Energy and Buildings 94 (2015) 189–199

Contents lists available at ScienceDirect

Energy and Buildings

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

Ensemble of various neural networks for prediction of heating energy consumption

Radiša Ž. Jovanović[a], Aleksandra A. Sretenović[b,*], Branislav D. Živković[b]

[a] University of Belgrade, Faculty of Mechanical Engineering, Department of Automatic Control, Kraljice Marije 16, Belgrade, Serbia
[b] University of Belgrade, Faculty of Mechanical Engineering, Department of Thermal Science, Kraljice Marije 16, Belgrade, Serbia

ARTICLE INFO

Article history:
Received 27 December 2014
Received in revised form 16 February 2015
Accepted 20 February 2015
Available online 3 March 2015

Keywords:
Heating consumption prediction
Neural networks ensemble

ABSTRACT

For prediction of heating energy consumption of a university campus, various artificial neural networks are used: feed forward backpropagation neural network (FFNN), radial basis function network (RBFN) and adaptive neuro-fuzzy interference system (ANFIS). Actual measured data are used for training and testing the models. For each neural networks type, three models (using different number of input parameters) are analyzed. In order to improve prediction accuracy, ensemble of neural networks is examined. Three different combinations of output are analyzed. It is shown that all proposed neural networks can predict heating consumption with great accuracy, and that using ensemble achieves even better results.

1. Introduction

The study of the building energy demand has become a topic of great importance, because of the significant increase of interest in energy sustainability, especially after the emanation of the EPB European Directive. In Europe, buildings account for 40% of total energy use and 36% of total CO_2 emission [1]. The estimation or prediction of building energy consumption has, therefore, played very important role in building energy management, since it can help to indicate above-normal energy use and/or diagnose the possible causes, if there has been enough historical data gathered. Scientists and engineers are lately moving from calculating energy consumption toward analyzing the real energy use of buildings. One of the reasons is that, due to the complexity of the building energy systems and behavior, non-calibrated models cannot predict well building energy consumption, so there is a need for real time image of energy use (using measured and analyzed data).

The classic approach to estimate the building energy use is based on the application of a model with known system structure and proprieties as well as forcing variables (forward approach). These engineering methods use physical principles to calculate thermal dynamics and energy behavior on the whole building level or for sub-level components [2]. A lot of different software tools have been developed for this purpose, such as DOE-2, EnergyPlus, TRNSYS, BLAST, ESP-r, HAP, APACHE. Using these tools requires detailed knowledge of the numerous building parameters (constructions, systems) and behavior, which are usually not available. Some simplified methods for building energy use prediction were developed. The steady-state method using degree-day was presented in [3]. A simple method of formulating load profile for UK domestic buildings has been introduced by Yao and Steemers in [4]. They used thermal dynamic model to predict daily breakdown energy demand load profile of appliance, domestic hot water and space heating.

A different approach for building energy analysis is based on the so-called inverse or data-driven models. In recent years, considerable attention has been given to data-driven based methods [5]. By a data-driven approach, building energy behavior is analyzed, while defining relationships with one or more different driving forces or parameters. It is required that the input and output variables are known and measured, and the development of the inverse model consists in determination of a mathematical description of the relationship between the independent variables and the dependent one. The data-driven approach is useful when the building (or a system) is already built, and actual consumption (or performance) data are measured and available. For this approach, different statistical methods can be used. Statistical regression models simply correlate the energy consumption with the influencing variables. Bauer and Scartezzini [6] introduced a regression model to predict heating and cooling load, dealing with internal and solar gains. Catalina et al. [7] developed the regression model for the prediction of the monthly heating need for the single-family residential sector. Katipamula et al. [8] found that a multilinear regression provides better accuracy than a single variable model for modeling energy

* Corresponding author. Tel.: +381 65 24 333 79.
E-mail address: asretenovic@mas.bg.ac.rs (A.A. Sretenović).

http://dx.doi.org/10.1016/j.enbuild.2015.02.052
0378-7788/

consumption. Lam et al. [9] used Principal Component Analysis (PCA) of prevailing weather conditions in sub-tropical Hong Kong.

Artificial neural networks (ANN) are the most used artificial intelligence models for the building energy consumption prediction. In [10] detailed model simulation (in EnergyPlus) was compared with a simple model based on ANN. Regarding the ANN models, the results for the simpler (temperature-only input) and the more complex (temperature/relative humidity/solar radiation inputs) neural networks showed a fair agreement between energy consumption forecasts and actual values, when different networks for working days and weekends are implemented. Ekici et al. [11] used a backpropagation three-layered ANN for the prediction of the heating energy requirements of different building samples. The inputs of the network for training and testing were considered as building transparency ratio, orientation and insulation thickness and the output is building heating energy needs. Hourly heating energy consumption for a model house calculated by degree-hour method was used for training and testing the ANN model in [12]. In [13] actual recorded input and output data that influence long-term energy consumption were used in the training, validation and testing process. The produced ANN results were compared with the results produced by a linear regression method, a support vector machine method and with real energy consumption records showing much better prediction accuracy. In [14] authors tested several training algorithms, calibrated in real conditions and used ANN to predict the energy consumption and PMV value for an indoor swimming pool. Afterwards ANN is utilized as a cost function engine to develop an intelligent energy and thermal comfort management system for this pilot. Li et al. in [15] proposed the hybrid genetic algorithm-adaptive network-based fuzzy inference system which combined the fuzzy if-then rules into the neural network-like structure for the prediction of energy consumption in the library building. The calculated results indicated better performance compared with ANN in term of forecasting accuracy. In [16] Karatasou et al. discussed how neural networks, applied to predict energy consumption in buildings, can advantageously be improved guided by statistical procedures. They deal with the identification of all potential relevant input, the selection of hidden units for this preliminary set of inputs, through an additive phase and the remove of irrelevant inputs and useless hidden units through a subtractive phase. A review of the different neural network models used for building energy prediction can be found in [17]. In most papers, use of single neural network was analyzed, while some authors, as in [15], compared different network architectures. In this paper, idea was to analyze possible application of various network topologies on the same case study. At first stage, prediction results achieved with different networks (feedforward, radial basis and adaptive neuro-fuzzy inference system) are compared. In the second stage qualitative specific and innovative use of created networks by combining them into ensemble is proposed. All these analyses are done using different number of input variables. The ensemble of neural networks is a very successful technique where the outputs of a set of separately trained neural networks are combined to form one unified prediction [18]. Since an ensemble is often more accurate than its members, such a paradigm has become a hot topic in recent years and has already been successfully applied to time series prediction [19], weather forecasting [20], load prediction in a power system [21]. The novelty in this paper would be creating ensemble of neural networks for prediction of heating energy consumption.

2. Artificial neural networks models

The main advantage of an ANN model is its self-learning capability and the ability to approximate a nonlinear relationship between the input variables and the output of a complicated system. In this study, the three different architectures of ANN are used for prediction of the heating energy consumption.

2.1. Feedforward backpropagation neural network (FFNN)

The feedforward neural network architecture consists of an input layer, an output layer, and one or more hidden layers of neurons. Each layer has a number of neurons and each neuron is fully interconnected with adaptable weighted connections to neurons in the subsequent layer. The nonlinear activation functions in the hidden layer neurons enable the neural network to be a universal approximator. The process of training network is the adjustment of the weights, so that the network can produce the desired response to the given inputs. Different training algorithms could be applied to minimize the error function, but the most widely used are the backpropagation algorithm and the algorithms derived from it. They use a gradient descent technique to minimize the cost function which is the mean square difference between the desired and the actual network outputs. In this study, a multilayer feedforward network with single hidden layer and backpropagation learning algorithm is used.

2.2. Radial basis function networks (RBFN)

A RBF network, as a type of feedforward neural network, consists of three layers including an input layer, a single hidden layer and an output layer. The input nodes are directly connected to the hidden layer neurons. The hidden layer transforms the data from the input space to the hidden space using a nonlinear function. RBFN uses the radially symmetrical function as an activation function in the hidden layer, and the Gaussian function, the most commonly used activation function, is adopted in this study. For a RBFN with an n-dimensional input $x \in \mathcal{R}^n$ the output of the jth hidden neuron is given by

$$h_j(x) = \phi_j(||x - c_j||), \quad j = 1, 2, \ldots, m \tag{1}$$

where c_j is the center (vector) of the jth hidden neuron, m is the number of neurons in the hidden layer and $\phi(\cdot)$ is the radial basis function. The neurons of the output layer have a linear transfer function. The kth output of the network is obtained by the weighted summation of the outputs of all hidden neurons connected to that output neuron:

$$\hat{y}_k(x) = \sum_{j=1}^{m} w_{kj} h_j(x) + w_{k0} \tag{2}$$

where w_{kj} is the connecting weight between the jth hidden neuron and the kth output unit, w_{ko} is the bias and m is the number of the hidden layer neurons. The training RBFN is aimed at adjusting parameters of Gaussian functions (centers and widths) and the weights between the hidden and the output layers. The weights are optimized using least mean square algorithm, and the centers can be chosen randomly or using some clustering algorithms.

2.3. Adaptive neuro-fuzzy inference system (ANFIS)

The adaptive network-based fuzzy inference system (ANFIS) proposed by Jang [22] is one of the most commonly used fuzzy inference systems, and its architecture is obtained by embedding the fuzzy inference system (FIS) into the framework of adaptive networks. The architecture of the ANFIS used in this study is based on the first-order Takagi–Sugeno model [23]. Typical rule set with base fuzzy if-then rules can be expressed as

$$\begin{aligned}&\textbf{If } x_1 \textbf{ is } A_1 \textbf{ and } x_2 \textbf{ is } B_1 \textbf{ then } f_1 = a_1x_1 + b_1x_2 + c_1\\ &\textbf{If } x_1 \textbf{ is } A_2 \textbf{ and } x_2 \textbf{ is } B_2 \textbf{ then } f_2 = a_2x_1 + b_2x_2 + c_2\end{aligned} \quad (3)$$

The ANFIS architecture is shown in Fig. 1. It is composed of five layers where each layer contains several nodes described by the node function. Let O_i^j denote the output of the i-th node in layer j.

Layer 1: In the first layer, all the nodes are adaptive nodes. The outputs of layer 1 are the fuzzy membership grades of the inputs, which are given by

$$O_i^1 = \mu_{A_i}(x_1), \quad i = 1, 2 \quad (4)$$

$$O_i^1 = \mu_{B_{i-2}}(x_2), \quad i = 3, 4 \quad (5)$$

where A_i and B_i are the linguistic labels and μ_{A_i} and μ_{B_i} are the membership functions for A_i and B_i linguistic labels, respectively. As node functions in this layer any continuous and piecewise differentiable functions, such as commonly used trapezoidal, triangular-shaped, Gaussian or generalized bell membership functions, can be used. Therefore, outputs of this layer form the membership values of the premise part and parameters contained in membership functions of fuzzy sets called premise parameters.

Layer 2: In contrast to layer 1 the nodes in this layer are fixed. The output O_i^2 of the node i can be computed as

$$O_i^2 = w_i = \mu_{A_i}(x_1)\cdot\mu_{B_i}(x_2), \quad i = 1, 2 \quad (6)$$

where w_i represents a firing strength of a rule.

Layer 3: In this layer where the normalization process is performed, the nodes are fixed. The ratio of the i-th rules firing strength to the sum of all rules firing strengths is calculated for the corresponding node and thus the outputs of this layer are called normalized firing strengths:

$$O_i^3 = \overline{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2 \quad (7)$$

Layer 4: The fourth layer deals with the consequent part of the fuzzy rule. Every node i in this layer is an adaptive node and it calculates the contribution of i-th rule in the model output function which is defined based on the first-order Takagi–Sugeno method as

$$O_i^4 = \overline{w}_i f_i = \overline{w}_i(a_ix_1 + b_ix_2 + c_i), \quad i = 1, 2 \quad (8)$$

where $\{a_i, b_i, c_i\}$ is the parameter set. Parameters in this layer are referred to as consequent parameters.

Layer 5: This is the summation layer, which consists of a single fixed node. It sums up all the incoming signals and produces the output:

$$O_i^5 = y = \sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (9)$$

From the proposed ANFIS architecture, it is observed that given the values of premise parameters, the overall output can be expressed as a linear combinations of the consequent parameters. More precisely, the output y can be rewritten as

$$\begin{aligned}y &= \frac{w_1}{w_1 + w_2}f_1 + \frac{w_2}{w_1 + w_2}f_2 = \overline{w}_1 f_1 + \overline{w}_2 f_2 = (\overline{w}_1 x_1)a_1\\ &+ (\overline{w_1}x_2)b_1 + \overline{w}_1c_1 + (\overline{w}_2x_1)a_2 + (\overline{w}_2x_2)b_2 + \overline{w}_1c_2\end{aligned} \quad (10)$$

In the training process, the least squares method (forward pass) is used to optimize the consequent parameters with the premise parameters fixed. Once the optimal consequent parameters are found, the backward pass starts immediately. The gradient descent method (backward pass) is used to adjust optimally the premise parameters corresponding to the fuzzy sets in the input domain. The output of the ANFIS is calculated by employing the consequent parameters found in the forward pass. The output error is used to adapt the premise parameters by means of a standard back propagation algorithm.

2.4. Artificial neural network ensembles

The generalization ability of a neural network system can be significantly improved through an ensemble of neural networks with respect to the best performing single neural network [24]. Generally, the main objective of combining ANNs in a redundant ensemble is to achieve better performance. Researchers have shown that simply combining the output of many neural networks can generate more accurate predictions. Good ensemble is the one where the individual networks have both accuracy and diversity, namely the individual networks make their errors on different parts of the input space. An important problem is, then, how to select the aggregate members in order to have an optimal compromise between these two conflicting conditions. The accuracy is achieved by proper training algorithms and parameters, depending of selected architecture type. The diversity of the ensemble members can be achieved by different methods, and most widely used [25,26] are:

(1) manipulating the set of initial random weights,
(2) varying the topologies: varying number of input and/or hidden nodes, or even networks with different types,
(3) varying the training algorithm, and

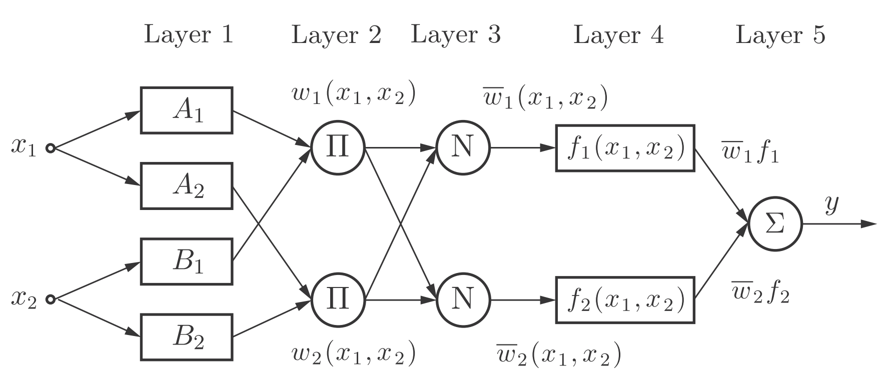

Fig. 1. The ANFIS architecture.

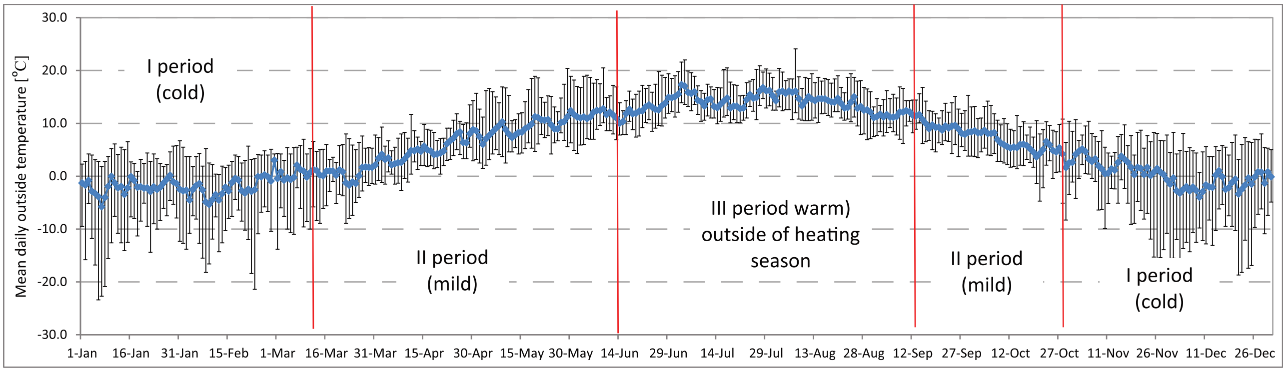

Fig. 2. Average mean daily temperature for years 2006–2014.

(4) manipulating the training set: combining networks trained with different sample data.

There are several methods of designing neural network ensembles, but most of them follow the two-stage design process [25]: first generating individual networks, which are usually trained independent of each other, and then combining them. In this paper, three different network architectures (FFNN, RBFN and ANFIS) are used as members, and the ensemble is created using three different combining methods (simple average, weighted average and median).

3. Case study

University campuses are specific groups of diverse buildings, with significant energy consumption [27]. They consist of many different buildings, representing small-scale town for itself. Therefore, they provide an excellent testbed to characterize and understand energy consumption of group of mixed use buildings. Norwegian University of Science and Technology (NTNU) campus Gløshaugen consists of 35 buildings, with total area of approximately 300,000 m^2. Building and Energy Management System (BEMS) and web-based Energy Monitoring System (Energy Remote Monitoring ERM) are available at NTNU. There are 46 heating meters and 79 electricity meters installed in the campus. Hourly heat and electricity consumption can be collected on ERM [28]. The Main meter installed by the district heating supplier was taken as relevant. Daily heating energy consumption of the campus is analyzed in this paper.

3.1. Data pre-processing

All weather data were gathered from the local meteorological station Skjetlein [29]. Heating season in Trondheim area lasts around 251–280 days [30]. The heating season is defined as the period from the day the mean daily temperature falls below 11 °C during the autumn and until the day it rises above 9 °C during the spring [31]. Based on the analysis for the period 1961–1990, the beginning of the heating season in Trondheim is usually between 29/08 and 17/09, and the end of the heating season 10/05–29/05 [30]. Considering that the outside temperature has the biggest influence on heating energy consumption, mean daily outside temperatures for years 2006 until 2014 were investigated in order to determine optimal number of neural networks. The average mean daily temperatures for the last 8 years is shown in Fig. 2. The error bars show the maximum and minimum mean daily temperature for the specific date in the same period (years 2006–2014). After the analysis, database is divided as follows:

- Cold period from January 1st to March 31st and from November 1st to December 31st.
- Mild period from April 1st to June 15th and from September 16th to October 31st.
- Warm period (outside of heating season) from June 15th to September 15th is excluded from the analysis.

It implicates that better prediction results can be obtained using separate network models for each period compared to using one network for all year. In this paper, only the cold period (with biggest heating energy consumption) will be analyzed.

The daily heating consumption was analyzed in terms of the type of the day. The correlation with mean daily outside temperature for each day of the week for the year 2012 is shown in Fig. 3. Analysis showed that there is no specific difference between the working days (heating consumptions for Monday to Friday have similar trendlines), while the regression lines for Saturday and Sunday are below them, as expected. In NTNU campus Gløshaugen, heating is not switched off during the weekends, only the design set-point is lowered, so the heating consumption on Monday is not significantly different than the other working days. The analysis of the daily heating consumption also showed that during the holidays and exam periods, heating operation is at the same level as for the working days (heating is maintained at the designed set-point). These conclusions implicate that there should be two separate networks created: one for the working days, and other for the weekend. In this paper, the network for the working days is analyzed.

4. ANN models development

The MATLAB with Neural Network Toolbox was used for designing, implementing, visualizing and simulating proposed prediction models.

4.1. Input variable selection

The most important task in building an ANN prediction model is the selection of input variables. Many different studies dealing with impact of various variables on energy consumption can be found in literature. Empirical research of the influence of hourly values of solar radiation and wind speed on heating demands of building complex heated by district heating system was conducted in [32]. The research results confirmed the influence of increasing heat demand in case of higher wind speeds and decreasing heat demand in cases of sunny days occurring during the heating season. The input variables for the neural network model, that are considered in this study, are: mean daily outside temperature (° C) (t_m), mean daily wind speed (m/s) (w_m), total daily solar radiation (W h/m^2) (*SR*), minimum daily temperature (°C) (t_{min}), maximum daily temperature (°C) (t_{max}), relative humidity (%) (φ), day of the week (*WD*), month of the year (MY) and heating consumption of the previous day (HC_p). HC_p is selected based on partial autocorrelation

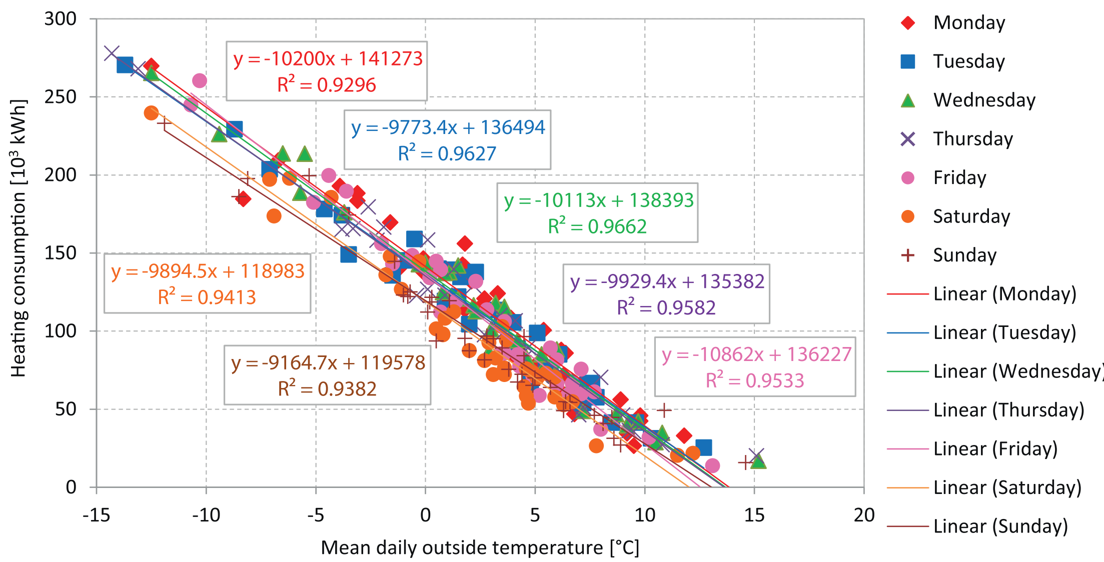

Fig. 3. Correlation of the daily heating energy consumption with mean daily outside temperature for the year 2012.

function (PACF) which indicates that the sample PACF has the most significant autocorrelation of 0.85 at lag 1 (previous day).

In data-driven modeling the identification of the most relevant influencing factors is necessary, so that the number of monitored parameters can be reduced. The focus of the investigations should be addressed, only or mainly, on those factors showing the highest impact. In order to analyze the influence of using different number of input variables on prediction results, three models are developed in this paper. For the first model (M1) all available input variables are used, for the second (M2) seven variables, selected using forward selection procedure, and for the third model (M3) three most influencing. Of course, using more inputs achieves better match with the reality, but the purpose of this comparison was also to show that even with less input variables satisfying prediction accuracy can be achieved. There are several reasons for this analysis:

- In reality, it is not always possible to measure and gather all the variables that are available for this case study, so the question is if using only three most influencing ones can predict heating consumption with good accuracy (M1).
- One of the most important issues in data-driven modeling is the identification of the most relevant influencing factors in order to reduce the number of parameters to be monitored. It is necessary to address the focus of the investigations, only or mainly, on those factors showing the highest impact. Using forward selection method is one way to determine these parameters (M2)
- Using less input variables can have several advantages, especially when it comes to prediction of consumption not just one day ahead, but for longer period in advance. It is necessary to have input variable database for that period. One way is to develop models to separately predict input variables (temperature, wind speed, etc.) and then use them to predict consumption. In that way, prediction is done day by day, and the error is accumulated at the end (especially with more previously predicted input variables). Therefore, selection of variables plays crucial role in prediction modeling.

There are different methods for reducing the number of input variables. In forward selection method used in this study, which is based on the linear regression model, the first step is to order input variables according to their correlation with the dependent variable (from the most to the least correlated variable). Then, the predictor (input) variable, which is best correlated with the dependent variable, is selected as the first input. Remaining variables are then added one by one as the second input according to their correlation with the output and the variable which most significantly increases the correlation coefficient (R^2) is selected as the second input, and the process is repeated until adding new variable does not significantly increase the R^2. In Table 1 the results of forward selection are shown.

Seven candidates according to their importance are selected as input variables for the second model (M2): outside mean daily temperature (t_m), heating consumption of the previous day (HC_p), day of the week (WD), maximum daily temperature (t_{max}), relative humidity (φ),total daily solar radiation (SR) and month of the year (MY). It was expected that the influence of wind speed and solar radiation would be higher. Taking into account that all performed analysis deal with daily consumptions, and mean daily meteorological parameters, many variations of wind speed and solar radiation are lost due to the 24-h averaging. The further work planned in this field will be focused on hourly data, where the real impact of these input variables can be expressed. The month of the year (MY) did not show significant impact, which was expected, considering that the whole dataset was divided into periods with relatively similar mean outside daily temperatures. Therefore, data pre-processing is necessary in order to find the optimal number of subsets, networks and variables. Considering the necessity to reduce the number of input variables, while obtaining acceptable prediction accuracy, the third model, taking into account three most dominant variables (M3) is proposed and analyzed. Input variables for M3 are: heating consumption of the previous day, mean daily outside temperature and day of the week. For all three models, each of the three proposed network architecture is analyzed. One of the goals of this paper is to compare the prediction accuracy of various network architectures, while using different number of input variables (M1, M2, and M3).

Table 1
Results of forward selection procedure.

Input variables subset	R^2	RMSE (kW h)
t_m	0.9369	13,898
t_m, HC_p	0.9533	11,970
t_m, HC_p, WD	0.9588	11,263
t_m, HC_p, WD, t_{max}	0.9615	10,908
t_m, HC_p, WD, t_{max}, φ	0.9636	10,618
t_m, HC_p, WD, t_{max}, φ, SR	0.9657	10,322
t_m, HC_p, WD, t_{max}, φ, SR, MY	0.9663	10,258

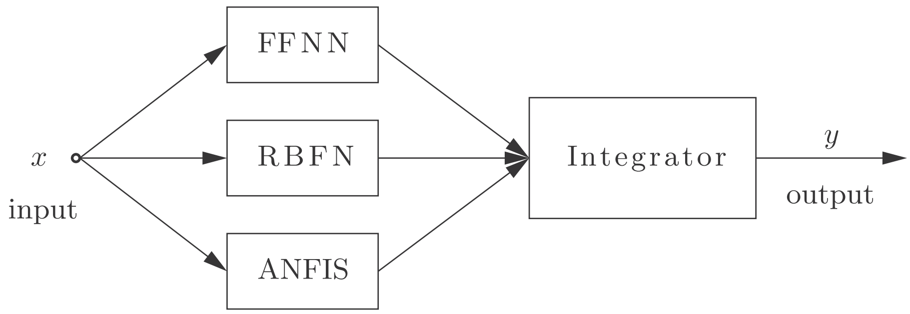

Fig. 4. The neural networks ensemble structure.

Table 2
Prediction indices for training networks.

	R^2			RMSE (kW h)			MAPE (%)		
	M1	M2	M3	M1	M2	M3	M1	M2	M3
FFNN	0.9836	0.9831	0.9723	7151.6	7771.4	9209.6	3.4402	3.6416	4.7947
RBFN	0.9764	0.9751	0.9715	8473.8	8699.8	9303.9	4.4663	4.3532	4.9405
ANFIS	0.9826	0.9813	0.9754	7200.3	7549.8	8651.1	3.8737	3.9458	4.6186
Ensemble									
Simple	0.9845	0.9812	0.9753	6885.4	7566.0	8662.9	3.5038	3.8282	4.6162
Weighted	0.9852	0.9816	0.9752	6737.5	7473.3	8686.8	3.3754	3.8427	4.6131
Median	0.9840	0.9809	0.9747	6994.7	7624.7	8775.4	3.5784	3.8634	4.6461

4.2. FFNN models

The FFNN model used in this study is a three-layer feedforward neural network composed of one input layer, one output layer and one hidden layer. The activation functions used for the hidden and output layers were the hyperbolic tangent (tansig) and linear (purelin) functions, respectively. In this study the Levenberg–Marquardt (LM) training algorithm, which is a variant of feedforward back-propagation algorithm is used. For the FFNN models the number of neurons in the hidden layer was found by a trial and error procedure. The optimum structure of the best FFNN models was found to be 10, 8 and 7 neurons in the hidden layer for models M1, M2 and M3, respectively.

4.3. RBFN models

A customized RBFN function available in MATLAB, which iteratively creates a radial basis network one neuron at a time, is used to develop the model. The number of neurons in the hidden layer is increased automatically until the error goal is achieved, or the maximum number of neurons in hidden layer has been exceeded. The radius value (known as spread) of the radial basis function was varied for the best performance of the RBF network. The values 0.34, 2.48 and 1.12 were obtained for the networks models M1, M2 and M3, respectively.

4.4. ANFIS models

In order to start the development of ANFIS prediction model, an initial fuzzy model has to be derived. This model is required to find the number of inputs, number of linguistic variables and hence the number of rules in the final fuzzy model. As the first step toward extracting the initial fuzzy model, the subtractive clustering method is applied to the input–output data pairs. The method proposed by Chiu [33] is a fast, one-pass algorithm for estimating the number of clusters and the cluster centers in an unsupervised way by measuring the potential of data points in the feature space. After clustering, the number of fuzzy rules and premise fuzzy membership functions are determined. Then, the least squares method and hybrid learning algorithm are used to identify the optimal values of these parameters, including consequent and premise parameters. For models M1, M2 and M3, using Gaussian membership function, after optimization procedure, 4, 5 and 9 fuzzy rules, respectively, are obtained.

4.5. Neural network ensemble

Many engineering problems, especially in energy use prediction, appeared to be too complex for a single neural network. The main goal of this paper is to examine the possible improvement of the prediction accuracy by using network ensemble. A neural-network ensemble is a very successful technique where the outputs of a

Table 3
Prediction indices for testing networks.

	R^2			RMSE (kW h)			MAPE (%)		
	M1	M2	M3	M1	M2	M3	M1	M2	M3
FFNN	0.9814	0.9793	0.9743	8496.1	9133.5	8869.6	5.6283	5.7914	5.2485
RBFN	0.9816	0.9813	0.9756	8849.1	9105.1	8797.7	5.6682	6.0097	5.4314
ANFIS	0.9783	0.9812	0.9748	9115.4	8860.9	8818.4	5.5778	5.6360	5.4300
Ensemble									
Simple	0.9845	0.9818	0.9770	8169.1	8751.9	8465.3	5.3204	5.8318	5.2496
Weighted	0.9843	0.9819	0.9766	8153.6	8745.9	8484.0	5.3686	5.7315	5.2274
Median	0.9830	0.9814	0.9773	8451.4	8932.4	8390.5	5.4820	5.8530	5.1124

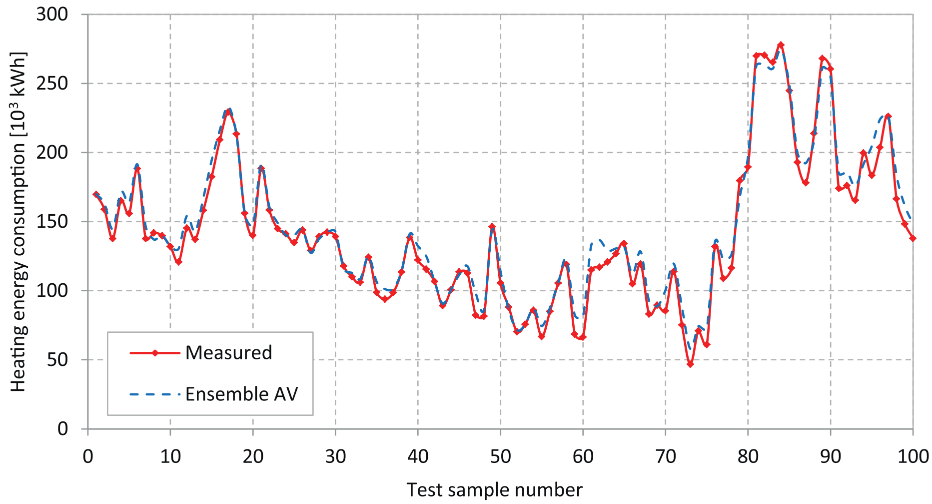

Fig. 5. Prediction results of ensemble with simple averaging for the test period for model 1.

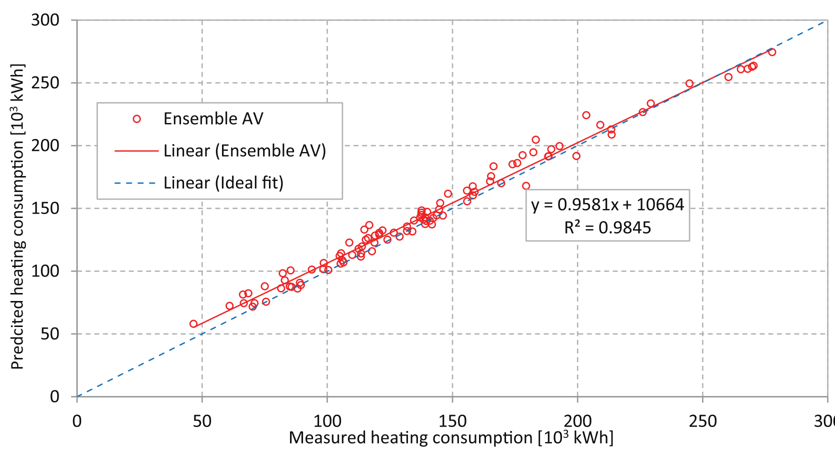

Fig. 6. Comparison of the measured and predicted values of ensemble with simple averaging for the test period for model 1.

set of separately trained neural networks are combined to form one unified prediction. The diversity of the ensemble members is achieved with using different network architectures. In this paper FFNN, RBFN and ANFIS are used as members, and the ensemble is created using three different combining methods. The accuracy of individual networks is achieved by proper training algorithms, selecting number of hidden layers and number of neurons in hidden layer by trial and error procedure (for FFNN), optimizing radius value of the radial basis function (for RBFN), and using adequate number of fuzzy rules (for ANFIS). The results show that using ensemble improves prediction accuracy compared to single network. In different data points, various networks show

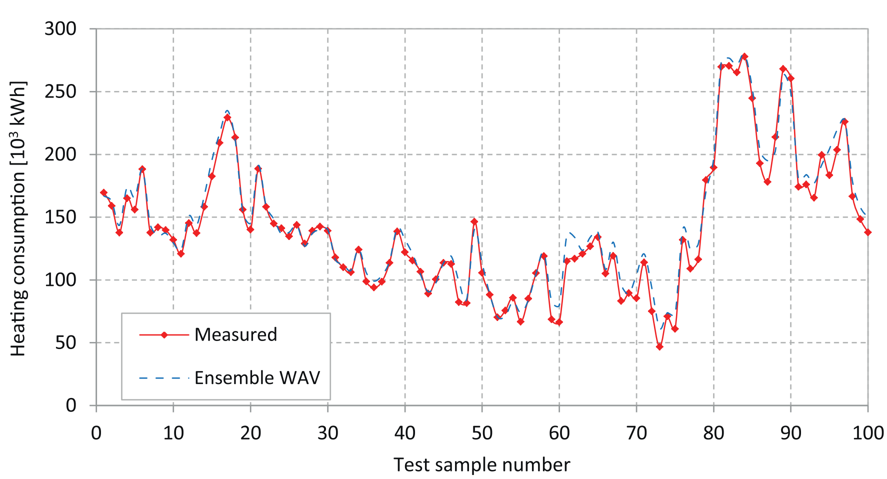

Fig. 7. Prediction results of ensemble with weighted averaging for the test period for model 2.

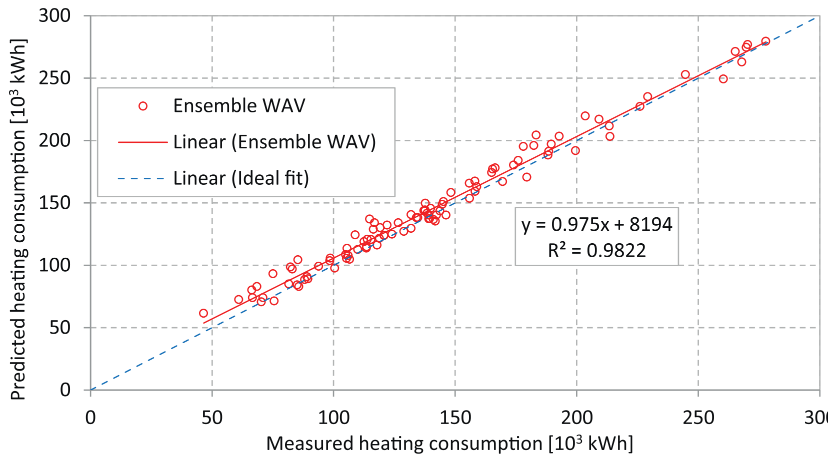

Fig. 8. Comparison of the measured and predicted values of ensemble with weighted averaging for the test period for model 2.

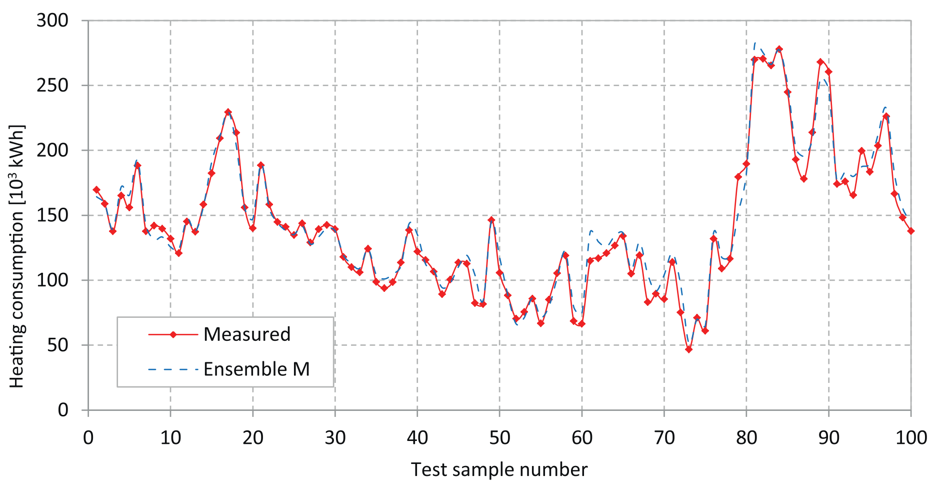

Fig. 9. Prediction results of ensemble with median based averaging for the test period for model 3.

better accuracy, and ensemble, by combining their outputs, can compensate the individual network errors, and therefore provide better result.

In Fig. 4 the network ensemble proposed in this paper is shown. Three different methods (combinations of output) for combining ensemble members are used:

- the simple average, determines the ensemble output by taking the average of all outputs provided by the individual members (ensemble AV);
- weighted average, the output of the ensemble is given by a weighted average of its components (ensemble WAV); and
- median based averaging (ensemble M).

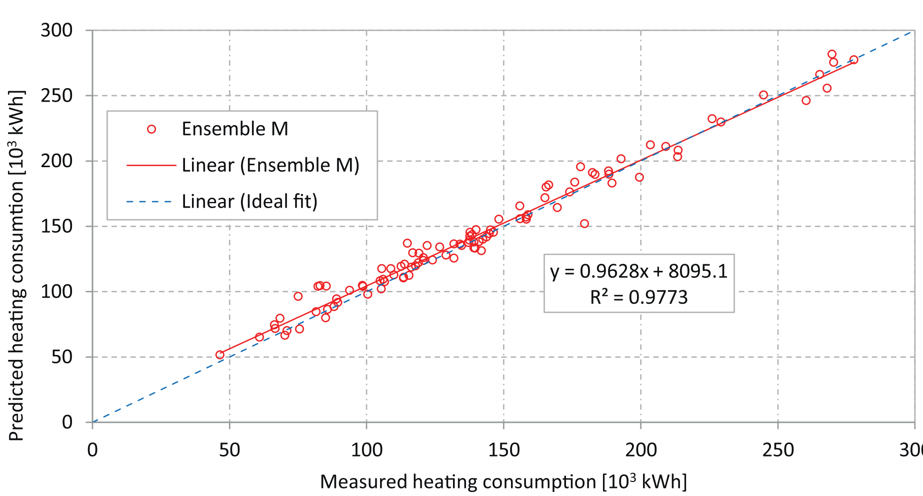

Fig. 10. Comparison of the measured and predicted values of ensemble with median based averaging for the test period for model 3.

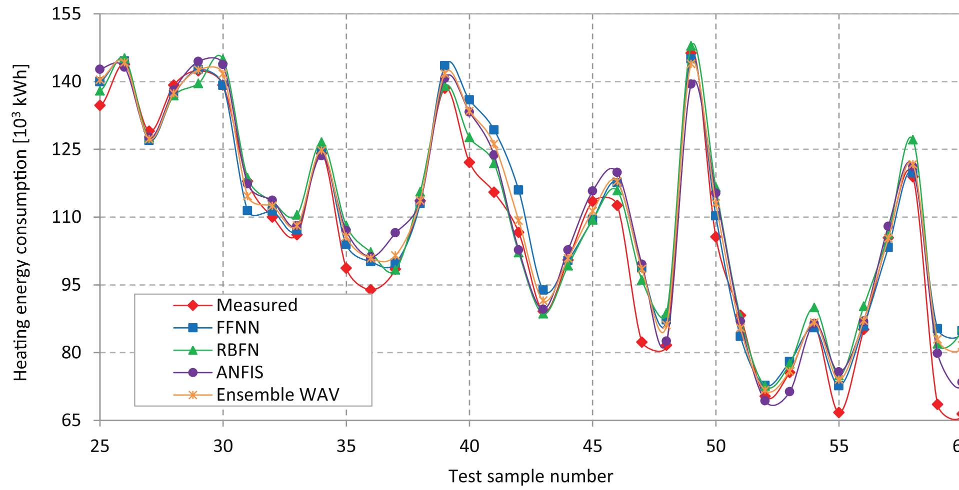

Fig. 11. Comparison of the three neural network architectures and their ensemble with weighted averaging for model 1.

5. Prediction results

For training the models, data for the working days in the cold period (from January 1st to March 31st and from November 1st to December 31st) for years 2009, 2010 and 2011 were used (318 samples in total), and for testing 2012 (100 samples). Data with obvious errors and heat meter malfunctions were removed from the dataset. To ensure that no special factor is dominant over the others, all inputs and outputs are normalized to the interval (0, 1) by a linear scaling function. The prediction accuracy is measured by the coefficient of determination (R^2), root mean square error (RMSE) and mean absolute percentage error (MAPE). In Tables 2 and 3, prediction indices for training and testing the models, respectively, are shown.

The results show that all analyzed network architectures can predict heating energy consumption with great accuracy. The coefficient of determination for all the network models is in the range of 0.9743 (for the FFNN for M3) and 0.9845 (for the ensemble AV for M1). Mean absolute percentage error is in the range from 3.44 (FFNN for M1) to 4.94 (RBFN for M3) for the training period and from 5.25 (FFNN for M3) to 6.01 (RBFN for M2) in the test period.

The other discussed issue is the influence of using different number of variables on prediction accuracy. In the analyzed case, reducing the number of input variables does not lead to significantly lower prediction quality. For example, for ensemble of neural networks with weighted average, when using all available input variables R^2 = 0.9843, while using only three most influencing variables gives coefficient of determination of 0.9766. Using less input variables has several advantages, especially when it comes to prediction of consumption not just one day ahead, but for example 2 months, or 1 year ahead. Besides the obvious, that less parameters needs to be monitored and measured, if input variable also needs to be predicted, the prediction error, which always occurs, will be accumulated for heating consumption prediction. Therefore, selection of variables plays crucial role in prediction modeling. It is shown that using ensemble of neural networks improves prediction accuracy. The improvement was expected, because the most probable is that the different neural networks will not make the mistakes at the same points. Then, using the ensemble, as a combination of ensemble members, errors can be compensated.

Figs. 5–10 show the comparison of the measured and predicted values for ensemble with simple averaging (AV) for model with nine predictor variables (M1), weighted average ensemble (WAV) for model with seven predictor variables (M2) and median based averaging ensemble for three most significant predictor variables (M3), respectively.

The part of the testing dataset, where three neural network architectures (FFNN, RBFN and ANFIS) and their ensemble with

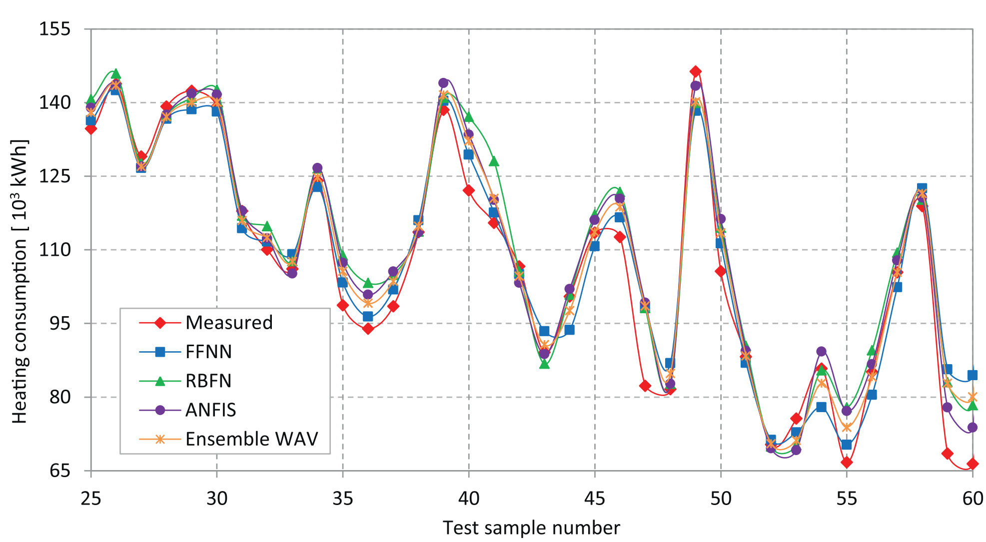

Fig. 12. Comparison of the three neural network architectures and their ensemble with weighted averaging for model 2.

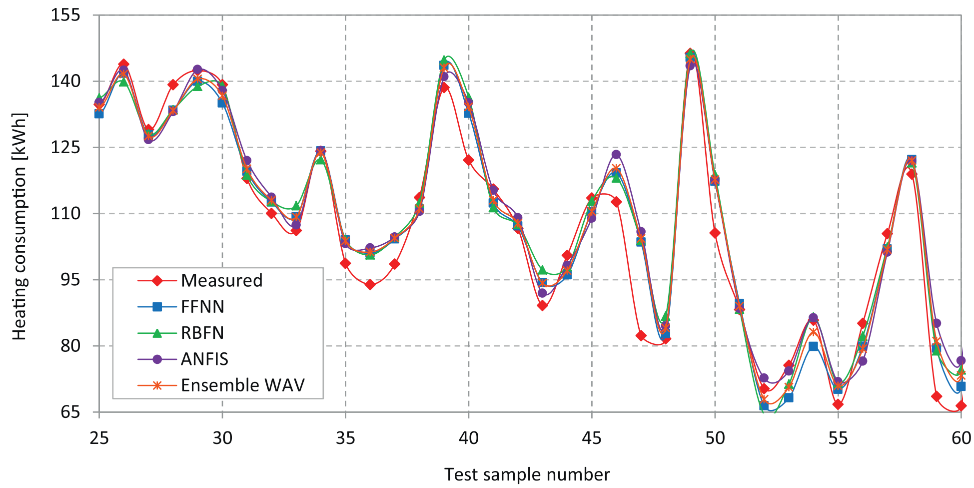

Fig. 13. Comparison of the three neural network architectures and their ensemble with weighted averaging for model 3.

weighted averaging can be compared, is presented in Figs. 11–13. It can be seen that in different data points, various networks show better accuracy. For example, at sample 45 in Fig. 12, measured heating consumption is 113,500 kW h, value predicted with RBFN is 117,144 kW h, with ANFIS 116,076 kW h, with FFNN 110,694 kW h, and with ensemble with weighted averaging 113,599 kW h. The advantage of the ensemble of neural networks is that it can compensate the errors of the networks, by combining their outputs, and therefore achieve better prediction than individual networks.

6. Conclusion

Application of three different neural network architectures (FFNN, RBFN and ANFIS) for heating energy consumption prediction for working days in the coldest period of the year (from January 1st to March 31st and from November 1st to December 31st) is analyzed. For model 1 all available input variables are used, for model 2 seven most influencing, and for model 3 only three. The results showed that all three different networks have excellent agreement with measured values. In order to improve the achieved prediction accuracy, ensemble of these three networks is created, while using three different combinations of members (simple average, weighted average and median). All networks and their ensembles were trained for 3 years period (2009–2011) and tested for the year 2012. The ensemble, by combining the outputs of member networks, achieves better prediction results. The relationship between input variables and heating consumption is very complex, especially when the observed object, such as university campus, consists of many different buildings. Hence, estimating building energy use, which is essential for energy supply strategy and capital investments, requires continuous development of new approaches for achieving more accurate results. It is shown that neural networks, and especially their ensembles, are very useful techniques that can be used for this type of predictions. In future work, the performance of each network could be improved by optimizing its architecture and parameters, for example by implementing some evolutionary algorithm; consequently the ensemble would be better.

Acknowledgments

The second author wants to thank Norwegian University of Science and technology for using data gathered during study visit to NTNU, as a part of the collaborative project Sustainable Energy and Environment in Western Balkans. The project was funded through the Norwegian Programme in Higher Education, Research and Development in the Western Balkans, Programme 3: Energy Sector (HERD Energy) for the period 2011–2013.

References

[1] E. Council, Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings, Off. J. Eur. Union 18 (201) (2010) 13–35.
[2] H.-X. Zhao, F. Magoulès, A review on the prediction of building energy consumption, Renew. Sustain. Energy Rev. 16 (6) (2012) 3586–3592.
[3] M.S. Al-Homoud, Computer-aided building energy analysis techniques, Build. Environ. 36 (4) (2001) 421–433.
[4] R. Yao, K. Steemers, A method of formulating energy load profile for domestic buildings in the UK, Energy Build. 37 (6) (2005) 663–671.
[5] A. Kusiak, M. Li, Z. Zhang, A data-driven approach for steam load prediction in buildings, Appl. Energy 87 (3) (2010) 925–933.
[6] M. Bauer, J.-L. Scartezzini, A simplified correlation method accounting for heating and cooling loads in energy-efficient buildings, Energy Build. 27 (2) (1998) 147–154.
[7] T. Catalina, J. Virgone, E. Blanco, Development and validation of regression models to predict monthly heating demand for residential buildings, Energy Build. 40 (10) (2008) 1825–1832.
[8] S. Katipamula, T. Reddy, D. Claridge, Multivariate regression modeling, J. Sol. Energy Eng. 120 (3) (1998) 177–184.
[9] J.C. Lam, K.K. Wan, S. Wong, T.N. Lam, Principal component analysis and long-term building energy simulation correlation, Energy Convers. Manage. 51 (1) (2010) 135–139.
[10] A.H. Neto, F.A.S. Fiorelli, Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption, Energy Build. 40 (12) (2008) 2169–2176.
[11] B.B. Ekici, U.T. Aksoy, Prediction of building energy consumption by using artificial neural networks, Adv. Eng. Softw. 40 (5) (2009) 356–362.
[12] Ö.A. Dombaycı, The prediction of heating energy consumption in a model house by using artificial neural networks in Denizli-Turkey, Adv. Eng. Softw. 41 (2) (2010) 141–147.
[13] L. Ekonomou, Greek long-term energy consumption prediction using artificial neural networks, Energy 35 (2) (2010) 512–517.
[14] B. Yuce, H. Li, Y. Rezgui, I. Petri, B. Jayan, C. Yang, Utilizing artificial neural network to predict energy consumption and thermal comfort level: an indoor swimming pool case study, Energy Build. 80 (2014) 45–56.
[15] K. Li, H. Su, J. Chu, Forecasting building energy consumption using neural networks and hybrid neuro-fuzzy system: a comparative study, Energy Build. 43 (10) (2011) 2893–2899.
[16] S. Karatasou, M. Santamouris, V. Geros, Modeling and predicting building's energy use with artificial neural networks: methods and results, Energy Build. 38 (8) (2006) 949–958.
[17] R. Kumar, R. Aggarwal, J. Sharma, Energy analysis of a building using artificial neural network: a review, Energy Build. 65 (2013) 352–358.
[18] Z.-H. Zhou, J. Wu, W. Tang, Ensembling neural networks: many could be better than all, Artif. Intell. 137 (1) (2002) 239–263.
[19] P. Melin, J. Soto, O. Castillo, J. Soria, A new approach for time series prediction using ensembles of ANFIS models, Expert Syst. Appl. 39 (3) (2012) 3494–3506.
[20] J.W. Taylor, R. Buizza, Neural network load forecasting with weather ensemble predictions, IEEE Trans. Power Syst. 17 (3) (2002) 626–632.

[21] K. Siwek, S. Osowski, R. Szupiluk, Ensemble neural network approach for accurate load forecasting in a power system, Int. J. Appl. Math. Comput. Sci. 19 (2) (2009) 303–315.
[22] J.-S. Jang, ANFIS: adaptive-network-based fuzzy inference system, IEEE Trans. Syst. Man Cybern. 23 (3) (1993) 665–685.
[23] T. Takagi, M. Sugeno, Fuzzy identification of systems and its applications to modeling and control, IEEE Trans. Syst. Man Cybern. 1 (1985) 116–132.
[24] L.K. Hansen, P. Salamon, Neural network ensembles, IEEE Trans. Pattern Anal. Mach. Intell. 12 (10) (1990) 993–1001.
[25] A.J. Sharkey, Multi-net systems, in: Combining Artificial Neural Nets, Springer, 1999, pp. 1–30.
[26] G.P. Zhang, V. Berardi, et al., Time series forecasting with neural network ensembles: an application for exchange rate prediction, J. Oper. Res. Soc. 52 (6) (2001) 652–664.
[27] A. Sretenovic, Analysis of energy use at university campus, Norwegian University of Science and Technology, Department of Energy and Process Engineering, 2013 (M.Sc. thesis).
[28] Energy Remote Monitoring. URL http://erm.tac.com/erm/
[29] Weather database. URL http://lmt.bioforsk.no/agrometbase/getweatherdata.php
[30] T. Tveito, Heating degree-days – present conditions and scenario for the period 2021–2050, 2002.
[31] T. Johannessen, Varmeutviklingen i bygninger og klimaet, 1956.
[32] K. Wojdyga, An influence of weather conditions on heat demand in district heating systems, Energy Build. 40 (11) (2008) 2009–2014.
[33] S.L. Chiu, Fuzzy model identification based on cluster estimation, J. Intell. Fuzzy Syst. 2 (3) (1994) 267–278.