

Rule-based system to detect energy efficiency anomalies in smart buildings, a data mining approach

Manuel Peña *, Félix Biscarri, Juan Ignacio Guerrero, Iñigo Monedero, Carlos León

Escuela Politécnica Superior, Universidad de Sevilla, Departamento de Tecnología Electrónica, C/Virgen de África 7, 41011 Seville, Spain

A B S T R A C T

The rapidly growing world energy use already has concerns over the exhaustion of energy resources and heavy environmental impacts. As a result of these concerns, a trend of green and smart cities has been increasing. To respond to this increasing trend of smart cities with buildings every time more complex, in this paper we have proposed a new method to solve energy inefficiencies detection problem in smart buildings. This solution is based on a rule-based system developed through data mining techniques and applying the knowledge of energy efficiency experts. A set of useful energy efficiency indicators is also proposed to detect anomalies. The data mining system is developed through the knowledge extracted by a full set of building sensors. So, the results of this process provide a set of rules that are used as a part of a decision support system for the optimisation of energy consumption and the detection of anomalies in smart buildings.

1. Introduction

Nowadays, the growing world energy use has already raised concerns over supply difficulties (Belussi, & Danza, 2012), exhaustion of energy resources and heavy environmental impacts (Pérez-Lombard, Ortiz, & Pout, 2008; Rashid et al., 2011). Worldwide trends in energy consumption in residential, public services and commercial sectors with a total consumption of 56.9% (EIA, 2013), which constitutes 24% of the world's CO₂ emission (Day, Jones, & Turton, 2013), are basically focused on improving the Energy Efficiency (EE), based on renewable energy use and systems that minimises the energy consumption.

Currently, there are many international organisations sponsoring initiatives to encourage the development of technologies in this area (de Alegría Mancisidor, Díaz de Basurto Uraga, Martínez de Alegría Mancisidor, & Ruiz de Arbulo López, 2009; EEA 225, 2013); International Energy Agency (IEA), U.S. Department of Energy – Energy Information Administration (DOE/EIA), Energy European Commission (EEC), Organism for Economic Co-operation and Development (OECD) have taken steps to facilitate and improve developments in this area with directives like Directive 2012/27/EU on energy efficiency (EU) or Executive Order (E.O.)

13514; Federal Leadership in Environmental, Energy, and Economic Performance; October, 2009(DOE/EIA).

The updated technology and cost parameters have possibly contributed to lower electricity consumption. Therefore, EE is rapidly increasing in world energy due to the exhaustion of energy resources and heavy environmental impacts. Substantial opportunities to improve energy efficiency have been expressed by the International Energy Agency (IEA) (DOE/EIA-0383, 2013).

The increase of smart cities (Lazaroiu, & Roscia, 2012; Yamagata, & Seya, 2013), smart and green building (Chou, Chua, Ho, & Ooi, 2004; Djevic, & Dimitrijevic, 2009; Singh, & Tiwari, 2010; GhaffarianHoseini, Dahlan, Berardi, GhaffarianHoseini, Makaremi, et al., 2013; Vadiee, & Martin, 2013) this is becoming the beginning of the future building construction trend (GhaffarianHoseini et al., 2013). The advantages of these buildings include a high level of comfort, high power efficiency, and environmental friendliness (Dounis, & Caraiscos, 2009; Day et al., 2013; Martin, Hernandez, & Valmaseda, 2015). Besides, the renewable energy sources utilised as the main power supply of the smart buildings (Omer, 2008; Xia, Zhu, & Lin, 2008; Tsiamitros et al., 2014; Gul & Patidar, 2015) help to reach an environmental friendliness, high-level comfort and power efficiency in buildings. These requirements used to be reached managing an effective control in the building.

At the present, current research trends in EE are focused on various scopes. In the last decade, there was considerable research concerning optimisation of EE in buildings (Weng, & Agarwal, 2012). Some trends go from building investment decision

* Corresponding author. Tel.: +34669381722.

E-mail addresses: manuelpm@dte.us.es (M. Peña), fbiscarri@us.es (F. Biscarri), juaguealo@us.es (J.I. Guerrero), imonedero@us.es (I. Monedero), cleon@us.es (C. León).

(Malatji, Zhang, & Xia, 2013) to building design (Pacheco, Ordóñez, & Martínez, 2012), but the main trend is focused on new technologies applied to smart buildings. In Noailly (2012), the impact of environmental innovation technologies to improve EE in general buildings are studied, and (Andrews, & Krogmann, 2009) specifically for commercial buildings. Another trend in EE is based on optimising the energy consumption keeping a high level of well-being for the occupants. In Ma, Qin, Salisbury, and Xu (2012), a demand reduction based on an economic slant is proposed. Also, in Marinakis, Doukas, Karakosta, and Psarras (2013) the studies are focused on tertiary sector and in large buildings through a simulation system (Colmenar-Santos, Terán de Lober, Borge-Diez, & Castro-Gil, 2013).

The increasing sophisticated Building Automation System (BAS) (Marinakis, Karakosta, Doukas, Androulaki, & Psarras, 2013) has become the cornerstone of modern intelligent buildings. Integrating energy supply and demand factors, often known as Demand-Side Management (DSM) has become an important energy efficiency policy concept (Azadeh, Saberi, Ghaderi, Gitiforouz, & Ebrahimipour, 2008). Much of this potential can be captured through policies for the advancement of the implementation of energy management and control systems (EMCS).

In EMCS (Haberl, Sparks, & Culp, 1996; Altwies, & Nemet, 2013), one of the most complex problems is the optimal energy management according to real-time environmental variables of the building. The EMCS manages the best energy policies for building in real time with the main goal of keeping the high-level comfort with the minimum power consumption in different operating conditions (Kolokotsa, Kalaitzakis, Antonidakis, & Stavrakakis, 2002; Dounis, & Caraiscos, 2009, Pang, Wetter, Bhattacharya, & Haves, 2012). Thus, at present, research groups works try to solve this problem from different points of view and with different techniques (Nguyen, & Aiello, 2013). Some studies go from basic actions to improve the EE in commercial buildings (Escrivá-Escrivá, 2011) to more sophisticated energy management based on a set of rules integrated in a SCADA based on predictive controller through a cost function system (Figueiredo, & Sá da Costa, 2012). In (Diakaki, Grigoroudis, & Kolokotsa, 2008; Diakaki et al., 2010), a multi-objective model based on a decision maker to improve the EE in buildings through multi-criteria decision analysis techniques are applied. Another proposal is considered in Oldewurtel et al. (2012), with a predictive control model joined with a weather forecast that advise how to control HVAC in the smart building. More complex solutions are based on a multi-agent system (MAS). In Doukas, Nychtis, and Psarras (2009), the energy consumption and the environment friendless criteria have optimised through a multi-objective model with the purpose of the investment. In Klein et al. (2012), the model is based on a Markov Decision Model (MDM). And in Yang, and Wan, (2012), an MAS based on building occupant behaviours is employed for an optimisation of energy consumption and maintaining high-level comfort.

Other research trends propose EE solutions through knowledge discovery from the data. A data analysis and classification to predict the energy consumption of the building is proposed in (Li, Bowers, & Schnier, 2010). In (Kim, Stumpf, & Kim, 2011; Capozzoli, Lauro, & Khan, 2015), an interesting analysis of energy efficient building design through DM techniques is provided. Also, in Yu, Haghghat, Fung, and Zhou (2012), the authors proposed discovering the knowledge based on mining associations between building operational data. However, the analysis of energy consumption at the use stage is very important, and it is obvious that the construction characteristics of the buildings strongly affect consumption during the life cycle of the building. Furthermore, the way in which the facilities are used during the use stage is also very important when determining the efficiency of the building (Yu, Haghghat, Fung, Morofsky, & Yoshino, 2011; Domínguez

et al., 2013). Within the current research trends in EE, other similar work done on EE anomalies detection in smart buildings using data mining techniques were not found. Most related papers found were (Li et al., 2010; Figueiredo & Sá da Costa, 2012; Oldewurtel et al., 2012; Yu et al., 2012).

Rule-based systems are traditionally useful in energy efficiency, implementing energy management and control systems, even in the new scenarios provided by smart grids (Liu et al., 2010). There are many references about rule-based system which support the energy efficiency optimisation process in different aspects, and some of these references usually are combined with other techniques to improve the accuracy and efficiency of the optimisation. SEMERGY.net (Fenz et al., 2014) proposed a web-based optimisation environment, which supports users in decision-making regarding energy-efficient building designs; this system is based on a rule-based system which uses an ontology of linked building product data. (Wang et al., 2012) proposed a rule-based algorithm for elevator group control, which focuses on design of dispatching rules for energy saving. Brooks and Barooah (2014) studied two control-oriented methods of improving energy efficiency in commercial buildings: rule-based, feedback controller that uses real-time occupancy measurements, and model predictive control (MPC). Multi-Agent Systems (MAS) are a technology related with RBS, because each agent implements a rule-based system. In this sense, Villar et al. (2009) proposed a real application of an MAS for coordinating the electrical heaters of a building in order to consider the comfort level and the energy efficiency in the building. Hurtado et al. (2014) proposed an agent-based control strategy for the operation of buildings, implementing a fuzzy rule based decision making strategy to monitor and control de energy flows in a building. Finally, Ryu et al. (2012) proposed a rule model where rules could be dynamically updated according to user feedback and applied to various situations, in order to maintain the energy efficiency and user comfort level.

The current work described in this paper is immersed in Project KnoHoEM, adhered to FP7 (FP7-285229 KnoHoEM). The main contents of Project KnoHoEM are based on an intelligent energy management solution relying on three main elements: knowledge modelling techniques and knowledge base, extended validation on various demonstration objects for the enhancement of the knowledge database, as well as on hardware implementation of the energy management system. Some of the most important objectives of the Project KnoHoEM are a functional energy-oriented building model, a specific building behavioural model completed by a building-specific ontology and a data-mining procedures for real-time detailed energy consumption analysis.

The aim of the present paper is to optimise the EE in a smart building, reducing energy consumption, keeping a high degree of comfort, and being environmentally friendly through a Data Mining (DM) approach. This work is focused on detecting EE anomalies in smart buildings. For this purpose, a system of Energy Efficiency Indicators (EIs) and a Rule-based System is presented. These systems are carried out after a DM process to extract the knowledge hidden in historical building data and EE experts. This knowledge is used in EIs and a set of "best policies" that help experts to achieve the primary goal. Subsequently, the anomalies detected are corrected through a set of policies that regulate these anomalous behaviours, as shown in Fig. 1. Nowadays the prototype is in test phase, and thanks to the progress made in this field and the results offered, a future goal is to integrate this framework in commercial an EE software called Eugene, owner of ISOTROL Company.

The paper is organised as follows: in Section 2, the smart building description and data sources are presented. The data sources description is divided into three subsections: indoor sensors, outdoor sensors and finally, energy analysers. In Section 3, a data mining-based DSS to increase the energy efficiency at the

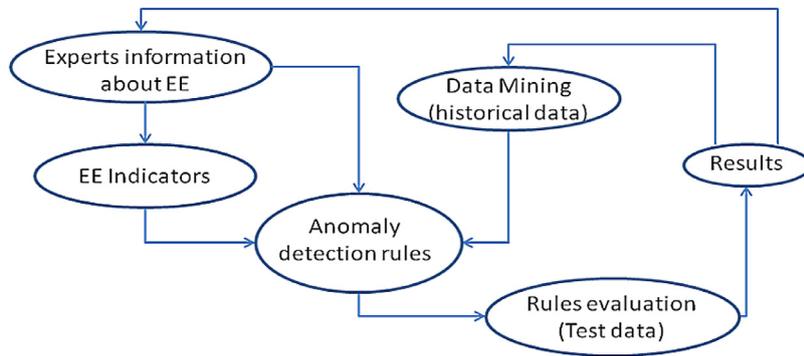


Fig. 1. Data Mining Energy Efficiency framework.

building is proposed. It is divided into three subsections: the data pre-processing (cleaning, filtering and transformation), the Energy Efficiency Indicators description, and finally, the Rule-based System to detect anomalies in the smart building. In Section 4, the results of the framework obtained from the smart building are presented. Section 5, has concluding remarks and future research lines are proposed.

In the related work, a case of study is proposed. The main goal of this case of study is how to detect EE anomalies in smart building automatically and with high effectiveness. Nowadays, most of the building needs to be analysed by EE experts to on one hand, analyse and understand the EE building behaviour, and on other hand, try to detect EE anomalies.

In this case of study a system which is able to analyse the EE behaviours of a building, detect all types of EE inefficiencies and check how of effective is proposed through a DM approach.

2. The smart building description and data collector

For anomaly detection it is a basics necessity to know the energy consumption behaviours in the building. Thus, in this section a description of BlueNet building and the technology employed in the smart building is described. Besides, a description of data sources used in the anomaly detection is presented. This knowledge is basic for understanding the efficiency of the energy consumption in the BlueNet building and therefore for the anomaly detection.

The BlueNet building is situated in the south of Spain (Seville, 37°24'29.97"N, 6°00'18.63"O). This building is used as an office building focusing on the main activity from 8:00 to 17:30, although sometimes there are different schedules. The building has continuous exterior and interior windows all around, 1.5 × 1.5 m. It is therefore a building with a glass Coolite cyan curtain wall. The windows are hinged to the outside of it and have shutters with double glazing. The frames are white aluminium interior doors and windows, while windows are aluminium exterior gray, and doors are solid white lacquered wood. The roof is flat.

The BlueNet building is composed of a luminary system Light Master of Philips (Luminary with ballast HF-R TD 318) connected by a Dali bus. The Dali bus is centralised by a LRC5141 controller that connects to a centralised server through a Long/ModBus gateway (IntesisBox). This server is responsible for sharing data and applying function policies.

In BlueNet, HVAC is based on a VRV System. This system is composed by a set of indoor units (Daikin FXSQ-M7V1B) and a set of outdoor units (Daikin RXYQ-MY1B VRV II inverter with heat pump). All units are connected by a DIII-Net. These connections are centralised in a DMS504B51 Daikin Lonworks Interface that is

connected in turn to a centralised server through a Long/ModBus gateway (IntesisBox) using a communication protocol I3E.

The power consumption in the BlueNet building mainly comes from IT equipment (PC's, Servers, media, etc.). The major constituent of power consumption used to be consumed in office hours with the exception of some work groups that are supported 24 hours a day.

Usually smart buildings are composed of several units of data metering. The BlueNet building covers three main areas of metering; indoor sensors, outdoor sensors, and energy analysers. The indoor sensors are composed of 16 ZigBee sensors per plant that collect indoors variables, ZigBee motes are centralised by a master ZigBee used as a gateway that assembles all the measurements and shares this data into the centralised server. Outdoor sensors are composed by a set of sensors placed on the BlueNet building to extract environmental variables and share all these measurements in a centralised server. Finally, energy analysers share all power consumption variables of the building (HVAC, air mixers, splits, lighting, power and other consumptions). These analysers are centralised through a Modbus master device that collects and transmits all the measurements to the centralised server that share all the BlueNet building information. Besides the information on these three areas, the information adhered to the calendar of working days is included. All this data is shared as quite valuable information to understand building consumption patterns and apply EE policies.

2.1. Data sources description

2.1.1. Indoor sensors

These sets of ZigBee sensors are placed in strategic locations in the BlueNet building and coordinated by a master sensor. This master sensor manages the other sensors sending the data metering to a device that aims to collect sensor information and sends it to the BlueNet centralised server. Indoor sensors are programmed to send a meter reading to the master sensor every 30 seconds. These sensors collect the following data:

- Temperature:* ZigBee sensors read the temperature in every room in the building, normally there are at least 2 or 3 sensors in every room to ensure obtaining the correct temperature.
- Humidity:* ZigBee sensors read the relative humidity in every room of the BlueNet building, it is very important to calculate the actual temperature in every room.
- Lux:* ZigBee sensors check the actual lighting system obtaining the amount of luxes in every room of the building.
- Presence:* These sensors obtain the data presence of every person that is placed in the BlueNet building identifying every person with a unique id and obtaining information about the time that every person is in the rooms.

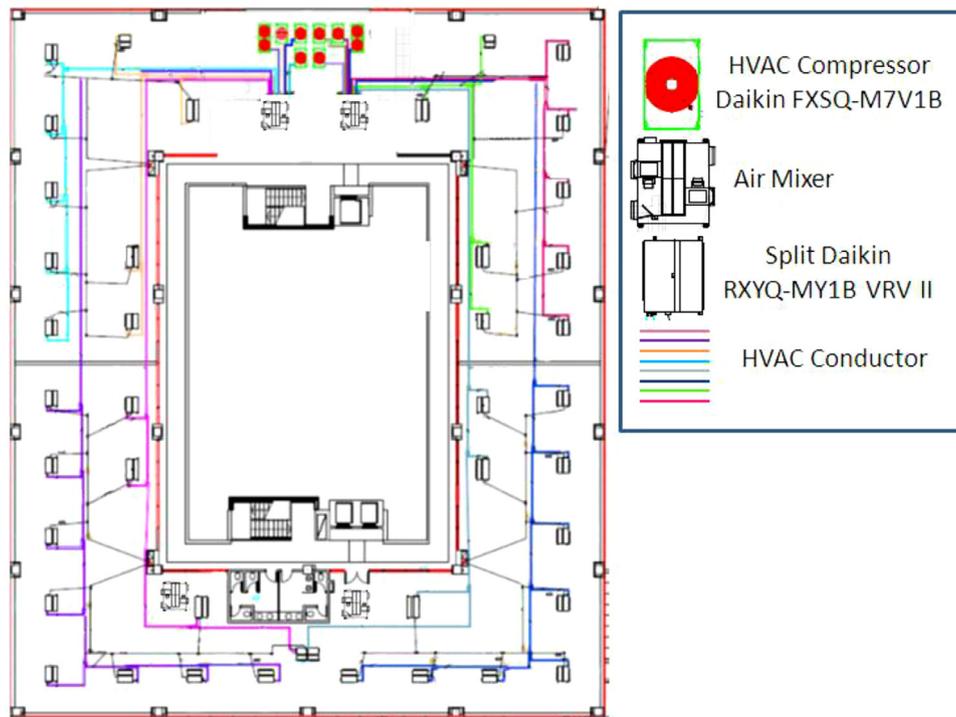


Fig. 2. Plane of BlueNet building architecture and HVAC distribution.

2.1.2. Outdoor sensors

Besides, there is a set of sensors that take data metering of the BlueNet environment. These sensors are responsible for collecting environmental data. Outdoor sensors are programmed to select a data metering every 10 min. The main measurements obtained from our work are:

- Temperature:** Sensors take the environment temperature every 10 min and also, maximum, minimum and mean temperature of each day.
- Humidity:** Outdoor sensors take the environmental humidity and the amount of rain that has fallen to calculate the actual of environmental temperature.
- Sunshine:** Sensors obtained calculate the amount of sunshine that irradiate to the building every day.
- Wind:** Outdoor sensors also obtain the amount of wind and its direction in the building environment.

2.1.3. Energy analysers

Energy analysers are placed in an energy distribution panel inside of BlueNet building, they are connected to a Master Modbus Device (PLC) responsible for managing the energy analysers. These analysers are connected to the four areas of energy consumption in the building (HVAC, Lights, Power and other consumptions). These analysers are programmed to send a meter reading of four areas every 5 min to the Master Modbus Device. The principal features of each area are:

(a) HVAC:

This area had been the primary area in this study due to the known influence in the BlueNet EE (Hou, Lian, Yao, & Yuan, 2006; Wu & Clements-Croome, 2007). Besides, the HVAC area is the largest area in energy consumption, and also the most controllable. HVAC management has the major margin for EE improvement with 143876.7 MWh in 509 days that represent 40.11% of the total energy consumption in the building. In this case, the different components in the HVAC field could be identified (compressors, parts, air mixer, Fig. 2), and the consumption of every item calculated as indicated:

- *AP_COMPRESSOR*: 3* *AP_CLI_FASE3* (kW)
- *AP_SPLITS*: *AP_CLI_FASE1* - *AP_CLI_FASE3* (kW)
- *AP_AIR_MIXERS*: *AP_CLI_FASE2* - *AP_CLI_FASE3* (kW)

After analysing every HVAC item, the following conclusions have been extracted; compressors consume the bulk of the HVAC energy and deliver a basal consumption of 1.7 kWh (as shown in Fig. 4). Below this, the active power compressor is considered switched off. In contrast, splits and air mixers work without a basal consumption.

(b) Lights:

Despite lighting only having a 19.03% of total energy consumption in BlueNet building, by 68281.7 MWh during a period of 509 days, it is difficult to reduce this consumption (lighting is a necessary building consumption and it associated to presences in rooms). There are new policies to utilise in future project extension with a dimmer and Lux sensor implementations that are able to widely reduce lighting consumption. The study of this field could provide interesting information in EE such as occupation patterns, anomalies consumptions or group behaviour.

(c) Other:

This area is the least in energy consumption by only 26170 MWh in total energy consumption in 509 days analysed besides having a strong fixed consumption which is difficult to manage. Thus, with only a 7.3% of total energy consumption this is the least relevant field in EE.

(d) Power:

This area is the second most important for this study with 120368.7 MWh in total consumption although had the second largest consumption with a 33.56% in the period of 509 days. The consumption manner is difficult to improve due to computers used in offices difficult to reduce that kind of consumption. Likewise, other policy taken from power consumption has been interesting to improve EE.

Thus, in the study the four areas of energy consumption were analysed. More importantly the BlueNet building EE were mainly focused on HVAC management, as well as, 143876.7 MWh of the

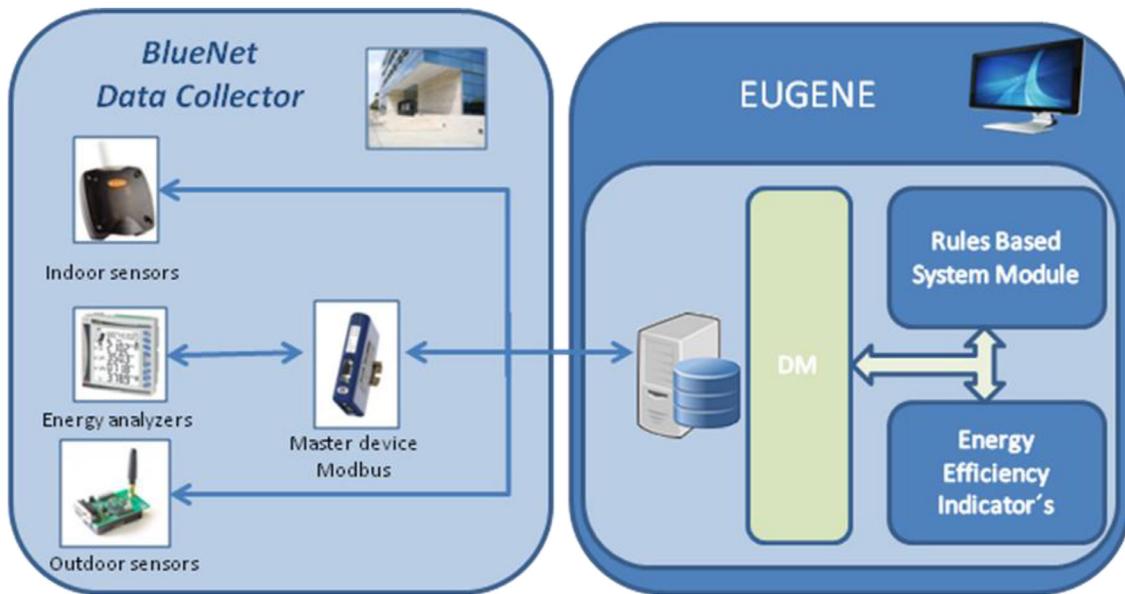


Fig. 3. BlueNet building software prototype.

energy building consumption is used in HVAC to provide sustainable and comfortable environment. Major HVAC consumption proceeds specifically from the compressor motor engine consumption with a 46.38%. Every area was carefully examined for detecting every possible EE anomaly that provides improved energy management.

On one hand, this first study aims to quantify the energy consumption in the BlueNet building and on the other hand, how the energy has been consumed (e.g. Regime changes in compressor, operational changes in the compressor, lighting patterns). Our aims in this analysis were to study the influences of these variables in the EE, as well as, quantify the importance of each one of these variables in every EE area (HVAC, lights, power and others). Hence, a set of EE Indicators has been deployed from these variables, and these EE Indicators help us to establish the system.

3. The data mining-based decision support system

Knowledge Discovery in Databases (KDD) refers to the overall data mining process of discovering useful knowledge from data. Before the DM process, the data were analysed, filtered, sorted and selected (Han, & Kamber, 2001; Maimon, & Rokach, 2010). In data pre-processing phase, BlueNet contains different data sources: energy consumption, indoor environmental data and environmental outdoor data. All of them have their own frequency of timing. Thus, in this pre-process the frequency is synchronized with a period of 10 min in data register. The time interval for these data sources are covered between January of 2011 and January of 2013. Once the data were standardised, the data sample contains 5087485 records corresponding to the set of variables of BlueNet building, where only the more relevant variables were selected for use in the study. (Described in Appendix A).

Once all the data was collected and analysed in BlueNet DB, a system based on the SCADA of BlueNet building were developed through DM techniques. The objective of this system was to extract knowledge hidden in BlueNet building data, and also, develop a Rule-based System Module and EELs based on this knowledge. This knowledge was used in real time with the BlueNet building data metering to detect EE anomalies. This knowledge became in the cornerstone of this work. Furthermore, Rule-based System Module and EELs are nowadays in test phase, and it will be connected to an

EE tool owner of ISOTROL company called Eugene. These modules provide all the knowledge extracted from BlueNet building data using Analytics and Big Data techniques, and improves the results of EE tool. This architecture is depicted in Fig. 3.

The data mining techniques used were based on statistic techniques and the study of the patterns extracted from BlueNet data. As a result of the patterns detected and the results of statistics, the knowledge in specifying EELs and establishing the set-points parameters for rules anomaly detection are provided. Also, these techniques have been applied mainly to assess the correct setting in EELs and rules. These data mining techniques were based on the study and valuation of energy behaviour graphics. All these processes were applied and explained in every EEL and rules that are described in Section 3.2 and 3.3.

For this work, we used SPSS Modeller (by SPSS Inc. an IBM company) a powerful and extended tool in the data mining area. This tool includes IA libraries using to create the anomaly detection system through DM techniques.

3.1. Data pre-processing

In the first step of pre-processing, the 5087,485 records of the data set were analysed and 636,310 missing data corresponding to null values of the distributed metering units were found. This problem used to happen in a distributed sensor's environment, so these data were filtered and deleted in the data sample. Once the outliers' exclusion of data had been completed the next requirement is guaranteed a basic data quality:

- Detect and fix the correct percentage of consumption dispersion.
- A valid sample of data requires at least one data each 30 min.

Under pre-processing, conditions were filtered 636310 records corresponding to outliers and after applying the basic rules of quality, 34 records were filtered. Finally, the data sample was simplified from 5087,485 records to 4451,141 records.

Once pre-processing was carried out through ETL methods, all the attributes described in Section 2.1 were analysed to infer the correlation between indoor and outdoor sensor data and energy consumption behaviours through PCA analysis techniques. PCA is a statistical procedure concerned with elucidating the covariance

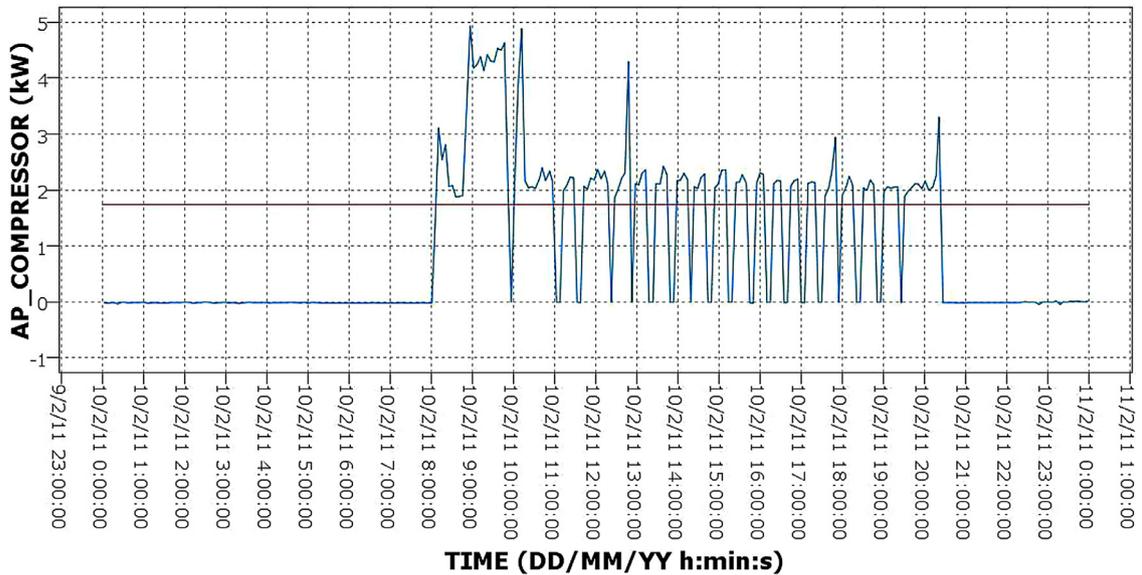


Fig. 4. Operational changes of HVAC compressor (OCC).

Table 1
Attribute selection.

Variable	Unit
AEMET_AT	°C
AP_CLI_FASE1	Kw
AP_CLI_FASE2	Kw
AP_CLI_FASE3	Kw
AP_LIG_FASE1	Kw
AP_LIG_FASE2	Kw
AP_LIG_FASE3	Kw
AP_POW_FASE1	Kw
AP_POW_FASE2	Kw
AP_POW_FASE3	Kw
PRESENCE_ID_EXT	ld
PRESENCE_ID_EXT	ld
ZIGBEE_AT	°C

structure of a set of variables. In particular, it allows us to identify the principal directions in which the data varies. As result of this analysis, an attribute selection were selected for data classification and prediction (Han, & Kamber, 2001; Maimon, & Rokach, 2010). The aim of this attribute selection was: study the influence of each attribute, as well as, to quantify the importance of every variable in EE. Before applying attribute selection, these variables were selected among all those that would be used as input parameters of the analysis (described in Appendix A). After filtered variables in data sample according to some similar studies (Hou et al., 2006) more important variables were selected after applying a Principal Components Analysis (PCA) study. As result, variables shown in Table 1 were selected for the EE study. Much of these variables were studied and selected in other similar researches studies (Wu, & Clements-Croome, 2007).

Once relevant variables were selected, the behaviour of historical data of these variables yield the following results; 40.11% of building consumption was spent in HVAC, 33.56% of energy consumption in power, a 19.03% spent on lighting and another 7.3% in other derivate consumption with a total accumulated energy consumption in the BlueNet building of 358697.1 MWh during the analysed period.

3.2. Energy efficiency indicators (EELs)

Based on experts' knowledge and historical data, a set of Energy Efficiency Indicators (EELs) was built in every relevant area

in BlueNet building. These EELs were in charge of the continuous evaluation of BlueNet building EE behaviour. Through Data Mining techniques, all the relevant information hidden in the data were extracted, and every EEL could provide the knowledge to detect EE anomalies.

(a) Operational changes of HVAC compressor (OCC) indicator:

A large number of daily On-Off operations were considered anomalous or inefficient. A large number of OCC could cause one of the most EE wastes or inefficiencies, greatly increasing power consumption. This EEL counts the number of daily On-Off operations in the compressor.

An example of a high value of this EEL is a high consumption variance, as shown in Fig. 4. This behaviour in HVAC compressors could warn about a possible anomaly in HVAC management (the HVAC is badly dimensioned for this room, the HVAC is too powerful for this room or a possible breakdown in temperature sensor is done). On the other hand, a compressor with a high OCC is prone to break with a lower lifetime period.

(b) Number of operational regime changes of the HVAC compressor (ORCC) indicator:

It considers abnormal or inefficient a large number of operational regime changes in the HVAC compressor (ORCC). An ORCC is defined as a change in the compressor power greater than 0.5 kW with respect to the previous measurement (10 min). These parameters have been specified based on the results of DM techniques under the consensus of Isotrol HVAC experts. This could mean that the HVAC machine is not properly calibrated (there is too much power for its require use) or the temperature of the room is forcing to the HVAC compressor to constantly change its operating mode, as shown in Fig. 5, where there is a high variation in the HVAC compressor consumption.

This EEL counts the number of daily ORCC periods, and during how many daily min the ORCC periods had been produced. Besides this, EEL indicates a possible improvement in HVAC management. This improvement consists in softening the HVAC consumption curve by applying EE policies in the Rule-based System module.

(c) Switch ON HVAC compressor and abnormal changes in indoor temperature (SONCCIT) indicator:

It is considered abnormal or inefficient if a compressor is running and the room temperature decreases sharply (winter) or increases sharply (summer). This could be due to a sudden leakage of high temperature (winter) or a sudden influx of heat

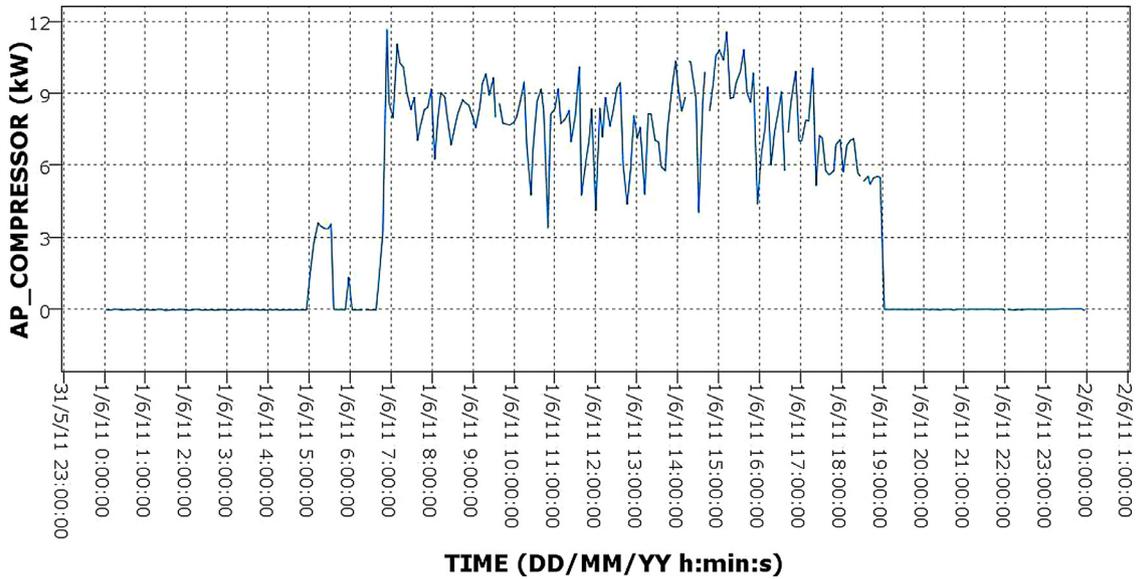


Fig. 5. Operational changes of HVAC compressor (ORCC).

(summer) in the room, which counteracts the effect of the HVAC system (i.e. Opened windows or doors). A SONCCIT is defined when the HVAC system is turned on (the HVAC compressor consumption is upper to 1.7 kWh) and had been produced a change in indoor temperature greater than or equal to 1 °C between samples (10 min).

This EEI counts the total minutes per day with a SONCCIT anomaly. Besides, this EEI averages the daily active power in the HVAC compressor during the compressor anomaly, and the number of periods per day in which a compressor-on anomaly had been observed (an anomalous period is specified by the aggregation of consecutive anomalous data points).

(d) *Switch OFF HVAC compressor and abnormal changes in indoor temperature (SOFFCCIT) indicator:*

It is considered abnormal or inefficient if the room temperature rises sharply (winter) whilst the compressor is not running. It could be due to a heat source (electric heater), replacing the HVAC system, it will indicate an inefficient use of power energy. From historical data, it is considered anomalous an increase of 1 °C between samples (10 min) during the winter period, or a decrease during the summer months.

This EEI counts the total minutes per day with a SOFFCCIT anomaly. Besides, this EEI averages the daily active power in the HVAC compressor, and the number of periods per day in which a compressor-off anomaly had been observed (like in SONCCIT, an anomalous period is defined by the adding of consecutive anomalous data points).

(e) *No persons in bluenet building and switch ON HVAC compressors (NPSONC) indicator:*

An anomalous function of compressor is considered when absence of any people or the lights are switched off and the compressor is switch on, NPSONC. It could be considered that the air conditioner is switched on due to ignorance, considering that there is no building occupancy without the lights switched on. This indicator can be restricted to non-daylight hours (night) or non-work days. To infer the presence of persons in the building were considered attributes, such as, *PRESENCE_ID_ENT*, *PRESENCE_ID_EXI*, *AP_LIG_FASE1*, *AP_LIG_FASE2*, *AP_LIG_FASE3*, *AP_POW_FASE1*, *AP_POW_FASE2* and *AP_POW_FASE3* for instance using the active power consumption (*AP_POW_FASE1*, *AP_POW_FASE2* and *AP_POW_FASE3*) to infer the basal consumption in the building and the consumption conducted by the presence of people.

This EEI counts the total daily minutes in which a NPSONC has been detected. The number of periods per day with a NPSONC anomaly (an anomalous period is defined by the aggregation of consecutive anomalous data points) and averages active power per day in the HVAC compressor during the anomaly.

3.3. Rule-based system to detect anomalies at the energy building consumption

Based on the knowledge extracted from EE experts, EELs and historical data, a Rule-based System to detect anomalies at BlueNet building was developed. The aim of this system is to predict inefficient consumption behaviours and provide correction policies. These objectives help achieve the main goals proposed; a high-level comfort with the minimum power consumption and environmental friendliness.

The Rule-based System was carried out through DM techniques. These techniques were applied on EELs and historical data behaviour. As a result of this study, a set of variables were defined, and after applying PCA analysis to determine those variables which have a high correlation with the BlueNet consumption patterns, only the more important variables were selected (*Appendix B*). Once the variable selection was carried out, Rule-based System was developed using the knowledge extracted.

The rule-base system module described in Fig. 3 was deployed on BlueNet servers. These servers are in charge of managing and control the sensor and consumption data supplied by master sensors and master Modbus device. These servers are as well in charge of monitoring and manage all the energy consumption in BlueNet building (EMCS). The rule-base system module was developed as an independent module in the BlueNet servers. The rule-base module is connected with the monitoring system and it supplies a set of methods, and subsequently, the outputs of these methods are shown in the front-end application. The rule-base system module was developed in Java, this module establishes a real time connection with sensor master devices and Modbus master devices. The rule-base system use time series techniques to analyse real time streaming channels from the sensors master devices and Modbus master device.

Once that rule-base system module is connected with the BlueNet real time data, a set of ETL methods are applied to aggregate and standardise the data to solve possible conflicts in the

data, analyse and apply the set of rules (R1–R7) described in this section. Finally, these results are shown in front end application.

Summarising, these rules help in a synthesised way, the decision-making process for experts in huge amounts of data. Secondly, this module allows experts to be advised with the best way to use energy consumption. Besides, this system allows the possibility to highlight anomalies leading to an inefficient use of the building.

As a conclusion, this system highlights the knowledge extracted from EEs, the company experts in EE area, and historical data from the BlueNet building. The rules applied to energy management system are as follows:

(a) *The use of compressor while the outside temperature is between temperature comfort range (R1):*

This rule controls the HVAC compressor use when the external temperature at the BlueNet kept inside the comfort range that advises “Regulation of Technical Systems in Buildings” (RTSB – Royal Decree 1826/2009) like best practice in EE. This range changes according to the season, in winter comfort ranges keep to the minimum of 21°C (AEMET_AT) and maximum of 26°C (AEMET_AT) in summer with a condition of relative humidity between 30% and 70%, and an interior wind velocity of 0.15 m/s inside of the area occupied. Thus, this rule notifies when the external temperature is more than 21°C and active power in HVAC compressor is greater than 1.7 kW (AP_COMPRESSOR). Similarly, in summer it notifies when the external temperature is less than 26°C and active power in HVAC compressor is greater than 1.7 kW (AP_COMPRESSOR). Winter and summer season are defined automatically by the calendar. In spring and autumn seasons the system takes the more convenient rule depending on inside sensor temperature and the set point temperature. Rule R1 is as follows:

If
 In winter – $AP_COMPRESSOR \geq 1.7 \text{ kW AND } AEMET_AT > 21^\circ\text{C}$
 In summer – $AP_COMPRESSOR \geq 1.7 \text{ kW AND } AEMET_AT < 26^\circ\text{C}$
 then
 HVAC excessive power used or problems with temperature probe.

(b) *Compressor ON and no temperature change (R2):*

This rule detects when the HVAC compressor is working (AP_COMPRESSOR) and there is no change in the temperature inside BlueNet building during a period of 30 min ($\Delta ZIGBEE_AT$). This rule reports a possible HVAC compressor malfunction caused by a compressor, if it doesn't work well or a breakdown. Rule R2 is as follows:

If
 $AP_COMPRESSOR \geq 1.7 \text{ kW AND } \Delta ZIGBEE_AT < 0.2^\circ\text{C}$
 then
 HVAC badly dimensioned or break down compressor or energy leak

(c) *Anomaly detected by a plurality of operational changes in the HVAC compressor (R3):*

This rule reports when an excessive number of operational changes in the HVAC compressor have happened during a day. Therefore, through DM techniques and an optimal function has been parameterised to these features; the number of daily On-Off in HVAC compressor (OCC_STOPPOINTS) and the daily average of active power in HVAC compressor (AP_COMPRESSOR_MEAN). These set up parameter values were obtained through a correlation and distribution analysis of all these parameters with a success rate of the 85% of the bad dimensioned and energy leaks cases. Rule R3 is as follows:

If
 $OCC_STOPPOINTS \geq 5 \text{ AND } AP_COMPRESSOR_MEN > 2 \text{ kW}$
 then
 HVAC bad dimensioned or excessive temperature changes due to energy leaks.

(d) *Anomaly detected by a large number of operational regime changes in HVAC compressor (R4):*

This rule detects when the HVAC compressor has an in-ordinate number of operational regime changes during a day. Once a study is carried out through DM techniques, the following features have been parameterised with optimal values. These features are three: First, the total number of ORCC periods (ORCC_PERIODNUMBER), a series of consecutive ORCCs constitutes a period. Second, the total daily minutes in which an ORCC has been detected (ORCC_MINUTES). And finally, the daily average of active power of HVAC compressor (AP_COMPRESSOR_MEAN). These set up parameter values were obtained through a correlation and distribution analysis of all these parameters with a success rate of the 85% of the bad dimensioned and energy leaks cases. Rule R4 is as follows:

If
 $ORCC_PERIOD_NUMBER \geq 50 \text{ AND } ORCC_MINUTES \geq 50 \text{ min AND } AP_COMPRESSOR_MEAN > 0.5 \text{ kW}$
 then
 HVAC bad dimensioned or excessive temperature changes due to energy leaks.

(e) *Anomaly detected by SONCCIT (R5):*

This rule detects anomalies caused by a SONCCIT with the aim to report every anomaly of a SONCCIT type that has been detected. The main feature for this rule is (SONCCIT_PERIODNUMBER) that it counts the total number of periods with a SONCCIT anomaly. A series of consecutive SONCCITs constitutes a period. Rule R5 is as follows:

If
 $SONCCIT_PERIODNUMBER \geq 1$
 then
 Bad thermal isolation or HVAC compressor breaks down.

A case with SONCCIT anomaly detection is shown in Fig. 6 where there are two anomalies detected at 20:40 P.M. and 21:30 P.M. As shown, SONCCIT detection alerts when among at least 10 min there is a drop of more than one degree centigrade in temperature while the HVAC compressor is going to heat the building in winter. These anomalies were due to an accidentally window opened, this conclusion was carried out after analyse the sensor data of temperature sensors. In summer, this situation is different, the anomaly is detected when the temperature goes up more than one degree centigrade for at least 10 min at least and HVAC compressor is running.

(f) *Anomaly detected by SOFFCCIT (R6):*

This rule detects the anomalies caused by a SOFFCCIT with the aim of reporting every anomaly of this type. The principal feature for this rule is counting the total number of periods with a SOFFCCIT anomaly (SOFFCCIT_PERIODNUMBER), a series of consecutive SOFFCCITs constitutes a period. Rule R6 is as follows:

If
 $SOFFCCIT_PERIODNUMBER \geq 1$
 then
 External source of heat in winter.

A situation with SOFFCCIT anomaly detection is shown in Fig. 7. In this graph, there are three anomalies detected at 00:00 A.M, 1:20 A.M. and 10:30A.M. SOFFCCIT detection alerts if for at least 10 min there is a rise of more than a degree centigrade in temperature while HVAC compressor is not running to heat the building in winter as shown in Fig. 7.

(g) *Anomaly detected by NPSONC (R7):*

The aim of this rule is to detect any inefficiency caused by the use of an HVAC compressor while there is nobody in the

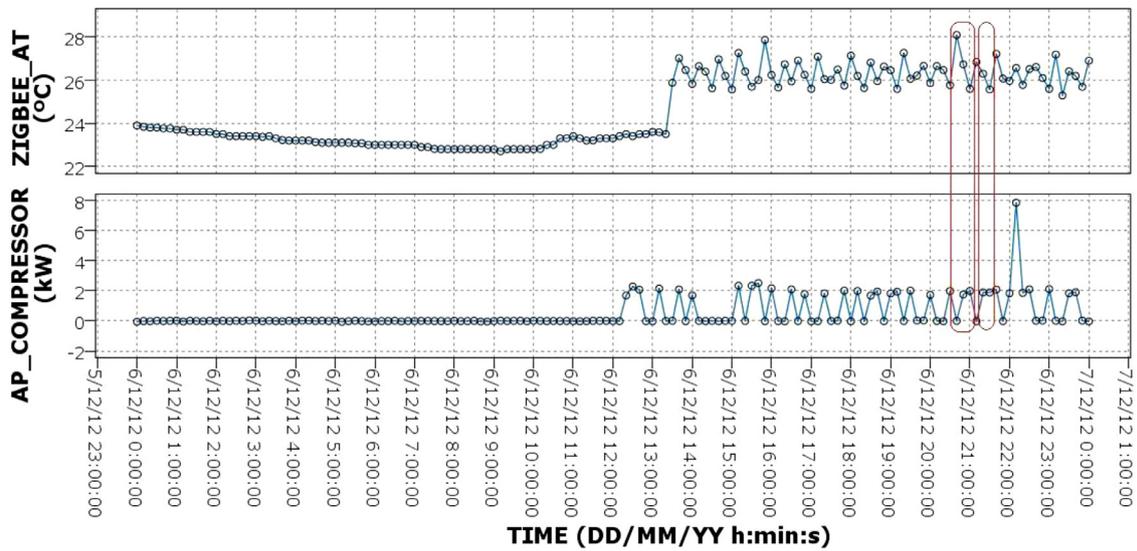


Fig. 6. SONCCIT anomaly detection.

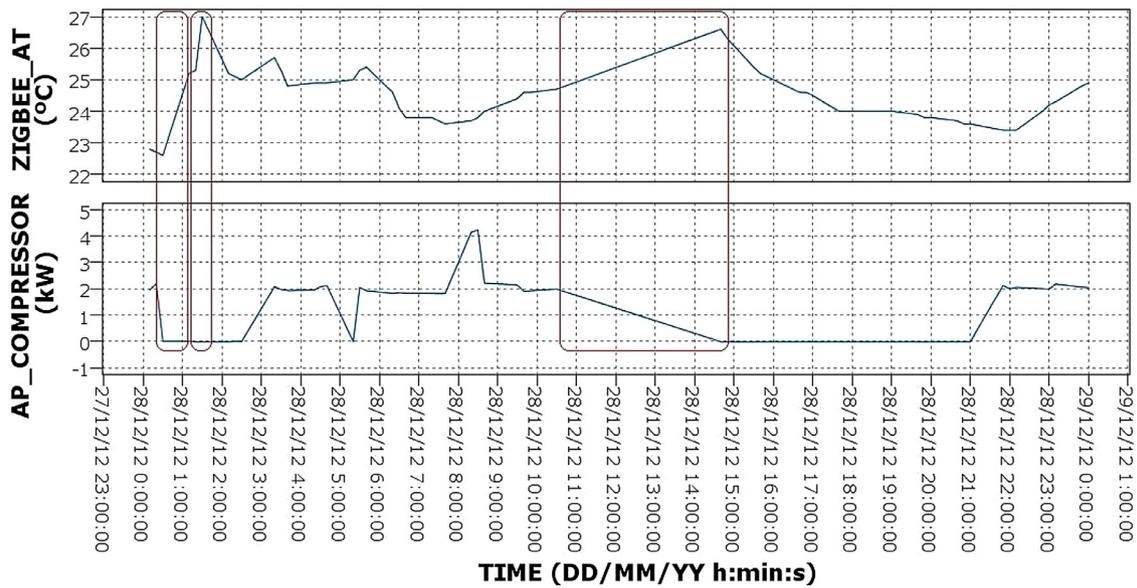


Fig. 7. SOFFCCIT anomaly detection.

BlueNet building. Thus, the following features have been considered; the total number of periods with a NPSONC during a day (NPSONC_PERIODNUMBER), a series of consecutive NPSONCs constitutes a period. Besides, this period with a NPSONC anomaly must have at least a period of 30 min (NPSONC_MINUTES). Rule R7 is as follows:

If	$NPSONC_PERIODNUMBER \geq 1$ AND $NPSONC_MINUTES > 30$ min
then	HVAC use without any one in the building.

A case with NPSONC anomaly detection is shown in Fig. 8 where there are two anomalies detected at 17:30 and 20:10 in February when lighting is especially necessary at this time in an office building. As shown, NPSONC detection alerts at least every 10 min when the lights are switched off (no one present either) and the HVAC is running, especially in hours without sunshine. This anomaly alerts about an unnecessary use of HVAC as shown in Fig. 8.

4. Results and discussion

The aim of this study was to optimise the Energy Management System (EMS) in buildings for any energy operation. Therefore, in order to optimise the EMS a rule-based system was developed based on EELs and historical data with the intent of detecting power consumption anomalies in building power management. All the results were analysed and validated with the EE experts with the following statistical measures of the performance of a binary classification test. The outcome of the test can be positive (classifying anomaly cases) or negative (classifying no anomaly cases). And the results for each case could match with the actual status of the case or not, as is detailed as follows. In that sense:

- *True positive (TP)*: anomaly case correctly identified as anomaly case.
- *False positive (FP)*: no anomaly case incorrectly identified as anomaly case.
- *True negative (TN)*: no anomaly case correctly identified as no anomaly case.

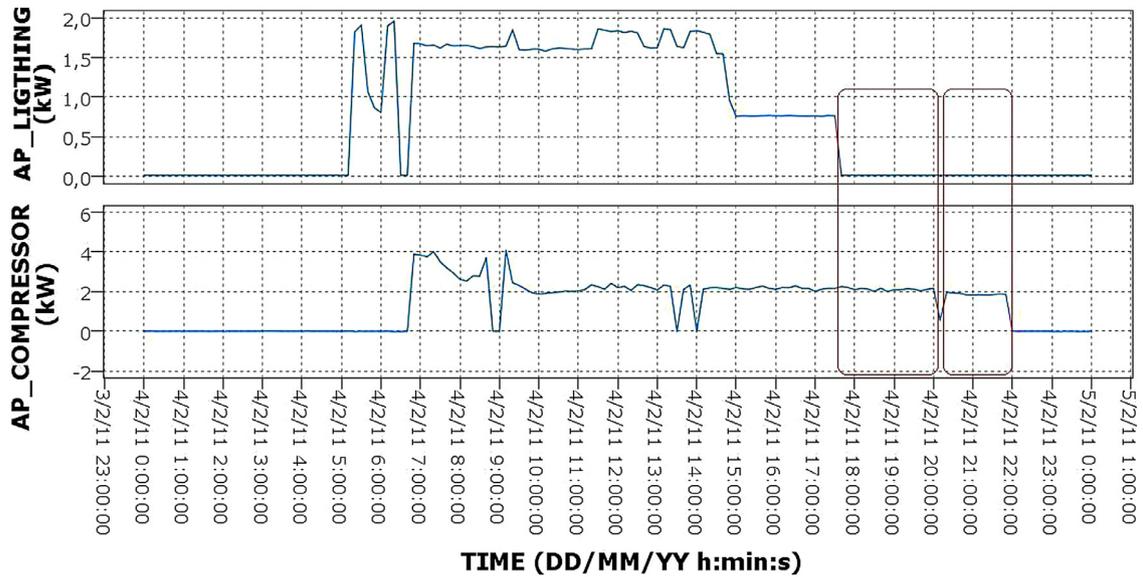


Fig. 8. NPSONC anomaly detection.

- *False negative (FN)*: anomaly case incorrectly identified as no anomaly case.

And in general, *positive* means *identified* and *negative* means *rejected*, therefore:

- *True positive* = correctly identified.
- *False positive* = incorrectly identified.
- *True negative* = correctly rejected.
- *False negative* = incorrectly rejected.

On other hand, other statistical measures were taken into account to analyse further the performance of the test. These statistical measures are; *sensitivity*, *FPR*, *specificity* and *FNR*, also known in statistics as classification functions:

- *Sensitivity*, also called *true positive rate (TPR)* or recall, it measures the proportion of positives that are correctly identified as such.

$$\text{Sensitivity} = \text{TPR} = \text{TP}/P = \text{TP}/(\text{TP} + \text{FN})$$

- *False positive ratio (FPR)* measures the proportion of positives that are incorrectly identified as such.

$$\text{FPR} = \text{FP}/N = \text{FP}/(\text{FP} + \text{TN})$$

- *Specificity*, also called the *true negative rate (TNR)*, it measures the proportion of negatives that are correctly identified as such.

$$\text{Specificity} = \text{TNR} = \text{TN}/N = \text{TN}/(\text{TN} + \text{FP})$$

- *False negative ratio (FNR)* measures the proportion of negatives that are incorrectly identified as such.

$$\text{FNR} = \text{FN}/(\text{FN} + \text{TP})$$

Thus, *sensitivity* quantifies the avoiding of false negatives, as *specificity* does for false positives ($\text{Sensitivity} + \text{FNR} = 1$; $\text{Specificity} + \text{FPR} = 1$).

The results were obtained over the sample of 509 days and the rates about the successful of the rules to detect anomalies were as follows:

- *Compressor use while the outside temperature is between temperature comfort range (R1) and compressor ON and no temperature change (R2)*:

Any anomalies type R1 and R2 were detected in the entire sample. This issue is due to a good HVAC calibration, without energy leaks and a good HVAC management.

- *Anomaly detected by a plurality of operational changes in the HVAC compressor (R3)*:

The 7.66% of days of the total sample had an anomaly of type R3. Concretely, in the last 2 months, 5 anomalies were detected due to excessive temperature changes caused by energy leaks. The results were double checked with EE experts, and successful rates of this test were as follows (Table 2):

Table 2
Statistical measures (R3).

	R3 anomaly cases	No R3 anomaly cases	Total cases
R3 anomaly detection	4	1	5
No R3 anomaly detection	3	57	60
Total	7	58	65

$$\text{TP} = 4/5; \text{FP} = 1/5; \text{TN} = 57/60; \text{FN} = 3/60$$

$$\text{Sensitivity} = 94.12\%; \text{FNR} = 5.88\%$$

$$\text{Specificity} = 82.61\%; \text{FPR} = 17.39\%$$

- *Anomaly detected by a large number of operational regime changes in HVAC compressor (R4)*:

The 13.75% of days of the total sample had an anomaly of type R4. In particular, in the last 2 months 9 anomalies of type R4 were detected due to excessive temperature changes caused by energy leaks. The results were double checked with EE experts, and successful rates of this test were as follows (Table 3):

Table 3
Statistical measures (R4).

	R4 anomaly cases	No R4 anomaly cases	Total cases
R4 anomaly detection	8	1	9
No R4 anomaly detection	2	54	56
Total	10	55	65

$TP=8/9$; $FP=1/9$; $TN=54/56$; $FN=2/56$
Sensitivity = 96.14%; *FNR* = 3.86%
Specificity = 89.67%; *FPR* = 10.33%

- *Anomaly detected by SONCCIT (R5):*

The 13.59% of days of the total sample had an anomaly of type R5. Specifically, in the last 2 months 8 anomalies of type R5 were detected due to bad thermal insulation. In this case, R5 successful rates were as follows (Table 4):

Table 4
Statistical measures (R5).

	R5 anomaly cases	No R5 anomaly cases	Total
R5 anomaly detection	8	0	8
No R5 anomaly detection	2	55	57
Total	10	55	65

$TP=8/8$; $FP=0/8$; $TN=55/57$; $FN=2/57$
Sensitivity = 96.61%; *FNR* = 3.39%
Specificity = 100%; *FPR* = 0.00%

- *Anomaly detected by SOFFCCIT (R6):*

The 23.72% of days of the total sample had an anomaly of type R6. In particular, in the last 2 months 14 anomalies of type R6 were detected caused by an external source of heat in winter or cold in summer. The results were double checked with EE experts, and successful rates of this test were as follows (Table 5):

Table 5
Statistical measures (R6).

	R6 anomaly cases	No R6 anomaly cases	Total
R6 anomaly detection	11	3	14
No R6 anomaly detection	3	48	51
Total	14	51	65

$TP=11/14$; $FP=3/14$; $TN=48/51$; $FN=3/51$
Sensitivity = 93.03%; *FNR* = 6.96%
Specificity = 81.45%; *FPR* = 18.54%

- *Anomaly detected by NPSONC (R7):*

The 5.69% of days of the total sample had an anomaly of type R7. Concretely, in the last 2 months 4 anomalies of type R7 were detected due to HVAC use without any one in the building. In this case, R7 successful rates were as follows (Table 6):

Table 6
Statistical measures (R7).

	R7 anomaly cases	No R7 anomaly cases	Total
R7 anomaly detection	3	1	4
No R7 anomaly detection	3	58	61
Total	6	59	65

$TP=3/4$; $FP=1/4$; $TN=58/61$; $FN=3/61$
Sensitivity = 92.73%; *FNR* = 7.27%
Specificity = 95.08%; *FPR* = 4.92%

These anomaly detections advise about possible problems that are taking place in the building (HVAC break down, incorrectly dimensioned and other problems). The detected anomalies could also notify about an unnecessary waste in power that could report significant profits, and on other hand, the high specificity on the or-

der of 90% and the sensitivity on the order of 95%, guarantee the quality of the EE anomalies detection.

With the results in the study that were shown above, the system integration in EE software could yields big benefits: an environmental friendly smart building management, an anomaly building detection and can yield large profits due to consumption reduction.

Regarding the set up stage, it was strategized an easy set up due to the modular architecture design of the system, in which the rule-base module obtains the streaming of master sensor devices and Modbus master device data as inputs, and a RestFull API which provides rule-base system services, and this services are connected with the front-end to provide all the knowledge extracted as outputs.

Some of the advantages of the system regarding to other systems are the hybrid knowledge extracted from the EE experts and data, which provides a customised solution for each scenario, furthermore, this customised solution supplies a deep know how provided by the EE experts, and the objective know how inferred from the point of view of the data. And finally, this system offers the possibility of analyse huge amount of data from distributed buildings, detecting EE anomalies with high precision. Some possible limitations of this system could be the bandwidth for several distributed data sources creating bottlenecks, but this problem was addressed using master devices which funnel the streaming data.

On other hand, this solution is able to scale easily only including inputs from others buildings in the rule-base system module. It is necessary apply the DM techniques described in Sections 3.2 and 3.3 to set-up and configure the set points of the EEI and rules regarding several features of the building; region, human consumption behaviours, the culture of the persons, etc. And on other hand, discover possible new rules, such as, freeze protection rules in those countries in which it is reached low temperatures.

5. Conclusions

Having contrasted the result of the study and corroborated the anomalies in the BlueNet building with Isotrol experts, the future goal is to integrate this framework in the commercial EE software called Eugene, whose owner is the ISOTROL company. Eugene is an EMS that obtains results in real-time of buildings through a SCADA and shown in real time through a graphic interface. This tool unites reports of accumulated consumption divided by zones of the building, monthly consumptions of every area of the building, active power of the building in real time and historical, etc. (Fig. 9). Adding this module to Eugene, has provided a general point of view that shows all anomalies happening in buildings controlled by Eugene in real time, and also quantifies EE based on an EEIs system and a Rule-Based System module. This tool advises experts about applying new policies for detection in real time and without analysing thousands of data. And it also advises about accurate consumption behaviour in order to improve the EE in smart buildings.

This paper tackles the EE problem in smart buildings from the point of view of power consumption. Smart buildings are equipped with an EMS that is able to intelligently control the activation of devices and accordingly to minimize power, taking into account the varying requirements of the users. This study describes the design and development of a system to improve EMS and detect anomalies in smart buildings based on EEIs and Rule-Based System module.

In the first instance, a bibliographic revision of research with the same objective has been carried out. Other similar work in smart buildings has been checked but no other work has been set up with the aim to render support in an EE detection of anomalies system. This system has been developed using DM techniques to

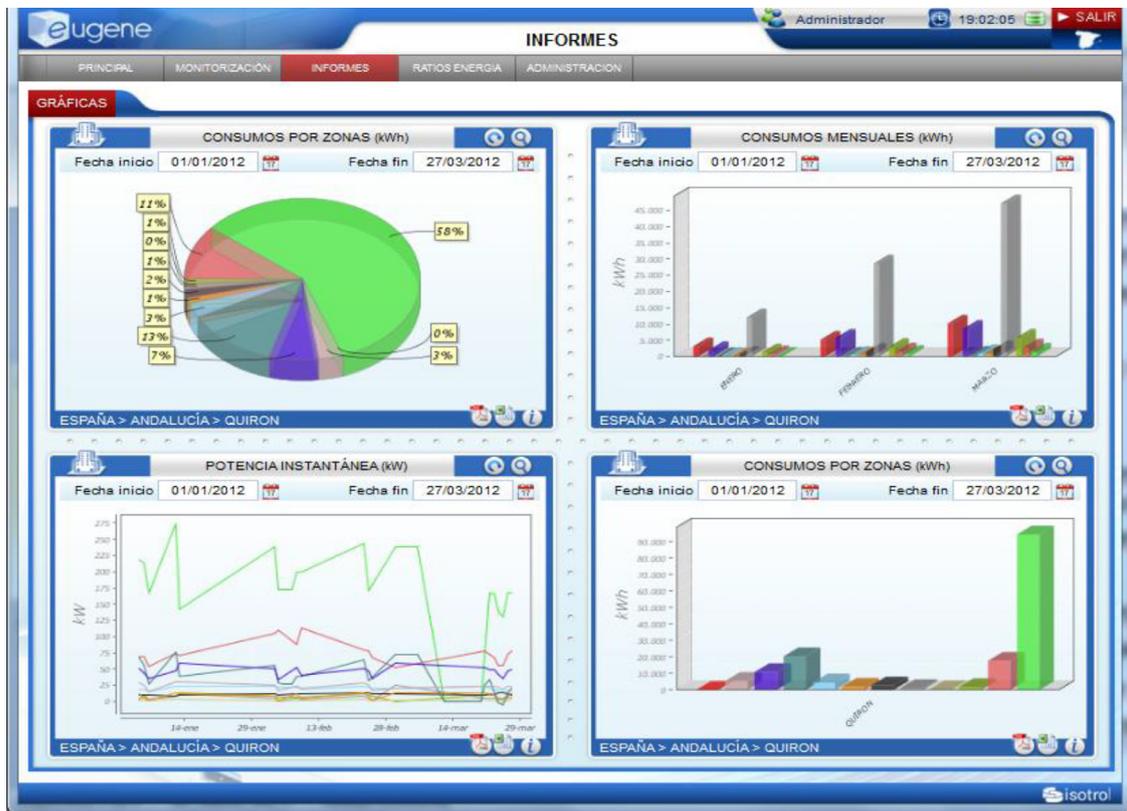


Fig. 9. Eugene EE software, ISOTROL company.

elicit the EE experts' knowledge and the knowledge hidden in the data. Once the knowledge in BlueNet consumption behaviour has been extracted, a module of EEs and a Rule-Based System module has been modelled and developed.

The quality of this work is illustrated by a case study that employs a real database and works in real time having excellent behaviour in anomaly detection. Also, this system helps to EE experts to analyse Big Data in smart buildings describing their behaviour, detecting anomalies and quantify this behaviour through EEs. In addition, this work is applied to a real project adhered to FP7 (FP7-285229 KnoholeM) demonstrating the feasibility of the approach, and also this work helps to reach the guidelines set by UE directives in future buildings (20-20-20 Objectives).

A future perspective of line to research could be the integration with a Demand Side Management (DSM), vehicles to the grid (V2G) management, system of storage of energy provided by renewable energy in the building, integration of a low consumption system of lighting managed by a dimmer, etc. all these research lines oriented to IoT with smart and autonomous decision making. With this argument new anomalies detected through DM techniques and expert knowledge could be integrated, and improve the intelligence of the edifice.

Acknowledgment

The authors thank the FP7-285229 KnoholeM, for providing funds for this project. The authors thank the ISOTROL company for their help and cooperation to extract the data from the corporative database.

APPENDIX A. MAIN attributes in bluenet building

AEMET_AT (°C): Outdoor air temperature in Celsius degrees outside of BlueNet building.

AP_CLI_FASE1 (kW): Active power of phase1 in HVAC.
AP_CLI_FASE2 (kW): Active power of phase2 in HVAC.
AP_CLI_FASE3 (kW): Active power of phase3 in HVAC.
AP_LIG_FASE1 (kW): Active power of phase1 in lighting.
AP_LIG_FASE2 (kW): Active power of phase2 in lighting.
AP_LIG_FASE3 (kW): Active power of phase3 in lighting.
AP_POW_FASE1 (kW): Active power of phase2 in power.
AP_POW_FASE2 (kW): Active power of phase2 in power.
AP_POW_FASE3 (kW): Active power of phase2 in power.
PRESENCE_ID_ENT (DD/MM/YY h:min:s): The entry timestamp of an employee in BlueNet building.
PRESENCE_ID_EXI (DD/MM/YY h:min:s): The exit timestamp of an employee in BlueNet building.
ZIGBEE_AT (°C): Indoor air temperature in Celsius degrees inside of BlueNet building.

**FASE1*, *FASE2*, *FASE3* variables for HVAC, lighting and power describes the active power for each phase of the three-phase AC.

APPENDIX B. EEs and rule-based attributes in bluenet building

AP_AIR_MIXERS (kW): Active power in air mixers of HVAC system.
AP_COMPRESSOR_MEAN (kW): Average of daily active power of the HVAC compressor.
AP_COMPRESSOR (kW): Active power of the HVAC compressor.
AP_LIGHTING (kW): Active power of the lighting system.
AP_SPLITS (kW): Active power in splits of HVAC system.
NPSONC_MINUTES (kW): Time in minutes in which lighting was working with a NPSONC anomaly is detected.
NPSONC_PERIODNUMBER: Total number of periods with a NPSONC anomaly during this day.
OCC_STOPPOINTS: Number of daily On-Off operations in HVAC compressor.

ORCC_PERIODNUMBER: Total number of periods with an operational regime change in HVAC compressor during a day.

ORCC_MINUTES (min): Time in minutes in which an operational regime change in HVAC compressor anomaly is detected.

SOFFCCIT_PERIODNUMBER: Total number of periods with a switch off in HVAC compressor and an abnormal change in interior temperature anomaly during a day.

SONCCIT_PERIODNUMBER: Total number of periods with a switch on in HVAC compressor and an abnormal change in interior temperature anomaly during a day.

$\Delta AP_COMPRESSOR$ (kW): Variation in active power of HVAC compressor during 10 min.

$\Delta ZIGBEE$ (°C): Temperature variation in a specific room of the BlueNet building during a period of 30 min.

References

- Altwies, J. E., & Nemet, G. F. (2013). Innovation in the U.S. building sector: an assessment of patent citations in building energy control technology. *Energy Policy*, 52, 819–831. doi:10.1016/j.enpol.2012.10.050.
- Andrews, C. J., & Krogmann, U. (2009). Explaining the adoption of energy-efficient technologies in U.S. commercial buildings. *Energy and Buildings*, 41(3), 287–294. doi:10.1016/j.enbuild.2008.09.009.
- Azadeh, a., Saberi, M., Ghaderi, S. F., Gitiforouz, a., & Ebrahimipour, V. (2008). Improved estimation of electricity demand function by integration of fuzzy system and data mining approach. *Energy Conversion and Management*, 49(8), 2165–2177. doi:10.1016/j.enconman.2008.02.021.
- Belussi, L., & Danza, L. (2012). Method for the prediction of malfunctions of buildings through real energy consumption analysis: holistic and multidisciplinary approach of Energy Signature. *Energy and Buildings*, 55, 715–720. doi:10.1016/j.enbuild.2012.09.003.
- Brooks, J., & Barooah, P. (2014). Energy-efficient control of under-actuated HVAC zones in buildings. In *American control conference (ACC)*, (pp. 424–429). <http://doi.org/10.1109/ACC.2014.6859151>.
- Capozzoli, A., Lauro, F., & Khan, I. (2015). Fault detection analysis using data mining techniques for a cluster of smart office buildings. *Expert Systems with Applications*, 42(9), 4324–4338. doi:10.1016/j.eswa.2015.01.010.
- Chou, S. K., Chua, K. J., Ho, J. C., & Ooi, C. L. (2004). On the study of an energy-efficient greenhouse for heating, cooling and dehumidification applications. *Applied Energy*, 77(4), 355–373. doi:10.1016/S0306-2619(03)00157-0.
- Colmenar-Santos, A., Terán de Lober, L. N., Borge-Diez, D., & Castro-Gil, M. (2013). Solutions to reduce energy consumption in the management of large buildings. *Energy and Buildings*, 56, 66–77. doi:10.1016/j.enbuild.2012.10.004.
- Day, A., Jones, P., & Turton, J. (2013). Development of a UK centre for efficient and renewable energy in buildings (CEREB). *Renewable Energy*, 49, 166–170. doi:10.1016/j.renene.2012.01.039.
- De Alegría Mancisidor, I. M., Díaz de Basurto Uruga, P., Martínez de Alegría Mancisidor, I., & Ruiz de Arbuló López, P. (2009). European Union's renewable energy sources and energy efficiency policy review: the Spanish perspective. *Renewable and Sustainable Energy Reviews*, 13(1), 100–114. doi:10.1016/j.rser.2007.07.003.
- Diakaki, C., Grigoroudis, E., Kabelis, N., Kolokotsa, D., Kalaitzakis, K., & Stavrakakis, G. (2010). A multi-objective decision model for the improvement of energy efficiency in buildings. *Energy*, 35(12), 5483–5496. doi:10.1016/j.energy.2010.05.012.
- Diakaki, C., Grigoroudis, E., & Kolokotsa, D. (2008). Towards a multi-objective optimization approach for improving energy efficiency in buildings. *Energy and Buildings*, 40(9), 1747–1754. doi:10.1016/j.enbuild.2008.03.002.
- Djevic, M., & Dimitrijevic, a. (2009). Energy consumption for different greenhouse constructions. *Energy*, 34(9), 1325–1331. doi:10.1016/j.energy.2009.03.008.
- DOE/EIA-0383. (2013). Annual Energy Outlook 2013.
- Domínguez, M., Fuertes, J. J., Alonso, S., Prada, M. a., Morán, A., & Barrientos, P. (2013). Power monitoring system for university buildings: architecture and advanced analysis tools. *Energy and Buildings*, 59, 152–160. doi:10.1016/j.enbuild.2012.12.020.
- Doukas, H., Nychtis, C., & Psarras, J. (2009). Assessing energy-saving measures in buildings through an intelligent decision support model. *Building and Environment*, 44(2), 290–298. doi:10.1016/j.buildenv.2008.03.006.
- Dounis, a. I., & Caraiacos, C. (2009). Advanced control systems engineering for energy and comfort management in a building environment—A review. *Renewable and Sustainable Energy Reviews*, 13(6–7), 1246–1261. doi:10.1016/j.rser.2008.09.015.
- EEA 225, E. C. (2013). Financial support for energy efficiency in buildings.
- EIA. (2013). U.S. energy information administration. annual energy outlook 2013. *U.S. Energy Information Administration. Annual Energy Outlook 2013, April 2013*, (February 2013).
- Escrivá-Escrivá, G. (2011). Basic actions to improve energy efficiency in commercial buildings in operation. *Energy and Buildings*, 43(11), 3106–3111. doi:10.1016/j.enbuild.2011.08.006.
- Fenz, S., Heurix, J., Neubauer, T., Tjoa, A. M., Ghiassi, N., Pont, U., & Mahdavi, A. (2014). SEMERGY.net: automatically identifying and optimizing energy-efficient building designs. *Computer Science - Research and Development, November 2014*, 1–6. <http://doi.org/10.1007/s00450-014-0285-5>.
- Figueiredo, J., & Sá da Costa, J. (2012). A SCADA system for energy management in intelligent buildings. *Energy and Buildings*, 49, 85–98. doi:10.1016/j.enbuild.2012.01.041.
- GhaffarianHoseini, A., Dahlan, N. D., Berardi, U., GhaffarianHoseini, A., & Makaremi, N. (2013). The essence of future smart houses: from embedding ICT to adapting to sustainability principles. *Renewable and Sustainable Energy Reviews*, 24, 593–607. doi:10.1016/j.rser.2013.02.032.
- Gul, M. S., & Patidar, S. (2015). Understanding the energy consumption and occupancy of a multi-purpose academic building. *Energy and Buildings*, 87, 155–165. doi:10.1016/j.enbuild.2014.11.027.
- Haberl, J., Sparks, R., & Culp, C. (1996). Exploring new techniques for displaying complex building energy consumption data. *Energy and Buildings*, 24(1), 27–38. doi:10.1016/0378-7788(95)00959-0.
- Han, J., & Kamber, M. (Eds.). (2001). Data mining. Concepts and techniques. MorganKaufmann.
- Hou, Z., Lian, Z., Yao, Y., & Yuan, X. (2006). Data mining based sensor fault diagnosis and validation for building air conditioning system. *Energy Conversion and Management*, 47(15–16), 2479–2490. doi:10.1016/j.enconman.2005.11.010.
- Hurtado, L. A., Nguyen, P. H., & Kling, W. L. (2014). Agent-based control for building energy management in the smart grid framework. In *Innovative smart grid technologies conference Europe (ISGT-Europe)*, 2014 IEEE PES (pp. 1–6). <http://doi.org/10.1109/ISGT-Europe.2014.7028937>.
- Kim, H., Stumpf, A., & Kim, W. (2011). Analysis of an energy efficient building design through data mining approach. *Automation in Construction*, 20(1), 37–43. doi:10.1016/j.autcon.2010.07.006.
- Klein, L., Kwak, J., Kavulya, G., Jazizadeh, F., Becerik-Gerber, B., Varakantham, P., & Tambe, M. (2012). Coordinating occupant behaviour for building energy and comfort management using multi-agent systems. *Automation in Construction*, 22, 525–536. doi:10.1016/j.autcon.2011.11.012.
- Kolokotsa, D., Kalaitzakis, K., Antonidakis, E., & Stavrakakis, G. (2002). Interconnecting smart card system with PLC controller in a local operating network to form a distributed energy management and control system for buildings. *Energy Conversion and Management*, 43(1), 119–134. doi:10.1016/S0196-8904(01)00013-9.
- Lazaroiu, G. C., & Roscia, M. (2012). Definition methodology for the smart cities model. *Energy*, 47(1), 326–332. doi:10.1016/j.energy.2012.09.028.
- Liu, H., Zabinsky, Z. B., & Kohn, W. (2010). Rule-based control system design for smart grids. In *2010 IEEE power and energy society general meeting* (pp. 1–5). <http://doi.org/10.1109/PES.2010.5589410>.
- Li, X., Bowers, C. P., & Schnier, T. (2010). Classification of Energy consumption in buildings with outlier detection. *IEEE Transactions on Industrial Electronics*, 57(11), 3639–3644. doi:10.1109/TIE.2009.2027926.
- Ma, J., Qin, J., Salsbury, T., & Xu, P. (2012). Demand reduction in building energy systems based on economic model predictive control. *Chemical Engineering Science*, 67(1), 92–100. doi:10.1016/j.ces.2011.07.052.
- Maimon, O., & Rokach, L. (Eds.). (2010). Data mining and knowledge discovery handbook. Springer, US.
- Malatji, E. M., Zhang, J., & Xia, X. (2013). A multiple objective optimisation model for building energy efficiency investment decision. *Energy and Buildings*, 61, 81–87. doi:10.1016/j.enbuild.2013.01.042.
- Marinakis, V., Doukas, H., Karakosta, C., & Psarras, J. (2013). An integrated system for buildings' energy-efficient automation: application in the tertiary sector. *Applied Energy*, 101, 6–14. doi:10.1016/j.apenergy.2012.05.032.
- Marinakis, V., Karakosta, C., Doukas, H., Androulaki, S., & Psarras, J. (2013). A building automation and control tool for remote and real time monitoring of energy consumption. *Sustainable Cities and Society*, 6, 11–15. doi:10.1016/j.scs.2012.06.003.
- Martin, S., Hernandez, J., & Valmaseda, C. (2015). A novel middleware for smart grid data exchange towards the energy efficiency in buildings. In *Proceedings of the 2015 international conference and workshops on networked systems (NetSys)*, IEEE (pp. 1–8). doi:10.1109/NetSys.2015.7089063.
- Nguyen, T. A., & Aiello, M. (2013). Energy intelligent buildings based on user activity: a survey. *Energy and Buildings*, 56, 244–257. doi:10.1016/j.enbuild.2012.09.005.
- Noailly, J. (2012). Improving the energy efficiency of buildings: The impact of environmental policy on technological innovation. *Energy Economics*, 34(3), 795–806. doi:10.1016/j.eneco.2011.07.015.
- Oldewurtel, F., Parisio, A., Jones, C. N., Gyalistras, D., Gwerder, M., Stauch, V., ... & Morari, M. (2012). Use of model predictive control and weather forecasts for energy efficient building climate control. *Energy and Buildings*, 45, 15–27. doi:10.1016/j.enbuild.2011.09.022.
- Omer, A. (2008). Renewable building energy systems and passive human comfort solutions. *Renewable and Sustainable Energy Reviews*, 12(6), 1562–1587. doi:10.1016/j.rser.2006.07.010.
- Pacheco, R., Ordóñez, J., & Martínez, G. (2012). Energy efficient design of building: a review. *Renewable and Sustainable Energy Reviews*, 16(6), 3559–3573. doi:10.1016/j.rser.2012.03.045.
- Pang, X., Wetter, M., Bhattacharya, P., & Haves, P. (2012). A framework for simulation-based real-time whole building performance assessment. *Building and Environment*, 54, 100–108. doi:10.1016/j.buildenv.2012.02.003.
- Pérez-Lombard, L., Ortiz, J., & Pout, C. (2008). A review on buildings energy consumption information. *Energy and Buildings*, 40(3), 394–398. doi:10.1016/j.enbuild.2007.03.007.
- Rashid, Y. R., Sulaiman, M. S., Aziz, A., Selamat, H., Yani, A. H. M., & Kandar, M. Z. (2011). Greening government's office buildings: PWD Malaysia experiences. *Procedia Engineering*, 21, 1056–1060. doi:10.1016/j.proeng.2011.11.2111.

- Ryu, M.-W., Kim, J., Lee, S.-S., Ahn, I.-Y., & Yun, J. (2012). Design of rule model for building energy efficiency base on user occupancy and spatial features. In T. Kim, C. Ramos, J. Abawajy, B.-H. Kang, D. Ślęzak, & H. Adeli (Eds.), *Computer applications for modeling, simulation, and automobile* (pp. 58–63). Berlin Heidelberg: Springer Retrieved from http://0-link.springer.com.fama.us/es/chapter/10.1007/978-3-642-35248-5_9.
- Singh, R. D., & Tiwari, G. N. (2010). Energy conservation in the greenhouse system: a steady state analysis. *Energy*, 35(6), 2367–2373. doi:10.1016/j.energy.2010.02.003.
- Tsiamitros, D., Stimoniaris, D., Poulakis, N., Zehir, M. A., Batman, A., Bagriyanik, M., & ... Dialynas, E. (2014). Advanced energy storage and demand-side management in smart grids using buildings energy efficiency technologies. In *Proceedings of the IEEE PES innovative smart grid technologies, Europe, IEEE* (pp. 1–6). doi:10.1109/ISGTEurope.2014.7028841.
- Vadiee, A., & Martin, V. (2013). Thermal energy storage strategies for effective closed greenhouse design. *Applied Energy*, 109, 337–343. doi:10.1016/j.apenergy.2012.12.065.
- Villar, J. R., Pérez, R., Cal, E. dela, & Sedano, J. (2009). Efficiency in electrical heating systems: an MAS real world application. In Y. Demazeau, J. Pavón, J. M. Corchado, & J. Bajo (Eds.), *Proceedings of the 7th international conference on practical applications of agents and multi-agent systems (PAAMS 2009)* (pp. 460–469). Berlin Heidelberg: Springer Retrieved from http://0-link.springer.com.fama.us/es/chapter/10.1007/978-3-642-00487-2_49.
- Wang, J., Mu, X., Xu, J., & Wang, S. (2012). Design and optimization of dispatching rules for elevator group control aiming at energy saving. In *Proceedings of the 2012 international conference on information science and technology (ICIST)* (pp. 124–127). <http://doi.org/10.1109/ICIST.2012.6221620>.
- Weng, T., & Agarwal, Y. (2012). From Buildings to smart buildings—sensing and actuation to improve energy efficiency. *IEEE Design & Test of Computers*, 29(4), 36–44. doi:10.1109/MDT.2012.2211855.
- Wu, S., & Clements-Croome, D. (2007). Understanding the indoor environment through mining sensory data—A case study. *Energy and Buildings*, 39(11), 1183–1191. doi:10.1016/j.enbuild.2006.07.011.
- Xia, C., Zhu, Y., & Lin, B. (2008). Renewable energy utilization evaluation method in green buildings. *Renewable Energy*, 33(5), 883–886. doi:10.1016/j.renene.2007.10.005.
- Yamagata, Y., & Seya, H. (2013). Simulating a future smart city: an integrated land use-energy model. *Applied Energy*. doi:10.1016/j.apenergy.2013.01.061.
- Yang, R., & Wang, L. (2012). Multi-objective optimization for decision-making of energy and comfort management in building automation and control. *Sustainable Cities and Society*, 2(1), 1–7. doi:10.1016/j.scs.2011.09.001.
- Yu, Z. (Jerry), Haghghat, F., Fung, B. C. M., Morofsky, E., & Yoshino, H. (2011). A methodology for identifying and improving occupant behavior in residential buildings. *Energy*, 36(11), 6596–6608. doi:10.1016/j.energy.2011.09.002.
- Yu, Z. (Jerry), Haghghat, F., Fung, B. C. M., & Zhou, L. (2012). A novel methodology for knowledge discovery through mining associations between building operational data. *Energy and Buildings*, 47, 430–440. doi:10.1016/j.enbuild.2011.12.018.