MULTISTAGE ENSEMBLE OF FEEDFORWARD NEURAL NETWORKS FOR PREDICTION OF HEATING ENERGY CONSUMPTION

by

Radiša Ž. JOVANOVIĆ, Aleksandra A. SRETENOVIĆ*, and Branislav D. ŽIVKOVIĆ

Faculty of Mechanical Engineering, University of Belgrade, Belgrade, Serbia

Original scientific paper
DOI: 10.2298/TSCI150122140J

Feedforward neural network models are created for prediction of heating energy consumption of a university campus. Actual measured data are used for training and testing the models. Multistage neural network ensemble is proposed for the possible improvement of prediction accuracy. Previously trained feed-forward neural networks are first separated into clusters, using k-means algorithm, and then the best network of each cluster is chosen as a member of the ensemble. Three different averaging methods (simple, weighted, and median) for obtaining ensemble output are applied. Besides this conventional approach, single radial basis neural network in the second level is used to aggregate the selected ensemble members. It is shown that heating energy consumption can be predicted with better accuracy by using ensemble of neural networks than using the best trained single neural network, while the best results are achieved with multistage ensemble.

Key words: *heating consumption prediction, neural networks, k-means clustering, multistage ensemble*

Introduction

Considering the constant growth of energy consumption, especially after the emanation of the EPB European Directive [1], energy sustainability and greenhouse gas reduction have become a world-wide challenge. In Europe, buildings account for 40% of total energy use and 36% of total CO_2 emission [1], so the estimation of building consumption plays a very important role in energy management. Besides the conventional methods based on solving of equations describing heat transfer, identification and prediction methods with statistical analysis of real energy use are nowadays important topics of research. 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 in buildings (using measured and analyzed data). The classical approach to estimate the building energy use is based on the application of a model with the known system structure and properties, as well as forcing variables (forward approach, *white box*). Using software tools available on the market (TRNSYS, BLAST, ESP-r, HAP, APACHE, *etc.*) re-

* Corresponding author; e-mail: asretenovic@mas.bg.ac.rs

quires extensive knowledge of numerous building parameters (constructions, systems) and behavior, which are usually not available or difficult to collect, especially for the existing buildings. In recent years, there is an increased interest in a different approach to building energy analysis, which is based on the so called *inverse* or data-driven models [2]. In a data-driven approach, the development of the *inverse* model involves determination of mathematical relationship between independent and dependent variables. All input and output variables are required to be known and measured. These models are often called *black box* models, since there is no need for detailed knowledge of physical characteristics of a building. 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. The main drawback is that significant number of measured data is required. Creating these models involves various statistical techniques. Artificial neural networks (ANN), with their self-learning capability and possibility to be a universal approximator, are the most used artificial intelligence models for different types of prediction. Vujić [3] used experimental and meteorological data to develop a feedforward neural network (FFNN) model for prediction of daily concentrations of air pollution in city of Subotica, Serbia. Ganapathy *et al.* [4] predicted various emissions of Diesel engine with satisfying accuracy. Ćirić *et al.* [5] compared different computational intelligence methodologies based on ANN used for forecasting the emission of CO_2 in the city of Nis, Serbia. Ozener *et al.* [6] showed that the ANN approach can be used for accurately predicting characteristic values of an internal combustion engine. Ekici and Aksoy [7] developed the backpropagation three-layered ANN for the prediction of the heating energy requirements of various buildings. Dombayci [8] used hourly heating energy consumption for a model house calculated by degree-hour method for training and testing the ANN model. Ekonomou [9] compared the ANN prediction results with the real long-term heating energy consumption and results produced by linear regression and support vector machine model. A review of the different neural network models used for building energy prediction can be found in [10]. 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 [11]. Since an ensemble is often more accurate than its members, in recent years there have been many examples of successful applications in various fields: time series prediction [12], weather forecasting [13], load prediction in a power system [14]. In this paper, possible improvement of prediction accuracy by using ensemble of FFNN is investigated.

Feedforward neural network

The ANN is a computational structure inspired by a biological neural system. The FFNN architecture consists of an input layer, an output layer, and one or more hidden layers built of processors called neurons, which are fully interconnected with neurons in the subsequent layer using adaptable weighted connections. The non-linear activation functions in the hidden layer neurons enable the neural network to be a universal approximator. During training process, the weights are adjusted so that the network can produce the desired response to the given inputs. The backpropagation learning algorithm and the algorithms derived from it are the most widely applied for minimization of error function. 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 a single hidden layer and backpropagation learning algorithm are used. In the first stage of backpropagation learning algorithm (forward pass), synaptic weights of the network are all fixed and actual response of the network is obtained based on input training dataset presented to the network input layer. The network output is compared with the desired values, the error signal is propagated backwards, and the weights are modified consequently (backward pass).

Artificial neural network ensembles

In order to improve efficiency and overcome some generalization issues of single neural network models in various engineering problems, the concept of neural network ensemble was introduced. It has been proven that a simple combination of outputs of many neural networks can generate more accurate predictions and significantly improve generalization ability than that of any of the individual networks [15]. As the individual networks tend to make errors on different parts of the input space, various studies showed that a good ensemble is the one where the single networks (ensemble members) have both accuracy and diversity [16]. The key issue is how to select the aggregate members in order to attain the optimal compromise between these two conflicting conditions [17]. The accuracy described by some prediction indicator (commonly the mean squared error) is achieved by proper training algorithms of neural networks. The most widely used approaches for obtaining diverse individual predictors (members) can be divided into three groups [18, 19]. The 1st group of methods refers to training individuals on different adequately-chosen subsets of the data set, including two important techniques: bagging [20] and boosting [21]. The 2nd group uses variation of topologies, by varying number of input and/or hidden nodes, initial weight sets, training algorithms, or even networks with different types. The 3rd group is named selective approach group, where the diverse components are selected from a number of accurately trained networks. Various algorithms for selecting ensemble components can be found in literature, such as: generic algorithm [22], pruning algorithm [23], selective algorithm based on bias/variance decomposition [24], *etc.* Clustering technology can be used to divide all networks into some groups (clusters) according to similarity of the networks. Then, the most accurate individual in each group of the validation set is selected, and finally, all selected individuals construct the ensemble.

The k-means for selecting ensemble members

Qiang *et al.* [16] created the clustering-based selective neural network ensemble using k-means and compared this method with two main ensemble approaches: bagging and boosting. Due to the ease of implementation, simplicity, efficiency and empirical success, k-means clustering, proposed by MacQueen [25], is one of the most popular methods for dataset partitioning. In selecting neural network ensemble members by using k-means, the goal is to divide prediction data achieved by individual networks $y = \{y_1, \ldots, y_r\}$ into m clusters, where number of elements in each cluster is n_i, and the center of cluster is c_i. So, clustering can be achieved by finding c_i which makes:

$$J_e = \sum_{i=1}^{m} \sum_{j=1}^{n_i} \left\| y_j^{(i)} - c_i \right\|^2 \qquad (1)$$

minimized. Consequently, after clustering, the diversity between networks in different cluster groups is greater than the one within the same group. The diversity is maintained by choosing the most accurate network in each group as a member of the ensemble. In k-means algorithm, cluster number m must be determined in advance, and the optimal value can be selected comparing ensemble prediction indices by trial and error methods. As one of the most popular approaches for combining the selected network outputs, the linear combination of the outputs of ensemble members is used (simple, weighted, and median based averaging). Different approach comprises using neural network for combining the selected ensemble members, as in [26] where authors proposed the system consisting of two ANN assembled in a hierarchical order. In this paper, radial basis function network (RBFN) is proposed for combining ensemble members selected by k-means clustering.

Radial basis function network

In this paper, a RBFN, as FFNN, is considered as multi-input, single-output system consisting of an input layer, one hidden layer, and an output layer. The RBFN uses the nonlinear radially symmetrical function as an activation function in the hidden layer, and the output of each hidden neuron depends only on the radial distance between the input vector and the center of the hidden neuron. The Gaussian function is adopted as activation function in this study. For a RBFN with an n-dimensional input $x \in R^n$, the output of the j^{th} hidden neuron is given by:

$$h_j(x) = \varphi_j(\|x - \vec{c}_j\|), \qquad j = 1, 2, ..., m \tag{2}$$

where $\vec{c}_j$ is the center (vector) of the j^{th} hidden neuron, m – the number of neurons in the hidden layer, and $\phi(\cdot)$ – the radial basis function. The output layer neurons have a linear transfer function. The k^{th} output of the network is obtained by the weighted summation of the outputs of all hidden neurons connected to that output neuron:

$$\hat{y}_k(x) = \sum_{j=1}^{m} w_{kj} h_j(x) + w_{k0} \tag{3}$$

where w_{kj} is the connecting weight between the j^{th} hidden neuron and the k^{th} output unit, w_{k0} – the bias, and m – the number of the hidden layer neurons. The training of RBFN usually comprises two stages: unsupervised procedure (for example some clustering algorithm or self-organizing map) to adjust parameters of radial basis function (centers and widths) in the 1st stage, and in the 2nd stage adopting of weights of the output layer by applying supervised training algorithm (such as gradient method or the least mean square algorithm).

Case study

University campuses, as specific groups of diverse buildings (classrooms, sport facilities, laboratories, kitchen, *etc.*) with significant energy consumption, can represent a small-scale town [27]. Also, because of their educational purpose, they are adequate examples for students to understand energy consumption patterns of group of *mixed use* buildings. Norwegian University of Science and Technology (NTNU) campus Gloshaugen consists of 35 buildings, with total area of approximately 300,000 m^2. Building and energy management system and web-based energy monitoring system (energy remote monitoring – ERM) are available at NTNU. There are 46 heating meters installed in campus. Hourly heat and electricity consumption from all meters can be collected on ERM [28]. The main meter is installed by the district heating supplier, so it was taken as relevant. Daily heating energy consumption was analyzed in this paper.

Data pre-processing

Meteorological data were gathered from the local weather station Skjetlein, Leinstrand, Norway [29]. The heating season is defined as the period starting from the day the mean daily temperature falls below 11 °C during the autumn and lasting until the day it rises above 9 °C during the spring [30]; it usually lasts 251-280 days [31]. The optimal number of neural network models is determined by the results of analysis of the mean daily outside temperatures for years 2006 until 2014, as the most influencing variable on heating energy consumption. The average mean daily temperatures, with maximum and minimum for the specific date [29] are shown in fig. 1. Consequently, database is divided in periods:

– cold period – from January 1st until March 31st and from November 1st until December 31st,
– mild period – from April 1st until June 15th and from September 16th until October 31st, and
– warm period (outside the heating season) – June 15th until September 15th is excluded from the analysis.

This database partitioning suggests that separate network models should be developed for each period, rather than using one for the whole year. In this paper, only the cold period (with biggest heating energy consumption) is analyzed.

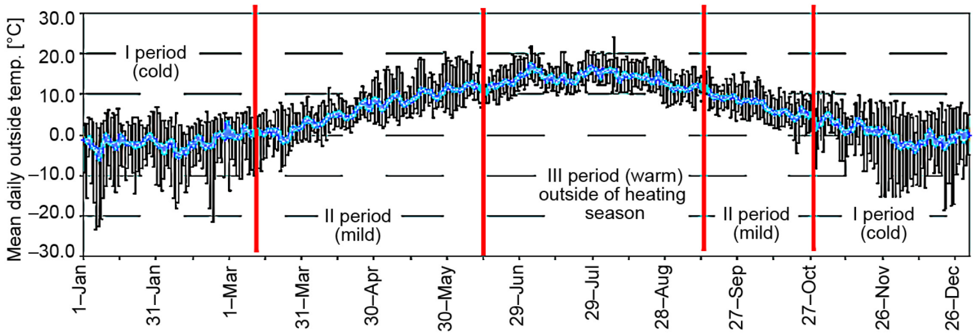


Figure 1. Average mean daily temperature for years 2006-2014

The influence of the day type on heating energy consumption is also analyzed. Based on the correlation of the daily heating energy use and the mean daily outside temperature for each day of the year 2012 (fig. 2), it can be seen that there is no significant difference among working days. Heating is not switched off during weekends, only the design set-point

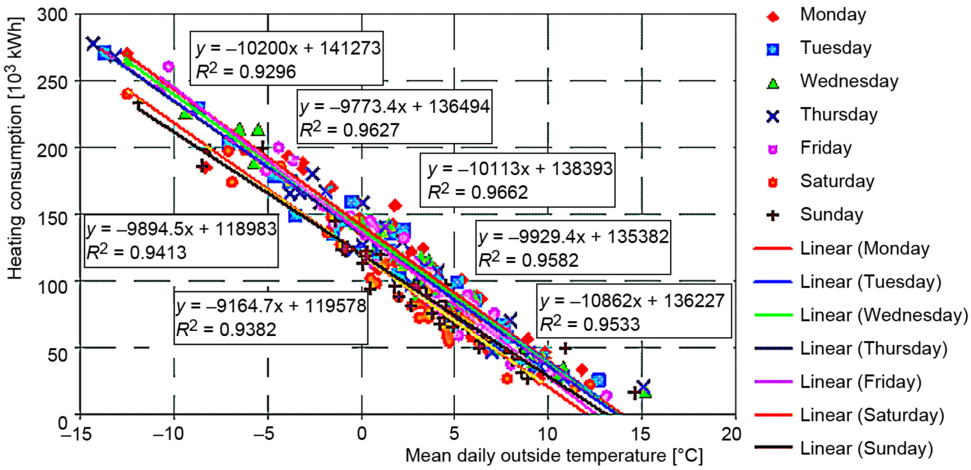


Figure 2. Correlation of the daily heating energy consumption with the mean daily outside temperature for the year 2012

is lowered, thus lower the trendlines for saturday and sunday. Holidays and exam periods have the same set-points as regular working days. These conclusions implicate that there should be created two networks: one for the working days, and other for the weekend. In this paper, the network for working days in the cold period is analyzed.

The ANN model development

The selection of the input variables plays a key role in building the ANN prediction model, so there are many different studies dealing with the impact of various variables on energy consumption. The influence of hourly values of solar radiation and wind speed on heating demands of building complex heated by district heating system was confirmed in the empirical research conducted by Wojdyga [32]. The input variables for the neural network model, considered in this study, are: the mean daily outside temperature [°C], the mean daily wind speed [ms^{-1}], total daily solar radiation [Whm^{-2}], the minimum daily temperature [°C], the maximum daily temperature [°C], relative humidity [%], day of the week, month of the year. As an additional input, it is selected the heating consumption of the previous day based on the results of partial autocorrelation, which measures how a series is correlated with itself at different lags. The ANN architecture used in this study is a three-layer FFNN composed of one input layer, one output layer with linear activation functions and one hidden layer with sigmoidal activation functions and Levenberg-Marquardt learning algorithm. Number of neurons in hidden layer is varied and the best results are achieved with ten neurons. For training all the networks, data for the working days in the cold period (from January 1st until March 31st and from November 1st until December 31st) of the years 2009, 2010, and 2011 were used (318 samples in total), and the year 2012 (100 samples) for testing. Data with obvious errors and heat meter malfunctions were removed from the dataset. All variables are normalized to values between 0 and 1. The prediction accuracy is measured by the coefficient of determination (R^2), root mean square error (RMSE), and the mean absolute percentage error (MAPE).

Neural network ensemble

The main idea of this paper, is to propose a multistage neural network ensemble. First, 50 separately trained FFNN with satisfying accuracy are chosen as possible members. Next task is to ensure also the diversity between individual members. The k-means clustering technology is employed to divide all networks into some groups (clusters) according to similarity of the network outputs. Then, the most accurate individual in each group on the validation set is selected. In that way, both accuracy and diversity are gradually achieved. Finally, all selected individuals construct the ensemble.

In this paper, four different methods (combinations of members) for creating ensemble are used:

– the simple average, determines the ensemble output by taking the average of all outputs provided by the individual classifiers (ensemble SAV),
– weighted average, the ensemble output is given by the weighted average of its components (ensemble WAV),
– median based averaging (ensemble MAV), and
– multistage ensemble: RBFN in the 2nd level.

The proposed algorithm for creating ensembles is shown in fig. 3. Besides first three conventional methods, that are most widely used, the multistage approach, which is expected to give an even better improvement in accuracy, is proposed. The 2nd level consists of a single RBFN, which is used as an ensemble`s integrator of selected members in order to

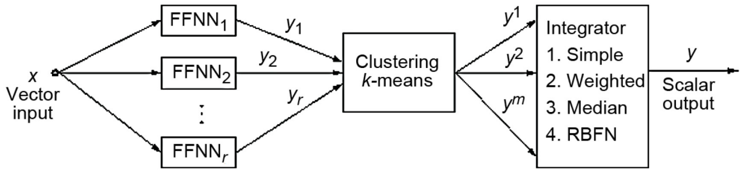


Figure 3. Proposed neural network ensemble

approximate the best combination function. A customized RBFN function available in MATLAB, which iteratively creates a radial basis network one neuron at a time, is used to develop the second stage network. The radius value (known as spread) of the radial basis function was varied for the best performance of the RBFN.

Results and discussion

The parameter to be predefined in k-means clustering is the number of clusters. Since there is no perfect mathematical criterion for selecting the optimal number of clusters, numerous heuristics and index methods are available in literature. The other way is to run k-means separately for various cluster numbers, and select the one where the ensemble prediction achieves the best results. Therefore, in this study, number of clusters is varied from 2-10 for all four methods. The number of networks in the ensemble is equal to cluster number because one best network in each cluster is selected to join the ensemble. In tabs. 1 and 2, prediction indices for training and testing the models, respectively, are presented.

Table 1. Prediction indices for training networks

Number of clusters		2	3	4	5	6	7	**8**	9	10
R^2 [–]	Best FFNN	0.9850	0.9850	0.9850	0.9850	0.9850	0.9850	0.9850	0.9850	0.9850
	Ensemble SAV	0.9866	0.9863	0.9863	0.9866	0.9866	0.9873	0.9870	0.9869	0.9867
	Ensemble WAV	0.9866	0.9863	0.9865	0.9870	0.9870	0.9877	0.9876	0.9875	0.9874
	Ensemble MAV	0.9866	0.9857	0.9858	0.9860	0.9860	0.9870	0.9867	0.9872	0.9868
	Multistage	0.9866	0.9864	0.9866	0.9870	0.9871	0.9877	**0.9877**	0.9874	0.9874
RMSE [kWh]	Best FFNN	6942.4	6942.4	6942.4	6942.4	6942.4	6942.4	6942.4	6942.4	6942.4
	Ensemble SAV	6445.8	6648.0	6524.1	6417.6	6407.0	6240.5	6298.2	6337.5	6415.2
	Ensemble WAV	6447.2	6550.4	6440.4	6312.2	6294.6	6130.3	6143.2	6180.2	6198.5
	Ensemble MAV	6445.8	6705.7	6619.1	6563.7	6543.5	6290.8	6379.0	6266.0	6388.3
	Multistage	6377.0	6421.2	6373.3	6279.7	6262.6	6119.1	**6128.0**	6200.3	6178.6
MAPE [%]	Best FFNN	3.5614	3.5614	3.5614	3.5614	3.5614	3.5614	3.5614	3.5614	3.5614
	Ensemble SAV	3.3706	3.3524	3.2418	3.2276	3.2403	3.1588	3.1065	3.1749	3.2514
	Ensemble WAV	3.3654	3.3193	3.2169	3.2278	3.2329	3.1504	3.1062	3.1883	3.1900
	Ensemble MAV	3.3706	3.3657	3.2375	3.2420	3.2602	3.1366	3.1070	3.1545	3.2301
	Multistage	3.3138	3.1980	3.1661	3.1655	3.1946	3.1354	**3.0870**	3.1699	3.1727

The best trained FFNN is used as a reference. The results show that no matter how many networks are in the ensemble, it is always better than the single FFNN. Comparing to

Table 2. Prediction indices for testing networks

Number of clusters		2	3	4	5	6	7	**8**	9	10
R^2 [–]	Best FFNN	0.9773	0.9773	0.9773	0.9773	0.9773	0.9773	0.9773	0.9773	0.9773
	Ensemble SAV	0.9806	0.9804	0.9803	0.9811	0.9816	0.9814	0.9818	0.9804	0.9804
	Ensemble WAV	0.9806	0.9801	0.9801	0.9814	0.9818	0.9815	0.9821	0.9811	0.9813
	Ensemble MAV	0.9806	0.9801	0.9809	0.9821	0.9821	0.9806	0.9818	0.9812	0.9807
	Multistage	0.9805	0.9803	0.9809	0.9818	0.9820	0.9818	**0.9821**	0.9814	0.9815
RMSE [kWh]	Best FFNN	9829.7	9829.7	9829.7	9829.7	9829.7	9829.7	9829.7	9829.7	9829.7
	Ensemble SAV	9101.2	9466.1	9044.5	8877.1	8710.5	8895.1	8622.6	8823.7	9108.2
	Ensemble WAV	9065.5	9228.0	8902.9	8720.5	8644.3	8812.2	8702.2	8991.4	9004.5
	Ensemble MAV	9101.2	9230.2	8870.8	8709.3	8582.1	8896.6	**8581.7**	8700.2	9085.7
	Multistage	8655.0	8650.1	8681.0	8449.7	8445.0	8625.0	8547.9	8715.4	8806.8
MAPE [%]	Best FFNN	6.3049	6.3049	6.3049	6.3049	6.3049	6.3049	6.3049	6.3049	6.3049
	Ensemble SAV	5.8460	6.2800	6.0255	5.8500	5.7336	5.7953	5.5400	5.6073	5.8726
	Ensemble WAV	5.8198	6.0304	5.9122	5.6944	5.6270	5.7372	5.6496	5.6831	5.7635
	Ensemble MAV	5.8460	6.0520	5.9033	5.6614	5.6158	5.6827	5.4982	5.5433	5.8601
	Multistage	5.5830	5.5774	5.6545	5.5134	5.5170	5.6088	**5.4934**	5.5277	5.5674

conventional averaging methods, multistage ensemble with RBFN in the 2nd level offers the improved performance, as can be seen in figs. 4 and 5. Multistage network ensemble prediction model can be viewed as a non-linear mapping that can be represented by:

$$\hat{y} = f(y^1, y^2, \ldots y^m) \tag{4}$$

where $y^1, \ldots y^m$ are the outputs of the selected single neural network predictors, and $\hat{y}$ – the aggregated output (ensemble prediction result), and f – the function determined by RBFN training.

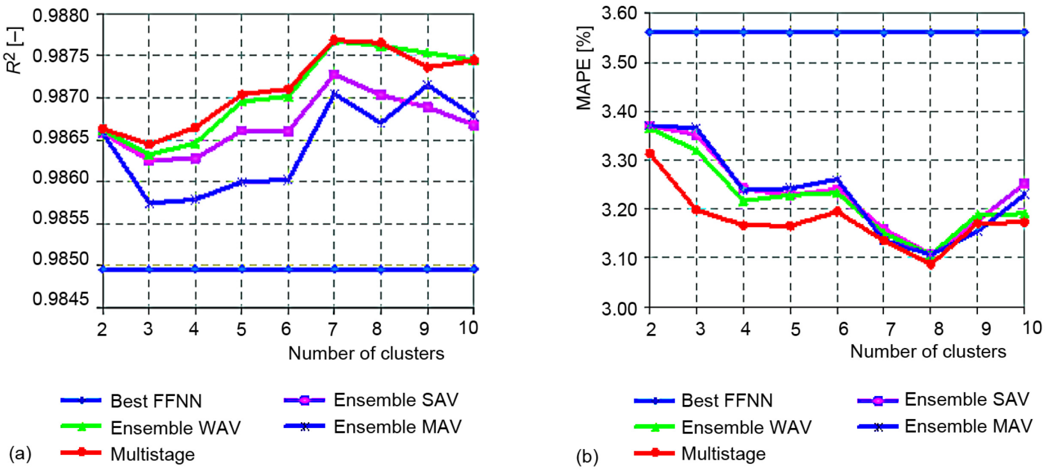


Figure 4. Prediction indices of all models with different number of clusters for training

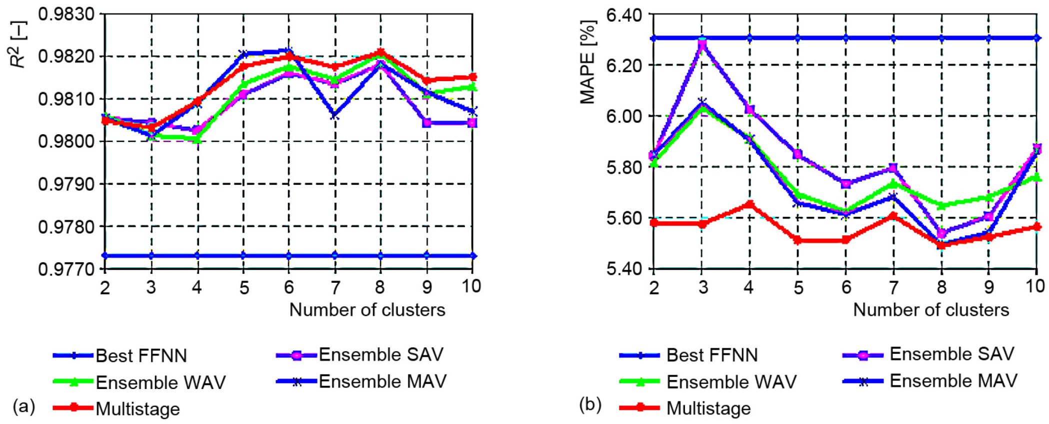


Figure 5. Prediction indices of all models with different number of clusters for testing

Best results are achieved using the multistage ensemble with eight clusters, R^2 = = 0.9877, RMSE = 6128 kWh, MAPE = 3.0870% for training, and R^2 = 0.9821, RMSE = = 8581.7 kWh, MAPE = 5.4934% for testing data. Figures 6 and 7 present the comparison of the prediction results using multistage ensemble of eight neural networks with the actual measured heating energy consumption data for the training and test period, respectively.

Figure 6. Prediction results of the multistage ensemble with eight clusters for the training period

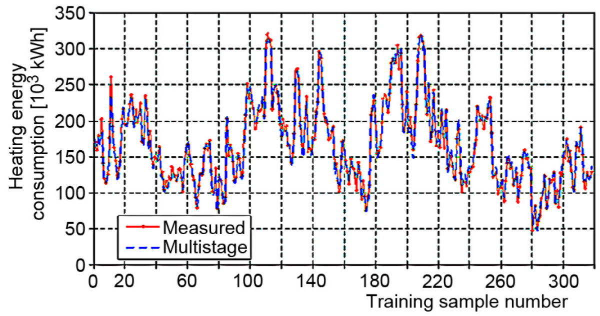
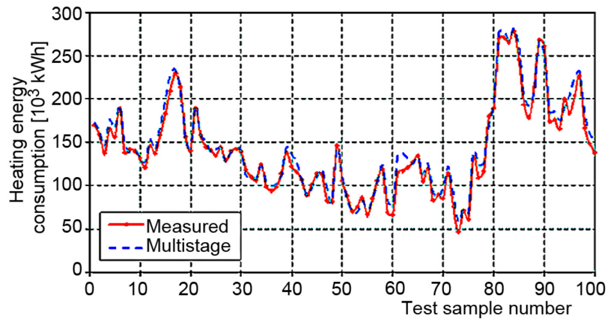


Figure 7. Prediction results of the multistage ensemble with eight clusters for the test period



Conclusion

The 1st stage of creating neural network ensemble for prediction of heating energy use of NTNU campus Gloshaugen consists of training 50 different FFNN based on the coldest period in the years 2009, 2010, and 2011 (318 samples), and tested for the year 2012 (100 samples). The main issue in ensemble technique is to achieve both accuracy and diversity of ensemble members. The accuracy is obtained by applying adequate training algorithm and selecting the number of neurons in the hidden layer by trial and error methods. The k-means, as one of the most used clustering technique, is used to separate previously trained networks into clusters, and the best network of each cluster is selected as the ensemble member. Ensemble is trained and tested for various numbers of clusters. Three different averaging methods are used for creating ensemble: simple, weighted, and median based. Averaging the predictions of these networks resulted in an improvement in accuracy over the predictions of the best trained individual FFNN. New individual neural network is used as an integrator, by taking FFNN outputs as input. Multistage model, using RBFN in the 2^{nd} level is proven to be most effective for various numbers of clusters. In this paper, we have demonstrated that multistage ensembles, where the adaptive properties of a second layer network are used to combine the outputs of the individual ensemble members, offer enhanced performance over conventional combining methods and best trained single network.

Acknowledgment

Data used for this paper were gathered during the 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] ***, Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the Energy Performance of Buildings, Official Journal of the European Union 2010, L153 (2010), pp. 13-35
[2] Kusiak, A., *et al.*, A Data-Driven Approach for Steam Load Prediction in Buildings, *Applied Energy*, *87* (2010), 3, pp. 925-933
[3] Vujić, B., *et al.*, Experimental and Artificial Neural Network Approach for Forecasting of Traffic Air Pollution in Urban Areas: The Case of Subotica, *Thermal Science*, *14* (2010), Suppl. 1, pp. S79-S87
[4] Ganapathy, T., *et al.*, Artificial Neural Network Modeling of Jatropha Oil Fueled Diesel Engine for Emission Predictions, *Thermal Science*, *13* (2009), 3, pp. 91-102
[5] Ćirić, I. T., *et al.*, Air Quality Estimation by Computational Intelligence Methodologies, *Thermal Science*, *16* (2012), Suppl. 2, pp. S493-S504
[6] Ozener, O., *et al.*, Artificial Neural Network Approach to Predicting Engine-Out Emissions and Performance Parameters of a Turbo Charged Diesel Engine, *Thermal Science*, *17* (2013), 1, pp. 153-166
[7] Ekici, B. B., Aksoy, U. T., Prediction of Building Energy Consumption by Using Artificial Neural Networks, *Advances in Engineering Software*, *40* (2009), 5, pp. 356-362
[8] Dombayci, O. A., The Prediction of Heating Energy Consumption in a Model House by Using Artificial Neural Networks in Denizli, Turkey, *Advances in Engineering Software*, *41* (2010), 2, pp. 141-147
[9] Ekonomou, L., Greek Long-Term Energy Consumption Prediction Using Artificial Neural Networks, *Energy*, *35* (2010), 2, pp. 512-517
[10] Kumar, R., *et al.*, Energy Analysis of a Building Using Artificial Neural Network: A Review, *Energy and Buildings*, *65* (2013), Oct., pp. 352-358
[11] Zhou, Z.-H., *et al.*, Ensembling Neural Networks: Many Could be Better Than All, *Artificial Intelligence*, *137* (2002), 1, pp. 239-263
[12] Melin, P., *et al.*, A New Approach for Time Series Prediction Using Ensembles of ANFIS Models, *Expert Systems with Applications*, *39* (2012), 3, pp. 3494-3506
[13] Taylor, J. W., Buizza, R., Neural Network Load Forecasting with Weather Ensemble Predictions, *Power Systems, IEEE Transactions*, *17* (2002), 3, pp. 626-632

[14] Siwek, K., *et al.*, Ensemble Neural Network Approach for Accurate Load Forecasting in a Power System, *International Journal of Applied Mathematics and Computer Science*, *19* (2009), 2, pp. 303-315
[15] Hansen, L. K., Salamon, P., Neural Network Ensembles, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12* (1990), 10, pp. 993-1001
[16] Qiang, F., *et al.*, Z. Clustering-Based Selective Neural Network Ensemble, *Journal of Zhejiang University SCIENCE A*, *6* (2005), 5, pp. 387-392
[17] Granitto, P. M., *et al.*, Neural Network Ensembles: Evaluation of Aggregation Algorithms, *Artificial Intelligence*, *163* (2005), 2, pp. 139-162
[18] Sharkey, A. J., *Types of Multi-Net Systems*, Springer, Berlin, 1999
[19] Zhang, G. P., Berardi, V., Time Series Forecasting with Neural Network Ensembles: An Application for Exchange Rate Prediction, *Journal of the Operational Research Society*, *52* (2001), 6, pp. 652-664
[20] Breiman, L., Bagging Predictors, *Machine Learning*, *24* (1996), 2, pp. 123-140
[21] Schapire, R. E., The Strength of Weak Learnability, *Machine Learning*, *5* (1990), 2, pp. 197-227
[22] Opitz, D. W., Shavlik, J. W., Actively Searching for an Effective Neural Network Ensemble, *Connection Science*, *8* (1996), 3-4, pp. 337-354
[23] Lazarevic, A., Obradovic, Z., Effective Pruning of Neural Network Classifier Ensembles, *Proceedings,* IJCNN'01-International Joint Conference on Neural Networks, Washington DC, USA, 2001
[24] Navone, H. D., *et al.*, Selecting Diverse Members of Neural Network Ensembles, *Proceedings,* 6^{th} Brazilian Symposium on Neural Networks, Rio de Janeiro, Brazil, 2000
[25] MacQueen, J., Some Methods for Classification and Analysis of Multivariate Observations, *Proceedings*, 5^{th} Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, Berkeley, Cal., USA, 1967.
[26] Ilić, S. A., *et al.*, Hybrid Artificial Neural Network System for Short-Term Load Forecasting, *Thermal Science*, *16* (2012), Suppl. 1, pp. S215-S224
[27] Sretenovic, A., Analysis of Energy Use at University Campus, Master, Norwegian University of Science and Technology, Trondheim, Norwey, 2013
[28] ***, Energy Remote Monitoring, http://erm.tac.com/erm/
[29] ***, AgroBase Weather Database, http://lmt.bioforsk.no/agrometbase/getweatherdata.php
[30] Johannessen, T. W., *Varmeutviklingen I bygninger og klimaet* (*Heat Transfer of Buildings and the Climate* – in Norwegian), Norges Byggforskingsinstitutt, Oslo, 1956
[31] Skaugen, T. E., *et al.*, Heating Degree-Days – Present Conditions and Scenario for the Period 2021-2050, DNMI Report 01/02KLIMA, Norwegian Meteorological Institute, Oslo, Norway, 2002
[32] Wojdyga, K., An Influence of Weather Conditions on Heat Demand in District Heating Systems, *Energy and Buildings*, *40* (2008), 11, pp. 2009-2014

Paper submitted: January 22, 2015
Paper revised: August 8, 2015
Paper accepted: September 7, 2015