Energy and Buildings 125 (2016) 45–54

Contents lists available at ScienceDirect

Energy and Buildings

journal homepage: www.elsevier.com/locate/enbuild

Application of neural networks for evaluating energy performance certificates of residential buildings

Fazel Khayatian, Luca Sarto, Giuliano Dall'O' *

Department of Architecture, Built Environment and Construction Engineering, Polytechnic of Milan, Milan, Italy

ARTICLE INFO

Article history:
Received 14 December 2015
Received in revised form 24 March 2016
Accepted 23 April 2016
Available online 30 April 2016

Keywords:
Energy performance certificate (EPC)
Artificial neural networks (ANN)
Energy certificate validation
Residential buildings

ABSTRACT

The Energy Performance Building Directive 91 of 2002, mandates Member States of the European Union to enforce energy certification of buildings through local legislation. Among the Italian regions, Lombardy has issued predicted energy performance certificates for buildings since 2007 which accumulate to over one million entries. The current study is an attempt to validate a dataset of energy certificates by benefitting from the magnitude of registered buildings. Considering that manual evaluation of every entry is exhaustive and time consuming, artificial neural network is used as a fast and robust alternative for predicting heat demand indicators. Various combinations of input features are compared for selecting a reliable model. The number of inputs and hidden neurons are also optimized in order to achieve better accuracy. Results show that using 12 variables from an energy certificate is sufficient for estimating the related heat demand indicator. Regarding the stochastic initialization of neural networks, a set of 100 models are trained for obtaining a frequency distribution and confidence interval. Final results indicate that about 95% of entries fall within ±3 confidence intervals.

1. Introduction

According to the Energy Performance Building Directive (EPBD), all member states of the European Union are obligated to pursue energy efficiency guidelines in buildings by implementing national regulations [1]. The Lombardy Region was the first in Italy to enforce the adapted national directive (Legislative Decree No. 192) [2] into regional regulations (Regional Law No. 24of12/11/06) [3]. As a result, the Certificazione Energetica degli Edifici (CENED) tool was developed in order to provide predicted Energy Performance Certificates (EPC) according to national and European standards [4,5]. Upon the request of a property owner, the procedure for building energy audit will be handled by a certified energy assessor through the CENED tool. The data implemented in energy labels consist of a great potential for statistical analysis of the energy efficiency trend in Lombardy region. However, the reliability of the self-declared certificates are unknown and therefore a validation study on the assigned EPCs merits to be investigated.

In order to promote transparency, the administration of the Lombardy region has provided an online database, from which all the predicted EPCs can be freely accessed. However, evaluating the declared predictions is challenged by two obstacles. First, the online database consists of more than one million entries from all over the Lombardy region. Therefore, the predictive model should be able to process large data in a reasonable amount time. Second, the data fed into the CENED software greatly differ from the ones provided online. The data regarding each certificate that is provided to public is only a sparse representation of all the data used in the calculation procedure. Therefore, the predictive model should be able to extract complex information from simple input data. Although multivariate regression is a useful tool for addressing similar problems, the performance of the model is greatly affected by the configuration of the features. When dealing with several variables, the procedure of manually defining suitable nonlinear features is exhaustive. In this particular study, the logic behind using Artificial Neural Network (ANN) for validating predicted EPCs is twofold. First, ANNs perform well with limited input data as they can extract sophisticated features from simple inputs, without the necessity of manually creating and/if/or logics. Moreover, the performance of ANNs is greatly affected by the number of samples which is a great potential for evaluating the intended database. The mentioned benefits along with its reliability and speed, justifies the application of ANN for evaluating predicted EPCs in large datasets.

* Corresponding author.
E-mail address: giuldal@polimi.it (G. Dall'O').

http://dx.doi.org/10.1016/j.enbuild.2016.04.067
0378-7788/

Nomenclature

Abbreviation

ANN	Artificial neural network
DEC	Display energy certificate
EPBD	Building energy performance directive
EPC	Energy performance certificate
MSE	Mean squared error
MAPE	Mean absolute percentage of error
MARE	Mean absolute relative error
CENED	Certificazione energetica degli edifici

Symbol

A_n	Net floor area (m^2)
CP	Construction period
CY	Construction year
DD_W	Winter degree days
EP_H	Heating energy indicator (kWh/m^2y)
ET_H	Heat demand indicator (kWh/m^2y)
F	Hybrid feature
A_H	Average height (m)
N_f	Number of flats
V_n	Net volume area (m^3)
S_D	Dispersant surface (m^2)
S_O	Opaque surface (m^2)
S_G	Glazed surface (m^2)
U_e	Average U-value of walls (W/m^2K)
U_r	Average U-value of roof (W/m^2K)
U_w	Average U-value of windows (W/m^2K)
U_b	Average U-value of basement (W/m^2K)

The CENED tool labels buildings within 8 different classes ranging from A+ to G. These energy classes represent the predicted energy efficiency of buildings based on an estimation of the building's energy consumption. However, the criteria for appointing an energy class varies among residential and non-residential buildings as well as their respective climatic zone [6]. While the required energy for heating, cooling, lighting and domestic hot water are separately calculated in the CENED software, the energy class is solely determined based on the heating energy consumption (EP_H). This issue is addressed in a recent update to the CENED software where all the mentioned loads are considered in the calculated energy performance index. However, the entire set of energy certificates collected at the time of the study are issued based on the previous version of CENED software. As a result, this study tends to validate self-declared certificates by focusing on the respective EP_H value.

2. Methodology

2.1. Machine learning and artificial neural networks

The problem of building energy prediction may be addressed by using different approaches. However, it is possible to divide all applied methods into three major categories: white-box methods (engineering approaches), black-box methods, (machine learning approaches) and grey-box methods (hybrid engineering/machine learning approaches) [7]. Each of the mentioned categories consist of specific advantages, yet extracting outputs from complex engineering models may be exhaustive and time consuming. The objective of using machine learning in prediction problems is to simulate a model based on observed data in order to estimate new outputs. Such approach provides the opportunity to perform predictions on sophisticated models in a reasonable amount of time.

In order to predict new outputs by using machine learning, a series of samples are required. These samples are individual observations of the intended model. Each sample consists of some inputs (features) and one or more outputs. The samples are used for training the machine learning model in order to find a relation between the inputs and targets. Studies have noted that by applying a correct machine learning algorithm and collecting sufficient samples, it is possible to replicate almost any desired model regardless of the level of complexity [8]. The application of machine learning in predicting building energy performance has been widely investigated. Studies have attempted to predict heating [9], cooling [10] and electric energy consumption [11] of buildings. Various machine learning algorithms have been applied to building energy prediction problems [12], and the extent of studies span from predicting annual energy demand [13] to hourly energy consumption [14].

Artificial neural networks (ANN) is a subcategory of machine learning which has repeatedly displayed reliable performances in various estimation problems. ANNs are frequently used to address building energy prediction problems. The application of ANN in different fields of building energy estimation cover a wide range of categories: building energy benchmarking [15–17], electric energy prediction [18–20] and heating/cooling loads prediction [21–23]. The concept of ANN is inspired by the behavior of biological neurons inside the brain, where each biological neuron is represented by a mathematical node also known as an artificial neuron. Generally, ANNs are divided into input layer, hidden layer and output layer (Fig. 1). All the layers of a simple neural network are fully connected, and therefore each artificial neuron is connected to all the values of the adjacent layer. An input layer is a matrix of all the entries and their features. The output layer is also a matrix which contains the intended values corresponding to each sample. Between the input and output layers, there is a middle layer (hidden layer) which consists of artificial neurons. Each neuron is composed from two main parts: the *"activation function"* and the *"transfer function"* (Fig. 1). The activation function contains weight and bias values from which the output is estimated. The transfer function is responsible for propagating the calculated values to the next layer.

In order to train a neural network, a measure of the network's performance is required. This measure can be calculated in different ways i.e. mean squared error (MSE), mean absolute percentage of error (MAPE), mean absolute relative error (MARE) and etc. This error which represents the accuracy of the network can be reduced by utilizing a minimization algorithms. Among various algorithms for measuring the performance of ANNs, back propagation algorithms have displayed promising performances [24]. In this method, along with the overall error of the output, the performance of each neuron is also calculated. Therefore, ANNs trained with back-propagation algorithm often display better performances compared to simple feed-forward ANNs. There are numerous optimization techniques for updating the weights of hidden neurons. However in this study, the speed of the algorithm is a priority when considering the magnitude of the dataset. The Levenberg-Marquardt algorithm is a robust optimization method and stands amongst the fastest techniques for updating weights in ANNs. Therefore, the ANN implemented in this study uses the Levenberg-Marquardt optimization algorithm for updating the weights of the hidden neurons.

2.2. Variable selection in energy performance estimation

Energy performance assessment of buildings follow two main objectives: energy classification and energy performance diagnostics [25]. The application of ANN in determining the energy use in Display Energy Certificates (DEC) has been studied in [26]. The study uses various combinations of inputs for estimating the actual

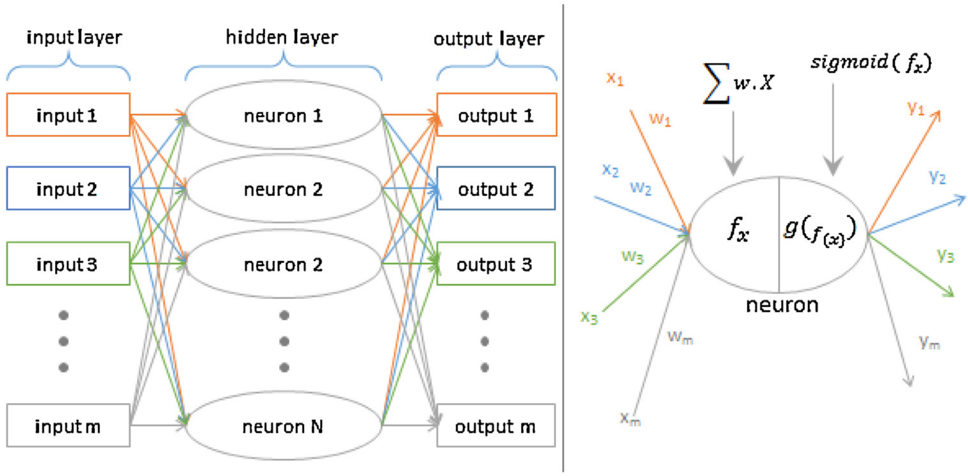

Fig. 1. Graphical representation of a neural network (left) and artificial neuron (right).

Table 1
Classification of utilized features.

Category	Features
Direct	degree days, net volume, net floor area, dispersant surface, opaque to glazed ratio, year of construction, thermal conductivity (walls, windows, roof, basement)
Derived	average floor height, opaque surface area, glazed surface area, construction period, non-linear features

energy consumption of university buildings and recommends similar applications in related problems. Melo et al. compared the performance of linear regression and ANN for predicting building energy labels which are assigned through a dynamic energy simulation [27]. The study achieved an acceptable error of ±16% over 90% of samples and recommended the application of ANN for complex energy labeling models. Buratti et al. used ANN in order to predict energy labels based on 2000 energy certificates which are issued by the Umbria Region in Italy [28]. The study used abstract information regarding the characteristics of each sample in order to replicate a quasi-steady-state approximation. Five different models with various numbers of features (12, 13, 15, 18 and 21) were studied. The preference for choosing the most suitable model was based on a tradeoff between performance and computation time. The selected neural network displayed an acceptable regression of 0.95 before optimization, along with a mean absolute error of 16.03 kWh/m^2 y while using 12 features and 50 neurons.

The online CENED database consists of various types of buildings based on the respective destination of use that is the type of occupancy. In case of residential buildings, the property may be audited by adopting one of the following approaches. One possible method is to request an energy certificate for a single residential unit i.e. a single apartment or flat. Another possibility is to request an energy certificate for the whole building regardless of the number of residential units. The current study is aimed at whole residential buildings as they share the largest portion of certificates among various types of building. The procedure for selecting only whole residential buildings was completed in two steps. First, entries labeled as E.1(1) under the category of "Destination of Usage" were extracted from the database. Afterwards, filters were applied to the 4 categories of thermal conductivity i.e. walls, windows, roof and basement. Entries with a thermal conductivity of zero in any of the mentioned categories were removed from the dataset.

The procedure for calculating the EP_H is a quasi-steady-state approximation described in [29,30]. The EP_H is obtained based on complex calculation of numerous parameters which can be divided into two main categories: the envelope characteristics and the plant specifications. Among the vast number of envelop characteristics for each building, a few parameters are available through the online database. These parameters include: floor area, volume, number of flats, dispersant area, window to wall ratio and four averaged thermal conductivity values for the categories of walls, roof, basement and windows. In case of the plant characteristics, the provided data is limited to: renewable energies and the area of installation, heating systems and their respective fuels, and mechanical ventilation along with the number of air changes per hour. However, the plant characteristics cannot be easily implemented in the ANN training data. This issue is related to simultaneous use of various heating systems as well as different renewable energies in a single building. Some samples consist of two or more heating systems with different characteristics while others have various types of photovoltaic systems with different efficiencies. Regarding that all input vectors should consist of the same number of features, it is not possible to implement all plant characteristics into the input dataset. Meanwhile, the online database does not provide any description regarding the share of each plant in every building. Therefore, using features that are representative of the plant characteristics is not possible.

As a result, the output of the training set was based on the Heat Demand Indicator (ET_H) rather than the EP_H. This value represents the heat demand of a building based on envelope characteristics and is calculated regardless of plant specifications and renewable energy systems. Table 1 describes the features which are used for predicting the ET_H value in this study. These features are separated into "direct" and "derived" categories. The first category consists of features that are extracted from the CENED database and implemented without any changes. The only exception is the winter degree days (DD_W) which was not available through the online database and was extracted from [31]. The second category contains features which are extracted from the database and modified before implementation in the model. The application of the second category is an attempt to achieve better weight distribution by using more complex input data. Combinations of both modified and unmodified features are used in different training models.

All hybrid features are developed by applying simple mathematical operations on the raw input data. Average height (A_H) is

Table 2
Classification of construction groups based on [30].

Construction Period	Acronym	Binary value
Before 1930	CP_1	1000000
1930–1945	CP_2	0100000
1946–1960	CP_3	0010000
1961–1976	CP_4	0001000
1977–1992	CP_5	0000100
1993–2006	CP_6	0000010
After 2007	CP_7	0000001

Table 3
Values of parameters utilized for training the ANN.

Parameter	Value
Maximum number of iterations	3000
Early stop validation errors	100
Training set portion	0.8
Validation set portion	0.1
Test set portion	0.1

extracted from unmodified features net volume (V_n) and net floor area (A_n) as displayed in the following:

$$A_H = V_n/A_n \quad (1)$$

Opaque surface (S_O) and glazed surface (S_G) are calculated by using dispersant surface (S_D) and glass to opaque surface ratio (GO_g) which are available from the database. Eqs. (2) and (5) display the calculation procedure of (S_O) and (S_G).

$$S_O + S_G = S_D \quad (2)$$

$$\frac{S_G}{S_O} = GO_g \quad (3)$$

Therefore by combining Eq. (1) and Eq. (2) we conclude that

$$S_O = \frac{S_D}{1 + GO_g} \quad (4)$$

$$S_G = S_D - S_O \quad (5)$$

Construction year (CY) is replaced with binary values regarding the period of construction (Table 2). Consequently, each input belongs to a specific group of construction period (CP).

The non-linear features are mainly complex representations of direct and derived features. Eqs. (6) and (11) display the calculation procedure of these hybrid features.

$$F_1 = (S_O)^2 \times DD_W \quad (6)$$

$$F_2 = (S_G)^2 \times DD_W \quad (7)$$

$$F_3 = (U_e)^2 \times S_O \quad (8)$$

$$F_4 = (U_r)^2 \times S_O \quad (9)$$

$$F_5 = (U_b)^2 \times S_O \quad (10)$$

$$F_6 = (U_w)^2 .S_G \quad (11)$$

2.3. Model implementation

A neural network is created in MATLAB software using predefined codes and some modifications in the training parameters (Table 3). The initial dataset extracted from the online database consisted of nearly 220000 energy labels for whole residential buildings. However a pre-processing phase was necessary to ensure the reliability of the input and target data. Neural networks adapt the weights of their hidden neurons based on the input and target data. Therefore, inclusion of defective data in the training set will alter the mean errors which are propagated back for weight optimization and result in inaccurate predictions. In this regard, a series of filters were applied to the dataset by using the conditions mentioned in Table 4. These filters were adopted from a recent study conducted on the same database [32]. The application of filters on the EPC characteristics resulted in removing about 32000 suspicious entries and consequently reducing the size of the dataset to 187587 entries.

Table 4
Validation conditions for sample properties [30].

Filter	Unit	Condition
ET_H	kWh/m^2 y	0 < value < 1000
U_e, U_r, U_b	W/m^2K	0 < value < 4
U_w	W/m^2K	0 < value < 6
Average U	W/m^2K	0.15 < value < 4
A_n	m^2	50 < value
V_n	m^3	150 < value
A_H	m	2.4 < value
S_G	m^2	1 < value
GO_g	–	0 < value < 0.9

Table 5
Properties of various training sets and their performance.

Model	Vector size	Features	MAPE	R
A	8	DD_W A_H S_O S_G U_e U_r U_w U_b	21.98	0.89
B	10	DD_W A_H A_n V_n S_D GO_g U_e U_r U_w U_b	15.66	0.93
C	18	DD_W N_f A_n V_n A_H S_O S_G U_e U_r U_w U_b CP	13.37	0.96
D	20	DD_W N_f A_n V_n A_H S_O S_G U_e U_r U_w U_b CP F_1 F_2	13.35	0.96
E	22	DD_W N_f A_n V_n A_H S_O S_G U_e U_r U_w U_b CP F_3 F_4 F_5 F_6	13.28	0.97

Various models with different combinations of input vectors are tested in order to develop a network which displays acceptable performance in a reasonable amount of time. All the models are trained with N number of neurons (where N is the number of features), while the target is always the ET_H value. The behavior of ANNs may be compared based on some performance indexes i.e. regression (R) and MSE. Monitoring the R value is one of the most common methods for comparing the obtained results and a perfect prediction, where R = 1 is considered as 100% accuracy. On the other hand, MSE describes a quantitative measure of the average prediction error over the entire dataset. However, in this particular study MSE is not a suitable measure as the range of ET_H values span from 1.7 to 989 kWh/m^2 y. MSE provides a scale dependent value which results in difficult interpretation of the actual performance. For instance, observing a deviation of 10 kWh/m^2 y in the prediction output corresponds to 0.011% prediction error for a sample with an ET_H of 900 kWh/m^2 y. Meanwhile, the same value will return 100% prediction error while dealing with an ET_H of 10 kWh/m2 y. Therefore, MAPE is considered as a more suitable measure for performance comparison as it is easier to interpret.

3. Results and discussion

3.1. Analysis of the trained neural networks

A total of 5 models (A–E) with different input vectors are compared in this study. Each model is trained 5 times and the best performance is documented. Table 5 displays the input features, vector size and performance of each model. It should be noted that the number of features may differ from the vector size. This difference is related to feature CP which is a binary vector with 7 inputs. Initially, more simple input vectors with less features are investigated. Each successive model is more complex compared to the previous one regarding the quantity and type of features. Model A consists of 8 features, 5 of which are extracted directly

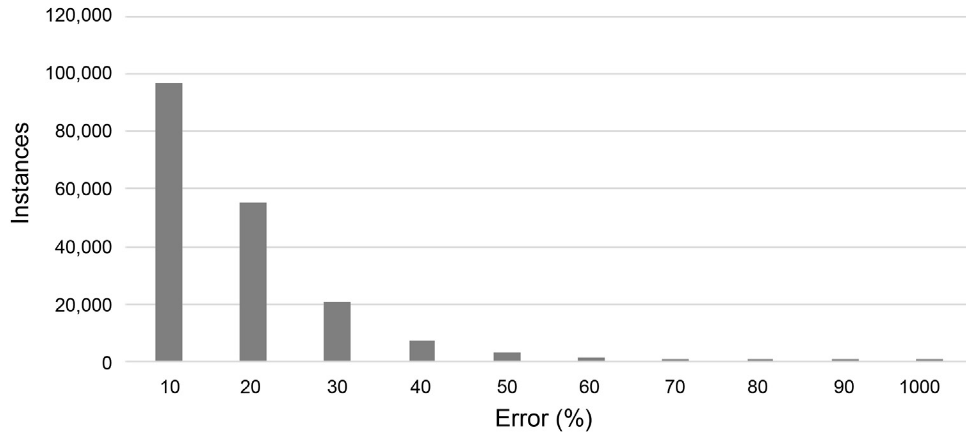

Fig. 2. Frequency of errors in model C (errors larger than 100% are neglected).

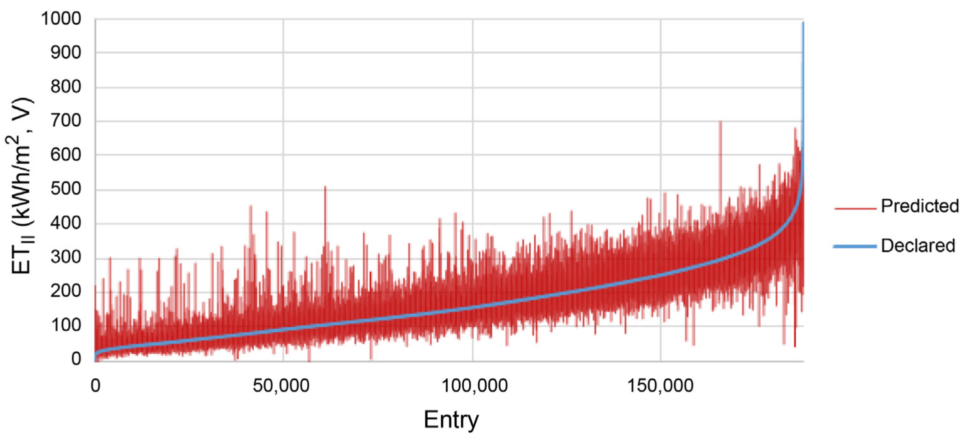

Fig. 3. Comparison of declared and predicted ET_H values (entries are sorted from small to large based on the magnitude of ET_H value).

from the database. The other three features H, S_O and S_G are hybrid inputs which are described in Eq. (1) to Eq. (3). Results from the model's performance indicate that 8 features will return a MAPE of 21.98. In order to increase the effectiveness of the network, other features such as net area and net volume are added to the input vector (model B). This model with 10 features displays a better performance (MAPE 15.66) compared to model A. Observing a 30% reduction in the MAPE indicates that the added features are vital for improving the performance of the network. Model C is constructed by including more complex features to the vector of the previous model (model B). Such modifications reduced the MAPE to 13.37, while increasing the overall regression to 0.96. In order to evaluate the possibility of reaching a better performance, hybrid features are added to the next models. Model D consists of 2 non-linear features F_1 (Eq. (6)) and F_2 (Eq. (7)) compared to the previous network. However, it is observed that the MAPE and overall regression do not experience a considerable change. Other complex features acquired from Eq. (8) to Eq. (11) are tested in model E. However, results indicated that increasing the number of features from 12 to 16 (model C to model E) does not greatly affect the performance of the model. Meanwhile, the computation time for model E (132 min average) is noticeably larger compared to model C (97 min average). Considering that the performance of models C–E are very close, and the fact that increasing the number of features results in higher computational cost, model C with 12 features and a vector size of 18 is considered as the optimum choice for implementing into the evaluation procedure.

Model C demonstrates a good performance as 80% of all the samples return an MAPE less than 20 (Fig. 2). Yet, the magnitude of error in a small portion of entries remains a concern. It is observed that the trained network returned an error above 100% for about 900 entries which corresponds to 0.5% of the entire dataset. Regarding this high value of error in rare instances, the relation between declared ET_H values and the corresponding errors is studied. A comparison between the predicted and declared values indicates that the magnitude of error is not related to the ET_H value. Fig. 3 displays a comparison between predicted outputs and target values. It is inferred that the trend of predictions and the magnitude of their deviation from the target values follow the same pattern throughout the comparison graph. Therefore it is concluded that the greatness of error is not related to the magnitude of the declared ET_H value.

3.2. Optimizing the model

Generally, neural networks may suffer from lack of samples (high variance) or lack of features (high bias). A bias-variance analysis is helpful for understanding the performance of a model. This analysis can be performed by comparing the errors observed from the training set and test set. Usually, high variance is perceived when the error of the training set is much smaller than the error of test set. This observation is related to insufficient number of samples. In such cases, increasing the number of entries is necessary. Appearance of high variance also means that the model will not perform well on new samples. On the other hand, witnessing high errors from both the training set and test set indicates a high bias. In such cases, increasing the complexity of the model by adding more features is often helpful for improving the predictions. Moreover,

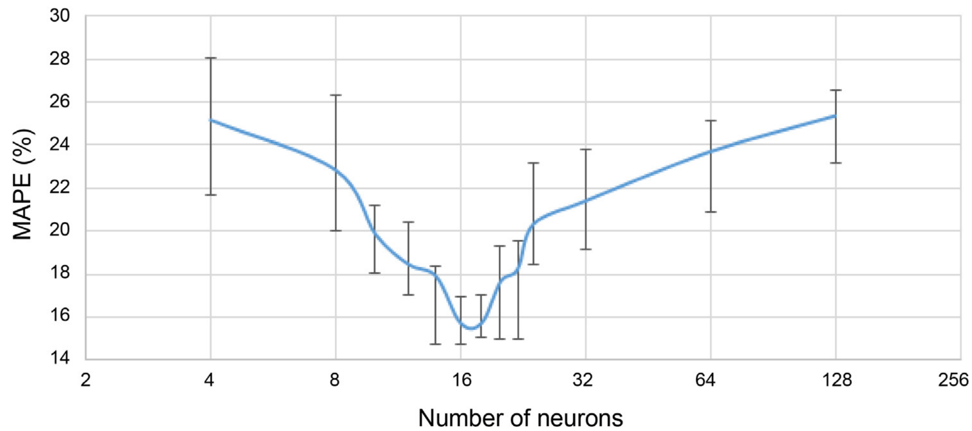

Fig. 4. Performance of the optimized model based on number of neurons.

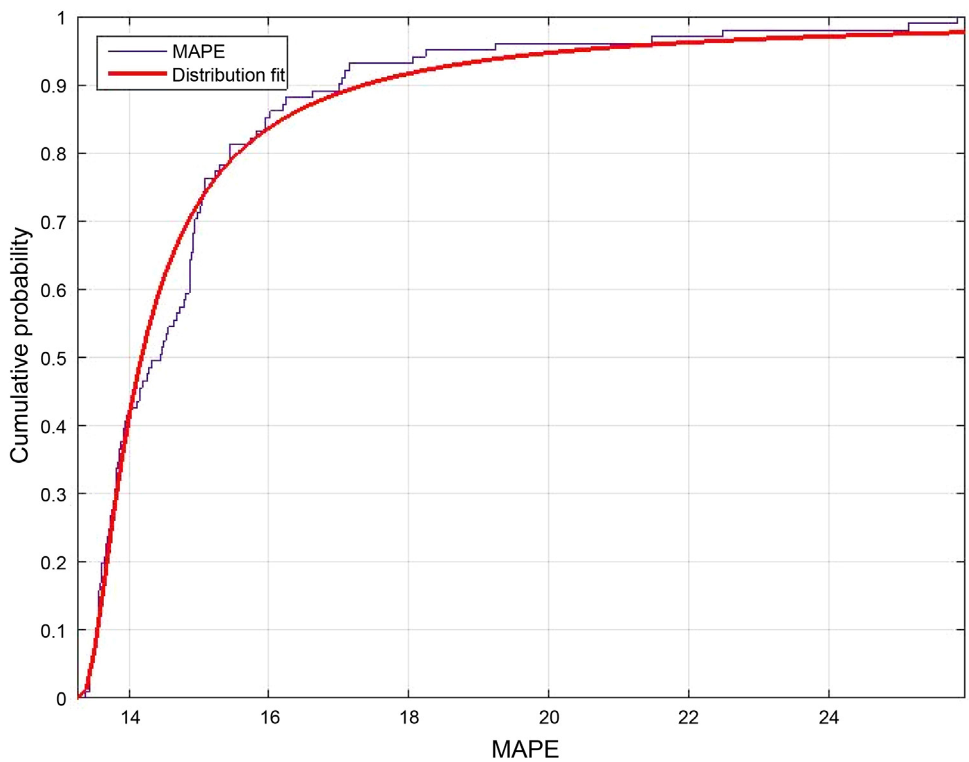

Fig. 5. Cumulative distribution (CDF) of MAPE for 100 models.

increasing the number of neurons may be helpful for decreasing the prediction error. The process of collecting new features is a time consuming activity, therefore investigating the optimum number of hidden neurons merits priority over additional feature collection.

Model C returned 13.37 and 13.78 MAPEs for the training set and test set respectively. Such observation implies that the model is not suffering from high variance and will perform similarly on new input data. As demonstrated in Table 5, all the available features from the CENED database are included in the input vectors. Moreover, linear and non-linear hybrid features are created in order to improve the performance of the model. Therefore, it can be implied that the trained model is currently at its best performance regarding the number of inputs and vector size.

As mentioned earlier, increasing the number of hidden neurons is an alternate approach for improving the performance of the model. Such method may result in better performance as the model is able to assign more complex weight configurations to each feature. However, ANNs with a large number of neurons are prone to overfitting. This means that the network will not perform well on new samples and suffer from high variance. Consequently, sequences of different models with various number of neurons are trained. The training process of each model is repeated 5 times while the average error and variances of the test set are documented (Fig. 4). The number of neurons are increased exponentially from 4 to 128. It is found that the model with 16 neurons displays the best performance among other samples. Therefore, increasing the number of neurons is not helpful for the current network and results in overfitting. In order to further optimize the model, various networks with 10–24 neurons are added to the previous tests. A final overview indicates that 16 neurons is the optimum configuration for the intended model C.

3.3. Applying neural networks for evaluating EPCs

Initialization of weight values is another factor which affects the performance of ANNs with back propagation. MATLAB soft-

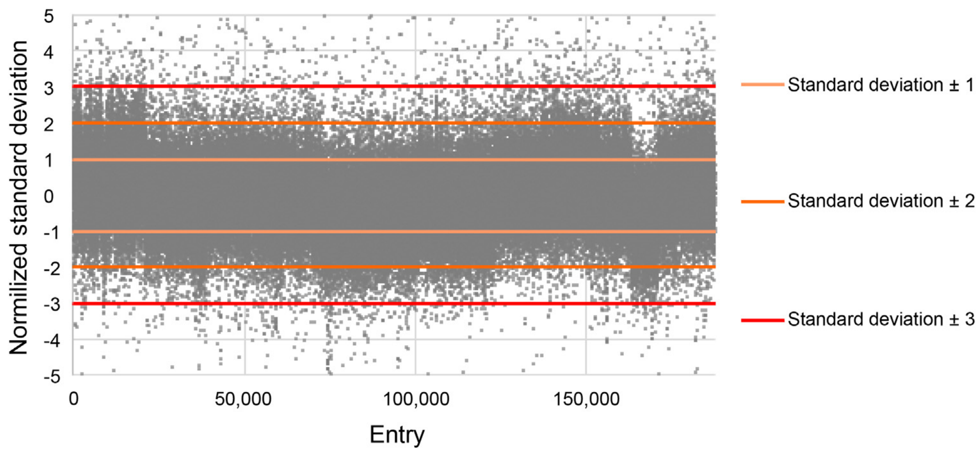

Fig. 6. A comparison of declared ET_H values and predicted confidence intervals.

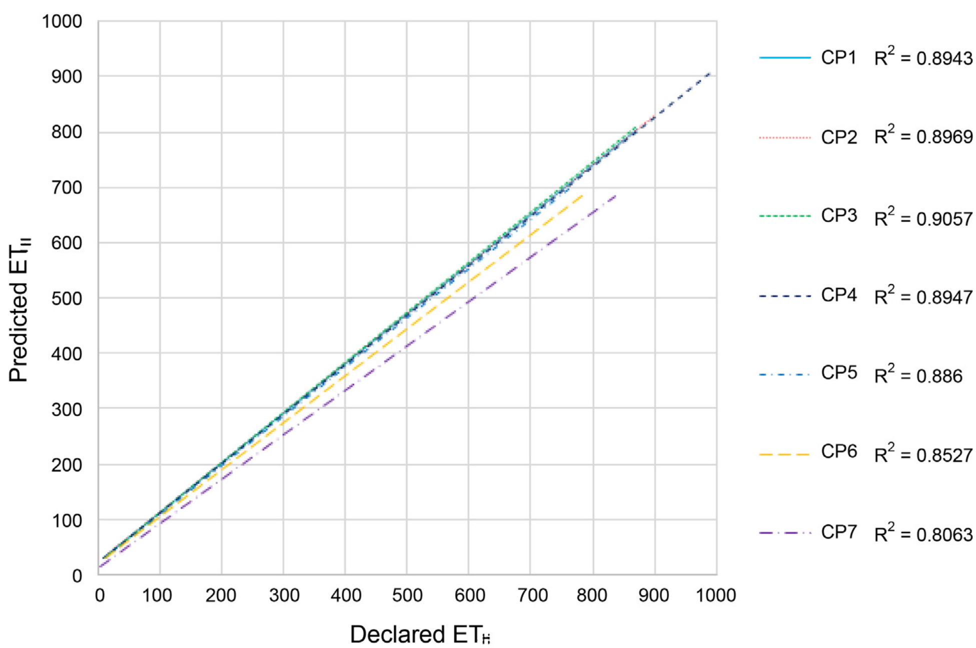

Fig. 7. Linear regression of predicted and declared ET_H values for each construction period.

ware uses a random initialization function in order to assign initial weight and bias values for each neuron of the model. Thus, for every time that the training process is repeated, a different performance is perceived. This stochastic approach in the prediction process is associated with a level of confidence and uncertainty. Therefore, a general perception of the most probable weights and biases is intended. In order to achieve a level of confidence over the predicted values, sequences of models with different initialization parameters are trained. The procedure of training the optimum model C is repeated 100 times; while the weight and bias values for each trained model are documented. Results indicate that the entire set of 100 trained models return an average MAPE of 14.44. Fig. 5 displays a cumulative probability distribution (CDF) of all 100 models. It is observed that 95% of the trained models have a MAPE less than 20, while 68% of the models return a MAPE below 15.

In order to evaluate an energy certificate, the related input features (model C) of the EPC are extracted from the CENED database and simulated through all 100 trained models. The output of these simulations is a variety of predicted ET_H values for each EPC. These predictions are represented in the form of a normalized Gaussian distribution. A confidence interval of 99% (±3 standard deviation intervals) defines the upper and lower boundaries of an acceptable declared ET_H. According to regulations, EPC calculators are permitted to have ±5% deviation in the results [33,34]. Consequently, declared ET_H values which fall within ±5% deviation of the confidence intervals are considered as valid EPCs. On the other hand, entries which fail to fulfill this requirement merit to be further investigated. A final survey over the entire dataset indicates that about 95% of the declared values are considered as valid EPCs (Fig. 6).

In order to obtain a better understanding of predicted EPCs with a large error, a survey over suspicious values is conducted. Further investigations on EPCs with high prediction errors reveal that they are likely to contain an abnormal input feature. For instance, some defective entries contained unreasonable surface to volume ratio. Others consisted of high thermal conductivity on all surfaces while declaring a low ET_H value. Fig. 7 displays the performance of model C on different portions of the dataset. Each regression line represents the accuracy of the model over a specific period of construction. It is perceived that the network returns a better precision for entries labeled as CP_1 to CP_4, while samples constructed after 1993 are predicted with less accuracy. Figs. 8 indicates that declared

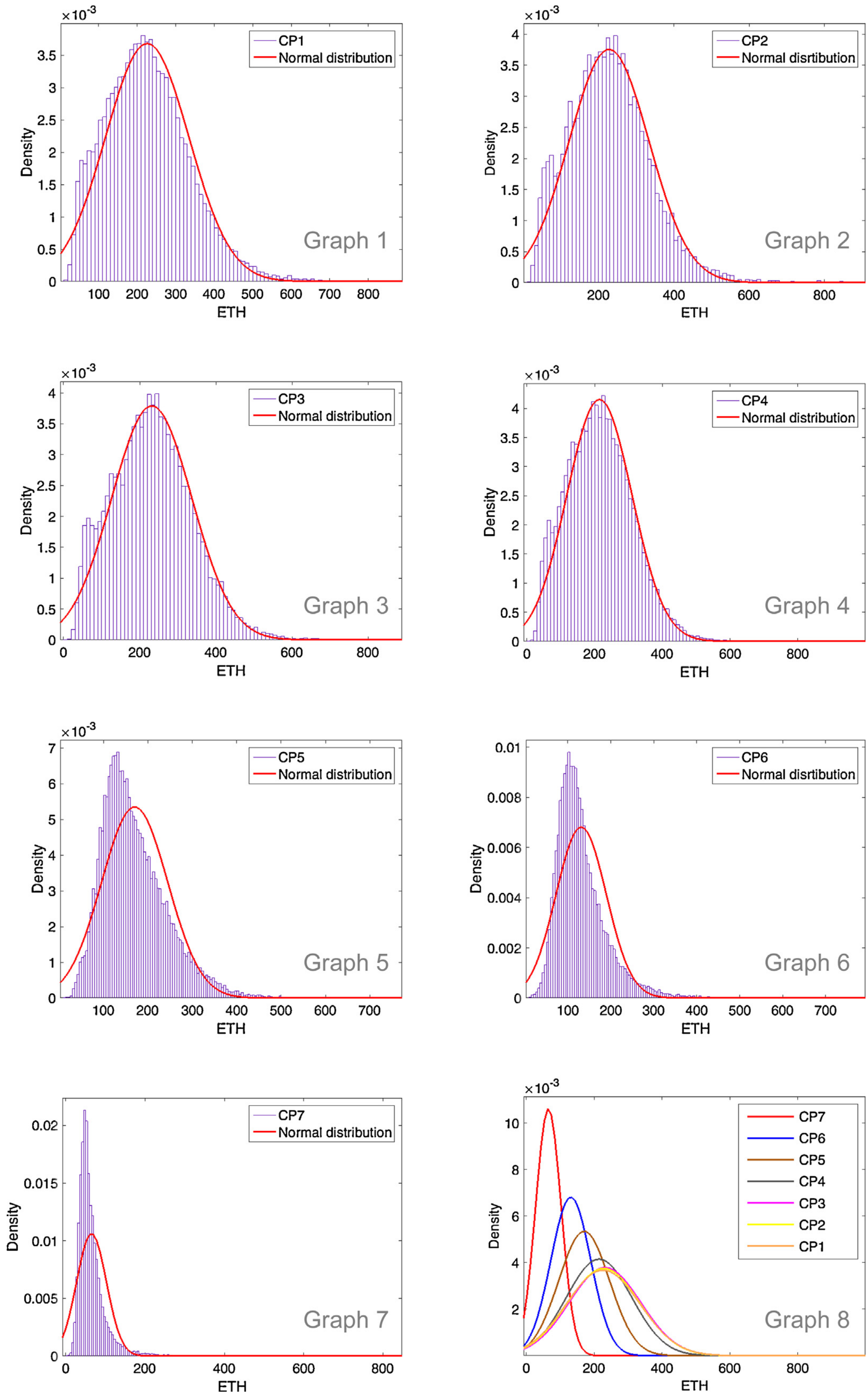

Fig. 8. Frequency distribution of declared ETH values according to each construction period (Graphs 1–7), and a comparison of all fitted normal distributions (Graph 8).

ET_H values of entries labeled as CP_6 and CP_7 form of an extremely skewed normal distribution. Neural networks perform minimization based on the mean of target values rather than the median. Consequently, the gap between the mean and median ET_H distribution of CP_6 and CP_7 forces the model to predict outputs based on a shifted target. It is concluded that a discrete labeling approach should be replaced with a more suitable feature for defining the year of construction (Fig. 8).

Some abnormalities were observed regarding the period of construction and the respective ET_H value. Considering that the ET_H is specifically related to the characteristics of the building envelope, observing a very high ET_H value for buildings that are constructed after 2007 or even 2000 is questionable. This suspicion is derived from the fact that the Italian law has defined multiple regulations at different stages which restrict the maximum thermal conductivity of building envelopes. Therefore, it is recommended to add a preprocessing phase before initiating the training procedure. In this phase, regulations in force by the regional law at the time of construction should be considered as a constraint. Consequently, entries which their envelope characteristics do not comply with the constraints should be removed from the dataset.

Although the current model returns an acceptable performance with small errors, yet increasing the network's accuracy is possible. Including features such as wall orientation, roof area and number of floors will be useful for improving the accuracy of the model. In this regard, inclusion of more detailed input features is a potential for future study.

4. Conclusion

The Lombardy region in Italy has implemented a calculation procedure in order to issue predicted EPCs for buildings. The certificates are issued by the CENED software which calculates the EP_H based on detailed building characteristics. Each energy certificate is uploaded to an online database that can be freely accessed through the CENED website.

The objective of this study was to fit an estimator to the CENED software by benefiting from the large number of energy certificates that are available online. Due to the complexity of the calculation procedure, a black-box approach using ANN was adopted for replicating the relation between the input parameters and the respective ET_H value. Several combinations of inputs were tested for defining the most suitable implementation of features. It was observed that using 12 features was sufficient for successfully predicting the ET_H value with an acceptable margin of error. Meanwhile, small abnormalities were observed in the overall performance of the network. Initial investigations on the predicted values revealed that the observed abnormalities were related to defective input data. Further explorations were also conducted in order to outline the most important factors which affect the model's performance. It was observed that the model consisted of sufficient number of samples, yet additional features were required for obtaining a better accuracy. A sequence of 100 models were trained due to the stochastic weight initialization in training ANNs. In order to evaluate each entry, a Gaussian distribution of 100 predictions were created based on the trained models. This approach provided a probability of occurrence for each certificate rather than a definitive statement. As a consequence, the procedure of monitoring each entry is prioritized based on its level of reliability i.e. probability of occurrence. Such approach may be replicated for similar EPC datasets with comparable data.

The described methodology of employing ANNs for EPC prediction offers an autonomous approach for detecting anomalies in building energy certificates. As a consequence, the process of validating self-declared energy certificates is facilitated with a fast estimator where suspicious entries have a priority to undergo a manual inspection. In order to evaluate the comprehensiveness of the current model, further explorations in applying the same procedure to other types of building is under study.

Expanding the developed model into a universal estimator for EPC prediction will provide the opportunity to compare the predicted energy performance index with dynamic energy simulation tools. This process can define a correlation between quasi-steady-state and dynamic simulations with the intention of obtaining a more realistic overview on the energy consumption trend in the region. Furthermore, the provided methodology can be adopted to predict the actual energy performance of buildings. It should be noted that the input features and target values described in this study are specifically aimed at estimating a calculation procedure, rather than the actual energy performance. However, the same pipeline of activities may be implemented while practicing ANN on the actual energy performance of buildings. The provided suggestions concerning pre-processing filters, feature selection, model optimization and post-processing evaluation of the trained network can be applied to similar predictive models aimed at actual building energy performance.

The application of machine learning in building energy estimation, provides the opportunity to extract predictions in a rapid manner. Therefore, it may be used to validate self-declared predicted energy performances in a reasonable amount of time. Furthermore, the effects of applying hypothetical modifications on the building's characteristics can be quickly observed. This particular advantage of black-box estimators is a suitable method for monitoring a common change over a large number of buildings, and therefore can be utilized to gain an early perspective of different energy policy strategies. Every few years, the calculation procedure of energy certificate tools may be improved in order to meet with new energy efficiency regulations. Application of powerful estimators such as ANN can provide the opportunity to update previously registered certificates based on new calculation procedures without the necessity of rerunning the entries through the updated energy certificate tool. Therefore, a quantitative assessment of possible changes in the assigned energy class due to new calculation procedures is a potential for future study. Moreover, similar approaches may be adopted in using ANN as a surrogate model for uncertainty and sensitivity analysis of energy certificate tools as rapid evaluation of entries is possible.

References

[1] E. Council, Ďirective 2002/91/ec of the European parliament and of the council of 16 December 2002 on the energy performance of buildings,Öfficial Journal of the European Communities. Directive, 91 (2002).

[2] Legislative Decree, No. 192 of 08/19/05, Implementation of Directive 2002/91/EC on the Buildings Energy Performance (2005).

[3] Regional Law, No. 24 of 12/11/06, Regulations for the Prevention and Reduction of Air Emissions to Protect the Health and the Environment (2006).

[4] Regional Council Decree No. VIII/5018 of 06/26/07, Determinations Regarding the Buildings Energy Certification. In: Implementation of Legislative Decree No. 192/05 and Art.9:25 of Regional Law No. 24/06 (2007).

[5] Regional Council Decree No.VIII/5773 of 10/31/07, Buildings Energy Certification –Amendments and Additions to Regional Council Decree No.VIII/5018, (2007).

[6] G. Dall'O, V. Belli, M. Brolis, I. Mozzi, M. Fasano, Nearly zero-energy buildings of the Lombardy region (Italy), a case study of high-energy performance buildings, Energies 6 (2013) 3506–3527.

[7] H.-x. Zhao, F. Magoulès, A review on the prediction of building energy consumption, Renew. Sustain. Energy Rev. 16 (2012) 3586–3592.

[8] A. Marenn, D. Jones, S. Franklin, Configuring and Optimizing the Back-Propagation Network: Handbook of Neural Computing Applications, Academic Press, 1990.

[9] R. Ž. Jovanović, A.A. Sretenović, B.D. Živković, Ensemble of various neural networks for prediction of heating energy consumption, Energy Build. 94 (2015) 189–199.

[10] Q. Li, Q. Meng, J. Cai, H. Yoshino, A. Mochida, Predicting hourly cooling load in the building: a comparison of support vector machine and different artificial neural networks, Energy Convers. Manage. 50 (2009) 90–96.

[11] A. Azadeh, S. Ghaderi, S. Tarverdian, M. Saberi, Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption, Appl. Math. Comput. 186 (2007) 1731–1741.
[12] A. Ahmad, M. Hassan, M. Abdullah, H. Rahman, F. Hussin, H. Abdullah, et al., A review on applications of ANN and SVM for building electrical energy consumption forecasting, Renew. Sustain. Energy Rev. 33 (2014) 102–109.
[13] M. Aydinalp-Koksal, V.I. Ugursal, Comparison of neural network, conditional demand analysis, and engineering approaches for modeling end-use energy consumption in the residential sector, Appl. Energ. 85 (2008) 271–296.
[14] P.A. González, J.M. Zamarreno, Prediction of hourly energy consumption in buildings based on a feedback artificial neural network, Energy Build. 37 (2005) 595–601.
[15] S.-M. Hong, G. Paterson, D. Mumovic, P. Steadman, Improved benchmarking comparability for energy consumption in schools, Build. Res. Inf. 42 (2014) 47–61.
[16] M. Yalcintas, An energy benchmarking model based on artificial neural network method with a case example for tropical climates, Int. J. Energy Res. 30 (2006) 1158–1174.
[17] M. Yalcintas, U. Aytun Ozturk, An energy benchmarking model based on artificial neural network method utilizing us commercial buildings energy consumption survey (CBECS) database, Int. J. Energy Res. 31 (2007) 412–421.
[18] R. Platon, V.R. Dehkordi, J. Martel, Hourly prediction of a building's electricity consumption using case-based reasoning, artificial neural networks and principal component analysis, Energy Build. 92 (2015) 10–18.
[19] M. Aydinalp, V.I. Ugursal, A.S. Fung, Modeling of the appliance, lighting, and space-cooling energy consumptions in the residential sector using neural networks, Appl. Energ. 71 (2002) 87–110.
[20] K. Li, C. Hu, G. Liu, W. Xue, Building's electricity consumption prediction using optimized artificial neural networks and principal component analysis, Energy Build. (2015).
[21] Z. Hou, Z. Lian, Y. Yao, X. Yuan, Cooling-load prediction by the combination of rough set theory and an artificial neural-network based on data-fusion technique, Appl. Energ. 83 (2006) 1033–1046.
[22] B.B. Ekici, U.T. Aksoy, Prediction of building energy consumption by using artificial neural networks, Adv. Eng. Software 40 (2009) 356–362.
[23] S. Ferlito, M. Atrigna, G. Graditi, S. De Vito, M. Salvato, A. Buonanno, et al., P̈redictive models for building's energy consumption: An Artificial Neural Network (ANN) approach,ïn AISEM Annual Conference, 2015 XVIII, 2015, pp. 1–4.
[24] E. Alpaydın, Introduction to Machine Learning, 2nd edition, MIT Press, Cambridge, Mass, 2010.
[25] S. Wang, C. Yan, F. Xiao, Quantitative energy performance assessment methods for existing buildings, Energy Build. 55 (2012) 873–888.
[26] D. Hawkins, S. Hong, R. Raslan, D. Mumovic, S. Hanna, Determinants of energy use in UK higher education buildings using statistical and artificial neural network methods, Int. J. Sustain. Built Environ. 1 (2012) 50–63.
[27] A. Melo, D. Cóstola, R. Lamberts, J. Hensen, Development of surrogate models using artificial neural network for building shell energy labelling, Energy Policy 69 (2014) 457–466.
[28] C. Buratti, M. Barbanera, D. Palladino, An original tool for checking energy performance and certification of buildings by means of artificial neural networks, Appl. Energ. 120 (2014) 125–132.
[29] Director General Decree, No. 5796 of 06/11/09, Updating of the Procedure for Calculating the Energy Certification of Buildings (2009).
[30] Regional Council Decree, No. IX/2554 of 11/24/11, Policy Address, Method of Investigation of Offenses and Sanctions Imposition of Regional Competence, provided by Art.27 of Law No.24, in the Field of Energy Certification (2011).
[31] T. a. C. Italian Government. Ministry of Industry, M̈inisterial Decree of 6/08/1994. Modifications and integration to the table of climatic zones for Italian municipalities, in the Presidential Decree of 26/08/1993 No. 412, regarding the energy saving of building thermal systems.,ëd. Official Gazette of 24/08/1994 No. 197, 6/08/1994.
[32] L. Sarto, N. Sanna, V. Tonetti, M. Ventura, On the use of an energy certification database to create indicators for energy planning purposes: application in northern Italy, Energy Policy 85 (2015) 207–217.
[33] 11300-1, Energy Performance of buildings: Determination of thermal energy demand of the building for air conditioning in summer and winter (2008).
[34] 11300-2, Energy performance of buildings: determination of the primary energy demand and yields for winter heating and the production of domestic hot water (2008).