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

Eduardo Dulce-Chamorro[1,*], Francisco Javier Martinez-de-Pison[a]

[a]*EDMANS Group, Department of Mechanical Engineering, University of La Rioja, Logroño, Spain.*

Abstract

Healthcare facilities consume massive amounts of energy. This study outlines a methodology to enhance energy efficiency and solve common problems in hospital cooling-water systems, since hospitals are the most energy-intensive type of building. Building Management Systems (BMS) are a widely used technique to control and monitor all the different energy facilities contained in hospitals. Proper setup and upgrades can resolve inefficiencies and existing problems. The methodology described herein addresses the general cooling system adjustments in three main areas: control system (CS), data acquisition system (DAS), and physical system (PS). An innovative feature incorporated in this methodology is the cooling demand model integrated into the CS, which is capable of forecasting and transmitting a schedule for maximum thermal energy requirements to the BMS a day in advance, thereby anticipating decisions and scheduling energy generation and maintenance operations. During the process of developing the cooling demand model, various machine learning models were trained. This process consisted of searching for low-complexity models using a methodology called `GAparsimony`. This methodology uses genetic algorithms to search for highly precise, robust models that use a low input. The final model consisted of a weighted combination of Artificial Neural Network (ANN) and Support Vector Regression (SVR) models. The energy savings obtained thanks to this methodology are estimated to be between 7% and 10% per year. The energy plant improved its performance and chiller starts were reduced by 82.5%. It should also be noted that this study was affected by the recommendations for increased ventilation due to the COVID-19 pandemic, which entailed at 22.4% increase in energy consumption in 2020. The methodology was developed and tested successfully in a real hospital BMS between 2017 and 2019; the model was finally integrated in 2020.

Keywords: Cooling demand forecasting, Building Management Systems (BMS), Energy efficiency, `GAparsimony`, Parsimonious modeling, Ensemble algorithms.

1. Introduction

The Paris climate accord (signed April 22th, 2016) 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]. This goal requires that global carbon emissions drop as soon as possible, in order to "achieve a balance between anthropogenic emissions by sources and removals by sinks of greenhouse gases". In accordance with this agreement, on November 28th, 2018 the European Commission published its Climate Strategies [2]: establishing greenhouse-gas-emissions reduction targets for 2020, key laws and measures to achieve their goals for 2030,

*Corresponding author
Email address: `edulce@riojasalud.es` (Eduardo Dulce-Chamorro)

and the long-term objective of a climate-neutral European Union (EU) by 2050. That is, by 2050, the EU will reduce its emissions by 80%, to below 1990 levels. More recently, the UN Climate Change Conference COP25 (2˘13 December, 2019) took place in Madrid with the purpose of following up on the implementation of the Paris agreement guidelines and build prospects ahead of 2020.

European buildings are the greatest consumers of energy and are responsible for approximately 40% of the EU's total energy consumption of its CO2 emissions [3] and 36%. Improving the efficiency of existing buildings could potentially lead to significant energy savings and lower CO2 emissions by about 5%, in the case of both total energy and CO2 emissions.

The International Energy Agency (IEA) has found that cooling is the fastest-growing end-use of energy in buildings, as 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 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].

This study focuses on a chilled-water installation because of its essential role in hospitals for healthcare 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 bank, kitchens, and pharmacies; pathology, the morgue, and laboratories. Computer and data center racks also require cooled water. Studies have shown that the energy required by chilled-water installations and AC in a medical building constitutes 40% to 45% of the total energy necessary [6, 7]. Hospitals can decrease their energy consumption by more than 20% by implementing a BMS, adequately zoning for AC, adding measurement sensors in different areas, analyzing historical data from those systems, planning proper use schedules, harnessing energy from extraction air and regulating the speed of fans and water pumps.

The methodology presented herein enhances energy efficiency and solves common problems in hospital cooling plants. Its foremost innovation is that it incorporates a predictive model of thermal cooling demand to the BMS that can forecast the activity of the water-cooled generators. The model integrated into the BMS creates a predicted schedule for the day ahead for the cooling generators. The optimized system is capable of reducing ineffective starts and stops that can otherwise lead to costly breakdowns and inefficient electrical starting peaks.

This study has shown that the methodology proposed is effective in improving the building's energy efficiency, optimizing the electrical consumption of cooling systems, decreasing CO_2 emissions, contributing to the thermal comfort of users, and minimizing maintenance costs through the use of machine learning techniques.

1.1. Related studies

In the past some interesting related studies have been conducted to predict thermal demand in buildings using different forecasting techniques: linear regression for estimating cooling energy of 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 hybrid approach for building stock energy prediction [12].

Likewise, electrical demand models have been designed to predict the energy consumption of Heating, Ventilation and Air Conditioning (HVAC) systems applying different techniques: algebraic modeling [13], ANN [14], an ANN comparison with Random Forest (RF) [15], and SVR [16]. In a field related to the present study, research has been conducted to forecast electrical consumption in hospital facilities based on ANN [17].

Model Predictive Control (MPC) applications for HVAC models have been tested with ANN models [18, 19], including a MPC formulation framework for Enhancing Building and HVAC System Energy Efficiency [20].

Some recently published methods have automated and facilitated modeling processes with hyperparameter optimization (HO) and feature selection (FS) in [21, 22]. The `GAparsimony` methodology used in this study is a genetic algorithm (GA) that conducts parsimonious model selection (PMS) [23]. It has been successfully applied in a range of contexts such as steel industrial processes [24], hotel room booking forecasting [25], mechanical design [26], and solar radiation forecasting [27].

This article presents a new methodology that has been successfully applied in an actual large-sized hospital to improve energy efficiency and performance by forecasting the thermal energy demand of the cooling-water system. The final model integrated into the BMS is an ensemble model comprised of the best SVR and ANN models built with the `GAparsimony` methodology for parsimonious modeling. The optimizations and the application of the improvements were carried out over the past three years. The results were integrated, tested and measured during 2020.

2. Case study description

2.1. Main hospital description

The San Pedro Hospital is the foremost hospital in the region of La Rioja (Spain) and belongs to the Spanish national public healthcare system. The area of the hospital is about $126,057.83\ m^2$ and has seven above-ground floors.

As the primary hospital in the area, it offers a wide range of medical services, the most energy-demanding of which are: over 600 beds for hospitalization (which were fully occupied during the first wave of COVID-19), a diagnostic imaging area, 23 operating rooms, an emergency area with 40 beds, hemodialysis, two ICUs with 32 beds (one of them adapted during the COVID-19 crisis), endoscopy, laboratories, pharmacy, sterilization, and other general services.

2.2. Technical description of the cooling system

Hospital's high voltage facilities, water tanks, emergency generators, storage of medical gases, cold-water production system and heating installations are centralized in a separate building and then distributed by a ring pipe around the basement of the building.

The cooling generation system under study consists of three centrifugal chiller units of 3.51 MW, namely the *Trane CVFG* model (herein designated as EF1, EF2, EF3), and another screw machine of 1 MW cooling capacity, namely the *Trane RTHD* model (designated as EF4). Figure 1 shows the hydraulic schema of the water-cooling facility.

The BMS is comprised primarily by controllers belonging to the *Sauter EY3600* family. The BMS interface is a SCADA application with a *novaPro Open 4.1.* environment. The existing BMS controlled the system on a real-time basis, using information captured by sensors and ordering actions to the actuators when temperature set-points exceeded pre-determined values.

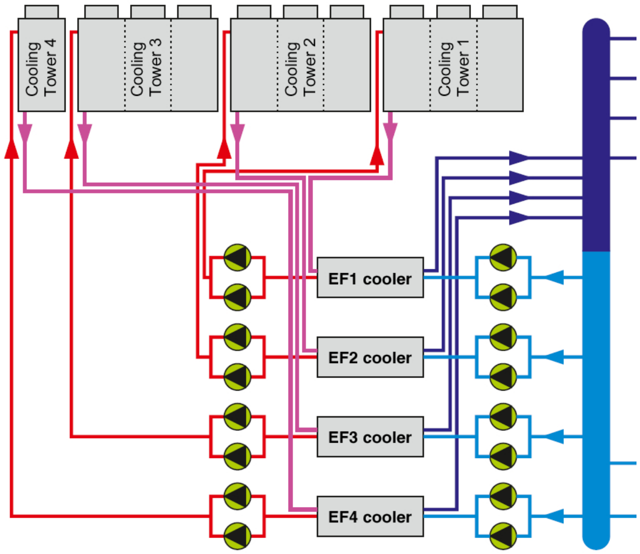

Figure 1: Hydraulic schema of cooling-water-generation system.

2.3. *Past problems in the cooling system*

Before starting the optimization process, the following malfunctions were detected:

- Inefficient and repeated starts and stops of the cooling generators, which were controlled exclusively by the water-distribution temperature set-point. This malfunction impacted energy efficiency negatively and can lead to significant breakdowns. The top manufacturers recommend that the maximum number of starts in scroll type compressors be under 12 per hour, and 6 in compressors equipped with inverters [28]. In addition, it is recommended that the working time after a chiller starts be at least 60 minutes.

- Inefficient maintenance expenses incurred because of a lack of a daily schedule. The system required having all the cold-water pumps ready for a start signal from the chillers. This fact entailed high maintenance costs because operating all the cooling towers required expensive antimicrobial and chemical treatments.

- Subcooling-water-ring temperature below established set-points diminished energy efficiency, e.g. two chillers begin operating simultaneously when only one of them was necessary.

- Overheating-water-ring temperature above established set-points owing to sudden chiller stops, which adversely affected healthcare services.

3. Methodology

The methodology provides a structured process to review all the main aspects related to the cooling-water system. As a final step, a forecasting demand model was implemented that can transmit the maximum energy required for the next day to the BMS. The process, which started in 2017, began with a deep optimization of the installation, the goal being to solve the problems described in Subsection 2.3. A timeline of the study and the stages of the methodology is depicted in Figure 2.

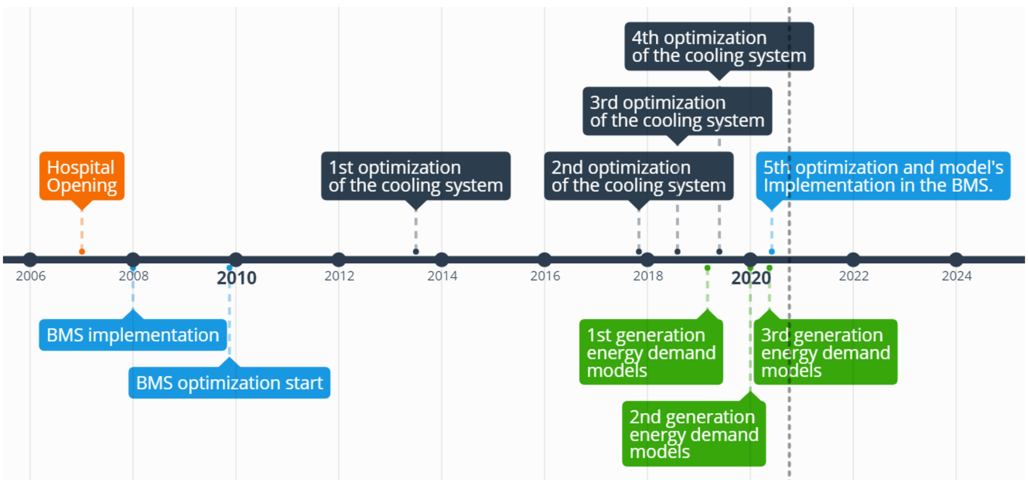

Figure 2: Case study timeline indicating the most influential improvements, model generations and implementation of the model inside the BMS.

3.1. Control system improvements

These enhancements are applied to the control system and affect the set-points, the behavior of the field elements and actuators, the operating schedules and the predictive control systems that could be integrated.

The main facilities of the hospital are controlled by the BMS system. It was implanted in 2008, one year after the hospital opening. Since 2010 on-going optimization of the BMS has been underway in three main areas: lighting (adding sensors and schedules), HVAC distribution adjustments (adding sensors, and implementing schedules in fan coils, air conditioners, and pumps), heating generation (implementing schedules in pumps, optimizing the system).

The 1^{st} optimization of the cooling system was developed before this study. It implemented a stepped set-point temperature of the cold-water ring (TCONSIG) that was calculated depending on the outside temperature, instead of a fixed value, as beforehand.

In the 2^{nd} optimization, a minimum working time for the water-cooled generators was established of at least one hour, and a cyclic order of use for the chillers was set up. This optimization significantly improved the behavior of the cooling-water generation system as can be observed in Figure 10 (from April 2018): the number of starts and stops in the chillers decreased dramatically. The improvement can also be appreciated in Figure 17, which depicts the number of starts of chillers.

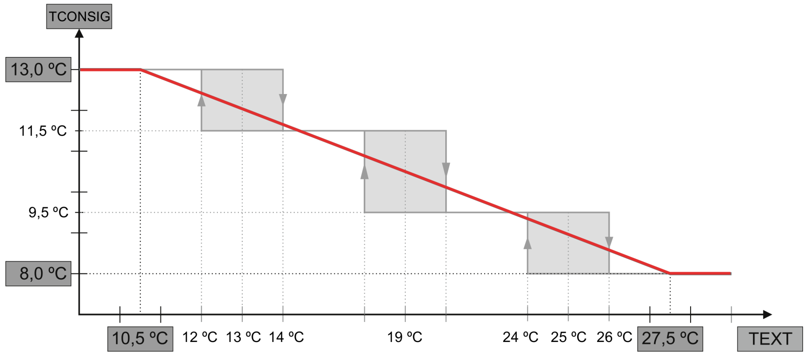

Figure 3: Set-point temperature of the cooling generation system (TCONSIG) calculated with the Exterior Temperature (TEXT) and modified in the third optimization from a stepped set-point to a linear one.

In the 3^{rd} optimization, a linear set-point temperature of the cold-water ring (TCONSIG) was applied and, in this optimization, calculated in proportion to the outside temperature (TEXT) instead of as a stepped variable, as can be seen in Figure 3.

A supervised control system was implemented in the 5^{th} optimization, described in Subsection 3.4. The 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. This also provides a schedule that allows planning for which chillers will be in operation, optimizing energy efficiency and planning maintenance operations. In Figure 4, the contribution of EF4 (1 MW) chiller to EF3 (3.5 MW) can be appreciated: these two chillers work in conjunction to best fit the power generation to the day's maximum demand, instead of starting two of the 3.5 MW chillers and subcooling the ring temperature.

Additionally, in this final step, a new generation schedule was implemented for Summer and Winter to adjust the demand to the appropriate chiller capacity. The scheduling establishes separate day and night programs in Summer. Night programming establishes the priority of use of the small chiller EF4 if the outside temperature is below 17 ºC. The winter schedule is similar to the summer-night program.

3.2. Improvements in the data acquisition system

These improvements affect the information acquisition and data processing system, as well as the measurement systems.

Local Operating Network (LON) communication cards were installed in every generator in the 4^{th} optimization to improve the internal adjustments of the chillers that the BMS had not been able modify before. These communication cards allow the BMS to monitor the internal operating parameters of the machine and modify the working conditions and limits. They allow the set-point to be modified and limit the electrical power of each generator. This enables the machine's consumption to be adjusted according to the immediate needs of the facility. The maximum power limitations of EF1, EF2, and EF3 were upgraded from 70% to 90%, thereby providing a maximum cooling power greater than 3 MW per chiller.

Furthermore, after these cards were set up, the quality and quantity of data recorded improved dramatically in terms of precision as compared to data recorded with external sensors, as can be observed in Figures 4 and 5.

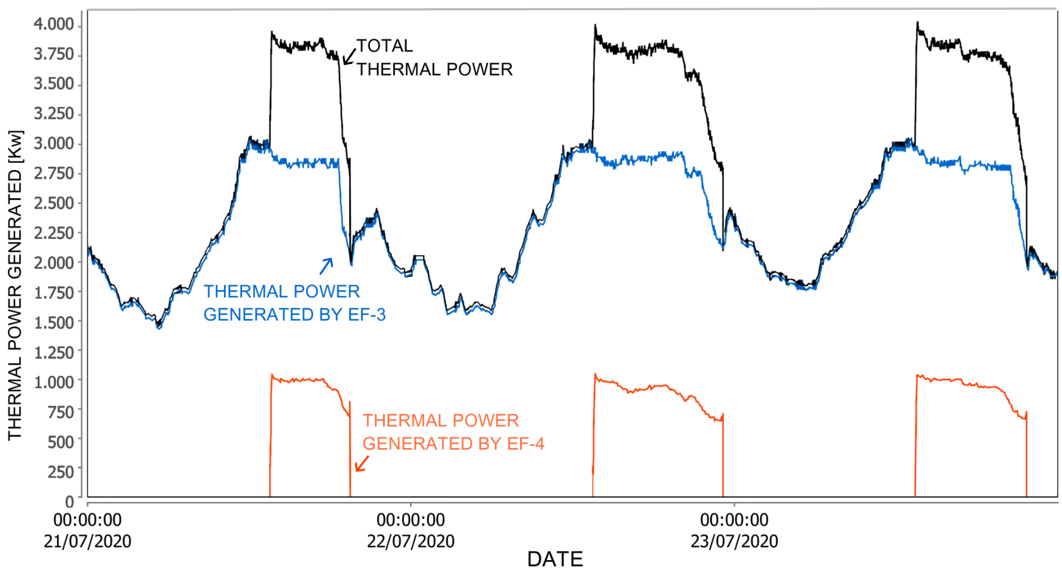

Figure 4: Cooling power generated July 21st – 23th , 2020. The reinforcement obtained by EF3 plus EF4 chiller to fit the demand can be observed. Data extracted from LON Cards installed in the 4^{th} optimization.

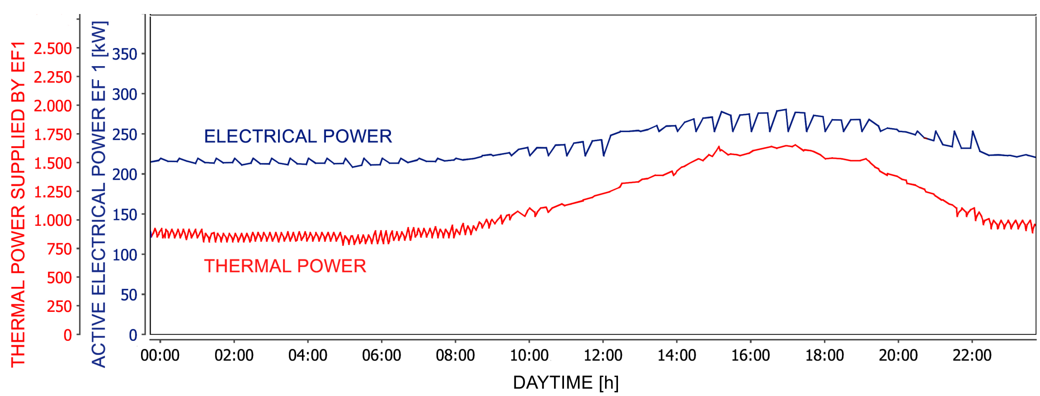

Figure 5: Thermal power generation of EF1 chiller data obtained with LON cards installed in the 5^{th} optimization.

In the 5^{th} optimization, electrical power meters were installed and integrated into the BMS system to measure the instantaneous consumption of each chiller. The impact of integrating these power meters can be visualized in Figure 5. This improvement made it possible to analyze the efficiency of each machine while in operation.

3.3. Improvements in the physical system

These improvements are made by integrating new physical systems into the existing installation.

A frequency inverter system was installed in the EF4 screw type generator in the 4^{th} optimization. The frequency inverter (AFD) can regulate the speed of the compressor with a partial load. In the EF1, EF2, and EF3, which are centrifugal chillers, AFDs could not be installed, nevertheless they still have a modulation with the refrigerant charge.

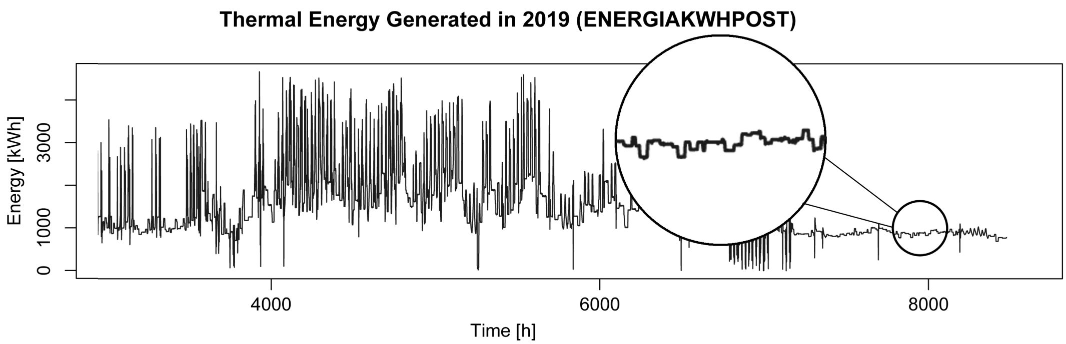

Figure 6: Detail of the thermal generation showing behavior after installation of frequency inverter in the EF4 chiller, after approx. 5000 hours, in 2019.

The operation of the screw type chiller is similar to a centrifugal chiller, in that operation ceases once it reaches the set-point. The main difference is that if the screw chiller has a frequency inverter installed in the modulation, it begins operating once this point is reached. The inverter acts directly on the power supplied to the compressor, reducing the electrical power injected and saving more energy. In addition, the minimum thermal power that the chiller can provide can be reduced before it has to be stopped. This optimization significantly reduces the number of starts and stops as shown in Figure 17, starting in the 35^{th} month (November 2019). The thermal energy graph is flattened, as can be appreciated in Figure 6, and the generation of cooling energy is adapted to the demand, especially during the periods before and after the Summer, in which EF4 is the pre-established chiller because of its power capacity.

3.4. Proposed predictive control schema

The activity of the hospital's water-cooled generators was improved by implementing a predictive model of cooling demand within the BMS control system that anticipates decisions. The incorporated control scheme is depicted in Figure 7. The prediction model was trained with real historical data from previous years.

Before implementation, the BMS controlled the starting and stopping of the chillers through the set-point temperature exclusively. With the new control scheme the prediction model foresees the maximum thermal energy demanded in the cooling system for the next

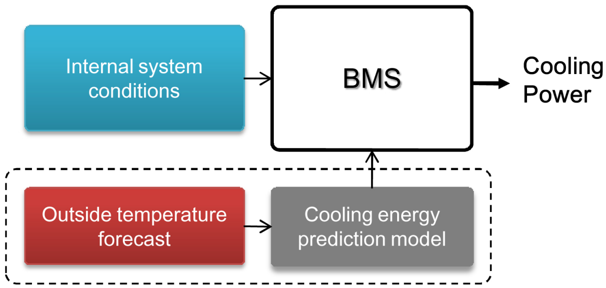

Figure 7: Control scheme. The cooling-energy prediction model communicates to the BMS the maximum thermal demand for the next day. The model reads the weather forecast conditions for the day ahead.

day. This allows decisions to be made ahead of time for the BMS, such as the maximum number of chillers necessary or scheduling the cooling towers. To do this, a script developed in R language is executed daily. This program reads the internal system and external variables, predicts the energy demand for the next 24 hours and communicates it to the BMS.

The weather conditions for the coming hours (temperature, relative humidity, etc.) are obtained from the climatological model of the Spanish National Meteorological Agency (AEMET). The data is loaded from an XML file that is updated daily [29]. It should be noted that these temperatures are predicted and subsequently can drag errors into the demand model results, as can be observed in the difference between the predicted and real temperatures in Figure 8.

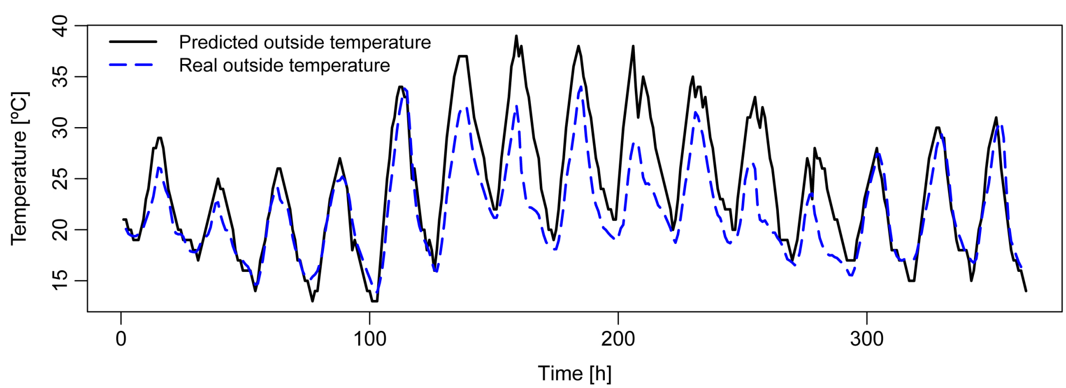

Figure 8: Outside temperature (TEXT) registered in BMS versus Forecasted temperature by AEMET model, 1st to 15th of August 2020.

Once the script reports the results of the forecasted cooling-energy demand for the day ahead, the maximum demand is predicted and communicated to the BMS by an analog signal transducer. Four possible system states have been established for the maximum demand of the next day, as depicted in Figure 9:

- STATE 1, covered by the 1 MW chiller (EF4).

- STATE 2, covered by one of the 3.5 MW chillers (EF1, EF2 orEF3).

- STATE 3, covered by a combination of the 1 MW and one of the 3.5 MW chillers(EF4 + EF1, EF2 or EF3).

- STATE 4, covered by two of the 3.5 MW chillers.

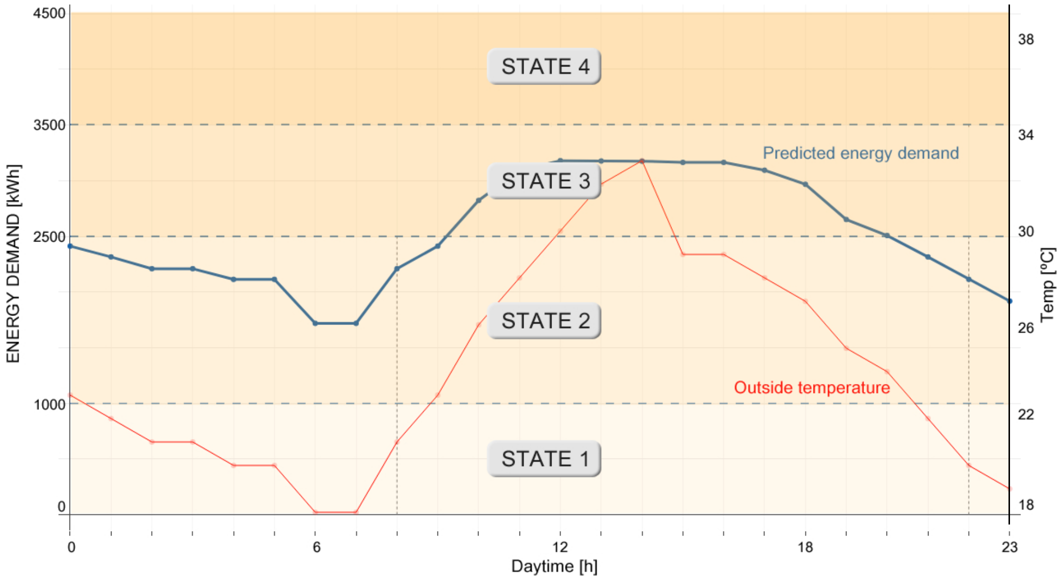

Figure 9: Energy Demand States based on the rank of the maximum energy demand for the next day.

In May 2020 the machine learning model that forecasts the maximum energy demand for the cooling system was included within the BMS. A logger software was also incorporated to register data and render graphs.

3.5. Dataset

The Knowledge Discovery in Databases (KDD) methodology was used to develop the forecasting model. During the process, three generations of models were created.

1. The first-generation models was tested in March 2019 [30].
2. A second generation of models was developed in December 2019 [31].
3. And finally, a third generation was created and included within the new control scheme in May 2020.

The following sections describe the processes for extracting information and creating the third-generation models, which are very similar to the previous generations.

3.5.1. Data extraction

The data was extracted from the *BMS Sauter NovaPro Open* software since the beginning of this study in 2017 (Table 1).

Table 1: Control system variables of the BMS

Short name	Description
EF1 to EF4	EF1 to EF4 - Status
TIMP	Cold-ring-drive temperature[ºC]
TEXT	Exterior temperature of power installation building [ºC]
TCONSIG	Calculated set-point of the regulation for cold-production drive [ºC]
RH	Relative humidity [%]
TENEF1 to 4	Water temperature at the inlet of the EF1 to EF4 [ºC]
TSALEF1 to 4	Water temperature at the outlet of the EF1 to EF4 [ºC]

3.5.2. *Data preprocessing*

The preprocessing entailed, among others, the following actions (common in KDD processes):

1. Data Cleaning: Filling in or dropping missing values.
2. Data Integration: Averaging measurements by hour.
3. Data Transformation: The generated cooling energy (ENERGYKWHPOST) was calculated from the combination of other control system variables. This new feature was converted to energy [kWh] rather than instantaneous power [kW]. To smooth the noise, the final target, ENE_GAUSSFILT11, was obtained by filtering ENERGYKWHPOST with a Gaussian function that used a window size of 11 hours.
4. Feature and Model Selection: `GAparsimony` R package was used to simultaneously select the most important attributes and algorithm's parameters. The objective was to obtain parsimonious models with high accuracy and low complexity (more robust against noise and process changes, and easier to maintain).

3.5.3. *Final dataset*

In order to improve the first two generations and, according to [9, 11, 15], two new variables were included in the third generation: time and relative humidity (RH).

The final selection of attributes was:

Table 2: Attributes selected for testing the forecast models.

Variable	Description
ENE_GAUSSFILT11	Target feature
time	Time of measurement
month	Month of measurement
day_of_week	Day of the week
Is_holiday	Boolean variable for holiday
TIMP	Instant impulsion temperature
TEXT	Instant outside temperature
TMEAN	Average daily temperature
TMAX	Maximum daily temperature
TMIN	Minimum daily temperature
RH	Relative humidity [%]

The set-point temperature of the cooling water (TCONSIG) was not selected because it was linearly related to the outside temperature (TEXT), as can be observed in Figure 3.

3.6. 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 is recommended, given that they 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 this study, training and selecting the best parsimonious models was 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 them, choosing those with the least complexity.

In this study, 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). The final selected model was a weighted blending of the two best models obtained with ANN and SVR. 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. All the experiments were implemented with the `GAparsimony` [32] package developed in the `R` language.

3.7. 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 λ_g^i chromosome formed by the combination of two vectors P and Q:

$$\begin{aligned} SVR(\lambda_g^i) &= [P(cost,\ gamma,\ epsilon),\ Q] \\ ANN(\lambda_g^i) &= [P(size,\ decay,\ num\_epochs),\ Q] \end{aligned} \tag{1}$$

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 $q_j \geq 0.5$.

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` [24, 25, 26, 27].

The optimization process with `GAparsimony` genetic algorithms 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 according to their 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 $RMSE_{val}$.

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 is created randomly, but with 90% of the characteristics of the individuals selected. This allows the search for models to start with models that have a high number of entries.

Finally, the mutation is 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 is performed. In the case of vector Q, the probability of changing 0 to 1 was set at 10% in order to facilitate the reduction of the number of attributes in subsequent generations.

3.8. Energy demand model

To calculate the 3^{rd} generation of models, data was collected from April 2018 to December 2019. Prior data was removed due the relevant improvements that went into effect at that time, as shown in Figure 10. The training dataset corresponded to the period between January 2018 and February 2019. The validation database was defined to the even weeks between March 2019 and December 2019, and the test database to the odd weeks of the same period.

`GAparsimony` was used to choose the best parsimonious models trained with SVR and ANN algorithms, by adjusting the algorithm's parameters, selecting the most relevant features and choosing the best parsimonious solution. Table 3 shows the best SVR and ANN models: $RMSE_{val}$ and $RMSE_{tst}$, selected features with the percentage of appearance in the elitist models during the last generations, model complexity (NFS), and the parameters of the best-tuned algorithm.

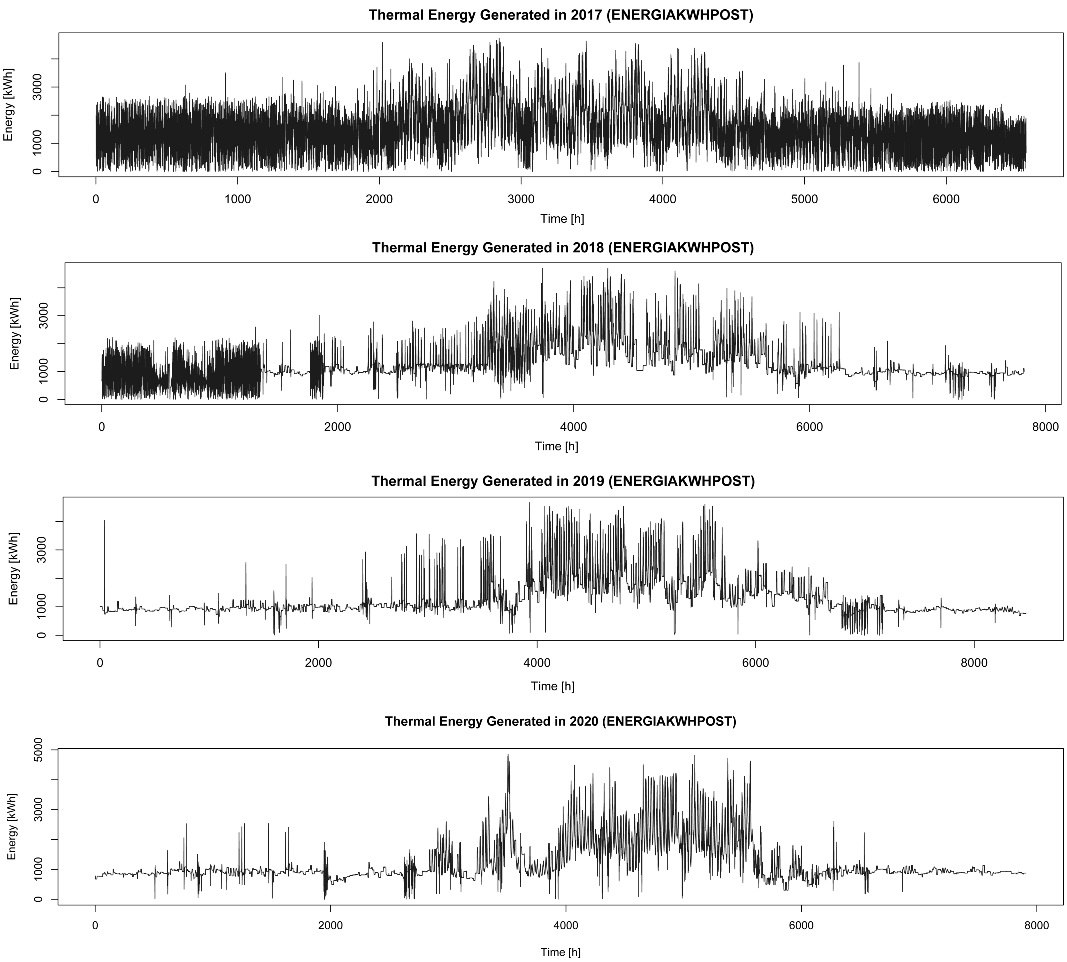

Figure 10: Evolution of energy generated in the cooling-water facility (ENERGYKWHPOST) from 2017 to 2020.

Table 3: Best models with RMSE errors, features used and their percentage of appearance in the group of elite models within the optimization process with genetic algorithms, complexity, generation and parameters.

	SVR		ANN	
$RMSE_{val}$	**222.95**		226.04	
$RMSE_{tst}$	**256.08**		264.02	
	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	
Parameters	expcost	-0.014	size	33.95
	gamma	0.331	decay	200.04
	epsilon	0.048	maxit	708.13

SVR Model: The best SVR model was obtained with 5 features: time (*time*); month (*month*); and the outside (*TEXT*), averaged (*TMEAN*), and maximum (*TMAX*) daily temperature. Figure 11 shows the evolution for the most elite population of the best `GAparsimony` iteration for SVR model. White and gray box-plots represent the $RMSE_{val}$ and $RMSE_{tst}$ evolutions respectively. Continuous and dash-dotted lines indicate the best individual error for validation and test of each population. The gray area covers the range of features of most elite individuals, and the dashed line indicates the minimum number of features N_{FS} (right axis).

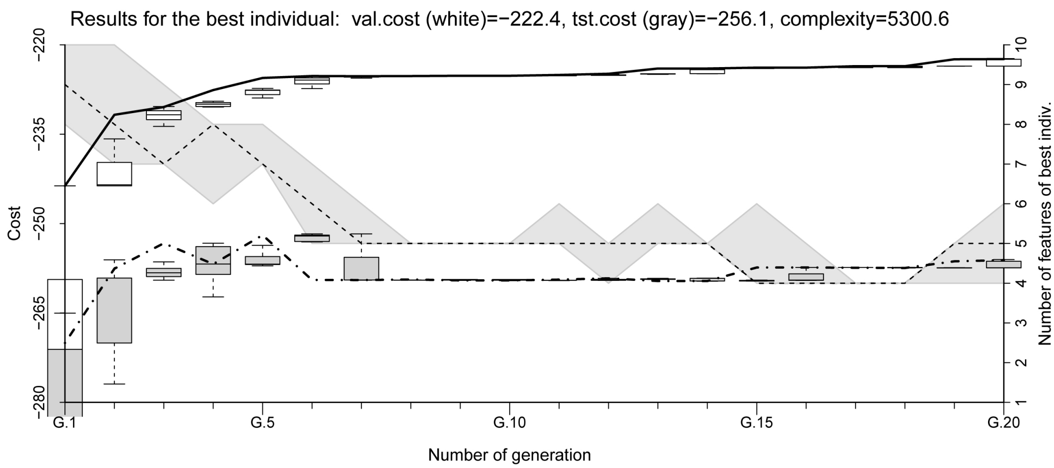

Figure 11: Evolution of the errors of the most elite solutions for SVR algorithm.

ANN Model: The best ANN model converged in 2 generations with 6 features: time (*time*); month (*month*); and the temperatures of the ring (*TIMP*), outside (*TEXT*), averaged (*TMEAN*), and maximum daily (*TMAX*). ANN errors were slightly higher than those of the SVR model. Figure 12 displays the evolution of the ANN model.

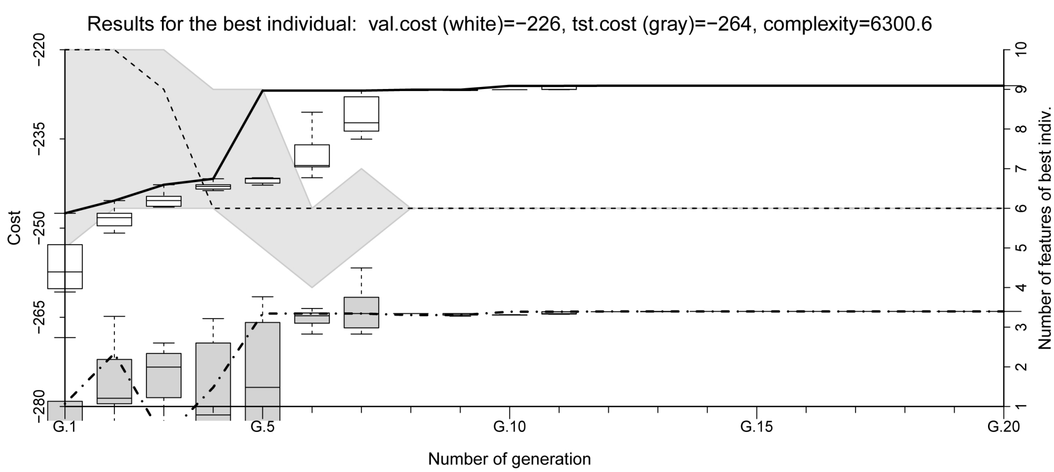

Figure 12: Evolution of the errors of the most elite solutions for ANN algorithm.

HYBRID Model: The best SVR and ANN models were combined to obtain a *blending model* by weighting the predictions as follows,

$$Hybrid\_Model = (w1 * SVR + w2 * ANN)/2 \tag{2}$$

The weights were optimized to reduce the $RMSE_{val}$ and obtain this solution:

$$Hybrid\_Model = (1.651619 * SVR + 0.348381 * ANN)/2 \quad (3)$$

Table 4 shows the improvement of $RMSE_{val}$ and $RMSE_{tst}$ of the hybrid model versus single models. The error rate was slightly better in the ensemble model than the best single model (SVR). This hybrid model reduces errors compared to the second generation hybrid model. And in addition, it was further simplified since the previous one was composed of 3 models (SVR, ANN, XGBoost), and it is less complex because it uses less features. Therefore, this model is easier to maintain and more robust against noise.

Table 4: Ensemble validation and test errors versus single models.

	SVR	ANN	HYBRID
$RMSE_{val}$	221.95	226.04	**220.86**
$RMSE_{tst}$	256.08	264.02	**253.20**
complexity	5	6	5+6

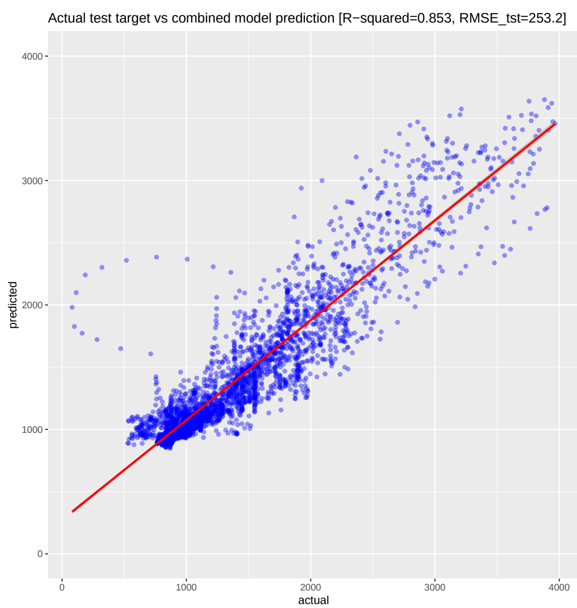

Figure 13: Combined prediction for the Hybrid model.

As described in Section 3.2, the quality of data logged improved significantly as a result of the installation of the LON cards. The graphs in Figures 14 compare the registered thermal energy generated (data obtained from the LON cards) to the energy demand predicted by the ensemble model (which uses the data predicted by AEMET as input). Furthermore, these graphs show the influence of the outside temperature on the demand in the dotted line: its impact increases when the outside temperature TEXT is more extreme (during July and August). Although the data that were finally used for modeling were extracted from the BMS sensors instead of that from LON cards, the prediction obtained is very close to the registered demand.

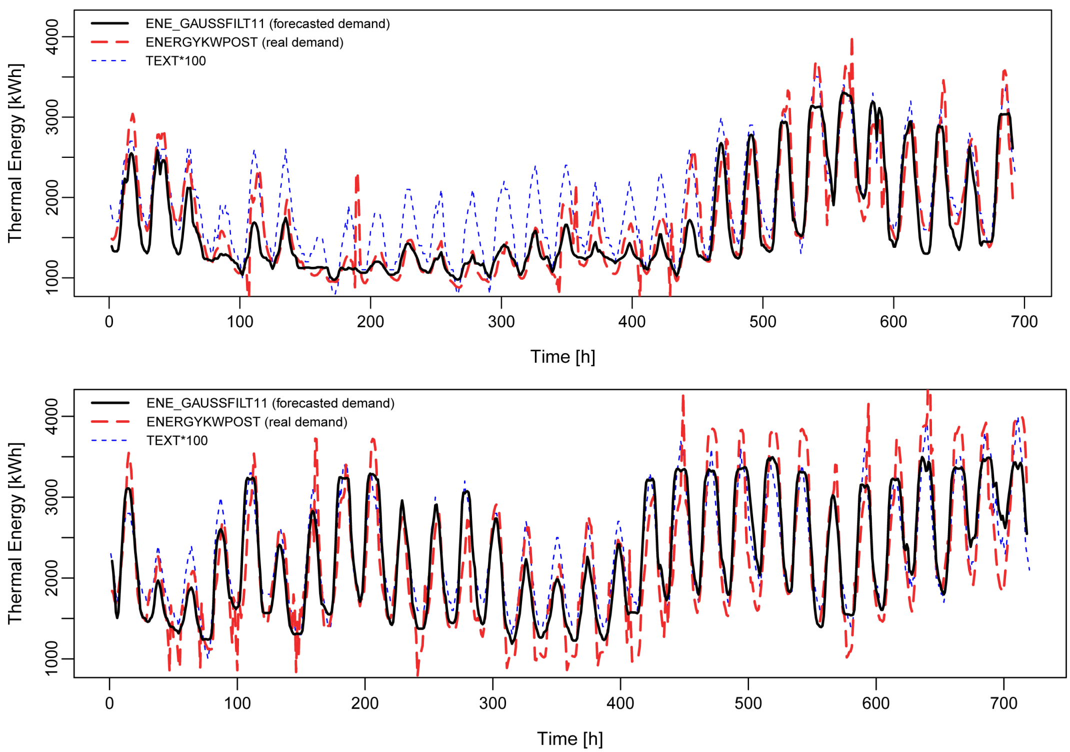

Figure 14: Forecasted energy demand (ENE_GAUSSFILT11) versus thermal energy generated (ENERGYKWPOST) obtained from LON cards, June and July of 2020.

4. Results

4.1. Energy savings

The electrical energy logged made it possible to compare the annual energy savings obtained before and after the model was implemented. The data was extracted from the readings of the electrical power meter located in the hospital's power plant. This meter also measured other electrical consumption from heating and lightning. Nevertheless this data is valid for this study because the cooling generation system is the most energy-intensive installation in the building, while heating or lightning have a stable demand over time. Moreover, electrical meters were installed in every chiller, but they were not integrated until the 5^{th} optimization of the system.

The method of cooling degree days (CDD) was used to normalize the consumptions for a more adequate comparison. For this calculation the temperature of 17 ºC was selected as the base temperature, Figure 15 shows the variation of energy that increases once this base temperature is exceeded. Below this temperature, the cooling system stabilizes at an almost constant power of less than 1 MW (approximately 800 kW). That is the reason why in the winter season the EF4 screw chiller is able to supply enough energy to the cooling system. The meteorological data for the study was obtained from the official La

Rioja Government weather station [33], with data validated in accordance with the Spanish UNE 500540 standard.

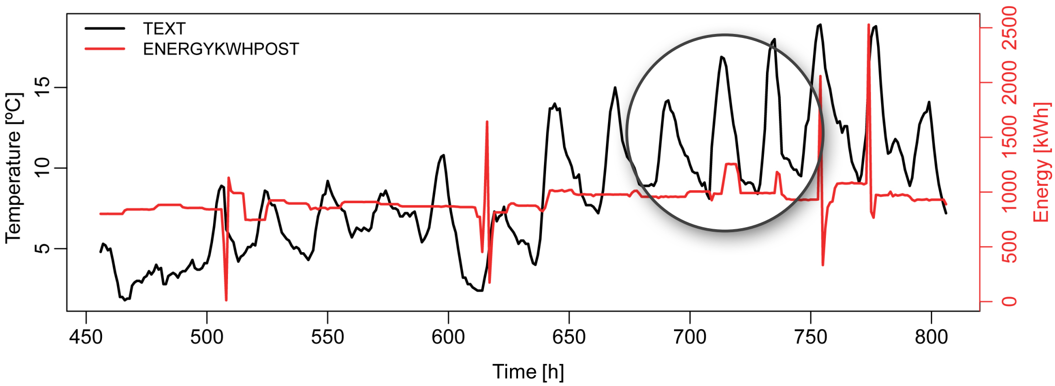

Figure 15: Outside temperature of 17 ºC was chosen as the base temperature for estimating CDD since it requires additional energy to hold the cooling system. The energy peak generated can be observed inside the circle.

Table 5 shows the normalization of the annual electricity consumption in the building from the year 2016 (prior to the study) to the year 2020. To perform this normalization, the average degree days $CDD17$ in the interval $2016 - 2020$ (which was 587.3 degree days), multiplied by each annual value of $Energy/CDD17$ provides the normalized energy for each year.

Table 5: Normalized energy per year [kWh] previous and over the course of the study, based on $CDD17$.

Year	**CDD17**	**Energy [kWh]**	**Energy/CDD17**	**Normalized E. [kWh]**
2016	590	5,968,990	10,119	5,942,682
2017	597	6,258,184	10,483	6,156,502
2018	576	6,124,609	10,629	6,242,594
2019	620	5,864,247	9,455	**5,553,164**
2020	553	6,400,075	11,569	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 is $0.0988521 EUR/kWh$. Depending on whether the comparison is between the year 2016, prior to the 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 $38,504 EUR$ and $59,641 EUR$ per year, as shown in Table 6.

Table 6: Estimated savings thanks to application of the methodology.

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

In order to evaluate the behavior of the plant during this year (2020, the year when the predictive system was implemented), the monthly evolution should be analyzed in the months of higher degree days $CDD17$, which are July and August, where the comparison of normalized data is more descriptive, Table 7.

Table 7: Normalized Energy of most demanding months [kWh], monthly $CDD17$.

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**

Higher electrical consumption (+22.3%, +122,716 EUR) was observed during the year 2020. The reason is that plant operations were atypical since all areas of the hospital equipped with Air Handling Units (AHU), see Figure 16, were configured to avoid air recirculation and increase ventilation flow to prevent the spread of COVID-19 [34].

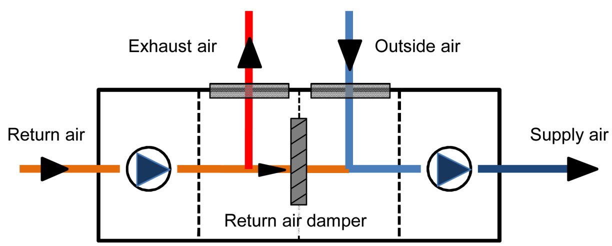

Figure 16: AHU internal scheme. The SARS-CoV-2 virus general recommendation is to avoid central recirculation by closing the recirculation dampers either using the BMS or manually.

4.2. *Measurement of the number of starts per chiller*

The number of starts should be reduced as much as possible, especially in chillers not equipped with inverter systems. An excessive number of starts can damage internal parts, and every start generates an electrical peak that may affect surrounding installations.

The results of measurements of the number of starts per chiller are shown in Table 8 summarizing measurements by year and chiller. If the year 2017 is compared with 2019, the total number of starts decreased by 82.7%.

Table 8: Number of starts per chiller from 2017 to 2020 (* Chiller EF4 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%
2019	91	177	192	427	**887**	77,4%

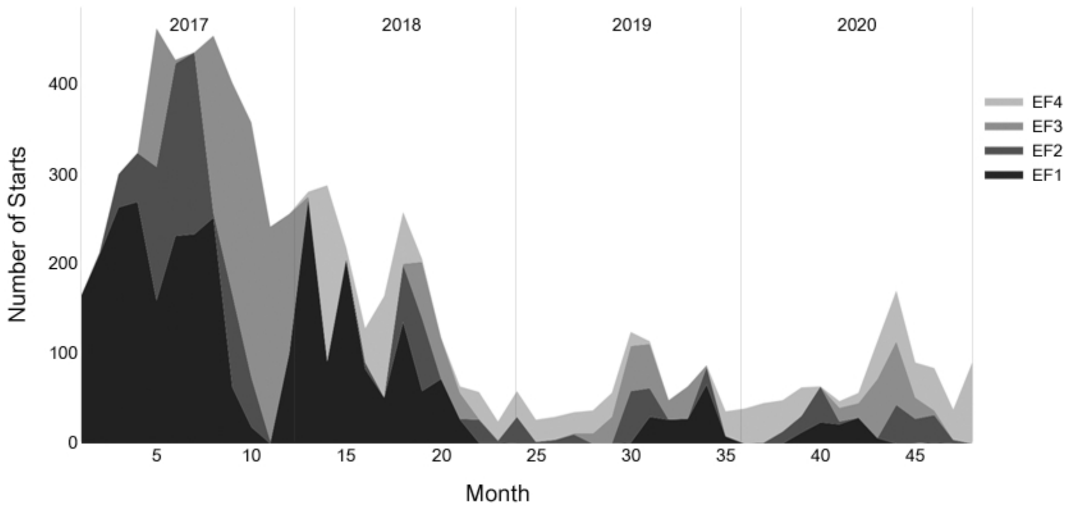

Figure 17: Number of starts per chiller from 2017 to 2020. The diagram shows the notable reduction in the number of chiller starts thanks to optimizations made to the system during the process.

In order to be able to compare the evolution of the number of starts during the current year 2020, Table 9 indicates the total sum of starts of all the chillers per month. As can be observed, in the year 2020 there have been more starts due to the night programming that the EF4 chiller activated. This action was carried out in a controlled manner and improves energy efficiency since this chiller gives its maximum Energy Efficiency Ratio (EER) in loads within that range.

Table 9: Total number of chiller starts for each month and each year. The number of starts since the model was implemented 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**

5. Conclusions

This methodology reworked the hospital's cooling system and solved problems that had plagued the system in the past. Optimizing the control system 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. Furthermore, implementing the BMS of a cooling-demand prediction model allowed plant operations and performance to be optimized. Thanks to this system, the maximum cooling energy demand for the next day can be forecasted, and therefore, the BMS system can establish the number of chillers necessary. In addition, this model provides a daily schedule for plant maintenance and a self-generated report in R script.

To develop the blended prediction model, the `GAparsimony` methodology facilitated optimization. In the final models, the XGBoost model was discarded because its high level of resource consumption was not compensated for by the improvements it offered. In the models that make up the final ensemble (SVR and ANN), it should be noted that the common features influencing the predictions were: time, month, outdoor temperature, average temperature and maximum daily temperature. 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 (the model would not be overtrained). On the other hand, it was observed that the external model that implements the weather-forecast information (outdoor temperature, average temperature and maximum daily temperature) can drag errors into the prediction results.

Improvements in the data acquisition system enhanced the accuracy of the data from the chillers. However, since this improvement occurred at the end of the optimization process, the last models made did not include the more accurate data. These acquisition systems have improved communication with the chillers, allowing the maximum working power to be fine-tuned, which contributes to expanding cooling power, and reducing the electrical demand of the chillers by improving modulation. What's more, the addition of electrical meters in each chiller would further enrich our knowledge of plant efficiency.

Regarding the improvements made to the physical system, it is worth highlighting the significant improvement in the modulation of the screw chiller after an inverter system was installed, which allowed the plant to work at maximum energy efficiency and significantly reduced the number of starts and electrical demand. In the last year of the study, the total number of starts was increased deliberately due to the implementation of time schedules for higher efficiency.

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

The optimization of the plant and the KDD process are long-term procedures; the present work was conducted over the course of more than 3 years. In order to apply this methodology in similar hospitals, it would be necessary to compile a database period of at least one year. Hence, it is exceedingly difficult to implement this methodology from scratch in a short period of time.

In terms of future ways to further improve the cooling plant within the same line of research, the forecasting model should be revisited using the data obtained from the LON cards installed in the chillers after a period of at least one year, and once the special measures implemented due to COVID-19 are lifted. The energy efficiency of the plant should be analyzed by studying the data provided by the electrical energy meters installed in the chillers. Such research would identify the most efficient conditions for each cooler. 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.

Acknowledgements

We are greatly indebted to *Banco Santander* for the APPI17/04, REGI2018/41 and REGI2018/43 fellowships. This study used the Beronia cluster (Universidad de La Rioja), which is supported by FEDER-MINECO grant number UNLR-094E-2C-225.

References

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. doi:https://doi.org/10.1016/j.egypro.2019.01.883; innovative Solutions for Energy Transitions.

7. IDAE, , Fenercom, . Guía de ahorro y eficiencia energética en hospitales. *Fenercom* 2010;:329URL: https://www.fenercom.com/wp-content/uploads/2010/11/Guia-de-Ahorro-y-Eficiencia-Energetica-en-Hospitales-fenercom-2010.pdf. doi:-.

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. doi:https://doi.org/10.1016/j.jobe.2018.01.011.

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. doi:https://doi.org/10.1016/j.apenergy.2018.07.085.

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. doi:https://doi.org/10.1016/j.jobe.2020.101504.

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. doi:https://doi.org/10.1016/j.enbuild.2016.11.009.

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. doi:https://doi.org/10.1016/j.enbuild.2021.110740.

13. Saeedi, M., Moradi, M., Hosseini, M., Emamifar, A., Ghadimi, N.. Robust optimization based optimal chiller loading under cooling demand uncertainty. *Applied Thermal Engineering* 2019;148:1081 – 1091. URL: http://www.sciencedirect.com/science/article/pii/S1359431118353547. doi:https://doi.org/10.1016/j.applthermaleng.2018.11.122.

14. Alamin, Y.I., Álvarez, J.D., [del Mar Castilla], M., Ruano, A.. An artificial neural network (ann) model to predict the electric load profile for an hvac system**this work has been funded by the grant from the spanish ministry of economy and competitiveness (enerpro dpi 2014-56364-c2-1-r). yaser i. alamin is a fellow of the marhaba, an erasmus mundus lot 3 project. josé domingo Álvarez is a fellow of the spanish 'ramón y cajal' contract program, co-financed by the european social fund. antonio ruano acknowledges the support of fct through idmec, under laeta grant uid/ems/50022/2013. *IFAC-PapersOnLine* 2018;51(10):26 – 31. URL: http://www.sciencedirect.com/science/article/pii/S2405896318305482. doi:https://doi.org/10.1016/j.ifacol.2018.06.231; 3rd IFAC Conference on Embedded Systems, Computational Intelligence and Telematics in Control CESCIT 2018.

15. Ahmad, M.W., Mourshed, M., Rezgui, Y.. Trees vs neurons: Comparison between random forest and ann for high-resolution prediction of building energy consumption. *Energy and Buildings* 2017;147:77 – 89. URL: http://www.sciencedirect.com/science/article/pii/S0378778816313937. doi:https://doi.org/10.1016/j.enbuild.2017.04.038.

16. Jain, R.K., Smith, K.M., Culligan, P.J., Taylor, J.E.. Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy. *Applied Energy* 2014;123:168 – 178. URL: http://www.sciencedirect.com/science/article/pii/S0306261914002013. doi:https://doi.org/10.1016/j.apenergy.2014.02.057.

17. 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. doi:10.1016/j.enbuild.2015.05.056.

18. 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. doi:https://doi.org/10.1016/j.enbuild.2016.03.043.

19. 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. doi:https://doi.org/10.1016/j.enbuild.2017.02.012.

20. 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. doi:10.3390/en11030631.

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

22. 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.

23. 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.

24. 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.

25. 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. doi:10.1007/978-3-319-19644-2_52.

26. 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. doi:http://dx.doi.org/10.1016/j.engstruct.2014.06.047.

27. 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. doi:http://dx.doi.org/10.1016/j.enconman.2015.02.086.

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

29. AEMET, . 7 days ahead weather forecasting for logroño (la rioja), spain. 2020. URL: http://www.aemet.es/xml/municipios/localidad_26089.xml.

30. 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. ISBN 978-3-030-29859-3; 2019:181–192.

31. Dulce-Chamorro, E., Javier Martinez-de Pison, F.. Parsimonious Modelling for Estimating Hospital Cooling Demand to Improve Energy Efficiency. *Logic Journal of the IGPL* 2021;URL: https://academic.oup.com/jigpal/advance-article-abstract/doi/10.1093/jigpal/jzab008/6139196. doi:10.1093/jigpal/jzab008; jzab008.

32. 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.

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

34. 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. doi:https://doi.org/10.1016/j.envint.2020.105832.