The conclusions reflect that frequent changes to control variables may lead to system instability, especially when simultaneous and having large magnitude, thus showing that in practice limiting the amount of changes and moderating their delta is desirable. Furthermore, another argument for considering the minimization of switching is that increased control changes may lead to energy losses and faster equipment degradation due to the dynamics of the equipment and their mechanical wear [18]. Finally, when suitable performance maps are available and an appropriate objective function is determined, an optimization algorithm needs to be applied to carry out the control of the group of chillers in an efficient manner. Mainly two types of methodologies can be found on the literature, i) generic global optimization tools and specific heuristics-based controller implementations. Generic global optimization tools are common in chiller control applications, a study of air-cooled chillers optimal control used random forests to implement an empirical model of the chillers and then applied generic algorithms to carry out the estimation of the optimal values of the control parameters [19].

Similarly, particle swarm optimization was employed to adjust the control parameters of a water-cooled chiller plant, simulated using Modelica models adjusted using empirical data in [20] . However, controller implementations of this type are slow and have a randomness component, thus both their results and runtime are undeterministic by nature. Instead, other researchers have focused on the design of control algorithms that implement specific heuristics, making the controllers more computationally efficient and robust, which is a desirable property of the system even though the global optimum may not consistently be achieved by this means [21].

## B. Innovative contribution

In consideration of the described shortcoming in the state of the art solutions for the chiller loading and sequencing problem, this research proposes a novel methodology for the optimal control of chiller groups by combining data-driven performance map modeling of the chillers and a thermal demand forecasting method-ology in order to implement a multi-objective model-predictive controller. Specifically, the originality of this work consists on the following key aspects of the developed control strategy: i The integration of data-driven COP maps of the involved equipment, defining their performance depending on their operating conditions and considering several affecting factors (multi- variate approach); ii The integration of a short-term forecasting model of the building's future thermal demand, considering affecting factors such as the weather and the building's occupancy patterns, to pro- vide accurate and reliable demand requirements to the optimization stage; iii The consideration of the thermal dynamics of the building, taking advantage of the distribution system's thermal capacity, re- calculating and adjusting their operating setpoints for scheduling the operation over the time horizon; iv The definition of a strategy considering multiple criteria for the determination of the optimal control sequence of the chillers, including the minimization of energy usage and switching.

## C. Paper organization

This paper is organized as follows. Section 2 present an overview of the energy system models employed, specifically the modeling of the performance of HVAC equipment and the forecasting of cooling demand. Section 3 defines the optimization problem formulation and presents the proposed control strategy, focusing on the implementation of the control algorithm that optimizes the sequencing and loading of the chillers by taking advantage of the performance map and future load. Section 4 describes the facility that was used as test environment. Section 5 shows the experimental results that were obtained from the implementation of the proposed control strategy in the test environment. Finally, the conclusions of this work are presented and discussed in Section 6.

#### 2. ENERGY MODELLING

For the development and implementation of this study, two kinds of energy models are employed: i) performance model of the HVAC equipment, used for estimating the expected power characteristics at a given environment conditions and control actions, and ii) load forecasting model, used for predicting the future cooling needs of the building. This section describes the motivation behind the selection of these models based on the modeling requirements and explains their implementation.

# 2.1. Performance modeling of HVAC equipment

A wide range of energy management strategies is dependent on the availability of HVAC information related to the operational performance of production equipment for the implementation of applications such as maintenance and control solutions [22, 23]. An important reduction of the operating costs through the optimal use of energy resources can be accomplished by these functionalities [24], but their success relies on operating performance data being accessible [25]. Recent studies have adopted a variety of approaches with the objective of incorporating equipment operating performance maps in their solutions. Nevertheless, the process of implementing models that can be interpreted in order to estimate the power characteristics of the equipment when subjected to a control input and a set of state variables is still a challenging problem [26]. Solutions based manufacturer-provided data, though being common in control applications, face issues relating to lack of accessibility, a mismatch between the standardized testing conditions and the real operating environment [27], lack of comprehensive consideration of the affecting variables and insufficient granularity [28]. An- other set of solutions is based on simulation tools, which support the accurate modelling of equipment contingent on the knowledge of the machine's physical equations, proper tuning of their parameters to achieve sufficient reliability [29] and requiring the participation of domain experts for their usage [30]. Furthermore, these approaches do not reflect the current state of operation, which is bound to drift from the initial parameters due to faults and aging during the system's lifecycle [31]. Accordingly, datadriven approaches have received an attention increase due to their capacity to offer general solutions that can provide suitable results regardless of the type of equipment and complexity of the installation, supporting the extraction of nonlinear relationships between signals [16]. Neural networkbased methods are a common solution, as corroborated by a review of modeling methods for HVAC systems which shows several related applications [32]. However, coming up with reduced and generic sets of variables that allow the characterization of the performance in all cases is not feasible, which requires the methodology to be able to consider a large set of influencing variables [33]. Thus, deep learning methods, in particular, are especially well suited to solve this problem due to their capacity for feature learning [34]. 2.1.1.

Performance modeling implementation A deep neural network trained in an unsupervised manner by means of stacked sparse autoencoders is an efficient way to dis-cover features in the operating behavior of a HVAC production installation, allowing the transformation of the data to a latent space that can then be mapped to the power characteristics of the ma- chines by implementing further supervised learning. For the implementation of the performance map models, initially a dataset is extracted from the database that stores the historical operation of the building's HVAC equipment, including consumed and produced power, control commands, state variables and weather conditions. The extracted dataset is preprocessed in order to remove gaps caused by acquisition interruptions, outlier removal and erroneous readings. Then, an unsupervised feature learning stage is implemented using a deep autoencoder structure built by iteratively stacking single layer autoencoders. This structure is trained to reconstruct the input space by means of the following layer-wise pre-training strategy: making each layer reconstruct the output of the previous layer while constraining the flow of information using a sparse unit-activation constraint. After the pre-training of each layer, the reconstruction performance variation is evaluated by comparing it to the previous layer in order to assess whether additional layers should be created. When the stacked layers have been obtained, a fine-tuning process further improves the reconstruction error by fitting the reconstruction problem on the full autoencoder.

The feature learning stage is completed with the fine-tuning step, leading to the supervised training to solve the performance mapping problem. In order to map the learned features with the desired outputs, the decoder half of the deep autoencoder is detached and a different structure which will perform the mapping is connected in its place.

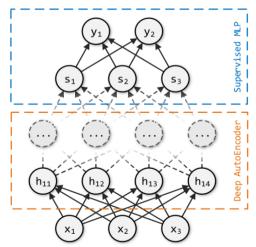


Fig. 1. Performance model architecture, composed of the encoder stage of an unsupervised deep autoencoder coupled to a supervised multiple-layer perceptron.

First, neural network layers are attached to the feature layer of the autoencoder in order to perform the mapping from the latent space to the target variables. Then, these additional layers are trained in a supervised manner, taking as in- puts the

autoencoder features in order to learn the relationship between the power consumption, production and coefficient of performance and the latent space. Finally, the performance map is obtained in the form of a deep neural network that is able to calculate the expected power consumption, production and coefficient of performance for a given set of operating conditions. The architecture of the performance map model is shown in Fig. 1, where an initial unsupervised feature learning is implemented as a deep autoencoder with layer-wise pretraining, and the supervised mapping is achieved using a multiple-layer perceptron with multiple outputs. The variables that compose the inputs of the model are the number of active compressors, chiller inlet temperature and chiller outlet temperature for each chiller, and the outdoor temperature and relative humidity. The outputs of the model are the instantaneous aggregate electrical power consumption, aggregate thermal power production and calculated COP, which can be estimated for a given set of state and control variables after the model has been trained.

#### 2.2 Load Forecasting in Buildings

Load forecasting takes an important role in the prevention of energy waste in buildings [35]. Thus, modeling and forecasting methodologies able to estimate the future consumption have be- come a meaningful concern of installation managers because of the useful knowledge that is obtainable from them [36] . Significant scientific effort is devoted to this topic, as evidenced by the different kinds of load forecasting methodologies employed in the literature [37]. Data-driven methods in particular are the most widespread solution for short-term forecasting applications due to its potentially complex dynamics, as indicated by the findings of a recent review [38]. However, although abundant methodologies exist, mainstream general-purpose forecasting solutions face limitations when applied to HVAC systems. These limitations are mainly tied to difficulties adapting the forecasts to the load demand variations caused by fluctuations in influencing variables like weather conditions or the behavior of the building's occupants throughout the day [39] . Indeed, accounting for occupant behavior is suggested to be a critical issue for achieving an increase in the degree of accuracy when performing load predictions in buildings [40]. Some authors indicate that the behavior of users should be a key concern in the research pertaining to the use of energy in buildings, due to the related potential for improvement regarding their energy efficiency [41]. This is supported by recent studies, which showcase the impact of occupant behavior, confirming there is a meaningful interrelationship between the regime changes of the operation of HVAC systems and the occupancy of the spaces in a building [42] Nonetheless, there is a lack of user behavior awareness in most of the modern load management tools, especially in simulation and forecasting methodologies. Additionally, those studies that deal with the consideration of occupancy data present crucial limitations and inadequate prediction accuracy [43]. To tackle this is- sue, proposals have surfaced for considering the occupancy data by means of occupant behavior attributes, stating that occupancy- based inputs should be considered due to their impact to energy consumption [44]. Another issue faced by optimization tools and controller implementers is that in general current methodologies are mainly aimed at solving the problem of energy consumption forecasting when used in HVAC systems [45]. Instead, the forecasting of the thermal power demand is better suited for supporting automation and control systems, which have the potential to benefit greatly

from this information, especially at the production stage for tackling supply planning concerns [46]. To summarize, the implementation of load demand models in the scope of buildings having reliable performance but requiring lower engineering effort is a complex concern being currently faced in installations [47]. 2.2.1. Cooling demand forecasting implementation In this study, a suitable thermal power load forecasting methodology is required for supporting the modelpredictive control, improving the controller response by anticipating power demand changes. The chosen approach is based on a data-driven methodology that is particularly well suited for this application due to the following key concepts: i) it allows the consideration of the thermal power demand of the building instead of the electrical power consumption of the chillers, allowing the decoupling of the machine performances and controller actions from the forecasted demand, and ii) it follows a hybrid approach that permits the incorporation of variables affecting the thermal power demand and their dynamics, specifically the occupant behavior of the building [48] . The main steps of the forecasting methodology are displayed in Fig. 2. As shown, the required signals are extracted from the building's historical database a), which are the signals corresponding to occupancy sensors b) and the power monitoring c). After loading the necessary signals, the modeling of the occupancy d) takes place, alongside the estimation of the thermal demand e) from the load profile of the distribution bus

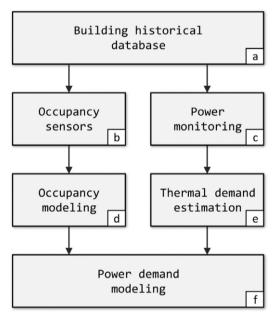


Fig. 2. Cooling power demand modeling methodology for short-term forecasting applications in buildings.

Finally, the modelling of the power demand f) is performed, combining the occupancy model and the estimated thermal demand. In order to improve the accuracy of the power demand forecasting, the methodology incorporates the level of activity in the building by implementing an artificial activity indicator. This indicator is built by aggregating presence detectors installed in each of the individual spaces and common areas in the building. Then, an activity model is implemented using recurrent neural networks to enhance the consideration of dynamic temporal patterns, while the power demand characterization is carried out by means of an adaptive neuro-fuzzy inference system structure. The activity indicator model provides a measure of the future occupancy level, which drives the HVAC power. The expected power demand is

calculated to obtain the final prediction, corresponding to this activity and the other influencing variables. In summary, the obtained partial models are combined to obtain the load forecasting model by the serialized composition of both inference systems, with the activity indicator forecast being fed to the power demand model to calculate the final prediction. Be- sides the activity indicator estimation procedure, the hybrid solution adopted in this study offers several advantages over traditional approaches. Namely, instead of fitting a single model using a general-purpose tool, a collaborative and modular structure is employed based on specialized models built for the activity and for the power demand. Such solution allows to fit and tune each method independently, adapting it to the dynamics of each signal and allowing to separately train the models with the use of different datasets.

#### 3. OPTIMIZATION AND CONTROL

This section describes optimization and control framework. First, the mathematical problem formulation is introduced, and then the implemented solution is described in detail

## 3.1. Optimization problem formulation

The mathematical formulation of the optimization problem can be considered as a multi-period nonlinear problem and it can be described as the determination of the optimal operating set-points ( v i ) of the HVAC equipment, for each time instant t of the optimization horizon, where {  $t \in N \mid 1 \leq t \leq K$  }, with objective to satisfy the thermal energy demand of the building (L), while minimizing a multi-criteria function ( f trans ) and satisfying the established bus temperature operating zone in order to provide the required comfort level.

Minimize: 
$$\sum_{t=1}^{K} \sum_{j=1}^{C} f_j^{trans,t}(v_i^t)$$
 (1)

Subject to: 
$$L_t - \sum_{i=1}^n (P_i^t * \eta_i) = 0$$
 (2)

$$T_{min} \le T_t \le T_{max} \tag{3}$$

$$\sum_{i=1}^{n} P_i^t \le P_{max}^{grid} \tag{4}$$

$$P_i \le P_i^t * \eta_i \le \overline{P_i} \tag{5}$$

In the above formulation, C describes the index of the optimization criteria, n is an index that describes the number of HVAC equipment, T min and T max indicate the lower and upper temperature bounds for the distribution bus, while P grid max describes the maximum electric power that the installation may supply to the cooling equipment, and P i , P i , P i and  $\eta i$  describe the energy consumption, minimum power generation, maximum power generation and COP of the equipment i , respectively. For the formulation of the multi-objective function, the weighted global criterion method has been used, in which all the objective criteria are combined to form a single criterion. In order to sum the results of each criterion, a unit normalization function is made, as described in (6).

$$f_{j}^{trans} = \frac{f_{j}(v_{i}) - f_{j}^{o}}{f_{j}^{max} - f_{j}^{o}}$$
 (6)

Whereas  $f_j^{trans}$  is described the transformed objective function value of criterion j, and as  $f_j^o$  and  $f_j^{max}$  are the utopia point of and maximum objective function values of criterion j, respectively. In this case, the optimization criteria are the combination of the maximization of the COP and the minimization of the switching cost. The COP metric is defined as the quotient between the output and input power, while the

switching cost is defined as the sum of differences in magnitude of the setpoints considered for each of the machines.

#### 3.2. Controller implementation

A step-by-step diagram of the proposed optimization and control framework is shown in Fig. 3, which is composed of three stacked control loops: 1) operating mode selection, 2) predictive control strategy and 3) production profile configuration.

## 1) Operating mode selection

This is the inner-most control loop, which relies on the performance map obtained by means of the deep learning methodology to estimate the cooling capacity, consumption and performance of viable setpoints, in order to select the optimal control state for the current instant. This stage is executed as follows. First, the current operating state is evaluated e) by acquiring the signals affecting the control: impulsion and return temperatures, external temperature, humidity, current cooling demand and previous operating state of the group of machines.

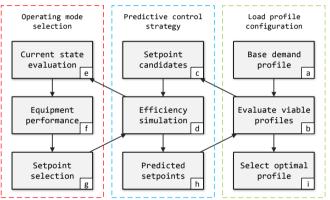


Fig. 3. Step diagram of the implementation of the proposed three-stage optimization and control framework.

Afterwards, the setpoint candidates are determined by establishing the possible control actions. This is achieved by exhaustive combinatorial of control states of the individual machines. For ma-chines having discrete operating modes, the full range of modes will be employed, but for machines having a continuous operating range this is not possible, so the operating range is discretized up to the desired granularity. In practice, coarse granularity shall be sufficient, given that HVAC equipment cannot be regulated with infinite precision. Each combination is then evaluated in conjunction with the current operating state in order to determine the equipment's performance f) by feeding it through the performance model. The exhaustive evaluation of the full solution space is made possible by the modeling approach, which implements the evaluation of the model as a series of matrix products and thus becomes very efficient. The process of training the performance model can be lengthy and take around 20 min, but the resulting model can be evaluated with an average runtime of 4.2 milliseconds per iteration. This exhaustive evaluation of the solution space would re- main feasible even in very large installations through partitioning and evaluation in concurrent batches. Finally, having obtained the cooling capacity, electrical consumption and COP of each control setting, these are scored and sorted in order to select the locally optimal setpoint g) according to the defined objective function, which considers performance and switching cost to apply the new setpoint.

#### 3.2.2. Predictive control strategy

This loop is proposed to operate on top of the previous one, iterating over N steps in the prediction horizon to determine the optimal future sequence of control actions that will satisfy the fore- casted load demand. Instead of selecting the best performing control action as determined in f), the top setpoint subset is selected and used as starting setpoint candidates c) which are then evaluated using a depth-first algorithm. Each of these setpoints is used as the starting point of a control sequence evaluated N steps into the future, which shall allow to determine which of the starting points leads to the most beneficial outcome. This is motivated by the fact that control setpoints that are locally optimal for the current control iteration may not be globally optimal due to switching to a state that is suboptimal, because it might be preventing that more efficient states are reached later on. This is likely to happen due to the switching being considered in the objective function, which penalizes control actions that make drastic changes to the setpoints, thus selecting the absolute best control action for the next step might lead to suboptimal situations. Therefore, the performance of a subset of the top performing candidate setpoints is applied over the optimization horizon to evaluate the full implications of taking that control action. How- ever, this presents one main difficulty: the evaluation of future control actions is not possible due to the operating state being un- known. The data necessary for evaluating a control setpoint using the performance map can be classified into three types:

- Known future state: part of the future operating state is known, i.e. the future cooling demand is obtained from the load forecasting model evaluation, and the external temperature and humidity can be queried from local weather services.
- Unknown future state: state that relates to the operating state of the machines, in this case each machine's inlet and outlet temperatures.
- Control action: the setting of each of the machines in the equipment group, which is the optimization variable and is always known for the previous control iteration.

Thus, knowing the future state of bus temperatures is required to be able to apply the operating mode control loop during future iterations in order to determine the future control sequence. To solve this problem, a simulation model of the distribution bus is employed, implemented in this case as a first order energy storage model with a single capacity coefficient, estimated using the historical data available. Using the bus model, a control sequence is determined by the described depth-first algorithm that iterates from each of the set-point candidates, simulating the application of the control action on the distribution bus d) while keeping track of the COP achieved by each sequence over the optimization horizon. Each of the obtained control sequences took locally optimal decisions at each iteration according to the objective function, thus finally the entire control sequences are evaluated using the objective function, which allows the final selection of the predicted setpoints h), i.e. the optimal control sequence over the optimization horizon for matching the predicted cooling load profile.

#### 3.2.3. Production profile configuration

The previous loop focused on the determination of the optimal control sequence to force the matching between the cooling demand and production, however this constraint can be relaxed on account of the thermal dynamics of the system. In

order to ensure that the equipment on the consumption stage of the HVAC system are able to effectively draw power from the distribution bus, production stage controllers focus on maintaining the temperature of the bus within a certain range, i.e. the goal of the production controller is to keep the temperature of the bus between a minimum and maximum threshold. Thus, the bus acts as the buffer between the production and the consumption stages, but due to the thermal dynamics of the building and associated thermal capacity, it allows the temporal decoupling of the production and consumption equipment, in a similar manner as dedicated energy storage equipment would al- low. Therefore, the production setpoints can be altered as long as the temperature of the bus is kept within the required thresholds, which is a property of the system that is proposed to be exploited in order to increase the performance. By shifting the load of the production equipment in time it is possible to take advantage of periods of time where conditions are more favorable, for example conditions such as better COP due to affecting operating state or varying cost of energy.

This control stage handles the determination of a production profile that shall take advantage of favorable production conditions while satisfying the temperature constraints. To solve this problem, a common approach is to employ global optimization tools like particle swarm optimization or genetic algorithm, mainly due to their capacity to overcome the non-derivable nature of the problem. However, these tools may present robustness issues and could be impractical for cases that involve running simulations, which makes them inadequate for solving this type of control problem [49]. Instead, a heuristics technique is employed which searches the possible production profile configuration space according to the following approach: i) an initial load profile is determined by considering the production uniform throughout the control horizon, with a value equal to the average load demand in the same period, ii) the maximum and minimum production rate that would lead the system to exceed the temperature thresholds are computed and used as upper and lower bounds, respectively, and iii) the space is binarysearched up to a predefined amount of iterations. The application of the binary search algorithm is especially useful because it allows to cover the full solution space while choosing a desired precision specified by the number of iterations, thus allowing to adjust it depending on the scale and requirements of the problem while ensuring that a solution is found in deterministic time. This is desirable property in control applications and very important in this case due to the fact that evaluation of each configuration depends on the evaluation of the previous stages which include simulation.

#### 4. TEST ENVIRONMENT

For the experimental validation of the proposed methodology, data from a real tertiary-sector building has been used as a test- bench. The selected building consists of a 3-floor university cam- pus building that contains offices and laboratories, with a total surface of 2.400m 2. The selected building is considered a research ecosystem of the Universitat Politècnica de Catalunya —BarcelonaTech (Spain), which includes an installation of renewable energy sources (photovoltaics), several energy production equipment, as well as a SCADA system instrumented in detail, permitting to be used as a pilot-plant for researches in the field of energy efficiency, smart-grids and industrial electronics among others. Fig. 4 depicts a 3D representation model of the building's

structure, consisting of 2 building blocks of 3 and 1 floors, respectively. The HVAC ma- chines that were used for the validation of the proposed method-ology are located on the upper part the building, on its deck. In terms of heating and cooling production equipment, the installation consists of two electric chillers, two heat pumps, one gas boiler and two air handling units, which manage the energy production, energy distribution, pre-conditioning and air-renewal for the building's spaces. The power characteristics of the HVAC are listed in Table 1. The existing controller implemented in the building prior to the development of the proposed strategy is a standard real-time controller with round-robin production equipment allocation. The supervision and control of all of the installation is made through a main SCADA system, which monitors the operation of the equipment, the condition of the heating and cooling distribution bus, as well as the environment conditions, such as the weather, the occupancy of the spaces and the temperature setpoints configured by occupants. For the development and validation of the proposed control strategy, a dataset of the equipment's operation has been acquired by recording the operation of the building's cooling equipment at 2-minute sampling frequency. The dataset comprises 120 days, from May 16th to October 27th of 2017, not including weekends. A summary of the acquired signals comprising the dataset is presented in Table 2.



Fig. 4 3D representation of the pilot plant, highlighting the location of the HVAC equipment on the building's deck

 $\begin{tabular}{l} TABLE~1\\ POWER CHARACTERISTICS~OF~THE~HVAC~MACHINES~OF~THE~VALIDATION\\ PLANT \end{tabular}$ 

Id	Туре	P <sub>elec</sub> [kW]	P <sub>thermal</sub> [kW]
CH1	Electric chiller	56.6	150
CH2	Electric chiller	56.6	150
HP1	Heat pump	56.7	130
HP2	Heat pump	66.2	150
B1	Gas boiler	2	430
AHU1	Air handling unit	5.5	n/a
AHU2	Air handling unit	7.5	n/a

 $\label{eq:table 2} TABLE~2\\ SUMMARY OF ACQUIRED SIGNALS FROM THE SCADA SYSTEM$ 

Name	Description	
Text	Outdoor temperature	
Hrel	Outdoor relative humidity	
Peleci	Electrical power consumption of equipment i	
Ptheri	Thermal power production of equipment i	
COPi	Coefficient of performance of equipment i	
Timpi	Bus impulsion temperature of equipment i	
Treti	Bus return temperature of equipment i	
Tinc	Temperature differential of equipment i	
Ncomp	Number of active compressors of equipment i	

#### 5. EXPERIMENTAL RESULTS

This section shows the implementation of the proposed methodology and discusses the obtained experimental results in the described test environment. The objective is to demonstrate the performance gained by taking advantage of the integration of data-driven models with the controller and the effectiveness of the developed control strategy in increasing the overall energy efficiency. For this purpose, the stages the control strategy are applied in steps, showing the effect of each stage relative to the base control strategy. Regarding the software implementation, all algorithms were developed using Matlab R2018b, except for the equipment performance modeling which was implemented using the TensorFlow library in Python 3.6. The implementation was carried out in a Windows 10 desktop workstation with an i5-3470 processor and 16 GB of memory.

## 5.1. Operating mode selection

This stage operates in current time, selecting the best operating mode for the next control iteration considering how the performance of the equipment is affected by the operating conditions. However, since at each control iteration a new mode can be selected, this control results in excessive commutation, as slight change in the state can lead to another mode surpassing its instantaneous performance. To overcome this issue, two mitigating actions are considered; first, the setpoint candidates are truncated post-evaluation to those causing up to a maximum switching cost; second, the control frequency is decreased in order to limit the switching.

The result of operating the group of chillers with this strategy is shown in Fig. 5, which shows how the selected control modes perform, compared to the base controller. A comparison between the results obtained by the application of the new controller and the historic behavior of the base controller are shown in Fig. 5, where labels marked with an asterisk \*refer to the value of variables with the application of the new controller, while others refer to the base controller. As it can be observed in the 2-hour time window represented in the figure, the result of the application of the new controller is similar to the result obtained by the base controller in terms of energy production and bus temperature. The accumulated production value was reduced by 1.16% in this period, causing the bus temperature to become warmer at the end of the period. However, the new controller selects operating modes that overall are more efficient than the base controller, which is unaware of the performance of the equipment. This behavior can be consistently observed throughout the day, except for periods of time where the base controller selects the optimal setpoint by chance, or the switching cost from the current state is too great.

# 5.2. Predictive control strategy

This stage considers the determination of control sequence that matches the forecasted cooling demand for the prediction horizon, in this case of 1 h because it allows to plan control actions with sufficient foresight given the dynamics observed in the building's dataset, which are in the range from two to three hours. The set- ting of the control action, i.e. operation mode of the equipment, is restricted to allow a change every 15 min to limit the amount of switching, so 4 control points are calculated within each control horizon. The improvement that this stage offers is based on consider- ing how the selection of an operating mode affects the outcome of the complete prediction horizon, instead of only the current control step.

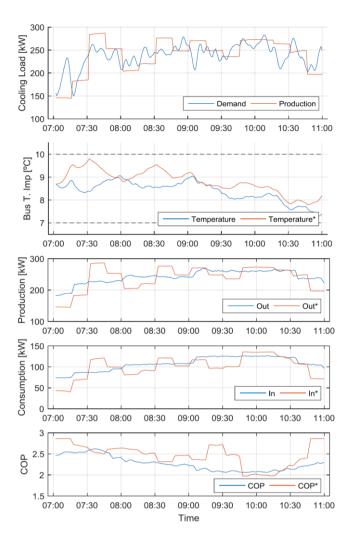


Fig. 5. Result of the application of the selected operating modes, compared to the operation of the base controller

The result of the application of this strategy is shows in Fig. 6, which shows the cumulative average performance of a set of initial control setpoint candidates as each one is evaluated over the next hour, which is the length of the prediction horizon in this case. As it can be observed, setpoints that initially lead to the largest instantaneous performance are not necessarily the best option once the prediction range is evaluated. For example, the operating mode that offers the best COP at the start of the prediction horizon leads to lower average performance when considering the behavior of the full time window than the setpoint that is initially ranked fifth best. This behavior is due to the imposed switching constraint which discourages drastic changes between operating modes. Therefore, this control stages achieves the desired effect, the anticipation of favorable and adverse conditions and the selection of a predictive control sequence that leads to the best average performance.

## 5.3. Load profile configuration

This stage comprises the determination of a better control se- quence by reconfiguring the load profile, considering the forecasted cooling demand and the operational constraints of the distribution bus. Thus, this strategy allows more freedom in the determination of the control sequence, since each of the selected control points do not need to match the instantaneous load, i.e. load shifting is allowed as long as the temperature thresholds are not exceeded.

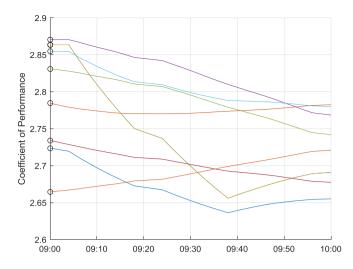


Fig. 6. Cumulative average performance of a set of setpoint candidates, evaluated over the prediction horizon

As described, the load profile reconfiguration process begins with the determination of the initial control sequence profile, and the maximum and minimum production sequences that would keep the system operating within thresholds. Then these profiles are evaluated and binary search is employed for determining the best profile configuration, allowing up to a maximum number of iterations. In this case 4 iterations are employed, as it is a good compromise between precision and computation time. The result of the implementation of this profile selection strategy is shown in Fig. 7 where it can be observed that different control sequences are evaluated within the temperature thresholds. The forecasted cooling demand and the base control sequence, which matches in magnitude the average demand on each control period, are shown highlighted in Fig. 7 a) for the next hour according to the prediction horizon, while the other bus load profiles correspond to some of the attempted control sequences, having varying magnitudes within the ranges from the minimum viable production to the base sequence and from the maximum viable production to the base sequence. The actual production value is curtailed to ensure the bus temperature does not exceed the limits, considering that a control action can only be carried out every 15 min. The simulated bus temperature response when each of the control sequences in Fig. 7 a) is applied is shown in Fig. 7 b). As it can be observed, each simulation begins at the current temperature reading, and gets cooler or warmer depending on whether the magnitude of the sequence is above or below the base control sequence, respectively. The base sequence is stabilized at the initial temperature value, with slight fluctuations corresponding to the changes in the load, because it's values are calculated to match the average cooling demand per control period, while other sequences follow different trajectories with varying steepness depending on the magnitude of the production.

## 5.4. Energy efficiency improvements

A comparative summary of the results achieved by all three control stages is presented in Fig. 8, which shows a histogram of the performance obtained when applying the relevant stage over single day periods, for each of the 120 days available in the dataset. The experimental results of the application of the three-stage control strategy consistently show a performance increase through- out the dataset. The application of the first controller stage, consisting on the selection of the operating mode based on current load demand, achieves an improvement

of between 3.92% and 17.05% of the COP respective to the base controller, with daily average increase of 10.88%, surpassing 10% performance increase in a significant part of the considered dates. However, there's cases where this stage is unable to achieve such improvements, which is due to the base controller already operating on moderately efficient modes, and the inability of this stage to realize large improvements due to being penalized by the consideration of the switching in the cost function. The application of the second stage, supported by the usage of the load forecasting capabilities to anticipate demand changes, achieves an improvement in the range of 9.84% to 24.11% respective to the base controller, with an average increase of 17.27%. This significant performance increase is due to the ability of this stage to consider the future control sequence, allowing the selection of control setpoints that lead to continued efficient operation while overcoming the switching minimization consideration. Finally, the proposed control strategy is realized with the incorporation of the third stage, the reonfiguration of the load profile to take advantage of favorable production conditions by taking ad-vantage of the thermal dynamics. The application of the full strategy achieves a performance improvement in the range of 12.39% to 24.30% respective to the base controller, and an average increase of 19.54%. As described, the performance gain respective to the second stage is not as extensive as when comparing the second stage respective to the first one, in part due to the second stage already accomplishing a highly efficient control sequence, but also due to this installation not including dedicated thermal storage.

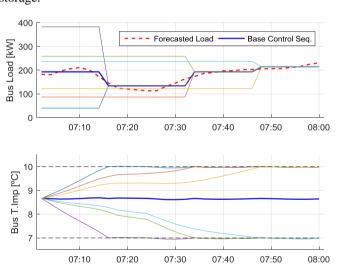


Fig. 7. Result of the application of the selected control sequences, compared to the operation of the base sequence that matches the averaged forecasting.

## 6. CONCLUSIONS

This paper introduces a framework for implementing a control strategy aimed at solving the optimal loading and scheduling problem in HVAC installations.

The framework is based on the selection of the control sequence that maximizes the performance of the production equip- ment in an HVAC installation by driving the machines to their most efficient setpoints, supported by the integration of data-driven models. A neural network-based model of the equipment's COP be- havior respective to operating conditions and implemented using a deep learning approach is used for evaluating the response to po- tential setpoint candidates, while

a model of the building's thermal load demand considering the building occupant's behavior is used for anticipating the energy production requirements.

The control sequence determination problem is formulated as an optimization problem, but finding a solution is a complex task, as the problem is not derivable and the exhaustive search of the solution space is infeasible. Instead of using a global optimization tool, a heuristics-based control method composed of three stages is designed and implemented achieving a substantial performance increase while maintaining a low and most importantly constant computational time. The obtained results consistently show a performance increase by the implementation of the control strategy, with the complete solution achieving a 19.54% daily average COP increase with 2.68% standard deviation. These results are coherent with the efficiency improvement potential inferred by related studies in the state- of-the-art literature over current control solutions in established HVAC systems. As future work, additional optimization criteria could be considered to further fine-tune the control sequence, for example the uniform utilization of the equipment could be enforced so that all of the machine's aging follows a similar rate, or the variable cost of energy which could be paired with the load profile determination to achieve greater economic savings.

#### **Declaration of Competing Interest**

We wish to confirm that there are no known conflicts of inter- est associated with this publication and there has been no signifi- cant financial support for this work that could have influenced its outcome.

## REFERENCES

- [1] K. Deng, Y. Sun, S. Li, Y. Lu, J. Brouwer, P.G. Mehta, M.C. Zhou, A. Chakraborty, Model predictive control of central chiller plant with thermal energy storage via dynamic programming and mixed-integer linear programming, IEEE Trans. Autom. Sci. Eng. 12 (2015) 565–579, doi: 10.1109/TASE.2014.2352280.
- [2] Z. Afroz, G.M. Shafiullah, T. Urmee, G. Higgins, Modeling techniques used in building HVAC control systems: a review, Renew. Sustain. Energy Rev. 83 (2018) 64–84, doi: 10.1016/j.rser.2017.10.044.
- [3] S. Huang, W. Zuo, M.D. Sohn, Amelioration of the cooling load based chiller sequencing control, Appl. Energy 168 (2016) 204–215, doi: 10.1016/j.apenergy. 2016.01.035.
- [4] F.W. Yu, K.T. Chan, Energy signatures for assessing the energy performance of chillers, Energy Build. 37 (2005) 739–746, doi: 10.1016/j.enbuild.2004.10.004.
- [5] Y. Wang, X. Jin, X. Fang, Rapid evaluation of operation performance of multi- chiller system based on history data analysis, Energy Build. 134 (2017) 162–170, doi: 10.1016/j.enbuild.2016.10.041.
- [6] Z. Liu, H. Tan, D. Luo, G. Yu, J. Li, Z. Li, Optimal chiller sequencing control in an office building considering the variation of chiller maximum cooling capacity, Energy Build. 140 (2017) 430–442, doi: 10.1016/j.enbuild.2017.01.082.
- [7] L. Dos Santos Coelho, V.C. Mariani, Improved firefly algorithm approach applied to chiller loading for energy conservation,

- Energy Build. 59 (2013) 273–278, doi: 10.1016/j.enbuild.2012.11.030.
- [8] B.M. Seo, K.H. Lee, Detailed analysis on part load ratio characteristics and cool- ing energy saving of chiller staging in an office building, Energy Build. 119 (2016) 309–322, doi: 10.1016/j.enbuild.2016.03.067.
- [9] X. Fang, X. Jin, Z. Du, Y. Wang, The evaluation of operation performance of HVAC system based on the ideal operation level of system, Energy Build. 110 (2016) 330–344, doi: 10.1016/j.enbuild.2015.11.020.
- [10] F. Sohrabi, M. Nazari-Heris, B. Mohammadi-Ivatloo, S. Asadi, Optimal chiller loading for saving energy by exchange market algorithm, Energy Build. 169 (2018) 245–253, doi: 10.1016/j.enbuild.2018.03.077.
- [11] Y. Liao, G. Huang, Y. Ding, H. Wu, C. Fan, Robustness analysis and enhancement of chiller sequencing control under uncertainties, Procedia Eng. 205 (2017) 1878–1885, doi: 10.1016/j.proeng.2017.10.270.
- [12] Y. Liao, G. Huang, A hybrid predictive sequencing control for multi-chiller plant with considerations of indoor environment control, energy conservation and economical operation cost, Sustain. Cities Soc. 49 (2019) 101616, doi: 10.1016/j. scs.2019.101616.
- [13] L. Wang, E.W.M. Lee, R.K.K. Yuen, A practical approach to chiller plants' opti- misation, Energy Build. 169 (2018) 332–343, doi: 10.1016/j.enbuild.2018.03.076.
- [14] C. Dullinger, W. Struckl, M. Kozek, A general approach for mixed-integer pre- dictive control of HVAC systems using MILP, Appl. Therm. Eng. 128 (2018) 1646–1659, doi: 10.1016/j.applthermaleng.2017.09.104.
- [15] P. Lee, P.T.I. Lam, W.L. Lee, E.H.W. Chan, Analysis of an air-cooled chiller re- placement project using a probabilistic approach for energy performance con-tracts, Appl. Energy 171 (2016) 415–428, doi: 10.1016/j.apenergy.2016.03.035
- [16] X. Wei, G. Xu, A. Kusiak, Modeling and optimization of a chiller plant, Energy 73 (2014) 898–907, doi: 10.1016/j.energy.2014.06.102.
- [17] H.S. Asad, R.K.K. Yuen, G. Huang, Degree of freedom based set-point reset scheme for HVAC real-time optimization, Energy Build. 128 (2016) 349–359, doi: 10.1016/j.enbuild.2016.07.013.
- [18] M.S. Soler, C.C. Sabaté, V.B. Santiago, F. Jabbari, Optimizing performance of a bank of chillers with thermal energy storage, Appl. Energy 172 (2016) 275–285, doi: 10.1016/j.apenergy.2016.03.099.
- [19] F.W. Yu, W.T. Ho, K.T. Chan, R.K.Y. Sit, Logistic regression-based optimal con- trol for air-cooled chiller, Int. J. Refrig. 85 (2018) 200–212, doi: 10.1016/j.ijrefrig. 2017.09.026.
- [20] M. Karami, L. Wang, Particle Swarm optimization for control operation of an all-variable speed water-cooled chiller plant, Appl. Therm. Eng. 130 (2018) 962–978, doi: 10.1016/j.applthermaleng.2017.11.037.
- [21] H. Satyavada, S. Baldi, An integrated control-oriented modelling for HVAC per- formance benchmarking, J. Build. Eng. 6 (2016) 262–273, doi: 10.1016/j.jobe. 2016.04.005.
- [22] T. Ma, J. Wu, L. Hao, W.-.J. Lee, H. Yan, D. Li, The optimal structure planning and energy management strategies of smart multi energy systems, Energy 160 (2018) 122–141, doi:

- 10.1016/J.ENERGY.2018.06.198.
- [23] Y. Sun, S. Wang, G. Huang, Online sensor fault diagnosis for robust chiller sequencing control, Int. J. Therm. Sci. 49 (2010) 589–602, doi: 10.1016/j. ijthermalsci.20 09.10.0 03.
- [24] W. Wang, M.R. Brambley, W. Kim, S. Somasundaram, A.J. Stevens, Automated point mapping for building control systems: recent advances and future re- search needs, Autom. Constr. 85 (2018) 107–123, doi: 10.1016/j.autcon.2017.09.013.
- [25] H.Z. Abou-Ziyan, A.F. Alajmi, Effect of load-sharing operation strategy on the aggregate performance of existed multiple-chiller systems, Appl. Energy 135 (2014) 329–338, doi: 10.1016/j.apenergy.2014.06.065.
- [26] M. Li, Y. Ju, The analysis of the operating performance of a chiller system based on hierarchal cluster method, Energy Build. 138 (2017) 695–703, doi: 10.1016/j. enbuild.2016.12.076.
- [27] S.R. Thangavelu, A. Myat, A. Khambadkone, Energy optimization methodol- ogy of multi-chiller plant in commercial buildings, Energy 123 (2017) 64–76, doi: 10.1016/j.energy.2017.01.116.
- [28] E. Salari, A. Askarzadeh, A new solution for loading optimization of multi- chiller systems by general algebraic modeling system, Appl. Therm. Eng. 84 (2015) 429–436, doi: 10.1016/j.applthermaleng.2015.03.057.
- [29] T. Dwyer, Knowledge is power: benchmarking and prediction of building en- ergy consumption, Build. Serv. Eng. Res. Technol. 34 (2013) 5–7, doi: 10.1177/0143624412471130.
- [30] R. Jovanovi 'c, A.A. Sretenovi 'c, B.D. Živkovi 'c, Ensemble of various neural net- works for prediction of heating energy consumption, Energy Build. 94 (2015) 189–199, doi: 10.1016/j.enbuild.2015.02.052.
- [31] A . Ghahramani, S.A . Karvigh, B. Becerik-Gerber, HVAC system energy optimization using an adaptive hybrid metaheuristic, Energy Build. 152 (2017) 149–161, doi: 10.1016/j.enbuild.2017.07.053.
- [32] A. Afram, F. Janabi-Sharifi, Review of modeling methods for HVAC systems, Appl. Therm. Eng. 67 (2014) 507–519, doi: 10.1016/j.applthermaleng.2014.03.055.
- [33] F.W. Yu, W.T. Ho, K.T. Chan, R.K.Y. Sit, Critique of operating variables importance on chiller energy performance using random forest, Energy Build. 139 (2017) 653–664, doi: 10.1016/j.enbuild.2017.01.063.
- [34] J. Yu, X. Zheng, S. Wang, A deep autoencoder feature learning method for process pattern recognition, J. Process Control 79 (2019) 1–15, doi: 10.1016/j. jprocont.2019.05.002.
- [35] J. Grant, M. Eltoukhy, S. Asfour, Short-term electrical peak demand forecast- ing in a large government building using artificial neural networks, Energies 7 (2014) 1935–1953, doi: 10.3390/en7041935.
- [36] J. Zhao, X. Liu, A hybrid method of dynamic cooling and heating load forecast- ing for office buildings based on artificial intelligence and regression analysis, Energy Build. 174 (2018) 293–

- 308, doi: 10.1016/j.enbuild.2018.06.050.
- [37] T. Hong, S. Fan, Probabilistic electric load forecasting: a tutorial review, Int. J. Forecast. 32 (2016) 914–938, doi: 10.1016/j.iiforecast.2015.11.011.
- [38] C. Kuster, Y. Rezgui, M. Mourshed, Electrical load forecasting models: a criti- cal systematic review, Sustain. Cities Soc. 35 (2017) 257–270, doi: 10.1016/j.scs. 2017.08.009.
- [39] A.H. Neto, F.A.S. Fiorelli, Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption, Energy Build. 40 (2008) 2169–2176, doi: 10.1016/j.enbuild.2008.06.013.
- [40] Z. Wang, R.S. Srinivasan, A review of artificial intelligence based building en- ergy use prediction: contrasting the capabilities of single and ensemble pre- diction models, Renew. Sustain. Energy Rev. 75 (2017) 796–808, doi: 10.1016/j. rser.2016.10.079.
- [41] T. Hong, S.C. Taylor-Lange, S. D'Oca, D. Yan, S.P. Corgnati, Advances in re- search and applications of energy-related occupant behavior in buildings, En- ergy Build. 116 (2016) 694–702, doi: 10.1016/j.enbuild.2015.11.052.
- [42] M. Peña, F. Biscarri, J.I. Guerrero, I. Monedero, C. León, Rule-based system to detect energy efficiency anomalies in smart buildings, a data mining approach, Expert Syst. Appl. 56 (2016) 242–255, doi: 10.1016/j.eswa.2016.03.002.
- [43] M. Jia, R.S. Srinivasan, A.A. Raheem, From occupancy to occupant behavior: an analytical survey of data acquisition technologies, modeling methodologies and simulation coupling mechanisms for building energy efficiency, Renew. Sustain. Energy Rev. 68 (2017) 525–540, doi: 10.1016/j.rser.2016.10.011.
- [44] J. Massana, C. Pous, L. Burgas, J. Melendez, J. Colomer, Short-term load fore- casting for non-residential buildings contrasting artificial occupancy attributes, Energy Build. 130 (2016) 519–531, doi: 10.1016/j.enbuild.2016.08.081.
- [45] G. Qiang, T. Zhe, D. Yan, Z. Neng, An improved office building cooling load prediction model based on multivariable linear regression, Energy Build. 107 (2015) 445–455, doi: 10.1016/j.enbuild.2015.08.041.
- [46] L. Wang, E.W.M. Lee, R.K.K. Yuen, Novel dynamic forecasting model for build- ing cooling loads combining an artificial neural network and an ensemble approach, Appl. Energy 228 (2018) 1740–1753, doi: 10.1016/j.apenergy.2018.07. 085.
- [47] X. Li, J. Wen, System identification and data fusion for on-line adaptive energy forecasting in virtual and real commercial buildings, Energy Build. 129 (2016) 227–237, doi: 10.1016/j.enbuild.2016.08.014.
- [48] E. Sala-Cardoso, M. Delgado-Prieto, K. Kampouropoulos, L. Romeral, Activity- aware HVAC power demand forecasting, Energy Build. 170 (2018) 15–24, doi: 10.1016/j.enbuild.2018.03.087.
- [49] T. Wortmann, Genetic evolution vs. function approximation: benchmarking al- gorithms for architectural design optimization, J. Comput. Des. Eng. 6 (2019) 414–428, doi: 10.1016/j.jcde.2018.09.001