An ANN-GA Semantic Rule-Based System to Reduce the Gap Between Predicted and Actual Energy Consumption in Buildings

Baris Yuce, *Member, IEEE*, and Yacine Rezgui

Abstract—This paper addresses the endemic problem of the gap between predicted and actual energy performance in public buildings. A system engineering approach is used to characterize energy performance factoring in building intrinsic properties, occupancy patterns, environmental conditions, as well as available control variables and their respective ranges. Due to the lack of historical data, a theoretical simulation model is considered. A semantic mapping process is proposed using principle component analysis (PCA) and multi regression analysis (MRA) to determine the governing (i.e., most sensitive) variables to reduce the energy gap with a (near) real-time capability. Further, an artificial neural network (ANN) is developed to learn the patterns of this semantic mapping, and is used as the cost function of a genetic algorithm (GA)-based optimization tool to generate optimized energy saving rules factoring in multiple objectives and constraints. Finally, a novel rule evaluation process is developed to evaluate the generated energy saving rules, their boundaries, and underpinning variables. The proposed solution has been tested on both a simulation platform and a pilot building – a care home in the Netherlands. Validation results suggest an average 25% energy reduction while meeting occupants' comfort conditions.

Note to Practitioners—This study presents a novel semantic rule generation process using GA and ANN with the objective to reduce the gap between predicted and actual energy consumption in public buildings. Due to the absence of historical energy consumption data, a theoretical simulation approach is used that takes into account a wide range of factors, including building fabric, occupancy patterns, and environmental conditions. Energy sensitive variables are then identified using PCA and MRA. These sensitive variables as well as available control variables (set points) are used to train an ANN to learn energy consumption patterns and behavior within the considered buildings. This trained network is then used as a cost function engine (predictor) for a GA-based optimization process to generate the optimized energy saving rules. Finally, a novel rule evaluation process is devised and implemented to assess energy saving rules quality and boundaries. All generated rules have been tested on both the proposed simulation environment and real buildings. Validation results suggest an average 25% energy reduction.

Index Terms—Artificial neural network (ANN), energy management, genetic algorithm (GA), multi regression analysis (RGA) ontology, principle component analysis (PCA), rule generation.

I. INTRODUCTION

OUR BUILT environment is responsible for some of the most serious global and local environmental changes [1], [2]. Creation and operation of the built environment account for at least 50% of all energy consumption in Europe [1], [2]. The EU is promoting conservation and rational use of energy in buildings as part of the Energy Performance Building Directive [3]. The demand for building energy consumption will continue to increase based on current predictions [2]. This forecasted growth has triggered stringent regulatory requirements to reduce energy demand and CO2 emission [3], [4].

Buildings consist of highly complex and interdependent systems. The holistic control of a building, and its systems, is necessary to achieve targeted regulatory performance objectives [4]. Current commercial products are supervisory control and data acquisition (SCADA) based and only provide a limited logic-based control capability, using predefined static schedules and set-points for available control points. These static schedules/set-points are limited in the way they address continuously changing environmental and building usage and occupancy conditions, requiring a (near) real-time inference capability.

The conceptualization of such a complex system and associated energy gap problem can be addressed through: (i) a detailed physical model; (ii) a black-box model; or (iii) a gray-box model [5], [6]. If the physical and mathematical relationships of the underpinning complex variables can be inferred, the choice of a detailed physical model would provide a most reliable and rigorous approach. However, the difficulty resides in the capability to determine the detailed and accurate mathematical relationship linking input and output variables [6]. A similar problem is posited for gray-box models, which also require prior information about the physical model of the building from a thermal perspective. Conversely, a black-box approach such as ANN does not require any prior information about the model, which makes it a preferred option to intelligently manage energy in buildings [7].

In recent years, energy optimization in buildings has become a popular research area as evidenced by the existing scientific

literature [7]. Reported techniques include artificial neural network (ANN) [4], fuzzy logic [8], ANFIS [9], expert systems [10], multi-agent systems [11], genetic algorithm (GA) [12], genetic programming [13], tabu search [14], and statistical predictive algorithms [15]. However, as noted earlier, the understanding of the complex energy behavior of a building can only be approached effectively using a holistic (systems engineering) approach [16], [17]. This is an area where ontology research offers interesting solutions.

Ontology is a computer and human readable conceptualization of a domain whereby knowledge is represented as sets of semantically related concepts through classes, their properties, and relationships [18]. Ontologies have attracted a lot of interest in the Building and Infrastructure domains [19]. Ontology development is a highly demanding task, requiring expert knowledge to identify and conceptualize underpinning domain artifacts [20]. The most popular ontology development languages are Resource Description Framework (RDF) schema (in a form of multigraph) and Web Ontology Language (OWL) [21]. RDF schemas can be used for a wide range of applications and they proceed by defining the types and properties of the domain resources. However, RDF lacks reasoning capability [22]. Conversely, OWL is designed to describe the structure of a domain with expressive modeling constructs using Boolean operations and universal quantification. OWL is better adapted for machine processing and interoperability compared with other ontology languages. However, both RDF and OWL are limited in the way they handle rules. Several rule extension frameworks have been proposed to address this limitation, including TRIPLE and SWRL [21]. Another challenge is faced when developing rules from several heterogeneous data sources or without domain expert knowledge [23]. Since data mining algorithms are capable of performing association, cluster generation and knowledge derivation using a large number of variables and datasets, they have been very popular in addressing the above issues [24]. Several data mining techniques have been developed to generate rules using classification and feature extraction methods such as decision trees [24], [25], ANN [26], inductive learning and rule family [27], [28], principle component analysis (PCA) [29], ANN-GP [30], Fuzzy systems [31], Neuro-Fuzzy systems [32], and SVM [33].

Our vision is that a holistic management of building energy can be addressed through a dedicated ontology developed with a language, OWL, with a high semantic expressiveness, augmented with self-updating semantic rules capable of dealing with heterogeneous data sources and (near) real-time conditions. A divide-and-conquer approach is used in that building energy complexity is broken down into a set of discrete but related scenarios, each of which formally described and detailed in terms of dependent and independent governing variables and their complex interactions through mathematical approximations. Our proposed approach addresses the limitations of current SCADA-based commercial energy systems.

Moreover, our approach factors in the intrinsic relationship between: (a) the network of sensing nodes present in a building and (b) the actuation system delivered through a BMS (building management system). This relationship is defined through rule-based semantic mappings that augment a purposely developed building ontology with real-time reasoning capabilities [16], [17].

This paper takes the overarching hypothesis that an in-depth understanding of building physics (through energy modeling and simulation) combined with recent advances in Information and Communication Technologies (in areas such as artificial intelligence, decision support science, optimization, and complex systems) can alleviate and help address the current energy deficiencies of our existing building stock.

In this study, a novel theoretical optimized semantic rule generation and implementation process is proposed to deal with all the complexities presented above. This consists of an optimized learning-based strategy. During the rule generation process, GA is utilized to generate optimized rules using ANN as a cost function predictor. The generated rules deliver optimized values to deal with multi-objective problems. In the building environment, these optimal values should be based on set points for control variables available in a given building. Hence, the theoretically generated rules will be able to respond to real world (multi-objective) requirements such as energy/$C0_2$ emission reduction, and thermal comfort enhancement.

This paper is organized as follows. In Section II, related work is presented, followed (Section III) by the methodology that underpins the research, involving a preprocessing stage (Section IV), the ANN-based learning (Section V), and GA-based optimized rule generation (Section VI). The boundary determination and rule selection stage are presented in Section VII. Testing and validation are presented in Section VIII. Section IX discusses the results and concluding remarks and directions for future research are provided in Section X.

II. RELATED WORK

Building control research has recently addressed limitations of proprietary communication protocols between the BMS and the various sensing and control nodes present in a building [34]–[36]. This paved the way to more advanced solutions for the control and management of energy in buildings [4], [12] that can deal with complex and stochastic problems [37], [38]. Conceptualization of energy problems through semantics (i.e., ontology) has been proposed in [17] and [39]. However, the proposed semantic rules required expert knowledge for the elicitation of the complex variables involved which revealed to be an intensive and time consuming effort [39]. Thus, several intelligent rule generation techniques have been proposed in the literature not requiring domain experts. Moreover, real-world problems involve both discrete and continuous input/output attributes in their rule formulation. Rule generation from mixed attribute data involves converting continuous type attributes to discrete attributes, with the drawback of reducing the accuracy of the rules [40]. However, a neural network-based solution can deal with this problem with a well-trained network [41].

Setiono *et al.* [40] proposed a recursive neural network (NN) with decision tree approach to extract the rules from data sets

with mixed attributes, a mix of discrete and continuous valued attributes of credit score data set. They have compared their results with several algorithms, and the NN-based solution performed best compared to the other algorithms.

Setnes [42] proposed an orthogonal transform and fuzzy clustering approach to extract fuzzy rules, which were applied to the identification of Takagi-Sugeno type rules. One of the major drawbacks with fuzzy rule generation is the need for domain expert(s) [42]. A data-driven approach with defined objectives is more appropriate. Setnes [42] determined the antecedent parts with fuzzy clustering, then the parameters of the rules were determined with least square estimation. One of the weaknesses of the approach is to determine the cluster number using a threshold value.

Pal and Chakraborty [43] proposed a fuzzy rule-based-classifier from a decision tree for real data. They utilized a gradient descent-based pruning stage to remove the non-necessary rules. They compared their results with a C4.5 decision tree and the results showed that the fuzzy-based solution performed better than a C4.5 decision tree.

Although the above techniques performed well with lower number of attributes, they fail in their majority to generate rules when faced with a higher number of attributed data sets [44]. Therefore, a preprocessing stage is required for the reduction of the number of attributes to deliver an accurate rule generation system. PCA is of the most popular and powerful attribute reduction algorithm which may increase the accuracy of the rule generation process. Qiu *et al.* [45] proposed a PCA-ANN-based solution to have a better learning process for higher dimensional data sets. Qiu *et al.* [45] recommend a PCA-based preprocessing stage to increase the accuracy of the learning process. Therefore, a rule generation process with higher dimensional attributed data set requires a dimension reduction to deliver a more accurate solution. PCA is well suited for the reduction of the dimension of attributes. However, one of the major drawbacks with PCA is that class separability is not supported. To overcome this difficulty, a PCA with multi regression analysis (MRA) helps accurately determine key attributes. While PCA can determine the required number of attributes, a MRA generates the highest interactional attributes with their coefficients and significance.

The academic literature provides rich insights into energy optimization algorithms. Deng *et al.* [46] proposed a dynamic programming and mixed-integer linear programming solution to control a chiller schedule. They modeled the problem with a predictive control problem. Although their addressed energy management problem is of a heuristic nonlinear type [9], a time dependent binary schedule of the chiller is developed. Also, the dynamic programming-based solution cannot handle large state spaced problems [36].

GA is one of the most popular stochastic optimization algorithms capable of dealing with complex problems as experienced in the building energy domain [47], [48]. GA has been used to generate the ontology. Wicaksono *et al.* [17] proposed a knowledge driven solution underpinned by an ontology framework. This methodology was implemented using historical data to determine energy consumption anomalies in domestic and public buildings. Although Wicaksono *et al.* [17] have
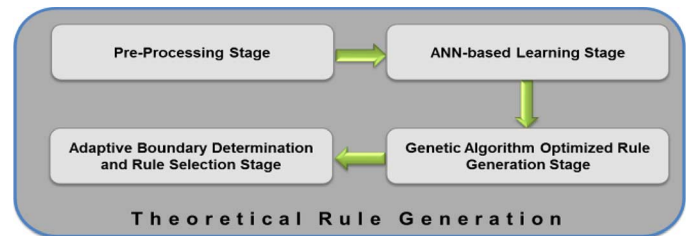

Fig. 1. The block diagram of the proposed method.

attempted to propose a knowledge-based (i.e., semantic) solution, the latter lacked the Artificial Intelligence (AI) dimension to fully exploit the (near) real-time readings (from sensor data) of their proposed solution.

The above limitations constitute a real gap that forms the focus of the present paper. This paper proposes a holistic user friendly negotiation approach using dynamically generated energy saving rules, underpinned by ontology. The proposed system is available through web-services that seamlessly and securely interface with the BMS to actuate negotiable energy reduction plans. This negotiation process allows the Facility Manager (FM) to control each objective (output) with a certain level of flexibility guided by his/her judgment. The rule generation has been completed through a high performance computing facility which includes a cluster of 96 cores.

III. METHODOLOGY

This section gives an overview of the optimized learning-based rule generation process used to implement negotiable energy saving plans in public buildings and its overall energy embedding framework.

The proposed rule generation method consists of four main stages, illustrated in Fig. 1: (a) preprocessing; (b) ANN-based learning; (c) GA-based optimized rule generation; and (d) an adaptive boundary determination process for the environmental variables and rule selection.

The preprocessing stage covers thermal model development, holistic scenario definition, data generation, sensitive environmental variables determination, and a semantic mapping stage. The ANN-based learning stage covers the development of a prediction engine. GA is used to generate optimized rules based on a negotiation process. Moreover, ANN is embedded into the GA to predict the optimization objectives. Finally, a rule selection and boundary determination stage is implemented to evaluate the boundary conditions and accuracy of the generated rules.

The proposed methodology has been successfully embedded and tested on a care home building in The Netherlands (the "forum" building) and validation results are presented in Section VIII. This is delivered in the form of an energy management framework, illustrated in Fig. 2, consisting of: (a) an OWL knowledge base (ontology) providing a semantic description of the building and its properties, augmented with the proposed rules; (b) a fuzzy logic real-time controller (RTC), which fires proposed energy saving rules when triggering conditions are met; (c) a graphical user interface (GUI) at the disposal of the FM to negotiate energy saving plans, and thus interact with the solution; (d) a RDF (Triple/quad store)/SPARQL mapper,
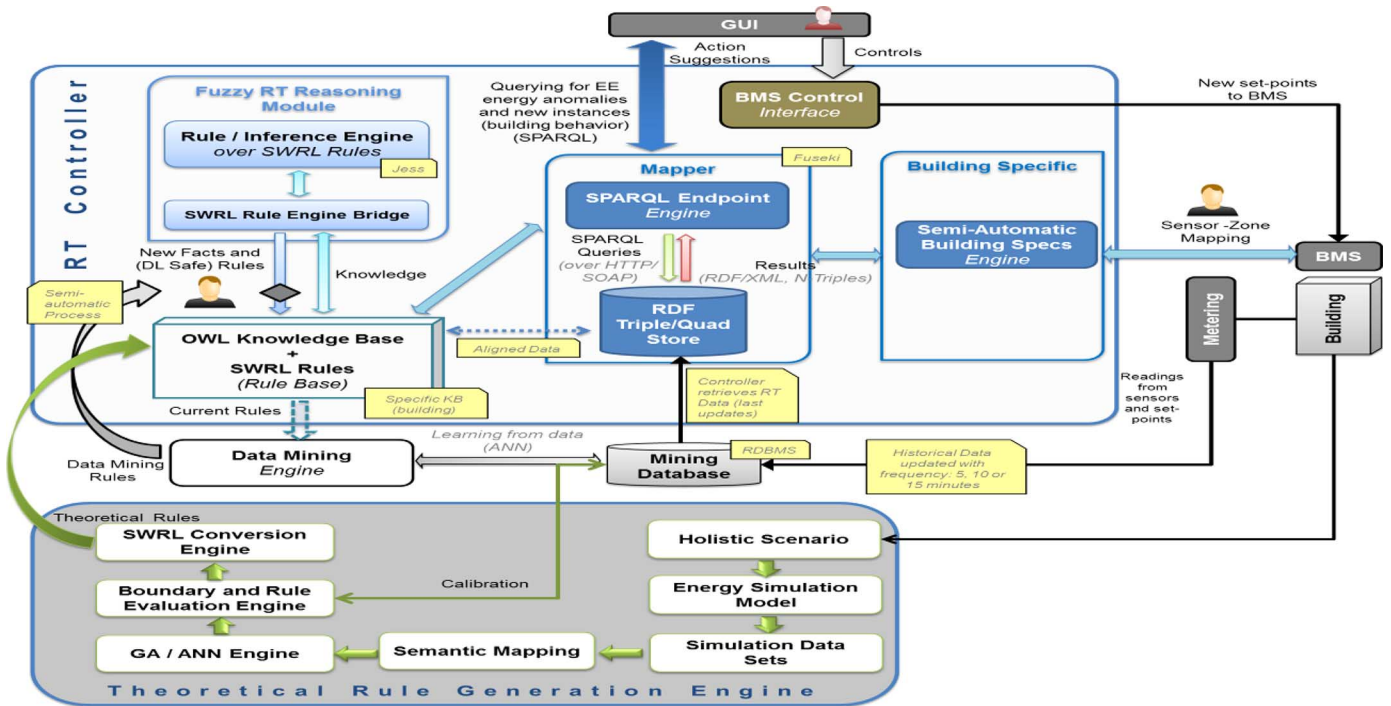
Fig. 2. Proposed system architecture.

which is used to query current energy consumption and CO_2 emission and suggested energy saving rules from the OWL knowledge base, and compare these with (i) baseline data obtained from our simulation model and (ii) mean historical values, and necessary warnings and recommendations to the FM via the GUI; and (e) a building-specific system, which provides rich information about the building as well as the location of each available sensor and actuator accessible through a 3D-Graphical User Interface (GUI). Fig. 2 illustrates the interfaces between these various components, whereby the simulated and metered data are utilized to provide a dynamic and system engineering characterization of the building energy dimension within the building. Each of the above stages is elaborated in the following sections.

IV. PREPROCESSING STAGE

The preprocessing stage is the first and most fundamental stage for this novel rule generation process. This process consists of thermal energy simulation model generation, a holistic use case scenario definition, simulation data generation, sensitive environmental variables determination, and a mapping of the sensitive variables to available sensing nodes. The overall preprocessing stage is illustrated in Fig. 3.

As highlighted above, energy management problems are highly complex and difficult, and an analytical solution is not convenient. We have therefore adopted a Cartesian (divide-and-conquer) approach through the identification and formulation of discrete use case scenarios. Use case scenarios help reduce the complexity and contextualize the addressed energy problem and the required optimal solution.

The preprocessing stage starts with the thermal energy model generation using a software platform, namely, DesignBuilder (a commercial 3D modeler and energy simulation software
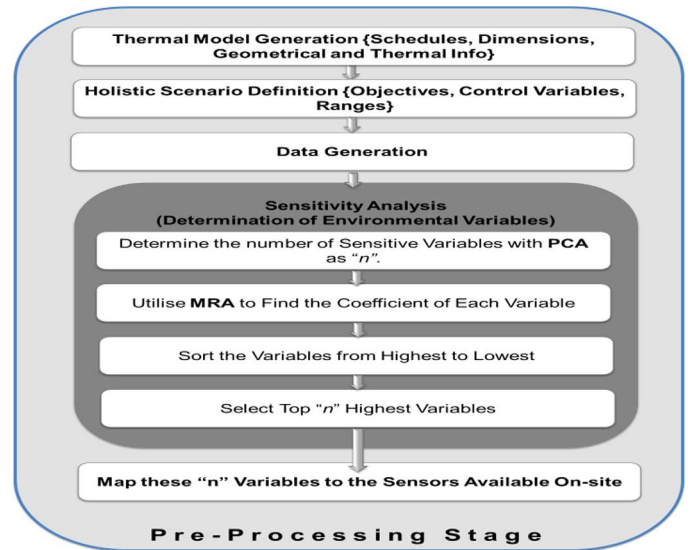

Fig. 3. The preprocessing stage.

platform). The Thermal energy model includes: geometrical information of the building enriched with occupancy information, material and envelope properties, overall intrinsic (including thermal properties) of the building, schedules for heating and cooling devices. The atrium zone (zone id: RC0.13) of the Forum building, situated on the ground floor, is selected as the pilot zone. The thermal model of the Forum building and the pilot zone is illustrated in Fig. 4.

Next, the generated thermal model is simulated based on a use case scenario. A system engineering approach requires a clear definition of the optimization objectives, while all available control variables and their ranges have to be considered.
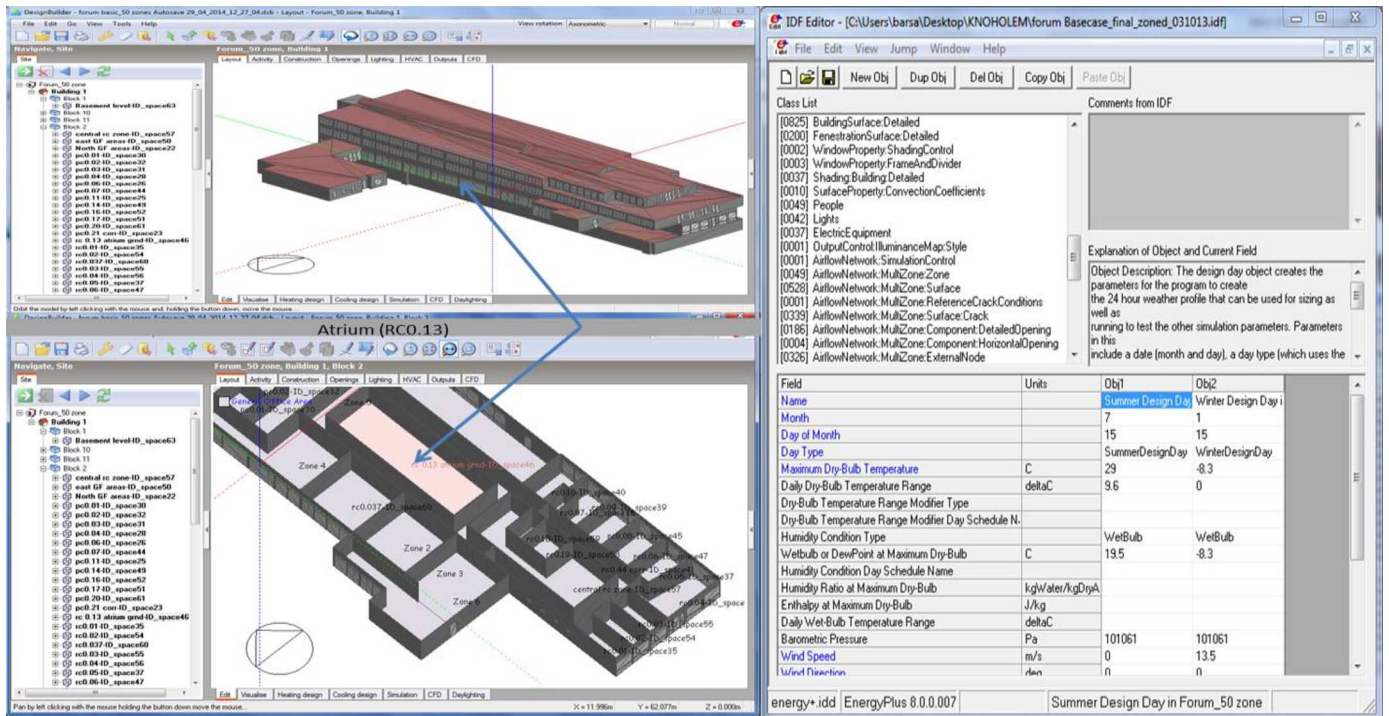
Fig. 4. Energy model of pilot building and floor plan.

TABLE I
A HOLISTIC SCENARIO DEFINITION FOR THE ATRIUM ZONE

Scenario Definitions	Holistic Solution
Scenario Definition	This scenario provides a negotiation based energy management solution to the FORUM building, the ATRIUM zone, by minimising energy consumption while maintaining required comfort conditions
Control Variables	• Heating temperature set point : [16°-24°],(incremental size =1)

Objectives	• Desired amount of minimisation for energy consumption
	   ○ Heating energy minimisation
	   ○ Atrium roof window set point (state): {Off=0, On =1}
	   ○ Lighting set point (state): {Off=0, On =1}
	   ○ Shading set point (state): {Off=0, On =1}
	• Comfort (Predicted Mean Vote (PMV)) optimisation

Effective Environmental Variables	The most effective variables will be determined after sensitivity analysis.
Control Rules	Legislation regarding required internal temperature will need to be adhered to at all times regardless of radiator optimisation scheme implemented.
Actors	BC5, Occupancy sensors, light automation system, automation system, facility manager, technician, TRV, actuators, temperature sensors, weather station, window actuator, shade actuator.
When Applicable	The scenario is applicable to optimise energy and comfort in the Atrium Zone of the FORUM building.

Fig. 5. PCA results of the FORUM building Scenario.

Our scenario involves minimizing energy consumption in the forum building, while maintaining acceptable comfort conditions of the elderly occupants of the care home. A description of the scenario is given in Table I. The next step of the preprocessing stage is to simulate the system based on the different combinations of control variables. Initially, the thermal model generated using DesignBuilder was exported into EnergyPlus (an open-source and cross-platform energy simulation environment). The simulation of the building was then carried out 40 times using 40 different combinations of set points (as an example: (16, 0, 1, 0) which means heating = 16°, window: off, light: on and shade: off).

Each simulation has been run for a year of data and recorded every 15 min. The main objective in generating this huge amount of data is to determine accurately the most sensitive variables (as 954 environmental variables have been found in every simulation) and then to utilize this data set to train a neural network which can then be used as a cost function predictor.
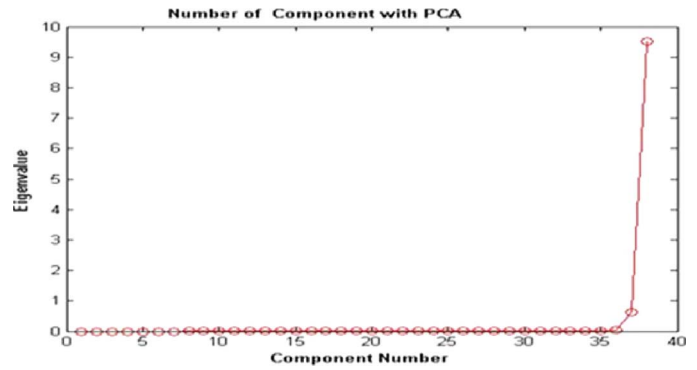
The next stage of preprocessing involves a sensitivity analysis to determine the most sensitive variables. A PCA and MRA-based sensitivity analysis process have been introduced. PCA is used to determine the number of most sensitive variables (expressed as "n"). It is a statistical approach using the covariance matrix and its Eigen values for dimensionality reduction.

The PCA result for the selected use case scenario is 38 components (based on its eigenvalue), as illustrated in Fig. 5 (sorted from lowest to highest).

A MRA is then used to determine the coefficient of each of the "n = 38" variables. This process is modeled as in (1) and (2)

$$F_1\left(\vec{X}\right) = \sum_{i=1}^{ntv} a_{1i}X_i \qquad (1)$$

$$F_2\left(\vec{X}\right) = \sum_{i=1}^{ntv} a_{2i}X_i \qquad (2)$$

where F1 and F2 denote heating (thermal) energy consumption and predicted mean vote (PMV – an indicator of human comfort with a scale of -3 to $+3$), respectively. $\vec{X}$ denotes all the variables generated from the simulation model, a_{1i} and a_{2i} denote the coefficient of variable X_i for heating rate and PMV, respectively, and ntv is the total number of variables in the simulation model ($ntv = 954$ for this case).

In this process, the absolute highest values of 38 a_{1i} and a_{2i} are determined, which represent the most sensitive variables.

Then, a mapping process is performed between the determined 38 variables and the existing sensors on the pilot zone. This process has been carried out using a cross comparison matrix between the list of existing sensors and the 38 environmental variables determined by the above sensitivity analysis. The mapped value of the variables' diagonal element value is 1 and the rest is found as 0. As a result, ten variables are mapped to the existing installed sensors: 1) outdoor air drybulb; 2) wind speed; 3) wind direction; 4) direct solar radiation; 5) solar azimuth angle; 6) solar altitude angle; 7) zone mean air temperature; 8) zone air system sensible heating rate; 9) zone ideal load total cooling rate; and 10) occupancy. The next stage of the proposed algorithm is to train a neural network to create a relationship between variables (environmental and control variables) and objectives in line with the scenario definition of Table I.

V. ANN-BASED LEARNING STAGE

The learning stage of the proposed algorithm is based on a multilayer perceptron (MLP) structure ANN used via MATLAB. ANN is one of the most popular techniques to deal with complex type problems such as energy management [4]. A well-trained ANN is able to create a relationship between inputs and outputs without having a mathematical relationship. In this section, several ANN learning algorithms and topologies have been tested to determine the best performing ANN. The topology of the proposed ANN is presented in Fig. 6 and consists of 17 inputs; the 10 sensitive environmental variables, the control variables (temperature set point, window set point, lighting set point and shading set point), and time information (month, day and hour). The outputs of the ANN are the objectives of the considered scenario, i.e., heating (thermal) energy consumption (KhW) and predicted mean vote (PMV) used to measure levels of comfort.

In order to find the best performing ANN, the first step is to find the most efficient learning algorithm, while keeping constant the number of process elements (five process elements) and transfer functions (both are tangent-sigmoid functions). The second step is to find the required number of process elements in the hidden layer, while keeping constant the best performing learning algorithm and transfer function (both are tangent-sigmoid functions). The last step is to find the transfer functions in the hidden layer and output layer, while keeping constant the best performing learning algorithm and number process elements in the hidden layer. All the analysis is carried out with a maximum of 4000 iterations; with the desired mean squared error level of 0.0001 set as a termination condition.
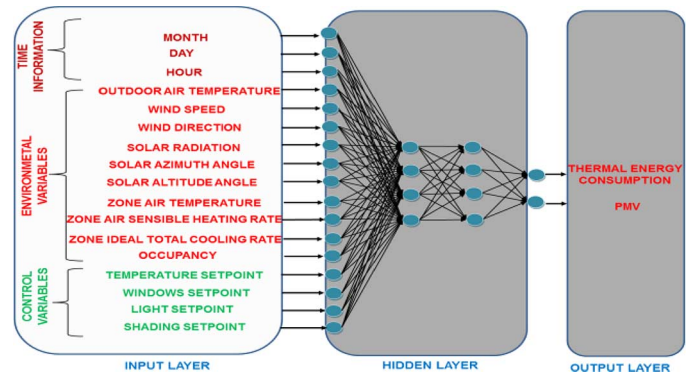

Fig. 6. The proposed topology of the ANN.

TABLE II
TRAINING RESULTS OF THE ANN BASED ON THE LEARNING FUNCTIONS

Training Algorithm	Expected Error Level	Error Level of Algorithm	Number of Epoch
trainbfg	0.0001	0.0008	4000
traincgb	0.0001	0.0005	4000
traincgf	0.0001	0.0006	4000
traincgp	0.0001	0.0005	4000
traingd	0.0001	0.0018	4000
traingda	0.0001	0.0009	4000
traingdm	0.0001	0.0010	4000
traingdx	0.0001	0.0009	4000
trainlm	0.0001	0.0001	830
trainscg	0.0001	0.0002	4000

TABLE III
TRAINING RESULTS OF THE ANN BASED ON THE LEARNING FUNCTIONS

Number of Process Elements in hidden layer	Expected Error Level	Error Level of Algorithm	Number of Epoch
10	0.0001	0.0001	3852
20	0.0001	0.0001	2617
30	0.0001	0.0001	890
50	0.0001	0.0001	1909

The flowing learning algorithms have been tested: 1) BFGS quasi-Newton backpropagation (trainbfg); 2) conjugate gradient backpropagation with Powell-Beale restarts (traincgb); 3) conjugate gradient backpropagation with Fletcher – Reeves updates (traincgf); 4) conjugate gradient backpropagation with Polak-Ribiere updates (traincgp); 5) gradient descent backpropagation (traingd); 6) gradient descent with adaptive learning rate backpropagation (traingda); 7) gradient descent with momentum backpropagation (traingdm); 8) gradient descent with momentum and adaptive learning rate backpropagation (traingdx); 9) Levenberg–Marquardt backpropagation (trainlm), and 10) scaled conjugate gradient backpropagation (trainscg). The results are illustrated in Table II.

According to Table II, Levenberg–Marquardt backpropagation (trainlm) provides the best performance; hence, the following analysis will be based on this algorithm. Once the most efficient learning algorithm is found then the number process elements is determined using a different number of process elements in the hidden layer. In this section, the number of process elements in each experiment was selected as (10-20-30-50), respectively. The results are illustrated in Table III.

TABLE IV
TRAINING RESULTS OF THE ANN BASED ON THE TRANSFER FUNCTION IN
HIDDEN AND OUTPUT LAYERS

Transfer Functions [hidden-layer output-layer]	Expected Error Level	Error Level of Algorithm	Number of Epoch
[tansig tansig]	0.0001	0.0001	863
[tansig logsig]	0.0001	0.0001	114
[tansig purelin]	0.0001	0.0004	4000
[logsig tansig]	0.0001	0.0001	70
[logsig logsig]	0.0001	0.0001	234
[logsig purelin]	0.0001	0.0009	4000
[purelin tansig]	0.0001	0.0002	4000
[purelin logsig]	0.0001	0.0008	4000
[purelin purelin]	0.0001	0.0027	4000
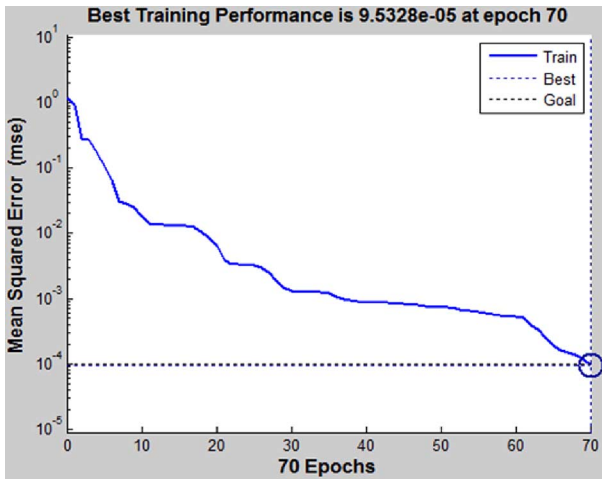
Fig. 7. The training performance of the best performed ANN.

According to Table III, the best performing ANN was found with 30 process elements in the hidden layer. The following experiments will therefore use this number as the number of process elements. Finally, the different transfer function types in the hidden layer and output layer are used to find the best performing ANN. The following functions have been analyzed: tangent-sigmoid (tansig), logarithmic-sigmoid (logsig), and linear function (purelin). The results are illustrated in Table IV.

According to Table IV, the best performing ANN topology was found using trainlm as the learning function, 30 process elements in the hidden layer, and the logarithmic sigmoid and tangent sigmoid function ([logsig tansig]) in the hidden layer and output layer, respectively. The training of ANN has been carried out using the best combination of the topology with defined inputs and outputs. To train ANN with this topology, the data set was divided into two parts as training and test data. The training data is 80% of the raw data, selected randomly. The rest of the data is used for the testing of the network. During the training, the error rate has been set to 0.0001. The proposed network achieved the target level in 70 epochs (Fig. 7). Following that stage, the trained network was tested using the test data (remaining 20% of the original data). According to testing results, 0.0083 MSE error rate is achieved from the network. This trained network was then selected as a prediction engine for GA to generate the semantic rules for the ontology.
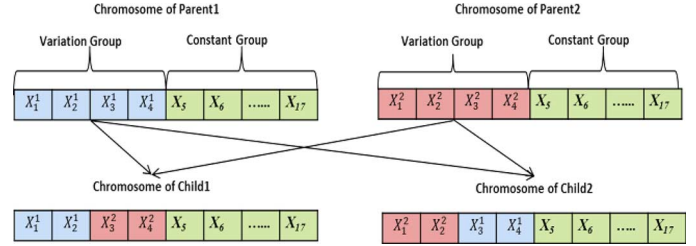

Fig. 8. The crossover operation of proposed GA model for the rule generation process.

VI. GA-BASED OPTIMIZED RULE GENERATION STAGE

This section presents the optimized rule generation process which allows the Facility Manager (FM) to reduce energy consumption while maintaining thermal comfort conditions in the selected scenario/zone. Energy management and control problems in the building environment are highly complex, multi-dimensional, non-linear and stochastic [50]. A GA-based optimization solution is used. GA is a population-based stochastic optimization algorithm which utilizes a global search process to find the optimum solution for complex and non-linear type optimization problems. It has successfully been implemented for building energy management problems [51], [52].

In the proposed novel rule generation process, each chromosome string consists of two main groups: variation group and constant group. The variation group contains the control variables and the constant group is the non-varied group which is composed of date and environmental variables. The crossover operation is carried out among the variation group of two parents chromosome strings with multi point gene exchange operation. The mutation operation is also carried out on the variation group of parent chromosome string. These operations are shown in Figs. 8 and 9.

GA utilizes crossover and mutation operations and carries out the search process until it finds the optimum value of set points, which are the control variables of the systems; and evaluation is carried out based on the objectives defined in the holistic scenario definition (Table I). The proposed optimization problem has four control variables, three date information and ten environmental variables which are given in (3)–(7). Moreover, ANN is utilized as a cost function predictor and embedded into GA. The overall process of the proposed GA is given in Fig. 10.

Subject to:

$$\text{Min } \{F_{objective}(\vec{X}^*)\}. \tag{3}$$

Contraints:

$$|F_{PMV}(\vec{X}^*)| \leq 1 \tag{4}$$
$$X_i^{c_{low}} \leq X_i^c \leq X_i^{c_{upper}} \tag{5}$$
$$X^{c_{low}} = \{16, 0, 0, 0\} \tag{6}$$
$$X^{c_{upper}} = \{24, 1, 1, 1\} \tag{7}$$

where $F_{objective}$ is the objective function for the selected case, i.e., "Heating Energy Consumption"; $\vec{X}^*$ represents overall variable vector which consists of control variables, time information and environmental variables,
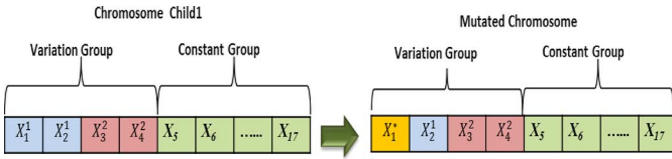
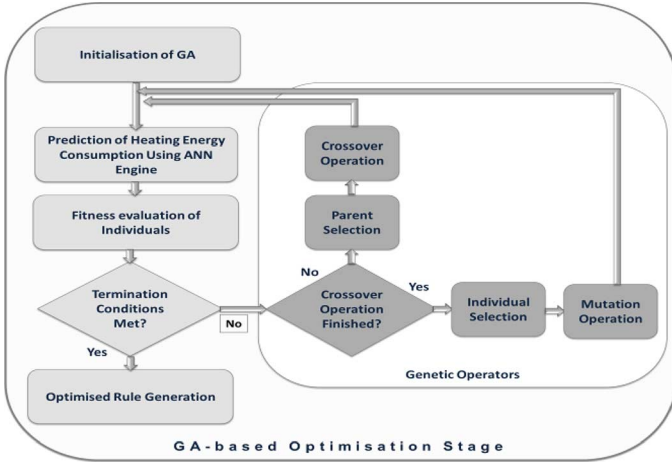Fig. 9. The mutation operation of proposed GA model for the rule generation process.

Fig. 10. The proposed GA model for the rule generation process.

Fig. 11. Variables utilized in the antecedent and consequent parts of the rules.

Fig. 12. Example of an optimized generated rule using GA and ANN.

Fig. 13. The fitness value of the heating energy consumption in each iteration.

$\vec{X}^* = \{X1, X2, X3, X4, \ldots, X17\}$; X^C represents control variable vector, $X^C = \{X_1, X_2, X_3, X_4\}$; $X1$, $X2$, $X3$; and $X4$ are control variables of the problem which are temperature set point, window set point, shading set point and lighting set point, respectively; $X5$, $X6$, and $X7$ are month, day and hour information, respectively. Finally, $X8 \ldots X17$ are the environmental variables. $X2$, $X3$ and $X4 \in \mathbb{Z}$, and $X1$, $X8$, $X9 \ldots X17 \in \mathbb{R}$. $|F_{PMV}(\vec{X}^*)|$ represents the absolute value of the predicted mean vote (i.e., thermal comfort level); $X^{c_{low}}$ represents the lower boundaries for the control variables; $X^{c_{upper}}$ represents the upper boundaries for the control variables.

The proposed rule generation is devised based on predetermined negotiation processes, designed for energy reduction levels of 5%, 10%, 15%, 20%, 25%, and 30%, while using thermal comfort as the constraint of the problem. Therefore, the proposed method does not allow the system to make reduction infinitely or to run continuously. Once the optimum set points are found, the optimized rules can be generated by placing the environmental variables in the input section (condition – antecedent part) and the optimized set points in the action section.

Each variable of the input section is designed with an initial β value range of ($\beta = \alpha/2 =$) 0.05. This represents the designed error condition for á during MRA (90% accurately determined). This value is then updated according to an adaptive boundary range process which will be explained in detail in the boundary determination section. Moreover, each rule has also been weighted based on the accuracy of the optimization process. When the target reduction value is found this value will be 1 (100%).

The antecedent and consequent parts of the rules are given in Fig. 11, and an example of rule formation is given in Fig. 12.

In the above example rule, 5% heating energy reduction is desired by the facility manager in the atrium zone RC0.13 (negotiated reduction level). The initial energy consumption was 3.13 KWh and the desired energy consumption level is 2.974 KWh. In this example, the GA seeks for the optimum value of the control variables to achieve the negotiated level. The optimization process has found an optimum solution and recommends the user to set control variables as 17.53° C, off (0), off (0) and off (0) for temperature set point, window set point, shading set point and lighting set point, respectively, to achieve this reduction level. The optimization process has found the global optimum solution in 20 iterations based on the fitness evaluation using ANN as a predictor engine. The fitness of the best individual is illustrated in Fig. 13.

VII. ADAPTIVE BOUNDARY DETERMINATION AND RULE SELECTION STAGE

This section presents the final stage of the proposed rule generation process. To implement rules in real-life, upper
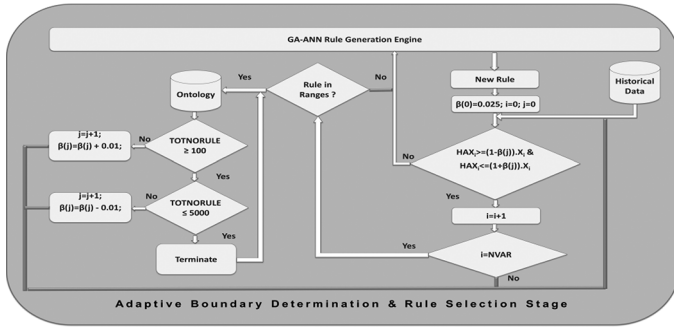
Fig. 14. Flow chart of the adaptive boundary determination rule selection process.

Fig. 15. The SWRL form of the example rule.

and lower boundaries of environmental variables have to be satisfied. However, the data generated through the simulation system may not be matched with real-life conditions. This is related to the unavailability of the most recent weather data used in the simulation model (EnergyPlus Epw-file). To start the rule generation process, the original raw data set is used and sensitivity analysis is carried out using this data set at a confidence rate of 90% and an error rate of 010. This error rate value is used as the initial the upper and lower boundaries of the rules as 0.05 for ($\beta = \alpha/2$) value. However, the rules may not be fired with this boundary range. Additionally, the large number of generated rules with the lower boundary range may hinder a (near) real-time response. Therefore, the generated rules had to be reduced, and factor in information of recent weather conditions. Thus, the value of β needs to be adaptive according to the most recent historical data. Therefore, the selected upper ($X_i + \beta$) and lower ($X_i - \beta$) boundaries of each variable in each rule has been updated by comparing the current range with the average upper and lower range of the historical data sets. In this study, the average values of seven consecutive day's weather conditions are selected to update the value of β. The average historical data value of environmental variable X_i ($i = 8, 9 \ldots 17$) is given according to (8)

$$\text{HAX}_i = \frac{1}{\text{nhd}} \sum_{j=1}^{\text{nhd}} \text{HX}_i^j \tag{8}$$

where HAX_i is the average weather data, HX_i^j is the historical data value of variable X_i ($i = 8 \ldots 17$), and nhd is the selected number of the historical data for environmental variables.

To keep the generated rules within an acceptable range, the total number of rules (TOTNORULE) is kept in the range of 100–2000, empirically. A lesser number of rules may not be able to give an efficient solution and a higher number of rules increase the response time of the overall solution. The overall process is illustrated in Fig. 14. After the rule generation process, these rules are implemented in the pilot zone after conversion into SWRL and then inclusion in the OWL ontology. The rules are fired by an inference engine when the user request is received. The inference engine seeks for the rule which contains the desired objective and the desired reduction level with the highest weight. Among the matched rules, if any of these rules' antecedent parts range match the sensor readings and has the highest weight, then that rule will be fired by the inference engine, and update the control set points value according to the rules consequent parts. In SWRL, the output has to be single. The SWRL form of the example rule given in Fig. 11 is shown in Fig. 15.

The implementation of the rules is elaborated in the experiments section.

VIII. EXPERIMENTS

The selected pilot zone (an atrium) has an area of 274.9 m2. Hence, the thermal comfort level (PMV) will be kept at ideal conditions. As highlighted in the optimization section, this condition is guaranteed by the generated rules to keep the thermal comfort level measured by PMV in the level of ($-1, 1$).

Generated rules were initially tested on a simulation environment using EnergyPlus. In order to evaluate the performance of the proposed approach, additional rules were also generated using three popular rule generation techniques, namely, C4.5 [25], Rule3, and Rule5 [27]. The antecedent and consequent parts of the rules are selected identical to the ones used in our proposed approach. The number of generated rules by each technique was found as 3830, 4894, 4012, and 3983 with the average accuracy of 98.6%, 76.5%, 92.3%, and 96.3% for the proposed rule generation method, C4.5, Rule3, and Rule5, respectively. An overall comparison, including from an accuracy perspective, between the proposed and above (C4.5, Rule3, and Rule5) algorithms is given in Table V.

Based on results illustrated in Table V, it is worth noting that while the computational time is on average comparable (not significantly different), the accuracy of the proposed algorithm is higher than all other algorithms with the lowest standard deviation. Experiments have initially been carried for 1 day (1st of September), then for three months (October–December) on a simulation environment. The control variables in all base cases were set to 21°, 0, 0 and 1 for temperature set point, windows set point, shading set point, and light set point. The test has been carried out on an Intel Pentium i-5 processor with 6 GB RAM desktop. Experimental results for simulation are given in Figs. 16 and 17.

According to the one day simulation results, total energy consumption has been found as 258.449 KWh, 259.155 KWh, 258.962 KWh, 258.292 KWh, and 201.3719 KWh for the

TABLE V
COMPARISON OF THE PROPOSED ALGORITHM WITH
OTHER RULE GENERATION ALGORITHMS

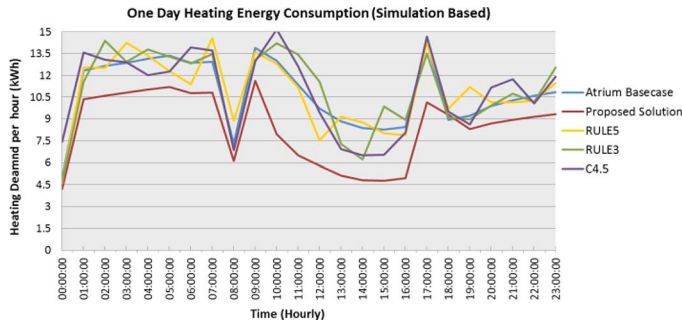| Algorithm	Minimum Accuracy (%)	Maximum Accuracy (%)	Standard Deviation	Average Accuracy (%)	Total Number of Generated Rules	Computational Time (Sec)
C4.5	51.5	100	11.551	76.5	4894	651
Rule3	78.4	100	6.084	92.3	4012	716
Rule5	86.8	100	2.092	96.3	3983	702
The Proposed Method	94.2	100	0.214	98.6	3830	698

Fig. 16. Energy consumption results comparison for the base-case, proposed solution, RULE5, RULE3, and C4.5.
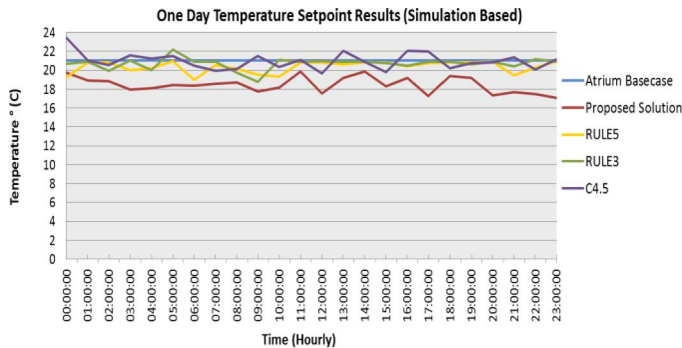

Fig. 17. Comparative analysis of temperature set point results (the base-case, proposed solution, RULE5, RULE3, and C4.5).
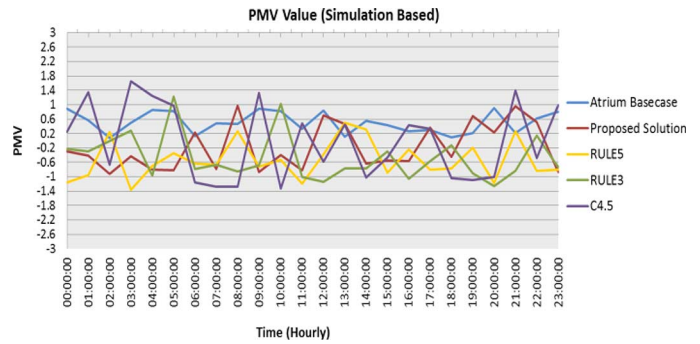

Fig. 18. PMV results comparison using the base-case, proposed solution, RULE5, RULE3, and C4.5 for a selected day.

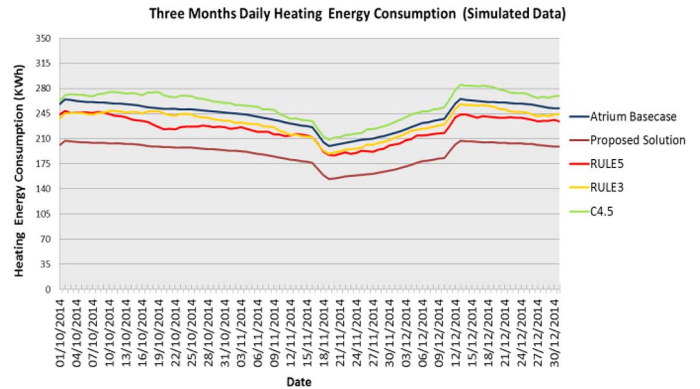Fig. 19. Comparative analysis of energy consumption (the base-case, proposed solution, RULE5, RULE3, and C4.5.).

base-case, C4.5, Rule3, Rule5 and the proposed solution, respectively. As illustrated in Fig. 16, the rules generated with the proposed solution achieve about 22.08% energy saving in one day. The thermal comfort level indicator, PMV, is kept in the range of −1 and 1 with the proposed solution (Fig. 18). The other control variables were found similar to the base-case and the consequent part of the rules generated with the benchmark techniques. Furthermore, the other control variables have been found as similar to the base-case and the consequent part of the rules generated with the benchmark techniques. This establishes the temperature set point as the key control. Therefore, this control variable has been considered as the major control variable while the others were kept constant.

The simulation program was run for three months between 1st of October and 31st of December. The heating energy consumption during this period was repeated for the base-case initially without the proposed rules and a 21 °C fixed temperature set point, and then with the fired rules. The results are illustrated in Fig. 19.

According to the three months simulation results, the total energy consumption in the atrium zone was found as 22327.34 KWh, 23621.271 KWh, 21371.62 KWh, 20605.93 KWh and 17442.33 KWh, for the base-case and the rules generated using C4.5, Rule3, Rule5 and the proposed technique, respectively. In accordance with these results, the proposed methodology allows the user to reduce energy consumption by 21.88%. However, the rules generated using C4.5 increase energy consumption by 5.7% in the simulation environment. During the simulation, 2208 rules have been fired from the proposed rules to achieve this figure. Generated rules have been fired every hour, on a continuous basis. The generated rules were also tested on site. Initially, two consecutive days were selected for comparison, 1st and 2nd of October. In the first day, the energy consumption of the building was measured between 8 am and 6 pm without using any rule. The following day the rules have been implemented on site for the same time scale. Finally, when the energy consumption is measured it has been corrected with outdoor temperature degree days to avoid comparison errors. The results are illustrated in Fig. 20.

Energy consumption of 76.74 KWh and 58.24 KWh have been recorded for the first and second day, respectively. 24.11% heating energy reduction has thus been achieved. Moreover, the analysis has been carried out for two months for both the base-case and rules generated with the proposed solution. The heating energy consumption for both cases is adjusted with degree days and illustrated in Fig. 21.
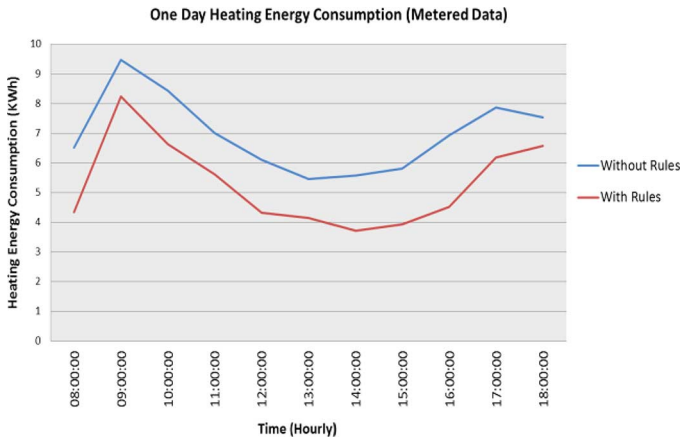
Fig. 20. The real energy consumption comparison in Forum-Atrium (first day without rules, the next day with rules of the proposed solution).
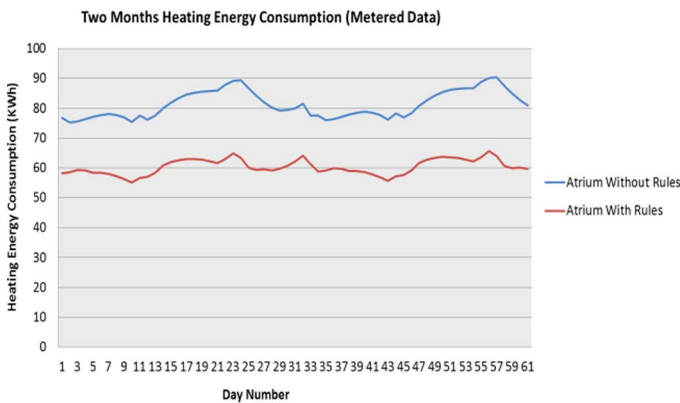

Fig. 21. Real energy consumption comparison (two months without rules, two months with rules).

The two months energy meter reading using the base case reported a heating energy consumption of 7512.57 KWh. For the following two months, the rules have been implemented and degree days corrected; energy consumption was reduced to 5581.30 KWh. The total energy saving within these two months came up to 25.71%.

Based on results from the simulation and real testing cases, our experimentation work reveals an energy reduction of 25% on average, taking into account acceptable comfort conditions for the elderly in the care home.

IX. DISCUSSION

A generic energy saving rule generation approach is proposed, demonstrated and validated on a public building, a care home in The Netherlands. Public buildings represent 25% of useful floor space in Europe, estimated at 6.25 billion m2 [52]. The selected home care provides a comprehensive case which involves: (i) complex user behavior, given the nature of the occupants (older people with health conditions that impose stringent comfort requirements); (ii) interesting building controllable features such as automatic blinds and windows; (iii) energy renewable generation sources; (iv) a wide range of sensing nodes deployed across the various zones of the building; and (v) a state-of-the-art building management system (BMS). These interesting characteristics are now common to
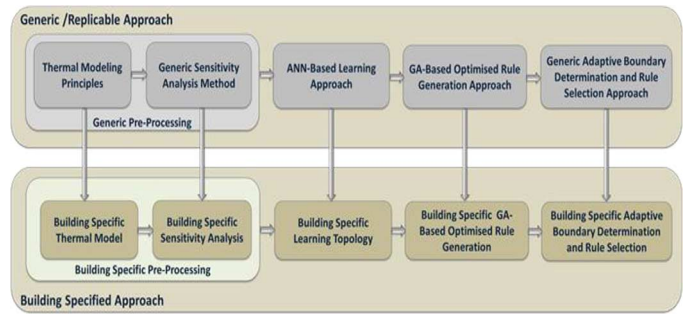

Fig. 22. The generic and specific dimension of the proposed approach.

new and energy retrofitted public buildings and provide an ideal context to develop a generic and scalable approach. Within the selected home care, various scenarios have been identified, following a divide-and-conquer approach, which helps break down the energy management complexity into discrete and manageable use cases. The selected scenario in this paper is generic in nature as it addresses the common objective of minimizing energy consumption, taking into account a wide range of sensed environmental and building variables, while maintaining required comfort conditions. It factors in all the above considerations (i.e., user behavior, building controllable components and systems, rich-sensed environment, and an open BMS).

Although the building case study is specific, the methodology and underpinning algorithms are generic and scalable requiring simple modifications/adaptations, such as determination of the number of input and output and boundary ranges. Further, from a scalability perspective, the proposed approach involves the following generic steps: (i) a preprocessing stage, including thermal energy modeling and sensitivity analysis; (ii) ANN-based learning; (iii) GA-based optimized energy saving rule generation; and (iv) adaptive boundary determination and rule selection. The transition from the generic to the building specific approach is illustrated in Fig. 22.

The preprocessing stage covers thermal modeling and sensitivity analysis. Thermal modeling principles are applied to a specific building to generate the building specific thermal model in line with the building specific features (including geometry, fabric, and occupancy patterns) conveyed by the building ontology. This specific thermal model is then utilized to generate data sets to perform a sensitivity analysis based on generic techniques, namely PCA and MRA, with a view to determine the most sensitive variables given the imposed environmental and building specific conditions. Based on the sensitive variables, a specific topology is determined to deliver an ANN-based learning process. Further, the sensitive variables with the specified trained network are embedded into a generic optimization algorithm with specific control variables and ranges. Moreover, a generic and adaptive boundary determination and rule selection process is introduced. Specific data boundary ranges and rules are then determined.

X. CONCLUSION

A hybrid rule-based energy-saving approach using Artificial Neural Network and Genetic Algorithm is presented.

This is underpinned by a comprehensive semantic referential (building ontology) that provides real-time characterization of the building and its zones. The proposed method utilizes a thermal energy model to identify the building governing variables through a PCA and MRA-based statistical process. The resulting information is used in the ANN to learn energy patterns and form the basis for a GA-based optimization process to reduce energy consumptions taking into account a wide range of objectives (including comfort) and constraints. The generated rules were initially tested on a simulation platform to evaluate their performance through a benchmarking process involving the base-case, C4.5, Rule3, and Rule5. The testing was carried out for four months in the winter term (two months without and two months with the proposed solution) and was thus focused on heating energy consumption. Initial results suggest that the generated rules were able to reduce heating energy consumption by 21.88% while satisfying occupants' (elderly people) comfort conditions. Next, the testing involved real-world experimentation where reduction by 25.71% within two months was recorded.

The rules generated via this methodology are underpinned by a conceptualization of the building (ontology) and form together a building specific energy knowledge base. The latter is routinely updated based on feedback via the installed sensors. Future development for the rule-based process is to utilize other optimization algorithms such as the Bees Algorithm; Ant Colony Technique, and Particle Swarm Optimization, to increase the performance of the generated rules. An enhanced fuzzy boundary determination process will also be explored to generate a rapid solution to adjust the boundaries of the generated rules. It is critical to note here that the number of generated rules can be reduced by proposing more generic rules which may cover more than one time step. This process will be explored by analyzing the historical data and implementing adapted data mining techniques such as enhanced fuzzy clustering approaches.

References

[1] C. A. Balaras, K. Droutsa, E. Dascalaki, and S. Kontoyiannidis, "Heating energy consumption and resulting environmental impact of European apartment buildings," *Energy Build.*, vol. 37, no. 5, pp. 429–442, Aug. 2004.

[2] L. Perez-Lombard, J. Ortiz, and C. Pout, "A review on building energy consumption information," *Energy Build.*, vol. 40, no. 3, pp. 394–398, Mar. 2007.

[3] "Directive 2002/91/EC of the European parliament and of the council of 16 December 2002 on the energy performance of buildings," *Official J. Eur. Communities*, pp. L 65–71, Dec. 2010.

[4] B. Yuce, H. Li, Y. Rezgui, I. Petri, B. Jayan, and C. Yang, "Utilizing artificial neural network to predict energy consumption and thermal comfort level: An indoor swimming pool case study," *Energy Build.*, vol. 80, pp. 45–56, Sep. 2014.

[5] B. Sun, P. B. Luh, Q.-S. Jia, Z. O'Neill, and F. Song, "Building energy doctors: An SPC and Kalman filter-based method for system-level fault detection in HVAC systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 11, no. 1, pp. 215–229, Jan. 2004.

[6] F. Amara, K. Agbossou, A. Cardenas, Y. Dubé , and S. Kelouwani, "Comparison and simulation of building thermal models for effective energy management," *Smart Grid Renewable Energy*, vol. 6, pp. 95–112, Apr. 2015.

[7] H. Y. Huang, J. Y. Yen, S. L. Chen, and F. C. Ou, "Development of an intelligent energy management network for building automation," *IEEE Trans. Autom. Sci. Eng.*, vol. 1, no. 1, pp. 14–25, Jul. 2004.

[8] C. H. Lo, P. T. Chan, Y. K. Wong, A. B. Rad, and K. L. Cheung, "Fuzzy-genetic algorithm for automatic fault detection in HVAC systems," *Appl. Soft Comput.*, vol. 7, no. 2, pp. 554–560, Mar. 2007.

[9] K. L. Ku, J. S. Liaw, M. Y. Tsai, and T. S. Liu, "Automatic control system for thermal comfort based on predicted mean vote and energy saving," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 1, pp. 378–383, Jan. 2015.

[10] B. Qela and H. Mouftah, "Observe, learn, and adapt (OLA)—An algorithm for energy management in smart homes using wireless sensors and artificial intelligence," *IEEE Trans. Smart Grid*, vol. 3, no. 4, pp. 2262–2272, Dec. 2012.

[11] M. Hadjiski, V. Sgurev, and V. Boishina, "HVAC control via hybrid intelligent systems," *Bulgarian Acad. Sci., Cybern. Inf. Technol.*, vol. 7, no. 1, pp. 77–94, 2007.

[12] C. Yang, H. Li, Y. Rezgui, I. Petri, B. Yuce, B. Chen, and B. Jayan, "High throughput computing based distributed genetic algorithm for building energy consumption optimization," *Energy Build.*, vol. 76, pp. 92–101, Jun. 2014.

[13] F. Doctor, H. Hagras, and V. Callaghan, "A fuzzy embedded agent-based approach for realizing ambient intelligence in intelligent inhabited environments," *IEEE Trans. Syst., Man, Cybern., Part A: Syst. Humans*, vol. 35, no. 1, pp. 55–65, Jan. 2005.

[14] L. D. Ha, S. Ploix, E. Zamai, and M. Jacomino, "Tabu search for the optimization of household energy consumption," in *Proc. IEEE. Inf. Reuse Integr. Conf.*, Waikoloa Village, HI, USA, Sep. 2006, pp. 86–92.

[15] Z. Liao and A. L. Dexter, "An inferential model-based predictive control scheme for optimizing the operation of boilers in building space-heating systems," *IEEE Trans. Control Syst. Technol.*, vol. 18, no. 5, pp. 1092–1102, Sep. 2010.

[16] M. Dibley, H. Li, Y. Rezgui, and J. Miles, "An ontology framework for intelligent sensor-based building monitoring," *Autom. Constriction*, vol. 28, pp. 1–14, Dec. 2012.

[17] H. Wicaksono, S. Rogalski, and E. Kusnady, "Knowledge-based intelligent energy management using building automation system," in *Proc. IEEE IPEC Conf.*, Singapore, Oct. 2010, pp. 1140–1145.

[18] J. Han, Y. K. Jeong, and I. Lee, "Building energy ontology for energy saving based on context-aware reasoning," in *Proc. Int. Conf. Inf. Knowl. Eng. (IKE). Steering Committee World Congr. Comput. Sci., Comput. Eng., Appl. Comput.*, LasVegas, NV, USA, Jul. 2013, pp. 1–2.

[19] Y. Rezgui, S. Boddy, M. Wetherill, and G. Cooper, "Past, present and future of information and knowledge sharing in the construction industry: Towards semantic service-based e-construction?," *Comput.-Aided Des.*, vol. 43, no. 5, pp. 502–515, May 2011.

[20] P. Lord, "The semantic web takes wing: Programming ontologies with Tawny-OWL," in *Proc. 10th Int. Workshop OWL: Experiences and Directions (OWLED'13) Co-Located With 10th Extended Semantic Web Conf. (ESWC'13)*, Montpellier, France, May 2013, vol. 1080.

[21] A. V. Zhdanova and U. Keller, "Choosing an ontology language," *WASET, Int. J Comput., Electr., Autom., Control, Inf. Eng.*, vol. 1, no. 4, pp. 996–999, Apr. 2007.

[22] X. Luo, Z. Xu, J. Yu, and X. Chen, "Building association link network for semantic link on web resources," *IEEE Trans. Autom. Sci. Eng.*, vol. 8, no. 3, pp. 482–494, Jul. 2011.

[23] M. Zakova, P. Kremen, F. Zelezny, and N. Lavrac, "Automating knowledge discovery workflow composition through ontology-based planning," *IEEE Trans. Autom. Sci. Eng.*, vol. 8, no. 2, pp. 253–264, Apr. 2011.

[24] Y. Gao, E. Tumwesigye, B. Cahill, and K. Menzel, "Using data mining in optimisation of building energy consumption and thermal comfort management," in *Proc. IEEE 2nd Int. Conf. Softw. Eng. Data Mining (SEDM)*, Chengdu, China, Jun. 2010, pp. 434–439.

[25] J. R. Quinlan, "Decision trees and decision-making," *IEEE Trans. Syst., Man, Cybern.*, vol. 20, no. 2, pp. 339–346, Mar. 1990.

[26] R. Setiono, W. K. Leow, and J. M. Zurada, "Extraction of rules from artificial neural networks for nonlinear regression," *IEEE Trans. Neural Netw.*, vol. 13, no. 3, pp. 564–577, May 2002.

[27] D. T. Pham, S. Bigot, and S. Dimov, "A rule merging technique for handling noise in inductive learning," *Proc. Inst. Mech. Eng., Part C: J. Mech. Eng. Sci.*, vol. 218, no. C, pp. 1255–1268, Oct. 2004.

[28] D. T. Pham and A. A. Afify, "RULES-6: A simple rule induction algorithm for supporting decision making," in *Proc. IEEE Ind. Electron. Soc. Conf. (IECON)*, Nov. 2005, pp. 2184–2189.

[29] J. Niu, Y. He, M. Li, and X. Zhang, "A comparative study on application of data mining technique in human shape clustering: Principal component analysis vs. factor analysis," in *Proc. IEEE 5th Ind. Electron. Appl. (ICIEA) Conf.*, Taichung, Taiwan, Jun. 2010, pp. 2014–2018.

[30] J. R. Rabunal, J. Dorado, A. Pazos, J. Perreira, and D. Rivero, "A new approach to the extraction of ANN rules and to their generalization capacity through GP," *Neural Comput.*, vol. 16, no. 7, pp. 1483–1523, Jul. 2004.

[31] N. R. Pal and S. Chakraborty, "Fuzzy rule extraction from ID3-type decision trees for real data," *IEEE Trans. Syst., Man, Cybern., Part B: Cybern.*, vol. 31, no. 5, p. 745,754, Oct. 2001.

[32] W. Duch, R. Adamczak, and K. Grabczewski, "A new methodology of extraction, optimization and application of crisp and fuzzy logical rules," *IEEE Trans. Neural Netw.*, vol. 12, no. 2, pp. 277–306, Mar. 2001.

[33] J. Ponni and K. L. Shunmuganathan, "Multi-agent system for data classification from data mining using SVM," in *Proc. IEEE 5th Green Comput., Comm., Conservation Energy (ICGCE) Conf.*, Chennai, India, Dec. 2013, pp. 828–832.

[34] S. Wang and J. Xie, "Integrating building management system and facilities management on the internet," *Autom. Construction*, vol. 11, no. 6, pp. 707–715, Oct. 2002.

[35] H. Doukas, K. D. Patlitzianas, K. Iatropoulos, and J. Psarras, "Intelligent building energy management system using rule sets," *Build. Environ.*, vol. 42, no. 10, pp. 3562–3569, Oct. 2007.

[36] Z. Xu, Q. S. Jia, X. Guan, and J. Shen, "Smart management of multiple energy systems in automotive painting shop," *IEEE Trans. Autom. Sci. Eng.*, vol. 10, no. 3, pp. 603–614, Jul. 2013.

[37] Y. Z. Lu, *Industrial Intelligent Control: Fundamentals and Applications*. West Sussex, U.K.: Willey, 1996.

[38] B. M. Flax, "Intelligent buildings," *IEEE Commun. Mag.*, vol. 29, no. 4, pp. 24–27, Aug. 1991.

[39] H. Grzybek, S. Xu, S. Gulliver, and V. Fillingham, "Considering the feasibility of semantic model design in the built-environment," *Buildings*, vol. 4, no. 4, pp. 849–879, Nov. 2014.

[40] R. Setiono, B. Baesens, and C. Mues, "Recursive neural network rule extraction for data with mixed attributes," *IEEE Trans. Neural Netw.*, vol. 19, no. 2, pp. 299–307, Feb. 2008.

[41] M. Sato and H. Tsukimoto, "Rule extraction from neural networks via decision tree induction," in *Proc. Int. Joint Conf. Neural Netw.*, Washington, DC, USA, Jul. 2001, vol. 3, pp. 1870–1875.

[42] M. Setnes, "Supervised fuzzy clustering for rule extraction," in *Proc. IEEE Fuzzy Syst. Conf.*, Seoul, South Korea, Aug. 1999, vol. 3, pp. 1270–1274.

[43] N. R. Pal and S. Chakraborty, "Fuzzy rule extraction from ID3-type decision trees for real data," *IEEE Trans. Syst., Man, Cybern., Part B: Cybern.*, vol. 31, no. 5, pp. 745–754, Oct. 2001.

[44] B. Soua, A. Borgi, and M. Tagina, "Attributes regrouping in fuzzy rule based classification systems," in *Proc. IEEE Signals, Circuits Syst. (SCS) Conf.*, Nov. 2009, pp. 1–6.

[45] J. Qiu, H. Wang, J. Lu, B. Zhang, and K. L. Du, "Neural network implementations for PCA and its extensions," *ISRN Artif. Intell.*, pp. 1–19, Jun. 2012.

[46] K. Deng, Y. Sun, S. Li, Y. Lu, J. Brouwer, P. G. Mehta, M. C. Zhou, and A. Chakraborty, "Model predictive control of central chiller plant with thermal energy storage via dynamic programming and mixed-integer linear programming," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 2, pp. 565–579, Apr. 2015.

[47] P. Agrawal and S. Rao, "Energy-aware scheduling of distributed systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 11, no. 4, pp. 1163–1175, Oct. 2014.

[48] S. M. Namburu, M. S. Azam, J. Luo, K. Choi, and K. R. Pattipati, "Data-driven modeling, fault diagnosis and optimal sensor selection for HVAC chillers," *IEEE Trans. Autom. Sci. Eng.*, vol. 4, no. 3, pp. 469–73, Jul. 2007.

[49] M. B. Maasoumy and A. Sangiovanni-Vincentelli, "Total and peak energy consumption minimization of building HVAC systems using model predictive control," *IEEE Des. Test Comput.*, vol. 29, no. 4, pp. 26–35, Aug. 2012.

[50] I. Petri, H. Li, Y. Rezgui, C. Yang, B. Yuce, and B. Jayan, "A HPC based cloud model for real-time energy optimisation," *Enterprise Inf. Syst.*, pp. 1–21, May 2014.

[51] D. Kolokotsa, G. S. Stavrakakis, K. Kalaitzakis, and D. Agoris, "Genetic algorithms optimized fuzzy controller for the indoor environmental management in buildings implemented using PLC and local operating networks," *Eng. Appl. Artif. Intell.*, vol. 15, pp. 471–428, Sep. 2002.

[52] B. Atanasiu, C. Despret, M. Economidou, J. Maio, I. Nolte, and O. Rapf, *Europe's Buildings under the Microscope: A Country-by-Country Review of the Energy Performance of Buildings*. Brussel, Belgium: BPIE, 2011.

Baris Yuce (M'15) received the B.Sc. and M.Sc. (First Class Hons.) degrees in engineering (industrial) from Sakarya University, Sakarya, Turkey, and the Ph.D. degree from Cardiff University, Cardiff, U.K., in 2012.

He worked as a Teaching and Research Assistant at the Industrial Engineering Department, Sakarya University, from 2006 to 2008. After completing his Ph.D., he work as a Research Associate at the School of Computer Science, Cardiff University, in 2012. In 2013, he joined the BRE Centre in Sustainable Engineering, School of Engineering, Cardiff University. His research interest focuses on intelligent systems such as artificial neural network, fuzzy logic, multiagent systems and their applications on buildings, energy management systems, robotics and supply chain management. He is also interested in optimization algorithms and their applications, in particular, swarm optimization algorithms such as the bees algorithm, genetic algorithm, ant colony optimization, hill climbing, and simulated annealing. He has published several papers on intelligent systems and their applications in the area of smart buildings, energy management, optimization, supply chain management, building management systems, and scheduling. He has worked on several EU and (Building Research Establishment (BRE) funded projects on intelligent building management systems including SPORTE2, KnoholEM, and Estate Energy Optimization.

Dr. Baris is a member of the IEEE Robotics and Automation Society and the IEEE Power and Energy Society.

Yacine Rezgui received the M.Sc. degree in architecture engineering from University Paris 6, Paris, France, and the Ph.D. degree in construction informatics from Ecole Nationale des Ponts et chaussees, Paris.

He is a Professor of Building Systems and Informatics, Cardiff University, Cardiff, U.K., and the Founding Director of the BRE Centre in Sustainable Engineering, sponsored by the Building Research Establishment (BRE), U.K. In 1995, he joined Salford University, U.K., as a Research Fellow, then Academic Lecturer (1996), Senior Lecturer (1998), and Professor (2001). He was the Founding Director (2003–2007) of the 5* (RAE 2001) rated Informatics Research Institute, Salford University. He has successfully completed over 40 research and development projects at a national (UK EPSRC and TSB) and international (European Framework Programs 5, 6, 7) level. He conducts research in informatics, including ontology engineering and artificial intelligence applied to the built environment.

Prof. Rezgui has over 150 refereed publications in the above areas, which appeared in international journals such as the IEEE TRANSACTIONS ON SERVICES COMPUTING, *Information Sciences*, *Data and Knowledge Engineering*, and *Computer-Aided Design*. He is also a qualified Architect.