UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI INGEGNERIA INDUSTRIALE DII

CORSO DI LAUREA MAGISTRALE IN INGEGNERIA ENERGETICA

CALIBRATION OF A GREY BOX MODEL USING PARTICLE SWARM OPTIMIZATION ON DIFFERENT BUILDING STRUCTURES

Relatore
Prof. Angelo Zarrella
Correlatore
Ing. Jacopo Vivian

Laureando
Mattia D'Incà
Matricola
1206194

ANNO ACCADEMICO 2020 - 2021

a nonna Gioconda

Abstract

L'obiettivo dello studio è stato quello di testare la calibrazione di un modello a parametri concentrati per lo sviluppo di un sistema BEMS (Building Energy Managment System) basato su MPC (Model Predictive Control), che ottimizzi il funzionamento di una pompa di calore accoppiata con moduli fotovoltaici. L'ottimizzazione, basata su approccio predittivo, permette di aumentare la flessibilità dell'edificio spostando i consumi energetici nelle finestre di maggior convenienza economica, minimizzando i costi.
La calibrazione viene effettuata con il metodo PSO (Particle Swarm Optimization), un algoritmo euristico ispirato al comportamento degli sciami, che sfrutta tecniche di apprendimento automatico (Machine Learning) per minimizzare lo scarto quadratico medio tra il profilo di temperature misurate nell'edificio e quelle calcolate dal modello 5R1C.
Per prima cosa il codice di calibrazione è stato applicato al laboratorio pilota dell'RSE di Piacenza, testandolo in tre dataset differenti e osservandone il comportamento al variare dei periodi di training e testing.
Dopodiché tramite il software EnergyPlus™, sono state simulate alcune strutture per testare la calibrazione in strutture di tipo residenziale con diversi isolamento e capacità termica. Dopo aver validato il modello EnergyPlus, il codice di calibrazione è stato applicato alle strutture simulate, confermando quanto dimostrato dalle prove sul laboratorio reale e permettendo di effettuare alcune ulteriori considerazioni sull'effetto di isolamento e capacità termica dell'involucro edilizio.
I risultati mostrano come sia conveniente effettuare la calibrazione con frequenza giornaliera e su un periodo di training compreso tra 2 e 4 giorni. Inoltre, è emerso che l'accuratezza dell'algoritmo è maggiore in strutture più massive e con maggior isolamento termico.
Infine si è indagato sul senso fisico dei parametri del modello RC a valle della calibrazione, dimostrando come non ci siano differenze sostanziali in termini di RMSE nei risultati ottenuti con i parametri iniziali scelti pari a quelli nominali corretti e in quelli con parametri iniziali scelti a caso in un intorno fisicamente coerente.

Contents

1 Introduction

1.1 Motivation and backgrounds

Buildings like homes, workplaces, schools, hospitals, libraries and public buildings in general are the largest energy consumer in the EU and one of the major contributors to CO_2 emissions. Overall, EU buildings and their constructions together account for 40 % of energy consumption and 36 % of energy-related greenhouse gas emissions annually [1].
Improving the energy efficiency of buildings is therefore crucial to achieving the ambitious carbon neutrality target by 2050, as defined in the European Green Deal [2].
Moreover, the European Directive 2018/44 requires to reduce greenhouse gas emissions by at least 40 % by 2030, with the aim of encouraging the construction of zero-emission buildings and promoting the use of home automation and intelligent technologies [3].
In recent years many research groups have worked with the aim of improving the energy efficiency of buildings, focusing not only on the retrofitting and modernization, but also on the application of advanced control systems. In particular, building energy management systems (BEMS) based on predictive control (MPC) are becoming increasingly attractive in both academy research and industry.

Building energy management systems can help to improve the energy performance of buildings by providing some flexibility to the building side, finding a compromise between the thermo-hygrometric comfort in the indoor environment, the integration of local renewable sources, the reduction of consumption and costs of the energy used by the end-user.
The energy management system usually consists of a hardware part, which receives signals from the building and a software part that develops solutions that coordinate the operation of heating systems, domestic hot water production, cooling and dehumidification installations and mechanical ventilation installations if present. This supervision and coordination activity carried out by BEMS therefore acts on a higher level than the controllers (actuators, relays, inverters, thermostats, etc.) of the individual plants and they must be able to receive signals from BEMS and react appropriately. These systems can also consider other electrical loads such as lights and other equipments and also any photovoltaic systems or other local thermal or electricity generation systems. They can also be coupled with shading devices to control the solar radiation incident on the building.
BEMS carries out this coordination activity by pursuing a certain objective, which can be either a local objective or a system objective. In the first case, the objective is linked to the

building in which the BEMS is installed, that could be the minimization of energy costs; in the second case, the interests of actors outside the building in question are considered. An example of this could be the provision of an auxiliary service to the electricity distribution network to which the building is connected.
BEMS control strategies can be in general resumed in two categories [4]: rule-based controls (RBC) and Model Predictive Control (MPC). Rule-based controls are simple methods which generally have the form 'if (condition is verified), then (action is made)'. RBC usually control a specific parameter (PV power, room temperature) on which a threshold value has been fixed. When the threshold is reached, the operation of the heat pump is changed, according to the predefined strategy. On the other hand, Model Predictive Control is a more complex strategy, which relies on a model of the building to projectits behavior in the future. These systems, to make optimal decisions, are based on forecasts of what will happen in the next few hours, so as to intervene in advance according to the forecasts of local energy demand and production in the next few hours. The signals received by BEMS MPC-based, come therefore both from the internal environment and from the outside, because it may be necessary to use information provided by third parties, such as weather forecasts. This makes it necessary to connect the BEMS not only to a local communication network to receive the signals of the sensors installed at the building, but also to the internet network through which the signals provided by third parties can arrive.

1.2 Overview and objectives

The research work concerns the development of an energy management system that will be tested in the RSE Laboratory of Piacenza. The aims of the research project are to analyze and exploit the potential of energy flexibility for different buildings, according to their construction characteristics, installations and their intended use (residential, commercial buildings, etc.).
In the pilot building of Piacenza, the plant will be managed in such a way as to ensure the comfort of indoor environments, where the heating energy in winter and cooling energy in summer are produced by an air-water heat pump. The core of the BEMS is the control algorithm of the heat pump developed in Python, based on a lumped capacitance model of the laboratory and able to decide autonomously the time windows of greater economic convenience, maximizing the self-consumption of electricity produced by a photovoltaic system. Continuous monitoring of the internal thermo-hygrometric conditions and weather forecast, will allow the self-calibration of the model parameters through machine learning techniques. The model parameters (resistances and thermal capacities) are first estimated from a limited number of information about the building and are then 'learned' during operation through the so-called calibration. The control system will then make autonomous decisions, based on weather forecast, users behavior and preferences [5].

The present study focuses on the calibration process, where the parameters of a lumped capacitance model are calibrated using Particle Swarm Optimization (PSO) iteratively to reproduce the average indoor air temperature pattern. In fact, Building Energy Managements Systems (BEMS) based on Model Predictive Control (MPC) must be calibrated periodically in order to accurately reproduce the dynamic thermal behavior of buildings [6]. The calibration process avoids that the actual indoor air temperature diverge from that predicted by the underlying building model, in order to have a more efficient optimization.

In the first part the study case will be described and the fundamental principles of BEMS systems based on predictive control will be illustrated, such as grey box concept, lumped capacitance models and PSO algorithm.

In the second part the experimental part will be carried on, where the code will be practically applied.
The investigation included the most appropriate length of the training period used by the PSO algorithm to produce the best indoor air temperature pattern.
The laboratory has been modeled through the software EnergyPlus and used to produce additional data for calibration, assuming heavyweight structures (masonry with external insulation as usual in nowadays building practice) instead of lightweight prefabricated structures. The combination of results from the real and simulated buildings, allowed to generalize the results. Finally, some considerations about the RC model parameters physical meaning have been made.

2 Study case

2.1 Laboratory description

The laboratory is property of RSE S.p.A (*Ricerca sul Sistema Energetico*) and is located in Piacenza industrial area. It consists of a prefabricated building oriented on the North-South axis. Experimental rooms are rooms A, B, C and D, with windows facing South. Rooms E and F will be used for other purposes. The net plan area of the considered rooms is 59.2 m^2. Tables 2.1 and 2.2 summarize the stratigraphy of the construction elements and their thermal transmittances. The layers are indicated from the outside in.
The perimeter walls and the roof are made up of polyurethane prefabricated panels with a declared conductivity of 0.021 W/(m K). This value was increased to 0.030 W/(m K) to take account of the ageing process of the material.
The net height of the rooms is 2.70 m, and they are separated from the attic by a false ceiling made by mineral fiber panels.
The laboratory is heated and cooled through an air-water heat pump and 4 fancoils, one for each room.

Figure 2.1: *RSE laboratory in Piacenza.*

LAB					
External wall	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
Polyurethane	0.1	0.03	120	1.00	3.33
Plasterboard	0.025	0.19	660	1.00	0.132

Ground floor	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
PVC floor	0.005	0.21	1300	1.45	0.024
Screed	0.05	1.4	2200	1.05	0.036
XPS insulation	0.05	0.037	40	1.45	1.35
Polyethylene sheet	0.001	0.4	90	1.00	0.003
Light concrete	0.07	0.1	520	1.05	0.7
Concrete base	0.15	2.3	2000	1.00	0.065

Roof	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
Polyurethane	0.1	0.03	120	1.00	3.33

False ceiling	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
Panels	0.1	0.03	120	1.00	3.33

Internal wall	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
Polyurethane	0.1	0.03	120	1.00	3.33

Table 2.1: *RSE laboratory stratigraphy.*

Total transmittances	LAB [W/(m^2 K)]
External wall	0.277
Roof	0.288
Ground floor	0.393

Fixture transmittances	LAB [W/(m^2 K)]
Windows	2.29
Door	1.539

Table 2.2: *Envelope and fixtures U values.*

2.1.1 Laboratory floor plan and views

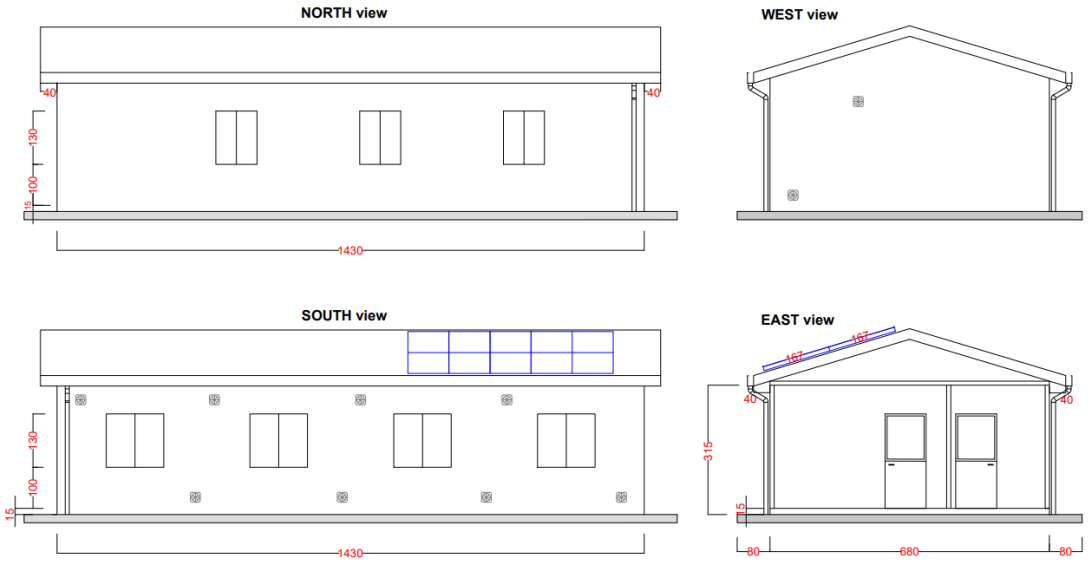

Figure 2.2: *RSE laboratory views.*

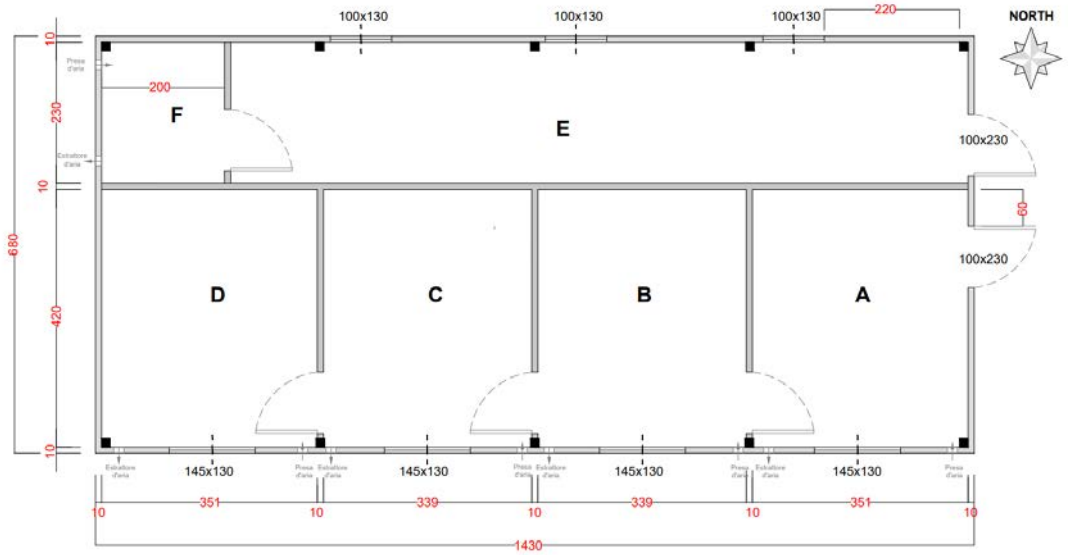

Figure 2.3: *RSE laboratory floor plan.*

2.2 Monitoring and control system

The monitoring and control system that has been realized is aimed to control the main plant components and real time detection of operating parameters and from these, to the evaluation of the energy performance of both the entire plant and individual most significant components. The designed monitoring system is made by field sensors and data related transducers, data collection and acquisition system, a personal computer dedicated to local and remote management of the data acquisition system and a UPS sized to feed the monitoring system. In order to achieve the monitoring objectives, the plant has been equipped with many measure probes to check the parameters of the plant components, the heat pump and the indoor and outdoor environment.
In this study the interest goes to the acquisition of environmental data and the calculation of fancoils power, therefore only the instruments and monitoring system used for our purposes will be explained.
The plant has equipped with:

- electromagnetic induction flow meters for the measurement of water flow rates in various circuits, mounted to insertion in the pipes, with an accuracy of 0.5 % of the bed value (fancoil distribution system) and with an accuracy of 0.8 % for the other circuits;
- platinum resistance thermometers for temperature measurement of fluids, with 1/5 platinum probes;
- 1/3 DIN platinum resistance thermometers for measuring ambient air temperature at center of each of the four premises of the laboratory;
- weather station composed by:
  - a pyranometer for the measurement of global horizontal solar radiation;
  - a pyranometer with shading ring for the measurement of diffuse solar radiation;
  - instruments for the measurement of temperature and humidity outside air.

The fancoils heating power is then given by:

$$P = q \cdot c_p \cdot \Delta T \quad [W] \tag{2.1}$$

where:
- q = water flow [kg/s];
- c_p = water specific heat [J/(kg K)];
- ΔT = temperature difference [K].

Table 2.3 shows the main features of the used instruments and data acquisition modules. For the acquisition of the data collected from the instruments and the control of the plant, has been realized a special software in National Labview environment based on a personal computer and acquisition modules Advantech ADAM 5000 and 4000. The communication architecture uses field buses Modbus ASCII and Modbus RTU over RS485 network. The acquisition program provides both the on-line presentation of measured data with a frequency of 1 minute in tabular and graphic form. At the end of each day the file is closed and stored on the hard disk; another copy is then opened automatically, assigning it an identifying name with the date to which it refers. The system also allows to view the synoptic pages of the devices/plants monitored with real-time indication of the values of the most representative quantities, measured directly or calculated from the data acquired as shown in Figure 2.6.

measure	**instrument**	**instrument model**
water flow	electromagnetic	E&H Promag P50
water temperature	RTD Pt100 4wire 1/5DIN	Italcoppie TRM
solar global radiation	thermopile pyrometer	Kipp&Zonen CMP11
solar diffuse radiation	photodiode sensor	Delta-T Device BF5
outdoor air temperature	Pt100 4wire	Siap+Micros
indoor air temperature	Pt100 4wire, 1/3DIN	Itsensor
data acquisition module	DA&C system	Advantech ADAM 4/5000

Table 2.3: *Features of main instruments.*

Figure 2.4: *Water temperature probes (Pt100) of the fancoils system (left and center) and indoor temperature probe (right).*

Figure 2.5: *Data acquisition modules Adavntech ADAM 5000 and 4000.*

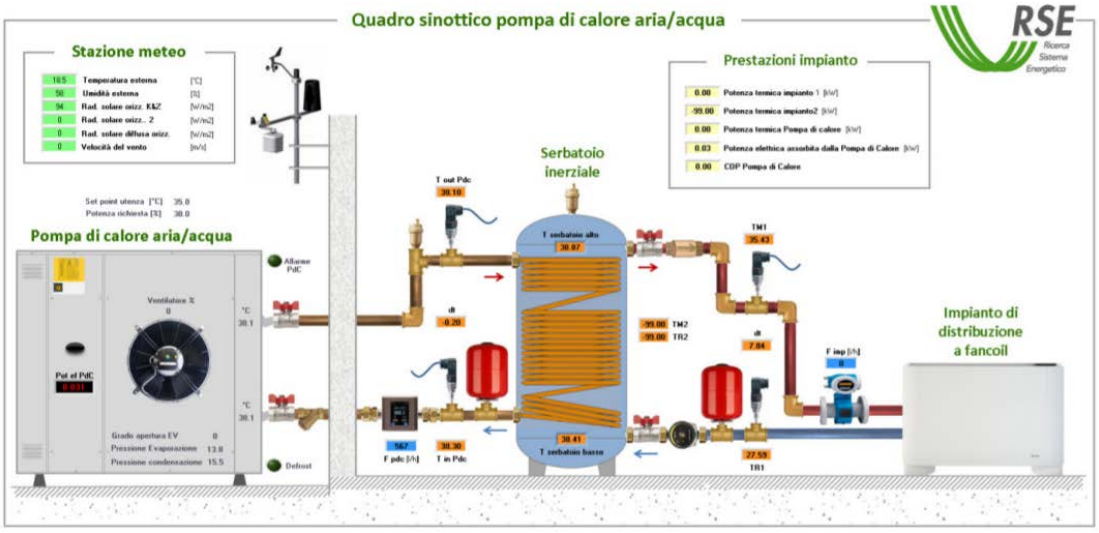

Figure 2.6: *Synopsis panel of the HVAC system.*

3 Models

3.1 Grey box models

The performance of the BEMS control requires a system reliable and a computationally light building model. According to ASHRAE, there are two different ways to model the building: 'forward' and 'data-driven' [7]. Forward models are also known as white box models or detailed dynamic models, which require a high knowledge of building parameters to be implemented, as they are based on purely physical equations. Programs like EnergyPlus and Trnsys are based on this type of approach [8]. These models are not used for control systems, because to be precise they require the manual insertion of a very large amount of data, often incompatible with the information coming from the sensors. This is usually too time consuming and not effective in terms of cost to collect such information [9].
Data driven models assume that there are some mathematical relationships between inputs and outputs but there is no interest in physical equations [10]. They are divided into 'black box' and 'grey box'. The black-box approach is based on statistical methods that are applied to the available data (energy consumption and temperatures) without any explicit reference to the equations that govern their energy balance. In these data-driven models appropriate algorithms are trained to describe the functioning of the considered system through a learning period. This method has been widely used in many areas and allows to learn the optimal functioning of a system based on the measured data, without the need to use equations that describe its physical principles. In this sense the machine 'learns' from the surrounding

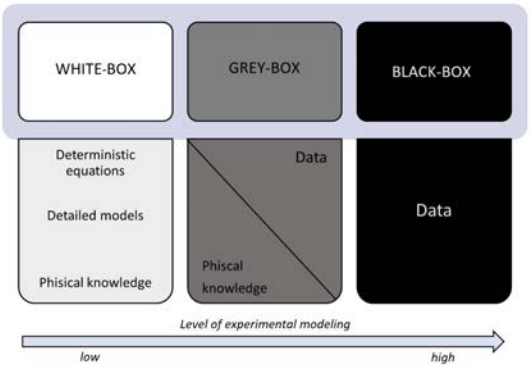

Figure 3.1: *Required level of experimental modeling for different model types.*

environment the best way to behave, this is called machine learning (ML). Grey box can be considered a hybrid between black box and white box, because they are based both on data and physical knowledge.
They consist of simplified physical models whose parameters are initially imposed thanks to the knowledge of physical processes. They are then calibrated during the operation of the systems, learning how to operate through interaction with the environment. These hybrid models therefore retain the typical learning phase of ML, but benefit from an approximate knowledge of the system in which they will operate.

3.1.1 Simplified models

The models most used by this type of approach are resistance and capacity models that exploit the electrical analogy with resistance R and capacitance C to describe the thermal behavior of the building. Capacitance represents the thermal capacity while resistance represents thermal resistance.
The temperature of the air, surfaces and structures are discretized in a set of nodes connected by thermal resistances and capacities that represent the parameters of the model. Their main advantage is that once their parameters are identified, such temperatures can be determined by the resolution of a simple linear system, with a very low cost in computational terms and avoiding convergence problems typical of non-linear systems that use complex numerical schemes. Since each thermal capacitance of the equivalent circuit corresponds to a state variable of the model, the number of thermal capacitances (nodes) leads to the order of the model. Reducing the model order curtails the number of parameters to be identified, thus achieving the objective of describing the building dynamics with a simple model. The values of R and C are estimated based on samples of inputs and outputs by applying an identification algorithm which typically minimizes a norm of either simulation errors or prediction errors [11].
The boundaries on the parameters in the identification process are normally estimated from a rough description of the building geometry and materials.

3.1.2 Lumped capacitance building model

The International Standard ISO 13790 [12], describes the simple hourly method to calculate the building's energy use with one hour time intervals.
The method is based on a RC circuit, made by 5 resistance and 1 capacity (5R1C) as shown in Figure 3.2. The heating and/or cooling need is found by calculating for each timestep, the need for heating or cooling power, Φ_{HC} (positive for heating and negative for cooling), that needs to be supplied to or extracted from, the internal air node, θ_{air}, to maintain a certain temperature. The Standard distributes the heat gains due to internal sources, Φ_{int} and the heat gain due to solar heat sources, Φ_{sol}, between the three temperature nodes of the thermal

network: Φ_i to the indoor air temperature node (θ_i), Φ_s to the surface temperature node (θ_s) and Φ_m to the thermal mass temperature (θ_m) node. The latter temperature represents the average temperature in the building structure. The corresponding thermal capacitance C_m, calculated in accordance with the ISO 13786 Standard [13], is used to reproduce the dynamic thermal behavior of the building. The other building parameters are the ventilation heat transfer coefficient H_{ve}, the coupling conductance between internal air and surface node $H_{tr,is}$, the thermal transmission coefficients of the windows $H_{tr,w}$ and of the opaque building components $H_{tr,op}$. The latter is divided into two components $H_{tr,em}$ and $H_{tr,ms}$. The model is also able to calculate the energy needs due to de-humidification of the indoor environment during the cooling season, by setting the internal relative humidity and performing the hygrothermal balance of the indoor air volume.

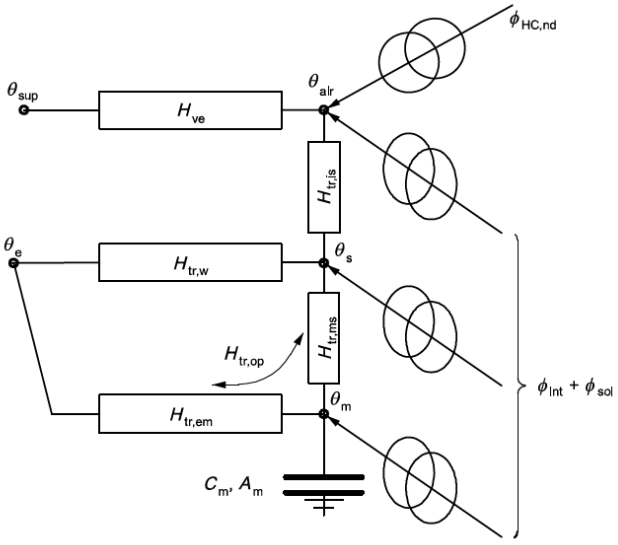

Figure 3.2: *Electric analogy.*

The resulting heating-cooling load Φ_{HC} is determined by solving the following linear system:

$$\boldsymbol{H_{ve}(\theta_{sup}-\theta_i)+H_{tr,is}(\theta_s-\theta_i)+\Phi_i+f_{conv}\Phi_{HC}=0} \tag{3.1}$$

$$\boldsymbol{H_{tr,w}(\theta_e-\theta_s)+H_{tr,is}(\theta_i-\theta_s)+H_{tr,ms}(\theta_m-\theta_s)+\Phi_s+(1-f_{conv})\Phi_{HC}=0} \tag{3.2}$$

$$\boldsymbol{H_{tr,em}(\theta_e-\theta_m^{\tau})+H_{tr,ms}(\theta_s-\theta_m^{\tau})+\Phi_m+\frac{C_m}{\Delta t}(\theta_m^{\tau-\Delta\tau}-\theta_m^{\tau})=0} \tag{3.3}$$

Where θ_s is the air supply temperature due to infiltration and/or ventilation (equal to the external air temperature θ_e in case there is no mechanical ventilation system) and f_{conv} is a parameter used to model different radiative and convective contributions of HVAC terminals (example: $f_{conv} = 1$ for fancoils, $f_{conv} = 0.5$ for radiators).
Another considered modification is the parameter f_{gc} which is a coefficient that multiply Φ_{sol}.
Thermal flows from internal and solar heat sources Φ_{int} and Φ_{sol}, are divided between the air node, θ_{air}, and the internal nodes, θ_i, θ_m, as follows:

$$\Phi_i = 0.5\Phi_{int} \tag{3.4}$$

$$\Phi_m = \frac{A_m}{A_t}(0.5\Phi_{int} + 0.5\Phi_{sol} f_{gc}) \tag{3.5}$$

$$\Phi_s = (1 - \frac{A_m}{A_t} - \frac{H_{tr,w}}{9.1A_t})(0.5\Phi_{int} + 0.5\Phi_{sol} f_{gc}) \tag{3.6}$$

Where A_t is the surface area [m^2] and A_m is the effective mass area [m^2] which represents the thermal capacitance of the building in terms of surface area and calculated as shown in Figure 4.3.
In this study the model does not consider the internal heat gains contribute, so $\Phi_{int} = 0$.
The solar heat gain Φ_{sol} includes here not only the solar radiation entering through external windows but also the short-wave radiation absorbed by the external walls and the long-wave radiation emitted by the external surfaces to the outdoor environment:

$$\Phi_{sol} = F_{so}(a_s R_{se} U_{op}) I_{sol} - F_r \Phi_r \tag{3.7}$$

$$\Phi_r = R_{se} U_{op} A_{op} \alpha_{rad} \Delta\theta_{er} \tag{3.8}$$

The first term is the short-wave radiation absorbed by opaque building components, where F_{so} is the shading reduction factor for the external obstacles, I_{sol} the solar irradiance, U_{op} and A_{op} the thermal transmittance and the projected area of the opaque building components, a_s and R_{se} their absorption coefficient and the surface heat resistance, respectively. All these variables refer to the external surface of the exterior walls and must be considered separately for each orientation.
The second term is the extra heat flow due to the thermal radiation to the sky, where F_r is the form factor between the building element and the sky (0.5 for vertical walls), α_{rad} is the radiative heat transfer coefficient and $\Delta\theta_{er}$ is the difference between the external air temperature and the apparent sky temperature.
The linear system includes four independent variables, the temperatures at the three nodes

(θ_i, θ_s, θ_m) and the heating-cooling load Φ_{HC}. Therefore, there are two ways to solve the system [14]:

- Φ_{HC} fixed: in this case, the heating-cooling load is set and, consequently, the air temperature of the thermal zone is calculated; this is how the RC model is used in this study;
- θ_i fixed: with this option the thermal zone heating-cooling load for the specified setpoint θ_i is calculated.

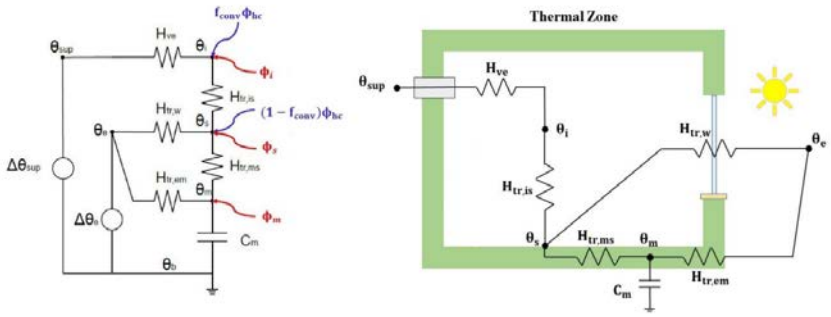

Figure 3.3: *5R1C model: physical scheme and equivalent electrical network.*

3.2 Detailed dynamic building model

For the reasons already explained, the white box models are not the most suitable for the control of a building. However, by modeling the building in EnergyPlus we can provide some preliminary information about is thermal behavior such as the energy consumptions and internal temperature trend. If the simulated behavior is sufficiently similar to the real one, we can use the software to test the calibration also on other types of structures without necessarily that they physically exist and without the need to install expensive sensors, saving money and time.

3.2.1 EnergyPlus

The EnergyPlus program is a collection of many program modules that work together to calculate the energy required for heating and cooling a building using a variety of systems and energy sources. It does this by simulating the building and associated energy systems when they are exposed to different environmental and operating conditions. The core of the simulation is a model of the building that is based on fundamental heat balance principles. EnergyPlus has its roots in both the BLAST and DOE-2 programs. BLAST (Building Loads Analysis and System Thermodynamics) and DOE–2 were both developed and released in the

late 1970s and early 1980s as energy and load simulation tools. Their intended audience is a design engineer or architect that wishes to size appropriate HVAC equipment, develop retrofit studies for life cycling cost analyzes, optimize energy performance, etc. Like its parent programs, EnergyPlus is an energy analysis and thermal load simulation program. Based on a user's description of a building from the perspective of the building's physical make-up, associated mechanical systems, etc., EnergyPlus will calculate the heating and cooling loads necessary to maintain thermal control setpoints, conditions throughout an secondary HVAC system and coil loads, and the energy consumption of primary plant equipment as well as many other simulation details that are necessary to verify that the simulation is performing as the actual building would. Many of the simulation characteristics have been inherited from the legacy programs of BLAST and DOE–2. Below is list of some of the features of the first release of EnergyPlus. While this list is not exhaustive, it is intended to give the reader and idea of the rigor and applicability of EnergyPlus to various simulation situations [15]:

- integrated, simultaneous solution where the building response and the primary and secondary systems are tightly coupled (iteration performed when necessary);
- sub-hourly, user-definable time steps for the interaction between the thermal zones and the environment; variable time steps for interactions between the thermal zones and the HVAC systems (automatically varied to ensure solution stability);
- ASCII text based weather, input, and output files that include hourly or sub-hourly environmental; conditions, and standard and user definable reports, respectively;
- heat balance based solution technique for building thermal loads that allows for simultaneous calculation of radiant and convective effects at both in the interior and exterior surface during each time step;
- transient heat conduction through building elements such as walls, roofs, floors, etc. using conduction transfer functions;
- thermal comfort models based on activity, inside dry bulb, humidity, etc..

This study exploits only a very small part of the potential of EnergyPlus to determine the energy demand of the RSE laboratory under various conditions and to reproduce the indoor temperature trend in some simulated buildings with the aim of developing the BEMS system and testing the calibration.

3.3 Heuristic methods and evolutionary algorithms

Often the determination of the optimal solution of an optimization problem can be too onerous in terms of calculation time, and in a industrial context, linked for example to production processes, control or logistic, the specifications on the execution times of a given algorithm are many times more restrictive than the specifications on the optimality of the solutions found. When dealing with real optimization problems in fact, all the theoretical aspects related to the rigid formulation of the problem, the verification of conditions necessary for the use of exact methods and the related validation of results obtained, are sacrificed in favor of the efficiency with which the solution is calculated; this fact is essentially due to a series of factors [16]:

- many of the parameters in real applications are estimates that may be subject to error, so it is not worth waiting too long to have a solution whose value is of uncertain evaluation; under consideration in order to quickly assess work scenarios;
- often works in real time, so it is necessary to have a good acceptable solution in very short times (minutes or seconds of time calculation);
- sometimes real applications have many constraints of a difficult nature, that is difficult to model with full linear programming models.

All these aspects explain why in real applications, it is very widespread the use of approximated or heuristic algorithms (from the greek *heuriskein* = *discover*), that is algorithms that do not guarantee to obtain the optimal solution, but in general they are able to provide a 'good' acceptable solution for the problem.
Among the most used heuristic algorithms, are mentioned:

- greedy algorithm;
- local research algorithm;
- Simulated Annealing;
- evolutionary algorithm such as PSO (*Particle Swarm Optimization*) or Genetic algorithm.

The evolutionary algorithms are therefore computer techniques inspired by biology that are based on a metaphor, how an individual of a population of organisms should be adapted to the environment to survive and reproduce, so a possible solution must be adapted to solve the problem. The problem is the environment in which a solution lives, within a population of others possible solutions; the solutions differ in quality, that is in cost or merit, which are reflected in the assessment of the objective function, as well as individuals in a population of organisms differ in degree adaptation to the environment.

3.3.1 PSO: Particle Swarm Optimization

Figure 3.4: *Flocks of bird in the sky.*

The collective behavior of groups of animals is a natural phenomenon from always very fascinating. Particularly interesting from the point of view of scientific research is the case of those aggregations of animals that self-organize in precise forms and complex dynamics. Examples of these behaviors are swarms of insects, flocks of birds or schools of fish. Understand from which assumptions emerge a collective behavior, is currently a transversal target and in recent years has fueled interest in those methods that are called PSO algorithms, heuristic techniques of research that belongs to the category of evolutionary algorithms. The basic idea is linked to the concept that each individual (one possible solution in the research set) of a population swarm, goes in search of an excellent solution to a certain optimization problem, with a dynamics affected by memory of one's best position and best location of the global swarm. In addition to an individual behavior what emerges is a collective behavior that lends itself well to the resolution of different optimization problems. Reynolds [17], proposed a behavioral model in which each agent follows three rules:

- separation: each agent tries to move away from its neighbors if they are too close;
- alignment: each agent steers toward of its neighbors;
- cohesion: each agent tries to go towards the average position of its neighbors.

From an algorithmic point of view, these ideas are taken up in the model that describes the individual's dynamics in the simplest and primitive PSO method version.

Given a particle population in a n-dimensional space, looking for a (sub-)optimal solution of a optimization problem, each individual x(t) = [x1(t), x2(t), . . , xn(t)] iteratively moves as follows:

$$\boldsymbol{x(t+1) = x(t) + v(t+1)\Delta t} \tag{3.9}$$

$$\boldsymbol{v(t+1) = w\,v(t) + c_1\rho_p(x_{p,best} - x(t)) + c_2\rho_g(x_{g,best} - x(t))} \tag{3.10}$$

where the vector $\mathbf{v(t) = [v1(t), v2(t), . . , vn(t)]}$ is the coming velocity vector updated linearly also in iterative way, influenced by different factors: it depends on the speed at the previous step $\mathbf{x_{v(t)}}$ (inertia), from best location $\mathbf{x_{p,best}}$ and best overall location $\mathbf{x_{g,best}}$. The latter are updated as soon as find better ones.
The coefficient $\boldsymbol{w}$ weighs the inertia of the system, $\boldsymbol{c_1}$ and $\boldsymbol{c_2}$ weigh the cognitive and social component, while $\boldsymbol{\rho_p}$ and $\boldsymbol{\rho_g}$ are random uniform distribution coefficient U[0, 1] of the two personal and global components respectively.
More specifically, the components to upgrade the velocity are analyzed:

- **Previous speed v(t)** acts as memory of the previous direction of movement (in the immediate past). This term can be seen as a moment that prevents the sudden change of direction of the particle;

- **Cognitive component** $\boldsymbol{\rho_p(x_{p,best} - x(t))}$ quantifies the dynamics of the particle relative to the performances of the past (not immediate). In a sense it represents the individual memory of the particle, and model the tendency of individuals to go back to their own best positions. Kennedy and Eberhart also refer to the term nostalgia of the particle [18];

- **Social component** $\boldsymbol{\rho_g(x_{g,best} - x(t)}$ quantifies the dynamics of particle relative to the performance of nearby particles (or of all swarm particles in general); each particle tends to move also in the direction of the best location of the swarm. In literature there are different variations, always inspired by social behavior observed in nature.

The inertial coefficient w is introduced as a control mechanism of the process of global and local exploration of space. In this sense such parameter adjusts tradeoff search from global to local: high values of inertia feed global research, increasingly lower values facilitate the local one, until completely eliminating the individual's ability to search. For $w = 1$ the velocity increases with each iteration leading to the instability of the system (the swarm diverges); for $w < 1$ the particles decelerate until the speed completely cancels. Given therefore a population of particles, we can find the following analogies between generic operators of evolutionary algorithms and operations of PSO algorithms: each particle represents an

individual, that is a possible solution to an arbitrary problem; the mutation operator can be seen as the process of perturbation of the position of each individual through the stochastic coefficients ρ_p and ρ_g; the recombination operator is instead identified in the way the total behavior of the particle emerges from three components of the speed, always due to the randomness of the parameters.

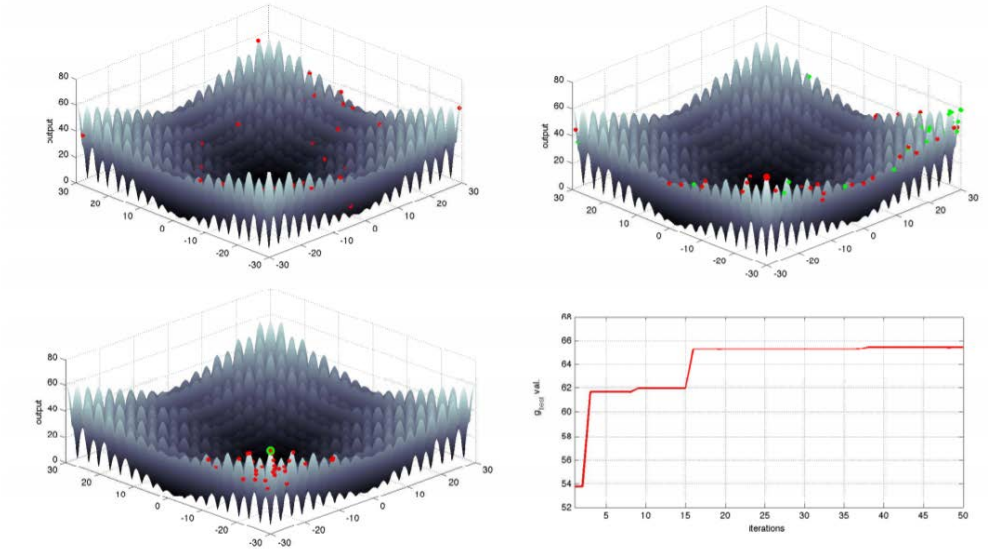

Figure 3.5: *Sequence of images showing the evolution of a 50-particle swarm in search of the maximum of an 'alpine' function, characterized by numerous peaks and minimum locals. It can be noticed in green the best positions personal and global, used by the algorithm to compute the speed vector.*

4 Methodology

4.1 BEMS code description

The building energy management system (BEMS) for the RSE laboratory has been developed in Python with the logic of object programming (OOP). In this way is possible to define classes or modules, associated with activities such as plant control, model calibration and weather forecast acquisition. Within each class, the variables are first initialized according to the parameters assigned by the user. Once initialized, classes are able to perform operations. For example, the calibration class will be able to update the parameters used by the optimizer. In the optimizer class, it will be possible to build the equations that define the optimization problem, the objective function, and finally launch the optimization to find the optimal operation of the heat pump.
The BEMS must coordinate the system and therefore must perform these tasks sequentially.

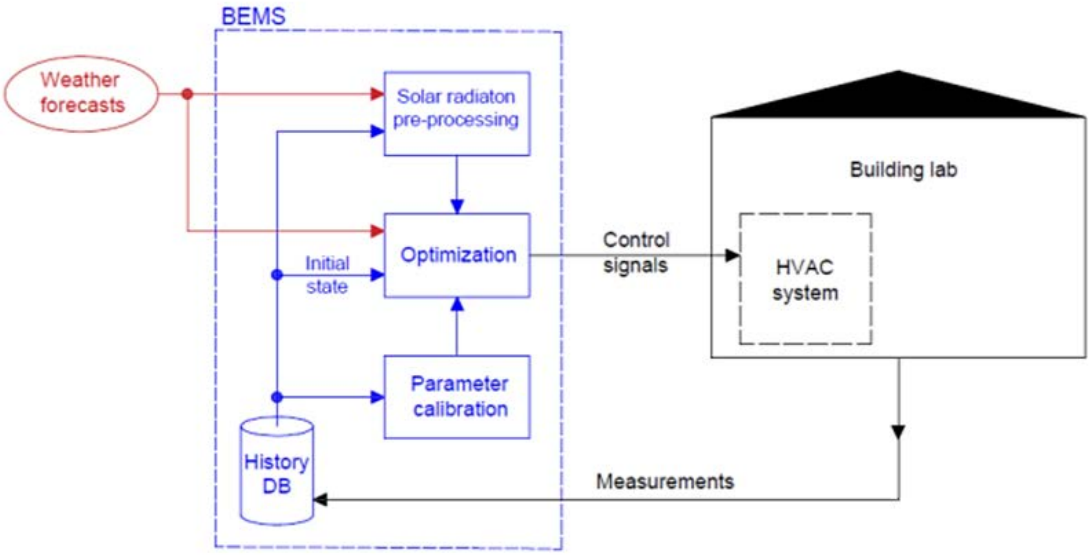

Figure 4.1: *Qualitative block diagram illustrating BEMS architecture.*

Four tasks are then defined which are essential for the operation of BEMS, each of which must be performed with a certain frequency:

- data acquisition from plant, internal and external environment;
- weather forecast and photovoltaics producibility data acquisition;
- calibration of model parameters;
- optimization of the heat pump.

4.1.1 Calibration of model variables

Calibration is the numerical process by which parameters describing the dynamic thermal behavior of the building in Equations 3.1 - 3.3, are initially estimated by an approximate knowledge of the building and then recalculated to ensure that the mathematical model matches as closely as possible to the physical quantities measured.
This operation is carried out by PSO, the optimization algorithm that minimizes the difference between the temperature profile measured in the building/laboratory and the temperature calculated by the RC model, in which the thermal power Φ_{HC} is set.
The objective function of the calibration minimizes the RMSE, which measures the differences between values predicted by a hypothetical model and the observed values. In other words, it measures the quality of the fit between the actual data and the predicted model.
It is one of the most frequently used statistic measures of the goodness of fit for machine learning models:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(x_i - x_{i,meas})^2}{N}} \tag{4.1}$$

where:

i = variable i;
N = number of data points;
x_i = actual observation time series;
$x_{i,meas}$ = estimated times series.

The objective function to minimize during the calibration is therefore:

$$min\sqrt{\frac{\sum_{t=1}^{T}(\theta_t^i(x) - \theta_t^{i,meas})^2}{T}} \tag{4.2}$$

where x is the parameters array to determine shown in Table 4.1, where the variables of the lumped capacitance model are described in 3.1.2.

x[n]	Variable	Unit
0	*C_m*	J/K
1	*H_tr_em*	W/K
2	*H_tr_is*	W/K
3	*H_tr_ms*	W/K
4	*H_tr_w*	W/K
5	*H_ve*	W/K
6	*k_conv*	-
7	*fgc*	-

Table 4.1: *RC model variables.*

PSO application in BEMS

Normally, this type of algorithms is initialized by assuming a random particle distribution within the domain in which the solution is sought. In this application, the parameters of the RC model can then take any value within the domain set by the user. The process is described in Figure 4.2:

- the user inserts in the BEMS, during installation, a series of information and physical characteristics of the building: heated area, perimeter or base lengths of the building divided by orientation, number of heated floors, transmittance of walls (external vertical walls), window transmittance, roof transmittance (attic, air layer and roof cover), floor to floor transmittance and window area divided by orientation, type of structure and type of plant;
- with this initial information, the model parameters are estimated as proposed by the Standard UNI ISO 13790;
- calibration module sets the domain as a range around the nominal parameters, $\pm DW$ % (domain width);
- the PSO algorithm is initialized and run to find parameters that minimize the target function 4.2;
- the optimal parameter set is the new midline of the domain;
- the PSO algorithm is run to find the optimal parameters in the new domain. The last two points of the process are repeated n times, as specified in the calibration

parameters. This approach, known as successive zooming, allows to deal with large-scale problems by first looking for interesting regions within the domain, and then to look for the best within these regions, doing precisely the next zooms. The goal of this approach is to look for solutions that maintain the physical sense of the model making however all the parameters of the model vary, which increases the size of the domain.

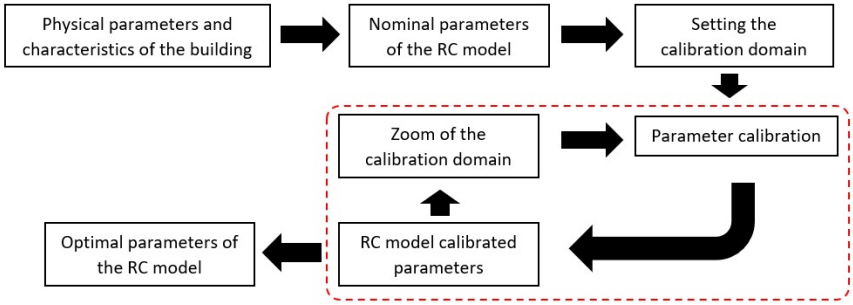

Figure 4.2: *Calibration of parameters logical sequence.*

Once the parameters have been calibrated, the linear system described in Equations 3.1 - 3.3 represents the first block of constraints of an optimization problem to determine the optimal operation of the heat pump. The optimization problem is outside the scope of this study, however being interesting to understand how the calibrated parameters are used in the optimization process, the working principle is described in Appendix A.

4.1.2 Calibration settings and parameters

The calibration function *run_calibration* requires information to be passed to it to initialize the parameters and calculate the optimal ones:

def run_calibration *(building_properties, loc_settings, cali_settings, logs)*

Where *loc_settings, cali_settings, building_properties* are called *dictionaries*. Dictionaries (dict) are a built-in, mutable and unordered type that contains elements (items) formed by a key and a value. Once the dictionary is created and valued with a set of <key, value> pairs, the key (which must be unique) can be used to get the corresponding value.

Building_properties

In *building_properties* are provided information about the building in order to calculate the RC model initial parameters of Table 4.1.
The idea is that the information supplied in building_properties are those that the hypothetical user goes to insert in the BEMS in initial phase through a graphical interface in order to initialize the RC model variables by calculating them as suggested by the UNI ISO 13790.

First of all, information about the envelope are provided, such as opaque and glazed elements transmittances and geometries. The calculation of the thermal capacity C_m is made following the simplified method described by the Standard UNI ISO 12379, as shown in the Standard Table 12 in Figure 4.3.
Then it is necessary to specify the type of terminal used for heating or cooling, being able to choose between *fancoils*, *radiant surfaces* and *radiators* that will determine the corresponding initial value of *k_conv*.

Class[a]	Monthly and seasonal method	Simple hourly method	
	C_m J/K [b]	A_m m²	C_m J/K
Very light	80 000 × A_f	2,5 × A_f	80 000 × A_f
Light	110 000 × A_f	2,5 × A_f	110 000 × A_f
Medium	165 000 × A_f	2,5 × A_f	165 000 × A_f
Heavy	260 000 × A_f	3,0 × A_f	260 000 × A_f
Very heavy	370 000 × A_f	3,5 × A_f	370 000 × A_f

[a] May be specified at national level.

[b] See discussion in Clause G.7 about whether or not a correction is needed for the internal heat capacity for the monthly and seasonal method, to take into account the surface resistance.

Figure 4.3: *C_m parameter calculation UNI ISO 12379.*

building_properties		
A_heated	Piacenza	surface heated [m^2]
perimeter	45.0477	perimeter length [m]
n_floors	1	floors heated
U_walls	0.277	external wall transmittance [W/(m^2 K)]
U_windows	2.29	windows transmittance [W/(m^2 K)]
U_roof	0.288	roof transmittance [W/(m^2 K)]
U_ground	0.393	ground transmittance [W/(m^2 K)]
A_windows	0, 0, 2.3, 0, 7.3, 0, 0, 0	windows surface for orientation [m^2]
ACR	0.2	air change rate [vol/h]
weight	light	type of building
HVAC	fancoils	HVAC type

Table 4.2: *Building_properties definition.*

Cali_settings

In *cali_settings* are set the PSO parameters for the simulation and some preliminary information. The chosen simulation timestep τ is 900 s. This means that the dataset with a minute timestep are resample with a quarter of an hour timestep, which is a good compromise between accuracy and simulation time.
The RC parameter T_{m0} is set equal to the mean of the four locals indoor temperature.
PSO parameters, described in 3.3.1, have been set as shown in Table 4.3, where *Population* represents the swarm particle number [19]. *Maxiter* and *maxloops* are set to make the algorithm doing 5 iteration 6 times. For each loop the domain is upgraded according to the DW thus increasing the search width and making the algorithm more exploring.
The heating season depends on the period of cooling or heating.

cali_settings		
τ	900	timestep of the simulation
training days	variable	days of the training period
testing days	variable	days of the testing period
$\mathbf{T_{m0}}$	mean of locals indoor temperature	parameter for the RC model [°C]
population	100	particles number
c1	1.05	individual coefficient
c2	1	social coefficient
w	0.8	inertial coefficient
DW	0.2,0.5,0.1,0.1,0.5,0.5,0.2,0.5	domain width
maxiter	5	max iterations number
maxloops	6	loop number
season	h	heating season

Table 4.3: *Cali_settings definition.*

Loc_settings and solarProcessor

In *loc_settings* information about the location and the time zone are provided to create datasets and make solar calculations.
The solar load Φ_{sol} contributes have been calculated using the function *solarProcessor* based on the Python library *pvlib*, which exploit the Ineichen and Perez clearsky model [20] to calculate the irradiances (W/m^2) on the opaque and glazed oriented surfaces of the building.

$$sp = solarProcessor(loc\_settings,\ date\_rng,\ bdf\_ext)$$

The model requires longitude, latitude and altitude of the location (*loc_settings*), date and time (*date_rng*) and the global solar horizontal radiation (*bdf_ext*) measured by the weather station[1].

loc_settings		
city	Piacenza	name
lat	45.0477	latitude [°]
lon	9.7004	longitude [°]
alt	45	altitude [m]
tz	Europe/Rome	time-zone

Table 4.4: *Loc_settings definition.*

Logs

The argument *logs* calls the function *read_logs* used to create the dataset:

def read_logs*(logs_folder, ndays, loc_settings)*

Where *logs_folder* opens the folder with saved data in *csv* format and *n_days* is the number of days of the dataset to consider. Every *csv* dataset is made by 8 columns, which report with a minute timestep the values of:

- **date/time**, [yy-mm-dd hh:mm:ss];
- **t_ext**: outdoor temperature [°C];
- **ghi**: global horizontal radiation [W/m^2];
- **ti_a**: thermal zone A indoor temperature [°C];
- **ti_b**: thermal zone B indoor temperature [°C];
- **ti_c**: thermal zone C indoor temperature [°C];
- **ti_d**: thermal zone D indoor temperature [°C];
- **q_fc**: sum of the fancoils power [W].

Data are provided by the weather station and measures from the laboratory instruments and probes as described in 2.2.
q_fc represents Φ_{HC} and is given by the sum of each fancoil thermal power.

[1]For the reader interested in going into greater details about the pvlib library: `https://pvlib-python.readthedocs.io/en/stable`

4.2 Training and testing length

4.2.1 Machine Learning dataset

The grey box models are the most interesting in MPC-oriented models, however these require a high attention in their implementation for various reasons [21].
The first reason concerns the accessibility of data. The problem concerns all phases of data acquisition: sensing, transmitting and storing.
The second challenge of identifying MPC-oriented models is calibrating them to real building operation. The accurate calibration of parameters is an example of Machine Learning. Since the scope is to predict the internal temperature, once is given to the model the nominal starting parameters, is necessary to give a dataset to 'train' the algorithm to match the objective function.
The machine learning procedure is composed by the following steps as shown in Figure 4.4:

- splitting dataset into training and testing;
- using training data to train the model;
- producing the model;
- testing the model and determine the accuracy.

Training data is kind of labeled data set used to train the machine learning algorithm to make it learn from such data sets and increase the accuracy while predating the results.
After using the training data sets, each machine learning model needs to be tested to check the accuracy and validate the model prediction, this is made exploiting the testing dataset.

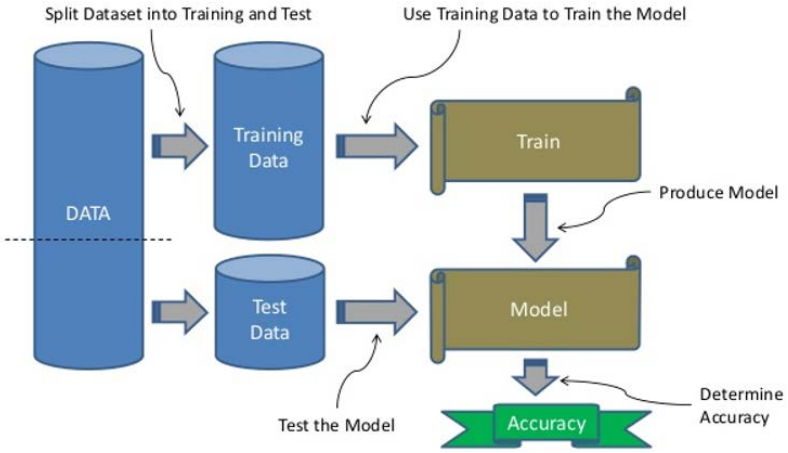

Figure 4.4: *Machine Learning dataset splitting.*

4.2.2 Training and testing length analysis

The scope of the experimental tests is to analyze the calibration accuracy on different training and testing periods. Three dataset from measures have been used to test the calibration in the heating season:

- **Winter 1**: from 09/11/2020 to 17/11/2020;
- **Winter 2**: from 08/12/2020 to 16/12/2020;
- **Winter 3**: from 20/01/2021 to 28/01/2021.

The choice of training days is important to verify the learning of the model, while the duration of testing period is interesting to understand with what frequency it will be appropriate to perform the calibration for practical purposes.
Through the Python code in *cali_settings*, it is possible for each dataset to chose the fraction of data dedicated to train and the fraction dedicated to test the model, starting from the last data of the dataset. So if a dataset is made by 30 days, if 15 days of training and 2 of testing are chosen, only the last 17 days of the dataset will be considered.
The parameter of the function *run_calibration* are set as shown in Table 4.2, Table 4.3 and Table 4.4. The Standard UNI ISO 12379 does not specify how to define the type of building, however being the laboratory a prefabricated structure, it has been chosen to consider it as 'light'. The considerations on the choice of these parameters and their effect in the calibration were subsequently analyzed and described in 5.5.
The calibration is made with a 15 minutes timestep.
Is important to take into account that being PSO an heuristic algorithm, by the nature of the model each run could potentially be different from the previous and the next.
For this reason every test has been performed 10 times through a *for* cycle, then it has been calculated the mean (Mn) and median (Me) RMSE and the standard deviation (σ).
When observing the standard deviation of the results, it was considered that 10 tests were enough to consider the results to be reliable.
Due to a significant number of outliers in some results, which result from convergence to local minimum during model identification, we focus mainly on the analysis of medians instead of means. However, the spread of data are discussed and presented as well in some case.
It has been chosen to test the calibration for every dataset always on 1 and 2 days of testing. This is because the forecast accuracy is optimal in a range of 1 or 2 days, therefore for practical uses is not interesting to extend the testing range.

4.3 EnergyPlus validation

Before proceeding with energy demand calculation an to extend the calibration tests in simulated buildings, it is important to verify the effectiveness of the software used. In fact, for the tests to have a practical validity, it is necessary that EnergyPlus will be actually able to predict the effective indoor temperature trend of the real building with sufficient accuracy. If the difference between the real indoor temperature and the simulated is sufficiently low, then the model can be used to calculate the energy demand and the design thermal power of the heat pump and the calibration code can be tested in various virtual environments of different characteristics and allows to extend the research field to real building typologies.
To validate the model, a period of 7 days has been chosen. That means that the indoor temperature measured from the laboratory is logged in a period of time of 7 days, in order to compare the results with the software simulation.
The best validation period has been find after many trials and finally, the one in which the boundary conditions were best known has been chosen.

4.3.1 Model setting

First of all the building has been drawn in Sketch Up (Figure 4.5) by reference to the geometrical characteristics described in Chapter 2. Here the main structures that divide the various thermal zones of the building are designed.
The thermal model shall consider as a heated zone that consists of rooms A, B, C and D. Rooms E and F shall constitute an adjacent zone not heated.

- Thermal zones A, B, C, D: rooms of RSE laboratory, heated and were measurement probes are placed;
- Thermal zones E, F: adjacent locals; belong to the same structure, but controlled by other people; because the use and the position of the two rooms are similar, they have been considered as a single thermal zone called Thermal zone EF;
- Attic: thermal zone between the roof and the ceiling tiles; not heated.

Once the building has been drawn and the thermal zones are matched, it is possible to upload the model in OpenStudio, which is a software tools to support the building energy modeling in EnergyPlus.
To generate the idf file and to edit it with the EnergyPlus editor is in fact necessary to run before the simulation in Openstudio.

In the period considered the laboratory was heated intermittently through 4 fancoils, arranged one for each thermal zone, controlled through a thermostat set to maintain the temperature between 20 °C ± 2 °C. The power of each fancoil have been inserted in EnergyPlus in the section *Schedule:file*, through which it is possible to create a schedule in *csv* format

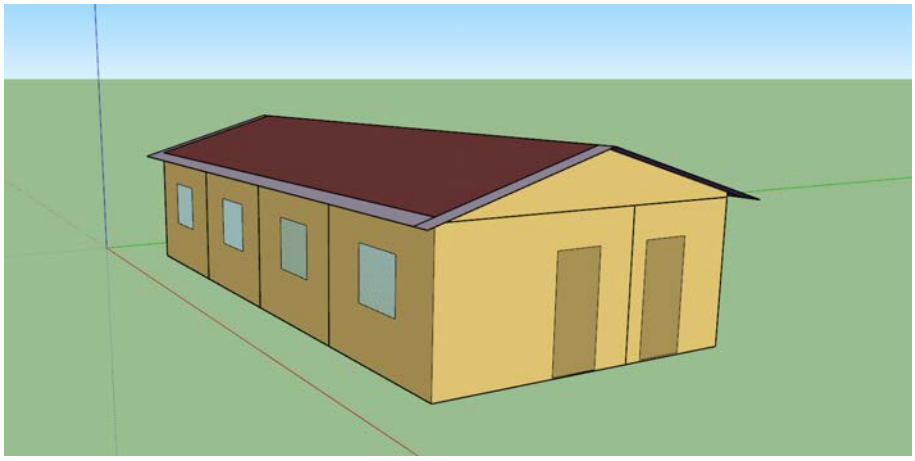

Figure 4.5: *SketchUp model, South-East view.*

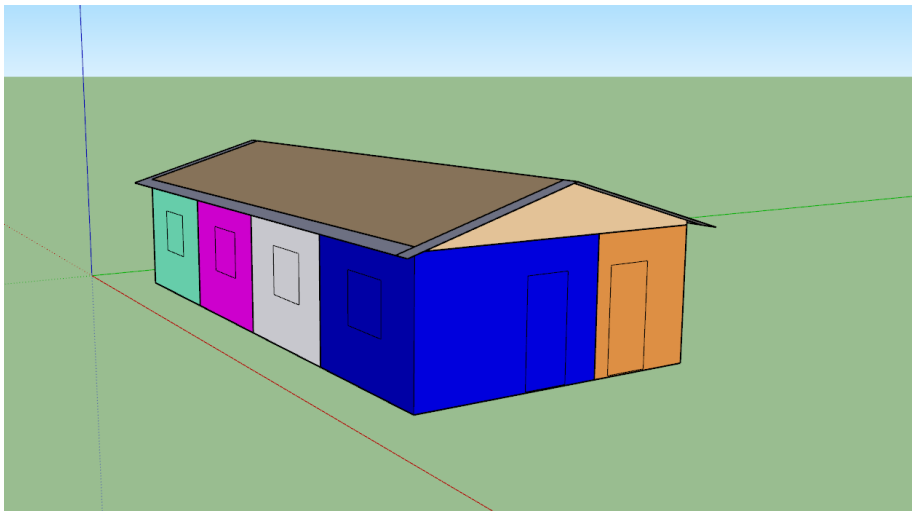

Figure 4.6: *Thermal zones rendering, South-East view.*

with the power values hour by hour, to be assigned to the internal load, defined in the section *Other Equipments*.

To make the test more accurate, so as not to have transients that could have influenced the evaluation, before the start of the validation period a setpoint equal to the actual temperature measured in rooms A, B, C and D has been set. For the adjacent thermal zone EF, on the basis of the information provided by the users, a constant setpoint of 17 °C has been set. These setpoints were then eliminated at the beginning of the validation at 10 am on 21 January to allow the temperature to vary in the rooms.

The logic used in the insertion of the various parameters of the model and the boundary conditions, was as follows: first of all, the parameters relating to the data actually measured or present in the technical reports as weather file, materials data and fancoils power has

inserted; after that, some modifications has made to the model in order to make it as much 'real' as possible.
In the next section are resumed all main parameters settings and choices that have been made to implement the boundary conditions.

4.3.2 Boundary conditions

Boundary conditions are the most important and critical aspect at the same time, because the more are carefully considered and set, the accurate will be the final result.
Let's see the parameters considered for the validation.

- weather conditions: air temperature, solar radiation, ground temperature, air humidity;
- internal heat gains: presence of people or electric equipments;
- thermal envelope: thermal and physical features of the glazing and opaque elements of the envelope;
- shading devices: presence and use of shading curtains.

Weather conditions

Weather conditions are accurately known thanks to the weather station of the lab; the Piacenza epw from EnergyPlus website has been edited thanks to the opensource software Elements[2], with those data that had the greatest influence on the internal temperature, which are:

- external temperature, T_e;
- global horizontal radiation, GH_i;
- external relative humidity, RH_e.

To these is added the radiation diffused on the horizontal, automatically obtained through Elements internal algorithm, who keeps preserved the psychrometrics and solar relationship between the variables.
The temperature of the ground is set in the object *Site:GroundTemperature:BuildingSurface* constant at 18 °C, how suggested by EnergyPlus.

[2] `https://bigladdersoftware.com/projects/elements`

Site Name: Piacenza
Latitude [degrees]: 44.92 Longitude [degrees]: 9.73
Time Zone: 1 Elevation [m]: 134

Header Chart

Tools: Offset Scale Normalize Normalize By Month Variables to Hold Constant:

Date/Time	Dry Bulb Temperature [C]	Wet Bulb Temperature [C]	Atmospheric Pressure [kPa]	Relative Humidity %	Dew Point Temperature [C]	Global Solar [Wh/m2]	Normal Solar [Wh/m2]	Diffuse Solar [Wh/m2]	Wind Speed [m/s]
2005/01/01 @ 00:00:00	0	-0.81	99.73	85	-1.95	1	0	1	0
2005/01/01 @ 01:00:00	0	0	99.73	100	0	0	0	0	0
2005/01/01 @ 02:00:00	0	0	99.73	100	0	0	0	0	0
2005/01/01 @ 03:00:00	0	-0.32	99.73	94	-0.74	0	0	0	0
2005/01/01 @ 04:00:00	0	-0.27	99.73	95	-0.62	0	0	0	0
2005/01/01 @ 05:00:00	0	-0.16	99.73	97	-0.37	0	0	0	0
2005/01/01 @ 06:00:00	0	0	99.73	100	0	0	0	0	0
2005/01/01 @ 07:00:00	0.2	0.2	99.73	100	0.17	0	0	0	0
2005/01/01 @ 08:00:00	0.6	0.6	99.73	100	0.58	22	0	22	0
2005/01/01 @ 09:00:00	1	1	99.73	100	0.98	75.39	2	75	0
2005/01/01 @ 10:00:00	1.1	1.1	99.73	100	1.08	117.12	14	113	0
2005/01/01 @ 11:00:00	1.1	1.1	99.73	100	1.08	145.26	26	136	0
2005/01/01 @ 12:00:00	1	1	99.73	100	0.98	153.23	30	142	0
2005/01/01 @ 13:00:00	1.7	1.7	99.73	100	1.69	140.36	24	132	1
2005/01/01 @ 14:00:00	2.7	2.7	99.73	100	2.7	109.08	11	106	3
2005/01/01 @ 15:00:00	4	4	99.73	100	4.01	65	0	65	5
2005/01/01 @ 16:00:00	4.3	4.3	99.73	100	4.31	13	0	13	5
2005/01/01 @ 17:00:00	4.3	4.3	99.73	100	4.31	0	0	0	4

Columns: Add Remove Move Left Move Right Units: SI IP

Figure 4.7: *Elements tool interface.*

Thermal envelope

Thermal and physical characteristics of the glazed and opaque elements has been obtained by the technical report of the building project. The stratigraphy with physical and thermal characteristics of the materials is reported in Table 4.6. In EnergyPlus is possible to provide additional information about the materials. To make the model more accurate, the solar absorbance value of the vertical external walls and of the roof has been corrected. EnergyPlus provides a default value of 0.7. As the cover is very dark, this value has been increased to 0.8; while the solar absorbance of the vertical external walls has been reduced to 0.2, these being white. In addition, to take into account the presence of the internal support structure consisting mostly of wood and steel beams, the density of the internal wall polyurethane, and of the ceiling panel has been increased respectively to 3000 kg/m^3 and to 2300 kg/m^3.[3]

Shading devices

The various internal contributions and the effective positions of the shading curtains has been decided and recorded by users and technicians of the laboratory. As the curtains remained in such a position to offer almost total window shading, to implement that, the glazed surfaces were assigned a coefficient *shg* equal to 0.1.

[3]This modification is the result of several tests, trying to make the inertia of the model more similar to the real one.

Internal gains and infiltrations

The only internal load in the rooms during the period considered was given by fancoils. The power of each fancoil is set in a *csv* file as a fraction of 1000 W. In this way is possible to set a fractional schedule as required by EnergyPlus and assign it to an hypothetical equipment with a design power of 1000 W. After many tries, it has been chosen appropriate to set a radiant fraction of 35 %.
No household appliance, light bulb or electronic appliance was turned on, just as no person entered and used the laboratory. The infiltrations have been kept low, equal to 0.3 for all the thermal zones as the windows and the doors have always remained closed during the validation.
The run period goes from 01/15 to 01/29, with a hourly timestep, while the validation period goes from 01/21 10:00 to 01/28 14:00. This difference is set to avoid the inaccuracy due to the possible initial transient, as explained in the previous section.

Field	Units	Obj1	Obj2	Obj3	Obj4
Name		Fan_coils_A	Fan_coils_B	Fan_coils_C	Fan_coils_D
Fuel Type		None	None	None	None
Zone or ZoneList Name		A	B	C	D
Schedule Name		Fan_coils	Fan_coils	Fan_coils	Fan_coils
Design Level Calculation Method		EquipmentLevel	EquipmentLevel	EquipmentLevel	EquipmentLevel
Design Level	W	1000	1000	1000	1000
Power per Zone Floor Area	W/m2				
Power per Person	W/person				
Fraction Latent		0	0	0	0
Fraction Radiant		0,35	0,35	0,35	0,35
Fraction Lost					

Figure 4.8: *EnergyPlus fancoils load setting.*

4.4 Laboratory energy demand

Once the model has been created and validated, to calculate the energy needs in heating season and cooling season, the set points and internal loads have been set through the schedules of Openstudio.
All the boundary conditions related to the model and previously seen have remained unchanged except for the internal loads that have been set following the indications of the Standard EN 16798 [25]. The Standard reports the parameters of environmental comfort, lighting and acoustic for design conditions and energy assessment. In the Appendix C (informative) of the norm they come brought back parameters and set-points with the relative profiles of use. The environments considered are: school, kindergarten, department store, meeting room, office, single office, restaurant, residential apartment and single-family residential house. For each type of environment are defined: the time of use, internal loads and set-points. For internal loads, the occupancy density, the thermal loads due to the occupants, the thermal loads due to the equipment, the production of steam and carbon dioxide are specified. As for setpoints, the values for operating temperature, ventilation rate, relative humidity and carbon dioxide concentration are defined. As for the residential sector, the Standard EN 16798 covers two categories: the apartment and the single-family house. In this case it has been chosen to refer to a residential apartment.
Below, the description of the internal loads and setpoints.

Thermostat setpoint

Piacenza is located in italian climatic Zone E, where the heating season lasts form 15 October to 15. The regulation actually provides a constant setpoint of 20 ºC, while the cooling setpoint has been set constant to 26 °C.

Internal load

Internal loads are given by people and electric equipments and they have been set as indicated in Table 4.5. The Figure 4.11 shows the daily building occupation during the week and the equipments use in percentage of the total power.
The sensible fraction in the 67 % of the total while de radiant fraction has ben set equal to 30 % of the total power.
The air changes are set to 0.3 vol/h for all rooms while the solar factor of the windows is set to 0.6. In the validation period this value has been reduced to take into account the actual use of the building during that period when the shading curtains where set to offer almost total window shading. In this case we are interested in an average behavior during the year, for which a higher solar factor is more appropriate.

Internal loads	
People	28.3 m^2/person
People (total)	4.2 W/m^2
People (sensible)	2.8 W/m^2
Equipments	3 W/m^2

Table 4.5: *EN 16798 apartment internal loads.*

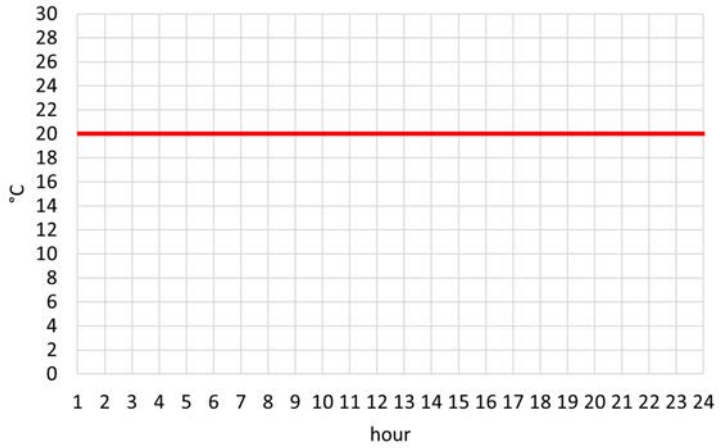

Figure 4.9: *Heating setpoint.*

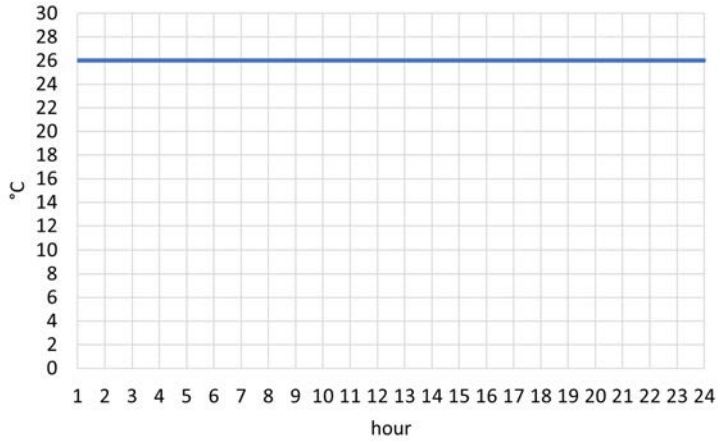

Figure 4.10: *Cooling setpoint.*

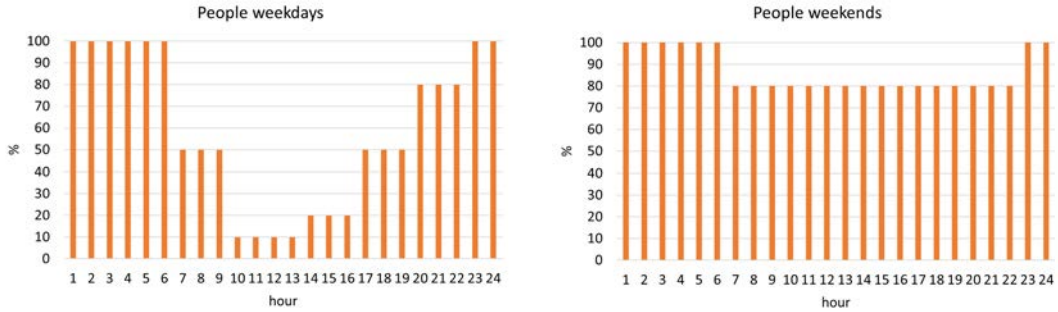

Figure 4.11: *People occupancy time profile.*

Figure 4.12: *Equipments use time profile.*

4.5 Simulated building structures application

After the analysis of the optimal length of training and testing period from the real measures of the laboratory, the aim is to deepen the study to other types of building in order verify the the behavior of the calibration and to extend the results.
Two types of building are considered, lightweight and heavyweight. The first type involves the RSE laboratory as it is and the RSE laboratory with a lower insulation. It has been chosen to run the simulation also for the laboratory in order to create a simulated environment also for the case study and to make some comparison with heavyweight structures with the same boundary conditions.
The heavyweight buildings have been chosen to represent typical situations of a single family houses built in different years and with different transmittance values. Therefore five different building type for the simulations have been set. Each code refers to a different stratigraphy of the envelope:

- lightweight structures:
  - **BLAB** refers to the RSE laboratory;
  - **BLAB_ni** refers to the RSE laboratory, but with a lower insulation;
- heavyweight structures:
  - **B70** refers to a typical single family house in Italy, built between 1976 and 1990;
  - **B90** refers to a typical single family house in Italy, built between 1991 and 2005;
  - **BN** refers to a typical single family house in Italy, built recently; where "N" stands for new.

Tables 4.6, 4.7, 4.8 and 4.9 show the stratigraphy of walls and horizontal surfaces for the buildings used in the simulations. In Table 4.11, the total transmittances of the opaque and glazed elements are shown. The values of the stratigraphy and the transmittances, except for the laboratory, has been chosen from Tabula Web tool [4]. Obviously, the more recent is the building, the lower will be the transmittance of the structure, mainly due to an increase of the insulating layer. The use of the thermal coat allows BN to have an external wall U value lower than 0.28 W/(m^2 K), which is the maximum admitted in climatic zone E in 2021 [26]. In Table 4.10 is also possible to see the fixtures U values. The same door for all the three type of building has been considered.
After validating the model, these stratigraphy are assigned to the same structures of the laboratory previously seen. The building geometry and all the boundary conditions are the same as before, except for the materials of the layers and for the weather. In this case the simulations have been carried out with epw file is Piacenza TRY, from EnergyPlus website.

[4] http://webtool.building-typology.eu/#bm

BLAB_ni					
External wall	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
Board formwork	0.02	0.14	500	2.52	0.143
Plasterboard	0.025	0.19	660	1.00	0.132
Board formwork	0.02	0.14	500	2.52	0.142

Ground floor	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
PVC floor	0.005	0.21	1300	1.45	0.024
Screed	0.05	1.4	2200	1.05	0.036
XPS insulation	0.05	0.037	40	1.45	0.135
Polyethylene sheet	0.001	0.4	90	1.00	0.003
Light concrete	0.07	0.1	520	1.05	0.7
Concrete base	0.15	2.3	2000	1.00	0.065

Roof	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
Polyurethane	0.1	0.03	120	1.00	3.33

False ceiling	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
Panels	0.1	0.03	120	1.00	3.33

Internal wall	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
Polyurethane	0.1	0.03	130	1.00	3.33

Table 4.6: *Simulated laboratory with a lower insulation.*

B70					
External wall	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
External plaster	0.015	0.9	1800	0.91	0.017
Solid brick	0.2	0.608	1800	0.84	0.329
Air gap	0.05				0.151
Hollow brick	0.08	0.36	1200	0.84	0.223
Internal plaster	0.015	0.7	1200	1.01	0.021
Internal wall	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
Internal plaster	0.015	0.7	1200	1.01	0.021
Hollow brick	0.08	0.36	1200	0.84	0.222
Internal plaster	0.015	0.7	1200	1.01	0.021
Ceiling	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
Screed	0.05	0.21	1300	1.45	0.24
Light concrete	0.07	0.1	520	1.050	0.7
Wooden boards	0.02	0.16	500	1.7	0.125
Roof	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
Cement mortar	0.06	0.619	1.200	1.00	0.097
Screed	0.05	0.21	1300	1.45	0.238
Polyethylene sheet	0.002	0.35	950	1.40	0.006
Slab	0.22	0.742	1800	0.920	0.296
Internal plaster	0.015	0.7	1.200	0.9	0.021
Ground floor	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
Gravel	0.2	0.7	1500	0.85	0.286
Bitumen	0.002	0.23	1400	1.00	0.009
Concrete base	0.15	2.3	2000	1.050	0.065
Light concrete	0.07	0.1	520	1.050	0.7
Cement mortar	0.06	0.619	1.200	1.00	0.097
Ceramic tile	0.015	1.2	2300	1.00	0.013

Table 4.7: *1976 to 1990 building stratigraphy.*

B90					
External wall	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
External plaster	0.015	0.9	1800	0.91	0.017
Solid brick	0.2	0.608	1800	0.84	0.329
External insulation	0.04	0.053	20	1.45	0.755
Hollow brick	0.08	0.36	1200	0.84	0.222
Internal plaster	0.015	0.7	1200	1.01	0.021
Internal wall	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
External plaster	0.015	0.9	1800	0.91	0.016
Hollow brick	0.08	0.36	1200	0.84	0.222
Internal plaster	0.015	0.7	1200	1.01	0.021
Ceiling	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
Screed	0.05	0.21	1300	1.45	0.24
Light concrete	0.07	0.1	520	1.050	0.7
Wooden boards	0.02	0.16	500	1.7	0.125
Roof	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
Cement mortar	0.06	0.619	1.200	1.00	0.097
Screed	0.05	0.21	1300	1.45	0.238
XPS insulation	0.03	0.037	40	1.450	0.81
Polyethylene sheet	0.002	0.35	950	1.40	0.006
Slab	0.22	0.742	1800	0.920	0.296
Internal plaster	0.015	0.7	1.200	0.9	0.021
Ground floor	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
Gravel	0.2	0.7	1500	0.85	0.286
Bitumen	0.002	0.23	1400	1.00	0.009
Concrete base	0.15	2.3	2000	1.050	0.65
Light concrete	0.07	0.1	520	1.050	0.7
XPS insulation	0.01	0.037	40	1.450	0.27
Polyethylene sheet	0.002	0.35	950	1.40	0.006
Cement mortar	0.06	0.619	1.200	1.00	0.097
Ceramic tile	0.015	1.2	2300	1.00	0.013

Table 4.8: *1991 to 2005 building stratigraphy.*

BN					
External wall	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
External plaster	0.015	0.9	1800	0.91	0.017
External insulation	0.16	0.053	20	1.45	3.00
Hollow brick	0.16	0.36	1200	0.84	0.444
Internal plaster	0.015	0.7	1200	1.01	0.021

Internal wall	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
External plaster	0.02	0.9	1800	0.91	0.022
Hollow brick	0.08	0.36	1200	0.84	0.222
Internal plaster	0.02	0.7	1200	1.01	0.029

Ceiling	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
Screed	0.05	0.21	1300	1.45	0.238
Light concrete	0.07	0.1	520	1.050	0.7
Wooden boards	0.02	0.16	500	1.7	0.125

Roof	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
Cement mortar	0.06	0.619	1.200	1.00	0.097
Screed	0.05	0.21	1300	1.45	0.238
XPS insulation	0.09	0.037	40	1.050	1.351
Polyethylene sheet	0.002	0.35	950	1.40	2.43
Slab	0.22	0.742	1800	0.920	0.296
Internal plaster	0.015	0.7	1.200	0.9	0.021

Ground floor	Thickness [m]	Conductivity [W/(m K)]	Density [kg/m^3]	Specific heat [kJ/(kg K)]	Resistance [(m^2 K)/W]
Gravel	0.2	0.7	1500	0.85	0.286
Bitumen	0.002	0.23	1400	1.00	0.009
Concrete base	0.15	2.3	2000	1.050	0.065
Light concrete	0.07	0.1	520	1.050	0.7
XPS insulation	0.08	0.037	40	1.050	2.16
Polyethylene sheet	0.002	0.35	950	1.40	0.006
Cement mortar	0.06	0.619	1.200	1.00	0.097
Ceramic tile	0.015	1.2	2300	1.00	0.013

Table 4.9: *Recent building stratigraphy (from 2006).*

heavyweight structures			
Total transmittances	B70 [W/(m^2 K)]	B90 [W/(m^2 K)]	BN [W/(m^2 K)]
External wall	1.175	0.687	0.277
Roof	1.291	0.631	0.312
Ground floor	0.751	0.622	0.286

Fixture transmittances	B70 [W/(m^2 K)]	B90 [W/(m^2 K)]	BN [W/(m^2 K)]
Windows	2.29	2.29	1.530
Door	1.539	1.539	1.539

Table 4.10: *Envelope and fixtures U values.*

Lightweight structures		
Total transmittances	BLAB [W/(m^2 K)]	BLAB_ni [W/(m^2 K)]
External wall	0.277	1.432
Roof	0.288	0.288
Ground floor	0.393	0.393

Fixtures transmittances	BLAB [W/(m^2 K)]	BLAB_ni [W/(m^2 K)]
Windows	2.29	5.67
Door	1.539	1.539

Table 4.11: *Envelope and fixtures U values.*

There are 4 types of simulated buildings for which the calibration has always been tested in the same period. The chosen period goes from **23-01 to 31-01** of the Test Reference Year. The data useful for the training of the model have been extracted from the epw file with **1 hour timestep** and uploaded to Python in order to be read by the calibration code. The data, as previously seen, concern the solar radiation, external temperature, indoor temperature of the thermal zones and the power provided by fancoils to environments.
The fancoils heating power is extrapolated from the data logs of the laboratory measurements in that period. In this way the internal temperature of the rooms assumes a behavior similar to the real one.

4.6 RC model physical parameters

In 3.1 have been described the advantages of using a grey box model in practical applications. These models, thanks to the data for machine learning, allow to reach a discreet precision despite a brief knowledge of the physical and thermal characteristics of the structures. What makes these models interesting and differentiates them from black boxes and purely statistical models is that they retain the physical sense of optimized parameters that minimize the objective function, reproducing the dynamic behavior of the building with a smaller number of data.
The code has been developed to make the user insert the nominal parameters of the building into an hypothetical graphical interface of a computer. However, physical parameters, correct transmittances, thermal capacities and solar factors are not always easy to calculate with high precision. The purpose of the investigation is therefore to determine whether the accuracy of the initial parameters influences calibration.
For these tests the initial parameters were set in 3 different ways and for each set 100 tests were carried out (run):

- **NOM**: the initial parameters are set equal to the nominal parameters (parameters calculated on the basis of the real transmittances and capacity of the building) as in the calibration tests;
- **RNDDW**: the initial parameters are set at each run randomly within a range defined by the previously defined DW; in this way the physical sense and the order of magnitude of the parameters is conserved;
- **RND100**: the initial parameters are set completely randomly by multiplying the values by a random number between 0 and 100; in this way the order of magnitude and physical sense of the parameters are deliberately lost.

All the simulation have been carried out with 4 days training and 1 day testing on the BN structure, which has shown the best calibration results and stability.
The average RMSE over the training and testing period and the standard deviation were calculated for each test set. By adimensioning the parameters, it was then possible to calculate the RMSE_parameters, which indicates for each run how much the optimized parameters differ from the nominals.

5 Results

5.1 Training and testing length

Winter 1, 1-21 training days, 1,2 testing days

First of all the dataset Winter 1 has been considered.
The survey started by analyzing for 1 and 2 days of testing, the optimal duration of the training period up to 21 days of training.
As expected, Table 5.1 and Figure 5.1 show that lower RMSE values are found in training periods. The trend of RMSE in the training period is increasing with the training days. The RMSE for 1 day testing is almost always lower than for 2 days testing. The trend is similar for both periods of testing, that is after high initial values the RMSE seems to lower and then return to rise as the training rounds increase. This result is consistent with [21] and therefore suggests to continue the tests focusing on shorter periods, limiting the training period to 1 week.
The standard deviation is also higher for 2 days testing, and lower for the training results.

1 day testing								
ndTR n.	**MeTR °C**	**MeTE °C**	**MnTR °C**	**MnTE °C**	**σTR °C**	**σTE °C**	**TeTR °C**	**TeTE °C**
1	0.64	0.8	0.63	0.78	0.01	0.09	11.08	8.2
2	0.65	0.74	0.65	0.72	0.02	0.1	11.27	8.2
3	0.65	0.62	0.65	0.64	0.02	0.07	11.29	8.2
5	0.68	0.71	0.68	0.74	0.03	0.08	11.2	8.2
7	0.66	0.72	0.67	0.74	0.03	0.13	10.82	8.2
10	0.7	0.8	0.7	0.79	0.01	0.08	10.85	8.2
14	0.76	1	0.76	0.97	0.02	0.13	11.42	8.2
21	0.87	0.93	0.88	0.94	0.02	0.18	11.72	8.2

2 days testing								
ndTR n.	**MeTR °C**	**MeTE °C**	**MnTR °C**	**MnTE °C**	**σTR °C**	**σTE °C**	**TeTR °C**	**TeTE °C**
1	0.64	0.82	0.64	0.8	0.01	0.13	11.47	9.64
2	0.65	0.69	0.65	0.7	0.02	0.06	11.39	9.64
3	0.63	0.6	0.64	0.61	0.03	0.04	11.21	9.64
5	0.66	0.74	0.66	0.73	0.01	0.07	10.78	9.64
7	0.66	0.81	0.66	0.83	0.01	0.11	10.99	9.64
10	0.7	0.99	0.7	0.98	0.02	0.15	10.98	9.64
14	0.87	0.71	0.88	0.79	0.02	0.14	11.36	9.64
21	0.83	0.94	0.84	0.95	0.02	0.14	11.84	9.64

Table 5.1: *Winter 1, 1-21 training, 1,2 testing*

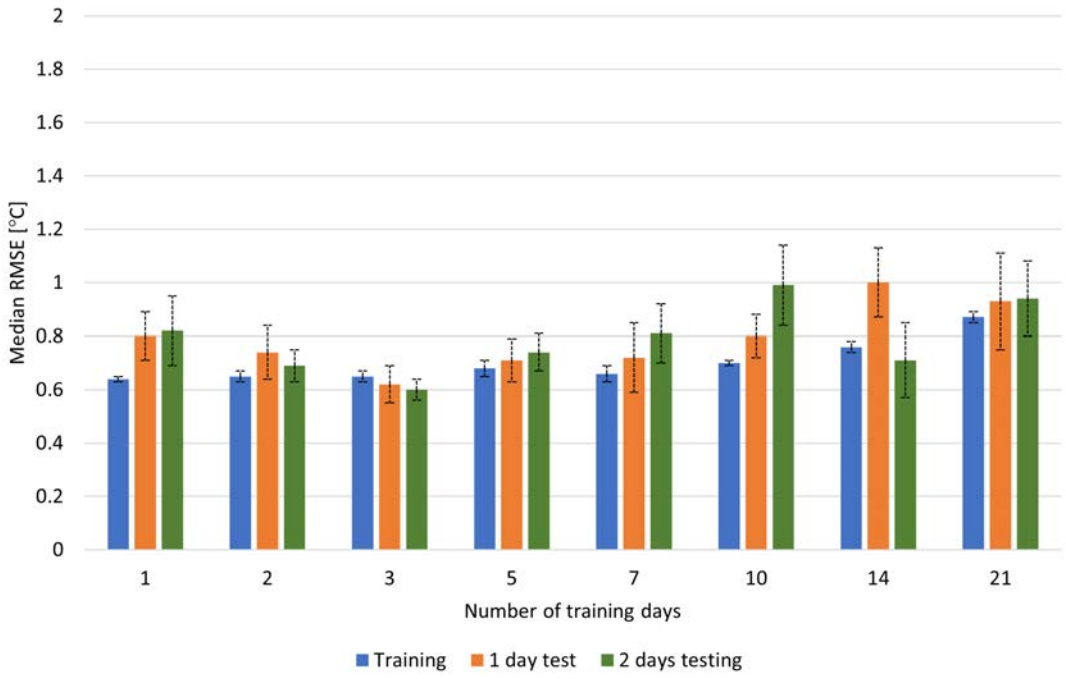

Figure 5.1: *Winter 1, 1-21 training, 1,2 testing.*

Winter 1, 1-7 training days, 1,2 testing days

Always considering the dataset Winter 1, Table 5.2 and Figure 5.2 show how the model behaves by testing it for 1, 2, 3, 4, 5, 6 and 7 days of training and 1 and 2 of testing.
The median values of RMSE have decreased compared to the previous case. This suggests that it makes no sense to push the training beyond 7 days. The results at 1 test day seem to be more interesting than those at 2 days. The training trend is always increasing with the days, while the test has always the tendency to lower to a minimum and then grow with oscillations.
There is a certain recurrence in trends; it seems that from 1 day of training the RMSE value is lowered to a minimum between 2 and 4 days, and then rises with oscillations. As before the standard deviation is lower in the training periods and higher values are given in testing results with 1 day of training.

1 day testing								
ndTR n.	**MeTR °C**	**MeTE °C**	**MnTR °C**	**MnTE °C**	**σTR °C**	**σTE °C**	**TeTR °C**	**TeTE °C**
1	0.63	0.77	0.63	0.77	0.01	0.12	11.08	8.2
2	0.65	0.64	0.65	0.67	0.02	0.09	11.27	8.2
3	0.65	0.63	0.65	0.68	0.01	0.08	11.29	8.2
4	0.64	0.57	0.65	0.59	0.01	0.04	11.18	8.2
5	0.67	0.77	0.67	0.77	0.02	0.09	11.2	8.2
6	0.65	0.69	0.66	0.73	0.03	0.08	10.83	8.2
7	0.68	0.66	0.7	0.68	0.04	0.11	10.82	8.2

2 days testing								
ndTR n.	**MeTR °C**	**MeTE °C**	**MnTR °C**	**MnTE °C**	**σTR °C**	**σTE °C**	**TeTR °C**	**TeTE °C**
1	0.65	0.82	0.65	0.82	0.01	0.09	11.47	9.64
2	0.63	0.63	0.64	0.66	0.01	0.07	11.39	9.64
3	0.65	0.63	0.65	0.63	0.02	0.03	11.21	9.64
4	0.66	0.81	0.66	0.81	0.02	0.05	11.23	9.64
5	0.66	0.84	0.66	0.81	0.02	0.08	10.78	9.64
6	0.68	0.7	0.68	0.72	0.03	0.08	10.78	9.64
7	0.66	0.76	0.66	0.8	0.02	0.15	10.99	9.64

Table 5.2: *Winter 1, 1-7 training, 1,2 testing.*

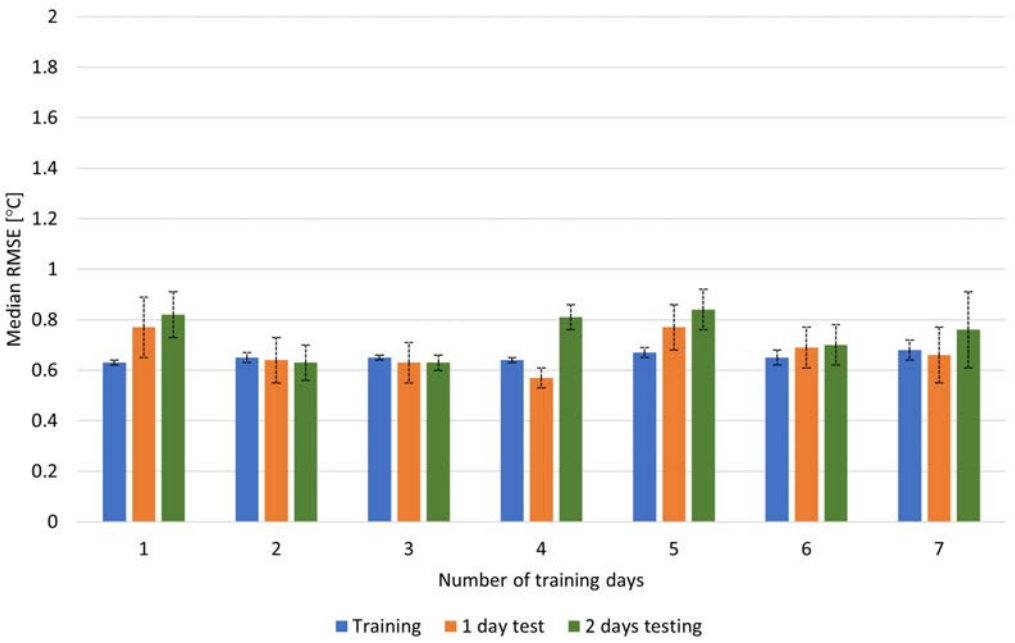

Figure 5.2: *Winter 1, 1-7 training, 1,2 testing*

Winter 2, 1-7 training days

Table 5.3 and Figure 5.3 show that for this dataset the RMSE values are quite higher than before, this is maybe due to an higher variability of weather conditions or operating conditions, probably due the presence of people in the building. However the trends seems to be the same as seen before. In fact, as before, the RMSE values are lower for one day testing and the curve has a minimum between 2 and for days of training.

1 day testing								
ndTR	**MeTR**	**MeTE**	**MnTR**	**MnTE**	**σTR**	**σTE**	**TeTR**	**TeTE**
n.	**°C**	**°C**	**°C**	**°C**	**°C**	**°C**	**°C**	**°C**
1	1.08	1.22	1.09	1.21	0.03	0.1	3.46	6.5
2	0.91	0.96	0.92	0.96	0.03	0.04	2.05	6.5
3	0.96	0.99	0.96	1	0.03	0.05	2.11	6.5
4	0.94	0.97	0.94	0.98	0.02	0.03	2.67	6.5
5	0.97	1.09	0.97	1.09	0.04	0.1	2.84	6.5
6	1.16	1.47	1.16	1.41	0.03	0.16	3.37	6.5
7	1.18	1.55	1.19	1.59	0.03	0.14	3.57	6.5

2 days testing								
ndTR	**MeTR**	**MeTE**	**MnTR**	**MnTE**	**σTR**	**σTE**	**TeTR**	**TeTE**
n.	**°C**	**°C**	**°C**	**°C**	**°C**	**°C**	**°C**	**°C**
1	0.78	1.09	0.78	1.09	0.02	0.13	0.65	4.98
2	0.86	1.03	0.86	1.05	0.02	0.06	1.43	4.98
3	0.9	1.11	0.9	1.12	0.03	0.05	2.4	4.98
4	0.92	1.25	0.93	1.23	0.03	0.12	2.68	4.98
5	1.09	1.63	1.1	1.68	0.05	0.18	3.36	4.98
6	1.13	1.75	1.13	1.75	0.02	0.16	3.59	4.98
7	1.17	2.27	1.17	2.21	0.02	0.15	3.65	4.98

Table 5.3: *Winter 2, 1-7 training, 1,2 testing.*

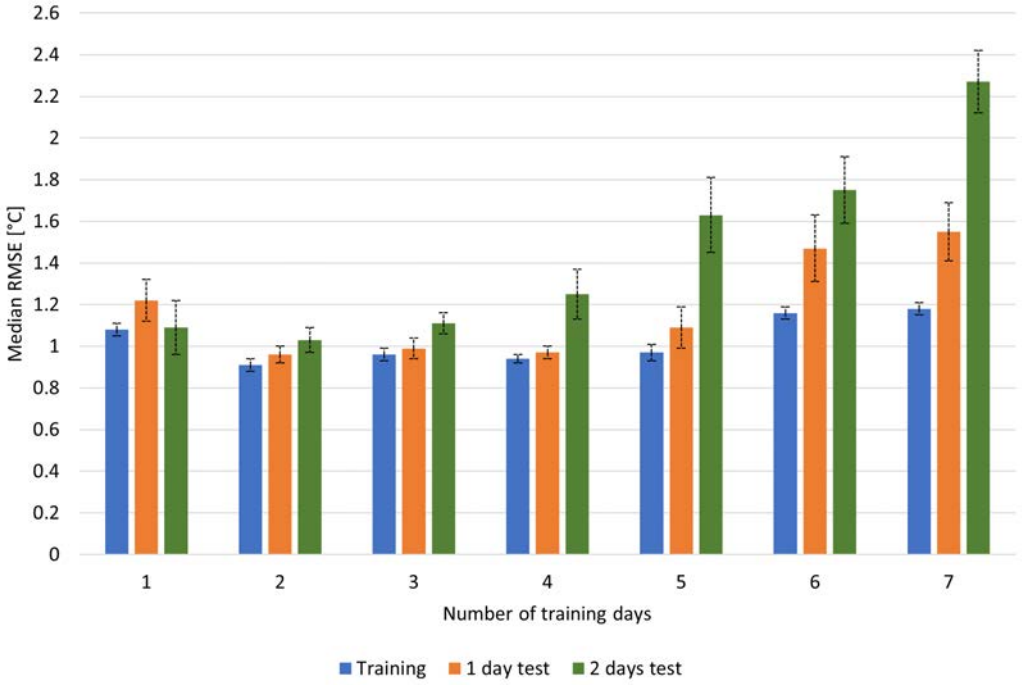

Figure 5.3: *Winter 2, 1-7 training, 1,2 testing.*

Winter 3, 1-7 training days

This datset is the earlier available from the laboratory measures.
Table 5.4 and Figure 5.4 show the lower RMSE values of the datasets until now. The prediction is in this case very accurate, with a minimum testing RMSE of 0.5 °C with 1 day testing and 4 of training. The RMSE value as before are lower with one day testing, and also the trend is quite similar.
Also in this case there is a high RMSE for 1 day of training, then the value stay low until 4-5 days and then starts increasing.

1 day testing								
ndTR n.	**MeTR °C**	**MeTE °C**	**MnTR °C**	**MnTE °C**	**σTR °C**	**σTE °C**	**TeTR °C**	**TeTE °C**
1	0.48	0.61	0.48	0.5	0.02	0.03	2.65	3.16
2	0.58	0.6	0.58	0.87	0.03	0.16	3.69	3.16
3	0.6	0.6	0.61	0.6	0.03	0.06	3.59	3.16
4	0.61	0.51	0.61	0.51	0.02	0.02	3.79	3.16
5	0.59	0.5	0.6	0.51	0.03	0.03	4.1	3.16
6	0.61	0.62	0.61	0.62	0.02	0.04	3.91	3.16
7	0.75	0.9	0.75	0.9	0.02	0.04	3.66	3.16

2 days testing								
ndTR n.	**MeTR °C**	**MeTE °C**	**MnTR °C**	**MnTE °C**	**σTR °C**	**σTE °C**	**TeTR °C**	**TeTE °C**
1	0.5	0.73	0.5	0.85	0.03	0.24	4.73	2.91
2	0.59	0.61	0.59	0.62	0.03	0.06	4.06	2.91
3	0.58	0.7	0.58	0.73	0.02	0.12	4.17	2.91
4	0.56	0.86	0.56	0.85	0.01	0.08	4.47	2.91
5	0.56	0.94	0.57	0.93	0.02	0.13	4.16	2.91
6	0.7	1.08	0.7	1.1	0.01	0.08	3.83	2.91
7	1.18	1.81	1.18	1.81	0.01	0.07	3.43	2.91

Table 5.4: *Winter 3, 1-7 training, 1,2 testing.*

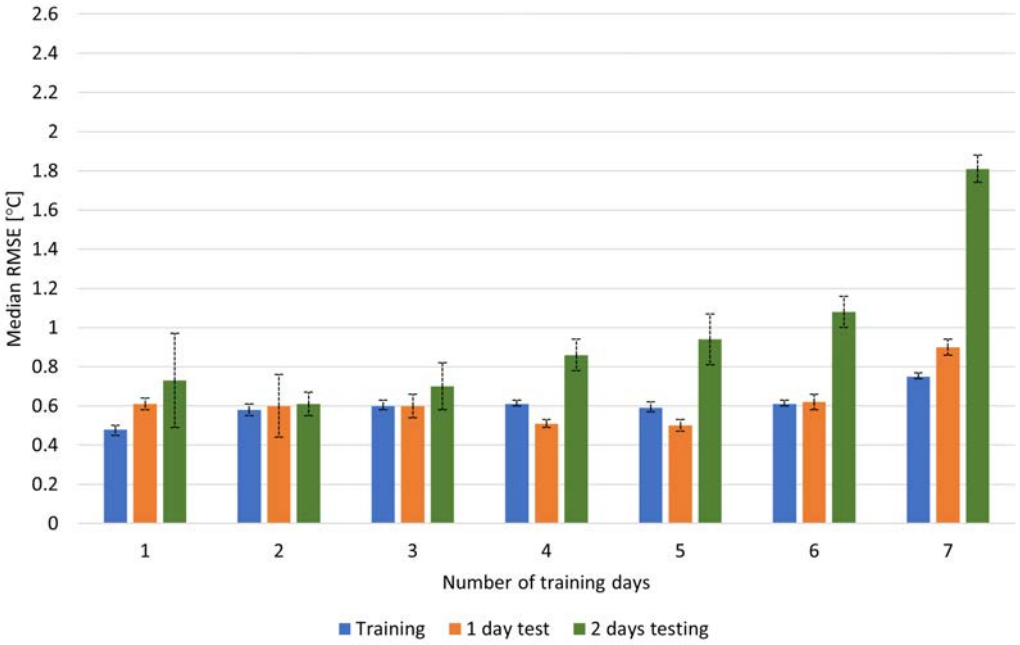

Figure 5.4: *Winter 3, 1-7 training, 1,2 testing.*

5.1.1 Interpretation

The purpose of the survey was to identify the best training and testing periods to minimize RMSE and achieve a better calibration process.
Analysis of the results shows that:

- the best RMSE are reached in training periods; this is an expected result that demonstrates a proper basic operation of the algorithm. The model is trained on the basis of training period data and for this reason it is expected that in that period the RMSE is better;
- training period RMSE increases with training length; this is because with each additional training day, the parameters are estimated for an average day instead of any one particular day;
- compare to 1 and 2 days, the best RMSE results are obtained with 1 day of testing; on average it was more convenient to have 1 day of testing for the algorithm;
- the trend of RMSE in the testing period seems to be similar for 1 and 2 days; the trend generally seems to be a curve that from 1 training day drops to reach a minimum value between 2 and 4 days and then return to rise.

In conclusion it can be said that for practical purposes, it seems to be convenient to carry out the calibration daily. In fact, for 1 day of testing, you get better results in terms of RMSE.
The training of the predictive algorithm should instead be done for a period of time between 2 and 4/5 days.
It is difficult in this case to decree a precise number of days, because it strongly depends on the variability of the boundary conditions in that given period.
Training the algorithm with too many days is likely to take into account even days with average temperatures and average solar radiation very different from the testing period and this is because the training period RMSE increases with training length.
This also proves that the algorithm works well if the boundary conditions in the training period are similar to those in the test period. This applies both to weather conditions and to operations inside the building.
Predictive models typically need a high number of data and the higher the number of data the greater the accuracy. This is true within certain limits due to ‘overfitting’.
Overfitting refers to a model that models the training data too well. It happens when a model learns the detail and noise in the training data to the extent that it negatively impacts the performance of the model on new data. This means that the noise or random fluctuations in the training data is picked up and learned as concepts by the model.
Overfitting is more likely with nonparametric and nonlinear models that have more flexibility when learning a target function.

In a model of this kind training the algorithm with too many days or with days in periods of change season where the conditions are very variable from day to day can cause a loss of accuracy.
It has generally been seen as the best results have been obtained for dataset Winter 3. This is reasonable because in this period there was no people in the building, no internal loads have been activated, the setpoint of the building has been set at a constant interval and the weather conditions have been repeated with more regularity in the periods. Obviously, the more is the control over the boundary conditions, the more the prediction will be accurate. Also opening the windows randomly, activating an internal electrical load or even just the presence of an extra person inside the building can create an 'anomaly' or imbalance in the learning process.
To show this, it is possible to compare the indoor temperature trends for two different datasets. The Figure 5.5 shows the mean indoor temperature in training and testing for the dataset Winter 1, while the Figure 5.7 shows the indoor temperature for the dataset Winter 3.
In Winter 1 with 10 days of training and 2 of testing the median RMSE is equal to 0.99 °C as shown in Figure 5.6 and the accuracy is quite bad, due to some noisy data and to a different thermostat setting for a different handling of the building during the periods.
In the second case, as shown in Figure 5.8, the average trend of the indoor temperature is very similar in the two periods, due to similar boundary condition, in fact in this case the RMSE is equal to 0.51 °C.

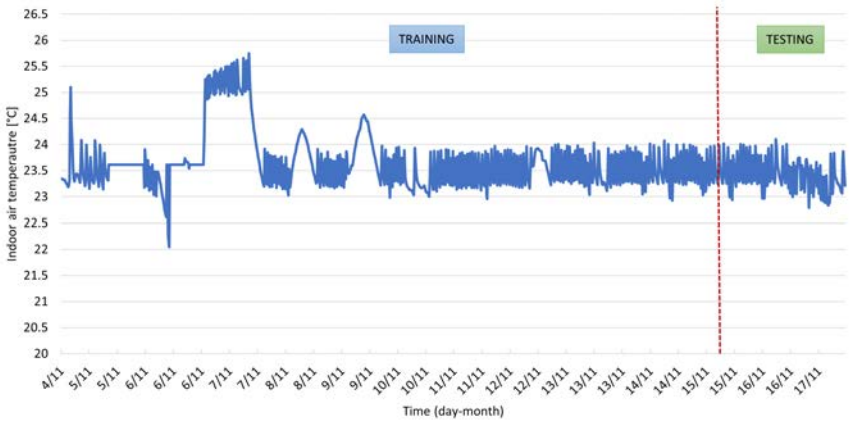

Figure 5.5: *Mean indoor temperature in Winter 1 in training (10 days) and testing period (2 days).*

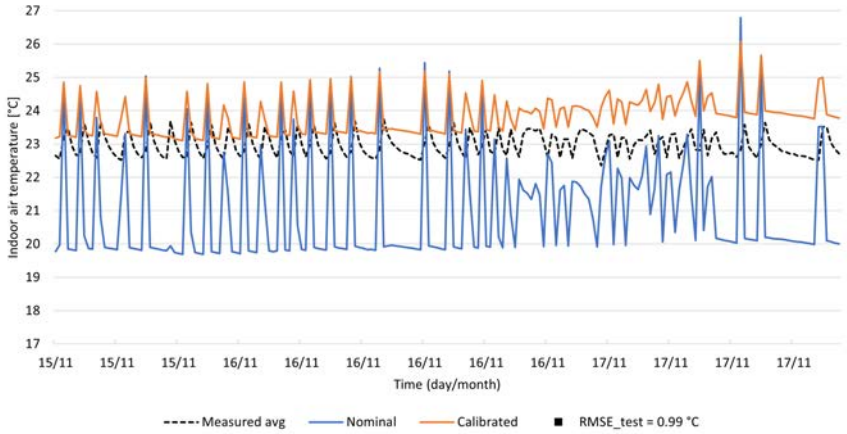

Figure 5.6: *Calibration accuracy in Winter 1 testing period (2 days).*

Figure 5.7: *Mean indoor temperature in Winter 3 in training (4 days) and testing period (1 day).*

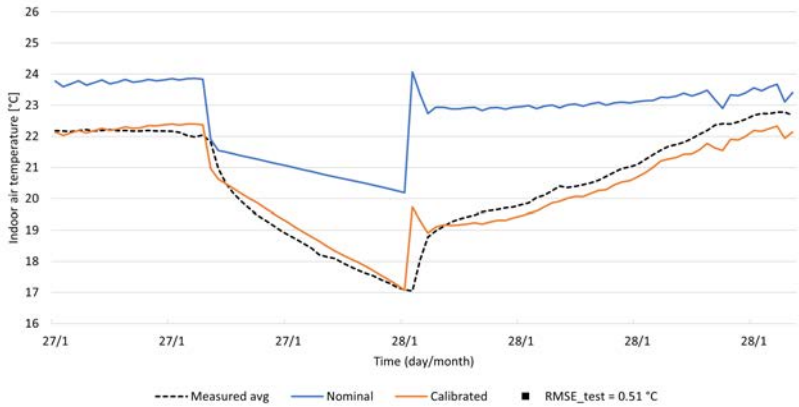

Figure 5.8: *Calibration accuracy in Winter 3 testing period (1 day).*

5.2 EnergyPlus validation

Validating the model means verifying that this closely behaves as the real building, matching the indoor temperature. Therefore in this part, the goal was to create a model of the laboratory as close to reality as possible in order to use it to make some energy demand consideration and to use it to extend the results to other building typologies. All the boundary conditions that could influence the internal temperature have been therefore implemented in detail.
The best result led to a significantly low RMSE equal to 0.75 °C.
Such a low value of RMSE means that the building is well modeled. Is it important to say that has been possible to reach this results thanks to the geometrical and physical simplicity of the building and thanks to the elevate knowledge and control of the boundary conditions. Moreover in the period form 21/01 to 28/01 there was not presence of people in the building and all electric loads were turned off; this obviously helps the accuracy of the model.
The small gap between the two curves shown in Figure 5.9 probably depends on the different thermal inertia of the structures, which is the most critical parameter to implement. This is because EnergyPlus is not optimized to model the internal frame of the building made of wood and steel beams. To take this into account, the density of internal structures has been arbitrarily raised to increase inertia and reduce RMSE.

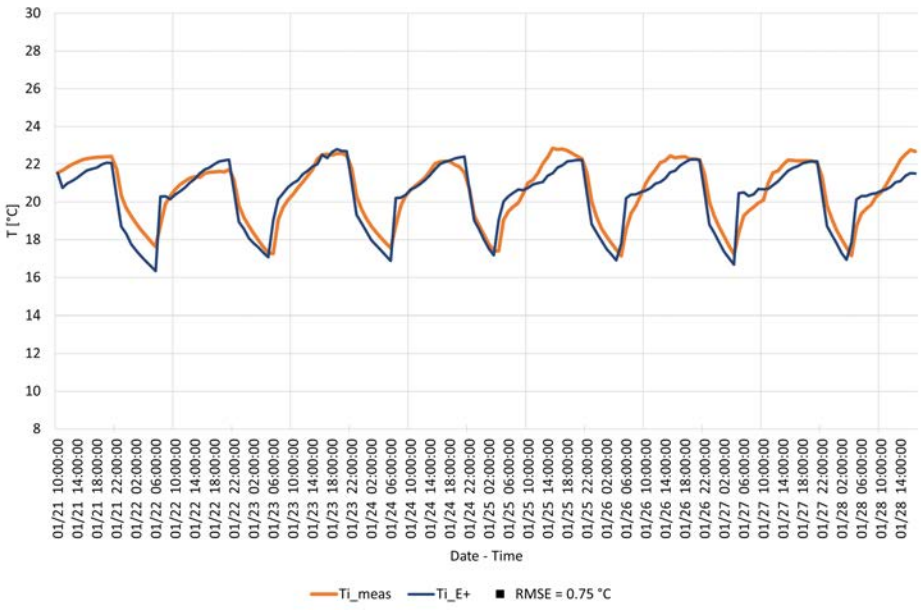

Figure 5.9: *Indoor temperature comparison between measured and simulated by EnergyPlus.*

5.3 Laboratory energy demand

5.3.1 Internal loads

The first simulation has been made with these boundary conditions:

- heating setpoint 20 °C; cooling setpoint 26 °C, latent + sensible;
- Internal loads Standard EN16798;
- windows solar heat gain = 0.6;
- ACR = 0.3 vol/h;
- heating season 15 October - 15 April;
- weather: Piacenza TRY from EnergyPlus website.

Table 5.5 shows heating and cooling energy demand and the peak power.
Figure 5.10 and 5.11 show the trend of the indoor temperature of the thermal zones and the timestep energy demand during the whole year.

Heating season		Cooling season	
Energy demand	**Peak power**	**Energy demand**	**Peak power**
[kWh/m2]	**[kW]**	**[kWh/m2]**	**[kW]**
68.3	2.22	13.1	1.71

Table 5.5: *RSE lab energy demand, internal loads*

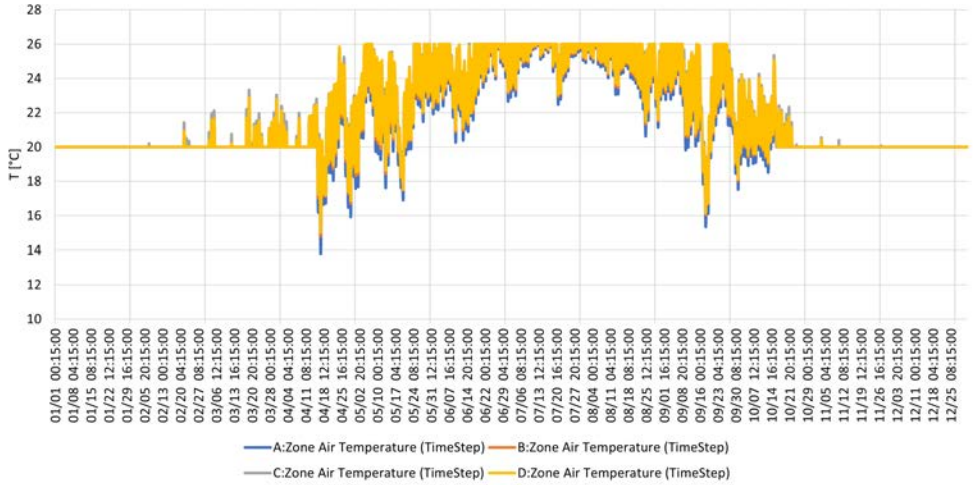

Figure 5.10: *Zone air temperatures, internal loads*

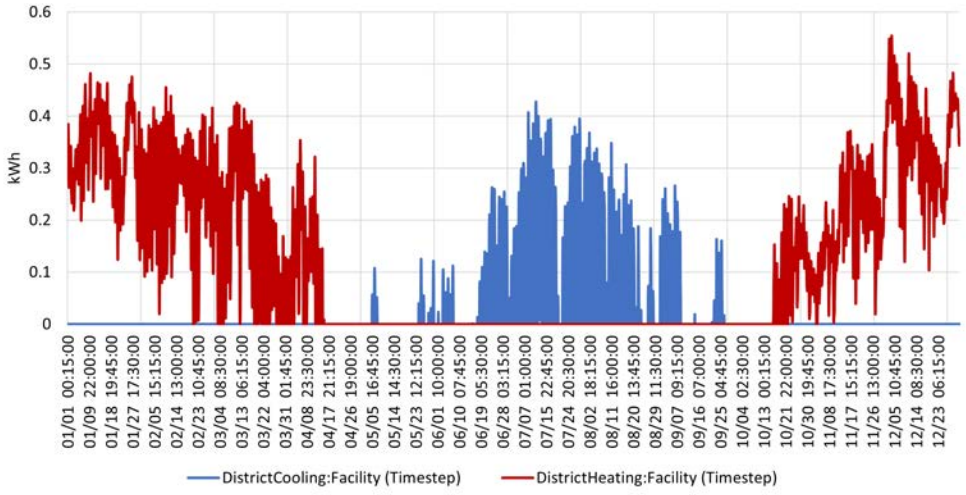

Figure 5.11: *Heating and cooling timestep energy demands, internal loads*

5.3.2 NO internal loads

The second simulation has been made with the same boundary conditions as before, but this time with no internal loads:

- heating setpoint 20 °C; cooling setpoint 26 °C, latent + sensible;
- Internal loads = 0;
- windows solar heat gain = 0.6;
- ACR = 0.3 vol/h;
- heating season 15 October - 15 April;
- weather: Piacenza TRY from EnergyPlus website.

Table 5.6 shows the heating and cooling energy demand and the peak power. Figure 5.12 and 5.13 show the trend of the indoor temperature of the thermal zones and the timestep energy demand during the whole year.

Heating season		**Cooling season**	
Energy demand	**Peak power**	**Energy demand**	**Peak power**
[kWh/m2]	**[kW]**	**[kWh/m2]**	**[kW]**
82.4	2.42	6.3	1.48

Table 5.6: *RSE lab energy demand, NO internal loads*

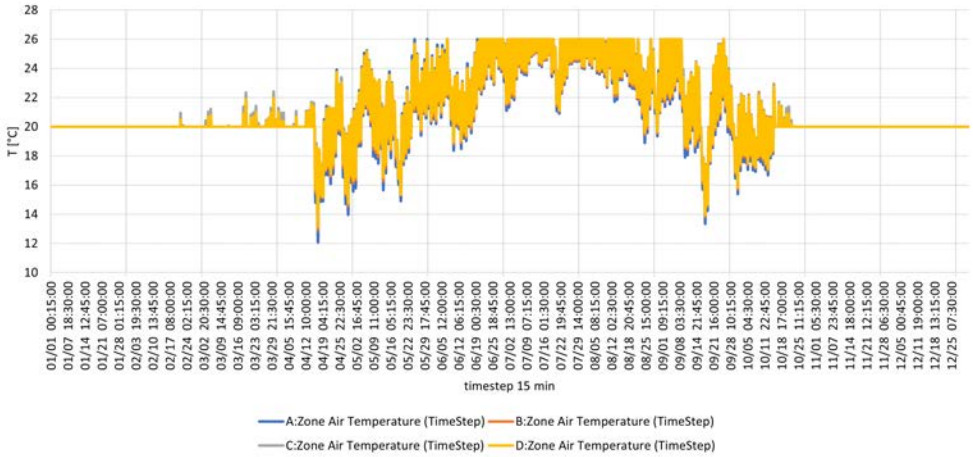

Figure 5.12: *Zone air temperatures, NO internal loads*

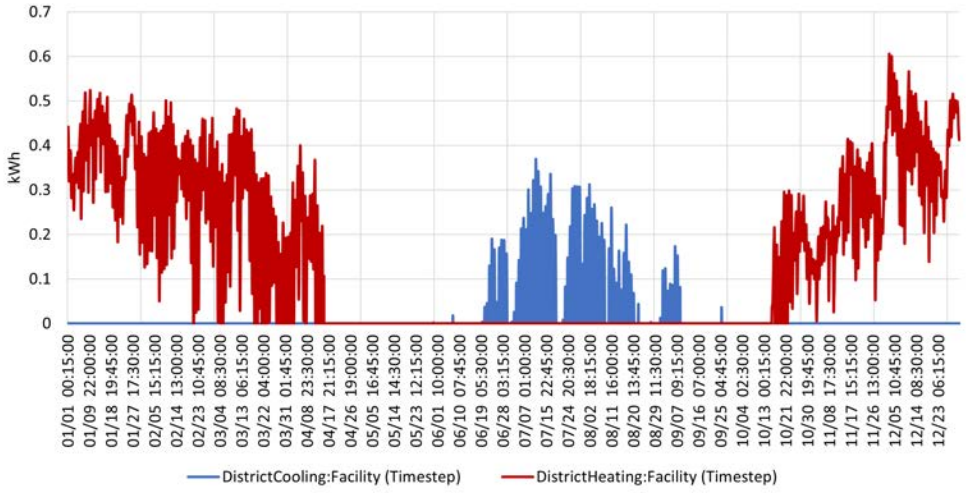

Figure 5.13: *Heating and cooling timestep energy demands, NO internal loads*

5.4 Simulated buildings training and testing

BLAB, 1-7 training, 1,2 testing

The first dataset is obtained by running the simulation on EnergyPlus with the laboratory model.
As for all the previous calibrations with real data, Table 5.7 and Figure 5.14, show lower RMSE values are reached in training periods and for 1 day of testing. Also the trend similar: one day is high to give accurate results, but then the median RMSE goes down in correspondence of the range 2 - 4 training days and then returns to rise.
The lower RMSE value reached in this simulation in testing period is 1.06 °C.

1 day testing								
ndTR n.	**MeTR °C**	**MeTE °C**	**MnTR °C**	**MnTE °C**	**σTR °C**	**σTE °C**	**TeTR °C**	**TeTE °C**
1	0.67	2.05	0.68	2.09	0.03	0.14	2.35	2.1
2	0.51	1.08	0.51	1.07	0.03	0.23	3.65	2.1
3	0.78	1.15	0.78	1.13	0.01	0.15	3.45	2.1
4	0.66	1.29	0.68	1.32	0.03	0.25	2.46	2.1
5	0.95	1.29	0.95	1.22	0.01	0.21	1.71	2.1
6	0.75	1.5	0.74	1.52	0.03	0.4	1.65	2.1
7	0.84	1.62	0.84	1.55	0.03	0.26	1.86	2.1

2 days testing								
ndTR n.	**MeTR °C**	**MeTE °C**	**MnTR °C**	**MnTE °C**	**σTR °C**	**σTE °C**	**TeTR °C**	**TeTE °C**
1	0.27	2.1	0.27	2.19	0.03	0.3	4.94	2.23
2	0.66	1.44	0.66	1.4	0.02	0.14	4	2.23
3	0.56	1.6	0.56	1.67	0.02	0.31	2.5	2.23
4	0.87	1.45	0.87	1.41	0.01	0.13	1.55	2.23
5	0.62	1.6	0.63	1.61	0.02	0.19	1.51	2.23
6	0.77	1.68	0.77	1.65	0.01	0.19	1.77	2.23
7	0.45	1.67	0.45	1.62	0.01	0.32	2.47	2.23

Table 5.7: *BLAB, 1-7 training, 1,2 testing.*

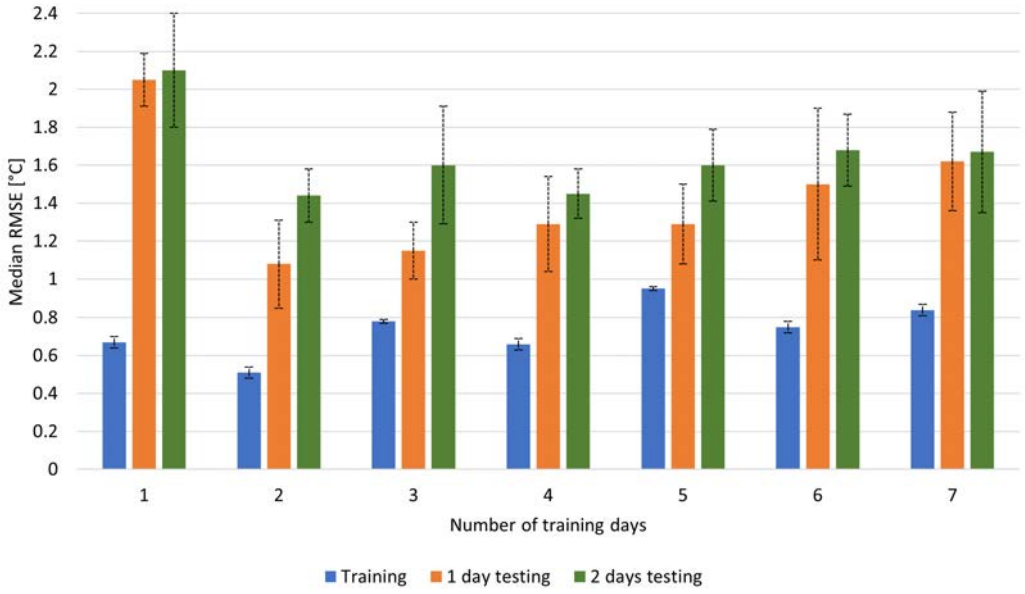

Figure 5.14: *BLAB, 1-7 training, 1,2 testing.*

BLAB_ni, 1-7 training, 1,2 testing

Table 5.8 and Figure 5.15 show the results of the simulation of the laboratory with a lower insulation.
As expected, better results are given in training and with 1 day testing. The gap between the results is this time higher than before, with the RMSE in 2 days testing simulation that stay always above 2 °C. The trends is similar to the previous simulations.

1 day testing								
ndTRn	**MeTR °C**	**MeTE °C**	**MnTR °C**	**MnTE °C**	**σTR °C**	**σTE °C**	**TeTR °C**	**TeTE °C**
1	0.6	1.59	0.6	1.64	0.02	0.17	2.35	2.1
2	0.62	0.94	0.62	0.96	0.03	0.13	3.65	2.1
3	0.54	0.89	0.55	0.87	0.03	0.15	3.45	2.1
4	0.65	1.19	0.66	1.27	0.03	0.18	2.46	2.1
5	0.88	1.14	0.89	1.13	0.02	0.13	1.71	2.1
6	0.75	1.71	0.75	1.75	0.04	0.28	1.65	2.1
7	0.94	1.89	0.95	2.01	0.05	0.36	1.86	2.1

2 days testing								
ndTRn	**MeTR °C**	**MeTE °C**	**MnTR °C**	**MnTE °C**	**σTR °C**	**σTE °C**	**TeTR °C**	**TeTE °C**
1	0.38	2.89	0.38	2.93	0.01	0.2	4.94	2.23
2	0.32	2.56	0.32	2.57	0.02	0.2	4	2.23
3	0.44	2.36	0.44	2.36	0.01	0.34	2.5	2.23
4	0.79	2.03	0.8	1.97	0.01	0.29	1.55	2.23
5	0.61	2.5	0.61	2.45	0.01	0.22	1.51	2.23
6	0.86	2.48	0.86	2.45	0.01	0.17	1.77	2.23
7	0.47	2.61	0.46	2.46	0.02	0.29	2.47	2.23

Table 5.8: *BLAB_ni, 1-7 training, 1,2 testing.*

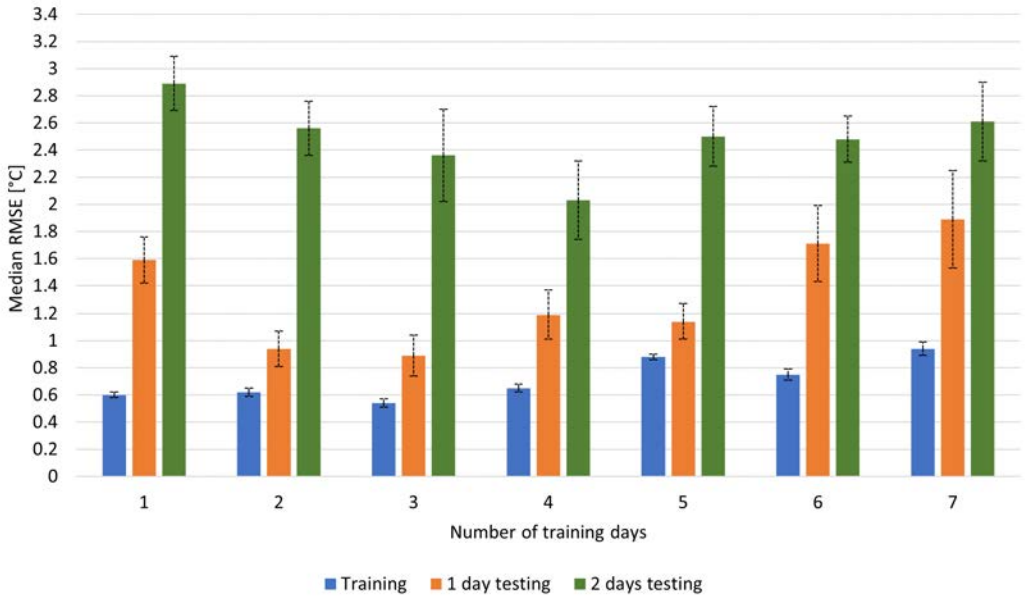

Figure 5.15: *BLAB_ni, 1-7 training, 1,2 testing.*

B70, 1-7 training, 1-2 testing

Table 5.9 and Figure 5.16 show the results of the simulation for the layer B70, the high-weight structure with the lowest insulation.
What can be immediately noticed are the low standard deviation values that are found in both training and testing with values very close to 0. In previous tests the standard deviation was lower in training, but in testing had never been found values so low.
Trends are similar to the previous for lightweight structures.

1 day testing								
ndTRn	**MeTR °C**	**MeTE °C**	**MnTR °C**	**MnTE °C**	**σTR °C**	**σTE °C**	**TeTR °C**	**TeTE °C**
1	0.52	2	0.52	1.99	0.02	0.18	2.35	2.1
2	0.4	0.73	0.4	0.74	0.01	0.05	3.65	2.1
3	0.39	0.38	0.39	0.38	0	0.04	3.45	2.1
4	0.45	0.97	0.45	0.98	0	0.04	2.46	2.1
5	0.53	1.03	0.53	1.03	0	0.07	1.71	2.1
6	0.46	1.2	0.46	1.18	0.01	0.08	1.65	2.1
7	0.44	1.11	0.44	1.09	0	0.08	1.86	2.1

2 days testing								
ndTRn	**MeTR °C**	**MeTE °C**	**MnTR °C**	**MnTE °C**	**σTR °C**	**σTE °C**	**TeTR °C**	**TeTE °C**
1	0.27	2.17	0.27	2.17	0.01	0.19	4.94	2.23
2	0.38	0.73	0.38	0.74	0	0.08	4	2.23
3	0.38	1.05	0.38	1.05	0	0.04	2.5	2.23
4	0.48	1.16	0.48	1.16	0	0.03	1.55	2.23
5	0.39	1.11	0.4	1.12	0.01	0.07	1.51	2.23
6	0.38	1.05	0.38	1.07	0.01	0.11	1.77	2.23
7	0.32	1.02	0.32	1.03	0.01	0.07	2.47	2.23

Table 5.9: *B70, 1-7 training, 1,2 testing.*

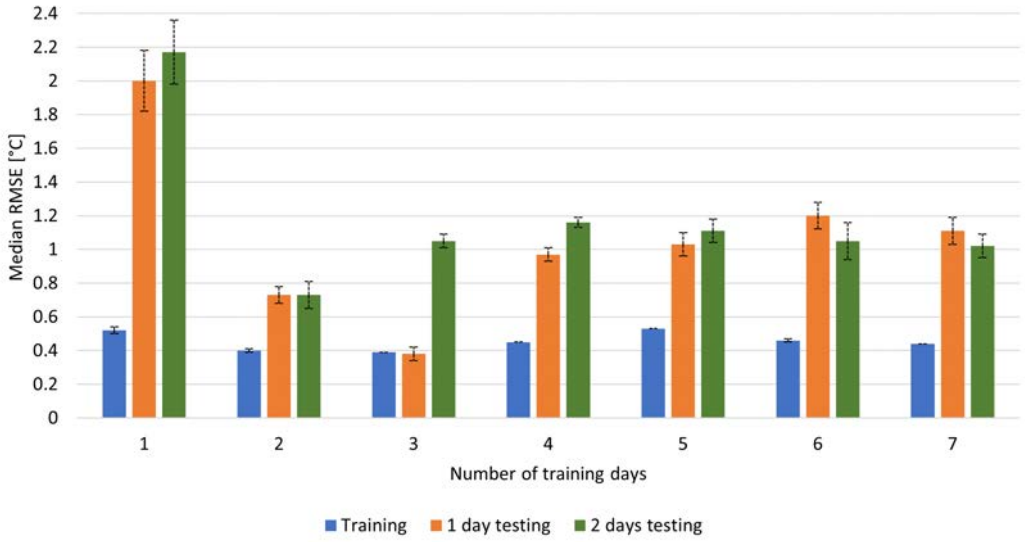

Figure 5.16: *B70, 1-7 training, 1,2 testing.*

B90, 1-7 training, 1,2 testing

Table 5.10 and Figure 5.17 show the results of the simulation for the layer B90.
RMSE values have become on average even lower, with standard deviation values very close to 0 except for the first simulation. RMSE for one day testing is always lower than one day and this time the values stay very low also for higher training days.
The standard deviation of results is low except for one day of training.

1 day testing								
ndTRn	**MeTR °C**	**MeTE °C**	**MnTR °C**	**MnTE °C**	**σTR °C**	**σTE °C**	**TeTR °C**	**TeTE °C**
1	0.44	1.61	0.44	1.61	0.02	0.17	2.35	2.1
2	0.27	0.51	0.27	0.51	0.01	0.06	3.65	2.1
3	0.45	0.38	0.45	0.37	0	0.03	3.45	2.1
4	0.42	0.8	0.42	0.79	0	0.05	2.46	2.1
5	0.5	0.72	0.5	0.72	0	0.05	1.71	2.1
6	0.41	0.78	0.41	0.78	0	0.08	1.65	2.1
7	0.4	0.79	0.4	0.81	0	0.07	1.86	2.1

2 days testing								
ndTRn	**MeTR °C**	**MeTE °C**	**MnTR °C**	**MnTE °C**	**σTR °C**	**σTE °C**	**TeTR °C**	**TeTE °C**
1	0.16	1.74	0.16	1.69	0.01	0.26	4.94	2.23
2	0.41	0.67	0.41	0.66	0	0.05	4	2.23
3	0.36	0.84	0.36	0.84	0	0.03	2.5	2.23
4	0.45	0.81	0.45	0.81	0	0.05	1.55	2.23
5	0.35	0.79	0.35	0.78	0.01	0.05	1.51	2.23
6	0.34	0.74	0.34	0.74	0.01	0.04	1.77	2.23
7	0.3	0.77	0.3	0.76	0	0.06	2.47	2.23

Table 5.10: *B90, 1-7 training, 1,2 testing.*

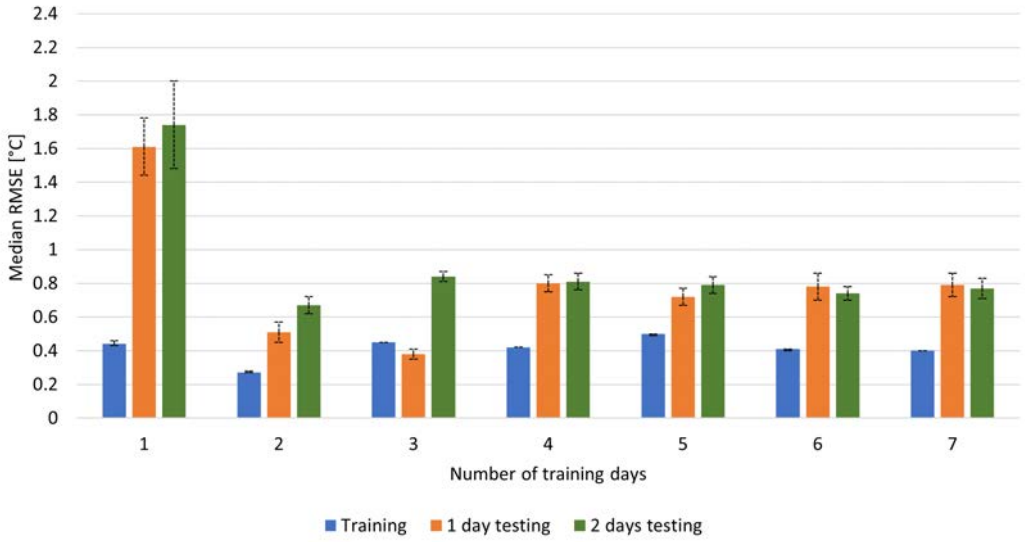

Figure 5.17: *B90, 1-7 training, 1,2 testing.*

BN, 1-7 training, 1,2 testing

Table 5.11 and Figure 5.18 show the results of the simulation for the layer BN.
The simulation run with the more insulating structure shows the lowest RMSE values both for training and testing. The trend of the RMSE is similar to the previous tests but the value stays always below 0.6 °C.

1 day testing								
ndTRn	**MeTR °C**	**MeTE °C**	**MnTR °C**	**MnTE °C**	**σTR °C**	**σTE °C**	**TeTR °C**	**TeTE °C**
1	0.39	1.46	0.39	1.43	0.02	0.13	2.35	2.1
2	0.21	0.41	0.22	0.42	0.02	0.06	3.65	2.1
3	0.4	0.44	0.4	0.45	0	0.06	3.45	2.1
4	0.36	0.54	0.36	0.54	0	0.07	2.46	2.1
5	0.42	0.46	0.42	0.47	0	0.04	1.71	2.1
6	0.32	0.62	0.32	0.59	0	0.08	1.65	2.1
7	0.31	0.52	0.31	0.56	0	0.1	1.86	2.1

2 days testing								
ndTRn	**MeTR °C**	**MeTE °C**	**MnTR °C**	**MnTE °C**	**σTR °C**	**σTE °C**	**TeTR °C**	**TeTE °C**
1	0.12	1.2	0.12	1.24	0	0.18	4.94	2.23
2	0.35	0.55	0.35	0.55	0	0.04	4	2.23
3	0.3	0.58	0.3	0.57	0.01	0.05	2.5	2.23
4	0.37	0.57	0.37	0.58	0	0.02	1.55	2.23
5	0.27	0.55	0.27	0.55	0.01	0.05	1.51	2.23
6	0.26	0.56	0.27	0.56	0.01	0.06	1.77	2.23
7	0.24	0.66	0.24	0.64	0.01	0.1	2.47	2.23

Table 5.11: *BN, 1-7 training, 1,2 testing.*

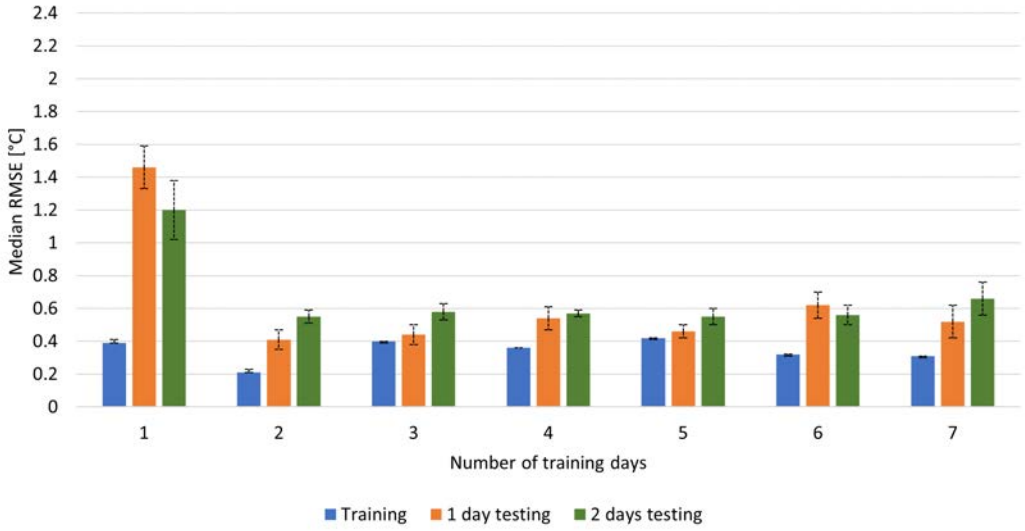

Figure 5.18: *BN, 1-7 training, 1,2 testing.*

5.4.1 Interpretation

After testing the calibration algorithm on the basis of data from the laboratory sensors, the tests were extended to more massive buildings simulated via Energyplus. Five stratigraphy were simulated divided into two groups, lightweight and heavyweight, and applied to the laboratory model. For each stratigraphy, tests were carried out with 1 and 2 days of testing for a training period between 1 and 7 days.
First of all it has been observed that the results have a similar tendency to the previous tests, from real data.
It is confirmed that:

- training RMSE increases with training length;
- best results are achieved with one day of testing;
- best results are achieved with a training period between 2 and 4 days;

To compare the results, the Figures 5.19, 5.20, 5.21 and 5.22 show for each group of structure the mean value of the RMSE obtained between 1 and 2 days of testing. The most interesting result is that lower RMSE are obtained for heaviest and better insulated structures. The external driving forces that have the greatest effect on the internal temperature are the solar radiation and the air temperature, while the internal driving forces are in this case given only by the thermal power provided by fancoils. In general it has been seen as such power, reaching peak values of 2300 W during the ON periods, has an important effect on the internal temperature and is preponderant compared to the external driving forces. This statement is the truest, the more insulated the building is. For badly insulated and light buildings there is a contrary tendency.
This leads to think that for more insulated structures the calibration becomes less effective because of the greater oscillations on the internal temperature due to thermal power. In reality, on equal thermal capacity, the tests show the opposite, namely that the greater the insulation of the structure, the greater the accuracy of the calibration. This is probably due to the fact that the trend of the thermal power supplied by fancoils is cyclical as shown in Figure 5.23, with values that are repeated periodically with the same hourly power in the period analyzed. This cyclicity of the predominant driving force facilitates machine learning of the algorithm and allows the achievement of better results in testing periods. For badly insulated buildings instead the external driving forces are predominant and as these are subject to greater variability, and worsen the quality of learning.
Also the standard deviation of results is lower for heaviest structure, which indicates more stability of the calibration.
It can then be said that on equal insulation the calibration is better for buildings more massive, as observed by comparing the LAB with BN in Figure 5.24. These two structure have a different thermal capacity but similar transmittances as described in Tables 4.10 and

4.11. This is because in buildings with a lower thermal capacity both internal and external driving forces cause more temperature fluctuations making calibration less accurate. Figure 5.23 shows how the indoor temperature in the BLAB structure is more affected by the solar radiation than in BN.

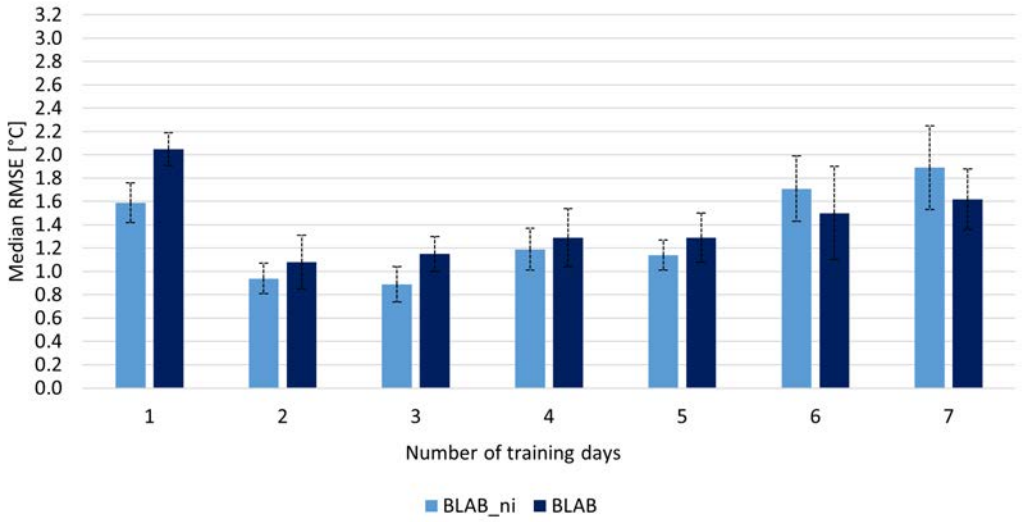

Figure 5.19: *Lightweight median testing RMSE, 1 testing day.*

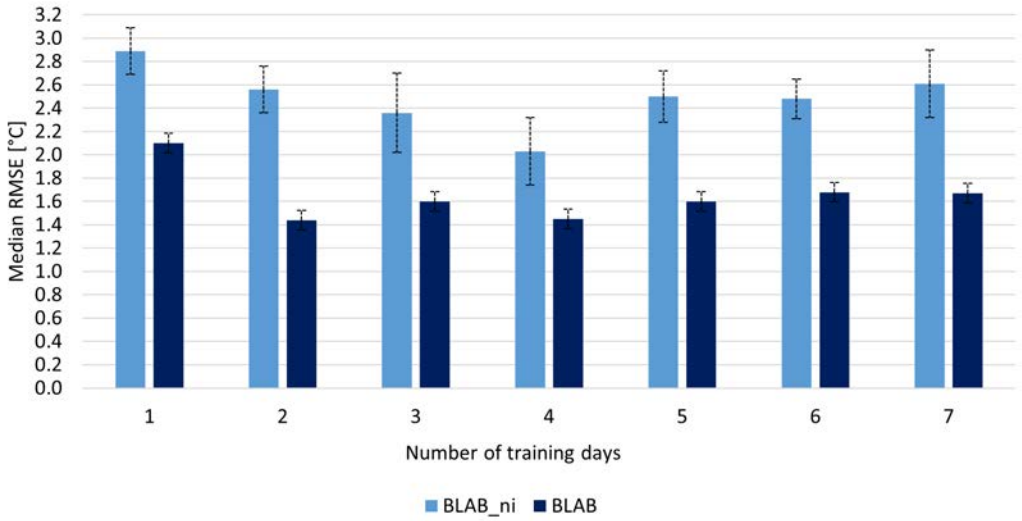

Figure 5.20: *Lightweight median testing RMSE, 2 testing days.*

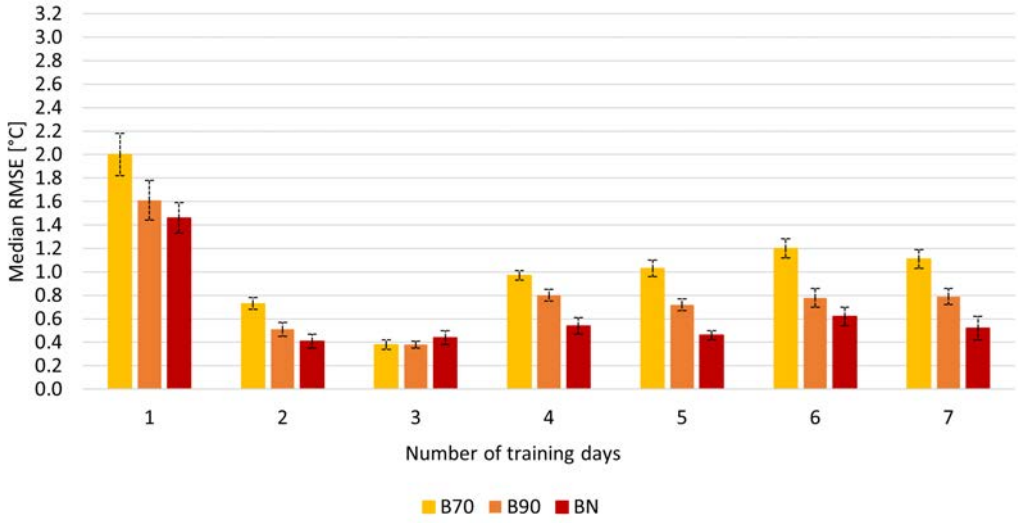

Figure 5.21: *Heavyweight median testing RMSE, 1 testing day.*

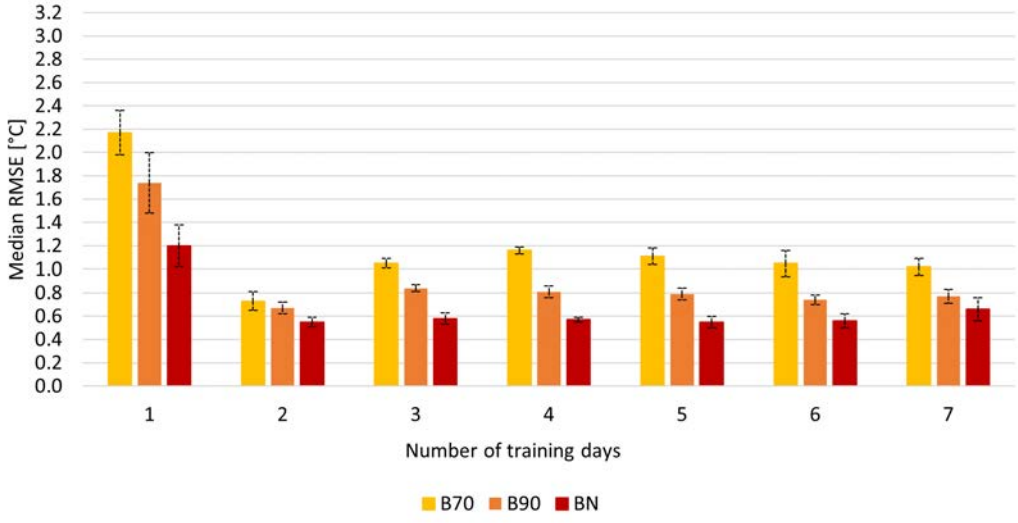

Figure 5.22: *Heavyweight median testing RMSE, 2 testing days.*

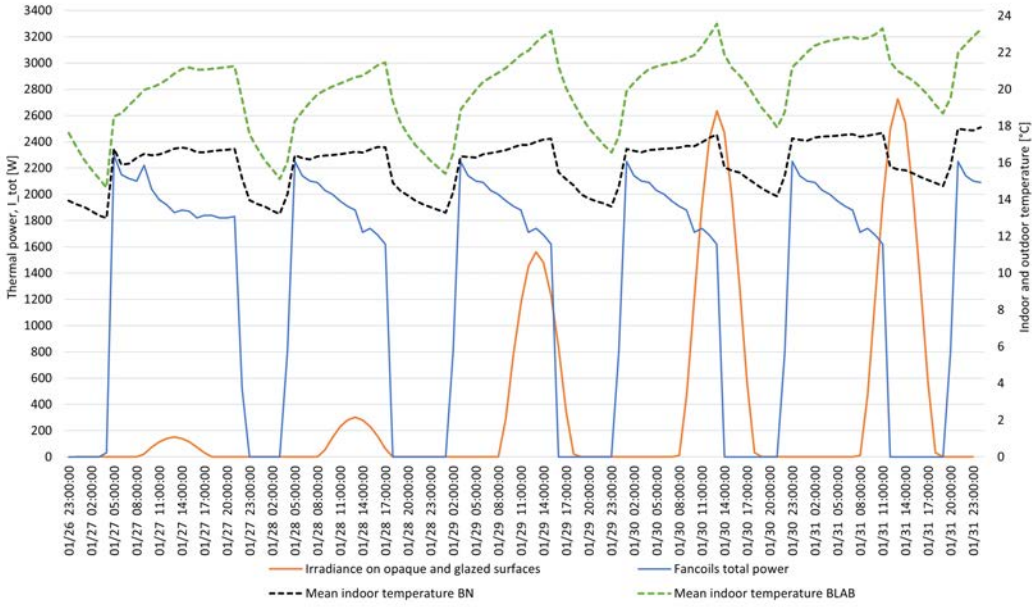

Figure 5.23: *Heavyweight BN and Lightweight BLAB indoor temperature comparison; external and internal driving force influence.*

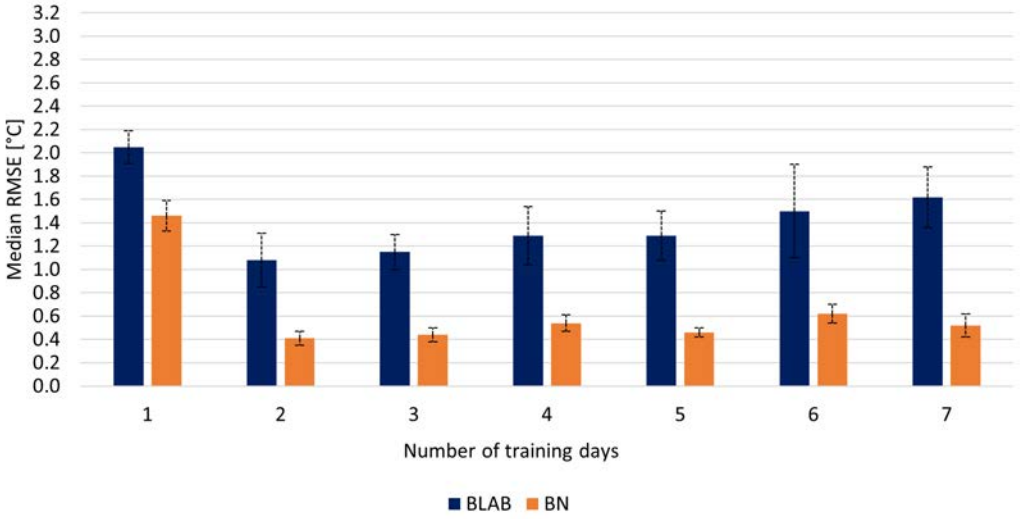

Figure 5.24: *Heavyweight BN and Lightweight BLAB; 1 day testing RMSE comparison.*

5.5 RC model physical parameters

The purpose of the investigation was to determine if the accuracy of the initial parameters influences the calibration.
So 3 different simulations has been made: **NOM**, **RNDDW** and **RND100** each with a different setting of the initial parameters.
The results of the tests are shown in Figures 5.26, 5.27 and 5.28 and resumed in Table 5.12 and Figure 5.25 where the RMSE on calibration and RMSE are resumed.

		RMSE_parameters	RMSE_train	RMSE_test
NOM	mean	0.422	0.359	0.545
	σ	0.166	0.004	0.062
RNDDW	mean	0.404	0.359	0.547
	σ	0.155	0.004	0.055
RND100	mean	46.000	0.931	1.406
	σ	18.840	0.840	0.945

Table 5.12: *Means and standard deviations RMSE for RNDDW, RND100 and NOM.*

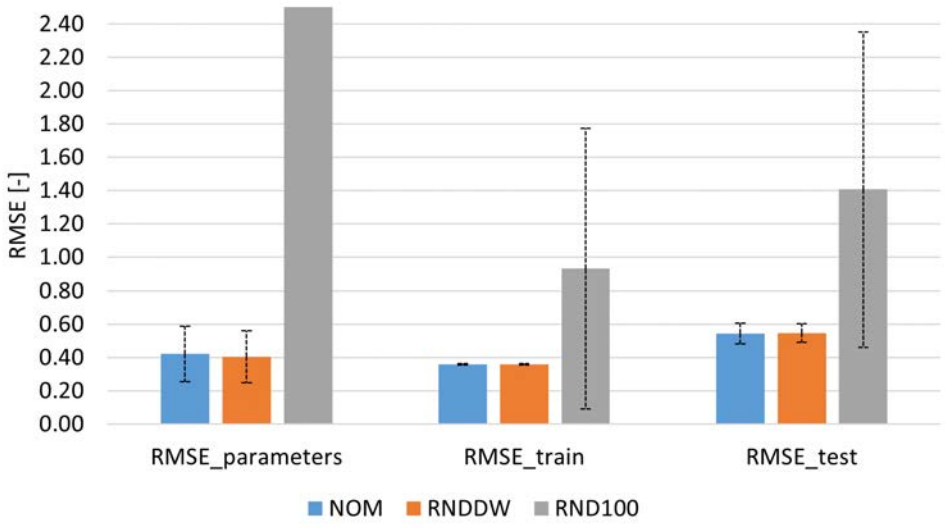

Figure 5.25: *Results comparison.*

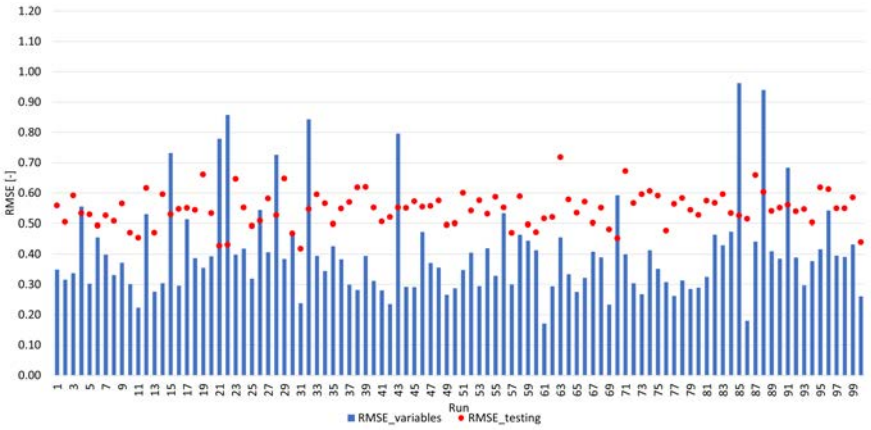

Figure 5.26: *1 day testing NOM, RMSE on variables and testing.*

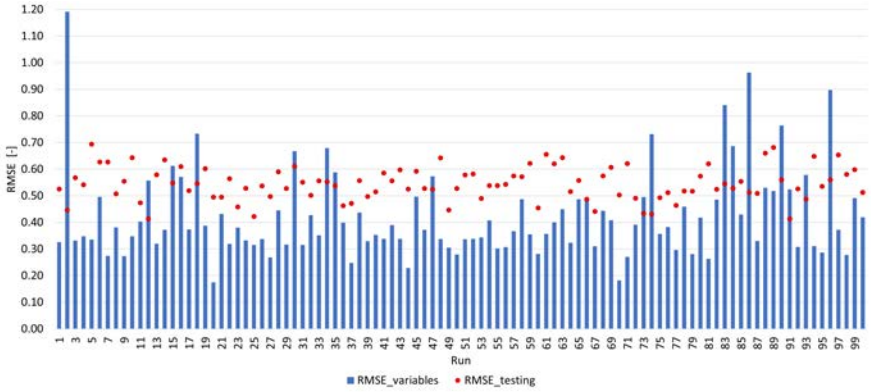

Figure 5.27: *1 day testing RNDDW, RMSE on variables and testing.*

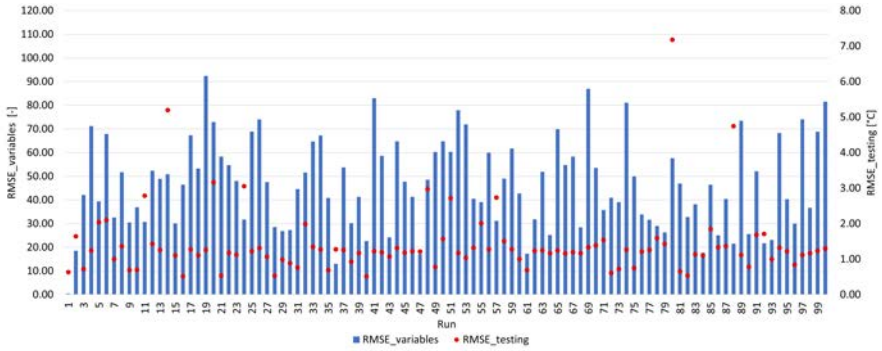

Figure 5.28: *1 day testing RND100, RMSE on variables and testing.*

RNDDW and NOM simulations show very similar results on both parameters and calibration, with equal averages RMSE and standard deviations. RND100 instead, where the initial variables do not respect at all the order of magnitude of the parameters, shows completely different and unsatisfactory results and RMSE much higher than in previous cases. The simulations show that the PSO, despite a small error in the initial parameters, it still manages to achieve the best results thanks to the nature of the algorithm itself and to the continuous updating of the limits of the domain to every iteration that makes it particularly exploratory. A low RMSE on the parameters indicates that the optimal solution and therefore the global minimum of the objective function is achieved preserving the physical sense of the parameters. If the error on the initial parameters is very high, the error on the optimal parameters will be very high to, and the calibration will be affected in accuracy.

6 Conclusions

The objective of the study was to develop and test the calibration of a lumped capacitance model for the prediction of buildings indoor temperature. The reproduction of the thermal behavior of the building is the first block of constraints of an optimization problem that aims to minimize the energy costs while maintaining the comfort of the indoor environment. The aim of the research project is to apply this optimization on the RSE pilot laboratory of Piacenza in order to develop a BEMS based on predictive control (MPC).
The calibration is the numerical process by which the parameters describing the dynamic thermal behavior of the building are initially estimated by an approximate knowledge of the building and then recalculated to ensure that the mathematical model matches as close as possible to the physical quantities measured. This operation is done by using the heuristic evolutive PSO, that minimizes the difference between the temperature profile measured in the building and those calculated by the RC model.
The greater the accuracy of the calibration, the greater the accuracy in the prediction of the internal temperature and the better the reliability and stability of the optimal controller.
First, the calibration code was applied to the RSE pilot building in Piacenza, using data from the sensors and the weather station installed in the laboratory. The results from the implementation of the code in three different periods, have shown that the best training period is between 2 and 4 days. Calibration yielded better results in the Winter 3 dataset, in which greater control and cyclicity of boundary conditions led to a better learning, with a minimum RMSE achieved with 4 days of training and 1 of testing equal to 0.51 °C. In general, a low training period avoids that the boundary conditions differ significantly from those of the testing period. On the other hand, a higher number of training days avoids problems caused by anormal, repentine change in weather conditions. In fact, the variability of the boundary conditions such as the random opening of a window or the activation of an electric load can lead to an anomaly in the data that can cause overfitting, thus resulting in a loss of accuracy. This observation suggests to incorporate indicators related to the boundary conditions in the numerical procedure of the calibration.
Then, in order to generalize these results, different structures has been simulated using EnergyPlus to test the calibration in buildings with different insulation and thermal capacity. The structures have been divided into light and heavy structures and include the simulated laboratory itself, a typical structure of the 70s, one from the 90s and a recent one with higher thermal insulation. The different structures were applied to the geometries of the building

modeled on the basis of the laboratory of Piacenza. Before proceeding with the process, the model was validated by verifying the internal temperature error in a 7 day interval, demonstrating a fair accuracy with an RMSE of 0.75 °C.
The application of the calibration code to simulated structures confirmed the evidence from the actual laboratory tests and prompted some further considerations. Better results have been obtained on the most insulated and massive building. In particular the most insulated haevyweight structure (BN) has reached a minimum median testing RMSE of 0.41 °C. This is because the periodical trend of the internal dominant driving force represented by the fancoils power, allowed an optimal learning of the algorithm. On the contrary, in lightweight and low insulated structures, external driving forces likes solar radiation and air temperature, become more important and their variability and uncontrollability leads to a worst learning of the algorithm. In fact the best testing RMSE achived by the structure BLAB is 1.08 °C.
In general, tests have shown that BEMS systems for minimizing energy costs, are well suited to be used in newly designed structures, very insulated and high thermal capacity.
In particular, this system has shown a high accuracy especially in cases where there was greater control and repetitiveness of the internal boundary conditions.
Finally, a further investigation was carried out to understand how the accuracy of the initial parameters influences the calibration. All tests were conducted on the BN structure with four training days and one testing day, setting into three different ways the initial parameters: initial parameters equal to the correct nominal parameters, initial parameters randomly in a physically meaningful domain so as to maintain the physical sense of the parameters and randomly multiplying the initial parameters by a number between 0 and 100 deliberately losing the order of magnitude and the physical sense of the parameters. The results showed that there are no substantial differences in terms of RMSE in the results obtained with the initial parameters equal to the correct nominal and those with the initial parameters randomly selected in a coherent physical neighborhood. In these cases, the average accuracy achieved by the calibration process and its variance are both unaltered. Conversely, unsatisfactory results were obtained by setting completely random parameters. Therefore, thanks to the exploratory nature of the PSO algorithm, a summary knowledge of the physical characteristics of the building is sufficient to obtain good results, which is a strong practical implication for the implementation of the proposed optimal controller as a commercial product.
The next step of the research work consists in quantifying the influence of the calibration accuracy throughout the whole optimization process. In fact, one of the biggest limitations of this work was the lack of a benchmark to compare the results, which did not make it possible to make deeper consideration about the calibration accuracy on the overall performance of the BEMS.
It might then be interesting to compare the grey box model with a neural network based black box model, which fully exploits the amount of data coming from the laboratory sensors.

Optimization problem description

A

The current version of the BEMS controls only the heat pump and the circulation pump of the secondary water loop (see Figure 2.6). The core of the BEMS is a mixed integer linear programming (MILP) problem solved in a rolling horizon scheme. This means that the optimization is repeated with a predetermined sample time. At each step, the operation of the HVAC system is planned for the next H hours, where H is the horizon of the optimization problem. The planning consists in determining the state of the heat pump, which consists of an on/off signal and a frequency signal communicated to the inverter-driven compressor.
The optimization problem consists in the minimization of the economic objective function:

$$min \sum_t (\lambda_t^b w_t^b - \lambda_t^s w_t^s) + \sum_t \gamma(\delta_t^{\uparrow} + \delta_t^{\downarrow}) \tag{A.1}$$

Where λ_t^b and λ_t^s are respectively the purchase and sale price of electricity; w_t^b and w_t^s are the quantities of electricity purchased and sold from/to the grid, and the second summation is a penalty that is introduced to limit the events in which the air temperature comes out of comfort limits. The variable $\delta_t^{\uparrow}$ quantifies the difference between the air temperature and the maximum temperature allowed when this limit is exceeded. The same happens with $\delta_t^{\downarrow}$ when the temperature drops below the minimum allowed temperature. The γ parameter serves only to weigh appropriately the penalty of the thermal discomfort compared to the economic objective function. The quantities of electricity bought and sold shall be in line with the electricity balance obtained by using the POD as the interface between the building and the electricity distribution network:

$$w_t^b + w_t^{pv} = w_t^s + w_t^{hp} + w_t^{od} \tag{A.2}$$

Equation A.2 expresses the equality between the energy entering the building, the result of the sum of energy bought from the grid and energy produced locally by the photovoltaic system w_t^{pv}, and the energy coming out of the building that is the sum of the electricity absorbed by the heat pump w_t^{hp} and the other devices w_t^{od} and the one sold to the grid. The indoor air temperature θ_t^i is always between the minimum limit $\theta_t^{i,min}$ and the maximum $\theta_t^{i,max}$, except when there are breaches by the limits of comfort $\delta_t^{\downarrow}$ and $\delta_t\uparrow$.
Two equations are then defined to define these variables:

$$\theta_t^{i,min} - \delta_t^{\downarrow} \leq \theta_t^i \tag{A.3}$$

$$\theta_t^i \leq \theta_t^{i,max} + \delta_t^{\uparrow} \tag{A.4}$$

The temperature overruns so defined shall be positive or zero. The amount of energy absorbed by the heat pump should therefore be defined. Note that normally the electrical energy absorbed by the compressor is expressed as the ratio of the thermal energy released by the condenser and the COP of the machine. Here it is expressed instead as a product between the thermal energy yielded by the condenser ϕ_t^{hc} and a factor F_t^{hp} equal to the inverse of the COP, as shown in Equation A.5. This way of expressing the COP allows to preserve the linearity of the constraint and therefore of the optimization problem. The factor F_t^{hp} is then expressed as a function of the external temperature θ_t^e as shown by Equation A.6.

$$w_t^{hp} = F_t^{hp} \phi_t^{hc} \tag{A.5}$$

$$F_t^{hp} = k_0 + k_1 \theta_t^e \tag{A.6}$$

The heat released by the heat pump to the ϕ_t^{hc} plant must always be less or equal to the nominal heat pump output and coincides, in this model, with the heat released by the plant to the building. The set of equations and inequalities constitutes the constraints of the optimization problem and must be repeated for each step within the horizon, that is $\forall t \in [1, H]$.

Objective function	**Minimizing energy costs and discomfort penality**	**Equation A.1**
1^{st} block of constraints	RC model of the building	Equations 3.1 - 3.3
2^{nd} block of constraints	Electric energy balance	Equation A.2
3^{rd} block of constraints	Thermal comfort limits	Equations A.3 - A.4
4^{th} block of constraints	Heat pump performance	Equation A.5

Table A.1: *Objective function and optimization problem constraints.*

Bibliography

[1] Abergel, T., Dean, B., Dulac, J., Hamilton, I., 2018 Global Status Report - Towards a zero-emission, efficient and resilient buildings and construction sector; International Energy Agency (IEA), 2018.

[2] European Commission, Communication and roadmap on the European Green Deal, 2019.

[3] European Commission, Directive (EU) 2018/844 of the European Parliament and of the Council of 30 May 2018 amending Directive 2010/31/EU on the energy performance of buildings and Directive 2012/27/EU on energy efficiency, 2018.

[4] Thibault Q. Péan, Jaume Salom, Ramon Costa-Castelló, Review of control strategies for improving the energy flexibility provided by heat pump systems in buildings, 2019.

[5] A.Zarrella, J.Vivian, Messa a punto e validazione di un sistema di Machine Learning (ML) per ottimizzare i consumi energetici di un singolo edificio (rif. RSE, 19006470 del 15 - 07 - 2019), Università degli studi di Padova, 2019.

[6] Borui Cui, Cheng Fanb, Jeffrey Munk, Ning Mao, Fu Xiao, Jin Dong, Teja Kuruganti, A hybrid building thermal modeling approach for predicting temperatures in typical, detached, two-story houses, 2019.

[7] American Society of Heating, Refrigerating and Air Conditioning Engineers, Inc., ASHRAE Handbook-Fundamentals, DeR International, Ltd, 2013.

[8] Coakley D, Raftery P, Keane M., A review of methods to match building energy simulation models to measured data, Renew Sustain Energy, 2014.

[9] Braun JE, Chaturvedi N., An inverse grey box model for transient building load prediction, 2002

[10] Dong B, Li Z, Rahman SM, Vega R., A hybrid model approach for forecasting future residential electricity consumption, 2016.

[11] Kim D, Cai J, Ariyur KB, Braun JE. System identification for building thermal systems under the presence of unmeasured disturbances in closed loop operation, lumped disturbance modeling approach, 2016

[12] International Organization for Standardization ISO 13790, Energy performance of buildings: calculation of energy use for space heating and cooling, 2nd ed. Geneva, Switzerland, 2008.

[13] International Organization for Standardization ISO 13786, Thermal performance of building components - Dynamic thermal characteristics - Calculation methods. Geneva, Switzerland, 1999.

[14] Angelo Zarrella, Enrico Prataviera, Pierdonato Romano, Laura Carnieletto, Jacopo Vivian, Analysis and application of a lumped-capacitance model for urban building energy modeling, University of Padua, 2020.

[15] EMS application guide, https://energyplus.net/documentation, 2021.

[16] Cosi Giovanni, Algoritmi evolutivi per la gestione ottimale di impianti HVAC, 2010.

[17] Craig W. Reynolds, Flocks, herds, and schools, A distributed behavioral model, ACM Computer Graphics, 1987.

[18] J. Kennedy, R.C. Eberhart, Particle swarm optimization, Proceedings of the IEEE International Conference on Neural Networks, 1995.

[19] Adam P.Piotrowski, Jaroslaw J.Napiorkowski, Agnieszk E.Piotrowska, Population size in Particle Swarm Optimization, 2020.

[20] Ineichen, P., Perez, R.,A new airmass independent formulation for the Linke turbidity coefficient, Solar Energy 73, 2002.

[21] D.H. Blum, K. Arendt, L. Rivalin, M.A. Piette, M. Wetter, C.T. Veje, Practical factors of envelope model setup and their effects on the performance of model predictive control for building heating, ventilating, and air conditioning systems, 2019.

[22] Muneer T, Younes S, Munawwar S., Discourses on solar radiation modeling, 2007.

[23] Žáčeková E, Vána Z, Cigler J., Towards the real-life implementation of MPC for an office building: identification issues, 2014.

[24] Cabrera DFM, Zareipour H. Data association mining for identifying lighting energy waste patterns in educational institutes, 2013.

[25] European Committee for Standardization. EN 16798-1, Energy performance of buildings - Ventilation for buildings. Part 1: Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics - Module M1-6, 2019.

[26] Decreto interministeriale 26 giugno 2015, Applicazione delle metodologie di calcolo delle prestazioni energetiche e definizione delle prescrizioni e dei requisiti minimi degli edifici - APPENDICE B, 2015.

'Considerate la vostra semenza:
fatti non foste a viver come bruti,
ma per seguir virtute e canoscenza'

Dante Alighieri, il Canto di Ulisse

Ringraziamenti

Vorrei ringraziare innanzitutto l'Ing. Jacopo Vivian, relatore della tesi, per la pazienza e l'aiuto fornitomi durante la stesura del lavoro e soprattutto per avermi dato la possibilità di partecipare ad un progetto cosi ambizioso e interessante.

Desidero poi ringraziare il mio coinquilino e amico lòzio Fex, per il sostegno informatico necessario a comprendere il linguaggio di Python e per le 1000 avventure vissute insieme in Via Ugo Foscolo.

Sarò grato per sempre al Petrarca Rugby e ai ragazzi della Serie A, con i quali ho trascorso momenti unici in campo e fuori dal campo, sacrificando volentieri parte della vita universitaria per vivere momenti indimenticabili e realizzare i sogni che avevo fin da bambino.

Ringrazio la mia fidanzata Balda, compagna di vita dal Liceo all'Università, sempre al mio fianco in ogni momento.

Ringrazio infine in modo speciale le mie nonne e i miei genitori.
A voi, che pur vivendo vite diverse siete sempre stati uniti, alimentando le mie passioni e sostenendomi con presenza e affetto in ogni situazione. A voi, che avete plasmato ogni mio successo e reso possibile questo giorno.

Belluno, 23 aprile 2021 *Mattia D'Incà*