UNIVERSIDAD DE LA RIOJA

TESIS DOCTORAL

Título
An Advanced Methodology to Enhance Energy Efficiency and Performance in a Hospital Cooling-Water System
Autor/es
Eduardo Dulce Chamorro
Director/es
Francisco Javier Martínez de Pisón Ascacíbar
Facultad
Escuela Técnica Superior de Ingeniería Industrial
Titulación
Departamento
Ingeniería Mecánica
Curso Académico

Doctoral Thesis

An Advanced Methodology to Enhance Energy Efficiency and Performance in a Hospital Cooling-Water System

Author:
Eduardo Dulce Chamorro

Supervisor:
Dr. Francisco Javier Martínez de Pisón Ascacíbar

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Engineering

in the

EDMANS Group
Department of Mechanical Engineering

June, 2021

To Kseniya, my wife

Declaration of Authorship

I, Eduardo Dulce Chamorro, declare that this thesis titled, “An Advanced Methodology to Enhance Energy Efficiency and Performance in a Hospital Cooling-Water System” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.

Signed:

Date:

Acknowledgements

First and foremost, I would like to thank my thesis supervisor, Dr. Francisco Javier Martínez de Pisón Ascacíbar for all the time he has dedicated to guiding my doctoral thesis work, collaborating with and supporting me over the past five years of this demanding endeavor. From the most nascent stages of this research, he has helped me to outline the work ahead and the thesis objectives addressed herein. His high expectations in terms of technical expertise when reviewing these articles inspired me to go to great lengths: writing countless drafts before crafting a final version. This whole process has shown me that an unwavering commitment to continual improvement is the only path to achieving excellence in research work.

All of my research would not have been possible without the unconditional support of my wife Kseniya. During these last years of my thesis work, we have grown together and created a beautiful family with two wonderful children, Andrés and Víctor. My gratitude is infinite for all the time you have dedicated to us so that I could write articles and this thesis. Thank you for trusting me and for being proud of our progress. Unfortunately, during the time I was developing my thesis, my father passed away, which has meant rebuilding our lives, so as to give my mother unconditional support.

Also, I would like to dedicate this thesis to my friend Sergio's memory, for having encouraged me to take this path of doctoral studies.

And lastly, I would like to thank all the engineers and maintenance staff at the San Pedro Hospital in Logroño (La Rioja, Spain) for their dedication to continuously improving the hospital facilities and for their support integrating the new energy efficiency systems. I wish to express my gratitude for your tireless dedication to serving patients in the Riojan public healthcare system. Thank you for your time and patience during the seemingly never-ending string of untimely breakdowns. You have made it possible for this work group to succeed when faced with challenges, which your interest and cooperation never allowed to overcome us. And in particular, my sincerest thanks to Óscar for collaborating and sharing his knowledge of these complex facilities with me.

Abstract

Hospitals are a type of building with especially high energy demands; and this is owing to the fact that they run life-saving services 24 hours a day, 365 days a year. Moreover, the healthcare services offered by hospitals are growing in number and complexity, which means that their energy demands increase every year.

In order to cover the energy needs of all this activity, a vast amount of technical installations are required. In addition, supplying energy and liquids increasingly necessitates greater control, precision, and quality. Due to the critical role of cooling-water systems, this thesis focuses on these installations that are vital for both the comfort they provide through air-conditioning and for healthcare activities.

The objective of this research is to improve the performance of hospital refrigeration plants to increase energy efficiency, while also reducing inefficiencies in generator start-ups and maintenance, which are commonplace problems in this type of facility.

By applying Machine Learning (ML) models to predict cooling demand, it has been possible to anticipate, adapt, and plan for actual thermal generation to meet, but not exceed, expected demand. To obtain said models, an already existing methodology based on genetic algorithms called GAparsimony was utilized. This methodology allows parsimonious models to be obtained in an automated fashion. The algorithms used include artificial neural networks (ANN), support vector machines for regression (SVR), and extreme gradient boosting machines (XGBoost).

Prior to the modeling phase, an extensive general optimization of the cooling-water facilities was carried out; and during this process a methodology was developed to be applied in the following areas: the control system, the data acquisition system, and the physical systems. The optimization culminated with a demand prediction model being implemented in the BMS (Building Management Systems). This feature enabled the BMS to anticipate generator programming a day in advance, thus exercising predictive management.

The research presented herein has been corroborated by the results obtained when the optimization methodology was applied, and by implementing the demand prediction model in the BMS as well.

Resumen

Los hospitales son edificios que tienen una gran demanda energética debido a que en su interior albergan servicios vitales las 24 horas del día, los 365 días del año. Así mismo, las carteras de servicios de procesos asistenciales que proporcionan los hospitales son cada vez más complejos y numerosos lo que hace incrementar anualmente esta demanda.

Para dar cobertura a cada una de las actividades se requiere un gran número de instalaciones técnicas. Además, el suministro de energía y de fluidos cada vez requiere mayor control, precisión y calidad. Esta tesis se ha centrado por su importancia crítica, en la instalación de generación de agua refrigerada que se requiere tanto para confort en los sistemas de aire acondicionado, como en los procesos asistenciales.

El objetivo de este trabajo de investigación consiste en mejorar el funcionamiento de las plantas de refigeración de los hospitales para aumentar la eficiencia energética, así mismo para reducir las ineficencias en los arranques de los generadores y en el mantenimiento, problemas comunes en este tipo de instalaciones.

Mediante el uso de modelos desarrollados mediante Machine learning (ML) aplicados en la predicción de la demanda de refrigeración, se ha conseguido anticipar, adaptar y planificar la generación térmica real a la demanda prevista. Para la obtención de los mismos se ha utilizado una metodología existente basada en algoritmos genéticos denominada GAparsimony. Esta metodología permite de una manera automatizada obtener modelos parsimoniosos. Entre los algoritmos utilizados se encuentran las artificial neural networks (ANN), las suport vector machines for regression (SVR) y las extreme gradient boosting machines (XGBoost).

Previamente, se realizó una extensa optimización general de las instalaciones de generación de agua refrigerada, desarrollándose una metodología de trabajo que se aplica en los siguientes ámbitos: el sistema de control; el sistema de adquisición de datos; y en los sistemas físicos. El proceso culminó con la implantación dentro del BMS (Building Management Systems) del modelo de predicción de demanda, lo que permite anticipar un día antes la programación de los generadores necesarios, realizándose así un control predictivo.

Este trabajo queda respaldado satisfactoriamente por los datos de los resultados reales obtenidos por la aplicación de la metodología de optimización, así como por la implementación en el BMS del modelo de predicción de demanda durante la duración del estudio.

Contents

List of Figures

List of Tables

List of Abbreviations

AEMET	Spanish State Meteorological Agency *(Agencia Estatal de Meteorología)*
AI	Artificial Intelligence
ANN	Artificial Neural Network
AHU	Air Handling Unit
BMS	Building Management System
CS	Control System
CDD	Cooling Degree Days
DAS	Data Acquisition System
DPC	Data Processing Center
EER	Energy Efficiency Ratio
EU	European Union
EDA	Exploratory Data Analysis
XGBoost	Extreme Gradient Boosting Machines
FE	Feature Engineering
FS	Feature Selection
GA	Genetic Algorithm
HVAC	Heating, Ventilating and Air Conditioning
HO	Hiperparameter Optimization
HW	Hot Water
ICU	Intensive Care Unit
IEA	International Energy Agency
LON	Local Operating Network
LEB	Low Energy Building
KDD	Knowledge Discovery in Databases
MAE	Mean Absolute Error
ML	Machine Learning
MPC	Model Predictive Control
NA	Data missing values (not available)
PMS	Parsimonious Model Selection
PS	Physical System
RBF	Radial Basis Functions
RH	Relative Humidity

RMSE	Root Mean Squared Error
SJR	Scimago Journal Rank
SVR	Suport Vector Machines for Regression
UPS	Uninterruptible Power Supply
UN	United Nations

List of Symbols

$Thermal\_Power$	Thermal Power	W
$Flow\_rate$	Flow rate	l/h
$Thermal\_jump$	Thermal jump	°C
Ce	specific heat of the water	Wh/kg°C
i	individual	-
g	generation	-
λ_g^i	Chromosome	-
P	Vector of training parameters	-
Q	Vector of probabilities	-
j	Input attribute	-
J	Fitness function	-
$RMSE$	Root Mean Squared Error	-
$RMSE_{val}$	RMSE error of validation database	-
$RMSE_{tst}$	RMSE error measured with the test database	-
y_i	Prediction value	-
x_i	True value	-
$EF1\ to\ EF4$	Chiller Status	boolean
RH	Relative humidity	%
$TIMP$	Cold-ring-drive temperature	°C
$TEXT$	Outside temperature	°C
$TCONSIG$	Set-point for cold-production	°C
$TMEAN$	Average daily temperature	°C
$TMAX$	Maximum daily temperature	°C
$TMIN$	Minimum daily temperature	°C
$TENEF1\ to\ 4$	Cooling generator water temperature at the inlet	°C
$TSALEF1\ to\ 4$	Cooling generator water temperature at the outlet	°C

$time$	Time of measurement	h
$month$	Month of measurement	-
day_of_week	Day of the week	-
$Is\_holiday$	Boolean variable for holiday	boolean

Subscripts

tst	test database
val	validation database
min	minutely/minute
h	hourly/hour
d	daily/day
m	monthly/month
y	annual/year

Chapter 1

Introduction

1.1 Background

The Paris Climate Accord, signed 22 April 2016, conveys the international awareness and commitment to reducing CO2 emissions at this point in time. The objective of this agreement was designed to keep global temperature rise below 2 °C above pre-industrial levels and to limit that increase even further to 1.5 °C [1]. Achieving this goal requires cutting down on CO2 and greenhouse gas emissions as soon as possible.

This goal was translated into our regional context on 28 November 2018, when the European Commission published its Climate Strategy [2] which establishes objectives for reducing greenhouse gas emissions by the year 2020. Likewise, it outlined regulatory guidelines to achieve these objectives by 2030, along with long-term objectives for a zero-emissions, or climate-neutral, European Union (EU) by the year 2050. According to this plan, by 2050, the EU will have reduced its CO2 emissions by 80%, which would make emissions approximately equivalent to the levels of 1990.

The most recent continuation of this initiative took place in December 2019, in Madrid where United Nations member countries held the "*Climate Change Conference COP25*". The aim of this gathering was to review the progress of their previously signed commitments and extend them beyond 2020.

Buildings are the EU's biggest energy consumers and greenhouse-gas emitters: they represent approximately 40% of total energy consumption and are responsible for 36% of the EU's total greenhouse gas emissions [3]. By improving the energy efficiency of buildings, very significant energy savings could be obtained and, in turn, CO2 emissions reduced. At present, 75% of buildings are inefficient in terms of energy, which means that much of the energy they consume goes to waste. It is estimated that both energy consumed and CO2 emissions could be reduced by 5% in the short term by renovating the building stock. Currently, less than 1% of existing buildings are renovated each year. At this pace, we are sure to fail to reach the objectives laid out in the EU climate strategy.

Cooling is the fastest-growing end-use of energy in buildings. The International Energy Agency (IEA) found that the energy demand of cooling systems more than tripled between 1990 and 2018, reaching around 2,000 TWh of electricity [4]. The increase in cooling demand is impacting power generation and distribution capacity, especially during peak-demand periods (for example, July and August in Spain) and extreme-heat events. Space cooling in buildings is responsible for 50% of peak electricity demand. $CO2$ emissions from space cooling are also rising rapidly, tripling between 1990 and 2018 to reach 1,130 million tons. Air conditioning accounts for nearly 20% of total electricity use in buildings around the world today [5].

Cooling performs a critical role in hospitals for many healthcare-related activities: air conditioning (AC) in operating rooms, intensive care units (ICU), emergency rooms, etc. It is also fundamental for operating medical equipment such as that used in radiology and diagnostic imaging, scanners, refrigeration storage in blood banks, kitchens, and pharmacies; pathology and laboratories. Computer and data center racks also require cooled water. Studies have shown that the energy required by chilled-water installations and AC used to create a comfortable environment and to support healthcare activities in a medical building constitute 40% to 45% of the total energy necessary for building operations [6, 7].

Hospitals can decrease their energy consumption in chilled-water installations by more than 20% by implementing and adjusting the BMS, adequately zoning for AC, adding activation sensors in different areas, measuring and analyzing historical data from those systems, planning proper-use schedules, harnessing energy from extraction air, and regulating the speed of fans and water pumps.

Chilled water plants for refrigerating technical equipment and air conditioning (AC) in hospitals are massive energy consumers whose operating and optimization problems are commonplace and widespread. Applying improvement measures to these facilities can lead to significant energy savings, along with the economic benefits that they entail. Generally speaking, these types of plants are controlled by centralized BMS systems or by their own cooling-plant systems; however, they lack the ability to predict demand.

Some interesting related studies have been conducted to predict thermal demand in buildings using different forecasting techniques: linear regression for estimating cooling energy in condominiums [8], combining ANN with an ensemble approach or clustering-enhanced adaptative ANN to forecast building cooling loads [9, 10], Artificial Intelligence (AI) to predict energy consumption of Low Energy Buildings (LEB) [11], and a hybrid approach for building stock energy prediction [12]. In a field related to this thesis, research was conducted to forecast electrical consumption in hospital facilities based on ANN [13].

Model Predictive Control (MPC) applications for HVAC have been tested with ANN models [14, 15], including an MPC formulation framework for Enhancing Building and Heating, Ventilating and Air Conditioning (HVAC) System Energy Efficiency [16].

Some recently published methods have automated and facilitated modeling processes with hyperparameter optimization (HO) and feature selection (FS) in [17, 18]. The `GAparsimony` methodology used in the articles comprising this thesis is a genetic algorithm (GA) that conducts parsimonious model selection (PMS) [19, 20, 21]. It has been successfully applied in a range of contexts such as steel industrial processes [22], mechanical design [23], to generate a landslide susceptibility map [24], and solar radiation forecasting [25].

The actions comprising the methodology developed herein affect these control systems by improving their performance, optimization, and energy efficiency.

1.2 Problem statement and motivation

All the studies comprising this thesis were conducted with a real cooling system in the San Pedro Hospital, which is the foremost hospital in the region of La Rioja (Spain) and belongs to the Spanish national public healthcare system. The collective know-how of the maintenance staff regarding the installation was insufficient to continue further optimizing the system; they found themselves in need of new techniques and tools. Since the BMS was installed in 2008, it has been logging data continuously. However, the existing data had not been analyzed prior to this study, nor had the possibility of using it to forecast energy demand been considered.

Before this study began, the cooling plant was not performing efficiently. The primary malfunctions were the following:

- Excessive energy consumption in the cooling plant. Especially in winter season and extremely hot summer periods.

- Inefficient and repeated starts and stops of the cooling generators, which were controlled exclusively by the water-distribution temperature set-point. This malfunction negatively impacted energy efficiency and led to significant breakdowns. Top manufacturers recommend that the maximum number of starts in scroll type compressors be under 12 per hour [26]. In addition, it is recommended that the working time after a chiller starts be at least 60 minutes.

- Inefficient maintenance expenses incurred due to the lack of a daily schedule. The system required that all the cold-water pumps be ready for a start signal from the chillers. This set-up entailed high maintenance costs because operating all the cooling towers required expensive antimicrobial and chemical treatments.
- Water-ring temperature below established set-points diminished energy efficiency, e.g. two chillers began operating simultaneously when only one of them was necessary.
- Water-ring temperature above established set-points owing to sudden chiller stops, which adversely affected healthcare services.

The principal motivation for this research is, on the one hand, to create a methodology capable of optimizing the operations and energy efficiency of chilled water plants. While on the other hand, state-of-the-art machine learning (ML) techniques are used and applied to create models that predict thermal demand for cooling and can subsequently be implemented in a BMS, thereby providing predictive control systems.

1.3 Scope of research and objectives

This thesis focused on improving the energy efficiency of buildings, optimizing the electrical consumption of cooling systems, decreasing CO_2 emissions, contributing to the thermal comfort of users, and minimizing maintenance costs, all through the use of machine learning techniques.

These prediction models must be trained with data obtained from optimized systems for accurate model learning, and to subsequently report useful predictions. The optimization methodology carried out prior to calculating the models allowed for operating problems in the original system to be detected and addressed. The entire methodology for preprocessing information and creating machine learning models was based on the well-known Knowledge Discovery in Databases (KDD) process. This process made it possible to conduct the work prior to modeling in an orderly and sequenced manner, which in turn facilitated the development of the optimization methodology.

The methodology created herein enhances energy efficiency by adding predictive models of thermal cooling demand to the existing BMS that can forecast the activity of the hospital's water-cooled generators. The final model was an ensemble model comprised of the best individual models. By integrating this model into the BMS, a predicted schedule for the day ahead for the cooling generators could be generated. This allows for supervised and predictive control of the installation. The optimized system is capable of reducing ineffective starts and stops that can otherwise lead to costly breakdowns and inefficient electrical starting peaks.

The main objective of the thesis is to improve the energy efficiency and performance of hospital cooling-water plants and solve the existing commonplace problems in these installations. The following is a complete list of the partial objectives of this thesis:

1. Improve energy efficiency of hospital cooling-water plants and minimize electrical consumption and maintenance expenses.
2. Improve plant performance by adjusting the BMS parameters and physical devices.
3. Fix the existing problems described in the Problem statement.
4. Explore and analyze how the cooling plant works to develop a methodology capable of solving the common problems of these plants.
5. Develop forecasting models for the current problem using KDD methodology.
6. Develop parsimonious models with the goal of demonstrating the utility of models applied to forecasting thermal energy demand.
7. Integrate these models into the BMS to obtain a system with model predictive control (MPC).
8. Fit expected demand to available generation by establishing a schedule for the chillers for the day ahead and programming them more efficiently with these schedules.
9. Minimize the number of starts and stops of the chillers.
10. Upgrade the data acquisition and metering systems to improve the forecasting capacity of the models and reporting.
11. Measure energy savings and chiller start-up savings.
12. Test the improvements and the performance of the plant in a real installation by applying the knowledge obtained through the use of this methodology and draw conclusions.

1.4 Thematic unit

The thematic unit of this thesis is centered on improving energy efficiency and optimizing the operation of cooling-water plants in hospitals. To this end, ML techniques were proposed to predict the maximum thermal demand for cooling on the following day. Therefore, it is possible to create the most optimal programming for the plant with the most efficient combination of cooling generators, in addition to forecasting the need for other equipment, such as the cooling towers.

During the process of searching for patterns in the data and for the most influential variables in prediction, the KDD process was utilized. This structured process made it possible for the pre-optimization phase of the dataset to bring to light existing problems in the plant, which are common in this type of facility. By classifying the action points, a methodology capable of achieving the objectives could be created.

1.5 Contributions presented in the thesis

The main findings of this thesis are presented in three scientific papers, published in journals listed in the Journal Citation Reports® and also in Scimago Journal Rank®.

1.5.1 Publication I

Dulce, E., Martinez-de Pison, F.J., Parsimonious modeling for estimating hospital cooling demand to reduce maintenance costs and power consumption. In: Pérez García, H., Sánchez González, L., Castejón Limas, M., Quintián Pardo, H., Corchado Rodríguez, E., eds. Hybrid Artificial Intelligent Systems. Cham: Springer International Publishing. ISSN: 0302-9743, 2019:181–192. https://doi.org/10.1007/978-3-030-29859-3_16

The publisher and copyright holder corresponds to Springer International Publishing®. The online version of this journal published for the Hybrid Artificial Intelligent Systems conference is the following URL: https://link.springer.com/conference/hais as part of the Lecture Notes in Computer Science book series (LNCS, volume 11734). https://www.springer.com/gp/book/9783030298586

This article was published in a journal ranked by the Scimago Journal Rank (SJR) as Q2 (best quartile in 2019), "Lecture Notes in Computer Science", Computer Science (miscellaneous). The SJR rank in 2019 is 0.427, https://www.scimagojr.com/journalsearch.php?q=25674&tip=sid

The main objective of this first article was to develop a predictive model to forecast the activity of the water-cooled generators. A procedure was used to search for low-complexity models through feature selection, parameter tuning, and parsimonious model selection. This methodology followed throughout the development of this thesis is called `GAparsimony` and can be found in the `R` application repository or on Github [27], and used for free in similar applications. The methodology was tested with neural networks, support vector machines, and gradient boosting algorithms.

The "Case study description" of this article relates the existing problems in the cooling plant (uncontrolled starts and stops, break-downs, sub-cooling ring water, etc.); and a series of improvements are proposed to be implemented and tested. In this first phase of the thesis research, an adequate dataset covering more than two years was not available; thus further conclusions, beyond model suitability based on reported errors, could not be made. The time period studied during this first phase was from 2017 to the beginning of 2019, as can be seen in Figure 1.1.

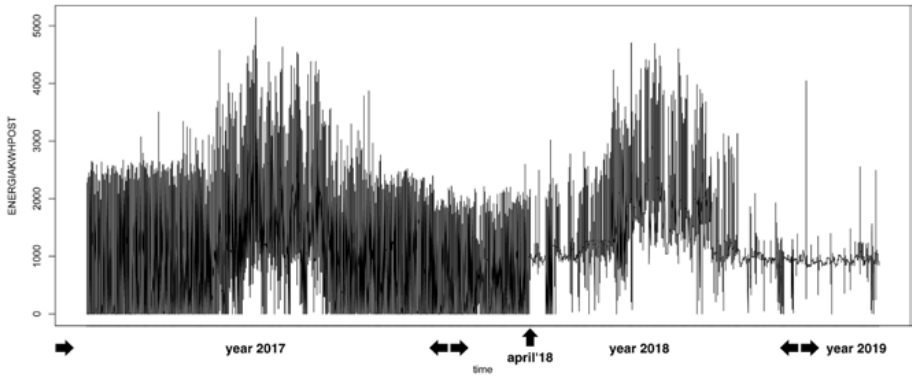

FIGURE 1.1: Evolution of generated thermal energy from 2017 to first early months of 2019.

The most significant progress of this initial research, which also provided the foundation for the subsequent studies, was the development of the scripts used to extract the BMS data, as well as the preprocessing conducted according to the KDD methodology. To this end, a script in `R` language was made to extract data from the BMS *Sauter NovaPro Open*. The BMS database stored only those records corresponding to when a variable value changed with respect to its previous recording. The script extracted the data and grouped it into hourly segments, a fact which further facilitated interpreting the model, since energy is commonly measured on an hourly basis [kWh].

Likewise, the script implemented: feature engineering (FE) tasks to create new variables such as one determining if a day “is a holiday" (*Is_holiday*); the calculation of the average daily temperature (*TMEAN*); and more importantly, the creation of a variable measuring thermal energy generation (ENERGIAKWHPOST). Since the installation did not measure generated thermal energy, this variable was obtained by calculating the instantaneous thermal power, and later by calculating its temporal grouping in energy during the time period of one hour.

Thanks to the other variables availables in the measurement system and the fact that the pump flow had a set value in this system; the thermal power could be calculated by the following formula:

$$Thermal\ Power = Flow * Thermal\ jump * Ce \tag{1.1}$$

Thermal power is expressed in watts [W]. Flow rate in l/h. Thermal jump in the chiller is expressed in degrees Celsius [ºC]. The specific heat of the water is 1.16 Wh/kg ºC. The specific weight is 1 kg/l.

Subsequently, a filter was applied to the calculated energy to smooth out the peaks due to untimely starts or stops, which would later facilitate more accurate learning in the model. Different functions were tested, as shown in Table 1.1. The Gaussian method was the method selected to filter the variable of thermal energy, ENE_GAUSSFILT7. This method was chosen for its low RMSE and MAE errors, and because the accumulated energy in the tested month was similar to the real amount of accumulated energy.

TABLE 1.1: Data filtering of the prediction variable thermal energy generated by the cooling system (ENERGYKWHPOST) tested with data from September, 2017.

Filter:	**accu.ENERGY [kWh]**	**RMSE**	**MAE**
ENERGIAKWHPOST	886,726.7	0	0
ENE_MEANFILT3	885,650.3	802.3	691.2
ENE_MEANFILT5	885,776.3	658.2	562.5
ENE_MEANFILT7	885,541.9	695.5	599.5
ENE_GAUSSFILT3	886,191.2	400.5	345.0
ENE_GAUSSFILT5	885,854.9	630.0	543.4
ENE_GAUSSFILT7	885,745.9	666.6	574.9

The root mean squared error (RMSE) is defined as the average prediction error (square root of mean squared error), where y_i is the prediction and x_i the true value:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - x_i)^2}{n}} \tag{1.2}$$

The mean absolute error (MAE) is defined as an arithmetic average of the absolute errors, where y_i is the prediction and x_i the true value:

$$MAE = \frac{\sum_{i=1}^{n}\left|y_i - x_i\right|}{n} \tag{1.3}$$

Likewise, the script cleaned and filtered the rest of the variables. Among these processes, a "*Not Available*" (NA) data resolution was realized, along with a filtering of null values and those out of the range of possibilities of maximum thermal generation.

Due to the prior adjustment of incorrect starts/stops and set-points in the generators, the generated variable of Thermal Energy, ENERGIAKWHPOST, exhibited a sawtooth graph during 2017, see Figure 1.2. Thus, the training, test, and validation of the model in this time range were not used.

And lastly, once an adequate database was available, the first models were made, as mentioned above, with SVR, ANN, and XGB algorithms using the `GAparsimony` methodology.

The author of this thesis contributed in all stages of this study. F.J. Martínez-de-Pisón assisted in developing the R scripts, applying the machine learning techniques, plotting the figures, and proofreading the text.

1.5.2 Publication II

Dulce-Chamorro, E., Martinez-de Pison, F.J., Parsimonious Modelling for Estimating Hospital Cooling Demand to Improve Energy Efficiency. Logic Journal of the IGPL, 2021, ISSN 1367-0751, https://doi.org/10.1093/jigpal/jzab008

The publisher and copyright holder corresponds to Oxford Academic®. The online version of this journal is the following URL: https://academic.oup.com/jigpal

The "Logic Journal of the IGPL" has a Journal Impact Factor in 2019 of 0.931 (ranked in Logic-Science: 3/21 (Q1), Mathematics-Science: 126/324 (Q2), Mathemathics applied-Science: 171/260 (Q3)). This

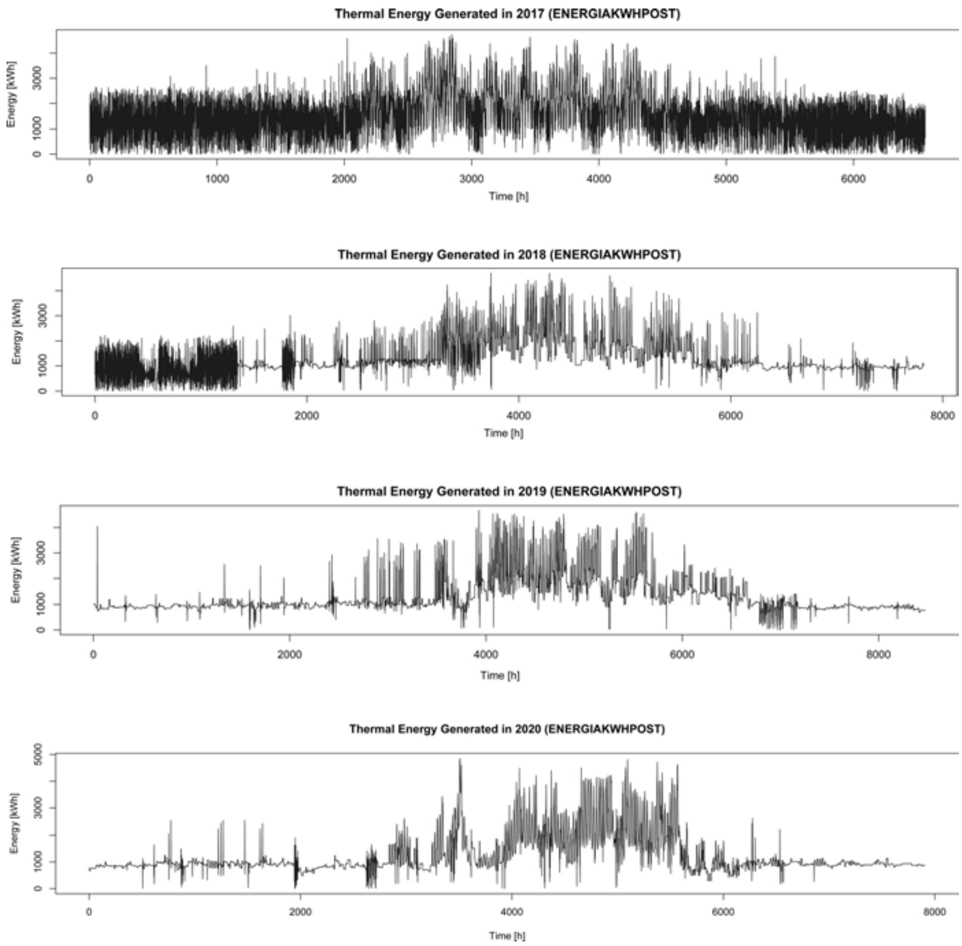

FIGURE 1.2: Evolution of the thermal energy generated by the cooling system (ENERGYKWHPOST) from 2017 to 2020.

information is taken from the Journal Citation Reports (JCR).

In this research article, there were several significant advances achieved in the modeling of the cooling water plant as compared to the first study. These improvements involved enhanced data pre-processing to reduce the noise that caused inadequate learning in the models. The decision to make this modification was based on the observation that the best model of the first generation (that of the SVR algorithm) used just three variables to adapt to the noise of the dataset during training. Among the improvement measures implemented in preprocessing, a new Gaussian filter with a larger window size was applied that further smoothed the real peaks of starts and stops in the calculated thermal generation curve ENERGYKWHPOST (see Figure 1.3).

In the present study, a hybrid model composed of the best models of each type of algorithm was used to create the final model, which improved upon the individual error of each of the initial models.

Different filters were tested, but the Gaussian filter with a window size of 11 (ENE_GAUSSFILT11) function was the method selected to filter and smooth thermal energy because of its slow error rate, as shown in Table 1.2, and because the accumulated energy in the tested year was similar to the real amount of accumulated energy. Therefore, ENE_GAUSSFILT11, was eventually selected as the target variable. This feature was considered close to the hospital's energy demand, which primarily depends upon weather conditions and the use of the facilities.

TABLE 1.2: Data smoothing of prediction variable with different filters, year 2018.

Filter:	**accu.ENERGY [kWh]**	**RMSE**	**MAE**
ENERGYKWHPOST	10.266.880,7	0	0
ENE_GAUSSFILT3	10.266.843,3	166,4	37,4
ENE_GAUSSFILT5	10.266.883,6	278,3	2,9
ENE_GAUSSFILT7	10.266.917,9	312,3	37,2
ENE_GAUSSFILT9	10.266.911,1	328,2	30,4
ENE_GAUSSFILT11	10.266.889,9	338,5	9,2

Likewise, a more exhaustive filtering of the data was conducted: especially the energy generated due to both abnormalities in plant operations and specific consumption peaks caused by chiller start-ups at maximum power due to programming errors or driver failures.

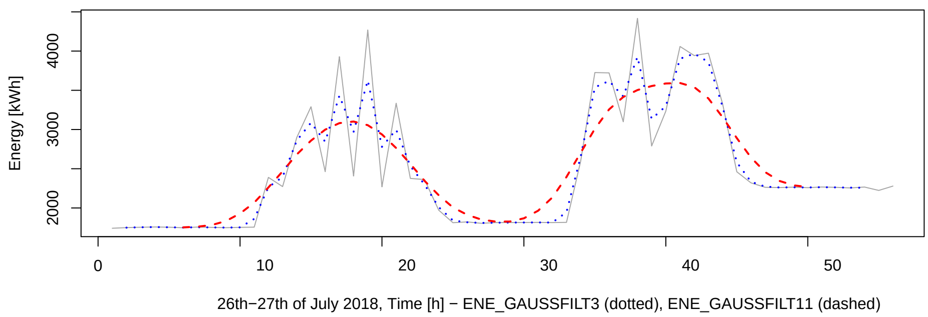

FIGURE 1.3: Filtering thermal energy generated by the cooling system (ENERGYKWHPOST) with different Gaussian steps.

The range of data available for this second study was much greater, with comprehensive data available for the years 2017, 2018 and 2019. This is why, thanks to improvements made during the optimization of April 2018, plant behavior changed significantly: the number of starts and stops decreased. This fact led to the training, validation, and testing of the models being carried out with the data collected between April 2018 and December 2019; while all previous data was discarded, as indicated in the published article.

In addition to the enhanced preprocessing, other improvements were proposed and implemented in the cooling water plant and are outlined in the third article.

The root mean squared error (RMSE) measure was chosen for model validation to evaluate the predictive model accuracy by training the model on a training dataset and testing on a test dataset. The dataset was splited into a training, validaton and test dataset, this method is known as "cross-validation".

Data modeling was done with SVR, ANN, and XGB models. These models were combined by means of an R script in such a way that takes into account the weight of each model to create an ensemble model with an error that is less than that of the best of the individual algorithm models, while also obtaining $RMSE_{val}$ and $RMSE_{tst}$ errors lower than the best errors of each model tested in the first study.

This article laid the foundation for the above described optimization methodology, organizing the measures in a conceptual framework without detailing the results and conclusions corresponding to each of action.

The author of this thesis contributed in all stages of this study. F.J. Martínez-de-Pisón assisted in developing the R scripts, applying the machine learning techniques, plotting the figures, and proofreading the text.

1.5.3 Publication III

Dulce-Chamorro, E., Martinez-de Pison, F.J., An Advanced Methodology to Enhance Energy Efficiency in a Hospital Cooling Water System. Journal of Building Engineering, 2021, 102839, ISSN 2352-7102, https://doi.org/10.1016/j.jobe.2021.102839

The publisher and copyright holder corresponds to Elsevier BV®. The online version of this journal is the following URL: https://www.journals.elsevier.com/journal-of-building-engineering

The "Journal of Building Engineering" has a Journal Impact Factor in 2019 of 3.379 (ranked in Construction and Building Technology: 15/63 (Q1), Civil Engineering: 22/134 (Q1)). This information is taken from the Journal Citation Reports (JCR).

This article outlines a complete methodology to enhance energy efficiency and solve common problems in hospital cooling-water systems based on the know-how acquired during the prior optimization conducted to create a thermal demand model of the system. The optimizations were determined through the prior KDD process and exploratory data analysis (EDA) applied in the modeling process. The methodology developed addresses the general cooling system adjustments in three main areas, which are described in Chapter II "Methodology". The implementation timeline can be observed in Figure 1.4.

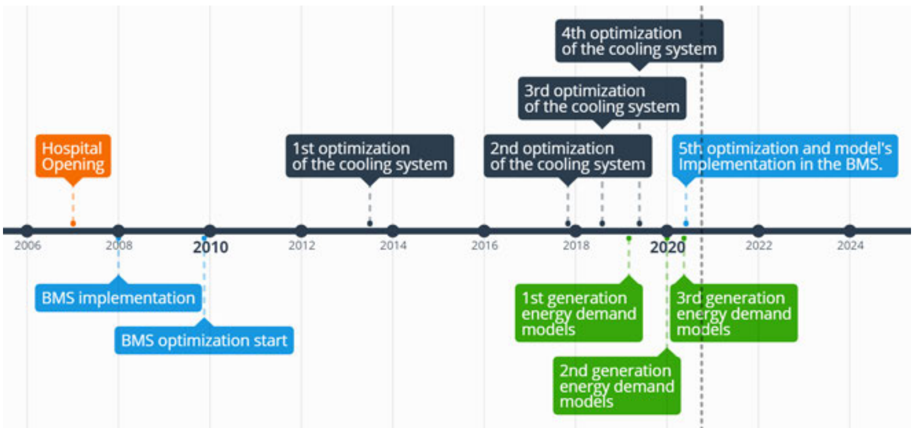

FIGURE 1.4: Case study methodology timeline indicating the most influential improvements, model generations, and implementation of the model inside the BMS.

As compared to the models of the previous articles, new variables were implemented herein to test and prove their influence on the final prediction models. The features implemented were time of measurement (*time*) 0-24 h and relative humidity (*RH*).

Various machine learning models were trained to implement the improvements during the process of updating the cooling demand model. The task of searching for low-complexity models was accomplished as in the previous studies following the `GAparsimony` methodology. In this study, the final model consisted of a weighted combination of Artificial Neural Network (ANN) and Support Vector Regression (SVR) models.

Another innovative feature was that the cooling demand model was integrated into the control system (CS). The system designed was capable of forecasting and transmitting a schedule for maximum thermal energy requirements to the BMS a day in advance. Figure 1.5 represents the classification of states depending on the maximum demand for the next day. The BMS was now able to preselect the maximum number of chillers necessary for the predicted demand and the system could also preselect the combination of chillers that best fit the demand, while also providing the greatest energy efficiency. As a result of these newfound capabilities, Figure 1.6 shows the real cooling power generation as a combination of the necessary chillers' cooling capacity. Maintenance operations also benefited from the improvements because operations in the water towers can be foreseen which reduces maintenance expenses like expensive antimicrobial and chemical treatments, etc.

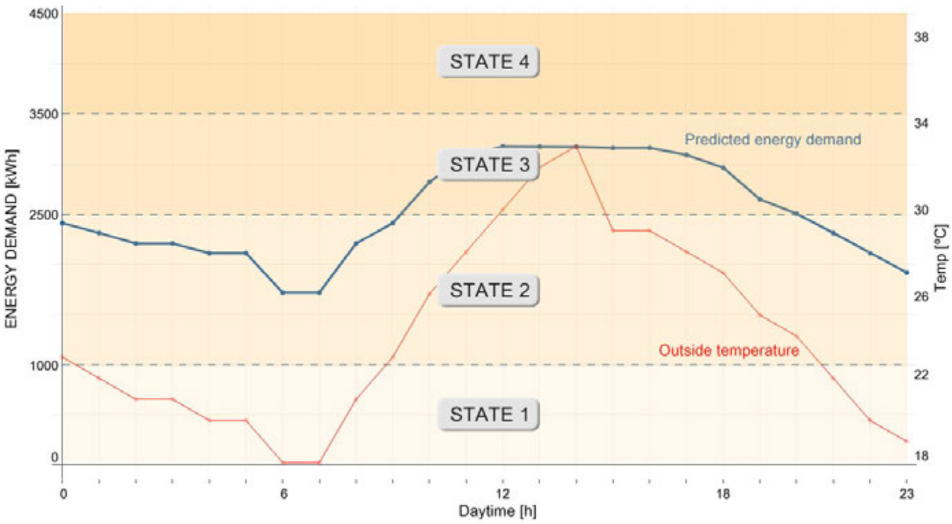

FIGURE 1.5: Energy demand classification based on the rank of maximum energy demand for the next day. Each state provides an appropriate schedule for the chillers.

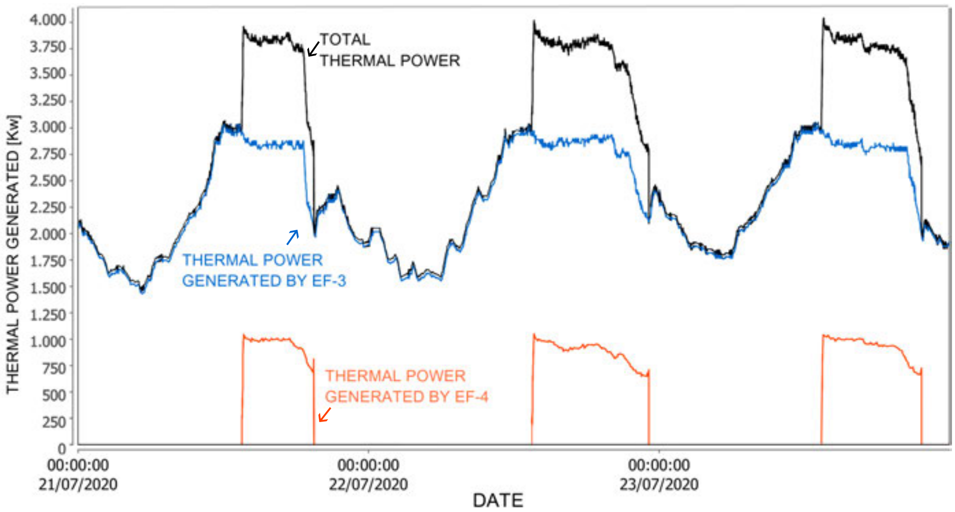

FIGURE 1.6: Cooling power generated July 21st – 23th, 2020. Note the reinforcement obtained by EF3 in addition to EF4 chiller to fit maximum demand.

The article quantified the estimated energy and percentage savings achieved by applying this methodology. In order to make an annual comparison, the method based on cooling degree days (CDD) was used. The base temperature for estimating the cooling degree days was derived by analyzing the outdoor temperature since as it increases, the cooling demand also grows from its averaged minimum value of 800 kWh, Figure 1.7. With the meteorological data and the energy recorded on a yearly basis, it was possible to normalize the energy each year to quantify the savings. The saving generated by reducing the number of chiller starts each year was also estimated after the aforementioned improvements had been implemented.

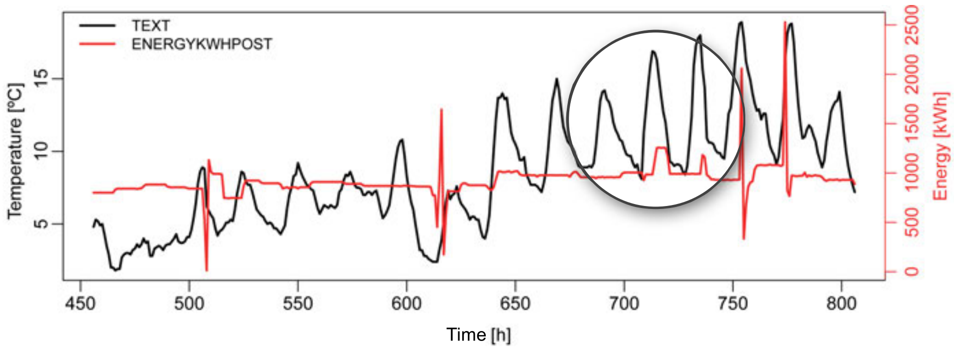

FIGURE 1.7: CDD base temperature determination graph. Outside temperature (*TEXT*) of 17 ºC was chosen as the base temperature for estimating CDD since it requires additional energy. The energy demand (red line) grows from that temperature.

This study was significantly impacted by the recommendations for increased ventilation issued to combat the COVID-19 pandemic, which entailed an increase in energy consumption in 2020, which was also the year in which the model was finally integrated. Energy consumption skyrocketed as plant operations were atypical since all areas of the hospital equipped with Air Handling Units (AHU), see Figure 1.8, were configured to avoid air recirculation and increase ventilation flow to prevent the spread of COVID-19 [28] by closing the return air dampers.

The author of this thesis contributed in all stages of this study. F.J. Martínez-de-Pisón assisted in developing the R scripts and applying the machine learning techniques.

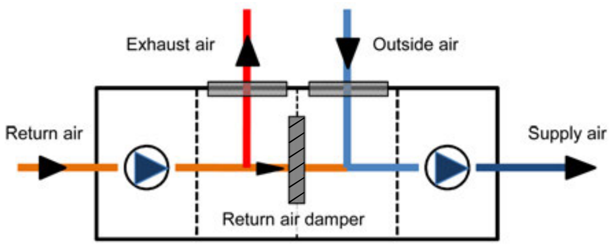

FIGURE 1.8: AHU internal scheme. A general recommendation for the SARS-CoV-2 virus is to avoid central air recirculation by closing recirculation dampers either using the BMS or manually.

1.6 Thesis outline

This dissertation is organised in seven chapters. The present chapter includes an introduction to energy efficiency in hospital cooling-water systems, and outlines the main objectives and issues to be addressed herein. The motivation of the thesis, its objectives, and a brief description of each article also can be found in this chapter. Chapter 2 describes the Methodology applied in the research and the scientific basis underlying the study. Chapters 3 through 5 contain the scientific publications contributing to this thesis. Chapter 6 presents the most remarkable results; and a general discussion of each publication, along with their limitations, is also included in this chapter. Finally, Chapter 7 summarizes the main conclusions and proposes future lines of research.

Chapter 2

Methodology

The methodology provided a structured process to review all the main aspects related to the hospital cooling-water system. The task began with a deep optimization of the installation, the goal being to solve the problems described in Subsection "Problem statement and motivation". A timeline of the study and the stages of the methodology is depicted in Figure 1.4.

The research was conducted in a structured fashion applying the KDD methodology, depicted in Figure 2.1, to develop the forecasting model and with a prior exploratory data analysis (EDA), thereby unveiling existing problems that otherwise would have not been discovered. As the final step, a forecasting demand model was implemented, which is capable of communicating the maximum energy required for the next day to the BMS.

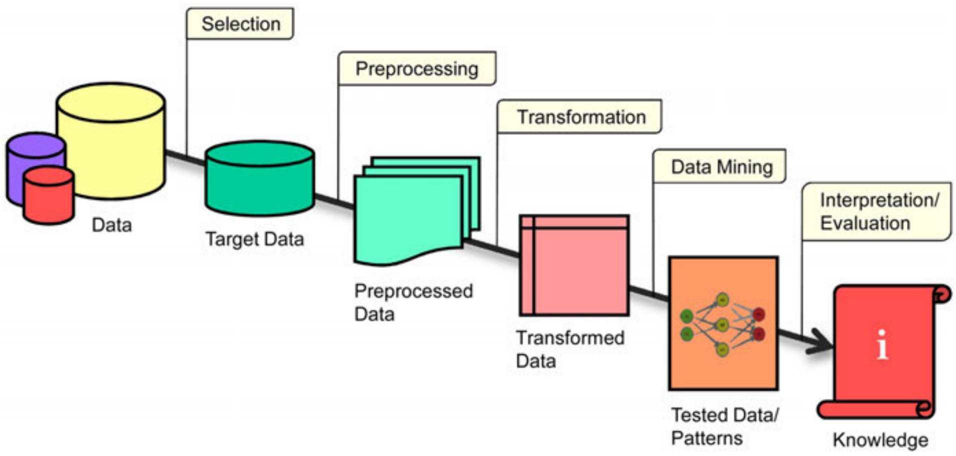

FIGURE 2.1: Knowledge discovery in databases (KDD) methodology schema.

With the knowledge acquired and with the improvements applied, this optimization methodology was developed to focus on three main areas of improvement: the control system (CS), the data acquisition system (DAS), and the physical system (PS). This methodology can be replicated in similar installations.

2.1 Control system improvements

These improvements were applied to the existing control system (CS) operated by the BMS. The CS manages the energy installations grouped in the following main areas: lighting, HVAC distribution, and heating and cooling generation. The methodology detailed herein supervised adjustments to better control the cooling system and incorporated some innovations. Among other improvements described in the aforementioned studies, the implementation of the following should be highlighted:

- A linear set-point temperature for the cooling plant.
- A minimum working time for the water-cooled generators
- A new generation schedule for Summer and Winter to adjust the demand to the appropriate chiller capacity.
- A supervised control system. The forecasting model communicates the maximum cooling-power demand for the next day to the BMS and allows the system to foresee how many chillers will be necessary. In Figure 1.6, the contribution of EF4 (1 MW) chiller to EF3 (3.5 MW) can be appreciated.

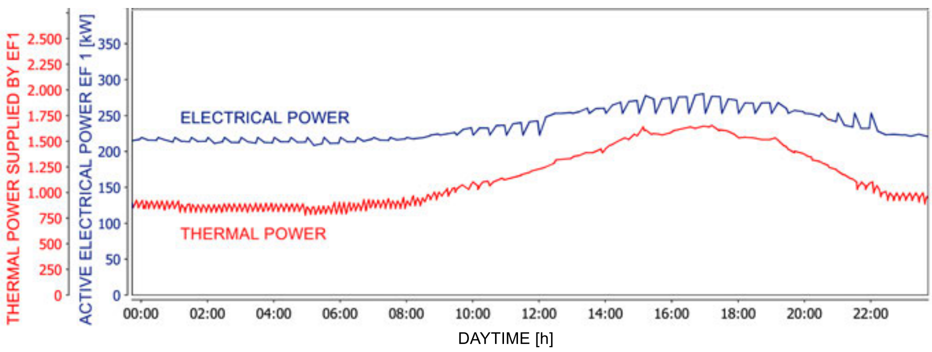

FIGURE 2.2: Thermal power generation and electrical power demand of EF1 chiller data obtained with LON cards installed during the optimization of the data acquisition system (DAS).

2.2 Improvements in the data acquisition system

These improvements affected the information acquisition and data processing system, as well as the measurement systems. Among other improvements described in the articles, the following should be noted:

- Installation of Local Operating Network (LON) communication cards in the chillers. These communication cards allow the BMS to monitor the internal operating parameters of the machine and modify the working conditions and limits.
- Installation and integration of electrical power meters into the BMS system. As an example can be visualized in Figure 2.2.

2.3 Improvements in the physical system

These improvements were made by integrating new physical systems into the existing installation, as an example:

- Installation of frequency inverter system in the screw type chillers. These systems allow that generation of cooling energy can be adapted to the demand. The effect over energy modulation can be appreciated in Figure 2.3.

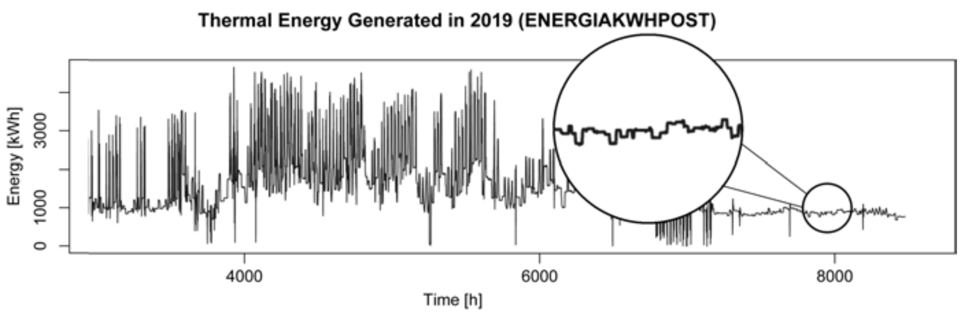

FIGURE 2.3: Detail of thermal generation showing behavior after installation of frequency inverter in the EF4 chiller in 2019.

2.4 Predictive control scheme

The activity of the hospital's water-cooled generators was improved by implementing a predictive model for cooling demand within the BMS control system that anticipates decisions, detailed in the Methodology as a Control system improvement. The incorporated control scheme is depicted in Figure 2.4.

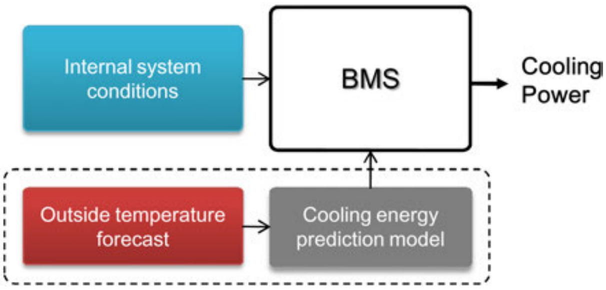

FIGURE 2.4: Control scheme. The cooling-energy prediction model communicates to the BMS the maximum thermal demand for the next day. The model reads the weather forecast conditions for the day ahead.

With this control scheme the prediction model foresees the maximum thermal energy demanded in the cooling system for the next day. This allows the BMS to anticipate the most efficient combination of chillers necessary to cover the demand or schedule the cooling towers. To do this, an R language script is executed daily. The code loads the updated external meteorological features from XML available on the Spanish State Meteorological Agency (*AEMET*) website and predicts the hourly energy demand for the next 24 hours and communicates it to the BMS.

2.5 Parsimonious Modeling

The search for parsimonious models (low complexity models) is one of the current challenges in the field of Machine Learning (ML). Among models of a similar degree of precision (accuracy), choosing those that are less complex (have less features) is recommended, given that these models will be better at generalizing the problem, perform more robustly against noise and disturbances; and they are easier for experts to interpret, and less expensive to maintain and update. Mechanisms used within KDD processes, such as regularization or feature selection, make valuable contributions in this regard.

In the studies included in this thesis, training and selecting the best parsimonious models were conducted using the `GAparsimony` methodology. This methodology performs a search for parsimonious machine learning models through optimization with genetic algorithms (GA). The final objective is to obtain models that are high in precision, yet low in complexity, using feature selection (FS), hyperparameter optimization (HO), and parsimonious model selection (PMS). In `GAparsimony`, the PMS of the best individuals of each generation is carried out in two steps: selecting the most accurate models and, from among them, choosing those with the least complexity.

The three ML algorithms that showed the best results in previous tests were selected: artificial neural networks (ANN), support vector machines for regression (SVR) with kernel based on radial basis functions (RBF), and extreme gradient boosting machines (XGB). For the third generation, the use of the XGB model was ruled out as the improvement it provided was minimal when compared to the significant computing effort it required. The final selected model was a weighted blending of the two best models obtained with ANN and SVR. All the experiments were implemented with the `GAparsimony` [27] package developed in the `R` language.

2.6 GAparsimony settings

To perform GA optimization with `GAparsimony`, it is necessary to define the chromosomes of each individual to be trained with the corresponding machine learning algorithm. In this methodology, the chromosome is defined by a combination of the algorithm's training parameters and the input attributes selected for that individual. In particular, for the SVR and ANN algorithms, each individual i of each generation g is defined by the λ_g^i chromosome, which is formed by the combination of two vectors P and Q, where the values of the vector P correspond to the training parameters of the algorithm, and Q corresponds to a vector of probabilities used for the selection of each input attribute j if $q_j \geq 0.5$:

$$\begin{aligned} SVR(\lambda_g^i) &= [P(cost,\ gamma,\ epsilon),\ Q] \\ ANN(\lambda_g^i) &= [P(size,\ decay,\ num\_epochs),\ Q] \end{aligned} \tag{2.1}$$

As a function of J (fitness function), `GAparsimony` uses the Root Mean Squared Error (RMSE) obtained with the validation database, $RMSE_{val}$. The RMSE error measured with the test database, $RMSE_{tst}$, is used to check the generalizability of the model. Finally, the complexity of the model is defined by N_{FS}, the number of attributes selected. This measure of complexity has proven to be very effective in past experiences when searching for parsimonious models with `GAparsimony` [22, 29, 23, 25].

The optimization process with `GAparsimony` genetic algorithms, represented in Figure 2.5, was defined with a population of 40 individuals evaluated in 100 iterations but with a stop criterion if the $RMSE_{val}$ error did not improve in 20 consecutive generations. The selection process used 20% of the best individuals (elite individuals) and was based on a two-step algorithm: In the first step, the selected models were ordered in an increasing manner based on the $RMSE_{val}$ error. In the second step, the individuals with similar values of $RMSE_{val}$ were re-ordered prioritizing lower complexity. This helped promote those less complex solutions (simpler because they have fewer variables) to the top positions. In this second step, two individuals were considered similar if the absolute difference of their $RMSE_{val}$ was less than a *ReRank* parameter, defined by the user. In this study, and after several experiments, *ReRank* was set at 0.1 as it showed a satisfactory balance between complexity and $RMSE_{val}$.

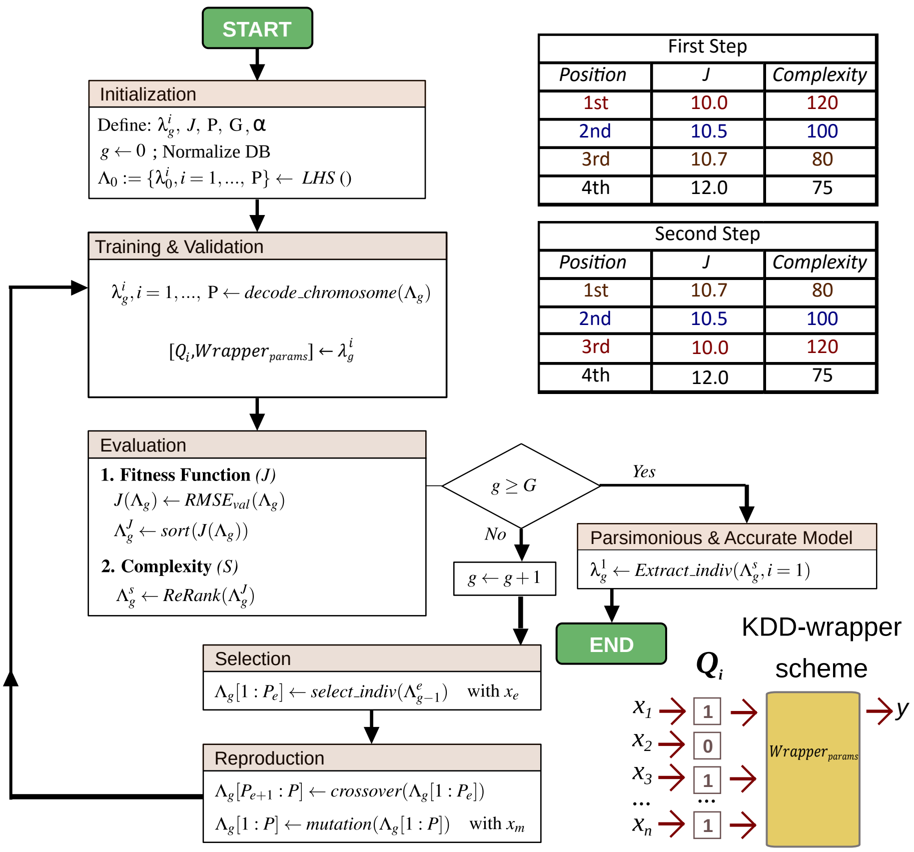

FIGURE 2.5: Flowchart of the GAparsimony with sample data.

After selecting the best individuals of a generation (the elite population), `GAparsimony` performs the traditional processes of crossing the chromosomes of the best individuals to create the next generation of individuals, as well as chromosome mutation to create more diversity of solutions in later generations. The crossover function for the P vector of the chromosomes was heuristic blending with $alpha = 0.1$. For the Q vector of the chromosomes, $random\ swapping$ was performed. In this case, the elite individuals of the previous generation also pass on to the new generation.

The first generation of individuals was created randomly, but with 90% of the characteristics of the individuals selected. This allowed the search for models to start with models that have a high number of entries.

Finally, the mutation was applied to the chromosomes of the new generation, except for the two best individuals. For the P vector of chromosomes, a random variation of 10% of the values was performed. In the case of vector Q, the probability of changing 0 to 1 was set at 10% in order to facilitate reducing the number of attributes in subsequent generations.

Chapter 3

Publication I

Dulce, E., Martinez-de Pison, F.J., Parsimonious modeling for estimating hospital cooling demand to reduce maintenance costs and power consumption. In: Pérez García, H., Sánchez González, L., Castejón Limas, M., Quintián Pardo, H., Corchado Rodríguez, E., eds. Hybrid Artificial Intelligent Systems. Cham: Springer International Publishing, 2019.
https://doi.org/10.1007/978-3-030-29859-3_16

The publisher and copyright holder corresponds to Springer. The online version of this journal is the following URL: https://www.springer.com/gp/book/9783030298586

Chapter 4

Publication II

Dulce-Chamorro, E., Martinez-de Pison, F.J., Parsimonious modelling for estimating hospital cooling demand to improve energy efficiency. Logic Journal of the IGPL, 2021.
https://academic.oup.com/jigpal/advance-article-abstract/doi/10.1093/jigpal/jzab008/6139196

The publisher and copyright holder corresponds to Oxford Academic. The online version of this journal is the following URL: https://academic.oup.com/jigpal

Chapter 5

Publication III

Dulce-Chamorro, E., Martinez-de Pison, F.J., An Advanced Methodology to Enhance Energy Efficiency in a Hospital Cooling Water System. The Journal of Building Engineering, 2021.

https://doi.org/10.1016/j.jobe.2021.102839

Chapter 6

Results and Discussion

This chapter summarizes and discusses the most relevant results included in the publications contributing to the thesis. Each of the three sections details the results of one of the publications and includes a general discussion on the implications and limitations of the study.

6.1 Results in Publication I

The complete results corresponding to this part of the research are found in the article "Parsimonious modeling for estimating hospital cooling demand to reduce maintenance costs and power consumption." (Dulce, E. and Martinez-de Pison, F.J., 2019).

6.1.1 Results and objectives

The objective of this study was to develop and evaluate an optimal and efficient model based on a genetic methodology that searches for low-complexity models to forecast the activity of water-cooled generators.

The study found that the model based on the SVR algorithm obtained the best validation and testing error with only 3 attributes:

- *month*,
- outside temperature (*TEXT*),
- minimum daily temperature (*TMIN*).

The ANN model came in second with 7 features and, finally, the XGB model had only 4 features.

Table 6.1 displays the validation and testing errors, and the final selected features of the best model from the last generation of SVR, ANN, and XGB, respectively.

TABLE 6.1: Best individual model for each algorithm obtained with GAparsimony in the first generation models.

	SVR	**ANN**	**XGB**
$RMSE_{val}$	**294.9**	327.4	347.8
$RMSE_{tst}$	**342.4**	363.3	371.1
month	1	1	1
day_of_week	0	1	1
Is_holiday	0	1	0
TIMP	0	1	1
TEXT	1	1	1
TMEAN	0	0	0
TMAX	0	1	0
TMIN	1	1	0
complexity	**3**	7	4

6.1.2 Discussion of Publication I

GAparsimony with the SVR algorithm was capable of obtaining a parsimonious model with only 3 attributes and acceptable validation and testing errors. An explication for these results can be found in the algorithm behaviour. The SVR algorithm obtained a model with a low-complexity solution that averaged 'the noise' and reduced the differences of the training database created based on the first year, and the validation/testing data compiled within the previous 12 months. This fact can be observed in Figure 1.1. Moreover, the ANN model overfits the learning process, which could explain why it used up to 7 variables to adapt to the noise of the dataset.

The study demonstrated that GAparsimony was an adequate method for selecting the best cooling demand model among different forecasting methodologies, and to adjust internal parameters as well.

6.2 Results in Publication II

The complete results corresponding to this part of the research are found in the article "Parsimonious modeling for estimating hospital cooling demand to improve energy efficiency" (Dulce, E. and Martinez-de Pisón, F.J., 2021).

6.2.1 Results and objectives

This study delved deeper into the research initiated in the first article by improving the system and updating the first generation cooling demand model. In order to create the second set of energy-demand models, the first dataset, which recorded a high level of noise produced by inefficient starts and stops prior to April 2018, was discarded given that significant optimizations had been implemented in the cooling system after this period.

The model was trained and tested with the information collected from April 2018 to December 2019. The training dataset corresponded to the period between January 2018 and February 2019. The validation data base corresponded to the even weeks between March 2019 and December 2019; and the testing database to the odd weeks of the same time period.

`GAparsimony` was used again to choose the best models among the different algorithms, adjust the internal parameters, and develop feature selection as well. Errors, parameters, and selected features are listed in Table 6.2. This table shows that the error values were moderately better than those obtained with the first generation models.

TABLE 6.2: Best models of the 2^{nd} generation with results and complexity.

	SVR	**ANN**	**XGB**
$RMSE_{val}$	**231.9**	233.2	239.8
$RMSE_{tst}$	**260.9**	268.2	267.7
month	1	1	1
day_of_week	0	0	1
Is_holiday	0	1	0
TIMP	0	1	0
TEXT	1	1	1
TMEAN	1	1	1
TMAX	1	0	0
TMIN	0	0	0
complexity	4	5	4

The best SVR model was obtained with 4 features: *month*, outside (*TEXT*), averaged (*TMEAN*), and maximum daily temperature (*TMAX*). Figure 6.1 shows, in white and gray box-plots, the $RMSE_{val}$ and $RMSE_{tst}$ SVR evolution for the most elite population of the best `GAparsimony` iteration.

This time, the best ANN model converged this time with 5 features (2 less than in the first generation model):*month*, if the day was a bank holiday (*Is_holiday*), ring temperature (*TIMP*), and outside (*TEXT*) and averaged daily temperature (*TMEAN*). ANN errors were only slightly superior to those of the SVR model.

The best XGB model was obtained with 4 features: *month*, day of week (*day_of_week*), and the external (*TEXT*) and averaged temperature (*TMEAN*) of the day.

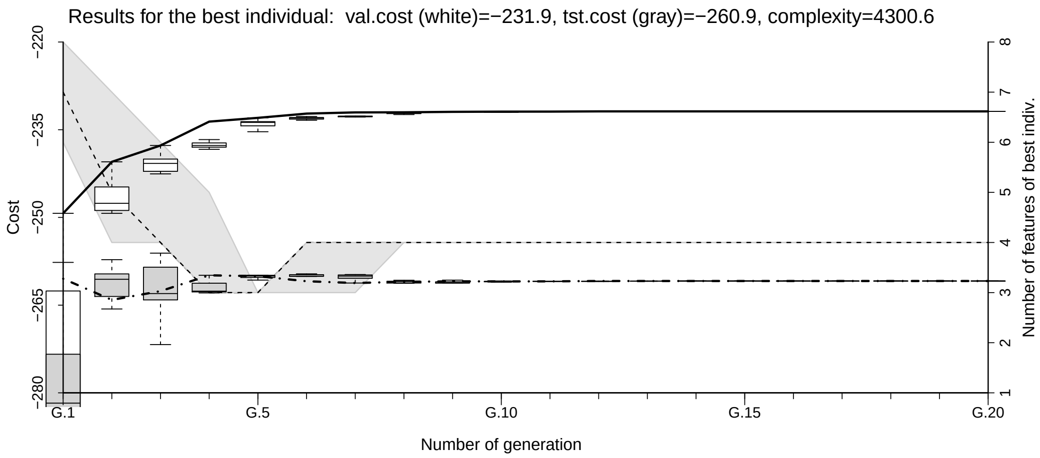

FIGURE 6.1: Evolution of errors of most elite solutions for SVR algorithm in 2^{nd} generation models. White and gray box-plots represent $RMSE_{val}$ and $RMSE_{tst}$ evolutions respectively; and continuous and shaded lines indicate the best individual of each population. The gray area covers the maximum and minimum number of features N_{FS} (righ-taxis).

The best SVR, ANN and XGB models were combined to obtain an ensemble model with an enhanced performance. Table 6.3 shows the $RMSE_{val}$ and $RMSE_{tst}$ of the weighted combined model. The process was conducted by weighting the predictions of each learner. The optimum model was comprised by the following weights:

$$Ensemble\ Model = (w1 * SVR + w2 * ANN + w3 * XGB)/3 \tag{6.1}$$

$$Ensemble\ Model = (1.36 * SVR + 1.41 * ANN + 0.23 * XGB)/3 \tag{6.2}$$

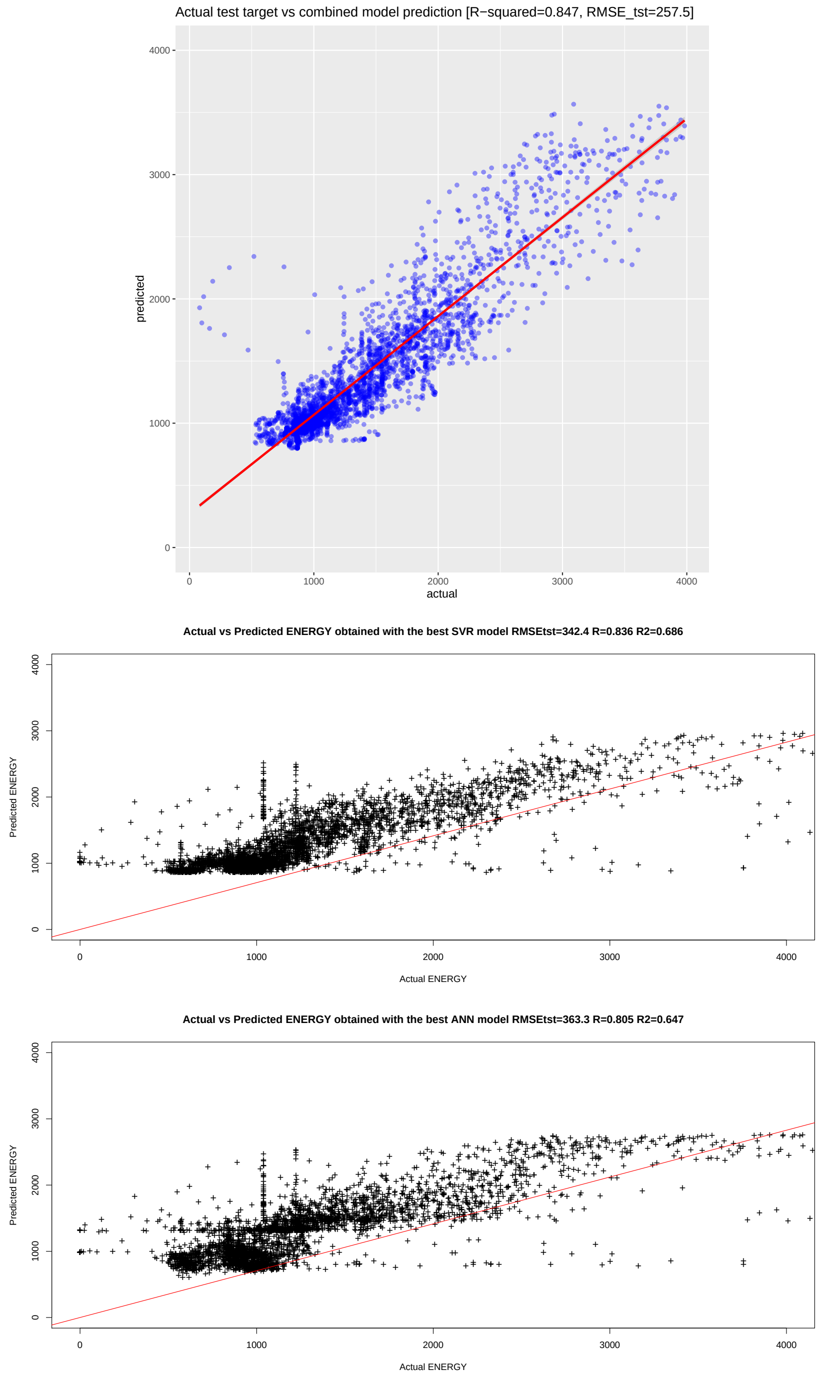

FIGURE 6.2: Ensemble, SVR and ANN combined predictions of the second generation models.

TABLE 6.3: Ensemble validation and test errors versus single models and their contribution to the hybrid model.

	SVR	ANN	XGB	HYBRID
$RMSE_{val}$	231.9	233.2	239.8	**224.8**
$RMSE_{tst}$	260.9	268.2	267.7	**257.4**
Weight %	45.3	47	7.7	-
complexity	4	5	4	7

6.2.2 Discussion of Publication II

In this phase of the research the model error values were moderately better than in the first generation models. Furthermore, the final ensemble model errors were slightly better than those of the best single model (SVR) as can be observed in Table 6.3. Complexity increased due to the number of features needed as input for the models comprising the hybrid model. The reduction in the number of features from 7 to 5 in the ANN model demonstrate that the overfitting found in the first generation model had been solved. Unlike the first generation models the variable of minimum daily temperature (*TMIN*) was not utilized in any model.

The models obtained with the ensemble model have similar errors and use similar features; and this fact demonstrates that the prior optimization process was a worthwhile endeavor.

The final ensemble model which combines the three best parsimonious models would be easy to maintain because information is directly available from sensors and external meteorological forecasting. The error rate, although not an insignificant error, does facilitate forecasting that will allow control engineers to program the chillers to supply the maximum demand for the coming hours. In addition, with the improvements made in modulating the cooling system, it will be able to buffer variations not programmed into the day-to-day activity.

The analysis conducted in the study over the course of more than three years demonstrates the pressing need to optimize cooling systems before creating effective prediction models. The datasets obtained from optimized systems avoid models learning from systems with noise, and therefore prevent model predictions from being erroneous.

6.3 Results in Publication III

The complete results corresponding to this part of the research are found in the article "An Advanced Methodology to Enhance Energy Efficiency in a Hospital Cooling-Water System" (Dulce, E. and Martinez-de Pison, F.J., 2021).

6.3.1 Results and objectives

The objective of this study was to implement the forecasting model within the BMS and then test the real response in order to validate the model and the improvements made to the cooling water plant. The model was updated with a third generation that improved upon the previous generations errors. To this end, new features such as time of measurement $0 - 24\ h$ (*time*) and relative humidity (*RH*), were added to the model.

The third generation of models were calculated with data collected from April 2018 to December 2019. The training dataset corresponded to the period between January 2018 and February 2019. The validation database consisted of the even weeks between March 2019 and December 2019; and the test database, the odd weeks of the same period. `GAparsimony` was used to choose the best models trained with SVR and ANN algorithms, discarding the XGB model because of its limited contribution. Table 6.4 shows the best SVR and ANN models: $RMSE_{val}$ and $RMSE_{tst}$, selected features with the percentage of appearance in the most elite models during the last generations, and model complexity (N_{FS}).

TABLE 6.4: 3^{rd} generation best models with RMSE errors, complexity and features used and their percentage of appearance in the group of elite models.

	SVR		ANN	
$RMSE_{val}$	**222.9**		226.0	
$RMSE_{tst}$	**256.1**		264.0	
	used	% appear.	used	% appear.
time	**1**	99.7	**1**	100
month	**1**	99.6	**1**	98.6
day_of_week	0	11.8	0	11.5
Is_holiday	0	1.9	0	7.7
TIMP	0	13.7	1	99.2
TEXT	**1**	99.6	**1**	100
TMEAN	**1**	99.5	**1**	96.4
TMAX	**1**	63.4	**1**	95.8
TMIN	0	8.5	0	11.9
RH	0	32.2	0	11.1
Complexity	5		6	

The best SVR model was obtained with 5 features: *time*; *month*, outside temperature (*TEXT*), averaged temperature (*TMEAN*), and maximum daily temperature (*TMAX*).

The best ANN model converged with 6 features: *time*, *month*, ring temperature (*TIMP*), outside temperature (*TEXT*), averaged temperature (*TMEAN*), and maximum daily temperature (*TMAX*). ANN errors were moderately higher than those of the SVR model. Figure 6.3 shows the evolution for the elitist population of the best `GAparsimony` iteration for the ANN model in the third generation models. The box-plots represent the $RMSE_{val}$ (white) and $RMSE_{tst}$ (grey) evolutions respectively on the left axis. The continuous line indicates the best individual error for validation, and the dash-dotted line indicates the best test error of each population. The gray area covers the range of features of the most elite individuals, and the dashed line indicates the minimum number of features N_{FS} on the right axis.

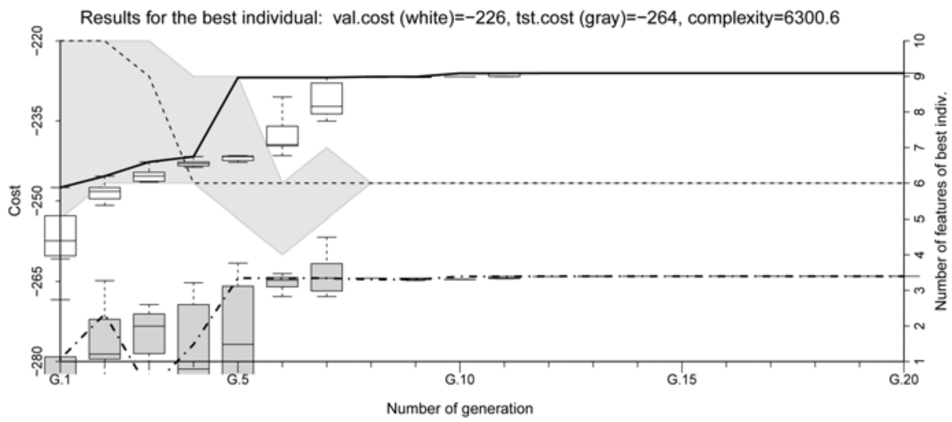

FIGURE 6.3: Error evolution of most elite solutions for ANN algorithm for 3^{rd} generation.

The SVR and ANN models were combined to reduce the $RMSE_{val}$ and to obtain the *blending model*.

Table 6.5 shows the improvement of the $RMSE_{val}$ and $RMSE_{tst}$ of the hybrid model as compared to the single models. The error rate was slightly better in the ensemble model than the best single model (SVR). Figure 6.4 shows the combined prediction for the hybrid model.

TABLE 6.5: Ensemble validation and test errors versus best single models of 3^{rd} generation and their contribution to the hybrid model.

	SVR	**ANN**	**HYBRID**
$RMSE_{val}$	222.9	226.0	**220.8**
$RMSE_{tst}$	256.1	264.0	**253.2**
Weight %	82.5	17.5	-
complexity	5	6	6

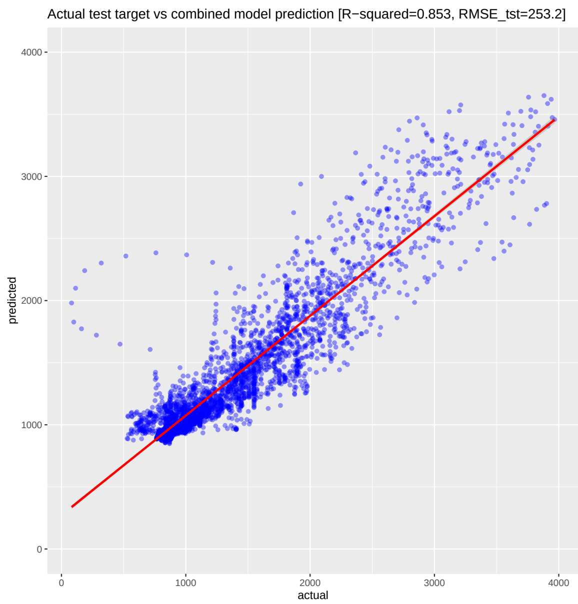

FIGURE 6.4: Combined prediction for the 3^{rd} generation hybrid model.

The graphs in Figures 6.5 compare the real registered thermal energy generated (ENERGYKWHPOST) to the energy demand forecasted by the ensemble model (ENE_GAUSSFILT11), which uses data predicted by the *Spanish State Meteorological Agency* as input). Furthermore, these graphs show the influence of the outside temperature (*TEXT**100) on demand in the dotted line: its impact increases when the outside temperature *TEXT* is more extreme (during July and August).

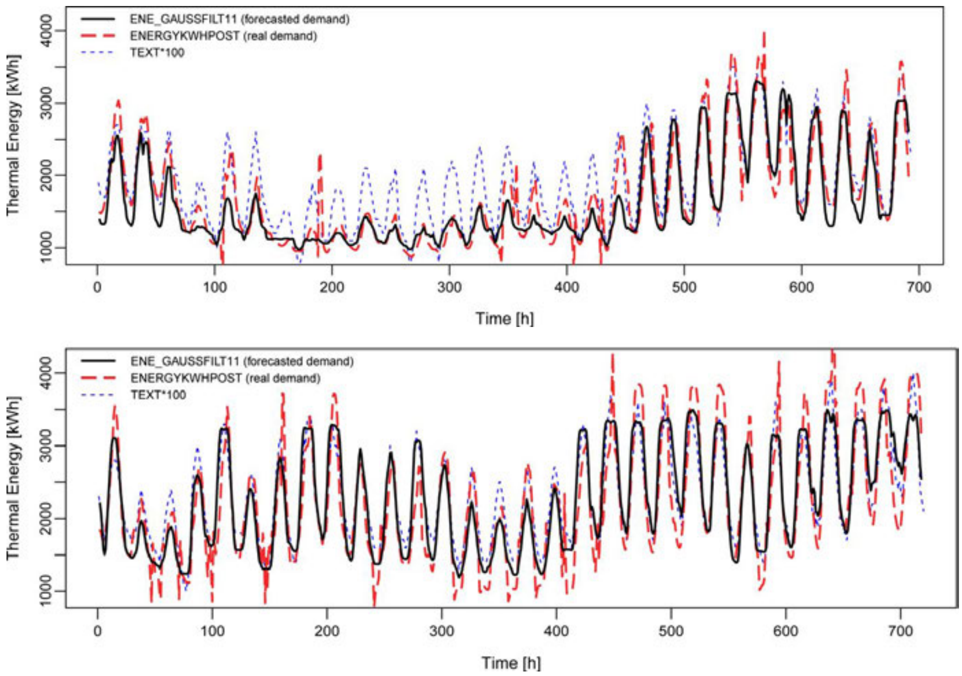

FIGURE 6.5: Energy demand forecasted (ENE_GAUSSFILT11) by 3^{rd} generation ensemble model versus real thermal energy generated (ENERGYKWH-POST) obtained from LON cards, 1-29 June and 1-30 July of 2020.

The results exhibited herein consisted of measuring the energy parameters that had been optimized, both during the initial phases and after all the updates implemented in 2020. To measure energy savings, the electrical energy records extracted from the electrical power meter located in the hospital's power plant were utilized. Annual energy savings were calculated by comparing energy demand before and after the optimizations and the model were installed. The method of cooling degree days (CDD) was used to normalize the energy consumptions for an adequate year-to year-comparison. The outside temperature of 17 ºC was chosen as the base temperature for estimating CDD since it requires additional energy to maintain the cooling system. The meteorological data for the study was obtained from the official La Rioja Government weather station [30].

Table 6.6 shows the normalization of the annual electricity consumption in the cooling system building from the year 2016 (prior to the study) to the year 2020. To perform this normalization, the annual average degree days $CDD17$ in the interval 2016 – 2020 (which was 587.3 degree days), multiplied by each value of $Energy/CDD17$ provides the normalized energy for each year (Norm.Energy).

TABLE 6.6: Normalized energy per year [kWh] prior to and over the course of the study, based on $CDD17$.

Year	CDD17	Energy [kWh]	Norm.Energy [kWh]
2016	590	5,968,990	5,942,682
2017	597	6,258,184	6,156,502
2018	576	6,124,609	6,242,594
2019	620	5,864,247	**5,553,164**
2020	553	6,400,075	6,794,584

The average cost of electrical energy during the 2017 – 2020 period for this building supplied from a 66 kV high voltage substation was 0.0988521 €/kWh.

Depending on whether the comparison is between the year 2016, prior to this study, or 2017, the first year of the study, and the year 2019, the energy savings obtained by implementing this methodology represent between 7% and 10%, which indicates economic saving of between €38,504.63 and €59,641.20 per year, as shown in Table 6.7.

TABLE 6.7: Estimated savings owing to the methodology applied.

Year	Saving (%)	Saving (€)
2016 vs 2019	7%	38,504.63
2017 vs 2019	10%	59,641.20

Unfortunatelly, during the 2020 period of the model's implementation, higher electrical consumption was registered than that of 2019 (+22.3%,+€122,716.97). The reason is that plant operations were atypical since all areas of the hospital equipped with Air Handling Units (AHU) had to be configured to avoid air recirculation and increase ventilation flow to prevent the spread of COVID-19 [28], notably reducing energy efficiency.

The monthly evolution behavior of plant energy demand during 2020 can be analyzed in Table 6.8. This table facilitates analysis of the months of higher degree days $CDD17$, which are July and August, where the comparison of normalized data is more suitable. The satisfactory evolution of the optimizations leading up to the pandemic in 2020 can be appreciated.

TABLE 6.8: Normalized energy of most demanding months [kWh], monthly *CDD*17.

Year	CDD17	July	CDD17	August
2016	177.5	691,762	180.2	627,752
2017	176.9	715,953	160.6	738,306
2018	183.8	713,190	185.9	671,673
2019	211.2	638,511	180.2	686,522
2020	180.7	**675,942**	164.0	**711,752**

The number of starts per chiller after the optimizations dropped significantly as can be observed in Figure 6.6. The results concerning the reduction in the number of starts are shown in Table 6.9, and presented by year and chiller. If the year 2017 is compared with 2019, the total number of starts decreased by 82.7%. During the year 2020, this number rose in a controlled manner due to the night-mode schedule of the EF4 chiller.

TABLE 6.9: Number of starts per chiller from 2017 to 2020, and % reduction compared to 2017 (* EF4 Chiller was damaged during 2017).

Year	EF1	EF2	EF3	EF4	TOTAL	Reduction
2017	1,911	783	1.234	0(*)	**3,928**	–
2018	971	210	137	498	**1,816**	53.8%
2019	155	122	196	206	**679**	**82.7%**
2020	91	177	192	427	**887**	77.4%

In order to be able to compare the evolution of the number of starts during the year 2020, when the forecasting models were implemented, and observe the influence of COVID-19, Table 6.10 indicates the total number of starts of all the chillers per month.

TABLE 6.10: Total number of chiller starts for each month and each year. The number of starts since the model was implemented in the BMS is marked in bold.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2017	159	207	292	315	450	416	424	442	391	348	235	249
2018	273	280	213	125	160	251	200	115	62	56	24	57
2019	26	29	34	36	55	121	111	47	62	85	35	38
2020	44	47	61	62	**46**	**55**	**111**	**166**	**88**	**82**	**37**	**88**

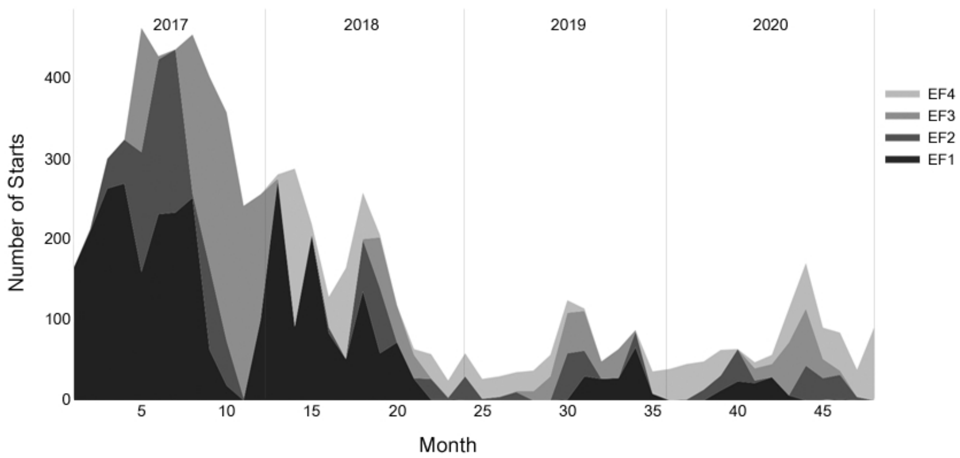

FIGURE 6.6: Number of starts per chiller from 2017 to 2020. The diagram shows the notable reduction in the number of chiller starts thanks to the optimizations.

6.3.2 Discussion of Publication III

Based on the aforementioned results, the methodology applied to optimize the cooling-water plant has demonstrated its ability to solve the preexisting problems and to improve energy efficiency, as well as to optimize chiller start-ups.

Optimizing the control system (CS) by adjusting parameters (such as set-point temperature and minimum machine working time) led to the most significant reduction in the number of chiller starts.

The number of starts per chiller decreased significantly, especially in chillers not equipped with inverter systems (centrifugal cooling generators EF1 to EF3) because of the adjustments to set-point temperature and the minimum cycle duration established as 1 hour. The number of starts decreased by 82.7% when comparing the year 2017 to 2019. During the year 2020 when the forecasting model was implemented, the total number of starts was greater than the year before. This fact can be explained by the night programming that had been activated. This modification was enacted in a controlled manner and improves energy efficiency since this chiller gives its maximum Energy Efficiency Ratio (EER) in loads within that range.

Implementing the BMS with a cooling-demand prediction model allowed plant operations and performance to be optimized. Thanks to the optimizations and features implemented in this 3^{rd} generation, errors decreased as compared to the 2^{nd} generation models. The model was simplified since the previous model was composed of 3 algorithms (SVR, ANN, XGBoost), and it is less complex as it uses less features. Therefore, this model is easier to maintain and more robust against noise. The XGBoost model was discarded because its high level of resource consumption was not compensated by the improvements it offered.

In the models comprising the final ensemble model (SVR and ANN), it should be noted that the common features influencing predictions were: *time*, *month*, outdoor temperature (*TEXT*), average temperature(*TMEAN*) and maximum daily temperature (*TMAX*). The prediction model behaves effectively, although in the months with the highest cooling energy demand (July and August), it is a conservative model and the feature "outside temperature" may have better correlation than the ensemble model (which would not be overtrained). On the other hand, it was observed that the external model that implements the weather-forecast information: outdoor temperature (*TEXT*), average temperature (*TMEAN*), and maximum daily temperature (*TMAX*), was dragging errors into the prediction results.

Improvements in the data acquisition system (DAS) enhanced the accuracy of the data collected from the chillers. However, the last models made did not include the more accurate data since this improvement occurred at the end of the optimization process.

Regarding the improvements made to the physical system (PS), it is worth highlighting the significant improvement in the modulation of the screw chiller after an inverter system was installed. This allowed the plant to work at maximum energy efficiency and significantly reduced the number of starts and electrical demand.

The methodology has achieved energy savings between 7% and 10%, but the most remarkable effect was the improvement in overall plant performance. It should also be noted that the unexpectedly greater energy demand due to increased ventilation to prevent the spread of COVID-19 obviously impacted this study. Hence, the electrical consumption data from 2020 (+22.3% as compared to 2019) cannot be factored in when calculating savings derived from implementing the prediction model.

Chapter 7

Conclusions and Future work

7.1 Conclusions

The methodology developed in this study has proven to be effective in meeting the objectives: increasing seasonal energy efficiency, reducing inefficient start-ups of cold generators, and improving overall cooling plant performance. The methodology has also innovated on the status quo in another significant way: by implementing a demand prediction model capable of communicating maximum daily demand to the BMS in advance so that it can preselect the most efficient grouping of thermal generation equipment.

As part of this thesis, three generations of demand prediction models were created until an optimal model was obtained. The final model is a simplified but more efficient ensemble model consisting of the best SVR and ANN individual models. The model could be further synthesized in practice because the regression generated by the SVR algorithm error is quite close to that of the ensemble model. Furthermore, as described in Article III, a linear regression model that used only the outdoor temperature variable was able to perform satisfactorily in certain seasonal climate conditions, though its performance was unsatisfactory at other times. Hence, the performance of a hybrid model is significantly superior thanks to its smaller seasonal error.

The features selected for the final third generation hybrid model that influence demand prediction are as follows: *time*, *month*, outdoor temperature (*TEXT*), average temperature (*TMEAN*), and maximum daily temperature (*TMAX*). Their influence on the model is detailed in Table 6.4. The model is easy to maintain and the variables are comprehensive and easy to implement. The fact that the variable *Is_holiday* was not selected as a variable for the final hybrid model attests to this assessment of the model. This variable is a Boolean indicator of whether a day is a holiday; thus it was eliminated from the model because the hospital BMS does not have a special schedule for weekends or holidays, as the desired temperatures for hospital rooms are determined by setpoints transmitted by individually-controlled thermostats or remote controls. All other refrigeration demands are programmed and, as in the case of operating rooms, demand is daily and continuous and does not depend on the day of the week or holidays, given that most healthcare services involve ongoing activity, with the exception of outpatient appointments and administrative activities. Other variables that were discarded are minimum daily temperature (*TMIN*) and relative humidity (*RH*). The latter variable did not influence this model as compared to its role in other models created in other studies using this variable. The explanation may be found in the regional inland climate of the area under study, in which *RH* is constant and in an optimal range as compared to the extreme humidity of coastal areas.

The meteorological features (outdoor temperature (*TEXT*), average temperature (*TMEAN*) and maximum daily temperature (*TMAX*)), are implemented in the demand forecasting model in a simple way: by importing models from the national weather agency's forecasting system. These models can therefore transmit their own errors into our model's predictions; nevertheless the errors are minor and restricted to certain situations.

The KDD methodology used to discover intrinsic knowledge within the data by making prediction models and the initial exploratory data analysis (EDA) enabled the research to be conducted in an organized manner, while also revealing existing problems that otherwise would not have been detected. Based on the knowledge acquired herein and the improvements achieved, this methodology was successful in implementing a prior optimization along with other improvements; and it can also be replicated in similar facilities.

The optimization of the plant and the KDD process are long-term processes: the present research was conducted over the course of more than 4 years. And in order to apply this methodology in similar hospitals, a database period of at least two years would need to be created. Hence, it is exceedingly difficult to implement this methodology from scratch in a short period of time. The analysis conducted in this thesis demonstrates the pressing need to optimize cooling systems before effective prediction models can be created. The datasets obtained from optimized systems avoid models learning from systems with noise, thereby preventing these models from making erroneous predictions.

The `GAparsimony` tool used to adjust the parameters of each algorithm and the number of variables to make parsimonious models was fundamental in this research. This tool has proven to be practical, saving enormous effort in testing, and selecting features.

The demand prediction model has been corroborated against real recorded data and graphs. Normalizing the energy made it possible to compare data across various years. Total energy savings based on implementing the complete methodology are estimated to be between 7% and 10%, as compared to the energy consumed prior to the implementation of this methodology. But, the most remarkable effect of the methodology was the improvement in the overall plant performance that can be appreciated in Figure 1.2.

The exceptional circumstances of COVID-19 that came about during the last year of this study (2020) spurred an increase in the electrical energy consumed by the cooling system, estimated at +22.3% as compared to 2019. On the other hand, the methodology managed to reduce the number of inefficient starts significantly: between 77.4% and 82.7% as described in publication III.

7.2 Future work

In terms of future ways to further improve the cooling plant within the same line of research, the forecasting model should be updated using the data obtained after the special measures implemented due to COVID-19 are lifted and a period of at least one year has passed.

Possible improvements to be incorporated into the data acquisition system include installing a physical system for measuring thermal energy generated by means of water flow meters and differential temperature probes in the chiller pipes. By referencing the information on electrical energy consumed and actual thermal energy generated, it is possible to examine actual plant efficiency and optimal working setpoints for each chiller to later transfer this information to the BMS.

Following an energy efficiency audit of the facility in 2020, new proposals emerged to enhance energy efficiency, including special standby mode programming for temperature and humidity in operating rooms during off-hours. Improving the insulation of the condensation return pipes in the towers was also proposed in order to prevent temperature increases due to solar radiation in the summer. Another possible measure would be to incorporate photovoltaic energy panels into the hospital's electrical systems for self-consumption. In this case, these systems do not require energy storage since all generated energy is immediately consumed by the large demand: the electricity generation and demand curves are correlated.

In terms of future physical improvements, there are plans to install a system that would capture surplus energy from the condensation cooling towers, which would reinforce the overall energy efficiency of the power plant. By means of a refrigerant gas evaporation system, these systems manage to capture excess heat from the pipes leading to the cooling towers, and subsequently transfer the energy acquired through condensation to the hot water storage tanks, which improves the building's overall energy efficiency by cutting down on gas consumption in the boilers.

This thesis focused on thermal generation for refrigeration. However, in 2020, during the period of study detailed in publication III, thermal energy meters were installed in the hospital boiler facilities. Thus, thermal generation data has been recorded since their date of installation. These meters will facilitate the creation of a methodology to optimize the efficiency of the heat generation plant, which would involve a demand prediction model and performance optimization.

7.3 Conclusiones

La metodología desarrollada en esta tesis doctoral ha demostrado ser eficaz en el cumplimiento de los objetivos, aumentando la eficiencia energética estacional, reduciendo los arranques ineficientes de los grupos generadores de frío, y mejorando el funcionamiento general de la planta de generación de agua refrigerada del hospital. La metodología aplicada ha innovado el estado del arte por medio de la implantación de un modelo de predicción de demanda capaz de comunicar al BMS anticipadamente la demanda máxima diaria para que este preseleccione la agrupación más eficiente de equipos de generación térmica disponibles.

Durante el desarrollo de la tesis se han desarrollado tres generaciones de modelos de predicción de la demanda hasta obtener un modelo óptimo. Este modelo final es un modelo híbrido simplificado y parsimonioso, que se compone de los mejores modelos individuales SVR y ANN. El modelo podría sintetizarse aun más en la práctica debido a que la predicción generada por el algoritmo SVR tiene un error bastante aproximado al error del modelo híbrido. Incluso como se describe en el artículo III un modelo de regresión lineal que empleara únicamente la variable temperatura exterior (TEXT) podría ser satisfactorio en ciertas condiciones climáticas temporales, aunque sería insatisfactorio en otras ya que el desempeño estacional de un modelo híbrido es muy superior si atendemos al error acumulado.

Los atributos seleccionados en el modelo híbrido final de la tercera generación y que influyen en la predicción de la demanda son: hora (*time*), mes (*month*), temperatura exterior (*TEXT*), temperatura media (*TMEAN*), y la temperatura máxima diaria (*TMAX*). Su influencia en el modelo ha quedado expuesta en la Tabla 6.4. El modelo es sencillo de mantener y las variables son coherentes y fáciles de implementar. Sirva como motivación de esta afirmación que la variable estudidada *Is_holiday*, que indica de manera booleana si el día es festivo no ha sido una variable seleccionada, lo que tiene sentido debido a que el BMS del edificio no dispone de una programación horaria especial para los fines de semana o festivos, siendo las temperaturas demandadas controladas por las consignas transmitidas por termostatos locales de las estancias. El resto de demandas de refrigeración son programadas y al igual que ocurre en quirófanos la demanda es contínua y diaria y no depende del día de la semana o de si nos encotramos en un festivo, ya que la mayor parte de la actividad sanitaria es de atención continuada a excepción de las consultas externas y de la actividad administrativa. Otras variables desechadas han sido la temperatura mínima (*TMIN*) y la humedad relativa (*RH*). Esta última variable no ha influido en este modelo frente a otros modelos estudiados en otros artículos referenciados en los que sí se emplea. La razón puede estar motivada en las características climáticas locales de la zona estudiada, muy estables y óptimas en lo que respecta a la humedad relativa si lo comparamos por ejemplo con lo extremo de las zonas costeras.

Las variables meteorológicas que emplea el modelo (*TEXT*, *TMEAN* y *TMAX*), son implementadas en el modelo de predicción de demanda de manera sencilla mediante la importación de los datos desde el sistema de predicción de la agencia nacional *AEMET*. Estos modelos a su vez pueden transmitir sus propios errores a la predicción realizada en nuestro modelo, no obstante son menores y restringidos a ciertas situaciones climáticas accidentales.

La metodología KDD para el descubrimiento de conocimiento intrínseco dentro de los datos y el exploratory data analysis (EDA) iniciales han dado como resultado realizar el trabajo de una manera organizada, y sacando a la luz problemas previos que de otra manera no hubieran sido detectados. Con el conocimiento adquirido y las mejoras efectuadas se ha desarrollado esta metodología de optimización previa y de revisión de la mejora energética que puede ser replicada en instalaciones similares.

La optimización de la planta de generación y el proceso KDD son procedimientos que requieren de mucho tiempo para su implementación. Este trabajo se ha realizado a lo largo de más de 4 años. Para la aplicación de esta metodología en hospitales similares sería necesario disponer de una base de datos que tuviera registros de más de dos años completos. Además, el análisis realizado en esta tesis demuestra la necesidad imperiosa de optimizar previamente la planta y sus subsistemas para obtener modelos de predicción efectivos. Las bases de datos obtenidas de sistemas optimizados evitan el proceso de aprendizaje de sistemas con ruido, lo que previene que las predicciones obtenidas de estos modelos sean erróneas.

La herramienta `GAparsimony` empleada para ajustar los parámetros de cada algoritmo y ajustar el número de variables para hacer modelos parsimoniosos ha resultado fundamental en el desarrollo del trabajo. Esta herramienta se ha mostrado práctica, ahorrando un enorme esfuerzo en la realización de pruebas, ajustes y en la selección de las variables.

Los datos obtenidos del modelo de predicción de demanda se han podido contrastar con la generación real de la planta de agua refrigerada por medio de los datos y gráficas registradas, realizándose una comparación y disponiéndose de resultados reales. La realización de una normalización de la energía por medio del sistema de grados día ha permitido porder comparar datos interanuales. La estimación de la mejora energética total por toda la metodología ha conseguido ahorrar anualmente entre un 7% y un 10% de la energía consumida previamente a la implementación de la metodología. Pero, la aportación más importante de la metodología es la mejora notable en el funcionamiento general de la planta de agua refrigerada.

La afección del COVID-19 en el estudio durante el último año de trabajo (2020), ha supuesto un aumento en la energía eléctrica consumida por el sistema de refrigeración estimada en un +22.3% si lo comparamos con el periodo del año 2019. Por otro lado la metodología ha conseguido reducir el número de arranques ineficientes en un alto porcentaje entre un 77.4% y un 82.7% como queda descrito en la publicación III.

7.4 Futuras líneas de investigación

En lo relativo a futuras líneas de investigación que permitan mejorar la eficiencia energética de la planta de frío, el modelo de predicción debería ser actualizado utilizando datos de una duración mínima de un año, y tomándolos posteriormente a las medidas especiales de ventilación impuestas por el COVID-19.

Entre las posibles mejoras a incorporar al sistema de adquisición de datos se encuentra la instalación de un sistema físico de medición de la energía térmica generada mediante medidores de caudal de agua y sondas de temperatura diferencial en las propias tuberías de las enfriadoras. Mediante la información de la energía eléctrica consumida y de la energía térmica real generada se podrá estudiar detalladamente la eficiencia real de la planta y los puntos óptimos de trabajo adecuados de cada enfriadora, para trasladarlos posteriormente al BMS.

Tras la auditoría de eficiencia energética realizada a la instalación durante el año 2020 han surgido nuevas propuestas de mejora de eficiencia energética como la programación especial de consignas de temperatura y humedad en modo de espera para los quirófanos, en horarios fuera de actividad quirúrgica. También se ha propuesto como mejora aumentar el aislamiento de las conducciones de retorno de agua de condensación de las torres para evitar que en la temporada estival aumente la temperatura por efecto de la radiación solar. Otra de las medidas propuestas es la incorporación en los sistemas eléctricos del hospital de paneles de energía fotovoltaica para su autoconsumo. Estos sistemas no requieren instalar dispositivos de almacenamiento de energía (baterías o similares), debido a que no son necesarios puesto que toda la energía generada es consumida instantáneamente debido a la gran demanda, y siendo además las curvas de generación y demanda eléctrica correlativas.

En lo que respecta a mejoras en el sistema físico (PS) de la planta, existen planes para instalar un sistema que capture la energía del agua que se envía a las torres de enfriamiento para su condensación. Estos nuevos sistemas extraen los excedentes de energía mediante un sistema de evaporación de gas refrigerante, consiguiendo capturar el calor excedente del agua de las tuberías que se dirigen a las torres de refrigeración, cediendo posteriormente la energía adquirida mediante condensación a los depósitos de acumulación de agua caliente sanitaria (ACS). Con ello además se mejora la eficiencia energética general del edificio al ahorrarse ese consumo de gas en las calderas.

Este estudio se ha centrado en la generación térmica para refrigeración, no obstante y durante la realización de la publicación III (en el año 2020), se instalaron en las instalaciones de las calderas del hospital contadores de energía térmica, registrando desde entonces datos de generación térmica real. Este hecho permitirá análogamente poder estudiar una metodología de optimización de la eficiencia de la planta de generación de calor, pudiendo realizarse igualmente un modelo de predicción de demanda y una optimización del funcionamiento de la misma.

Bibliography

[1] United-Nations. The paris agreement. 2016. URL: https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement.

[2] European-Commission. Eu climate action. 2018. URL: https://ec.europa.eu/clima/policies/eu-climate-action_en.

[3] European-Commission. Energy performance of buildings. 2014. URL: https://ec.europa.eu/energy/en/topics/energy-efficiency/energy-performance-of-buildings/overview.

[4] IEA. Cooling. 2019. URL: https://www.iea.org/reports/tracking-buildings/cooling.

[5] IEA. The future of cooling. 2019. URL: https://www.iea.org/reports/the-future-of-cooling.

[6] Shen, C., Zhao, K., Ge, J., Zhou, Q.. Analysis of building energy consumption in a hospital in the hot summer and cold winter area. *Energy Procedia* 2019;158:3735 – 3740. URL: http://www.sciencedirect.com/science/article/pii/S1876610219309270; innovative Solutions for Energy Transitions.

[7] IDAE, Fenercom. Guía de ahorro y eficiencia energética en hospitales. *Fenercom* 2010;:329URL: https://www.fenercom.com/publicacion/guia-de-ahorro-y-eficiencia-energetica-en-hospitales-2010/.

[8] Geekiyanage, D., Ramachandra, T.. A model for estimating cooling energy demand at early design stage of condominiums. *Journal of Building Engineering* 2018;17:43 – 51. URL: http://www.sciencedirect.com/science/article/pii/S2352710217303741.

[9] Wang, L., Lee, E.W., Yuen, R.K.. Novel dynamic forecasting model for building cooling loads combining an artificial neural network and an ensemble approach. *Applied Energy* 2018;228:1740 – 1753. URL: http://www.sciencedirect.com/science/article/pii/S0306261918311103.

[10] Luo, X.. A novel clustering-enhanced adaptive artificial neural network model for predicting day-ahead building cooling demand.

Journal of Building Engineering 2020;32:101504. URL: http://www.sciencedirect.com/science/article/pii/S235271022030752X.

[11] Paudel, S., Elmitri, M., Couturier, S., Nguyen, P.H., Kamphuis, R., Lacarrière, B., Corre], O.L.. A relevant data selection method for energy consumption prediction of low energy building based on support vector machine. *Energy and Buildings* 2017;138:240 – 256. URL: http://www.sciencedirect.com/science/article/pii/S0378778816314700.

[12] Li, X., Yao, R.. Modelling heating and cooling energy demand for building stock using a hybrid approach. *Energy and Buildings* 2021;235:110740. URL: http://www.sciencedirect.com/science/article/pii/S0378778821000244.

[13] Bagnasco, A., Fresi, F., Saviozzi, M., Silvestro, F., Vinci, A.. Electrical consumption forecasting in hospital facilities: An application case. *Energy & Buildings* 2015;103(Complete):261–270.

[14] Ruano, A.E., Pesteh, S., Silva, S., Duarte, H., Mestre, G., Ferreira, P.M., Khosravani, H.R., Horta, R.. The imbpc hvac system: A complete mbpc solution for existing hvac systems. *Energy and Buildings* 2016;120:145 – 158. URL: http://www.sciencedirect.com/science/article/pii/S0378778816301979.

[15] Afram, A., Janabi-Sharifi, F., Fung, A.S., Raahemifar, K.. Artificial neural network (ann) based model predictive control (mpc) and optimization of hvac systems: A state of the art review and case study of a residential hvac system. *Energy and Buildings* 2017;141:96 – 113. URL: http://www.sciencedirect.com/science/article/pii/S0378778816310799.

[16] Serale, G., Fiorentini, M., Capozzoli, A., Bernardini, D., Bemporad, A.. Model predictive control (mpc) for enhancing building and hvac system energy efficiency: Problem formulation, applications and opportunities. *Energies* 2018;11(3). URL: https://www.mdpi.com/1996-1073/11/3/631.

[17] Husain, H., Handel, N.. Automated machine learning. a paradigm shift that accelerates data scientist productivity. 2017. URL: https://medium.com/airbnb-engineering/.

[18] Feurer, M., Klein, A., Eggensperger, K., Springenberg, J., Blum, M., Hutter, F.. Efficient and robust automated machine learning. In: Cortes, C., Lawrence, N.D., Lee, D.D., Sugiyama, M., Garnett, R., eds. *Advances in Neural Information Processing Systems 28*. Curran Associates, Inc.; 2015:2962–2970. URL: http://papers.nips.cc/paper/5872-efficient-and-robust-automated-machine-learning.pdf.

[19] Sanz-Garcia, A., Fernandez-Ceniceros, J., Antonanzas-Torres, F.,

Pernia-Espinoza, A., Martinez-de Pison, F.J.. GA-PARSIMONY: A GA-SVR approach with feature selection and parameter optimization to obtain parsimonious solutions for predicting temperature settings in a continuous annealing furnace. *Applied Soft Computing* 2015;35:13–28.

[20] de Pison, F.M., Ferreiro, J., Fraile, E., Pernia-Espinoza, A.. A comparative study of six model complexity metrics to search for parsimonious models with gaparsimony r package. *Neurocomputing* 2020;URL: https://www.sciencedirect.com/science/article/pii/S0925231220317525.

[21] de Pison, F.M., Gonzalez-Sendino, R., Aldama, A., Ferreiro-Cabello, J., Fraile-Garcia, E.. Hybrid methodology based on bayesian optimization and ga-parsimony to search for parsimony models by combining hyperparameter optimization and feature selection. *Neurocomputing* 2019;354:20–26. URL: https://www.sciencedirect.com/science/article/pii/S092523121930459X; recent Advancements in Hybrid Artificial Intelligence Systems.

[22] Sanz-García, A., Fernández-Ceniceros, J., Antoñanzas-Torres, F., Martínez-de Pisón, F.J.. Parsimonious support vector machines modelling for set points in industrial processes based on genetic algorithm optimization. In: *International Joint Conference SOCO13-CISIS13-ICEUTE13*; vol. 239 of *Advances in Intelligent Systems and Computing*. Springer International Publishing; 2014:1–10.

[23] Fernandez-Ceniceros, J., Sanz-Garcia, A., Antonanzas-Torres, F., de Pison, F.M.. A numerical-informational approach for characterising the ductile behaviour of the T-stub component. Part 2: Parsimonious soft-computing-based metamodel. *Engineering Structures* 2015;82:249 – 260.

[24] Hong, H., Tsangaratos, P., Ilia, I., Loupasakis, C., Wang, Y.. Introducing a novel multi-layer perceptron network based on stochastic gradient descent optimized by a meta-heuristic algorithm for landslide susceptibility mapping. *Science of The Total Environment* 2020;742:140549. URL: https://www.sciencedirect.com/science/article/pii/S0048969720340717.

[25] Antonanzas-Torres, F., Urraca, R., Antonanzas, J., Fernandez-Ceniceros, J., de Pison, F.M.. Generation of daily global solar irradiation with support vector machines for regression. *Energy Conversion and Management* 2015;96:277 – 286.

[26] Danfoss. Guías de Selección y Aplicación. Performer Compresores scroll Sencillos, 20 a 110 kW 50 - 60 Hz; 2005.

[27] Martínez-De-Pisón, F.J.. GAparsimony: GA-based optimization R package for searching accurate parsimonious models.; 2017. URL:

https://github.com/jpison/GAparsimony; R package version 0.9-1.

[28] Morawska, L., Tang, J.W., Bahnfleth, W., Bluyssen, P.M., Boerstra, A., Buonanno, G., Cao, J., Dancer, S., Floto, A., Franchimon, F., Haworth, C., Hogeling, J., Isaxon, C., Jimenez, J.L., Kurnitski, J., Li, Y., Loomans, M., Marks, G., Marr, L.C., Mazzarella, L., Melikov, A.K., Miller, S., Milton, D.K., Nazaroff, W., Nielsen, P.V., Noakes, C., Peccia, J., Querol, X., Sekhar, C., Seppänen, O., ichi Tanabe, S., Tellier, R., Tham, K.W., Wargocki, P., Wierzbicka, A., Yao, M.. How can airborne transmission of covid-19 indoors be minimised? *Environment International* 2020;142:105832. URL: http://www.sciencedirect.com/science/article/pii/S0160412020317876.

[29] Urraca-Valle, R., Sanz-García, A., Fernández-Ceniceros, J., Sodupe-Ortega, E., de Pisón Ascacibar, F.J.M.. Improving hotel room demand forecasting with a hybrid GA-SVR methodology based on skewed data transformation, feature selection and parsimony tuning. In: Onieva, E., Santos, I., Osaba, E., Quintián, H., Corchado, E., eds. *Hybrid Artificial Intelligent Systems - 10th International Conference, HAIS 2015, Bilbao, Spain, June 22-24, 2015, Proceedings*; vol. 9121 of *Lecture Notes in Computer Science*. Springer. ISBN 978-3-319-19643-5; 2015:632–643.

[30] de La Rioja, G.. Agroclimatic information website la rioja spain. 2020. URL: https://www.larioja.org/agricultura/es/informacion-agroclimatica/.