UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

PhD program in Electronics Engineering

Advanced energy management strategies for HVAC systems in smart buildings

Doctoral thesis by:

Enric Sala Cardoso

Thesis advisors:

Dr. Luis Romeral

Dr. Miguel Delgado Prieto

Department of Electronics Engineering

Terrassa, September, 2019

Abstract

The efficacy of the energy management systems at dealing with energy consumption in buildings has been a topic with a growing interest in recent years due to the ever-increasing global energy demand and the large percentage of energy being currently used by buildings. The scale of this sector has attracted research effort with the objective of uncovering potential improvement avenues and materializing them with the help of recent technological advances that could be exploited to lower the energetic footprint of buildings. Specifically, in the area of heating, ventilating and air conditioning installations, the availability of large amounts of historical data in building management software suites makes possible the study of how resource-efficient these systems really are when entrusted with ensuring occupant comfort. Actually, recent reports have shown that there is a gap between the ideal operating performance and the performance achieved in practice.

Accordingly, this thesis considers the research of novel energy management strategies for heating, ventilating and air conditioning installations in buildings, aimed at narrowing the performance gap by employing data-driven methods to increase their context awareness, allowing management systems to steer the operation towards higher efficiency. This includes the advancement of modeling methodologies capable of extracting actionable knowledge from historical building behavior databases, through load forecasting and equipment operational performance estimation supporting the identification of a building's context and energetic needs, and the development of a generalizable multi-objective optimization strategy aimed at meeting these needs while minimizing the consumption of energy.

The experimental results obtained from the implementation of the developed methodologies show a significant potential for increasing energy efficiency of heating, ventilating and air conditioning systems while being sufficiently generic to support their usage in different installations having diverse equipment. In conclusion, a complete analysis and actuation framework was developed, implemented and validated by means of an experimental database acquired from a pilot plant during the research period of this thesis. The obtained results demonstrate the efficacy of the proposed standalone contributions, and as a whole represent a suitable solution for helping to increase the performance of heating, ventilating and air conditioning installations without affecting the comfort of their occupants.

Keywords: *chiller sequencing; deep learning; energy efficiency; energy management; load forecasting; machine learning; model-predictive control; neural networks; optimal chiller loading; unsupervised learning*.

Contents of the document

Abstract 3

Contents of the document 4

Index of figures 6

Index of tables 9

Acronyms and their definitions 10

1. Introduction 12

- **1.1 Research topic** 12
- **1.2 Research problem** 17
- **1.3 Hypotheses** 21
- **1.4 Objectives** 22
- **1.5 Description of chapters** 25

2. State of the art in HVAC energy management 28

- **2.1 Technological advances** 28
- **2.2 Data mining and modeling** 29
- **2.3 Control applications** 30
- **2.4 Supervision and maintenance** 31

3. Power demand forecasting in buildings 34

- **3.1 Introduction** 34
  - 3.1.1 Background and motivation 34
  - 3.1.2 State of the art 35
  - 3.1.3 Innovative contribution 37
- **3.2 Modeling for thermal demand forecasting** 39
  - 3.2.1 Activity indicator modeling 40
  - 3.2.2 Power demand modeling 42
- **3.3 Experimental implementation and validation** 45
  - 3.3.1 Activity indicator modeling 45
  - 3.3.2 Power demand modeling 48
- **3.4 Discussion and conclusions** 56

4. Performance modeling of HVAC equipment 60

- **4.1 Introduction** 60
  - 4.1.1 Background and motivation 60
  - 4.1.2 State of the art 61
  - 4.1.3 Innovative contribution 64
- **4.2 Deep Learning with AutoEncoders** 65

4.2.1 Base autoencoder implementation ... 66
4.2.2 Sparse autoencoder implementation ... 68

4.3 Operational performance modeling methodology ... 71
4.3.1 Unsupervised feature learning ... 72
4.3.2 Supervised target mapping ... 74

4.4 Experimental implementation and validation ... 76
4.4.1 Unsupervised feature learning ... 76
4.4.2 Supervised performance mapping ... 80

4.5 Discussion and conclusions ... 86

5. Predictive control of chiller groups ... 90

5.1 Introduction ... 90
5.1.1 Background and motivation ... 90
5.1.2 State of the art ... 91
5.1.3 Innovative contribution ... 93

5.2 Formulation of the optimization problem ... 94

5.3 Chiller group control methodology ... 96
5.3.1 Operating mode selection ... 96
5.3.2 Predictive control strategy ... 97
5.3.3 Load profile reconfiguration ... 98

5.4 Experimental implementation and validation ... 100
5.4.1 Operating mode selection ... 100
5.4.2 Predictive control strategy ... 103
5.4.3 Load profile reconfiguration ... 105
5.4.4 Energy efficiency improvements ... 106

5.5 Discussion and conclusions ... 108

6. Conclusions and future work ... 110

6.1 Conclusions ... 110

6.2 Future work ... 114

7. Thesis results dissemination ... 118

7.1 Publications: thesis contributions ... 118
7.1.1 Journal publications: ... 118
7.1.2 Conference publications: ... 118

7.2 Publications: collaborations and other works ... 120
7.2.1 Journal publications: ... 120
7.2.2 Conference publications: ... 120

References ... 124

Annex 1. Test environment ... 134

Index of figures

Fig 1.1 Distribution of total energy consumption in the European Union. 13
Fig 1.2 Annual market revenue projection of years 2014-2023 for energy efficient products and services in Europe. 14
Fig 1.3 Simple Demand-Side Management implementation schema. 16
Fig 3.1 Steps of the proposed power demand forecasting methodology divided into activity modeling stage and power demand modeling stage. 39
Fig 3.2 Structure of the closed-loop recurrent neural network, composed by an input layer with memory units, a hidden layer, and an output layer with a feedback loop. .42
Fig 3.3 Activity indicator estimated from the aggregate of the individual occupancy signals during a week in March of 2016 46
Fig 3.4 Performance of the activity indicator model when used for multi-iteration predictions using the validation set. Root mean squared error, RMSE. Mean absolute percentage error, MAPE. Mean absolute error, MAE. Maximum error, MAX. Determination coefficient, R^2. Evaluation time 48
Fig 3.5 Normalized power demand signal drivers. a) Bus impulsion and return temperatures. b) Bus power production and estimated demand...............................49
Fig 3.6 Cross-correlation between each model input candidate and the forecasting target. 50
Fig 3.7 Selected input variables during a period of 5 days in August 51
Fig 3.8 Frequency spectrum comparison between the power demand model input candidates and the power demand target signal...52
Fig 3.9 Comparison between the power demand signal and a prediction obtained using the trained power demand model 53
Fig 3.10 Results of the cross-validation process when splitting the data into 11 subsets, corresponding to the weeks in the dataset. 54
Fig 4.1 Sample architecture of a multiple-layer autoencoder composed of two encoder and two decoder steps. 66
Fig 4.2 Single-layer autoencoder architecture. 67
Fig 4.3 Single-layer autoencoder with sparse-activated neurons. 69
Fig 4.4 Steps of the proposed performance modeling methodology divided into two main stages: unsupervised feature learning and supervised target mapping. 71
Fig 4.5 Coupling of a supervised multiple-layer perceptron to the encoder stage of an unsupervised deep autoencoder. 74

Fig 4.6 Comparison between the reconstruction error and the number of units present in each autoencoder layer with a sparsity of 0.07 and a concurrency of 5................77
Fig 4.7 Representation of the activation of a subset of neurons over the first two components of a principal component analysis..78
Fig 4.8 Comparison between the activation a subset of the eighty hidden units in the first layer and the target production and coefficient of performance. 80
Fig 4.9 Comparison between the real targets (consumption, production, COP) and the deep neural network model for the single machine case. 81
Fig 4.10 Error distribution comparison between a) the proposed deep autoencoder approach and b) a traditional neural network, for each of the outputs: consumption, production and coefficient of performance. 82
Fig 4.11 Comparison between the real targets (consumption, production, COP) and the deep neural network estimation for the multiple machine case...........................83
Fig 4.12 Error distribution comparison between a) the proposed deep autoencoder approach and b) a traditional MLP with GA feature selection, for each of the outputs: consumption, production and coefficient of performance. 85
Fig 5.1 Step diagram of the implementation of the proposed three-stage chiller operation optimization and control framework...96
Fig 5.2 Actual cooling demand compared to the cooling power production resulting from the application of the resulting from the application of the selected sequence of operating modes..101
Fig 5.3 Operating modes of the cooling equipment, obtained from the application of the first stage controller: the instantaneous operating mode selection. 101
Fig 5.4 Impulsion temperature of the distribution bus after the application of the selected sequence of operating modes*, compared to the actual bus temperature during the same period of time...102
Fig 5.5 Power production, consumption and coefficient of performance resulting from the application of the selected sequence of operating modes*, compared to the operation of the base controller during the same period of time. 103
Fig 5.6 Cumulative average performance of a set of setpoint candidates, evaluated over the prediction horizon. 104
Fig 5.7 Evaluation of candidate control sequences that reconfigure the forecasted cooling demand. 105
Fig 5.8 Histogram of the daily average performance achieved by the different control stages, relative to the base controller..106

Fig A1.1 3D representation of the pilot plant, highlighting the location of the HVAC equipment on the building's deck. 134

Index of tables

Table 3.1 Performance comparison between the proposed method (A), and the classical single ANFIS approach with the activity Indicator (B), and the classical single ANFIS approach without the activity indicator (C)..55

Table 4.1 Single machine case: accuracy comparison between the proposed method (A), and a classical NN based implementation (B). 82

Table 4.2 Multiple machines case: accuracy comparison between the proposed method (A), and a classical NN implementation with GA feature selection (B)..........84

Table A1.1 Power characteristics of the HVAC machines of the pilot plant. 135

Table A1.2 Summary of acquired machine signals from the SCADA system. 136

Table A1.3 Summary of acquired space signals from the SCADA system. 136

Table A1.4 Summary of acquired weather signals from the SCADA system. 137

Acronyms and their definitions

ABM	Agent-Based Modeling
AE	Autoencoder
AHU	Air Handling Unit
ANFIS	Adaptive Neuro-Fuzzy Inference System
BEMS	Building Energy Management System
COP	Coefficient of Performance
DSM	Demand-Side Management
DNN	Deep Neural Network
DSS	Decision Support System
FDD	Fault Detection and Diagnosis
GA	Genetic Algorithm
HMM	Hidden Markov Model
HVAC	Heating Ventilating and Air Conditioning
IT	Information Technology
MAE	Mean Average Error
MAPE	Mean Absolute Percentage Error
MAX	Maximum Error
MLP	Multi-Layer Perceptron
NN	Neural Network
OPC	Open Platform Communications
PCA	Principal Component Analysis
PLR	Partial Load Ratio
PSO	Particle Swarm Optimization
R^2	Coefficient of Determination
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
SCADA	Supervisory Control and Data Acquisition
TRNSYS	Transient System Simulation Tool

1.

Introduction

Introduction of the research topic focusing on the research problem that defines the scope of the thesis, the formulation of the hypotheses and finally the specification of the objectives and methodology to accomplish them during the development of the thesis.

CONTENTS:

1.1 Research topic

1.2 Research problem

1.3 Hypotheses

1.4 Objectives

1.5 Description of chapters

1. Introduction

This chapter introduces the research topic focusing on the research problem that defines the scope of the thesis, the formulation of the thesis hypotheses and finally the objectives and methodology to accomplish them during the development of this thesis.

1.1 Research topic

The manner in which we deal with power consumption has changed in recent years. This change has mainly been driven by the continued increased in energy demand, the difficulties associated with remaining competitive as the marketplace becomes global and the threat of instability in securing energy sources [1].

The efficiency of energy processes has been identified as a growing concern at the full range of the power spectrum, encompassing very large power consumers such as factories, medium power consumers such as buildings in the tertiary sector, or even the smaller consumers such as households. Consequently, global efforts are underway directed at raising awareness of the urgency to enable new developments and promoting further research in the field of energy efficiency to tackle current but also upcoming challenges. This is reflected by the weight of related topics in Europe's main research and innovation programs, which contains initiatives like the Horizon 2020 Programme [2].

Being the largest research and innovation program that the European Union has ever conducted, the Horizon 2020 Programme operates with a budget of approximately 80 billion euros in available funding to be allocated between 2014 and 2020 for different topics. Approximately 7.5% of the available funding is directed towards the *Secure, Clean and Efficient Energy* research and innovation track, while 4% is directed towards the *Climate Actions, Environment and Resource Efficiency* track. Both this resource allocation and the proclaimed aim of the *Energy Efficiency* track of the Programme of reducing the primary energy consumption by 20% by 2020, showcases the commitment of the European Union to this goal [3].

The considered avenues for action include the realization of transformational changes at all the stages of the energy consumption chain, from generation to transformation and distribution, but having a special focus on the final stage of the energy chain: the consumption of energy, where the buildings and the transportation sector are highlighted. It is precisely at the last stage of the energy chain, also called the demand side, where the greatest potential for energy savings is attributed. This is because there has been a continued growth in power demand, so the best way best way for increasing the overall energy efficiency is considered to be addressing the reduction of the this demand.

Regarding the allocation of the global energy demand, there have been several studies that show how it is distributed by sectors. These show that the energy consumption in buildings, including the residential, commercial and public sector buildings, accounts on average for 40% of the worldwide energy consumption, and 50% of the total electric consumption [4,5].

The distribution of the energy consumption in the European Union is shown in Fig 1.1. As it can be observed it presents a consumption share in buildings similar to the worldwide average. However, these figures can become larger on different countries, for example in the United States up to 70% of the consumed electric energy could be associated with buildings, where half of that amount is attributed to commercial buildings and is projected to keep increasing even further respective to buildings in the residential, industrial or transportation sectors [6,7].

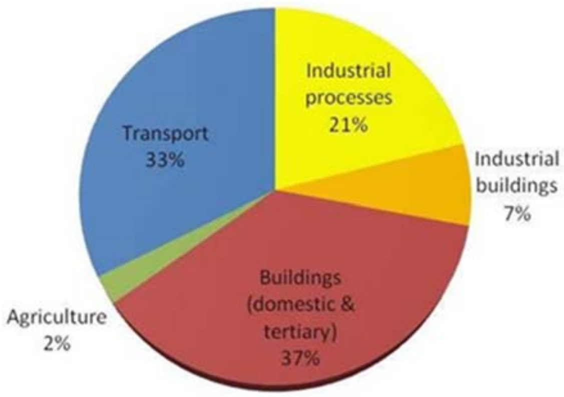


Fig 1.1 *Distribution of total energy consumption in the European Union[1].*

Nevertheless, recent studies point out that research focused on this topic could offer a great improvement potential specifically in buildings due to technological advances developed in the last two decades in a variety of areas, including engineering and materials science, but also in other areas such as in computer and data science. It is

[1] European Commission, "EU Energy and Transport in figures 2010. Statistical pocket book," Publications Office of the European Union, 2010.

estimated that by researching and implementing energy saving strategies, it could generally be possible to achieve energy savings in the range from 10% to 30% [8,9].

Aligned with this estimate, Fig 1.2 shows the projected trend in annual market revenue for the sector of energy efficient products and services in Europe, which is expected to keep growing in the coming years, responding to a demand of solutions for tackling the energy efficiency problem.

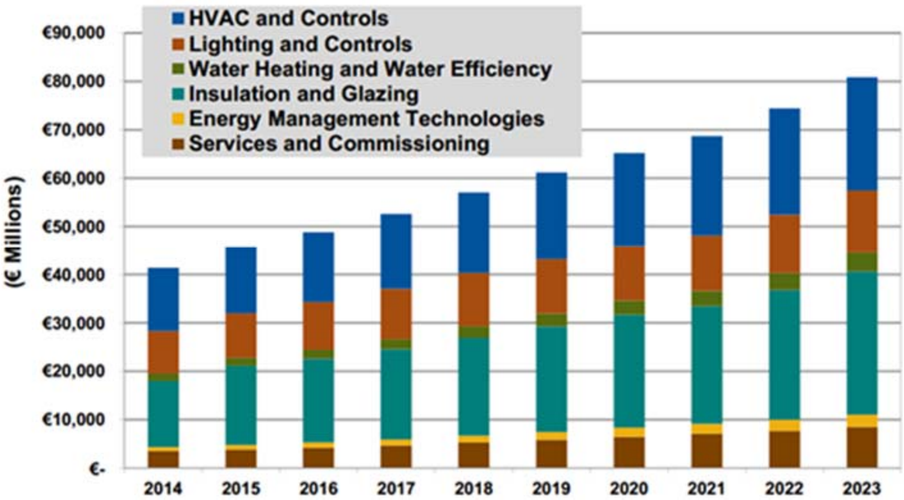


Fig 1.2 *Annual market revenue projection of years 2014-2023 for energy efficient products and services in Europe[2].*

The four major groups that are responsible for the energy consumption in modern, mixed-use, buildings are typically the Heating, Ventilating and Air Conditioning (HVAC) system, lighting system, IT equipment and assorted electrical plug-in equipment [6,10].

Out of these four groups, HVAC systems are comparatively the largest consumer, accounting for a share in the range of 30% to 50% of the building's total energy consumption [11,12]. Thus, HVAC systems may account for 10% to 20% of the world's total energy consumption. However, HVAC systems are also one of the groups that currently offer the greatest unrealized potential for improvement.

Current research has established that by closely monitoring the operation of HVAC systems and by improving HVAC control strategies, it could be possible to detect energy waste, correct energy misuse circumstances and to prevent their occurrence.

[2] Navigant Research, "Spending on Energy Efficient Buildings in Europe is expected to total nearly 800 billion from 2014 through 2023", Building innovations – Energy Efficient Buildings, 2014.

Combined, these actions may achieve a potential energy savings of up to 40% in HVAC systems in buildings [13].

This large savings potential is due to the fact that practically all HVAC facilities have hidden efficiency problems that difficult the efficient use of energy. The most common of these problems are i) misconfigured controls such as inefficient system configurations relating to incorrect operating setpoints or overall inefficient operating strategies, or ii) faulty equipment caused by compressed air leakages, dirty air filters, aging-related problems or others. Furthermore, such problems are commonly neglected unless dedicated tools for power demand management and energy efficiency analysis on the demand side are employed [14]. The framework that encompasses these and other tools is called Demand-Side Management (DSM).

The DSM topic includes a wide range of applications and methodologies, which can include from behavior-modifying strategies like the implementation of smart power tariffs aimed at incentivizing certain consumption patterns, to load shifting strategies than focus on peak shedding to prevent grid overloads, but also control applications aimed at increasing the efficient use of resources on the demand side. Amidst the different DSM approaches, it becomes evident that the most advantageous from the point of view of tackling the problem of reducing the use of resources, are those aimed increasing the energy efficiency of consumer-side processes. Methodologies aimed at load shifting, for example, are appropriate for increasing consumption flexibility, thus solving grid stability or production cost issues, but have little to no effect in terms of amount of energy consumed.

The application of different flavors of methods within the DSM framework for the purpose of managing energy demand and energy efficiency in buildings has led to the concept of Smart Buildings, which can be considered as regular buildings with a layer of control, supervision and maintenance solutions built into the Energy Management systems that operate them [15].

The idea of Smart Buildings is based on a dual concept, i) the comprehensive sensing and monitoring of the events within the building, including the monitoring of all relevant equipment and technical systems, but also other factors such as the local weather conditions and occupant behavior, and ii) intelligent actuation strategies that interpret and consider the high-level information provided by the aggregated low-level sensor signals and are able to detect inefficient behaviors and implement actions aimed at correcting these inefficiencies but also to further increase efficiency where possible while maintaining occupant comfort [16]. According to the state of the art in building operation, this kind of solution has a significant potential for energy savings in buildings when properly tuned and controlled [17–19].

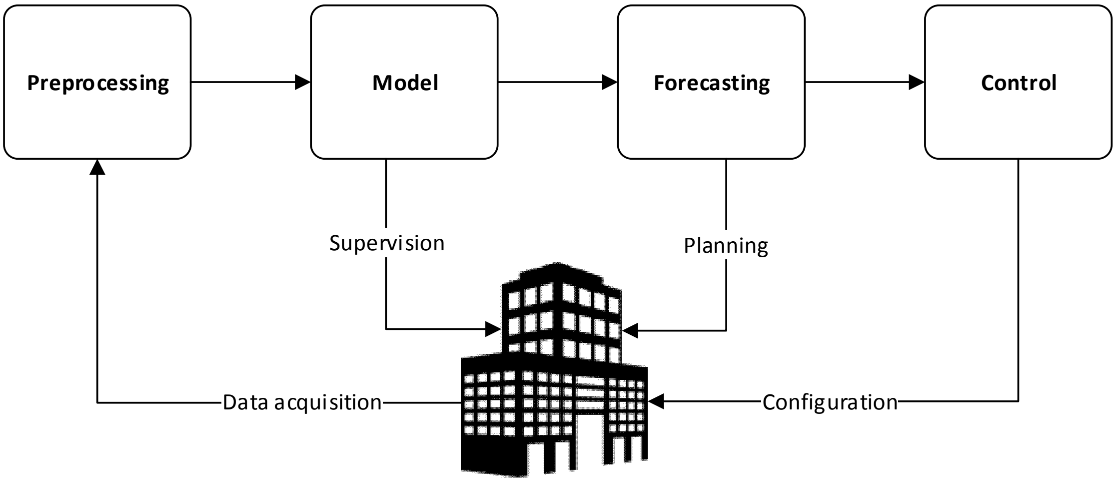


Fig 1.3 *Simple Demand-Side Management implementation schema.*

A simple schema depicting the implementation of a DSM application for the operation of a building is shown in Fig 1.3, which includes the most common blocks of recently proposed approaches to sweeping reform of current EMS in buildings: i) the preprocessing of acquired data from building sensors to extract knowledge from the operating state, ii) the modeling of technical processes to provide a baseline for both supervision and implementation of control solutions, iii) the forecasting of energy consumption patterns to allow to efficiently plan the use of resources, and iv) the implementation of control strategies that take advantage of the increased context awareness regarding the current operating state, the modelled technical systems, and the forecasted consumption patterns to determine an operation strategy that maximizes the overall energy efficiency of the building.

In this framework, this research thesis defines a path towards the development of novel strategies for supporting and achieving an increase of energy efficiency by means of the advancement of the state of art in HVAC system management for the advancement of Smart Buildings.

1.2 Research problem

The trend towards including more and more instrumentation systems in buildings has been increasing during the recent years. The instrumented systems include the various HVAC equipment such as chillers, pumps or air handling units, but also other aspects such as the indoor room temperatures and humidity, other state variables such as occupancy, light levels, outdoor weather conditions and in some cases solar irradiance sensors when buildings have a large surface of windows [20–22].

Moreover, there has been an increase in the incorporation of field data acquisition systems using standardized building automation systems [23]. Some examples include BACnet, Modbus, KNX, LonWorks, OPC, among others. These have allowed the gathering and historical accumulation of a huge quantity of operational information made traditionally available to the building's EMS or SCADA systems. This increased visibility into how the systems function allows to monitor their operation in a more detailed manner and encourages the study of their performance [24–26].

This increase of the in-depth visibility into the technical systems in buildings laid the foundation of Building Energy Management Systems (BEMS), which consist of computer-based tools that may integrate DSM techniques into software packages to support the management, supervision and control of a building's technical systems, from the point of view of managing the energy usage required to allow the building to fulfill its purpose [15,27]. BEMS can support building managers by providing insight and tools to understand the energy usage of the building and to properly control and further improve the energy performance of the building without degrading the thermal comfort of the occupants [28].

The scientific interest on this topic has grown during recent years, as evidenced by the increase in the number of studies addressing the incorporation of tools for DSM within the context of the Smart Building. These have undoubtedly been motivated by these elevated monitoring and auditing capabilities found in modern BEMS. The derived gain in insight regarding the operation of the different technical systems has accelerated the research and development of solutions that are aware of the operating context of the building and are able to distill actionable knowledge that may help improve energy efficiency by a significant degree [29].

Nevertheless, limited tools are available to building engineers or building energy managers to actually analyze the energy usage patterns in relation to the operating state of the equipment and behavior of the occupants, and to actually carry out improvements. BEMS solutions currently available require being operated manually by expert users and offer limited capabilities. Furthermore, recent surveys have highlighted that the state of the art in BEMS for control of energy systems at the consumer side present different shortcomings, therefore being a path with significant potential for further research and development [30].

Classical building management is focused on monitoring and controlling the indoor environment, mainly the temperature and humidity, by means of performing corrective actions when these deviate from predefined value ranges. However, this concept is shifting towards more advanced applications that take advantage of a more detailed view of the operating context of the building to select a more adequate actuation strategy based on this view [31]. In particular, the current trends in BEMS for modern buildings are contemplating the embedding of advanced monitoring systems capable of observing the environment and operating conditions, potentially permitting an intelligent management system to understand their state and rapidly focus on correcting and preventing energy inefficient behaviors. Consequently, further research is being carried out to develop data processing algorithms capable of extracting augmented information from large historical databases existing in current BEMS with the objective of extracting knowledge to support the creation of higher-level controllers [32].

From the perspective of monitoring the operation of the HVAC systems in buildings, the potential to identify inefficiencies and anomalous behavior by analyzing historical energy consumption data is often not fulfilled [17]. An approach to data analysis of the acquired energy consumption patterns must be effective and should be able to identify anomalous behavior, deviations from proper operation and to establish the desired baseline operation, yet it must be reliable, robust and must minimize the amount of expert knowledge required to properly operate it in order to achieve the expected impact by facilitating its widespread adoption.

The information extracted by such approaches in HVAC systems can be a very valuable resource for example to alert the managers of a building regarding unexpected deviations from nominal operation with limited delay [33]. Furthermore, the detected anomalies but also the observation of trends in the consumption patterns may be used to estimate wear or faults in equipment, and may also play a role in scheduling maintenance operations by being a warning sign of potential equipment fault [34].

Nevertheless, both the complexity and the vast amounts of acquired historical data make it difficult to establish the baseline for nominal operating state, thus causing the classification of abnormal behavior to become challenging and identifying resolution actions impractical [35].

> **Having an effective means of establishing or modeling the relationship between the operating state and the energy consumption and efficiency of an HVAC facility becomes an essential milestone. Such a model may be employed to detect inefficient or anomalous behavior in the energy consumption patterns and to introduce appropriate corrective or preemptive actions to overcome them.**

On the operation side, this extracted information may also be employed to adjust the control strategies of HVAC equipment, regulating and scheduling their operation by considering the operational state of the building and avoiding inefficient operating ranges. However, most related research focuses on saving energy by implementing room temperature control solutions without considering the global operating state of the facility, its dynamics and relationships between equipment operating conditions and achieved energy performance [30,36]. Furthermore, part of the studies implement on/off controllers that operate equipment by reacting to the real-time demand by the users, but do not consider the historical operating patterns of energy demand and how a given control strategy may affect the efficiency of the HVAC system [37]. Some studies have focused on the implementation of scheduling actions for HVAC equipment for temperature control, but few have approached the link between these actions and the energy efficiency side of the problem, thus in-depth investigations regarding the optimal control of HVAC equipment are not available in the literature. As outlined in a review of integrated control of HVAC systems, possible reasons might be that 1) it is difficult to establish models that offer an acceptable tradeoff between accuracy and simplicity, making them convenient for optimization problems; 2) models are difficult to calibrate and; 3) the interactions and coupling between equipment may make it difficult and time consuming to search for the optimal or effective control strategy [12]. Therefore, there is a gap in the state of the art of BEMS control:

> **Control strategies that consider the operational state of the equipment and the dynamics of the load demand of the building, conditioned by its operating environment and comfort constraints, are well positioned to fulfill significant potential savings by adapting the operation of the**

system according to its most efficient configuration while satisfying the demand.

In summary, considering the previously mentioned problems and limitations of current BEMS (relationship between state of operation and energy consumption and efficiency, performance supervision for maintenance purposes and control of the HVAC equipment of the building to achieve optimal energy efficiency configurations) further research is necessary with the objective of proposing a solid framework of analysis and actuation in order to tackle these issues and to strive towards greater energy efficiency. This framework shall be composed of novel data processing algorithms able to model the energy consumption and efficiency of the system regarding the state of operation of the equipment and to steer their operation towards the fulfilling the performance increase potential.

1.3 Hypotheses

In order to address the described research problems, the following hypotheses were formulated as a starting point for this research work:

- The performance of existing modelling methodologies for load forecasting can be improved by providing them with perception and knowledge about the internal state of the plant.

- The internal states and conditions of the plant under analysis can be described by the variables which define the operation of the processes therein.

- If the internal states of the plant are not directly observable through the variables exposed by the monitoring system, they can be estimated by means of data mining and data fusion techniques.

- A modelling methodology with the aim to aid the solution of the supervision and control problem shall consider the state of the installation and its equipment of predominant influence to the energy consumption and efficiency of the whole system.

- The correlation between the internal states of the plant, the actuations of the control system and the energy consumption shall allow to classify operating ranges depending on their efficiency and to steer towards high efficiency modes in control applications.

In conclusion,

- The implementation of advanced HVAC energy management strategies in the context of the smart building shall allow to identify the actual conditions of the underling subsystems in order to reach an optimal comfort-energy consumption tradeoff trough the introduction of models and control schemes by taking advantage of the wealth of historical operation information available.

1.4 Objectives

The objective of this thesis is to progress the HVAC energy management state of the art in the context of the Smart Building by the proposal of an integral information management methodology and the definition of a framework to address the complex problem of increasing their energy efficiency.

Specifically, the three main work areas of this thesis are declared as follows:

- Power demand forecasting: the development of a load modeling and forecasting methodology specifically tailored for predicting the power demand in buildings, considering operating state variables such as weather conditions or occupancy of the building.

- Operating performance modeling: the development of a generic methodology for modeling the energy characteristics of HVAC equipment, in particular their energy consumption, expected production and coefficient of performance, as a function of their operating state and control setpoints.

- HVAC production equipment control: the development of a framework for implementing the control of HVAC production installations, coordinating the operation of different equipment and considering their estimated operating performance and future demand of the building.

To successfully accomplish these contributions, the following specific objectives shall be achieved:

- The selection and extraction of information in building databases for the purpose of obtaining useful knowledge regarding their operation and internal states of the building and its equipment.

- The proposal of new approaches on the topic of operational state estimation and correlation with energy demand and efficiency.

- The proposal of a novel methodology and algorithms, for demand and efficiency modelling and forecasting, considering the operational state of the plant instead of heavily relying on historical data.

- The proposal of new approaches for multi-objective and predictive control strategies for optimal point of operation based on the steering towards efficient operation modes, which could take into account the foreseen behavior to regulate the aggressiveness of the control.

- The validation of the complete analysis and actuation framework by means of building operation databases acquired during the research period from a pilot plant.

In order to allow the research, development of the contributions and validation of the outcomes of the declared objectives, a test environment was selected which includes the most common components in prevailing HVAC installations in modern buildings. The chosen test environment is a complete building that acts as a pilot plant and was fitted with additional instrumentation in order to acquire the relevant data and allow the validation of the proposed energy management tools. A description of the pilot plant is provided in *Annex 1. Test environment*.

The methodology employed for meeting the declared objectives is defined as follows:

- Detailed literature review of the state of the art related to energy management systems in buildings, specifically focused on building energy demand and HVAC equipment operation.

- Definition of the research problem and objectives, proposal of a methodology to accomplish them and identification of a pilot plant to act as a test environment.

- Instrumentation of the pilot plant and execution of data acquisition campaign to build a database of the historical behavior relating to i) the HVAC equipment in terms of operating modes and energetic performance, and ii) the building's energy demand and operating context.

- Research, development and implementation of novel techniques focused on solving the identified problems and expanding the associated state of the art.

- Evaluation of the attained novel techniques and comparison of their performance and trade-offs with state of the art solutions by means of implementation using experimental data.

- Analysis and discussion of the research outcomes, drafting of conclusions and dissemination of the methods and results on indexed peer-reviewed journals and notable conferences.

1.5 Description of chapters

This section describes the content of each of the remaining chapters in this document.

A general literature review of energy management approaches for HVAC systems in buildings is presented in chapter ***2. State of the art in HVAC energy management***. This chapter outlines the main types of solutions being currently employed on modern buildings and actively being researched for further increasing the efficacy of BEMS and for increasing energy efficiency while ensuring occupant comfort.

In chapter ***3. Power demand forecasting in buildings***, a study of the main factors influencing HVAC power demand is presented, leading to the development of a methodology aimed at creating forecasting models capable of accurately predicting the short-term thermal needs of a building considering factors such as the weather or the occupancy patterns.

In chapter ***4. Performance modeling of HVAC equipment***, the most employed methods for establishing the relationship between operating state, control actions and energy performance are reviewed and their drawbacks identified in order to introduce a novel generic method for modeling HVAC equipment based on a deep learning approach.

In chapter ***5. Predictive control of chiller groups***, the state of the art relating to solving the optimal chiller loading and sequencing problems is critically reviewed in order to highlight the potential avenues for improvement, and a novel control strategy is developed by taking advantage of the previous thermal demand forecasting model and equipment performance model to implement a multi-objective predictive control solution.

The general conclusions of the research work carried out in this thesis and the possible future work are presented in chapter ***6. Conclusions and future work***.

A summary of the publications derived from this work is included in chapter ***7. Thesis results dissemination***, including also collaborations in related research projects.

Finally, the pilot plant used as a test environment for the development of this thesis is described in ***Annex 1. Test environment***.

2.

State of the art in HVAC energy management

This chapter outlines the main types of solutions being currently employed on modern buildings or actively being researched for further improving the energy efficiency of HVAC systems in buildings while ensuring occupant comfort.

CONTENTS:

2.1 Technological advances

2.2 Data mining and modeling

2.3 Control applications

2.4 Supervision and maintenance

2. State of the art in HVAC energy management

The magnitude of the worldwide energy demand in buildings for the operation of HVAC systems, the increasing sustainable energy production challenges and the impact of potential improvements have attracted the attention of ample research efforts. Consequently, energy management of HVAC systems has grown to be a very wide topic, and ongoing efforts are directed at multiple areas and tackle different problems, like HVAC machine technology, data mining and modeling, control or supervision and maintenance.

This chapter outlines the main types of solutions being currently employed on modern buildings and actively being researched for further improving the energy efficiency of HVAC systems in buildings while ensuring occupant comfort.

2.1 Technological advances

In order to achieve better energy efficiency on new installations, one of the solutions is to investigate new technical systems from the point of view of HVAC components in order to replace current systems with technologically more advanced versions offering better characteristics. A review of strategies for saving energy in HVAC systems presented some of the paths for further developing air conditioning technologies, but highlighted that local weather conditions may have a significant impact when considering the types of HVAC technologies [38]. For example in the case of vapor compression systems, where better COP can be achieved by using water-cooled systems rather than air-cooled, this is contingent on whether installing a cooling tower might be feasible due to the local relative humidity, which may impact the performance of entire system. Furthermore, even though investing on new HVAC technologies is a worthwhile effort, the bulk of the benefits will be observed when these become cost-effective and begin to be widely installed in new HVAC facilities, thus there is still a need for solutions to help increase the efficiency of existing installations.

Retrofitting of existing installations is another avenue for increasing energy efficiency of HVAC systems, for example by replacing inefficient equipment to introduce improved versions, or by replacing machines or components that have deteriorated due to aging, presenting lower performance and possibly originating faults. Nevertheless, there is still a need to help identify the machines within an installation that could provide the best return on investment when considering the retrofitting of equipment [39]. Helping to determine what equipment should be replaced and when to do it in order to maximize savings could be supported by data mining processes that analyze the state of operation of HVAC installations [40].

Similarly, a recent practical case approached the implementation of a case-based reasoning solution for supporting the decision making process on new retrofitting projects. However, instead of focusing on the actual target building, the study started with the collection of successful retrofitting cases and analyzing the attributes of the involved buildings, energy characteristics and cost parameters, then carrying out a weight optimization of the decision attributes by identifying similarities to existing cases [41].

2.2 Data mining and modeling

Approaches that include data mining processes could help take advantage of the large amounts of operational data available in existing BEMS. A recent review of the potential of data mining strategies for energy efficiency enhancement in buildings pointed out that extracting insights from BEMS databases could provide awareness into avenues for implementing energy saving actions, and specifically underlined the capacity of unsupervised methods to extract knowledge from large datasets [42]. There are ongoing efforts to achieve the improvement potential reiterated by different reports. For example another review of recent advances and current challenges being faced in building engineering revealed the areas that are being currently targeted by researchers by means of data mining applications. These include the prediction of energy demand, the effects of building occupants and their behavior, the modeling of buildings for integration in optimal control, and fault detection and diagnosis applications [43].

A study of the operational signatures of HVAC systems attempted to evaluate the characteristics of the equipment available in an installation and match their values with the energy consumption patterns, focusing on the identification of what parameters caused the different behaviors observed in the energy usage [44]. The outcome included the automation of the selection of custom configuration for a group of air handling units in a building.

Indeed, the usage of data mining systems for supporting the selection of settings is a common approach. A project participating in the FP7 Framework of the H2020

Programme focused on the development of a decision support system based on data mining methods that allowed the discovery of rules from different data sources including weather, energy usage and energy prices, with the objective of helping facility managers by producing future action suggestions in the short term [45]. However, the application of the generated suggestions is still left up to the manager, and further research into automated solutions could help close the loop for including them in control systems.

However, not all data mining and modeling solutions are suitable for control applications. A review of modeling methods for HVAC systems in buildings specifically aimed at supporting control applications discussed that the modeling of these systems is a current unresolved problem due to difficulties caused by the characteristics of HVAC systems, including non-linearity [46]. A comprehensive taxonomy of the existing approaches is included in the review, providing recommendations for consideration when selecting a type of model and reiterating that implementing a modeling method for HVAC processes that matches the case is concern of upmost importance. Some of the recommendations include avoiding the over-simplification of processes, the validation in real systems and the consideration of the ease of usage of a method besides the features it provides.

2.3 Control applications

Regarding the implementation of control applications, the development of effective strategies specifically for HVAC system applications is a crucial aspect of improving energy efficiency in buildings [47]. However, these present unique challenges due to non-linear behavior, time-varying disturbances and interactions between systems, among others. Fortunately, the costs of data storage have been steadily decreasing while processing capacity has been increasing, which makes it possible to create more advanced control strategies with increased knowledge of the operating behavior of the installation extracted from historical data, and implementing more complex optimization methods. In particular, the integration of predictive solutions in the control process, the usage of multi-objective cost functions and the development of advanced control algorithms could be some of the more promising features of modern control solutions [48].

For example, a recent experimental study implementing a model-predictive controller for an academic building considered some of these aspects to reduce the energy consumption of the building while maintaining occupant comfort [49]. For the evaluation of control actions, the study employed a calibrated EnergyPlus model of the building but had to resort to a simplified model in order to overcome computational constraints, falling back to a random forest regression to estimate the average indoor temperature and aggregated consumption of the whole building for the next control iteration. Regarding the optimization method, a genetic algorithm was used for calculating permutations of the predictive control setpoint for the next few iterations. Solutions such as this one commonly achieve the reduction of energy consumption successfully at the expense of occupant comfort, which could become compromised because there is a direct tradeoff between the thermal energy delivered to a space and the comfort conditions within [50].

Besides the usage of models appropriate for control and the adoption of algorithms better suited for control applications than generic global optimization methods like the genetic algorithm, focusing on other aspects of the building's operation could also provide savings without sacrificing occupant comfort. For example a solution not aimed towards the minimization of the energy used for space conditioning, but instead targeting the optimal utilization of the HVAC equipment to meet the required energy demand while minimizing the consumption of the equipment could be worth investigating [51].

Accordingly, another investigation topic corresponds to the optimal chiller loading and sequencing, which typically has no effect on the comfort due to being aimed at controlling the operation of the machines to maximize their performance, rather than minimizing energy demand.

2.4 Supervision and maintenance

Another research area within energy management consists on the detection and diagnostic of faults to aid in the supervision and maintenance of HVAC installations. Considerable studies focus on the development and implementation of automated fault detection and diagnosis (FDD) solutions for integration in BEMS, mainly due to existing research having documented that a large portion of HVAC equipment present in existing buildings may require repairs or contain faults, and that their early detection and rectification could help to considerably reduce the amount of energy waste [52].

FDD can provide a solution to this problem, as it consists on the automated detection of faults in equipment and the diagnosis of the component causing them. This is typically implemented in one of two ways, either by using supervised learning when the failure modes are known and there is sufficient data available to classify operation samples to determine if they match a known fault, or by using unsupervised or novelty detection solutions when this is not the case. Indeed, when the failure modes are not known, or there is no data to compare them, it is still possible to implement fault detection for example by monitoring whether the system operates within known ranges or by defining and tracking the value of specific features [53].

An application of FDD is shown in a study consisting on the selection of the main features for allowing the detection and classification of faults with sufficient accuracy in a case study where the failure modes are well documented, having a focus on minimizing the amount of sensors required, thus reducing costs, and employing a multi-class support vector machine for classification [54]. Alternatively, another study used an array of nonlinear repressors for estimating the behavior of the HVAC system, implementing a Gaussian model of the process and using the variance of the prediction error to classify whether the behavior had been observed or the system is operating outside of known ranges [55]. These studies represent two very different approaches for implementing FDD, but both are based on the idea that finding and fixing faults shall lead to the reduction of energy waste, thus increasing the efficiency of the HVAC installation.

In conclusion, multiple avenues for improving the energy efficiency are being explored and significant efforts are underway to develop novel solutions that tackle this issue from different perspectives. Current research effort targets a broad range of studies, from research on technical systems and HVAC equipment technologies, to artificial intelligence-based management solutions. However, it is likely that not one solution will solve the problem and that future building energy management systems will need to adopt a combination of these solutions to close the performance gap between ideal energetic performance and performance achieved in practice.

3.

Power demand forecasting in buildings

Load forecasting is an important topic in the context of building energy management due to the numerous applications it facilitates, including control, planning and supervision. This chapter studies the main factors influencing HVAC power demand with the objective of proposing and developing a methodology aimed at creating forecasting models capable of accurately predicting the short-term thermal needs of a building.

CONTENTS:

3.1 Introduction

3.2 Modeling for thermal demand forecasting

3.3 Experimental implementation and validation

3.4 Discussion and conclusions

3. Power demand forecasting in buildings

This chapter describes the role of power demand forecasting in buildings and carries out a study of the main factors influencing HVAC power demand, leading to the proposal and development of a methodology aimed at creating forecasting models capable of accurately predicting the short-term thermal needs of a building considering factors such as the weather or the occupancy patterns.

3.1 Introduction

This section introduces the background and motivation for pursuing this line of research, reviews the state of the art related to load forecasting applied to HVAC systems in buildings and describes the innovative contributions of this work.

3.1.1 Background and motivation

As previously discussed, recent advances in the functionalities of modern BEMS in terms of monitoring and supervision [13,56] have paved the way in the framework of smart buildings for the introduction of DSM practices [57], which are one of the most important methods for achieving energy savings [58]. The increased insight derived from this progress has been instrumental in the further study of context-aware solutions that are capable of improving the energy efficiency of technical services in BEMS by building on the expanded knowledge available [59]. By accounting for up to 40% of the power consumed in buildings, heating ventilating and air conditioning (HVAC) systems, in particular, have attracted a substantial share of current research efforts [6,12].

In modern buildings, load modeling and forecasting methodologies able to predict the future power demand of HVAC systems are an important concern of installation managers due to the useful knowledge that they provide [60], since real-time demand information plays a role in mitigating energy waste [61]. Several types of methodologies exist, being data-driven approaches the most prevalent. However, when applied to HVAC systems, these methodologies are mostly aimed at forecasting the consumption load [62]. Instead, focusing on the thermal power demand may help abstract from performance differences caused by regulation systems and to better reflect the power needs of the facility. Automation systems can benefit from this information in order to make decisions autonomously by following energy-saving optimization strategies. This is especially true for the control of HVAC equipment, where the predicted load could be used for implementing model-predictive control strategies. Multiple control approaches applied to HVAC systems that could benefit from this information can be found in the literature, such as the planning of energy storage during off-peak periods using cooling storage systems [63]. Others also include the planning of adequate startup and shutdown times for heating and cooling equipment in order to save energy by meeting the right amount of power demands, and for the orchestration of machine actuations in installations where multiple machines are available [64]. Furthermore, the combination of HVAC load forecasting with equipment efficiency maps represents an underexploited avenue of improvement with a high potential for the optimization of the operation of the system. That is, the demand anticipation and the utilization of the most suitable machine for each situation would provide a positive affectation to the overall equipment's performance, which is a significant present-day problem in building management and maintenance. Indeed, the overall efficiency of the installation could be improved, since the current most common method for allocation HVAC capacity is based on setting the same water temperature thresholds on all the available machines [65].

Even though this framework represents one of the main current research interests stated by the related scientific community, the obstacles to its implementation are double-sided. First, the efficiency maps are difficult to obtain when precision beyond the manufacturer's sparse figures is desired, as they would require extensive testing of the unit in each installation, and would likely drift over time as the equipment deteriorates with aging. Secondly, the methodologies for obtaining load predictions in HVAC systems are not mature enough and their implementation can be quite challenging due to the potential complexity of energy systems [66].

3.1.2 State of the art

In the recent literature, considerable scientific effort has been committed to the research of load forecasting algorithms and methodologies, as seen in the latest review papers [67]. A comprehensive review of more than one hundred papers on electrical load forecasting defined a general taxonomy for selecting modeling algorithms from the point of view of their popularity in different applications, indicating that data-driven approaches are mainly used in short-term forecasting applications due to their complex dynamics [68]. In contrast, a comparative analysis studied eleven modeling algorithms from the point of view of their performance when applied to the same dataset, revealing their applicability in different scenarios including cases with limited data or high variability [69]. However, even though numerous general-purpose approaches exist for the implementation of load forecasting, their limitations are revealed when applied to real HVAC systems as opposed to simulations or synthetic data. These limitations are mainly related to the difficulty of adapting the predictions to the power demand changes caused by fluctuations of influencing parameters, such as the weather and the occupant's behavior during the day [66].

In this regard, recent studies as the one presented in [70], confirm the significant correlation between the occupancy of the building's spaces and the HVAC equipment's actuations and consequent operational regime changes. This, as promoted by different authors, for example in [71], indicates that the occupancy should be a key aspect in the research of energy usage in buildings, because of its potential contributions to efficiency improvements. Actually, a recent review of energy efficient ventilation strategies concluded that large amounts of energy are being wasted because of conditioning building areas that have effectively empty periods of time, and that accounting for these may help to greatly increase efficiency [72]. Indeed, most of the current load simulation and forecasting methodologies show a lack of occupancy awareness, while the available studies dealing with the integration of occupancy data into load forecasting systems to enhance the accuracy of power demand predictions present critical limitations and insufficient proficiency [73].

Similarly, a recent review of artificial intelligence methods for load forecasting in buildings suggested that the integration of occupancy data has the potential for improving energy predictions [74]. Moreover, it was stated in a study of the application of neural networks for building energy forecasting, that occupancy-based inputs should be taken into consideration in future studies because of the impact that the occupancy can have on the building's thermal energy usage. This is shown in [75] and further developed in [76], where several attributes were studied, concluding that it would be useful to create occupancy indicators for improving the prediction capabilities.

On this subject, some methodologies for the modeling and forecasting of occupancy in buildings exist, being Agent-Based Modeling (ABM), and Hidden Markov Models (HMMs) the most common. ABM approaches try to mimic the behavior of occupants of a building in order to simulate either occupancy patterns or their effects at the occupant level [77], hence being too fine-grained for full building applications. Alternately, HMMs are stochastic processes that naturally fit the problem of modeling occupancy patterns, because they treat occupancy as a series of transitions between states and attempt to estimate and simulate the probabilities of transitions among such states [78]. HMMs are useful at low aggregation levels, for example for assessing the probability of a given space becoming occupied, but are not a good fit for big scenarios, as the complexity grows exponentially with the number of zones [79]. Another disadvantage of HMMs at high aggregation levels is that their future state is a function of their current state, not taking into consideration past states. This property could neglect important features of the aggregated occupancy, such as the ratio of change. Indeed, complete and viable solutions are yet to be investigated, and the proper way to monitor the occupancy, to define the indicators and to integrate them into a load forecasting system remains to be established.

3.1.3 Innovative contribution

In this chapter, an HVAC thermal power demand forecasting methodology composed by the integration of a power demand model and an activity indicator model is studied. The methodology aims to extract the occupancy patterns in order to determine the level of activity in the building and thus to improve the accuracy of the power demand forecasting. With this objective, the building's historical database is divided into occupancy and load data for separate preprocessing. Then, an activity indicator is built and a model is implemented using Recurrent Neural Networks (RNN) to enhance the consideration of dynamic temporal patterns, while the power demand characterization is carried out by means of a state-of-the-art Adaptive Neuro-Fuzzy Inference System (ANFIS) structure. Finally, a reliable and robust power demand forecasting model is obtained by the serialized fusion of both inference systems.

The main contribution of this study lies in a new data-driven short-term load forecasting methodology for the prediction of the thermal power demand of HVAC systems in buildings, and the introduction and verification of an activity indicator estimation procedure to support the prediction of the power demand.

Aligned with the current research challenges in the field, the methodology takes advantage of real-time occupancy data in order to predict an activity indicator, providing accurate insight regarding the thermal needs of the building in terms of the volume of consumption endpoints in operation. Furthermore, due to the difficulty in directly measuring the thermal power demand signal, which would involve the use of extensive instrumentation installed in consumption endpoints throughout the building, an estimation method is proposed in order to calculate the actual power draw, derived from the measurement of the thermal power output of the HVAC energy production equipment in the building.

The novelty of this work includes the implementation of a new hybrid solution that offers major advantages over traditional approaches. In particular, the collaborative model structure, comprehending the separate modeling of the activity indicator’s dynamics and the thermal power demand characterization, differs from classical single model approaches in that it allows the selection, tuning and fitting of each structure independently, increasing its adaptability to the dynamics of each signal and improving the resulting accuracy through the specialization of its modeling process.

It should be noted that this is the first time that this methodology, as well as this activity indicator modeling, is used in building automation and energy management for providing accurate insight regarding the thermal needs of the building, with the objective of supporting the enhancement of resource management and the optimization of the operation of local HVAC equipment.

3.2 Modeling for thermal demand forecasting

A step-by-step diagram of the complete methodology is shown in Fig 3.1, which is divided into three stages: the activity indicator modeling stage, where an artificial activity indicator is defined and modelled, the power demand modeling stage, where the power demand of the HVAC system is estimated and modelled separately and finally the demand forecasting stage, where predictions are obtained by means of the evaluation of the models.

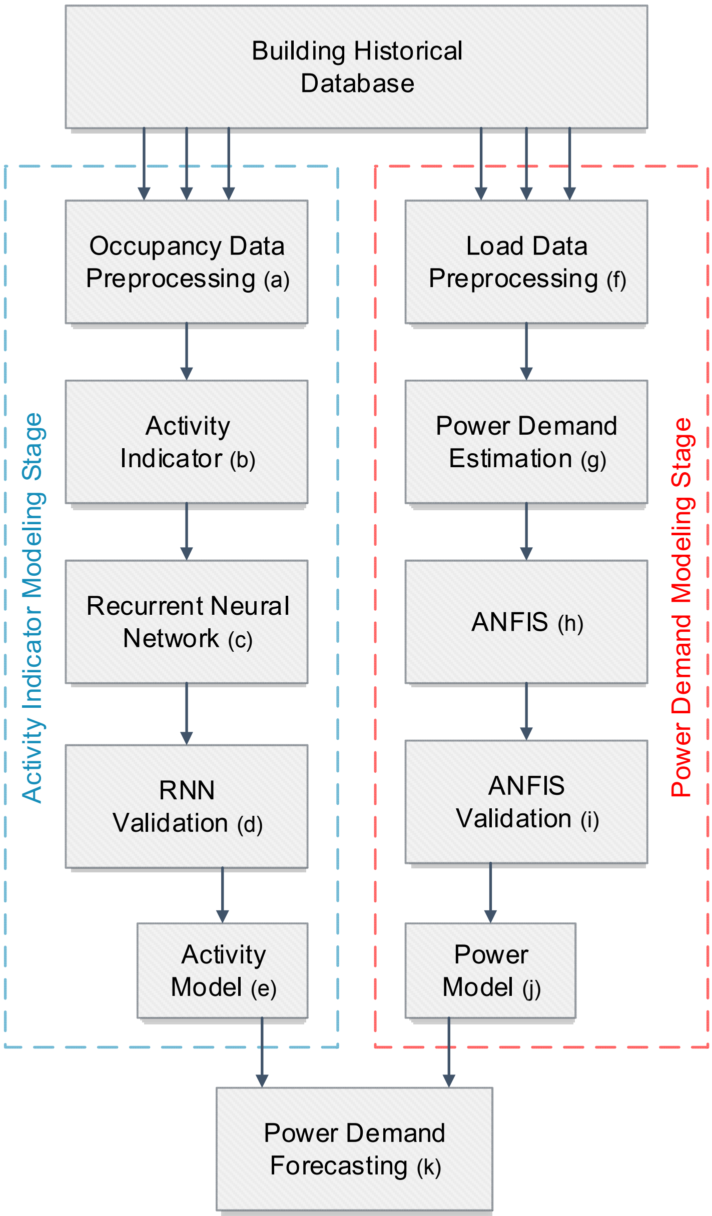


Fig 3.1 *Steps of the proposed power demand forecasting methodology divided into activity modeling stage and power demand modeling stage.*

Initially, on the activity indicator modeling stage, the occupancy data is extracted from the building's historical database and is preprocessed in order to remove gaps due to acquisition interruptions, outliers and erroneous readings (a). The activity indicator is then defined as the aggregation of the binary occupancy signals (b) and the obtained indicator is modeled by means of a recurrent neural network with global feedback (c). The trained network's performance is evaluated over a test dataset in order to validate that it has properly learned the indicator's behavior (d).

Afterward, during the power demand modeling stage, power data plus auxiliary signals are loaded and preprocessed in a similar manner (f). Then, a power demand estimation method (g) allows the calculation of the total power demand corresponding to the consumption endpoints in the building, decoupling the effect of the distribution bus capacity and the control strategy. Next, an ANFIS model is built for the forecasting of the obtained thermal power consumption signal by selecting the most suitable set of input variables and training the inference structure (h). After the model is trained, it is validated (i) in a similar manner as the activity indicator model in order to ensure its accuracy.

Finally, the activity indicator model (e) is combined with the obtained power demand model (j) to support the calculation of power demand predictions (k). The combination is performed in series, where the output of the activity model is used as an input of the power model.

The following subsections describe the main stages of the methodology in detail.

3.2.1 Activity indicator modeling

In the literature, some studies use timetables as a rough estimation of occupancy, exploring the potential energy savings that could be achieved by implementing management strategies that take advantage of personalized occupancy schedules [80], schedules of the temperature settings of the building [81], or occupancy patterns derived by mining the energy consumption of appliances [82]. However, a recent review of occupancy modeling approaches concluded that schedule-based methodologies are not suitable for applications aimed at improving energy efficiency in buildings, in favor of more sophisticated methods that are able to learn and predict the behavior of occupants [73]. Accordingly, the implementation of a new model of the occupancy pattern of a building is introduced in this methodology.

Thus, in the proposed methodology the concept of an activity indicator is introduced with the aim of incorporating the information relating to the occupancy of the building into the load forecasting system. The proposed activity indicator is defined as the percentage of active spaces in a building, given that the spaces are monitored with presence detectors, which are common in modern buildings for climate and lighting control purposes. The percentage of active spaces is not intended to be a direct measurement of the occupation as the number of present occupants, instead, it is used as a measurement of the amount of activity in the building in terms of spaces where the HVAC system is in operation. The integration of this indicator into the load forecasting system may lead to more accurate predictions because the amount of rooms with an operating local air handling unit (AHU) is likely to significantly affect the load of the HVAC equipment (chillers, heat pumps, etc.) at the energy production stage. However, information regarding this or any other artificial activity indicator is unknown beforehand, as opposed to variables such as weather conditions, which can be pulled from a local weather service with reasonable accuracy. In consequence, a dedicated activity modeling system is integrated into the methodology in order to independently obtain a model of the dynamics of this signal so it can be used for improving the accuracy of the subsequent power demand forecasting.

The modeling of the activity indicator is based on an RNN, which is a data-driven technique that is well suited for cases where the target signal does not present a direct correlation with other signals that could have been used as model inputs, and instead depends on learning the target signal's own dynamics. This is possible because RNNs introduce the time element through their internal states, which allow the network to remember information about the past and to use it for the calculation of predictions, facilitating the learning of the temporal dynamics of the target, instead of relying solely on the current inputs [83]. This feature of RNNs makes them suitable for modeling the activity indicator, which is not strongly correlated with other signals, thus the modeling relies on the accumulated state for learning its temporal dynamics, in this case, complemented with the time of the day and the day of the week for increased robustness. Additionally, memory units have been incorporated into the network in order to provide auto-regressive behavior; this allows the network to not only take into account the previous recurrent state, but past states as well.

The RNN is trained in open-loop form by means of backpropagation, where its coefficients are tuned with the objective function corresponding to the minimization of the mean-squared error of the prediction of the state of the next iteration. After the modeling process is carried out using the open-loop network, the feedback loop is closed to allow the calculation of predictions taking advantage of the recurrent nature of the network. Using the closed-loop form, prediction iterations are calculated based on the value of the previous state, the inputs and past states provided by the memory units. The structure of the complete closed-loop RNN is shown in Fig 3.2. The trained network is then validated in terms of accuracy using several error metrics, evaluating its performance as more iterations are calculated. The results of the validation ascertain whether the performance is sufficient at the desired prediction horizon.

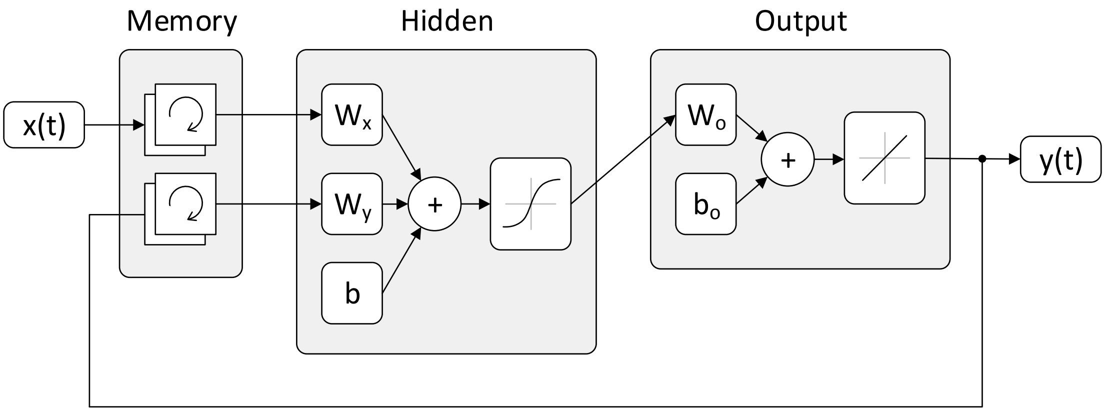


Fig 3.2 *Structure of the closed-loop recurrent neural network, composed by an input layer with memory units, a hidden layer, and an output layer with a feedback loop.*

3.2.2 Power demand modeling

The implementation of the power demand model begins with the initial step of preprocessing the signals to interpolate possible gaps and filter noisy signals acquired by sensors. In addition, a final step is considered for the validation of the trained model structure. However, the core of the proposed power demand modeling is composed of the following two main steps: the power demand estimation, and the fitting of the ANFIS model.

3.2.2.1 Power demand estimation

The power consumption of HVAC systems is a form of instrumentation that is frequently found in buildings, especially in modern smart buildings that incorporate BEMS, which are the main target environment of novel methodology proposals. Thermal power demand, however, is not a variable that is commonly monitored directly due to the high cost of installing sensors in consumption endpoints, even though it is the most useful signal to support the optimization of local resources. The reasoning is based on the fact that when load forecasting systems are implemented for demand response programs or other applications in the context of the smart grid, it makes sense to provide the power consumption of the complete system, because these applications are focused on the optimization and planning of upstream resources. Instead, the proposed method is aimed at providing a forecasting model of the thermal power demand, which can be used to optimize the operation of on-site resources such as HVAC machines.

Since directly measuring the thermal power consumption of the building in real-time is not a commonly affordable option, which would limit the applicability and impact of the methodology, an indirect solution is proposed. The method follows a grey-box approach to allow the estimation of the power demand observed in the thermal distribution bus of the building, implemented as described next.

The energy balance of the bus (Eq. 3.1) is calculated for each time sample, where Q_{in} is the thermal power produced by the HVAC equipment, measured using an ultrasonic flow meter plus a differential temperature sensor, and Q_{out} is the power drawn from the bus, which is not known. The energy accumulated in the bus Qbus during each cycle is described by (Eq. 3.2) where C_p is the specific heat of the fluid in the bus, ΔT_{bus} is the increment of the temperature of the bus, and m is the total mass of the fluid.

$$\Delta Q_{bus} = Q_{in}(t) - Q_{out}(t) \qquad \text{Eq. 3.1}$$

$$Q_{bus} = C_p \cdot m \cdot \Delta T_{bus} \qquad \text{Eq. 3.2}$$

Once the energy balance is defined by the input energy flow Q_{in} and the energy accumulated in the bus Q_{bus}, the resulting power flow being drawn by the consumption endpoints Q_{out} can be calculated by subtraction.

3.2.2.2 Power demand model fitting

After the thermal power demand is obtained, a forecasting model is built for this new signal. The method used in this study for the implementation of the load forecasting is the Adaptive Neuro-Fuzzy Inference System (ANFIS). Even though neural networks are the most popular data-driven methods, mainly due to their accuracy and non-linear mapping capabilities [84], they present drawbacks such as falling on local minima and requiring large datasets [85]. Instead, ANFIS combines the advantages of neural networks with fuzzy systems to better handle complex and adaptive systems, having been validated in multiple load forecasting studies [86].

For the implementation of the ANFIS model, several input signal candidates are considered besides the previously built activity indicator, including weather parameters and other variables commonly available in BEMS, as described in the test environment section. In order to select the set of input signals that allows the proper characterization of the power demand, an input selection process is carried out, which is based on the cross-correlation analysis between each of the input candidates and the target signal to rule out uncorrelated signals, and the study of their dynamics by means of the frequency analysis of each variable. Having considered the candidate inputs and obtained the final selection, an ANFIS model is trained and then evaluated using common performance indicators: the Root Mean Squared Error (RMSE), the Mean Absolute Percentage Error (MAPE), the Mean Absolute Error (MAE), the Determination Coefficient (R^2) and the Maximum Error (MAX).

3.2.2.3 Power demand forecasting

Finally, the power demand of the HVAC system of the building can be predicted using the combination of the trained models obtained following the previous steps. The activity indicator model provides a measure of the future occupancy level, which drives the HVAC power. Then, the expected power demand is calculated to obtain the final prediction, corresponding to this activity and the other influencing variables. In summary, the obtained models are combined in series, with the activity indicator forecast being fed to the power demand model to calculate the final prediction.

Besides the activity indicator estimation procedure, the hybrid solution adopted in this study offers several advantages over traditional approaches. Namely, instead of fitting a single model using a general-purpose tool, a collaborative and modular structure is proposed based on specialized models built for the activity and for the power demand. Such solution allows to fit and tune each method independently, adapting it to the dynamics of each signal and allowing to separately train the models with the use of different datasets.

3.3 Experimental implementation and validation

This section shows the implementation of the proposed thermal load forecasting methodology and discusses the obtained experimental results in the test environment described in *Annex 1. Test environment*.

Two separate datasets are used for the experimental validation of the proposed methodology. For the activity indicator model, the available dataset comprises 8 months of data, from March to October of 2016, including the individual occupancy signal of each of the spaces of the building. Separately, the dataset for the power demand model comprises 11 weeks of data, from late June to early September of 2016, as the dataset corresponds to the cooling power demand, which is only relevant during summer. The power demand dataset contains the power output of the energy production equipment, the bus impulsion and return temperatures, and the external temperature and solar irradiation, measured by a local weather station.

The forecasting horizon is set to one hour in this case, as a shorter horizon would limit the applicability of the load forecasting methodology, and would not allow optimization systems to plan actions with sufficient foresight. Furthermore, a one-hour forecast horizon is sufficient to adapt the predictions to the significant dynamics observed in the building's datasets, which are in the range of two to three hours.

3.3.1 Activity indicator modeling

After the preprocessing of the dataset's signals to remove gaps and to filter out erroneous out-of-range samples, the activity indicator is built using the sum of the individual occupancy signals obtained from the presence detector associated to each space. The resulting activity indicator is shown in Fig 3.3. The pattern presented by the resulting signal follows an expected trend, the indicator rises in the morning as more spaces in the building become occupied and their presence detector is triggered, some drops are observed at midday as people leave for lunch, and finally, most people leave during the evening. However, being a research facility, some remnant occupation can routinely be observed in the building, even during nighttime.

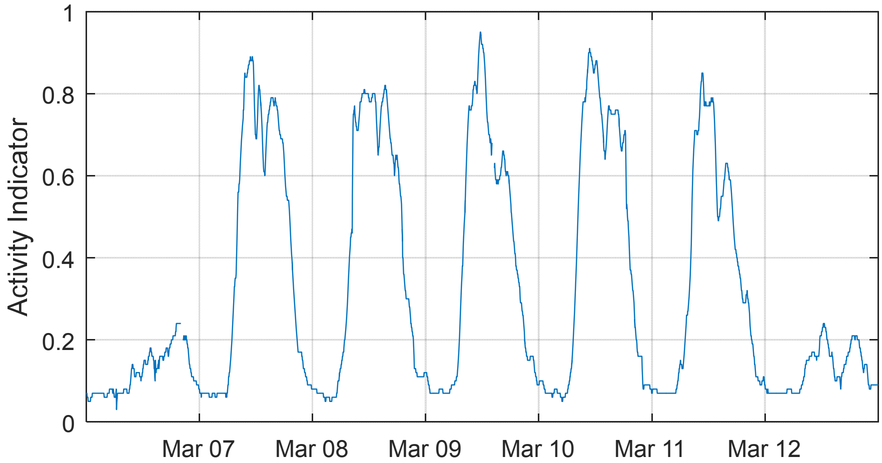


Fig 3.3 *Activity indicator estimated from the aggregate of the individual occupancy signals during a week in March of 2016*

Next, the activity indicator model is built using an RNN, which must be configured before the training. The parameters to be configured are the time step between the recurrent iterations, the number of memory units on the inputs and on the output feedback loop, and finally the number of neurons in the hidden layer.

Considering the temporal aspect of RNNs, it is necessary to properly configure the iteration time step according to the dynamics present in the signal and the desired prediction horizon. Thus, a small time step value in the range of minutes is required in order to capture the dynamics for the next hour horizon. Further experimentation was performed in order to characterize the effect of increasing the iteration time step value. This improves the performance of the network when predicting the activity indicator several hours ahead. In fact, it was possible to predict the activity of the next 8 hours with slightly over 10% RMSE. However, even though increasing the time step lead to expanding the forecasting horizon where the model was still usable, the performance decreased in the short-term, which is precisely when maximum performance is required in order to feed the power demand model. Thus, the value of the iteration time step of the RNN was configured at 4 minutes, which is the minimum acquisition-step available in this case.

Regarding the number of memory units, this amount is set to zero for the inputs, since the dynamics of the input signals of the activity indicator model, which are the day of the week and the time of the day, are not relevant. Instead, the number of memory units in the output feedback loop is set to 15, which at 4 minutes per iteration step matches the one-hour forecasting horizon desired. Therefore, the past states in the last hour are used when forecasting the next hour. Additional experiments were conducted, confirming that including too few units resulted in poor performance, while including too many units did not improve the prediction accuracy, but severely increased the training time due to the added parameters.

Concerning the number of neurons in the hidden layer, related studies recommend using a number of neurons bigger than the number of inputs in order to contribute to an information expansion prior to the output convergence. Subsequently, further empirical experiments were carried out in order to select an optimal configuration. An amount of 16 neurons is finally selected for the hidden layer, as fewer neurons were not able to fully estimate the dynamics of the signal, and more neurons increased the training time while actually decreasing performance.

After the training of the network with the selected configuration, the performance of the resulting model was evaluated over a reserved validation dataset, which accounted for 30% of the available data. The selected performance indicators are the defined for the power demand model: the root-mean-square error (RMSE), the mean absolute percentage error (MAPE), the mean average error (MAE), the maximum error (MAX) and the coefficient of determination (R^2). Because of the iterative nature of the evaluation of the recurrent network, where each prediction is fed back into the model to generate the next state, it is not enough to evaluate the forecasting performance of a single step, as the error is accumulated at each iteration. Thus, the multi-iteration performance must be evaluated to find out if the model is suitable. Fig 3.4 shows the progression of the selected performance indicators as the prediction horizon is expanded. As it can be observed, all of the considered error indicators exhibit a performance decrease as more iterations are applied to the RNN. At 1-hour prediction horizon the mean absolute error is 2.3%, which is a very accurate response taking into account the apparently random behavior of the occupancy in buildings, therefore the model is deemed acceptable for the further implementation of the methodology. It is also observed that the evaluation time increases in a linear trend as more feedback loops are applied in order to increase the prediction horizon.

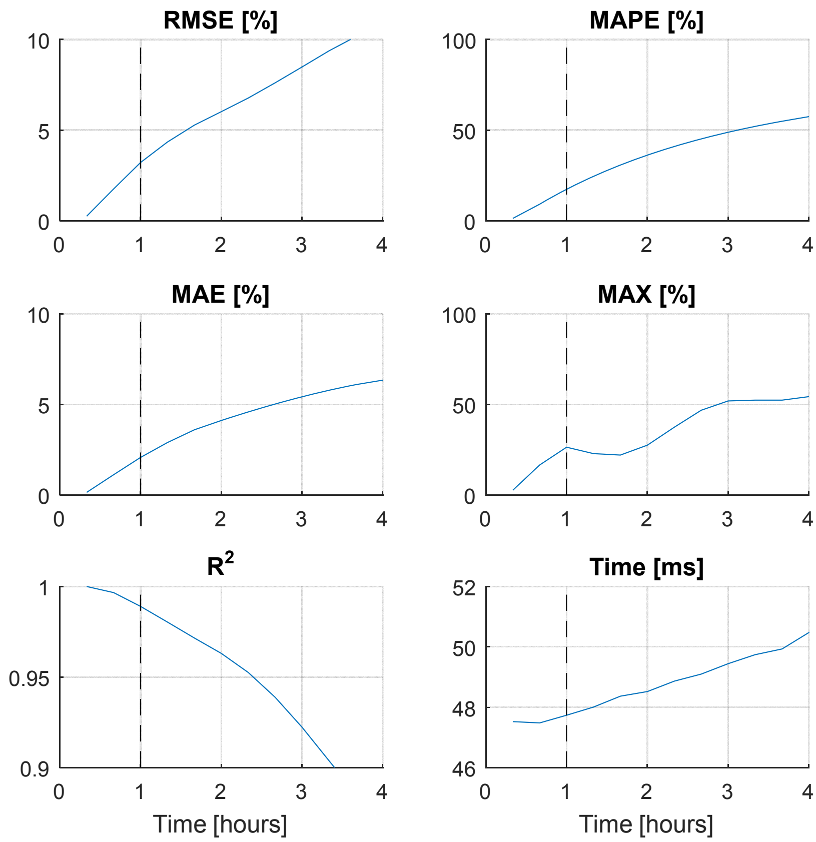


Fig 3.4 *Performance of the activity indicator model when used for multi-iteration predictions using the validation set. Root mean squared error, RMSE. Mean absolute percentage error, MAPE. Mean absolute error, MAE. Maximum error, MAX. Determination coefficient,* R^2*. Evaluation time*

3.3.2 Power demand modeling

Having accomplished the activity indicator modeling stage and having obtained an activity model suitable for use, the next step is to carry out the power demand modeling stage, where the activity forecasting is integrated with ANFIS in order to model the power demand of the HVAC system.

A dataset was extracted from the building's historical database, comprising the variables defined in the test environment section. After the preprocessing of these signals, the first step was to calculate the power demand signal from the measured power output of the machines and the bus temperatures by means of the estimation of the bus dynamic behavior. The bus temperature signals and the estimated power demand compared to the measured power production are shown in Fig 3.5 for a period of three days in August.

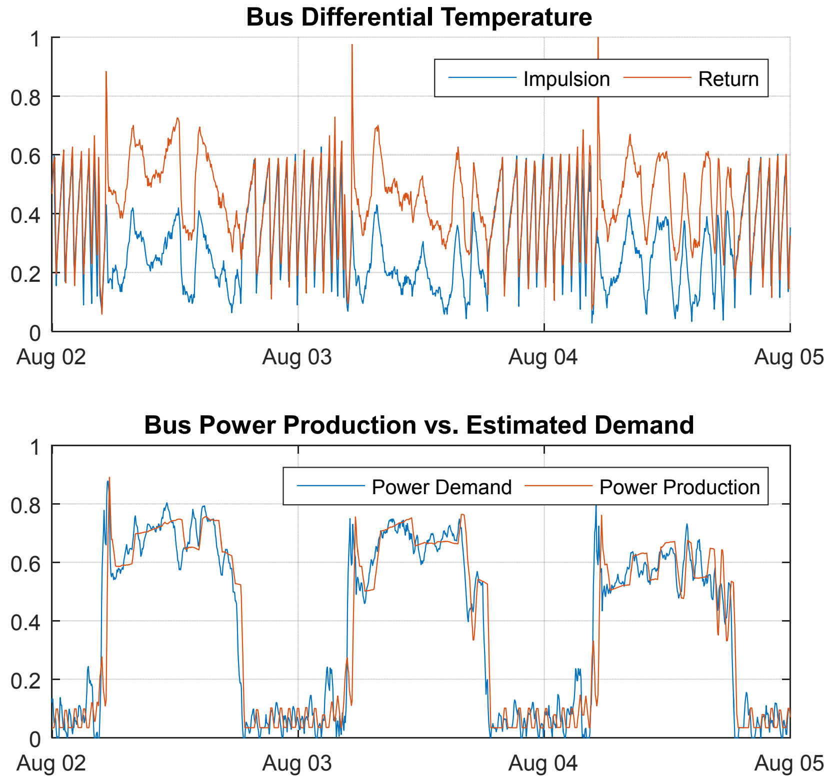


Fig 3.5 *Normalized power demand signal drivers. a) Bus impulsion and return temperatures. b) Bus power production and estimated demand.*

As it can be observed in Fig. 5 b), the power demand signal, corresponding to the aggregated power drawn by the consumption endpoints in the building, presents higher dynamics than the power production, corresponding to the aggregated power generated by the production equipment, while having the same integral value, as the consumed energy must be equal to the production. It is worth mentioning that there is a delay between the risings and fallings of the power demand compared to the power production. This is due to the control scheme implemented in this HVAC system, which does not take into account power demand, and instead focuses on maintaining the bus temperature between thresholds. The difference between the power production and the power demand at the end of each workday is energy that is wasted and will not be consumed by the HVAC system. This energy remains in the distribution bus until it is dissipated because of insulation losses. Having a power demand forecast, this could be improved by producing the minimal energy that is required to match the power demand.

In order to build the power demand model, a set of variables are selected as the inputs for the model from the available signals in order to facilitate the work of the training algorithm. The following signals were considered as inputs: external temperature, solar irradiation, bus impulsion temperature, bus return temperature, bus differential temperature and finally the estimated activity indicator. To select the model's inputs, the cross-correlation between the target signal and each of the input candidates is calculated in order to rule out uncorrelated signals.

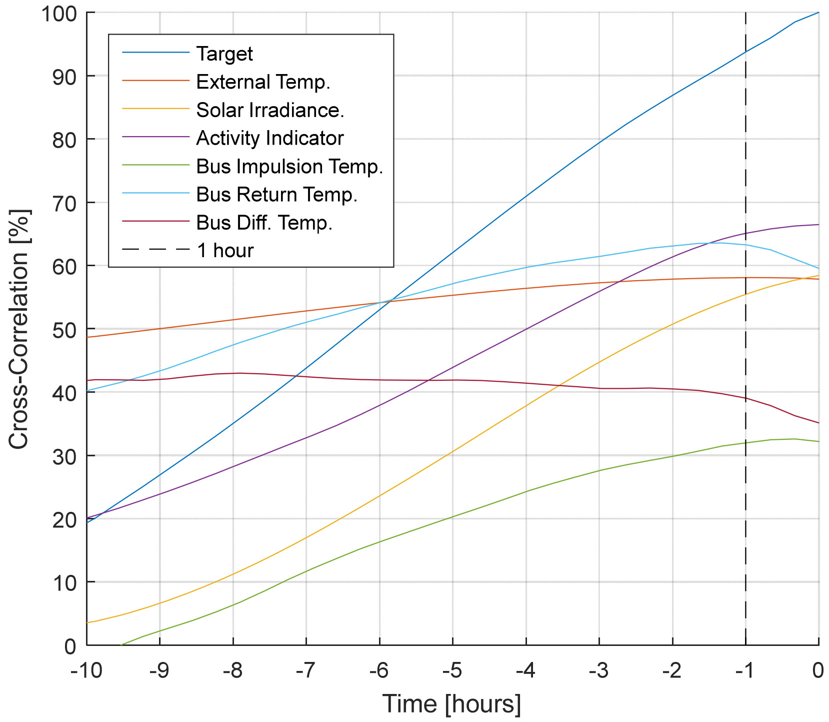


Fig 3.6 *Cross-correlation between each model input candidate and the forecasting target.*

The different cross-correlation pairs are shown in Fig 3.6, where each series shows the correlation between an input candidate and the thermal power demand as a time shift is applied between the two signals. It is desirable that the selected inputs show a high correlation with the target signal at the forecasting horizon, which is set to 1 hour in this case. As it can be observed, the most strongly correlated input candidates when the offset between each pair is 1 hour are the external temperature, the solar irradiance, the activity indicator and the bus return temperature. On the other hand, the bus impulsion temperature and the bus temperature differential present low correlation with the target. Finally, it is noticeable that the target shows a strong correlation with itself when a 1-hour offset is applied, therefore the current power demand value was also considered as an input for the model. A sample of the preselected input variables is shown in Fig 3.7.

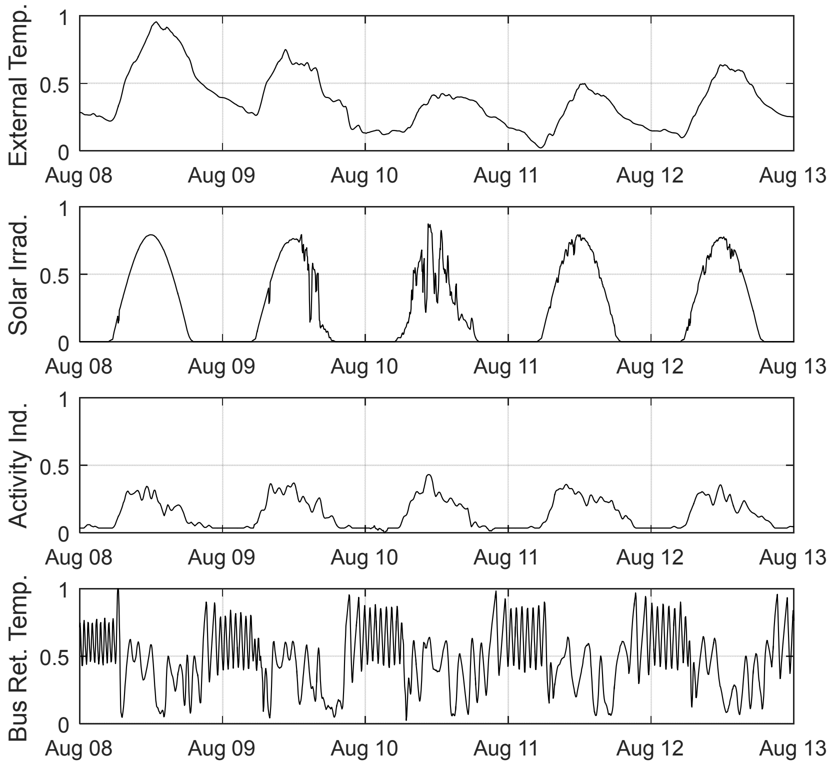


Fig 3.7 *Selected input variables during a period of 5 days in August*

The study of the signal's frequency components, shown in Fig 3.8 as the frequency spectrum analysis, revealed the magnitude of the signal's dynamics. As it can be observed, the solar irradiation and the external temperature present rather slower dynamics than the power demand, which is expected as they mostly follow a daily pattern. Instead, the activity indicator presents significant dynamics up to sub-hour period frequencies, which is more aligned with those observed in the power demand, as is the case of the bus return temperature, which presents even higher frequency components. Thus, the inclusion of the activity indicator and the bus return temperature may help the model to better adapt to the power demand's dynamics, as these signals present more similar frequency components.

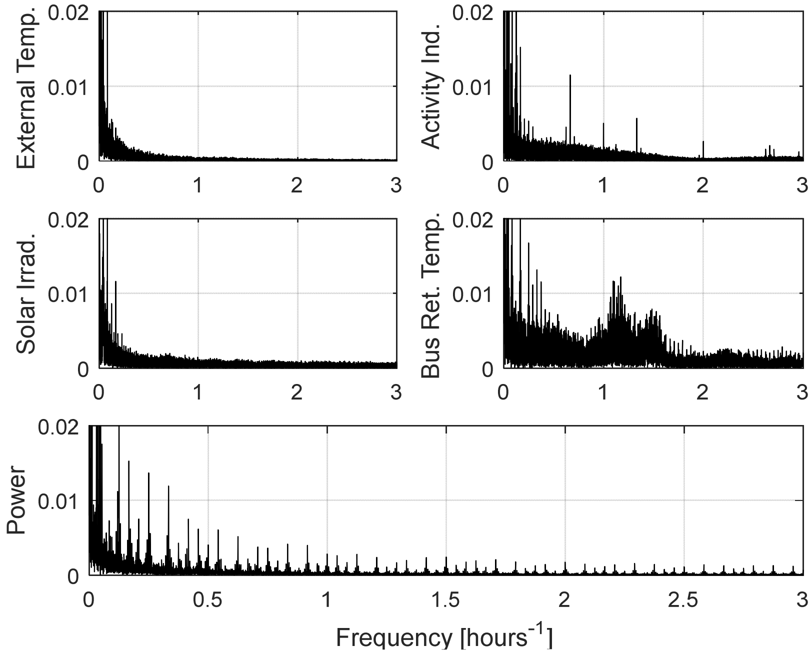


Fig 3.8 *Frequency spectrum comparison between the power demand model input candidates and the power demand target signal.*

Additional empirical analyses carried out with the available signals, reveal that the use of both the external temperature and the solar irradiance do not improve the modeling performance, as these two signals present correlation between them and introduce redundant information into the model. As the external temperature presents a smoother behavior than the solar irradiance, which has very steep peaks due to passing clouds, and a forecast of the external temperature is readily available through a local weather service provider, but not for the case of the irradiance, the latter was discarded and only the former was used. Regarding the current value of the target, it was noticed that it improved the forecasting accuracy when included, as it provided a reference point to calculate the next values. Concerning the bus temperature signals, only the return temperature was used, as it provides feedback about the state of the production/demand match. The bus differential temperature was considered, even though it presented low correlation with the target, in an attempt to increase the accuracy of the model during rapid changes, as the bus differential presents high dynamics. This helps the model perform better in some cases, but overall introduces noise and is finally discarded. Finally, another considered variable is the day of the week, which was included as it helps the ANFIS rule inference step to properly characterize the behavior of the power demand during weekends. In summary, the study revealed that the most appropriate set of signals to characterize the power demand of the building is: the external temperature, the activity indicator, the bus return temperature, the current power demand value and the day of the week.

The result of the model training is shown in Fig 3.9, where it can be observed that the model closely matches the target on most of the signal, presenting a low average error. However, there are also error peaks that occur when the target signal presents the fastest dynamics, causing error spikes due to steep changes, but having very short duration.

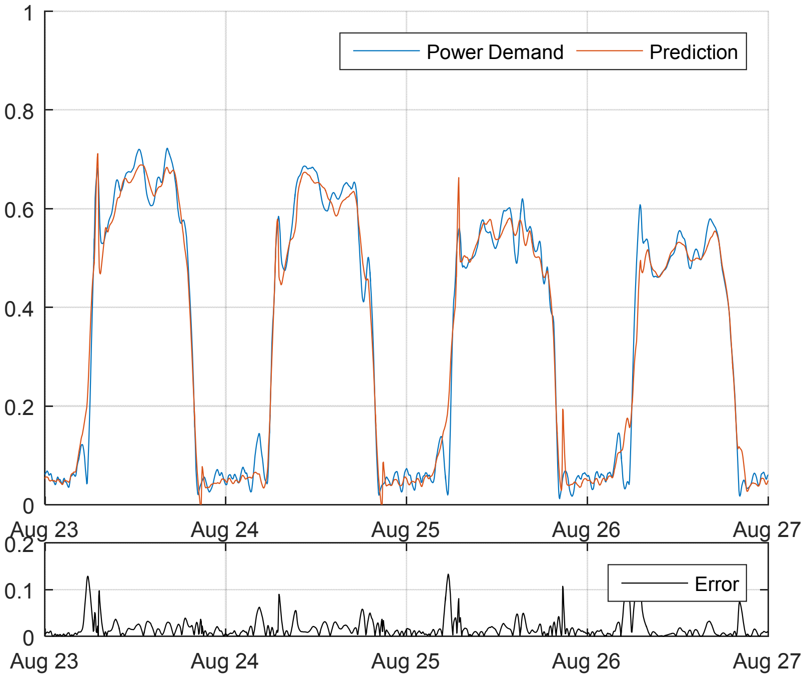


Fig 3.9 *Comparison between the power demand signal and a prediction obtained using the trained power demand model*

In order to validate the methodology and to evaluate its performance and generalization capabilities, a cross-validation strategy was followed. The cross-validation implementation removes one week of data at a time from the dataset, builds a model using the remaining data and validates the model against the removed subset. Thus an 11-fold cross-validation is considered. The results of the cross-validation are shown in Fig 3.9, where several performance indicators were calculated when the model is applied to the training set and separately over the validation set. As it can be observed, the error indicators are quite low, with the mean absolute error being the most compelling at an average value of 2% during training and 3% during validation. The maximum error shows an average of 13%, which is acceptable due to the occasional rapid changes observed in the signal, but reaches a value of 26% when week 3 is not present in the training set. In fact, the other error indicators are also noticeably higher when week 3 is used as validation and is excluded from the training. This observation indicates that week 3 presents a behavior that differs from the rest of the data, as the resulting model achieves worse prediction performance when learning from the other cases.

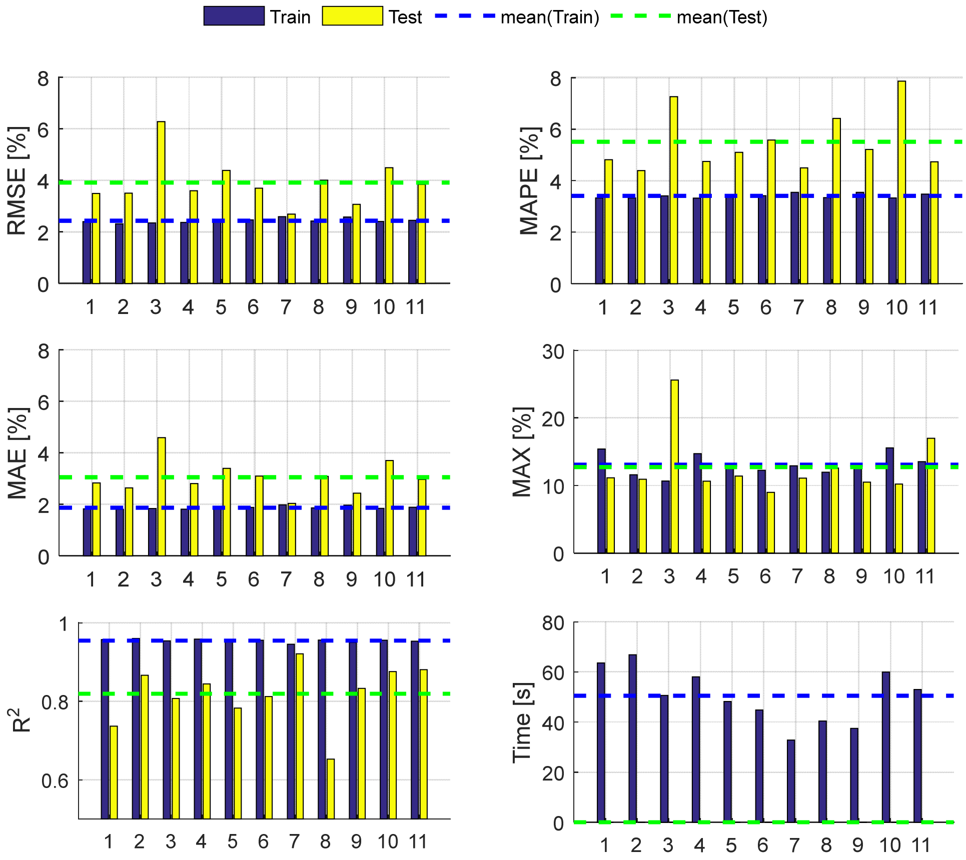


Fig 3.10 *Results of the cross-validation process when splitting the data into 11 subsets, corresponding to the weeks in the dataset.*

Finally, in order to quantify the increase of performance provided by the application of the proposed methodology, the obtained results have been compared with a classical load forecasting implementation based on ANFIS. The evaluation of the power demand modeling stage using the proposed methodology resulted in decreased error metrics. The following Table 3.1 shows the performance change when comparing the average performance metrics of the proposed methodology to the ones obtained from the cross-validation of the classical approach both including and withholding the activity information.

	A	B		C	
	Value	Value	Δ%	Value	Δ%
RMSE	3.932	4.239	-7.81	5.196	-32.15
MAPE	5.571	6.205	-11.38	7.745	-39.02
MAE	3.055	3.384	-10.77	4.198	-37.41
MAX	13.200	12.941	+1.96	13.352	-1.14
R^2	0.821	0.775	+5.58	0.704	+14.25
TIME	51.21	51.57	-0.70	30.260	+40.91

Table 3.1 *Performance comparison between the proposed method (A), and the classical single ANFIS approach with the activity Indicator (B), and the classical single ANFIS approach without the activity indicator (C)*

As it can be observed, the introduction of the activity indicator causes a significant improvement in most of the performance metrics over a classical ANFIS approach that does not take into account occupancy data, except for the training time, which is almost halved. This reduction in the duration of the training time is likely due to the reduction in the size of the data and the loss of one dimension in the input space by not considering the activity indicator, which allowed the modeling to speed up convergence at the cost of increased error. Additionally, the integration of the occupancy forecasting in the proposed methodology in order to provide more updated activity values helped to further increase the performance metrics over a classical approach that used the activity indicator.

3.4 Discussion and conclusions

A short-term activity-aware thermal power demand forecasting methodology is studied in this chapter, aligned with the state of the art on load forecasting in buildings for energy management applications. The proposed methodology consists in a hybrid modeling process where a dedicated recurrent neural network learns the dynamics present in an activity indicator developed for this study, and an adaptive neuro-fuzzy inference system correlates activity predictions obtained in this manner with the outdoor temperature and the bus return temperature in order to characterize the thermal power demand of the building's HVAC system.

The integration of the activity assessment into the modeling process, through the definition of an indicator that reflects the occupancy state of the whole building, has been shown to increase the accuracy of the power demand forecasting. The error metrics are significantly decreased when the activity is used as an additional input for the power demand forecasting, but they are further diminished when the neural network is included as a dedicated means to learn the activity's dynamics, providing an estimation of the use that the building shall receive in the following hour. To this end, the implementation of the activity modeling with a recurrent neural network is validated as a suitable approach in order to consider the temporal patterns of the building's activity, as the proposed activity modeling process exhibits an important performance increase compared with state-of-the-art approaches, achieving a mean absolute error below 10%.

The proposed thermal power demand estimation procedure allows the modeling of the total power being drawn by the consumption endpoints in the building, instead of modeling the consumption of the entire installation as is done in most related studies. The estimation is achieved by means of an energy meter that monitors the aggregate output of the production stage equipment and the simulation of the bus capacity in order to calculate the difference. The main benefit of this change is to allow the decoupling of the effect of the capacity of the distribution bus and the effect of the management strategy followed by the HVAC energy production equipment. Therefore, future studies may build on this methodology for implementing production management strategies that optimize the operation of the equipment according to the forecasted power demand in order to increase the energy efficiency.

A study of the available input candidates for implementing the power demand model was carried out in order to obtain a set of variables that allows the accurate modeling of the target signal. This study helped identify the set that achieves the best results: the current power demand, the activity indicator, the external temperature, the bus return temperature and the day of the week. The developed methodology can be generalized to other cases, extending its applicability.

Besides increased accuracy, the proposed methodology presents other advantages, such as the possibility of using separate datasets of potentially different sizes for the activity indicator model and for the power demand model, which allowed the selection of representative datasets for each case. Additionally, this decoupling allowed the separation of concerns, promoting the specialization during the selection of the best modeling algorithm for each signal and the independent tuning of the configuration of each model, including the use of different inputs and dynamics to match each target signal's behavior. The proposed structure also decouples the model tuning process, allowing to update a model independently of the other when necessary, since the activity model may need to be updated more often due to the changing behavior of the activity of the building.

4.

Performance modeling of HVAC equipment

Obtaining models establishing the relationship between the operating performance of HVAC equipment and their control settings and operating conditions is an essential milestone for the implementation of different energy management applications that rely on them. This chapter introduces a novel operating performance modeling method for HVAC production equipment based on a deep learning approach.

CONTENTS:

4.1 Introduction

4.2 Deep Learning with AutoEncoders

4.3 Operational performance modeling methodology

4.4 Experimental implementation and validation

4.5 Discussion and conclusions

4. Performance modeling of HVAC equipment

In this chapter, the most employed methods for establishing the relationship between operating state, control actions and energy performance of HVAC equipment are reviewed and their drawbacks identified, leading to the development of a novel generic method for modeling HVAC equipment based on a deep learning approach.

4.1 Introduction

This section introduces the background and motivation for pursuing this line of research, reviews the state of the art related to how the relationship between operating state, control actions and resulting energy performance is established in practice in current energy management applications and describes the innovative contributions of this work.

4.1.1 Background and motivation

As has been previously described, buildings are one of the largest energy consumers world-wide, with a substantial share of energy being used for operating HVAC systems [87]. Operational improvements aimed at increasing the efficiency of buildings may lead to considerable savings, potentially up to 20-30% according to different studies [88,89], considering that the control of HVAC systems is consistently suboptimal [90].

In this regard, the operating performance modeling and tracking of HVAC equipment represents an enabling tool for the implementation different energy management schemes [91], including optimization and maintenance functionalities [92–94], which perform an essential part in achieving the optimal use of energy resources and for reducing operating costs [95]. Such energy management functionalities often rely on accurate operational performance data being available [96]. However, obtaining performance data of specific equipment and providing performance maps suitable for: i) predicting the expected behavior, and ii) selecting the control parameters to achieve a desired state, currently represents a challenging and costly process that requires scientific and technical attention [97].

4.1.2 State of the art

By means of a thorough review of the state-of-the-art related to HVAC equipment performance characterization, modeling and applications requiring performance maps, three main strategies can be identified based on i) manufacturer data, ii) simulation tools and iii) empirical methods.

In the first case, the use of performance tables from equipment manufacturers is common, especially in control applications [98]. Its main advantage is that the characterization effort is done by the manufacturer, meaning that no further step is necessary because the data is tabulated and ready for use. However, even when such data is accessible, it still faces a crucial issue: inaccuracy due to not reflecting that equipment performance is highly dependent on operating conditions [99]. Indeed, manufacturer-provided tables do not reflect the real operation context because they are obtained from testing equipment in controlled settings, while operating on a standard test environment and near design load [100]. Therefore, it is not comprehensive, often does not contain the full range of operation and does not provide the correlation with the full set of influencing variables, most of the time presenting only the correlation between the Coefficient of Performance (COP) and the Partial Load Ratio (PLR), while presenting insufficient granularity [101]. Thus, applications relying in such performance data may not fulfill the energy savings potential [102].

In the second case, the characterization of HVAC equipment behavior through its simulation is also a common strategy. Powerful specialized tools like TRNSYS [103], or general-purpose toolkits like Simulink [104], are often used for building thermodynamic models of HVAC systems. The simulation of their internal variables in order to study their response to control settings and to influencing external conditions, like the weather, is a common way of use [105]. This approach achieves great precision and data granularity, as it is not restricted by the cost of empirical experimentation, and only incurs in computational and expert time cost. However, they are complex tools that require ample domain knowledge, information and time for both implementation and usage [106], and require fine-tuning of the physical parameters of the equipment to be simulated in order to guarantee a certain degree of reliability [107]. Furthermore, simulation tools, likewise manufacturer tables, do not reflect the state of performance degradation due to aging or faults [108], which is a very important effect that requires the models to be able to adapt during the system's lifecycle [109]. Indeed, this kind of models are more adequate during the design and test stages of the equipment's lifecycle [110].

Finally, in the third case, empirical performance characterization approaches can be sub-classified into gray-box and black-box modeling methods. Gray-box methods rely on physical equations that define the underlying physical processes and use empirically obtained data to estimate and tune the equations' parameters [111]. Nonetheless, gray-box methods also present drawbacks, such as the requirement of a comprehensive dataset that sufficiently represents the operating conditions of the equipment, which is also a problem in black-box methods. However, even though this obstacle can be overcome due to the pervasiveness of monitoring solutions in Building Energy Management Systems (BEMS), which often track and store large amounts of operational information [112], ample domain knowledge is still necessary for defining the physical equations and for tuning their parameters [113].

Several recent studies are favoring black-box approaches to achieve accurate and practical performance characterization and modeling of HVAC equipment, because they provide a general solution that can be applied with a higher degree of independence from installation type or complexity. The main advantage of black-box approaches is that they are capable of extracting signal relationships from historical datasets, based on the behavior observed in the data, which may be nonlinear, as in the case of chiller performance maps [114]. Numerous black-box methods exist and some have been applied to solve the performance modeling problem, from simple approaches consisting in regression methods [115], to more complex solutions based on Neural Networks (NN) [116,117]. Indeed, a recent comparison among black-box methods concluded that NN represents a viable method because of its adaptability and capacity for learning non-linear relationships between the input variables and the target [118], which is a required feature in this problem given the non-linear behavior of the chiller's power properties, dependent on the operating temperatures [51]. However, as stated by different authors, NN presents some limitations that make their application in a particular case non-trivial, mainly: i) they require a comprehensive dataset that properly represents the behavior of the system to be modeled, ii) the process of selecting the hyper-parameters for a specific application is a subject of ongoing discussion [119,120], and iii) the training process by means of the back-propagation (BP) algorithm can lead to sub-optimal solutions [121].

Although a few techniques have been adopted for lessening these issues [122], NN still present limitations to their applicability in specific cases. From the engineering perspective, extensive work is required in terms of dimensionality reduction and feature engineering, which is usually application-specific and needs to be performed by an expert [123,124]. In addition, from the practical perspective, performance maps are required on a per-machine basis, which means that empirical approaches rely on extensive instrumentation at the machine level. This leads to high costs when carried out beyond small-scale installations.

In this regard, a recent advancement in the field of artificial intelligence and machine learning approaches, is the introduction of deep learning for overcoming such limitations [125]. Deep learning refers to non-linear algorithmic approaches that extract a hierarchical abstraction model from raw data. Most of deep learning deployments use neural networks architectures. Thus, each layer of a deep neural network (DNN) represents a non-linear data mapping fed by the output of a previous layer using a set of computation elements [126]. These data-layer representations depict extracted features from the data for its characterization. Indeed, deep learning is commonly referred to as feature learning, emphasizing the fact that the feature extraction itself is learned from data during the training. In fact, standard machine learning approaches requires specific feature reduction assistance, since the direct introduction of all available features leads to poor convergence during the learning phase [127]. In this regard, optimization techniques, as genetic algorithm (GA) based, are often considered to locally maximize the model accuracy by selecting a subset of the available features. These approaches, however, lead to a lack of generalization capabilities in front of slight variations of the input data, the so-called overfitting effect [128]. Deep learning approaches avoid the preceding feature selection, instead, dealing directly with data, allowing the so called end-to-end learning. This concept, although through a more complex structure, leads to three potentialities: i) increase of generalization capabilities, since higher order spatial interpolation is available, ii) visualization of the features significance through the analysis of the resulting layers, and iii) transfer learning, since additional data can be incrementally learnt without the need of complete model re-training and previous feature selection constraints. Although deep learning theory has been available for decades, three recent advancements are helping to increase its adoption, that is: i) training process improvements, ii) large amounts of data required for deep learning have become attainable, in this case facilitated by

BEMS, and iii) the increase of computational power that has allowed the utilization of training algorithms to become feasible.

4.1.3 Innovative contribution

In this chapter, a novel methodology for performance modeling of HVAC equipment based on deep learning is introduced. The originality of this work includes the development of a deep neural network based scheme that overcomes the drawbacks of approaches found in the state of the art and delivers a solution to model the coefficient of performance, electrical consumption and thermal power production of a multi-machine based HVAC system. The main contributions of this research work include:

- The proposal of a methodology based on the performance analysis of a set of machines as a group by means of a deep learning approach, lowering implementation costs by permitting to reduce the amount of required sensors.

- The estimation of the expected electric power consumption and the expected thermal power production besides the expected operating performance of each machine, while enclosed in a single modeling process.

- The visualization of the feature significance resulting from the modeling process and its adaption during an incremental learning process.

It must be noted that this work represents an important step to the introduction of deep learning techniques to the development of HVAC modeling procedures, being the first time that is has been applied at the equipment level. Moreover, the suitability and accuracy of the proposed method is validated in a real case study in a tertiary sector building, and the results are compared with a classical modeling approach.

4.2 Deep Learning with AutoEncoders

Traditional NN applications implement shallow architectures, being the feed-forward network with three layers the most common, where the first layer corresponds to the inputs, the middle layer is a hidden layer and the final layer corresponds to the outputs. In such shallow networks, the inputs are a carefully curated set of signals obtained by means of feature reduction and/or feature engineering, because they are unable to learn complex features and relationships from the raw data [129]. In contrast, deep learning takes a feature learning approach, i.e. features are discovered instead of being given, by taking advantage of the properties of deep networks, where the initial layers extract meaningful features in an unsupervised manner and the final layers map these features to the target [130]. By following this approach, the resulting network is able to work with a wider set of inputs, where the training stage identifies features and assigns weights autonomously [131].

Unfortunately, the larger number of layers contained in deep networks poses a crippling issue in practice, where the classical backpropagation algorithm fails to update the weights through the layers during the training process as the gradient becomes too small to influence a change and prevents further learning, a problem known as the vanishing gradient [132]. Several techniques have been shown to be effective in alleviating this issue in recent years, including the use of different network architectures, or the use of regularization techniques such as the dropout method [133]. However, one of the best performing solutions is the implementation of layer-wise pre-training schemes, where each layer is trained separately instead of presenting a target to the network and using a supervised learning algorithm such as backpropagation to update all the weights of the network at once. Still, it is not possible to train each layer in a supervised manner, as the values observed on each intermediate layer are unknown. Instead, an unsupervised approach can be adopted by considering the network as a deep autoencoder where each layer is treated as single layer autoencoder, as shown in Fig 4.1. In this sample network there are two consecutive encoder layers that transform the input to a latent space representation located at the center of the network, which is then followed by two decoder layers that transform it back to the input data space.

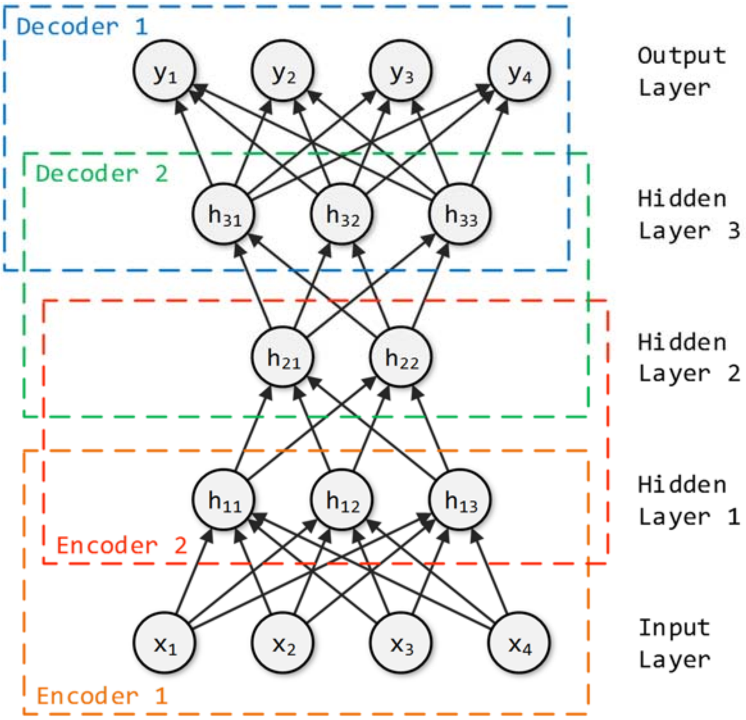


Fig 4.1 *Sample architecture of a multiple-layer autoencoder composed of two encoder and two decoder steps.*

Single autoencoders are trained individually, where each subsequent layer uses the result of the prior encoder as the input. After each autoencoder has been trained in this manner, the set of autoencoders is unrolled in order to build the deep autoencoder, which can then be further fine-tuned [134]. The result is a deep autoencoder that is only capable of reconstructing the input space on the output of the network, as it has not been shown any labelled data yet, but it has learned an internal representation of the data and relationships between the input signals. This trained autoencoder can then be adapted for classification or regression applications by extending it and applying a further supervised learning stage so it can learn to map the internal representation to a target, an approach that has been shown to provide notably better results in practice [135].

4.2.1 Base autoencoder implementation

Autoencoders are neural network based structures that are trained in an unsupervised manner, usually by backpropagation, where the objective is to learn a transformation that approximates the identity function [136]. By means of the introduction of constraints, the information flowing through each encoder becomes compressed while the autoencoder is forced to reconstruct the input at the output, leading to the discovery of signal relationships and the internal structure of the data [137].

A single-layer autoencoder is can be defined as shown in Fig 4.2, where $\boldsymbol{x}$ is the vector of length $\boldsymbol{k}$ containing the set of inputs signals.

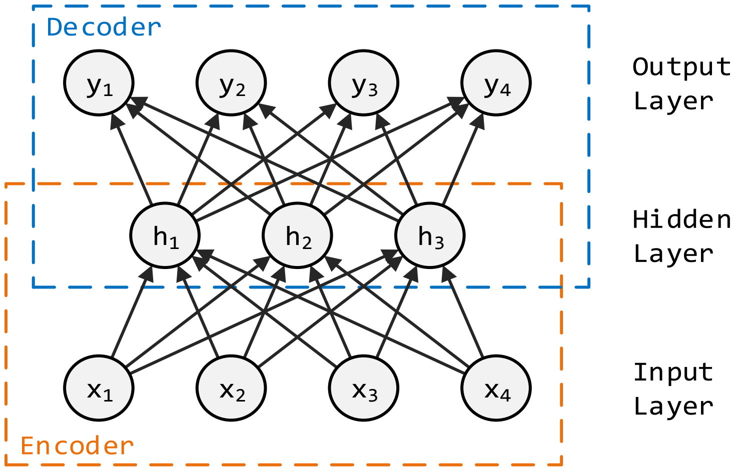


Fig 4.2 *Single-layer autoencoder architecture.*

In order to transform the input data to the hidden layer representation $\boldsymbol{h}$ consisting of $\boldsymbol{n}$ sparse-activated neurons, the encoder transformation (Eq. 4.1) is applied to the input vector, where $\boldsymbol{W_e}$ and $\boldsymbol{b_e}$ are the weights and biases matrices respectively, and $\boldsymbol{f}$ is the chosen activation function.

$$h = f(W_e x + b_e) \qquad \textit{Eq. 4.1}$$

Different activation functions exist, being the sigmoid function (Eq. 4.2) the most commonly used.

$$S(t) = 1/(1 + e^{-t}) \qquad \textit{Eq. 4.2}$$

Then, the encoded hidden layer is transformed again in order to obtain the output of the autoencoder by applying the decoder transformation (Eq. 4.3), where $\boldsymbol{W_d}$ and $\boldsymbol{b_d}$ are the weights and biases matrices of the decoder, respectively, and $\boldsymbol{y}$ is the output of the network, having the same dimension as $\boldsymbol{x}$.

$$y = f(W_d h + b_d) \qquad \textit{Eq. 4.3}$$

The parameters of the autoencoder can then be tuned by training the network with the aim of reconstructing the inputs vales at the output layer. This can be achieved by defining the loss function as the L2-norm error between the inputs and their reconstruction (**¡Error! No se encuentra el origen de la referencia.**).

$$J_{\varepsilon} = \|x_i - y_i\|^2 = \sum_{i=1}^{k} (x_i - y_i)^2 \qquad \text{Eq. 4.4}$$

In order to limit overfitting, a weight-decay regularization term (**¡Error! No se encuentra el origen de la referencia.**) is added to prevent large weights from appearing, where $\boldsymbol{\beta_\omega}$ is a parameter that controls the weight decay, $\boldsymbol{a}$ is the number of weight parameters, $\boldsymbol{b}$ is the number of rows and $\boldsymbol{c}$ is the number of columns in each weight matrix.

$$J_{\omega} = \beta_{\omega} \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} W_i(j,k)^2 \qquad \text{Eq. 4.5}$$

4.2.2 Sparse autoencoder implementation

There are two ways to achieve information compression in autoencoders: i) constrictive autoencoders, which limit the size of subsequent hidden layers in the network, resulting in a form of dimensionality reduction, or ii) restrictive autoencoders, which enforce a sparsity constraint on the activations of the neurons in the hidden layers, allowing an autoencoder to learn relationships in the data even when subsequent layers present a higher dimension that the inputs. In fact, the units present in the hidden layer of sparsely-activated autoencoders, as shown in Fig 4.3, are tuned so that they seldom activate, effectively functioning as specialized feature detectors. Thus, sparsity is a useful property that offers great potential for the purpose of feature learning [138].

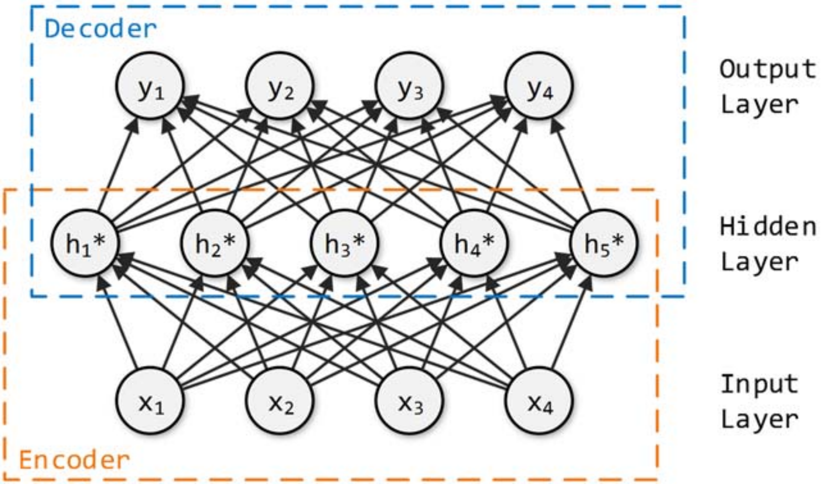


Fig 4.3 *Single-layer autoencoder with sparse-activated neurons.*

In order to coerce the activation of the neurons into becoming sparse when a larger number of units is present at the hidden layer of the autoencoder than at the input layer, an additional penalty term is introduced to restrict the tuning of the parameters. This sparsity penalty term (Eq. 4.6) forces the neurons to activate sparsely throughout the dataset. The term is implemented using the Kullback-Leibler divergence (Eq. 4.7), which is a standard measure of how a probability distribution diverges from the expected distribution, where $\boldsymbol{\rho}$ is the desired sparsity parameter and $\hat{\boldsymbol{\rho}}_{\boldsymbol{i}}$ is the effective sparsity of a given hidden unit.

The characterization of the sparsity is defined as the mean activation of the hidden unit over all the training samples in the dataset. In general, a small sparsity parameter is desired so that the units seldom activate.

$$J_{sp} = \beta_{sp} \sum_{i=1}^{n} KL(\rho \| \hat{\rho}_i) \qquad \text{Eq. 4.6}$$

$$KL(\rho \| \hat{\rho}_i) = \rho \log \frac{\rho}{\hat{\rho}_i} + (1-\rho) \log \frac{1-\rho}{1-\hat{\rho}_i} \qquad \text{Eq. 4.7}$$

However, several different units may activate in common sections of the data. In order to prevent feature over-representation, a concurrent activation penalty is introduced (Eq. 4.8), where $\boldsymbol{h_j}(\boldsymbol{x_i})$ denotes the activation value of a given hidden unit when presented with a data sample $\boldsymbol{x_i}$, the maximum allowed concurrency is represented by $\boldsymbol{\gamma}$,while $\boldsymbol{\beta_\gamma}$ is an adjustable weight. This term's value increases when overlapping is observed in the activation of hidden units, forcing them to become differentiated. A small concurrency value is recommended so that the neurons become distinct indicators.

$$J_{\gamma} = \beta_{\gamma} \sum_{i=1}^{N} \left(\max \left(\sum_{j=1}^{n} h_j(x_i) - \gamma, 0 \right) \right) \quad \text{Eq. 4.8}$$

Finally, the loss function to be minimized is defined as the sum of the error term plus the regularization penalty terms, so the parameter-tuning problem can be stated as an optimization problem where the networks parameters are adjusted in order to minimize the resulting loss function (Eq. 4.9).

$$\min \left\{ J(W_e, b_e, W_d, b_d) = J_{\varepsilon} + J_{\omega} + J_{sp} + J_{\gamma} \right\} \quad \text{Eq. 4.9}$$

4.3 Operational performance modeling methodology

A step-by-step diagram of the proposed performance modeling methodology is shown in Fig 4.4, which is divided into two main stages: the feature learning stage, where the raw input signals are consolidated into an internal representation through an unsupervised layer-wise autoencoder pre-training approach, and the performance mapping stage, where the internal representation is mapped to the targets by means of a supervised training scheme.

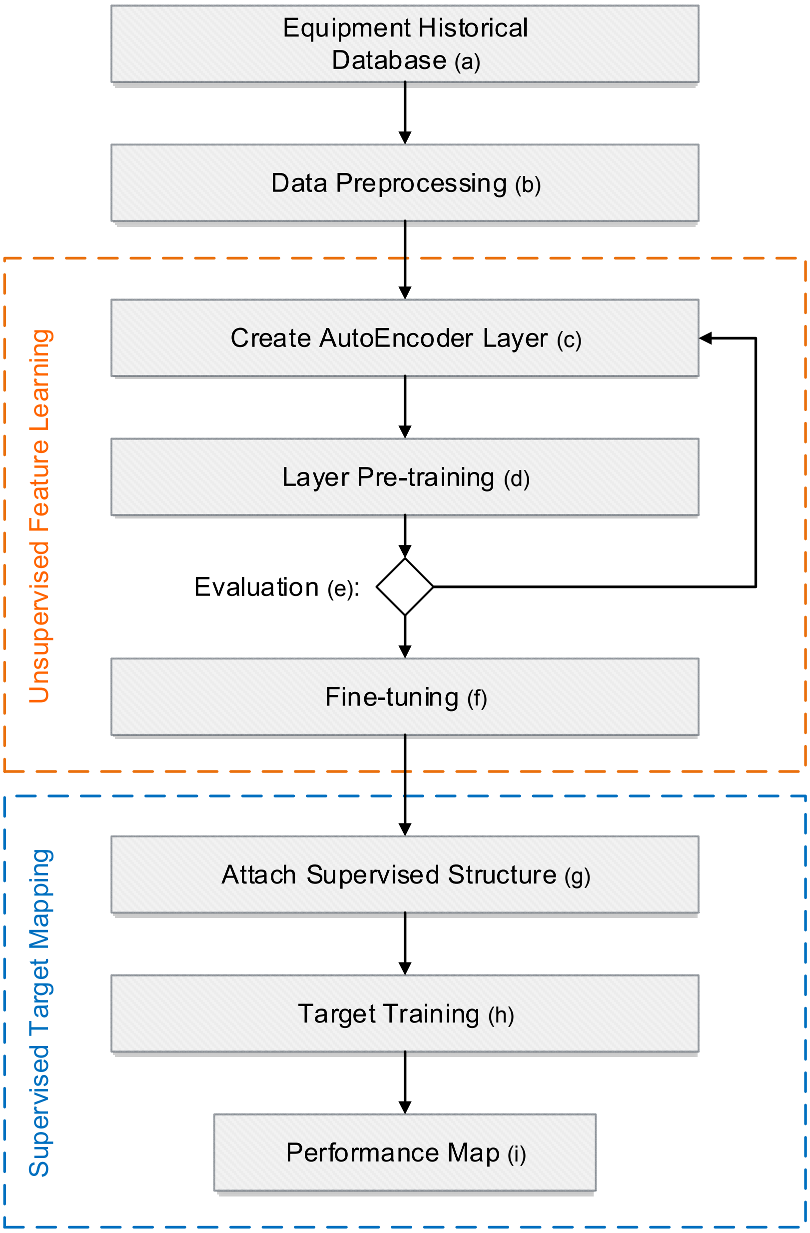


Fig 4.4 *Steps of the proposed performance modeling methodology divided into two main stages: unsupervised feature learning and supervised target mapping.*

Initially, a dataset is extracted from the database that stores the historical operation of the building's HVAC equipment (a), including for instance consumed and produced power, control signals, state variables and weather conditions. The extracted dataset is preprocessed in order to remove gaps caused by acquisition interruptions, outlier removal and erroneous readings (b).

Following the proposed methodology, during the unsupervised feature learning stage, a deep autoencoder structure is built by iteratively stacking layers (c), and trained to reconstruct the input space by means of the layer-wise pre-training strategy (d). After the pre-training of each layer, the reconstruction is evaluated by comparing it to the previous layer in order to assess whether additional layers should be created (e). When the stacked layers meet the evaluation criteria, a fine-tuning process (f), further improves the reconstruction error. After the feature learning stage is completed, the decoder half of the deep autoencoder is detached to proceed with the supervised target mapping stage. First, additional layers are attached to the central layer of the network (g), in order to perform the mapping from the latent space to the target variables. Then, the additional layers are trained in a supervised manner (h), in order to learn the relationship between the power consumption, production and coefficient of performance and the latent space.

Finally, the performance map (i), is obtained in the form of a deep neural network that is able to calculate the expected power consumption, production and coefficient of performance for a given set of operating conditions. The accuracy of the resulting performance map is then evaluated using several error metrics.

The following subsections describe the main stages of the methodology in detail.

4.3.1 Unsupervised feature learning

The aim of this stage is to construct a deep neural network able to discover features in the input data space in an unsupervised manner. This is achieved by means of a stacked autoencoder architecture where features are extracted at each layer and composed through the layers in order to perform complex feature learning.

Single-layer sparse autoencoders are created and trained following the implementation described in the previous section. The advantage of following a layer-wise pre-training scheme is that the parameters of each layer in the network are tuned separately by means of backpropagation, where the vanishing gradient problem is non-existent due to each autoencoder being composed of a single layer. However, in order to build the multi-layer network by stacking single-layer autoencoders, care must be taken to ensure the proper sizing of the intermediate layers, as the amount of necessary hidden units is difficult to ascertain beforehand due to the intrinsic relationship of this magnitude with the complexity of the data. Thus, a dimensioning strategy is implemented in this methodology by means of the evaluation of the activations of the units in the hidden layer.

Firstly, an initially large amount of hidden units is configured, the layer is trained and its reconstruction error is evaluated. After the initial training, due to the large amount of units and the training constraints imposed, the evaluation of the activations of the units throughout the dataset shall reveal the presence of neurons that never activate, being effectively inert, which indicates that the layer is too large. The next step is to reduce the hidden layer and to retrain the autoencoder to evaluate the change in the unit's activations and its effect on the reconstruction error. The number of units is decreased according to a fixed percentage of the initial size of the layer in order to follow a linear rate. This process can be expressed as an optimization problem (Eq. 4.10), where $\boldsymbol{J_e(n)}$ is the reconstruction error, $\boldsymbol{\delta(n)}$ corresponds to the presence of inactive units, determined by their mean activation value not surpassing an activation threshold and $\boldsymbol{n}$ is the parameter to be selected, representing the number of units. By repeating this process, the optimal dimension of the hidden layer is determined empirically, conducting to the equilibrium between the presence of inactive units and the error.

$$\min \{ \; f(n) = J_e(n) + \delta(n) \; \}, \;\; n \in \mathbb{R} \qquad \textit{Eq. 4.10}$$

When the training of the first autoencoder has resulted in a suitable hidden layer, its decoder stage is detached and a new autoencoder is defined where the hidden layer of the first autoencoder becomes the input layer of the second autoencoder. The input data used in the first autoencoder is transformed using its encoder stage in order to obtain the dataset for the next autoencoder. Then, this training process is carried out sequentially to tune new layers until sufficient layers have been created. In order to evaluate if sufficient layers are present, the following criteria are employed: as successive layers become smaller due to the implemented dimensioning strategy, the creation of new layers is stopped when the next layer is unable to further reduce its dimension without increasing the reconstruction error, or when a pre-established maximum number of layers is reached.

The result is a deep autoencoder composed of stacked sparse autoencoders, where the data flows through the layers causing the activation of the internal feature detectors that have adapted to the internal structure of the data.

4.3.2 Supervised target mapping

The deep autoencoder has performed the discovery of features in the structure of the data. Nevertheless, this process was carried out in an unsupervised manner, which means that the features are not correlated yet to the performance of the machines. The next stage of the methodology is to take advantage of the feature learning stage in order to map the target variables to these features. In this case, the target variables are defined as the power consumption and power production of the equipment, and the coefficient of performance at the current operating point.

This mapping of the discovered features to the targets is implemented as a feedforward multilayer perceptron (MLP) that transforms the latent space data representation to the required output values. To train this additional structure, the input data of the network is passed through the deep autoencoder to calculate the value of the latent space for each input sample. The new dataset composed of the transformed inputs and the targets is then used for training the MLP by means of standard backpropagation.

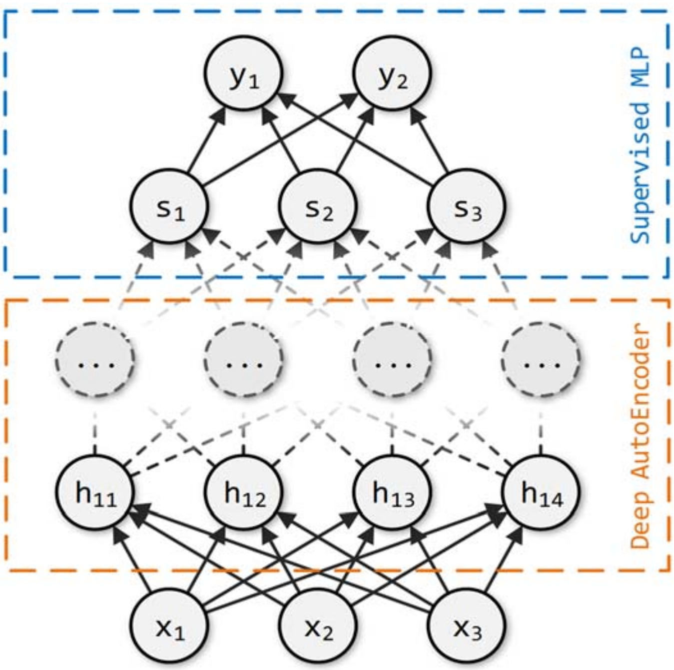


Fig 4.5 *Coupling of a supervised multiple-layer perceptron to the encoder stage of an unsupervised deep autoencoder.*

After this supervised tuning step, the MLP is coupled to the encoder stage of the deep autoencoder so that the inputs flow through the encoding layers to calculate the features and then through the MLP to map them to the targets. The final structure is shown in Fig 4.5, where $\boldsymbol{x}$ are the inputs of the deep network, $\boldsymbol{h}$ are sparse-activated neurons of the unsupervised autoencoder composed of $\boldsymbol{N}$ layers, $\boldsymbol{s}$ are the neurons of the $\boldsymbol{M}$-layer MLP and $\boldsymbol{y}$ are the outputs of the network, which now correspond to the dataset's target variables. After the coupling of both networks, supervised backpropagation is then also applied to the complete set of parameters of the resulting structure in order to fine-tune them. This last step may improve the modeling's accuracy in some cases, as the initial layers that performed the feature learning in an unsupervised manner, may experience slight adjustments when presented with the target data.

Finally, the accuracy of the performance mapping model obtained is evaluated using the following set of error metrics: the Root Mean Squared Error (RMSE), the Mean Absolute Percentage Error (MAPE), the Mean Absolute Error (MAE), the Determination Coefficient (R^2) and the Maximum Error (MAX).

4.4 Experimental implementation and validation

This section shows the implementation of the proposed operational performance modeling methodology and discusses the obtained experimental results in the test environment described in *Annex 1. Test environment*.

As it has been aforementioned, the objective is to build a deep neural network based model to characterize the equipment's power performance in terms of electric power consumption, thermal power production and coefficient of performance of the machines as a function of their operating conditions. Two cases are defined for the evaluation of the proposed methodology. A complex case is studied where multiple machines are present in an installation but only general power metering is available, which is a common scenario due to instrumentation costs. Additionally the considered equipment is heterogeneous, presenting different efficiency curves, which further complicates the problem but is also a general issue [139]. A simpler case deals with a single machine scenario, having dedicated power consumption and power production meters installed. Thus, the complex case considers all the cooling equipment, consisting of two chillers (CH1, CH2) and two heat pumps (HP1, HP2), while the simple case considers only chiller CH1.

A four-month dataset was employed for the implementation and validation of the methodology by recording the operation of the building's cooling equipment over the summer of 2017. Since the acquired period encompasses the full cooling season, the resulting dataset comprises a wide range of operating conditions, from mild weather and low power demand to warm weather and high power demand, at various machine operation combinations and different levels of occupancy in the building.

4.4.1 Unsupervised feature learning

During the feature learning stage, the unsupervised layer-wise pre-training strategy is followed in order to build the deep autoencoder. This process involves creating and training the individual layers, which must be properly dimensioned in order to facilitate the reduction of the reconstruction error. The defined dimensioning strategy is followed, setting an initially high number of units and evaluating the activations of the hidden layer units in order to determine the optimal size for each layer, according to the sparsity and concurrency constraints. Regarding the sparsity constraint, a value between 0.05 and 0.10 is generally recommended, while in the case of concurrency, a suitable setting is between 2 and 10, but should be determined by an expert user as it relates to the complexity of the problem, given by the amount of machines and their modes of operation.

The result of the dimensioning strategy is shown in Fig 4.6, where the reconstruction errors when varying the number of units in the hidden layers are presented. As it can be observed, the initial size of the layers is set to 100 neurons, which is progressively decreased in steps of 5 units (5% of the total), causing little effect to the reconstruction error at first, which is stable during the first layer reduction iterations, due to the presence of inactive units. As the layer's size is reduced, the training process is further constrained in terms of the amount of available units to distribute the structure of the data, leading to the increase of the reconstruction error.

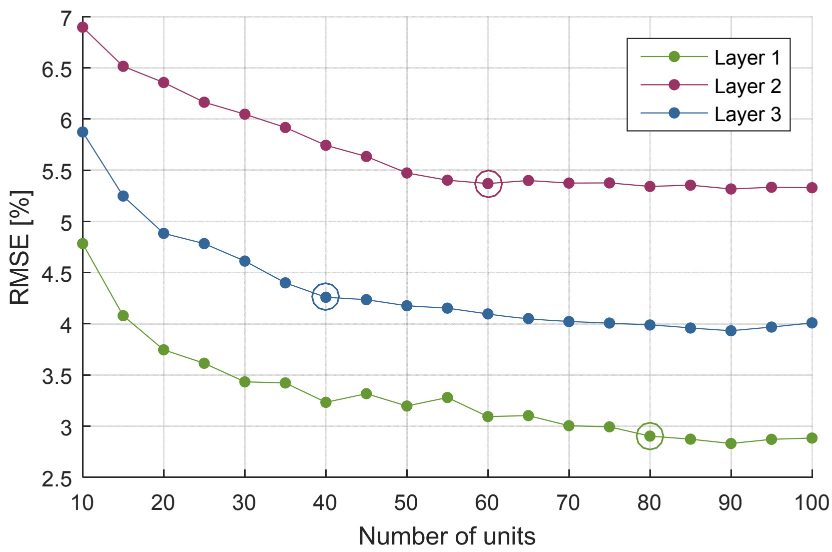


Fig 4.6 *Comparison between the reconstruction error and the number of units present in each autoencoder layer with a sparsity of 0.07 and a concurrency of 5.*

In the multiple-machine case, shown in Fig 4.6, the examination of the presence of inactive units established the threshold at 80 units for the first autoencoder layer, at 60 units for the second autoencoder layer and at 40 units for the third layer. A slight slope change can also be observed at this point, where furtherly reducing the size of the layer results in a larger error increase. To assess the necessity of additional layers, the behavior of each new layer is compared to the previous layer. In this case, the process was stopped at the third layer because the next layer presented the same behavior, considering that inactive neurons disappeared when 35 units were configured, and the reconstruction error steeply increased. In the case of a single machine, three layers were also used, with sizes 33, 21 and 15 for the first, second and third layers, respectively, with a decrease step of 3 units per iteration from a starting size of 60.

Thus, the number of hidden units in the single-machine case was smaller, which is expected due the number of units necessary being closely related to the complexity of the data.

After the training of each layer individually, the full autoencoder can be evaluated in order to determine the end-to-end reconstruction error, 13.70% RMSE and 10.49% MAE in the multi-machine case, which are indeed poor results when compared to the direct training of a deep autoencoder with the same layer configurations, resulting in 10.38% RMSE and 7.07% MAE. However, a fine-tuning step is applied to the layer-wise pre-trained deep autoencoder using the tuned parameters as a initialization values, which further reduces the reconstruction error to 5.64% RMSE and 3.96% MAE for the validation set, leading to a great improvement of the reconstruction performance. Similar results were obtained in the single-machine case, where the error was reduced from 10.07% RMSE and 7.09% MAE during pre-training to 6.12% RMSE and 3.82% MASE after fine-tuning, compared to the direct training of the network without the pre-training implementation, which resulted in 8.26% RMSE and 6.09% MAE.

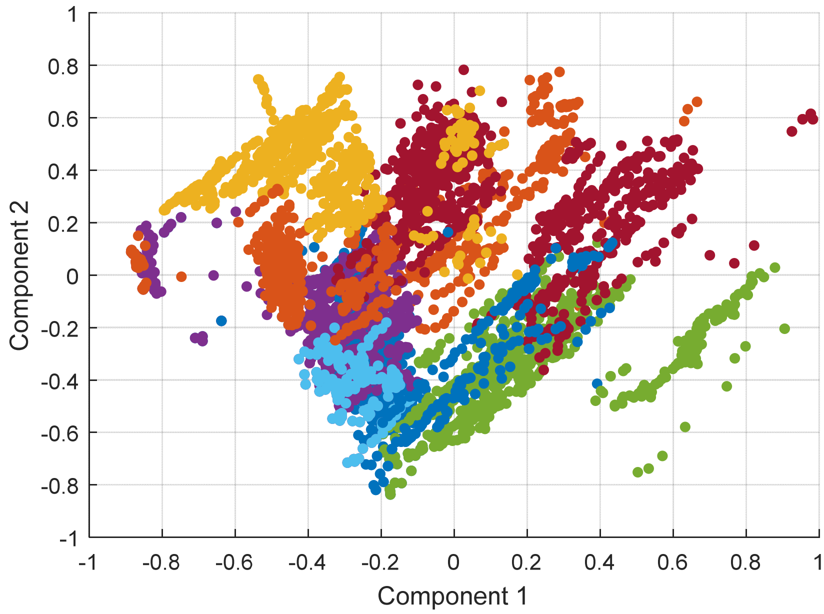


Fig 4.7 *Representation of the activation of a subset of neurons over the first two components of a principal component analysis.*

The trained autoencoder can then be evaluated in order to study the features that the unsupervised learning has discovered. To demonstrate this, the color-coded activations of a subset of hidden units from the first autoencoder layer in the multi-machine case are rendered over the samples of a principal component analysis (PCA) in Fig 4.7, where each color represents the activation of a single unit across the data space. The first two PCA components are used for illustration purposes, accounting for 72% of the accumulated variance. As it can be observed, several clusters of samples materialize, each corresponding to the activation of a different hidden unit. Therefore, the desired effect is achieved, where the units have scattered over the dataset according to two key concerns: the sparsity constraint forces the units to seldom activate, restricting their activation throughout the dataset, and the concurrency constraint discourages simultaneous activations in the same area of the data space, making the units disperse and cover different regions. It is important to note the reversibility of this transformation, which allows the reconstruction of the input data.

The activations of the hidden units are compared to the target variables in Fig 4.8. As it can be observed, many of the units' activation events can be correlated to changes in the target variables, even though no target information has been shown to the network at this point. Furthermore, while the meaning of some units may only make sense in composition, their activation can be interpreted in some cases, such as in regions where neurons are active during similar target values. For example, the activations of unit 7 can be traced to samples with high production values, while the activations of unit 4 can be traced to samples with low production values in the multiple machine case. It is important to note that this effect can consistently be observed, even though the units are repositioned each time the training process is repeated, indicating that the hidden units are able to discover features in the data.

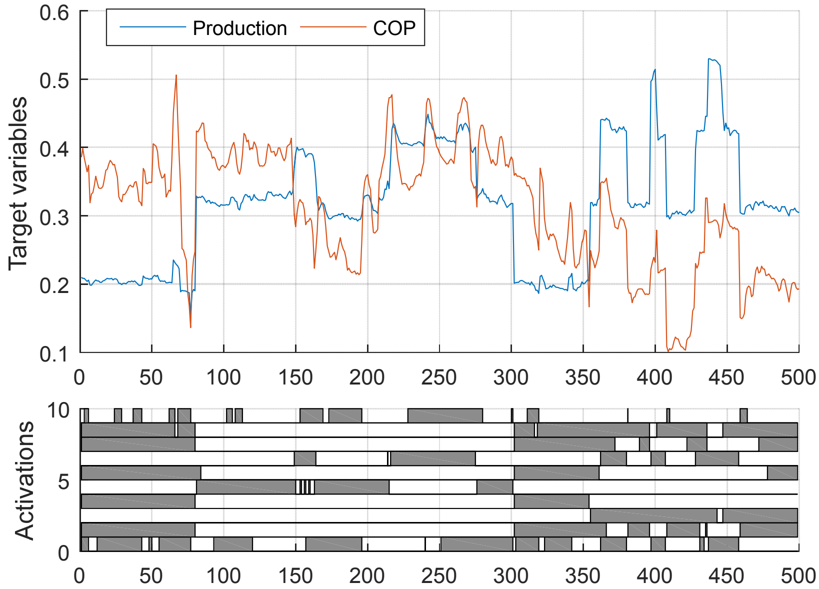


Fig 4.8 *Comparison between the activation a subset of the eighty hidden units in the first layer and the target production and coefficient of performance.*

4.4.2 Supervised performance mapping

During the supervised performance mapping, the feature learning stage is coupled to an MLP in order to map the discovered features to the desired outputs. The supervised structure is connected at the middle layer, representative of the latent space learned by the unsupervised training, replacing the decoder stage of the deep autoencoder. First, the encoded representation of the input data is obtained by means of the application of the encoder stage. Then the supervised layers are trained with the objective of mapping the encoded data to the performance data, consisting of electric consumption, thermal production and coefficient of performance. Finally, the full structure is fine-tuned in order to improve the accuracy further.

4.4.2.1 Case with a single machine

The result of the supervised training consisting in a two-hidden layer MLP applied to the single-machine case is shown in Fig 4.9, where it can be observed that the network model accurately follows the real performance characteristics of the machine.

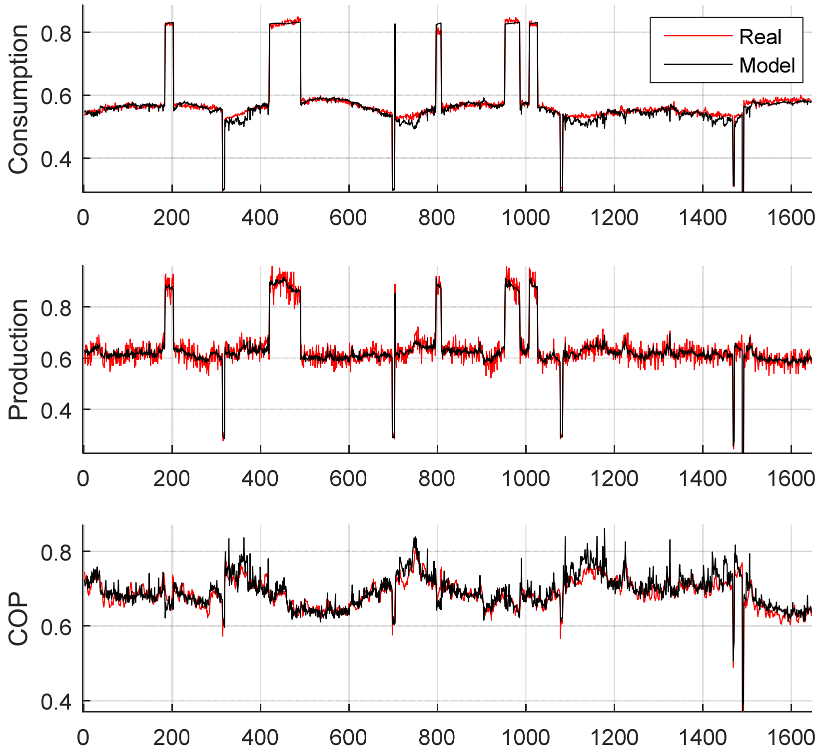


Fig 4.9 *Comparison between the real targets (consumption, production, COP) and the deep neural network model for the single machine case.*

The accuracy of the proposed methodology has been compared to a traditional neural network approach consisting in a MLP with two hidden layers, of dimensions 6 and 4, which directly mapped the inputs to the performance targets without the unsupervised autoencoder pre-training. The accuracy of both approaches has been evaluated by means of the defined error metrics. A comparison of the results is presented in Table 4.1. As it can be observed, the proposed method achieves a substantial improvement of all the error metrics, decreasing the mean RMSE and MAE by 20.11% and 23.41%, respectively, and increasing the coefficient of correlation by 2.70%. Another observation is that the COP prediction is not as accurate as the consumption or production, which is expected due to the COP being calculated as a quotient instead of measured.

	Consumption		Production		COP		Mean	
	A	B	A	B	A	B	A	B
RMSE	2.24	3.02	1.73	2.40	4.72	5.46	2.90	3.63
MAE	1.58	2.30	1.30	1.69	2.89	3.57	1.93	2.52
MAPE	15.06	13.96	6.28	6.61	5.90	6.77	9.08	9.11
MAX	21.55	34.35	10.42	29.66	67.59	75.12	33.18	46.37
R^2	0.9863	0.9713	0.9913	0.9831	0.8886	0.8366	0.9554	0.9303

Table 4.1 *Single machine case: accuracy comparison between the proposed method (A), and a classical NN based implementation (B).*

A histogram representation of the absolute value of the error is shown in Fig 4.10, where the distribution of the error can be observed. It is worth mentioning that besides the reduction of the overall error metrics, the standard deviation of the error was also decreased, presenting fewer samples with large errors. This could be due to the nature of the sparse autoencoder implementation, where a large amount of neurons is distributed over the dataset, therefore broadening the focus of the network, and successfully learning more complex features of the data.

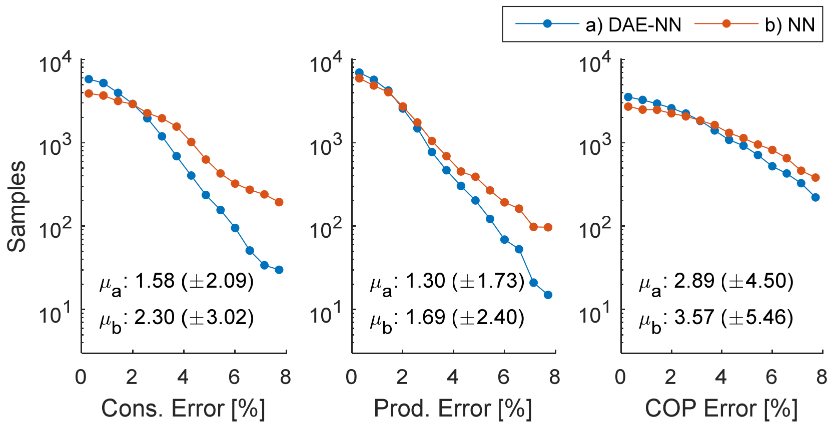


Fig 4.10 *Error distribution comparison between a) the proposed deep autoencoder approach and b) a traditional neural network, for each of the outputs: consumption, production and coefficient of performance.*

4.4.2.2 Case with multiple machines

In this case, the methodology is applied to the group of four machines as a whole, assuming that partial energy metering is unavailable. The inputs are defined as the pooling of the inputs of each of the machines, while the outputs are defined as the electric consumption, thermal production and coefficient of performance, as observed by the general meters. The result of the supervised training of the additional two-hidden layer MLP is shown in Fig 4.11, where it can be observed that the deep network also models the real performance characteristics of the machine with good accuracy, even in samples that present drastic changes, which are a typical affliction of modeling methodologies.

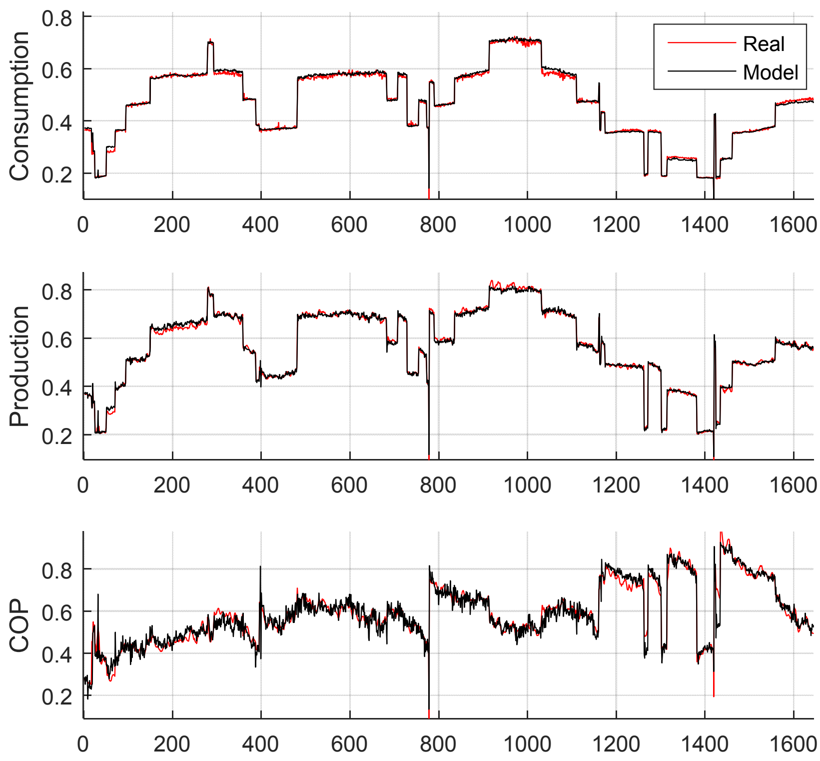


Fig 4.11 *Comparison between the real targets (consumption, production, COP) and the deep neural network estimation for the multiple machine case.*

An important observation can be made by examining the COP signal, which shows the normalized value of the performance. It can be observed that the effective COP of the group of machines presents high variability in different periods, even between periods that present similar values in the consumption and production signals. This effect is caused by the composite operation of machines of different technologies which are being controlled at various load ratios by a controller that does not take optimal operating points into consideration. In fact, the control strategy of the HVAC controller in this test environment is designed with the objective of evenly distributing the time of operation so that all the machines age at an equal rate. Thus, the availability of accurate performance maps that reflect the real characteristics of the installation could be a great asset in the development of control strategies that account for the variations in performance at different operating conditions.

In this case, the accuracy of the proposed methodology was also compared to a traditional neural network approach, consisting in a MLP with three hidden layers, of dimensions 14, 10 and 6. However, due to the large number of input signals, consisting in the pooling of the inputs of the four cooling machines, a feature reduction process was implemented by means of a genetic algorithm pruning the input space, which is a common approach in feature selection applications [140]. This process resulted in the elimination of all but one bus return temperature, as its behavior is largely the same for all machines with the addition of an offset, and a few other signals related to modes of operation that had low relative representation in the dataset. This is due to the characteristics of a traditional NN, which may have difficulty learning those behaviors that are less represented, in favor of those more common in the dataset. Thus, approaches like sparse autoencoders, where feature detectors are distributed over the data, may pose an advantage in cases that aggregate different patterns and modes of operation.

	Consumption		**Production**		**COP**		**Mean**	
	A	B	A	B	A	B	A	B
RMSE	1.59	2.21	1.47	2.69	5.98	7.91	3.01	4.27
MAE	1.07	1.64	1.05	1.97	4.33	6.07	2.15	3.23
MAPE	2.91	5.34	3.09	4.94	11.56	15.34	5.85	8.54
MAX	16.11	24.27	20.13	39.09	61.41	74.25	32.55	45.87
R^2	0.9946	0.9852	0.9956	0.9837	0.8957	0.8029	0.9619	0.9239

***Table 4.2** Multiple machines case: accuracy comparison between the proposed method (A), and a classical NN implementation with GA feature selection (B).*

The accuracy of both approaches has been evaluated by means of the same error metrics. A comparison of the results is presented in Table 4.2. As it can be observed, the proposed method achieves a greater improvement of all the error indicators, decreasing the mean RMSE and MAE by 29.51% and 33.43%, respectively, while increasing the coefficient of correlation by 4.11%.

The error distribution analysis, shown in Fig 4.12, reveals similar results in this case. The deep autoencoder-based methodology outperforms the traditional neural network approach with feature selection implemented by means of genetic algorithm, both in terms on mean absolute error and the standard deviation of the error distribution. In this case, the consumption and production outputs achieve great accuracy, even though this case presents more complexity because each of this signals is the aggregate composition of several machines.

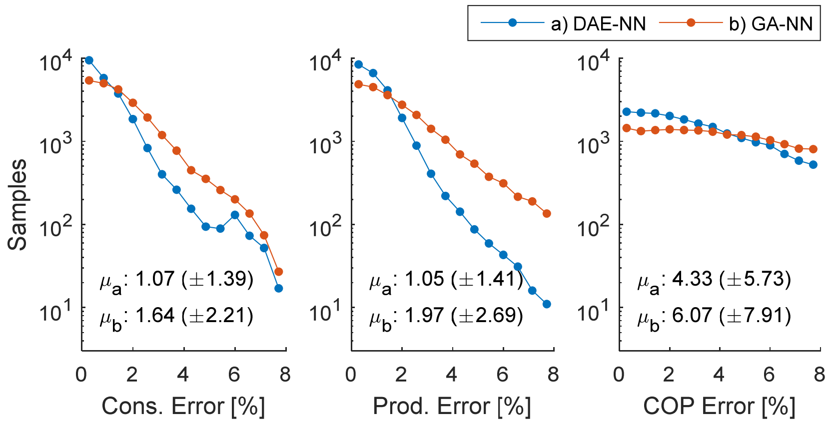


Fig 4.12 *Error distribution comparison between a) the proposed deep autoencoder approach and b) a traditional MLP with GA feature selection, for each of the outputs: consumption, production and coefficient of performance.*

4.5 Discussion and conclusions

A data driven operating performance modeling methodology for the characterization for HVAC production equipment based on a deep learning approach is presented in this chapter, aligned with the state of the art regarding the analysis of performance of equipment for modeling and control applications. The proposed method consists in the implementation of a feature learning stage by means of stacked sparse autoencoders, which are able to pre-train the initial layers of a deep neural network in an unsupervised manner. The autoencoders are then followed by a supervised stage where the learnt features are mapped to the target variables.

The implementation of the proposed methodology has been shown to improve the accuracy of the modeling of the behavior of a machine when compared with a traditional neural network approach. The defined error metrics were improved up to a 23% in the single machine case during the estimation of the electrical consumption, thermal power production and coefficient of performance based on the machine's operation conditions. Besides the single machine case, a more complex multi-machine case was also studied, where a group of machines was treated as a whole with the objective of estimating the group's total consumption, production and performance, therefore including cross-effects that would otherwise not be accounted for. The improvement over the traditional neural network, this time complemented by a feature selection stage implemented by means of genetic algorithms due to the large amount of input variables, was larger than in the single machine case, with up to a 33% improvement of the defined error metrics, indicating that the feature learning approach is better suited when complexity grows.

In addition to increased accuracy, the proposed methodology presents other advantages, such as the ability for feature introspection, consisting in the study of the sparse detectors in order to determine the conditions that caused their activation, and leading to the potential uncovering of the causes of a given behavior. Furthermore, this approach helps to reduce the initial feature selection and engineering effort by embedding it in the network, decreasing the pre-processing needs to a minimum. However, deep learning approaches do present a significant drawback, consisting in the increase of hyper-parameters that need to be properly adjusted for the modeling to be successful, such as the number of layers and the amount of units per layer. The tuning of such hyper-parameters is a non-obvious task, which was mitigated in this case by the implementation of a dimensioning strategy that supported the selection a proper network configuration. For these reasons, the deep learning-based modeling by means of stacked sparse autoencoders has been validated as a suitable approach for the modeling of the performance of HVAC equipment in real operating conditions.

5.

Predictive control of chiller groups

The state of the art relating to solving the optimal chiller loading and sequencing problems is critically reviewed in order to highlight the potential avenues for increasing energy efficiency, and a novel control strategy is developed by taking advantage of the previous thermal demand forecasting modeling and equipment performance modeling to implement a multi-objective predictive control solution.

CONTENTS:

5.1 Introduction

5.2 Formulation of the optimization problem

5.3 Chiller group control methodology

5.4 Experimental implementation and validation

5.5 Discussion and conclusions

5. Predictive control of chiller groups

This chapter performs a review of the state of the art relating to solving the optimal chiller loading and sequencing problems in order to highlight the potential avenues for improving energy efficiency. Then, a novel control strategy is developed by taking advantage of the previous thermal demand forecasting model and equipment performance model to implement a multi-objective predictive control solution.

5.1 Introduction

This section introduces the background and motivation for pursuing this line of research, reviews the state of the art related to control applications for chiller groups focusing on the chiller loading and sequencing problem, and describes the innovative contributions of this work.

5.1.1 Background and motivation

Considering the goal of increasing the energy efficiency of buildings, production equipment such as chillers are one of the areas that present the greatest potential for causing a significant impact [141] since these are one of the largest consumers in buildings, especially in the residential and tertiary sector where they can account for up to 40% of the building's energy consumption [142]. This is especially true in multi-chiller plants commonly found in medium to large buildings, where proper chiller loading and coordination is critical for increasing efficiency, and achieving an optimal control solution is of great interest with many studies attempting to tackle this issue with a variety of methods [143].

The potential for improvement is due to the fact the Coefficient of Performance (COP) of production equipment is not uniform throughout their operation range, meaning that differences in the control strategy of multi-chiller systems can lead to significant changes in the resulting energy consumption [64,144]. Indeed, a recent study concluded that up to 70% of the annual power consumption occurred while the chillers were operating at low Partial Load Ratio (PLR) which is the range were the coefficient of performance of chillers is typically at its lowest [145], while a study evaluating ideal operation levels concluded there was a potential for improvement of up to 23.4% [146]. Thus, a strategy able to optimally control the chillers' operation has the potential to increase the operational efficiency of the overall system, minimizing the energy consumption by ensuring that each machine operates at its optimal COP while considering the affecting operating state, like weather conditions, inlet temperature from the distribution bus and future load demand.

5.1.2 State of the art

The chiller sequencing and optimal loading problems refer to the necessity to find a strategy for coordinating the operation of a group of chillers in order to meet the cooling demand while minimizing energy consumption, in terms of the loading ratio of each chiller [147] and the sequence at which they should be turned on or off according to the cooling load requirements [148]. A thorough review of the related state-of-the-art reveals that this is a complex problem having several facets that need to be carefully considered in order to achieve a performant solution. The main key aspects identified are: i) the need of chiller performance characterization; ii) the objective function of the controller and iii) the choice of optimization strategy and implementation.

Regarding the chiller performance characterization, chiller sequencing and optimal loading methodologies rely on the capability of mapping control actions to expected performance in order to select the chiller’s settings that lead to their operation at optimal COP, given a set of operating conditions [149]. The most common approach in recent methodologies is the utilization of manufacturer-provided datasheets that specify the performance according to specific operating conditions or the simulation and modeling of the equipment using software tools like TRNSYS [150]. For example, performance data obtained from datasheet lookup tables were employed for implementing a predictive HVAC controller based on a mixed-integer approach, which has the downside of providing only static figures and not considering all affecting parameters [98]. Instead, a study focused on the determination of a probabilistic approach for chiller replacement discussed the utilization of simulation compared to other methods, concluding that these could be viable for energy estimation if calibrated using measured data [151]. Even though the employment of such solutions in control applications is useful for research purposes, in practice, it is essential to have performance maps that accurately reflect the actual behavior of the equipment considering the multiple affecting factors besides the PLR, such as weather conditions, operating state and aging. In contrast, a methodology for optimizing the operation of a chiller plant employed a data-driven approach based on the modeling of the chiller group using a neural network and the implementation of a two-level algorithm, which allowed to achieve energy savings of 14% under simulation [116].

Another key aspect of the problem is the selection of a proper objective function for the optimization process. The de facto standard objective function is based on the aggregate power consumption of the group of chillers, being a useful strategy when the minimization of the used energy is the ultimate goal of the solution [150]. However, the need to tackle other concerns indicate that a multi-objective strategy may be preferred. In particular, some studies make reference to the switching problem, which can be described as the frequency and magnitude of changes to the relevant control settings. An application of real-time HVAC optimization looked into minimizing the disturbances caused by controller actions, comparing rate-limited setpoint reset to controlled step-changes [152]. The conclusions reflect that frequent changes to control variables may lead to system instability, especially when simultaneous and having large magnitude, thus showing that in practice limiting the amount of changes and moderating their delta is desirable. Furthermore, another argument for considering the minimization of switching is that increased control changes may lead to energy losses and faster equipment degradation due to the dynamics of the equipment and their mechanical wear [153].

Finally, when suitable performance maps are available and an appropriate objective function is determined, an optimization algorithm needs to be applied to carry out the control of the group of chillers in an efficient manner. Mainly two types of methodologies can be found on the literature, i) generic global optimization tools and ii) specific heuristics-based controller implementations. Generic global optimization tools are common in chiller control applications, a study of air-cooled chillers optimal control used random forests to implement an empirical model of the chillers and then applied generic algorithms to carry out the estimation of the optimal values of the control parameters [154]. Similarly, particle swarm optimization was employed to adjust the control parameters of a water-cooled chiller plant, simulated using Modelica models adjusted using empirical data in [155]. However, controller implementations of this type are slow and have a randomness component, thus both their results and runtime are nondeterministic by nature. Instead, other researchers have focused on the design of control algorithms that implement specific heuristics, making the controllers more computationally efficient and robust, which is a desirable property of the system even though the global optimum may not consistently be achieved by this means [156].

5.1.3 Innovative contribution

In consideration of the described shortcoming in the state of the art solutions for the chiller loading and sequencing problem, this chapter proposes a novel methodology for the optimal control of chiller combining data-driven performance map modeling, a thermal demand forecasting methodology and a multi-objective model-predictive controller implementation.

Specifically, the originality of this work consists of the following key aspects:

- The integration of data-driven COP maps of the involved equipment, defining their performance depending on their operating conditions and considering several affecting factors (multi-variate approach);

- The short-term forecasting of the building’s future thermal demand, considering affecting factors such as the weather and the building’s occupancy patterns, to provide accurate and reliable demand requirements to the optimization stage;

- The consideration of the thermal dynamics of the building, taking advantage of the distribution system’s thermal capacity, recalculating and adjusting their operating setpoints for scheduling the operation over the time horizon;

- The definition of a strategy considering multiple criteria for the determination of the optimal control sequence of the chillers, including the minimization of energy usage and switching.

5.2 Formulation of the optimization problem

The mathematical formulation of the optimization problem can be considered as a multi-period nonlinear problem and it can be described as the determination of the optimal operating set-points $\boldsymbol{v_i}$ of the HVAC equipment, for each time instant $\boldsymbol{t}$ of the optimization horizon, where $\{\boldsymbol{t} \in \mathbb{N} \mid \mathbf{1} \leq \boldsymbol{t} \leq \boldsymbol{K}\}$, with objective to satisfy the thermal energy demand of the building (L), while minimizing a multi-criteria function ($\boldsymbol{f^{trans}}$) and satisfying the established operating bus temperature thresholds ($\boldsymbol{T_{min}}$, $\boldsymbol{T_{max}}$).

Minimize: $$\sum_{t=1}^{K} \sum_{j=1}^{C} f_j^{trans,t}(v_i^t) \qquad \text{Eq. 5.1}$$

Subject to: $$L_t - \sum_{i=1}^{n} (P_i^t * \eta_i) = 0 \qquad \text{Eq. 5.2}$$

$$T_{min} \leq T_t \leq T_{max} \qquad \text{Eq. 5.3}$$

$$\sum_{i=1}^{n} P_i^t \leq P_{max}^{grid} \qquad \text{Eq. 5.4}$$

$$\underline{P_i} \leq P_i^t * \eta_i \leq \overline{P_i} \qquad \text{Eq. 5.5}$$

In the above formulation, $\boldsymbol{C}$ describes the number of the optimization criteria, $\boldsymbol{n}$ describes the number of the HVAC equipment, $\boldsymbol{T_{min}}$ and $\boldsymbol{T_{max}}$ indicate the lower and upper temperature bounds for the distribution bus, while $\boldsymbol{P_{max}^{cap}}$ describes the maximum cooling capacity that the equipment can supply, and $\boldsymbol{P_i}$, $\underline{\boldsymbol{P_i}}$, $\overline{\boldsymbol{P_i}}$ and $\boldsymbol{\eta_i}$ describe the energy consumption, minimum power generation, maximum power generation and COP of the equipment $\boldsymbol{i}$, respectively.

For the formulation of the multi-objective function, the weighted global criterion method has been used, in which all the objective criteria are combined to form a single criterion. In order to sum the results of each criterion, a unit normalization function is made, as described in (Eq. 5.6).

$$f_j^{trans} = \frac{f_j(v_i) - f_j^o}{f_j^{max} - f_j^o} \quad \text{Eq. 5.6}$$

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

5.3 Chiller group control methodology

A step-by-step diagram of the proposed optimization and control procedure is shown in Fig 5.1, which is composed of three stacked control loops: 1) operating mode selection, 2) predictive control strategy and 3) load profile reconfiguration.

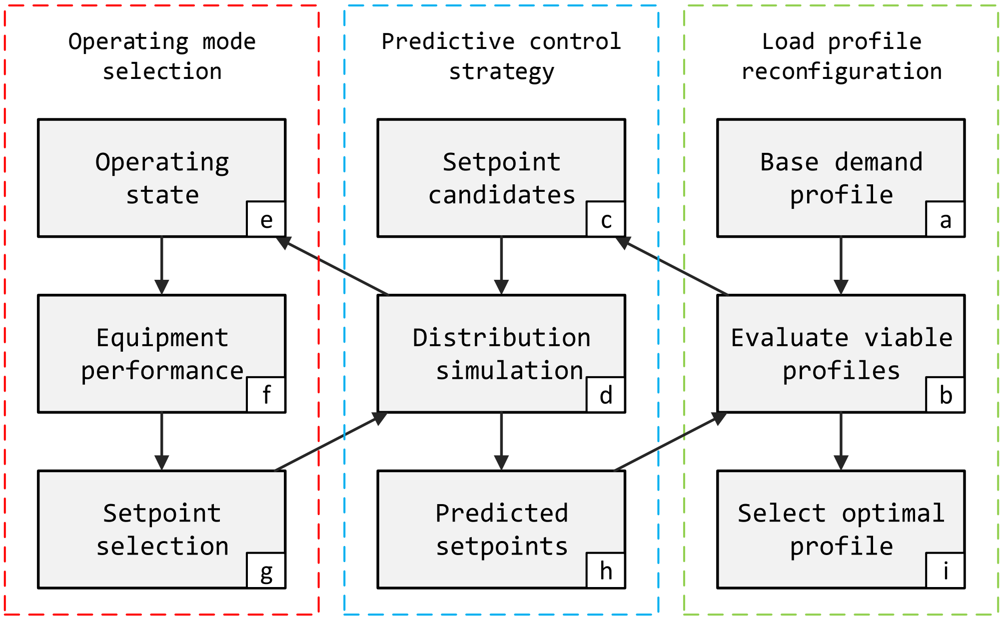


Fig 5.1 *Step diagram of the implementation of the proposed three-stage chiller operation optimization and control framework.*

5.3.1 Operating mode selection

This is the inner-most control loop, which relies on the performance map obtained by means of the neural network model to estimate the cooling capacity, consumption and performance of viable setpoints in order to select the optimal control state for the current instant.

This stage is executed as follows. First, the current operating state is evaluated (e) by acquiring the signals affecting the control: inlet and outlet temperatures, external temperature, humidity, current cooling demand and previous operating state of the group of machines.

Afterwards, the setpoint candidates are determined by establishing the possible control actions. This is achieved by exhaustive combinatorial of control states of the individual machines. For machines having discrete operating modes, the full range of modes will be employed, but for machines having a continuous operating range this is not possible, so the operating range is discretized up to the desired granularity. In practice, coarse granularity shall be sufficient, given that HVAC equipment cannot be regulated with infinite precision. Each combination is then evaluated in conjunction with the current operating state in order to determine the equipment's performance (f) by feeding it through the performance model. The full evaluation of the exhaustive combinations is possible due to the runtime of this evaluation being instantaneous, even though the implementation and training of the performance model being a lengthy process.

Finally, having obtained the cooling capacity, electrical consumption and COP of each control setting, these are scored and sorted in order to select the locally optimal setpoint (g) according to the defined objective function, which considers performance and switching cost to apply the new setpoint.

5.3.2 Predictive control strategy

This loop is proposed to operate on top of the previous one, iterating over N steps in the prediction horizon to determine the optimal future sequence of control actions that will satisfy the forecasted load demand. Instead of selecting the best performing control action as determined in (f), the top setpoint subset is selected and used as starting setpoint candidates (c). Each of these setpoints is used as the starting point of a control sequence evaluated N steps into the future, which shall allow to determine which of the starting points leads to the most beneficial outcome.

This is motivated by the fact that control setpoints that are locally optimal for the current control iteration may not be globally optimal due to switching to a state that is suboptimal, because it might be preventing that more efficient states are reached later on. This is likely to happen due to the switching being considered in the objective function, which penalizes control actions that make drastic changes to the setpoints, thus selecting the absolute best control action for the next step might lead to suboptimal situations.

Therefore, the performance of a subset of the top performing candidate setpoints is applied over the optimization horizon to evaluate the full implications of taking that control action. However, this presents one main difficulty: the evaluation of future control actions is not possible due to the operating state being unknown.

The data necessary for evaluating a control setpoint using the performance map can be classified into three types:

- Known future state: part of the future operating state is known, i.e. the future cooling demand is obtained from the load forecasting model evaluation, and the external temperature and humidity can be queried from local weather services.

- Unknown future state: state that relates to the operating state of the machines, in this case each machine’s inlet and outlet temperatures.

- Control action: the setting of each of the machines in the equipment group, which is the optimization variable and is always known for the previous control iteration.

Thus, knowing the future state of bus temperatures is required to be able to apply the operating mode control loop during future iterations in order to determine the future control sequence. To solve this problem, a simulation model of the distribution bus is employed, implemented in this case as a first order energy storage model with a single capacity coefficient, estimated using the historical data available.

Using the bus model, a control sequence is determined by iterating from each of the setpoint candidates, simulating the application of the control action on the distribution bus (d) while keeping track of the COP achieved by each sequence over the optimization horizon. Each of the control sequences obtained is then evaluated using the objective function, which allows the final selection of the predicted setpoints (h), i.e. the optimal control sequence over the optimization horizon for matching the predicted cooling load profile.

5.3.3 Load profile reconfiguration

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

In order to ensure that the equipment on the consumption stage of the HVAC system are able to effectively draw power from the distribution bus, production stage controllers focus on maintaining the temperature of the bus within a certain range, i.e. the goal of the production controller is to keep the temperature of the bus between a minimum and maximum threshold.

Thus, the bus acts as the buffer between the production and the consumption stages, but due to the thermal dynamics of the building and associated thermal capacity, it allows the temporal decoupling of the production and consumption equipment, in a similar manner as dedicated energy storage equipment would allow. Therefore, the production setpoints can be altered as long as the temperature of the bus is kept within the required thresholds, which is a property of the system that is proposed to be exploited in order to increase the performance. By shifting the load of the production equipment in time it is possible to take advantage of periods of time where conditions are more favorable, for example conditions such as better COP due to affecting operating state or varying cost of energy.

This control stage handles the determination of a load profile that shall take advantage of favorable production conditions while satisfying the temperature constraints. To solve this problem, a global optimization tool like particle swarm optimization or genetic algorithm could be used to overcome the non-derivable nature of the problem. However, these tools may present robustness issues and could be impractical for cases that involve running simulations, which makes them inadequate for solving this type of control problem [157]. Instead, a heuristics technique is employed which searches the possible production profile configuration space according to the following approach: i) an initial load profile is determined by considering the production uniform throughout the control horizon in step (a), with a value equal to the average load demand in the same period, ii) the maximum and minimum production rate that would lead the system to exceed the temperature thresholds are computed and used as upper and lower bounds, respectively, and iii) the space is binary-searched up to a predefined amount of iterations evaluating viable demand profiles in step (b), ensuring that a solution is found in deterministic time. This is desirable property in control applications and important in this case due to the fact that evaluation of each configuration depends on the evaluation of the previous stages which includes simulations. Finally, the best performing profile and associated production control sequence are selected and applied only to the current instant in step (i), because the next control iteration shall recalculate and determine a new control sequence starting from this iteration.

5.4 Experimental implementation and validation

This section shows the implementation of the proposed loading and sequencing chiller control methodology and discusses the obtained experimental results in the test environment described in *Annex 1. Test environment*.

The objective is to demonstrate the performance gained by taking advantage of the integration of data-driven models with the controller and the effectiveness of the developed control strategy in increasing the overall energy efficiency. For this purpose, the stages the control strategy are evaluated in steps, showing the effect of each stage relative to the base control strategy.

For the implementation and validation of the proposed control strategy, a dataset of the equipment's operation was acquired by recording the operation of the building's cooling equipment during the summer period of 2017. The dataset comprises 120 days, from May 16th to October 27th of 2017, not including weekends, where each day is treated as a standalone validation case.

5.4.1 Operating mode selection

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

The result of operating the group of chillers with this strategy is shown in Fig 5.4, which shows a comparison between the actual cooling demand and the cooling power produced by the cooling equipment implementing the selected control sequence.

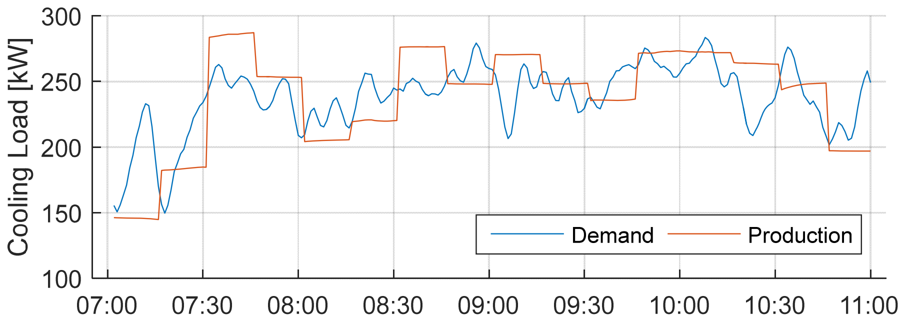


Fig 5.2 *Actual cooling demand compared to the cooling power production resulting from the application of the resulting from the application of the selected sequence of operating modes.*

As it can be observed in the figure, the equipment switches between operation modes to match the instantaneous cooling demand. At each control iteration, the operating mode that offers the best tradeoff between instantaneous performance and minimal switching is selected.

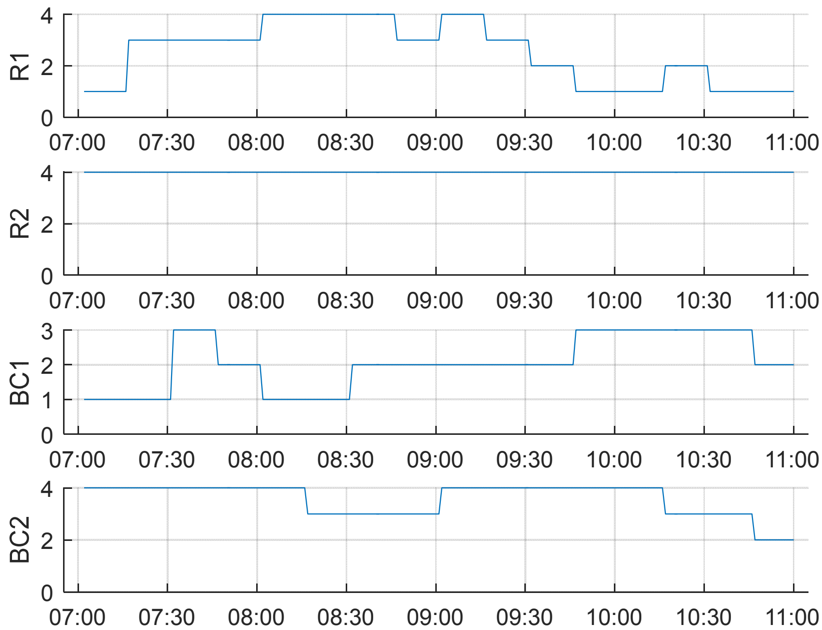


Fig 5.3 *Operating modes of the cooling equipment, obtained from the application of the first stage controller: the instantaneous operating mode selection.*

The commutations between operating modes can be observed in Fig 5.3. As it can be observed, the machines switch between operating modes to supply the instantaneous cooling power demand, however the switching between operating modes happens in a controlled manner, limiting both the rapid succession of switching and restricting the magnitude of changes. The resulting control sequence prevents the equipment from being forced to rapidly switch between low and high power production modes, which is desirable to avoid energy losses due to transient states and to lengthen the lifespan of the equipment.

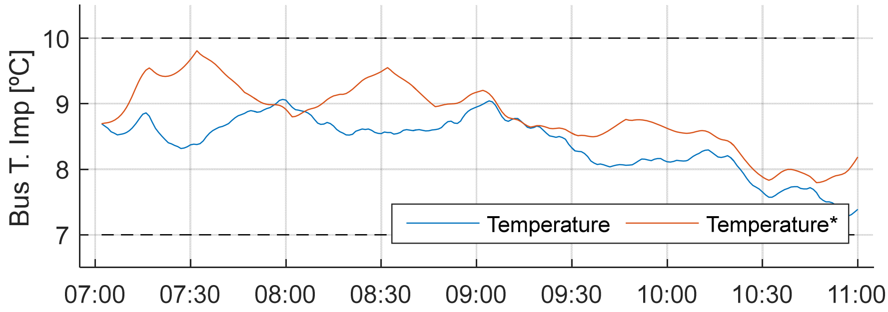


***Fig 5.4** Impulsion temperature of the distribution bus after the application of the selected sequence of operating modes*, compared to the actual bus temperature during the same period of time.*

The application of the selected control sequence results in a cooling power production curve being introduced into the bus. This production curve, combined with the cooling load causes the temperature of the bus to fluctuate. This temperature is shown in Fig 5.4, where the new, simulated, behavior of the temperature of the bus is compared to prior, measured, behavior of the temperature of the bus. This new temperature curve presents a similar behavior, still being maintained within operational constraints, but presenting deviations from the original curve. This is expected due to the new selection of equipment control actions, which selects different operating points that produce cooling power in a similar range but increased COP. As a result, the accumulated cooling power produced may not be exactly the same, thus the temperature of the bus at the end of the shown period becoming warmer in this case.

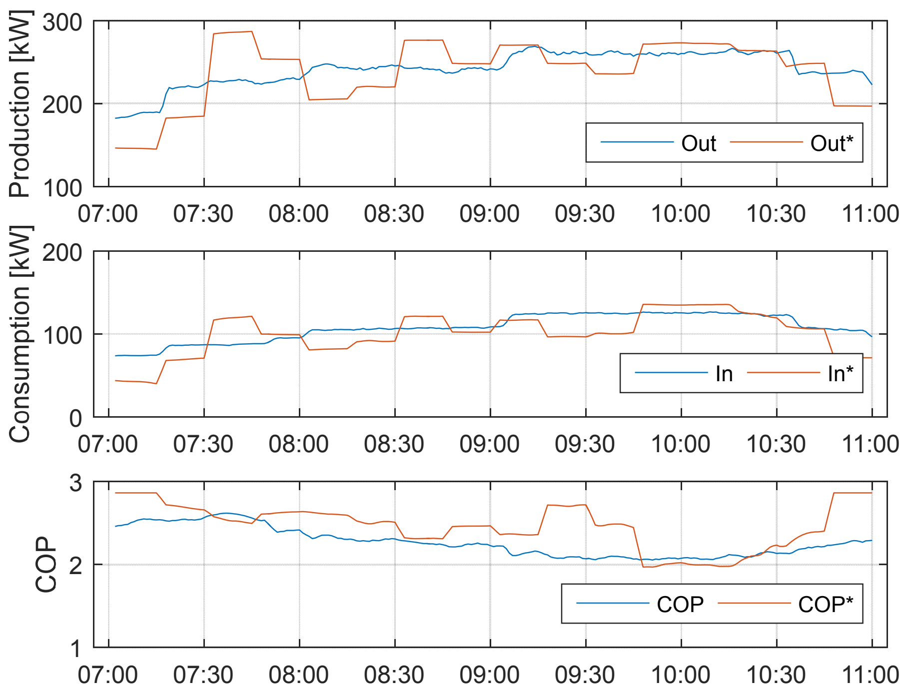


Fig 5.5 *Power production, consumption and coefficient of performance resulting from the application of the selected sequence of operating modes*, compared to the operation of the base controller during the same period of time.*

The comparison between the application of the old and new control strategies, in terms of energy production, consumption and COP are shown in Fig 5.5. During this period, the accumulated production value was reduced by 1.16%, which is the cause of the bus temperature becoming warmer at the end of the period. However, the new controller selects operating modes that overall are more efficient than the base controller, which is unaware of the performance of the equipment.

In summary, this controls strategy manages to supply cooling power to the system with increased average COP by using different control actions that lead to a different cooling profile. This new production profile performs differently in terms of bus temperature but is kept within operational constraints. This behavior can be consistently observed throughout the day, except for periods of time where the base controller selects the optimal setpoint by chance, or the switching cost from the current state becomes too great.

5.4.2 Predictive control strategy

This stage considers the determination of control sequence that matches the forecasted cooling demand for the prediction horizon, in this case of 1 hour because it allows to plan control actions with sufficient foresight given the dynamics observed in the building's dataset, which are in the range from two to three hours. The setting of the control action, i.e. operation mode of the equipment, is restricted to allow a change every 15 minutes to limit the amount of switching, so 4 control points are calculated within each control horizon.

The improvement that this stage offers is based on considering how the selection of an operating mode affects the outcome of the complete prediction horizon, instead of only the current control step. The result of the application of this strategy is shows in Fig 5.6, which shows the cumulative average performance of a set of initial control setpoint candidates as each one is evaluated over the prediction horizon.

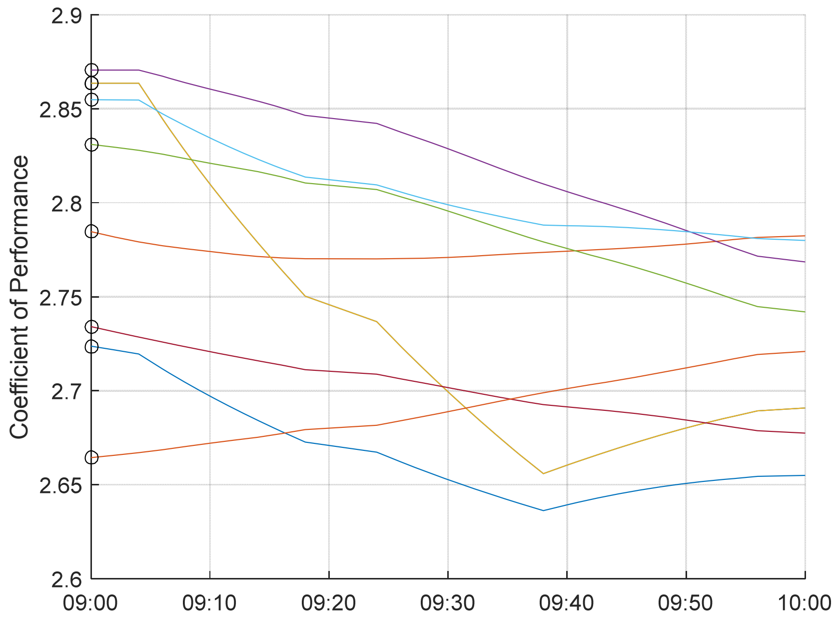


Fig 5.6 *Cumulative average performance of a set of setpoint candidates, evaluated over the prediction horizon.*

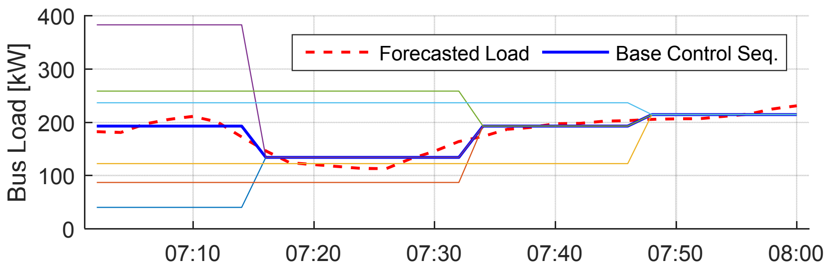
As it can be observed, setpoints that initially lead to the largest instantaneous performance are not necessarily the best option once the prediction range is evaluated. For example, the operating mode that offers the best COP at the start of the prediction horizon leads to lower average performance when considering the behavior of the full time window than the setpoint that is initially ranked fifth best. This behavior is due to the imposed switching constraint which discourages drastic changes between operating modes.

Therefore, this control stages achieves the desired effect, the anticipation of favorable and adverse conditions and the selection of a predictive control sequence that leads to the best average performance.

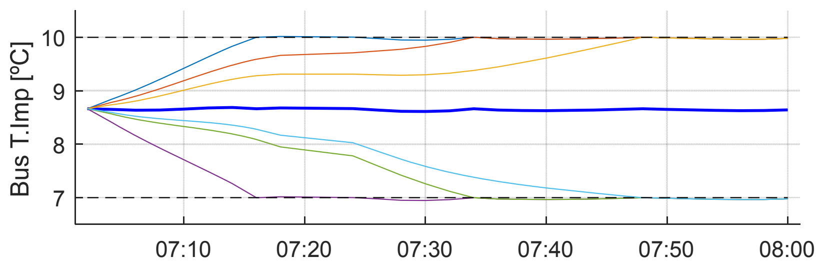
5.4.3 Load profile reconfiguration

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

As described, the load profile reconfiguration process begins with the determination of the initial control sequence profile, and the maximum and minimum production sequences that would keep the system operating within thresholds. Then these profiles are evaluated and binary search is employed for determining the best profile configuration, allowing up to a maximum number of iterations. In this case 4 iterations are employed, as it is a good compromise between precision and computation time. The result of the implementation of this profile selection strategy is shown in Fig 5.7 where it can be observed that different control sequences are evaluated within the temperature thresholds.



a) Control sequence candidates.



b) Bus impulsion temperature simulation.

Fig 5.7 *Evaluation of candidate control sequences that reconfigure the forecasted cooling demand.*

The forecasted cooling demand and the base control sequence, which matches in magnitude the average demand on each control period, are shown highlighted in Fig 5.7 a), while the other bus load profiles correspond to some of the attempted control sequences, having varying magnitudes within the ranges from the minimum viable production to the base sequence and from the maximum viable production to the base sequence. The actual production value is curtailed to ensure the bus temperature does not exceed the limits, considering that a control action can only be carried out every 15 minutes. The simulated bus temperature response when each of the control sequences in Fig 5.7 a) is applied is shown in Fig 5.7 b). As it can be observed, each simulation begins at the current temperature reading, and gets cooler or warmer depending on whether the magnitude of the sequence is above or below the base control sequence, respectively. The base sequence is stabilized at the initial temperature value, with slight fluctuations corresponding to the changes in the load, because it's values are calculated to match the average cooling demand per control period, while other sequences follow different trajectories with varying steepness depending on the magnitude of the production.

5.4.4 Energy efficiency improvements

A comparative summary of the results achieved by all three control stages is presented in Fig 5.8, which shows a histogram of the performance obtained when applying the relevant stage over single day periods, for each of the 120 days available in the dataset. The experimental results of the application of the three-stage control strategy consistently show a performance increase throughout the dataset.

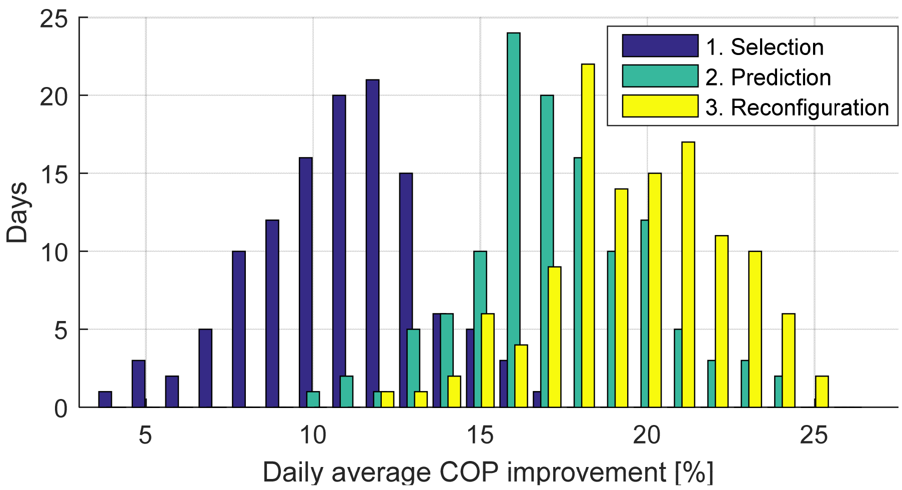


Fig 5.8 *Histogram of the daily average performance achieved by the different control stages, relative to the base controller.*

The application of the first controller stage, consisting on the selection of the operating mode based on current load demand, achieves an improvement of between 3.92% and 17.05% of the COP respective to the base controller, with daily average increase of 10.88%, surpassing 10% performance increase in a significant part of the considered dates. However, there's cases where this stage is unable to achieve such improvements, which is due to the base controller already operating on moderately efficient modes, and the inability of this stage to realize large improvements due to being penalized by the consideration of the switching in the cost function.

The application of the second stage, supported by the usage of the load forecasting capabilities to anticipate demand changes, achieves an improvement in the range of 9.84% to 24.11% respective to the base controller, with an average increase of 17.27%. This significant performance increase is due to the ability of this stage to consider the future control sequence, allowing the selection of control setpoints that lead to continued efficient operation while overcoming the switching minimization consideration.

Finally, the proposed control strategy is realized with the incorporation of the third stage, the re-configuration of the load profile to take advantage of favorable production conditions by taking advantage of the thermal dynamics. The application of the full strategy achieves a performance improvement in the range of 12.39% to 24.30% respective to the base controller, and an average increase of 19.54%. As described, the performance gain respective to the second stage is not as extensive as when comparing the second stage respective to the first one, in part due to the second stage already accomplishing a highly efficient control sequence, but also due to this installation not including dedicated thermal storage.

5.5 Discussion and conclusions

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

The framework is based on the selection of the control sequence that maximizes the performance of the production equipment in an HVAC installation by driving the machines to their most efficient setpoints, supported by the integration of data-driven models. A neural network-based model of the equipment's COP behavior respective to operating conditions and implemented using a deep learning approach is used for evaluating the response to potential setpoint candidates, while a model of the building's thermal load demand considering the building occupant's behavior is used for anticipating the energy production requirements.

The control sequence determination problem is formulated as an optimization problem, but finding a solution is a complex task, as the problem is not derivable and the exhaustive search of the solution space is infeasible. Instead of using a global optimization tool, a heuristics-based control method composed of three stages is designed and implemented achieving a substantial performance increase while maintaining a low and most importantly constant computational time.

The obtained results consistently show a performance increase by the implementation of the control strategy, with the complete solution achieving a 19.54% daily average COP increase with 2.68% standard deviation. These results are coherent with the efficiency improvement potential inferred by related studies in the state-of-the-art literature over current control solutions in established HVAC systems.

As future work, additional optimization criteria could be considered to further fine-tune the control sequence, for example the uniform utilization of the equipment could be enforced so that all of the machine's aging follows a similar rate, or the variable cost of energy which could be paired with the load profile determination to achieve greater economic savings.

6.

Conclusions and future work

This chapter presents the general conclusions of the research conducted in this thesis, outlining potential avenues for improvement in future works.

CONTENTS:

6.1 Conclusions

6.2 Future work

6. Conclusions and future work

This chapter presents the main conclusions of the thesis and outlines the future work.

6.1 Conclusions

This section presents the main conclusions of this thesis relating to the initial hypotheses and stated objectives.

As the introductory research topic section established, buildings take a large share of the world's total energy consumption and specifically HVAC systems account for a large percentage of that share. This, in part, is the reason for the growing interest in researching ways to increase energy efficiency in this kind of systems, as it could have a significant impact in lowering energy consumption.

The main hypothesis of this thesis states that by taking advantage of the wealth of historical operation information available in building management software suites it could be possible to identify the actual operating context and conditions of the underling subsystems. This information could be exploited in order to reach an optimal tradeoff between energy consumption and comfort trough the introduction of models and control schemes to tackle the performance gap documented in the literature, relating to the ideal operating performance and the performance observed in practice.

Initially, a review of the state of the art in the topic of HVAC energy management in buildings was conducted to find potential avenues for improvement. This prompted the identification of three main contribution areas, and the definition of a research plan to investigate how to improve the current solutions. After the definition of the research plan, a pilot plant was identified to be used as a test environment for the development of the thesis. The plant's HVAC installation and energy management software suite was studied and supplementary instrumentation was installed with the purpose of facilitating the development and validation of the thesis. Afterwards, the work concentrated on the contribution on these areas:

The first contribution belongs to the topic of load forecasting of thermal demand for HVAC systems in buildings. The review of the state of the art in this topic revealed shortcomings in terms of taking into account the operating context of the building for the calculation of energy demand forecasts, and the lack of solutions for estimating the thermal demand instead of the electrical consumption. In particular, the consideration of the behavior of the users of the buildings was found to be a concern with rising interest, and a modeling approach was proposed to take advantage of real-time occupancy data in order to build an activity indicator that could be integrated with the demand forecasting. Furthermore, an estimation method was proposed for the calculation of the actual power draw by the air handling units from the distribution bus, allowing the determination of the thermal power needs of the building, decoupling the effects of the control strategy and thermal capacity of the distribution bus.

The implementation of the forecasting methodology in the test environment included the study of the input candidates for modeling the power demand in order to obtain the set of variables that allows to accurately model this signal. This study was performed by comparing different aspects like the cross-correlation and frequency analysis of the signals. The comprehensive process resulted in a modeling methodology specifically tailored for thermal demand estimation in buildings which achieves significant accuracy and could support different applications.

The second contribution belongs to the topic of operation performance modeling of HVAC equipment in buildings. The review of the state of the art of this topic revealed that a wide range of applications rely on operating performance maps of HVAC equipment being available in order to perform their function. However, in frequent instances these applications are based on inaccurate or insufficient data due to the unavailability of proper equipment models that establish the relationship between their operating state, the control setpoints and their expected energetic characteristics, i.e. energy consumption, production and performance. A generic solution is required to approach this issue, as most HVAC installations are different and may include various types of equipment. Thus, a novel modeling methodology based on deep learning was developed which drastically reduces the feature selection and engineering steps by providing a feature learning stage. The unsupervised learning process supports a large number of input signals due to being capable of uncovering relationship between them and discovering features. Therefore, the model supports the analysis of a set of machines as a group, lowering implementation costs by reducing the amount of instrumentation required. Furthermore, it is possible to visualize the feature significance resulting from the modeling process and its adaption during an incremental learning process, which can help in understanding the reasons for variance in the performance of the equipment depending on their operating state and control setpoints.

The implementation of the performance modeling methodology in the test environment successfully permitted the modeling of a group of chillers as standalone machines but also as a group, and achieved high accuracy for estimating the energy characteristics of these machines in both cases. Thus, it was verified that the modeling process is suitable for this use case, but that it is also compatible with different types of equipment, asserting that it can be an effective implementation for modeling HVAC equipment performance.

The third contribution belongs to the topic of optimal chiller loading and sequencing of chiller groups in buildings. The review of the state of the art of this topic described how the chiller loading and sequencing problems are being faced in the literature, and highlighted their drawbacks and avenues for improvement, which are threefold: i) the insufficient accuracy of equipment performance models employed, being in many cases simple lookup tables showing the expected COP at certain points of the partial load ratio, ii) the need for consideration of additional criteria during the optimization process besides simply the amount of consumed energy, and iii) the optimization process implementation, which often relies on generic global optimization tools like genetic algorithms or particle swarm optimization, which are useful but inconvenient for control applications. Accordingly, a control strategy was developed with the main focus of integrating the developed operating performance model of the equipment which allows to simulate equipment setpoints at given operating conditions and determine the expected consumption, production and performance, the integration of the thermal demand forecasting which allows to estimate the future energy needs of the building to anticipate changes, and the implementation of an efficient control strategy that is capable of also considering the switching criteria.

The implementation of the chiller control strategy with experimental data from the test environment successfully permitted to significantly reduce the energy consumption required to meet the energetic demands of the HVAC system, due to being able of taking advantage of the highest performing modes of operation of the equipment, and to being capable of anticipating changes in the thermal demand and planning accordingly. Furthermore, the generic formulation of the optimization problem allowed the consideration of the switching criteria in order to ensure stability and reduce losses and aging, but resulted in an extensible solution that could accommodate other criteria if required. Thus, the control strategy was shown to be an effective solution for increasing the performance of an HVAC installation while having a minimal to no effect on the comfort of its occupants due to its focus being on optimizing the operation of the production stage to meet the demands of the consumption stage.

In conclusion, a complete analysis and actuation framework was developed, implemented and validated by means of an experimental database acquired from the pilot plant during the research period of this thesis. The obtained results demonstrate the efficacy of the proposed standalone contributions, and as a whole represent a suitable solution for helping decrease the energy consumption footprint of buildings.

The methodology for load forecasting, designed specifically for usage in HVAC systems in buildings, together with the general operating performance modeling methodology for HVAC equipment, and finally the development of the extendable framework for implementing control strategies for chiller groups considering the maximization of their performance, extend the state of the art in energy management in buildings and represent an advancement in this topic.

6.2 Future work

The outcomes of the research work conducted in this thesis include the development and validation of a novel equipment modeling technique, load forecasting methodology and HVAC installation control strategy that surpass the state of the art regarding their characteristic features and performance. The presented methods offer increased modeling accuracy, generalization capabilities and applicability to different types of HVAC equipment and installation configurations, leading to a potential increase of the energetic efficiency in buildings with the application of the accomplished solution.

However, certain improvements could be considered to further improve the methodologies resulting from the contributions of this thesis:

- Load forecasting in buildings: the forecasting horizon was determined by analyzing the experimental data pertaining to the practical case study, but future work could involve the investigation of an automated method for establishing the maximum horizon given a predetermined accuracy requirement. Furthermore, future work could include the integration of a confidence estimation process to provide prediction confidence intervals.

- Operational performance modelling: this modelling methodology is supported by an unsupervised approach that is able to discover structure and relationship between the input signals of the model. This could permit the further analysis of the determined features and of what conditions lead to the activation of the sparse feature detectors, which could provide insight into what causes a certain behavior and to help quantify their effect, besides being used as a model to evaluate control setpoint scenarios. Additionally, the methodology includes a network dimensioning strategy to decrease the hyper-parameter tuning, but future work could focus on further alleviating this issue.

- Chiller group control strategy: the concluding strategy allows the consideration of multi-objective criteria, thus it could be extended to consider other factors if necessary. For example it could be adapted to enforce the uniform utilization of the HVAC equipment so that the aging of the equipment follows a similar rate, or it could be made to consider the variable cost of energy to achieve greater economic savings.

Finally, the equipment modeling and load forecasting methods are the advancements that have supported the implementation of the control strategy, but they can also provide value as standalone tools, or supporting other types of applications. For example, the operational performance modeling of HVAC equipment could be employed to implement supervision and maintenance solutions by establishing a baseline of a given installation and detecting deviations or tracking trends. Similarly, the load forecasting for buildings could support other applications such as demand response schemes in the context of the smart grid.

7.

Thesis results dissemination

Summary of the publications that disseminate the results obtained in this thesis, including publications originating from direct contributions of this thesis plus other publications resulting from collaborations in related research projects.

CONTENTS:

7.1 Publications: thesis contributions

7.2 Publications: collaborations and other works

7. Thesis results dissemination

This chapter summarizes the published results and dissemination activities carried out in the context of the thesis, and related collaborations.

7.1 Publications: thesis contributions

This section shows dissemination results and activities that are directly related to the contributions of this thesis.

7.1.1 Journal publications:

- **E. Sala-Cardoso**, M. Delgado-Prieto, K. Kampouropoulos, L. Romeral, “Activity-Aware HVAC Power Demand Forecasting,” Energy and Buildings. 170 (2018) 15–24. doi:10.1016/j.enbuild.2018.03.087.

- **E. Sala-Cardoso**, M. Delgado-Prieto, L. Romeral, “Operational Performance Modeling of HVAC Equipment using a Deep Learning Approach,” IEEE Transactions on Automation Science and Engineering. **Under review**.

- **E. Sala-Cardoso**, M. Delgado-Prieto, K. Kampouropoulos, L. Romeral, “Predictive Chiller Operation: A Data-Driven Loading and Scheduling Approach,” Energy and Buildings. **Under review.**

7.1.2 Conference publications:

- **E. Sala-Cardoso**, K. Kampouropoulos, M. Delgado and L. Romeral, "Intelligent monitoring of HVAC equipment by means of aggregated power analysis," IECON 2016 - 42th Annual Conference of the IEEE Industrial Electronics Society, Florence, 2016, pp. 1-6.

- **E. Sala-Cardoso**, D. Zurita, K. Kampouropoulos, M. Delgado and L. Romeral, "Occupancy forecasting for the reduction of HVAC energy consumption in smart buildings," IECON 2016 - 42th Annual Conference of the IEEE Industrial Electronics Society, Florence, 2016, pp. 1-6.

- **E. Sala-Cardoso**, K. Kampouropoulos, M. Delgado and L. Romeral, "Disaggregation of HVAC load profiles for the monitoring of individual equipment," Proceedings of the 2016 IEEE Emerging Technology and Factory Automation (ETFA), Berlin, 2016, pp. 1-6.

- **E. Sala-Cardoso**, D. Zurita, K. Kampouropoulos, M. Delgado-Prieto and L. Romeral, "Enhanced load forecasting methodology by means of probabilistic prediction intervals estimation," Industrial Technology (ICIT), 2015 IEEE International Conference on, Seville, 2015, pp. 1299-1304.

- **E. Sala-Cardoso**, K. Kampouropoulos, F. Giacometto and L. Romeral, "Smart multi-model approach based on adaptive Neuro-Fuzzy Inference Systems and Genetic Algorithms," IECON 2014 - 40th Annual Conference of the IEEE Industrial Electronics Society, Dallas, TX, 2014, pp. 288-294.

7.2 Publications: collaborations and other works

This section shows collaborative dissemination results and activities that belong to a topic related to the contributions of this thesis.

7.2.1 Journal publications:

- K. Kampouropoulos, F. Andrade, **E. Sala-Cardoso**, A. Garcia Espinosa and L. Romeral, "Multiobjective Optimization of Multi-Carrier Energy System using a Combination of ANFIS and Genetic Algorithms," in IEEE Transactions on Smart Grid, vol. PP, no. 99, pp. 1-8.
- F. Giacometto, F. Capelli, L. Romeral, J.-R. Riba, **E. Sala-Cardoso**, "Thermal Response Estimation in Substation Connectors Using Data-Driven Models," Advances in Electrical and Computer Engineering, vol. 16, n. 3, pp. 25-30, 2016.

7.2.2 Conference publications:

- K. Kampouropoulos, F. Andrade, **E. Sala-Cardoso**, A. Garcia and L. Romeral, "Multi-Carrier Optimal Power Flow of Energy Hubs by means of ANFIS and SQP," IECON 2016 - 42th Annual Conference of the IEEE Industrial Electronics Society, Florence, 2016, pp. 1-6.
- E. Cortez, M. Moreno-Eguilaz, F. Soriano and **E. Sala-Cardoso**, "Estimation of Fuel Consumption in a Hybrid Electric Refuse Collector Vehicle using a Real Drive Cycle," IECON 2016 - 42th Annual Conference of the IEEE Industrial Electronics Society, Florence, 2016, pp. 1-6.
- D. Zurita, **E. Sala-Cardoso**, J. A. Cariño, M. Delgado and J. A. Ortega, "Industrial Process Monitoring by means of Recurrent Neural Networks and Self Organizing Maps," Proceedings of the 2016 IEEE Emerging Technology and Factory Automation (ETFA), Berlin, 2016, pp. 1-6.
- F. Giacometto, F. Capelli, **E. Sala-Cardoso**, J. Riba and L. Romeral, "Temperature rise estimation of substation connectors using data-driven models: Case: Thermal conveccion response," Industrial Electronics Society, IECON 2015 - 41st Annual Conference of the IEEE, Yokohama, 2015, pp. 3957-3962.

- F. Giacometto, **E. Sala-Cardoso**, K. Kampouropoulos and L. Romeral, "Short-term load forecasting using Cartesian Genetic Programming: An efficient evolutive strategy: Case: Australian electricity market," Industrial Electronics Society, IECON 2015 - 41st Annual Conference of the IEEE, Yokohama, 2015, pp. 5087-5094.

- C. López, **E. Sala-Cardoso**, A. Espinosa and L. Romeral, "Constrained-size torque maximization in SynRM machines by means of genetic algorithms," Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED), 2015 IEEE 10th International Symposium on, Guarda, 2015, pp. 338-344.

- D. Zurita, **E. Sala-Cardoso**, J. A. Carino, M. Delgado and J. A. Ortega, "Enhanced time series forecasting by means of dynamics boosting for industrial process monitoring," Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED), 2015 IEEE 10th International Symposium on, Guarda, 2015, pp. 212-218.

- D. Zurita, J. A. Carino, **E. Sala-Cardoso**, M. Delgado-Prieto and J. A. Ortega, "Time series forecasting by means of SOM aided Fuzzy Inference Systems," Industrial Technology (ICIT), 2015 IEEE International Conference on, Seville, 2015, pp. 1772-1778.

- K. Kampouropoulos, F. Andrade, **E. Sala-Cardoso** and L. Romeral, "Optimal control of energy hub systems by use of SQP algorithm and energy prediction," IECON 2014 - 40th Annual Conference of the IEEE Industrial Electronics Society, Dallas, TX, 2014, pp. 221-227.

References

This chapter documents the resources referenced within the scope of this work.

References

[1] N. Busch, D. Pooley, G. Faninger, Further tasks for future European energy R&D - Research and Innovation (European Commission), 2006.

[2] A. Martínez-Molina, I. Tort-Ausina, S. Cho, J.L. Vivancos, Energy efficiency and thermal comfort in historic buildings: A review, Renew. Sustain. Energy Rev. 61 (2016) 70–85. doi:10.1016/j.rser.2016.03.018.

[3] F. Pacheco-Torgal, Eco-efficient construction and building materials research under the EU Framework Programme Horizon 2020, Constr. Build. Mater. 51 (2014) 151–162. doi:10.1016/j.conbuildmat.2013.10.058.

[4] World Business Council for Sustainable Development, Energy efficiency in buildings: Transforming the market, (2009).

[5] Buildings Performance Institute Europe, Europe's Buildings under the Microscope. A country-by-country review of the energy performance of buildnings., 2011.

[6] A. Costa, M.M. Keane, J.I. Torrens, E. Corry, Building operation and energy performance: Monitoring, analysis and optimisation toolkit, Appl. Energy. 101 (2013) 310–316. doi:10.1016/j.apenergy.2011.10.037.

[7] T. Weng, Y. Agarwal, From buildings to smart buildings-sensing and actuation to improve energy efficiency, IEEE Des. Test Comput. 29 (2012) 36–44. doi:10.1109/MDT.2012.2211855.

[8] L. Pérez-Lombard, J. Ortiz, C. Pout, A review on buildings energy consumption information, Energy Build. 40 (2008) 394–398. doi:10.1016/j.enbuild.2007.03.007.

[9] Directorate-General for Energy and Transport (European Commission), EU Energy and Transport in figures 2010, Statistical pocket book, 2010.

[10] Y. Agarwal, B. Balaji, S. Dutta, R.K. Gupta, T. Weng, Duty-Cycling Buildings Aggressively: The Next Frontier in HVAC Control Categories and Subject Descriptors, in: Inf. Process. Sens. Networks (IPSN), 2011 10th Int. Conf., 2011: pp. 246–257.

[11] T. Kalamees, K. Jylhä, H. Tietäväinen, J. Jokisalo, S. Ilomets, R. Hyvönen, S. Saku, Development of weighting factors for climate variables for selecting the energy reference year according to the en ISO 15927-4 standard, Energy Build. 47 (2012) 53–60. doi:10.1016/j.enbuild.2011.11.031.

[12] B. Sun, P.B. Luh, Q.S. Jia, Z. Jiang, F. Wang, C. Song, Building energy management: Integrated control of active and passive heating, cooling, lighting, shading, and ventilation systems, IEEE Trans. Autom. Sci. Eng. 10 (2013) 588–602. doi:10.1109/TASE.2012.2205567.

[13] D. Wijayasekara, O. Linda, M. Manic, C. Rieger, Mining building energy management system data using fuzzy anomaly detection and linguistic descriptions, IEEE Trans. Ind. Informatics. 10 (2014) 1829–1840. doi:10.1109/TII.2014.2328291.

[14] P. Palensky, D. Dietrich, Demand side management: Demand response, intelligent energy systems, and smart loads, IEEE Trans. Ind. Informatics. 7 (2011) 381–388. doi:10.1109/TII.2011.2158841.

[15] G.T. Costanzo, G. Zhu, M.F. Anjos, G. Savard, A system architecture for autonomous demand side load management in smart buildings, IEEE Trans. Smart Grid. 3 (2012) 2157–2165. doi:10.1109/TSG.2012.2217358.

[16] D. Dietrich, D. Bruckner, G. Zucker, P. Palensky, Communication and Computation in Buildings: A Short Introduction and Overview, IEEE Trans. Ind. Electron. 57 (2010) 3577–3584. doi:10.1109/TIE.2010.2046570.

[17] X. Li, C.P. Bowers, T. Schnier, Classification of energy consumption in buildings with outlier detection, IEEE Trans. Ind. Electron. 57 (2010) 3639–3644. doi:10.1109/TIE.2009.2027926.

[18] A. Ahmed, J. Ploennigs, K. Menzel, B. Cahill, Multi-dimensional building performance data management for continuous commissioning, Adv. Eng. Informatics. 24 (2010) 466–475. doi:10.1016/j.aei.2010.06.007.

[19] K. Whitehouse, J. Ranjan, J. Lu, T. Sookoor, M. Saadat, C.M. Burke, G. Staengl, A. Canfora, H. Haj-Hariri, Towards occupancy-driven heating and cooling, IEEE Des. Test Comput. 29 (2012) 17–25. doi:10.1109/MDT.2012.2201439.

[20] X. Cao, J. Chen, Y. Xiao, Y. Sun, Building-environment control with wireless sensor and actuator networks: Centralized versus distributed, IEEE Trans. Ind. Electron. 57 (2010) 3596–3605. doi:10.1109/TIE.2009.2029585.

[21] M. Ruta, F. Scioscia, E. Di Sciascio, G. Loseto, Semantic-based enhancement of ISO/IEC 14543-3 EIB/KNX standard for building automation, IEEE Trans. Ind. Informatics. 7 (2011) 731–739. doi:10.1109/TII.2011.2166792.

[22] S. Runde, A. Fay, Software support for building automation requirements engineering-An application of semantic web technologies in automation, IEEE Trans. Ind. Informatics. 7 (2011) 723–730. doi:10.1109/TII.2011.2166784.

[23] H. Dibowski, J. Ploennigs, K. Kabitzsch, Automated design of building automation systems, IEEE Trans. Ind. Electron. 57 (2010) 3606–3613. doi:10.1109/TIE.2009.2032209.

[24] T. Novak, A. Gerstinger, Safety- and security-critical services in building automation and control systems, IEEE Trans. Ind. Electron. 57 (2010) 3614–3621. doi:10.1109/TIE.2009.2028364.

[25] W. Granzer, F. Praus, W. Kastner, Security in building automation systems, IEEE Trans. Ind. Electron. 57 (2010) 3622–3630. doi:10.1109/TIE.2009.2036033.

[26] M. Ruta, F. Scioscia, G. Loseto, E. Di Sciascio, Semantic-based resource discovery and orchestration in home and building automation: A multi-agent approach, IEEE Trans. Ind. Informatics. 10 (2014) 730–741. doi:10.1109/TII.2013.2273433.

[27] S. Mohagheghi, N. Raji, Managing industrial energy intelligently: Demand response scheme, IEEE Ind. Appl. Mag. 20 (2014) 53–62. doi:10.1109/MIAS.2013.2288387.

[28] D. Lee, C.C. Cheng, Energy savings by energy management systems: A review, Renew. Sustain. Energy Rev. 56 (2016) 760–777. doi:10.1016/j.rser.2015.11.067.

[29] H.E. Brown, S. Suryanarayanan, G.T. Heydt, Some characteristics of emerging distribution systems considering the smart grid initiative, Electr. J. 23 (2010) 64–75. doi:10.1016/j.tej.2010.05.005.

[30] S. Lee, B. Kwon, S. Lee, Joint energy management system of electric supply and demand in houses and buildings, IEEE Trans. Power Syst. 29 (2014) 2804–2812. doi:10.1109/TPWRS.2014.2311827.

[31] J.E. Seem, Using intelligent data analysis to detect abnormal energy consumption in buildings, Energy Build. 39 (2007) 52–58. doi:10.1016/j.enbuild.2006.03.033.

[32] G. Zucker, A. Sporr, S. Kollmann, A. Wendt, L. Siafara Chaido, A. Fernbach, A Cognitive System Architecture for Building Energy Management, IEEE Trans. Ind. Informatics. 14 (2018) 2521–2529. doi:10.1109/TII.2018.2815739.

[33] R. Velik, G. Zucker, Autonomous perception and decision making in building automation, IEEE Trans. Ind. Electron. 57 (2010) 3645–3652. doi:10.1109/TIE.2009.2038985.

[34] M.A.A. Pedrasa, T.D. Spooner, I.F. MacGill, Coordinated scheduling of residential distributed energy resources to optimize smart home energy services, IEEE Trans. Smart Grid. 1 (2010) 134–143. doi:10.1109/TSG.2010.2053053.

[35] Tsair-Fwu Lee, Ming-Yuan Cho, Ying-Chang Hsiao, Pei-Ju Chao, Fu-Min Fang, Optimization and Implementation of a Load Control Scheduler Using Relaxed Dynamic Programming for Large Air Conditioner Loads, IEEE Trans. Power Syst. 23 (2008) 691–702. doi:10.1109/tpwrs.2008.919311.

[36] M.A.A. Pedrasa, T.D. Spooner, I.F. MacGill, Scheduling of demand side resources using binary particle swarm optimization, IEEE Trans. Power Syst. 24 (2009) 1173–1181. doi:10.1109/TPWRS.2009.2021219.

[37] C.M. Chu, T.L. Jong, A novel direct air-conditioning load control method, IEEE Trans. Power Syst. 23 (2008) 1356–1363. doi:10.1109/TPWRS.2008.926432.

[38] V. Vakiloroaya, B. Samali, A. Fakhar, K. Pishghadam, A review of different strategies for HVAC energy saving, Energy Convers. Manag. 77 (2014) 738–754. doi:10.1016/j.enconman.2013.10.023.

[39] C. Deb, S.E. Lee, M. Santamouris, Using artificial neural networks to assess HVAC related energy saving in retrofitted office buildings, Sol. Energy. 163 (2018) 32–44. doi:10.1016/j.solener.2018.01.075.

[40] A.S. Javid, F. Aramoun, M. Bararzadeh, A. Avami, Multi objective planning for sustainable retrofit of educational buildings, J. Build. Eng. 24 (2019) 100759. doi:10.1016/j.jobe.2019.100759.

[41] X. Zhao, Y. Tan, L. Shen, G. Zhang, J. Wang, Case-based reasoning approach for supporting building green retrofit decisions, Build. Environ. 160 (2019) 106210. doi:10.1016/j.buildenv.2019.106210.

[42] C. Fan, F. Xiao, Z. Li, J. Wang, Unsupervised data analytics in mining big building operational data for energy efficiency enhancement: A review, Energy Build. 159 (2018) 296–308. doi:10.1016/j.enbuild.2017.11.008.

[43] Z. Yu, F. Haghighat, B.C.M. Fung, Advances and challenges in building engineering and data mining applications for energy-efficient communities, Sustain. Cities Soc. 25 (2016) 33–38. doi:10.1016/j.scs.2015.12.001.

[44] X. Yu, S. Ergan, G. Dedemen, A data-driven approach to extract operational signatures of HVAC systems and analyze impact on electricity consumption, Appl. Energy. 253 (2019) 113497. doi:10.1016/j.apenergy.2019.113497.

[45] A. Capozzoli, M.S. Piscitelli, A. Gorrino, I. Ballarini, V. Corrado, Data analytics for occupancy pattern learning to reduce the energy consumption of HVAC systems in office buildings, Sustain. Cities Soc. 35 (2017) 191–208. doi:10.1016/j.scs.2017.07.016.

[46] Z. Afroz, G.M. Shafiullah, T. Urmee, G. Higgins, Modeling techniques used in building HVAC control systems: A review, Renew. Sustain. Energy Rev. 83 (2018) 64–84. doi:10.1016/j.rser.2017.10.044.

[47] M. Royapoor, A. Antony, T. Roskilly, A review of building climate and plant controls, and a survey of industry perspectives, Energy Build. 158 (2018) 453–465. doi:10.1016/j.enbuild.2017.10.022.

[48] A. Afram, F. Janabi-Sharifi, Theory and applications of HVAC control systems - A review of model predictive control (MPC), Build. Environ. 72 (2014) 343–355. doi:10.1016/j.buildenv.2013.11.016.

[49] T. Hilliard, L. Swan, Z. Qin, Experimental implementation of whole building MPC with zone based thermal comfort adjustments, Build. Environ. 125 (2017) 326–338. doi:10.1016/j.buildenv.2017.09.003.

[50] N. Radhakrishnan, S. Srinivasan, R. Su, K. Poolla, Learning-Based Hierarchical Distributed HVAC Scheduling With Operational Constraints, IEEE Trans. Control Syst. Technol. PP (2017) 1–9. doi:10.1109/TCST.2017.2728004.

[51] D. Zhang, P. Luh, J. Fan, S. Gupta, Chiller Plant Operation Optimization: Energy-Efficient Primary-Only and Primary-Secondary Systems, IEEE Trans. Autom. Sci. Eng. 15 (2018) 341–355. doi:10.1109/TASE.2017.2751605.

[52] Z. Shi, W. O'Brien, Development and implementation of automated fault detection and diagnostics for building systems: A review, Autom. Constr. 104 (2019) 215–229. doi:10.1016/j.autcon.2019.04.002.

[53] Y. Zhao, T. Li, X. Zhang, C. Zhang, Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future, Renew. Sustain. Energy Rev. 109 (2019) 85–101. doi:10.1016/j.rser.2019.04.021.

[54] K. Yan, L. Ma, Y. Dai, W. Shen, Z. Ji, D. Xie, Cost-sensitive and sequential feature selection for chiller fault detection and diagnosis, Int. J. Refrig. 86 (2018) 401–409. doi:10.1016/j.ijrefrig.2017.11.003.

[55] P.M. Van Every, M. Rodriguez, C.B. Jones, A.A. Mammoli, M. Martínez-Ramón, Advanced detection of HVAC faults using unsupervised SVM novelty detection and Gaussian process models, Energy Build. 149 (2017) 216–224. doi:10.1016/j.enbuild.2017.05.053.

[56] I. Lampropoulos, W.L. Kling, P.F. Ribeiro, J. Van Den Berg, History of demand side management and classification of demand response control schemes, in: IEEE Power Energy Soc. Gen. Meet., 2013: pp. 1–5. doi:10.1109/PESMG.2013.6672715.

[57] K.O. Aduda, T. Labeodan, W. Zeiler, G. Boxem, Y. Zhao, Demand side flexibility: Potentials and building performance implications, Sustain. Cities Soc. 22 (2016) 146–163. doi:10.1016/j.scs.2016.02.011.

[58] G. Graditi, M.G. Ippolito, R. Lamedica, A. Piccolo, A. Ruvio, E. Santini, P. Siano, G. Zizzo, Innovative control logics for a rational utilization of electric loads and air-conditioning systems in a residential building, Energy Build. 102 (2015) 1–17. doi:10.1016/j.enbuild.2015.05.027.

[59] I. Sartori, A. Napolitano, K. Voss, Net zero energy buildings: A consistent definition framework, Energy Build. 48 (2012) 220–232. doi:10.1016/j.enbuild.2012.01.032.

[60] A.H. Neto, F.A.S. Fiorelli, Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption, Energy Build. 40 (2008) 2169–2176. doi:10.1016/j.enbuild.2008.06.013.

[61] J. Grant, M. Eltoukhy, S. Asfour, Short-term electrical peak demand forecasting in a large government building using artificial neural networks, Energies. 7 (2014) 1935–1953. doi:10.3390/en7041935.

[62] G. Qiang, T. Zhe, D. Yan, Z. Neng, An improved office building cooling load prediction model based on multivariable linear regression, Energy Build. 107 (2015) 445–455. doi:10.1016/j.enbuild.2015.08.041.

[63] Q. Zhou, S. Wang, X. Xu, F. Xiao, A grey-box model of next-day building thermal load prediction for energy-efficient control, Int. J. Energy Res. 32 (2008) 1418–1431. doi:10.1002/er.1458.

[64] Z. Liu, H. Tan, D. Luo, G. Yu, J. Li, Z. Li, Optimal chiller sequencing control in an office building considering the variation of chiller maximum cooling capacity, Energy Build. 140 (2017) 430–442. doi:10.1016/j.enbuild.2017.01.082.

[65] Y.C. Chang, Application of Hopfield Neural Network to the optimal chilled water supply temperature calculation of air-conditioning systems for saving energy, Int. J. Therm. Sci. 48 (2009) 1649–1657. doi:10.1016/j.ijthermalsci.2009.01.002.

[66] B. Yildiz, J.I. Bilbao, A.B. Sproul, A review and analysis of regression and machine learning models on commercial building electricity load forecasting, Renew. Sustain. Energy Rev. 73 (2017) 1104–1122. doi:10.1016/j.rser.2017.02.023.

[67] T. Hong, S. Fan, Probabilistic electric load forecasting: A tutorial review, Int. J. Forecast. 32 (2016) 914–938. doi:10.1016/j.ijforecast.2015.11.011.

[68] C. Kuster, Y. Rezgui, M. Mourshed, Electrical load forecasting models: A critical systematic review, Sustain. Cities Soc. 35 (2017) 257–270. doi:10.1016/j.scs.2017.08.009.

[69] S. Ferlito, G. Adinolfi, G. Graditi, Comparative analysis of data-driven methods online and offline trained to the forecasting of grid-connected photovoltaic plant production, Appl. Energy. 205 (2017) 116–129. doi:10.1016/j.apenergy.2017.07.124.

[70] M. Peña, F. Biscarri, J.I. Guerrero, I. Monedero, C. León, Rule-based system to detect energy efficiency anomalies in smart buildings, a data mining approach, Expert Syst. Appl. 56 (2016) 242–255. doi:10.1016/j.eswa.2016.03.002.

[71] T. Hong, S.C. Taylor-Lange, S. D'Oca, D. Yan, S.P. Corgnati, Advances in research and applications of energy-related occupant behavior in buildings, Energy Build. 116 (2016) 694–702. doi:10.1016/j.enbuild.2015.11.052.

[72] B. Chenari, J. Dias Carrilho, M. Gameiro da Silva, Towards sustainable, energy-efficient and healthy ventilation strategies in buildings: A review, Renew. Sustain. Energy Rev. 59 (2016) 1426–1447. doi:https://doi.org/10.1016/j.rser.2016.01.074.

[73] M. Jia, R.S. Srinivasan, A.A. Raheem, From occupancy to occupant behavior: An analytical survey of data acquisition technologies, modeling methodologies and simulation coupling mechanisms for building energy efficiency, Renew. Sustain. Energy Rev. 68 (2017) 525–540. doi:10.1016/j.rser.2016.10.011.

[74] Z. Wang, R.S. Srinivasan, A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models, Renew. Sustain. Energy Rev. 75 (2017) 796–808. doi:10.1016/j.rser.2016.10.079.

[75] J. Massana, C. Pous, L. Burgas, J. Melendez, J. Colomer, Short-term load forecasting in a non-residential building contrasting models and attributes, Energy Build. 92 (2015) 322–330. doi:10.1016/j.enbuild.2015.02.007.

[76] J. Massana, C. Pous, L. Burgas, J. Melendez, J. Colomer, Short-term load forecasting for non-residential buildings contrasting artificial occupancy attributes, Energy Build. 130 (2016) 519–531. doi:10.1016/j.enbuild.2016.08.081.

[77] Y.S. Lee, A.M. Malkawi, Simulating multiple occupant behaviors in buildings: An agent-based modeling approach, Energy Build. 69 (2014) 407–416. doi:10.1016/j.enbuild.2013.11.020.

[78] J. Virote, R. Neves-Silva, Stochastic models for building energy prediction based on occupant behavior assessment, Energy Build. 53 (2012) 183–193. doi:10.1016/j.enbuild.2012.06.001.

[79] Z. Chen, J. Xu, Y.C. Soh, Modeling regular occupancy in commercial buildings using stochastic models, Energy Build. 103 (2015) 216–223. doi:10.1016/j.enbuild.2015.06.009.

[80] Z. Yang, B. Becerik-Gerber, The coupled effects of personalized occupancy profile based HVAC schedules and room reassignment on building energy use, Energy Build. 78 (2014) 113–122. doi:10.1016/j.enbuild.2014.04.002.

[81] Y.T. Chae, R. Horesh, Y. Hwang, Y.M. Lee, Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings, Energy Build. 111 (2016) 184–194. doi:10.1016/j.enbuild.2015.11.045.

[82] J. Zhao, B. Lasternas, K.P. Lam, R. Yun, V. Loftness, Occupant behavior and schedule modeling for building energy simulation through office appliance power consumption data mining, Energy Build. 82 (2014) 341–355. doi:10.1016/j.enbuild.2014.07.033.

[83] Ö.F. Ertugrul, Forecasting electricity load by a novel recurrent extreme learning machines approach, Int. J. Electr. Power Energy Syst. 78 (2016) 429–435. doi:10.1016/j.ijepes.2015.12.006.

[84] A. Laouafi, M. Mordjaoui, S. Haddad, T.E. Boukelia, A. Ganouche, Online electricity demand forecasting based on an effective forecast combination methodology, Electr. Power Syst. Res. 148 (2017) 35–47. doi:10.1016/j.epsr.2017.03.016.

[85] S. Barak, S.S. Sadegh, Forecasting energy consumption using ensemble ARIMA-ANFIS hybrid algorithm, Int. J. Electr. Power Energy Syst. 82 (2016) 92–104. doi:10.1016/j.ijepes.2016.03.012.

[86] C. Deb, F. Zhang, J. Yang, S.E. Lee, K.W. Shah, A review on time series forecasting techniques for building energy consumption, Renew. Sustain. Energy Rev. 74 (2017) 902–924. doi:10.1016/j.rser.2017.02.085.

[87] K. Deng, Y. Sun, S. Li, Y. Lu, J. Brouwer, P.G. Mehta, M.C. Zhou, A. Chakraborty, Model Predictive Control of Central Chiller Plant With Thermal Energy Storage Via Dynamic Programming and Mixed-Integer Linear Programming, IEEE Trans. Autom. Sci. Eng. 12 (2015) 565–579. doi:10.1109/TASE.2014.2352280.

[88] N. Li, G. Calis, B. Becerik-Gerber, Measuring and monitoring occupancy with an RFID based system for demand-driven HVAC operations, Autom. Constr. 24 (2012) 89–99. doi:10.1016/j.autcon.2012.02.013.

[89] C. Fan, Y. Sun, K. Shan, F. Xiao, J. Wang, Discovering gradual patterns in building operations for improving building energy efficiency, Appl. Energy. 224 (2018) 116–123. doi:https://doi.org/10.1016/j.apenergy.2018.04.118.

[90] N. Zibin, R. Zmeureanu, J. Love, Automatic assisted calibration tool for coupling building automation system trend data with commissioning, Autom. Constr. 61 (2016) 124–133. doi:10.1016/j.autcon.2015.10.004.

[91] P. Wang, R.X. Gao, Automated Performance Tracking for Heat Exchangers in HVAC, IEEE Trans. Autom. Sci. Eng. 14 (2017) 634–645. doi:10.1109/TASE.2017.2666184.

[92] T. Ma, J. Wu, L. Hao, W.-J. Lee, H. Yan, D. Li, The Optimal Structure Planning and Energy Management Strategies of Smart Multi Energy Systems, Energy. 160 (2018) 122–141. doi:10.1016/J.ENERGY.2018.06.198.

[93] Y. Sun, S. Wang, G. Huang, Online sensor fault diagnosis for robust chiller sequencing control, Int. J. Therm. Sci. 49 (2010) 589–602. doi:10.1016/j.ijthermalsci.2009.10.003.

[94] X. Liu, B. Akinci, M. Bergés, J.H. Garrett, Extending the information delivery manual approach to identify information requirements for performance analysis of HVAC systems, Adv. Eng. Informatics. 27 (2013) 496–505. doi:10.1016/j.aei.2013.05.003.

[95] W. Wang, M.R. Brambley, W. Kim, S. Somasundaram, A.J. Stevens, Automated point mapping for building control systems: Recent advances and future research needs, Autom. Constr. 85 (2018) 107–123. doi:10.1016/j.autcon.2017.09.013.

[96] H.Z. Abou-Ziyan, A.F. Alajmi, Effect of load-sharing operation strategy on the aggregate performance of existed multiple-chiller systems, Appl. Energy. 135 (2014) 329–338. doi:10.1016/j.apenergy.2014.06.065.

[97] M. Li, Y. Ju, The analysis of the operating performance of a chiller system based on hierarchal cluster method, Energy Build. 138 (2017) 695–703. doi:10.1016/j.enbuild.2016.12.076.

[98] C. Dullinger, W. Struckl, M. Kozek, A general approach for mixed-integer predictive control of HVAC systems using MILP, Appl. Therm. Eng. 128 (2018) 1646–1659. doi:10.1016/j.applthermaleng.2017.09.104.

[99] K. Shan, S. Wang, D.C. Gao, F. Xiao, Development and validation of an effective and robust chiller sequence control strategy using data-driven models, Autom. Constr. 65 (2016) 78–85. doi:10.1016/j.autcon.2016.01.005.

[100] S.R. Thangavelu, A. Myat, A. Khambadkone, Energy optimization methodology of multi-chiller plant in commercial buildings, Energy. 123 (2017) 64–76. doi:10.1016/j.energy.2017.01.116.

[101] E. Salari, A. Askarzadeh, A new solution for loading optimization of multi-chiller systems by general algebraic modeling system, Appl. Therm. Eng. 84 (2015) 429–436. doi:10.1016/j.applthermaleng.2015.03.057.

[102] H. Kazmi, F. Mehmood, S. Lodeweyckx, J. Driesen, Gigawatt-hour scale savings on a budget of zero: Deep reinforcement learning based optimal control of hot water systems, Energy. 144 (2018) 159–168. doi:https://doi.org/10.1016/j.energy.2017.12.019.

[103] Q.P. Ha, V. Vakiloroaya, Modeling and optimal control of an energy-efficient hybrid solar air conditioning system, Autom. Constr. 49 (2015) 262–270. doi:10.1016/j.autcon.2014.06.004.

[104] H. Karlsson, C.E. Hagentoft, Application of model based predictive control for water-based floor heating in low energy residential buildings, Build. Environ. 46 (2011) 556–569.

doi:10.1016/j.buildenv.2010.08.014.

[105] A. Afram, F. Janabi-Sharifi, Gray-box modeling and validation of residential HVAC system for control system design, Appl. Energy. 137 (2015) 134–150. doi:10.1016/j.apenergy.2014.10.026.

[106] S. Singaravel, J. Suykens, P. Geyer, Deep-learning neural-network architectures and methods: Using component-based models in building-design energy prediction, Adv. Eng. Informatics. 38 (2018) 81–90. doi:10.1016/j.aei.2018.06.004.

[107] T. Dwyer, Knowledge is Power: Benchmarking and prediction of building energy consumption, Build. Serv. Eng. Res. Technol. 34 (2013) 5–7. doi:10.1177/0143624412471130.

[108] B. Mu, Y. Li, J.M. House, T.I. Salsbury, Real-time optimization of a chilled water plant with parallel chillers based on extremum seeking control, Appl. Energy. 208 (2017) 766–781. doi:10.1016/j.apenergy.2017.09.072.

[109] A. Ghahramani, S.A. Karvigh, B. Becerik-Gerber, HVAC system energy optimization using an adaptive hybrid metaheuristic, Energy Build. 152 (2017) 149–161. doi:10.1016/j.enbuild.2017.07.053.

[110] J. Labus, J.C. Bruno, A. Coronas, Performance analysis of small capacity absorption chillers by using different modeling methods, Appl. Therm. Eng. 58 (2013) 305–313. doi:10.1016/j.applthermaleng.2013.04.032.

[111] A. Afram, F. Janabi-Sharifi, Black-box modeling of residential HVAC system and comparison of gray-box and black-box modeling methods, Energy Build. 94 (2015) 121–149. doi:10.1016/j.enbuild.2015.02.045.

[112] N. Cotrufo, R. Zmeureanu, Virtual outdoor air flow meter for an existing HVAC system in heating mode, Autom. Constr. 92 (2018) 166–172. doi:10.1016/j.autcon.2018.03.036.

[113] X. Li, J. Wen, E.-W. Bai, Developing a whole building cooling energy forecasting model for on-line operation optimization using proactive system identification, Appl. Energy. 164 (2016) 69–88. doi:https://doi.org/10.1016/j.apenergy.2015.12.002.

[114] F.W. Yu, W.T. Ho, K.T. Chan, R.K.Y. Sit, Critique of operating variables importance on chiller energy performance using random forest, Energy Build. 139 (2017) 653–664. doi:10.1016/j.enbuild.2017.01.063.

[115] T.S. Lee, K.Y. Liao, W.C. Lu, Evaluation of the suitability of empirically-based models for predicting energy performance of centrifugal water chillers with variable chilled water flow, Appl. Energy. 93 (2012) 583–595. doi:10.1016/j.apenergy.2011.12.001.

[116] X. Wei, G. Xu, A. Kusiak, Modeling and optimization of a chiller plant, Energy. 73 (2014) 898–907. doi:https://doi.org/10.1016/j.energy.2014.06.102.

[117] A. Afram, F. Janabi-Sharifi, Review of modeling methods for HVAC systems, Appl. Therm. Eng. 67 (2014) 507–519. doi:10.1016/j.applthermaleng.2014.03.055.

[118] M. Mohanraj, S. Jayaraj, C. Muraleedharan, Applications of artificial neural networks for refrigeration, air-conditioning and heat pump systems - A review, Renew. Sustain. Energy Rev. 16 (2012) 1340–1358. doi:10.1016/j.rser.2011.10.015.

[119] P. Frey, S. Fischer, H. Drück, Artificial Neural Network modelling of sorption chillers, Sol. Energy. 108 (2014) 525–537. doi:10.1016/j.solener.2014.08.006.

[120] A. Kusiak, G. Xu, Z. Zhang, Minimization of energy consumption in HVAC systems with data-driven models and an interior-point method, Energy Convers. Manag. 85 (2014) 146–153. doi:10.1016/j.enconman.2014.05.053.

[121] N.M. Nawi, A. Khan, M.Z. Rehman, Levenberg Marquardt based Back Propagation Algorithm Optimized with Cuckoo Search, J. ICT Res. Appl. 7 (2013) 103–116. doi:10.5614/itbj.ict.res.appl.2013.7.2.1.

[122] C.L. Chen, Y.C. Chang, T.S. Chan, Applying smart models for energy saving in optimal chiller loading, Energy Build. 68 (2014) 364–371. doi:10.1016/j.enbuild.2013.04.030.

[123] F. Jia, Y. Lei, J. Lin, X. Zhou, N. Lu, Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data, Mech. Syst. Signal Process. 72–73 (2016) 303–315. doi:10.1016/j.ymssp.2015.10.025.

[124] M. Liu, J. Zhang, X. Guo, L. Cao, Hypergraph regularized sparse feature learning, Neurocomputing. 237 (2017) 185–192. doi:10.1016/j.neucom.2016.10.031.

[125] G. Zhong, L.-N. Wang, X. Ling, J. Dong, An overview on data representation learning: From traditional feature learning to recent deep learning, J. Financ. Data Sci. 2 (2016) 265–278. doi:10.1016/j.jfds.2017.05.001.

[126] G. Suryanarayana, J. Lago, D. Geysen, P. Aleksiejuk, C. Johansson, Thermal load forecasting in district heating networks using deep learning and advanced feature selection methods, Energy. 157 (2018) 141–149. doi:10.1016/j.energy.2018.05.111.

[127] A. Kusiak, F. Tang, G. Xu, Multi-objective optimization of HVAC system with an evolutionary computation algorithm, Energy. 36 (2011) 2440–2449. doi:10.1016/j.energy.2011.01.030.

[128] S. Silva, L. Vanneschi, A.I.R. Cabral, M.J. Vasconcelos, A semi-supervised Genetic Programming method for dealing with noisy labels and hidden overfitting, Swarm Evol. Comput. 39 (2017) 323–338. doi:10.1016/j.swevo.2017.11.003.

[129] H. Shao, H. Jiang, H. Zhang, W. Duan, T. Liang, S. Wu, Rolling bearing fault feature learning using improved convolutional deep belief network with compressed sensing, Mech. Syst. Signal Process. 100 (2018) 743–765. doi:10.1016/j.ymssp.2017.08.002.

[130] C. Fan, F. Xiao, Y. Zhao, A short-term building cooling load prediction method using deep learning algorithms, Appl. Energy. 195 (2017) 222–233. doi:10.1016/j.apenergy.2017.03.064.

[131] L. Yin, T. Yu, X. Zhang, B. Yang, Relaxed deep learning for real-time economic generation dispatch and control with unified time scale, Energy. 149 (2018) 11–23. doi:10.1016/j.energy.2018.01.165.

[132] C. Methods, A. Mech, A. Oishi, G. Yagawa, Computational mechanics enhanced by deep learning, Comput. Methods Appl. Mech. Eng. 327 (2017) 327–351. doi:https://doi.org/10.1016/j.cma.2017.08.040.

[133] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: A Simple Way to Prevent Neural Networks from Overfitting, J. Mach. Learn. Res. 15 (2014) 1929–1958. doi:10.1214/12-AOS1000.

[134] G.E. Hinton, Reducing the Dimensionality of Data with Neural Networks, Science (80-.). 313 (2006) 504–507. doi:10.1126/science.1127647.

[135] Xue-Wen Chen, Xiaotong Lin, Big Data Deep Learning: Challenges and Perspectives, IEEE Access. 2 (2014) 514–525. doi:10.1109/ACCESS.2014.2325029.

[136] D.B. Araya, K. Grolinger, H.F. Elyamany, M.A.M. Capretz, An ensemble learning framework for anomaly detection in building energy consumption, Energy Build. 144 (2017) 191–206. doi:10.1016/j.enbuild.2017.02.058.

[137] R. Liu, B. Yang, E. Zio, X. Chen, Artificial intelligence for fault diagnosis of rotating machinery: A review, Mech. Syst. Signal Process. 108 (2018) 33–47. doi:10.1016/j.ymssp.2018.02.016.

[138] Q. Zhang, L.T. Yang, Z. Chen, P. Li, A survey on deep learning for big data, Inf. Fusion. 42 (2018) 146–157. doi:10.1016/j.inffus.2017.10.006.

[139] R.M. Vignali, F. Borghesan, L. Piroddi, M. Strelec, M. Prandini, Energy Management of a Building Cooling System with Thermal Storage: An Approximate Dynamic Programming Solution, IEEE Trans. Autom. Sci. Eng. 14 (2017) 619–633. doi:10.1109/TASE.2016.2635109.

[140] H. Dong, T. Li, R. Ding, J. Sun, A novel hybrid genetic algorithm with granular information for feature selection and optimization, Appl. Soft Comput. J. 65 (2018) 33–46. doi:10.1016/j.asoc.2017.12.048.

[141] S. Huang, W. Zuo, M.D. Sohn, Amelioration of the cooling load based chiller sequencing control, Appl. Energy. 168 (2016) 204–215. doi:10.1016/j.apenergy.2016.01.035.

[142] F.W. Yu, K.T. Chan, Energy signatures for assessing the energy performance of chillers, Energy Build. 37 (2005) 739–746. doi:https://doi.org/10.1016/j.enbuild.2004.10.004.

[143] Y. Wang, X. Jin, X. Fang, Rapid evaluation of operation performance of multi-chiller system based on history data analysis, Energy Build. 134 (2017) 162–170. doi:10.1016/j.enbuild.2016.10.041.

[144] L. Dos Santos Coelho, V.C. Mariani, Improved firefly algorithm approach applied to chiller loading for energy conservation, Energy Build. 59 (2013) 273–278. doi:10.1016/j.enbuild.2012.11.030.

[145] B.M. Seo, K.H. Lee, Detailed analysis on part load ratio characteristics and cooling energy saving of chiller staging in an office building, Energy Build. 119 (2016) 309–322. doi:https://doi.org/10.1016/j.enbuild.2016.03.067.

[146] X. Fang, X. Jin, Z. Du, Y. Wang, The evaluation of operation performance of HVAC system based on the ideal operation level of system, Energy Build. 110 (2016) 330–344. doi:10.1016/j.enbuild.2015.11.020.

[147] F. Sohrabi, M. Nazari-Heris, B. Mohammadi-Ivatloo, S. Asadi, Optimal chiller loading for saving energy by exchange market algorithm, Energy Build. 169 (2018) 245–253. doi:10.1016/j.enbuild.2018.03.077.

[148] Y. Liao, G. Huang, Y. Ding, H. Wu, C. Fan, Robustness Analysis and Enhancement of Chiller Sequencing Control under Uncertainties, Procedia Eng. 205 (2017) 1878–1885. doi:10.1016/j.proeng.2017.10.270.

[149] Y. Liao, G. Huang, A Hybrid Predictive Sequencing Control for Multi-chiller Plant with Considerations of Indoor Environment Control, Energy Conservation and Economical Operation Cost, Sustain. Cities Soc. 49 (2019) 101616. doi:10.1016/j.scs.2019.101616.

[150] L. Wang, E.W.M. Lee, R.K.K. Yuen, A practical approach to chiller plants' optimisation, Energy Build. 169 (2018) 332–343. doi:10.1016/j.enbuild.2018.03.076.

[151] P. Lee, P.T.I. Lam, W.L. Lee, E.H.W. Chan, Analysis of an air-cooled chiller replacement project using a probabilistic approach for energy performance contracts, Appl. Energy. 171 (2016) 415–428. doi:10.1016/j.apenergy.2016.03.035.

[152] H.S. Asad, R.K.K. Yuen, G. Huang, Degree of freedom based set-point reset scheme for HVAC real-time optimization, Energy Build. 128 (2016) 349–359. doi:10.1016/j.enbuild.2016.07.013.

[153] M.S. Soler, C.C. Sabaté, V.B. Santiago, F. Jabbari, Optimizing performance of a bank of chillers with thermal energy storage, Appl. Energy. 172 (2016) 275–285. doi:10.1016/j.apenergy.2016.03.099.

[154] F.W. Yu, W.T. Ho, K.T. Chan, R.K.Y. Sit, Logistic regression-based optimal control for air-cooled chiller, Int. J. Refrig. 85 (2018) 200–212. doi:https://doi.org/10.1016/j.ijrefrig.2017.09.026.

[155] M. Karami, L. Wang, Particle Swarm optimization for control operation of an all-variable speed water-cooled chiller plant, Appl. Therm. Eng. 130 (2018) 962–978. doi:10.1016/j.applthermaleng.2017.11.037.

[156] H. Satyavada, S. Baldi, An integrated control-oriented modelling for HVAC performance benchmarking, J. Build. Eng. 6 (2016) 262–273. doi:10.1016/j.jobe.2016.04.005.

[157] T. Wortmann, Genetic evolution vs. function approximation: Benchmarking algorithms for architectural design optimization, J. Comput. Des. Eng. 6 (2019) 414–428. doi:10.1016/j.jcde.2018.09.001.

Annexes

The annexes include the description of the pilot plant was used as the experimental test environment for the development of this thesis.

CONTENTS:

A1 Annex 1. Test environment

Annex 1. Test environment

For the experimental validation of the proposed methodology, data from a real tertiary-sector building has been used as a test environment. The selected building consists of a 3-floor university campus building that contains spaces with different usages, with a total surface of 2.400m2. The building is considered a research ecosystem of the Universitat Politècnica de Catalunya, which includes an installation of renewable energy sources (photovoltaics), several energy production equipment, as well as a SCADA with extensive instrumentation, permitting to be used as a pilot-plant for research in the field of energy efficiency, smart-grids and industrial electronics among others. Fig A1.1 shows a 3D representation model of the building's structure, consisting of 2 building blocks of 3 and 1 floors, respectively. The HVAC machines that were used for the validation of the proposed methodology are located on the upper part the building, on its deck.

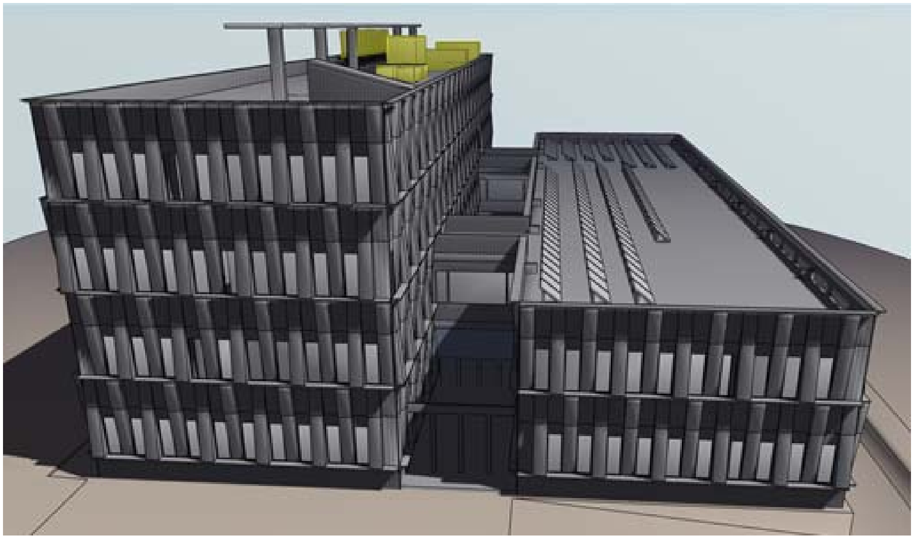


Fig A1.1 *3D representation of the pilot plant, highlighting the location of the HVAC equipment on the building's deck.*

In terms of HVAC equipment, the installation consists of two electric chillers, two electric heat pumps, one gas boiler and two air handling units, which manage the energy production, energy distribution, pre-conditioning and air-renewal for the building's spaces.

No cooling tower is present in this installation, and secondary devices such as such as valves, dampers fans or pumps are considered out of the scope of this study, with the main focus being on the equipment at the production stage

The power characteristics of the HVAC machines are listed in Table A1.1.

Id	Type	P_{elec} [kW]	$P_{thermal}$ [kW]
CH1	Electric chiller	56.6	150
CH2	Electric chiller	56.6	150
HP1	Heat pump	56.7	130
HP2	Heat pump	66.2	150
B1	Gas boiler	2	430
AHU1	Air handling unit	5.5	n/a
AHU2	Air handling unit	7.5	n/a

Table A1.1 *Power characteristics of the HVAC machines of the pilot plant.*

Most instrumentation required was already installed, configured and accessible through the SCADA's OPC server, where it could be easily queried periodically, but additional instrumentation was installed in order to provide an exhaustive view of the operation of the HVAC systems and to support the validation of the developments of this thesis. In particular, the following additional sensors were installed and incorporated into the building's SCADA:

- Individual electric power meters Circutor CVM Mini-MC-ITF-HAR-RS485-C2 plus MC3-125 current transformers were installed to measure the electrical consumption of each of the cooling production machines (CH1, CH2, HP1, HP2).

- Individual thermal power meters Kamstrup MULTICAL 602 plus ultrasonic flow meters LTM 100E to measure the produced thermal energy and allow the calculation of the coefficient of performance were also installed on each of the cooling production machines.

A summary of the signals acquired for from the HVAC equipment is presented in Table A1.2.

Name	Description
${P_{elec}}^i$	Electrical power consumption of equipment i
${P_{ther}}^i$	Thermal power production of equipment i
COP^i	Coefficient of performance of equipment i
${T_{in}}^i$	Inlet or return temperature of equipment i
${T_{out}}^i$	Outlet or impulsion temperature of equipment i
${T_{inc}}^i$	Temperature differential of equipment i
${C_n}^i$	Command signal for compressor n of equipment i

Table A1.2 *Summary of acquired machine signals from the SCADA system.*

All signals are measured using field sensors except for the COP, which is calculated as the quotient between the thermal power production and the electric power consumption, and ${T_{inc}}^i$ which is the difference between the outlet and inlet temperatures.

At the consumption stage, besides the two general distribution AHUs at the building level, each of the spaces of the building also includes one or two terminal AHUs depending on their surface, installed in offices, meeting rooms, laboratories and common areas. These units are wired to passive infrared presence detectors and use their feedback for the regulation of the temperature in each space, which allows the fine-grained control of the internal temperatures in the building. Each space is allowed to define its own comfort range, within global constraints.

A summary of the signals acquired from the AHUs present in the spaces of the building is presented in Table A1.3.

Name	Description
${T_{in}}^s$	Indoor temperature of space s
${T_{set}}^s$	Indoor temperature setpoint of space s
Occ^s	Occupancy detector signal of space s
${L_n}^s$	Load of AHU n of space s

Table A1.3 *Summary of acquired space signals from the SCADA system.*

Additionally, the building's SCADA has access to a local weather station that provides the following weather parameters shown in Table A1.4:

Name	Description
T_{ext}	Outdoor temperature
H_{rel}	Outdoor relative humidity
Sol	Solar irradiance

Table A1.4 *Summary of acquired weather signals from the SCADA system.*

The supervision and control of all of the installation is made through the main SCADA system, which monitors the operation of the equipment, the condition of the heating and cooling distribution bus, as well as the environment conditions, such as the weather, the occupancy of the spaces and the temperature setpoints configured by occupants. In order to operate the equipment, a Modbus communication bus reads status variables such as temperatures and operation modes and delivers control signals to the HVAC installation. Additionally, these signals are also exposed on an OPC server linked to the SCADA. The control of the overall HVAC system performed through the SCADA also supports manually setting up priorities and schedules for the machines as well as supervising their state in real time.

In order to acquire a dataset including all of the required signals from the building, a desktop application was implemented which allowed to select the signals of interest from the OPC server and to periodically scan and store them in a dedicated time-series database for posterior processing and analysis.