SUPPORT VECTOR MACHINE FOR THE PREDICTION OF HEATING ENERGY USE

Aleksandra A. SRETENOVIĆ [a,*], Radiša Ž. JOVANOVIĆ[a], Vojislav M. NOVAKOVIĆ[b], Nataša M. NORD[b], Branislav D. ŽIVKOVIĆ[a]

[a]University of Belgrade, Faculty of Mechanical Engineering, Serbia

[b]Norwegian University of Science and Technology, Trondheim, Norway

Prediction of the building energy use for heating is very important for adequate energy planning. In this paper the daily district heating use of one university campus was predicted using the Support Vector Machine (SVM) model. SVM is the artificial intelligence method that has recently proved that it can achieve comparable, or even better prediction results than the much more used artificial neural networks. The proposed model was trained and tested on the real, measured data. The model accuracy was compared with the results of the previously published models (various neural networks and their ensembles) on the same database. The results showed that the SVM model can achieve better results than the individual neural networks, but also better than the conventional and multistage ensembles. It is expected that this theoretically well known methodology finds wider application, especially in prediction tasks.

Key words: *heating use prediction, support vector machine, artificial intelligence models*

1. Introduction

In Europe, building sector is responsible for 40% of total energy use and 36% of total CO_2 emission [1], so constant improvement of energy efficiency is the key action for the researchers. The popular axiom that „you cannot improve what you cannot measure“ points out the importance of the accurate estimation of energy consumption. There are various methodologies for classification of the methods for energy use prediction [2]. In [3] authors proposed a detailed review of the existing techniques, classifying them into three main categories: „black box“, „white box“ and „grey box“. The classical approach to estimate the building energy use which is based on the application of a model with the known system structure and properties, as well as the forcing variables (forward approach) belongs to the first category ("white box"). It requires extensive knowledge of the building and systems and involves using some of the available softwares for building performance simulation. The "inverse" or data-driven models, as a new methodology for the analysis of energy use, are recently gaining popularity [4]. These methods determine the mathematical relationship between independent and dependent variables. In order to develop these models it is necessary that all input and output variables are identified and measured, which is their main drawback (substantial amount of data). They are the representatives of „black box" modeling category, without requiring detailed knowledge of the

* Corresponding author: asretenovic@mas.bg.ac.rs

physical characteristics of a building. The data-driven approach is useful when the building is already built, and actual consumption data are measured and available. Creating these models involves various statistical techniques, from the simplest using regression up to very complex artificial intelligence. The “grey box” methods are the combination of these two methods, by using both physical and statistical techniques. Artificial intelligence involves relatively new techniques that have been used for solving problems in various engineering fields in the last couple of decades. The majority of these methods were originally developed for solving classification problems, which is still their most widely used application range. Their application is later extended to regression problems, such as prediction of building energy consumption, which are complex, especially due to significant number of influencing variables, as well as the non-linear relationships. Artificial neural networks (ANN) belong to the “black box” methods, and because of their self-learning capability and possibility to be an universal approximator, are the most used artificial intelligence models for different types of prediction. In [5] feedforward neural network (FFNN) is used for the prediction of daily concentrations of air pollution. Ćirić et al. in [6] compared various computational intelligence methodologies based on ANNs for forecasting the emission of CO_2. Özener et al. [7] used ANN for the prediction of the characteristic values of an internal combustion engine. ANN can be successfully used for modelling ground couple heat-pump [8]. A review of the different neural network models used for building energy prediction can be found in [9]. In [10] k-means clustering was used to group 50 separately trained FFNNs and the best trained network of each cluster was selected as a member for creating multistage ensemble for prediction of campus heating energy use. In [11] different NN architectures were compared: FFNN, Radial basis function network (RBFN) and Adaptive neuro-fuzzy inference system (ANFIS) and their ensembles are created. Ensemble, as a technique of combining individual network`s outputs, achieves higher accuracy. In [12] various multistage ensembles were compared. The k-means clustering was used for resampling training dataset before creating the ensembles in [13]. Support Vector Machine (SVM) is the artificial intelligence method that is nowadays gaining popularity. Similar like other artificial intelligence methods, SVM is more often used as a classifier: in [14] the authors used SVM classifier to determine whether the segmented object is human or not, and used this information for control of person-following robot platform. In [15] the ANN prediction results were compared with the real heating energy consumption and results produced by linear regression and SVM. At earlier stage of the theory development, SVM has been considered not so successful method for solving practical problems, but recently it has proven that it can achieve same, or even better results comparing to the widely used neural networks [16]. SVM model was successfully used so far to predict diesel engine performance [17]. In [18] authors compared the SVM with various NN models for the prediction of the cooling load in the office building. SVM was successfully used for modeling the performance of heat pump systems in [19] and [20] and for prediction of the efficiency of the solar air heater in [21]. A review on applications of ANN and SVM for building electricity consumption forecasting can be found in [22]. Although it is gaining popularity, there are not much examples of successful models for prediction of building`s energy use that can be found in recent literature. Therefore, the main idea of this paper is to examine the possibility of using SVM for solving this important engineering problem. In this paper, SVM model for the prediction of daily heating energy use of the university campus was developed and it was compared with the previously published results for the same case study using the same database. The results may be used heating energy prediction and for developing a reliable heating energy billing system.

2. Support Vector Machine

SVM is the universal approximator of any multivariate function to any desired degree of accuracy [22], same (or similar) as the Neural Network (NN). There is a difference in the development between these two most common used statistical techniques. SVM was developed from theory, and later it was implemented in practice and experiments, while the NN followed the path in reverse order: from application and extensive experiments to theory. The foundations of SVM have been developed by Vapnik and co-workers in 1964/65 [16], but until recently it was widespread the opinion that it was not suitable for practical application. Today, SVMs show better (or comparable) results than NNs and other statistical models. Traditional NN approaches have suffered difficulties with generalization, producing models that can overfit the data [23]. This is a consequence of the optimization algorithms used for parameter selection and the statistical measures used to select the optimal model. SVMs are the so-called nonparametric model, which means that their "learning" (training) is the crucial issue. The parameters are not predefined and their number depends on the training data used. In other words, parameters that define the capacity of the model are data-driven in such a way as to match the model capacity to data complexity [22]. This is a basic paradigm of the Structural Risk Minimization (SRM) introduced by Vapnik which has been shown to be superior to traditional Empirical Risk Minimization (ERM) principle, employed by conventional neural networks. SRM of the generalization error consisting of the sum of the training error and a confidence level based on Vapnik–Chernoverkis (VC) dimension, as opposed to ERM that minimizes the error only on the training data. Due to this, SVM has greater ability to generalize, which is the goal in statistical learning. Another key characteristic of SVM is that training of SVM is equivalent to solving a linearly constrained quadratic programming problem so that the solution of SVM is always unique and globally optimal, unlike other network's training which requires non-linear optimization with the danger of getting stuck into local minimum. In SVM, the solution to the problem is only dependent on a subset of training data points which are referred as support vectors. Using only support vectors, the same solution can be obtained as using all the training data points. SVMs were developed to solve the classification problem, but recently they have been extended to the domain of regression problems, which is often referred as Support Vector Regression (SVR) [24]. Consider a set of training points $\{(x_1, y_1), \ldots, (x_l, y_l)\}$, where $x_i \in \Re^d$ is a feature vector and $y_i \in \Re$ is the target output. The support vector regression for the linear case finds a linear function that can best approximate the actual output vector y, with a tolerance ε, and is as flat as possible. The regression function can be expressed as:

$$f(w,b) = w \cdot x + b \tag{1}$$

where w and b are the parameter vectors of the function. If the optimization problem is written in its dual form (which is not detailed herein), the input vectors are multiplied as dot product, which enables the use of the „kernel trick" for extending SVM to nonlinear case. In order to expand the problem to the non-linear case, one possible idea is to map the input data in a higher dimensional feature set, and perform linear regression in the feature set. Let x_i be mapped into a feature space by a non-linear function $\varphi(x)$, than the Equation (1) becomes:

$$f(w,b) = w \cdot \varphi(x) + b \tag{2}$$

The non-linear regression problem can be expressed as the following optimization problem:

$$\min_{w,b,\xi,\xi^*} \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}\left(\xi_i + \xi_i^*\right)$$

$$\text{subject to} \quad \begin{cases} y_i - w\cdot\varphi(x_i) - b \le \varepsilon + \xi_i \\ w\cdot\varphi(x_i) + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases} \tag{3}$$

where ξ_i and ξ_i^* are slack variables that specify the upper and the lower training errors subject to an error tolerance ε, and C is the positive constant, which determines the trade-off between the flatness of function and the empirical error. Minimizing the first term is equivalent to minimizing the confidence interval of the learning machine and minimizing the second term corresponds to minimizing the empirical risk. By introducing a dual set of Lagrange numbers optimization problem can be solved more easily in dual form, which is not detailed herein (more details can be found in [24]). The nonlinear Support Vector Regression is illustrated in Figure 1. The region enclosed by the tube is called ε -insensitive zone. The values with excess positive and negative deviations are denoted with ξ_i and, ξ_i^*, respectively and they are called the „slack variables“. The optimization criterion penalizes those data points whose values of y lie more than ε distance away from the fitted function $f(x)$. If the predicted value is within the tube, the loss is zero, while if the predicted point is outside of the tube, the loss is magnitude of the difference between the predicted value and the radius ε of the tube.

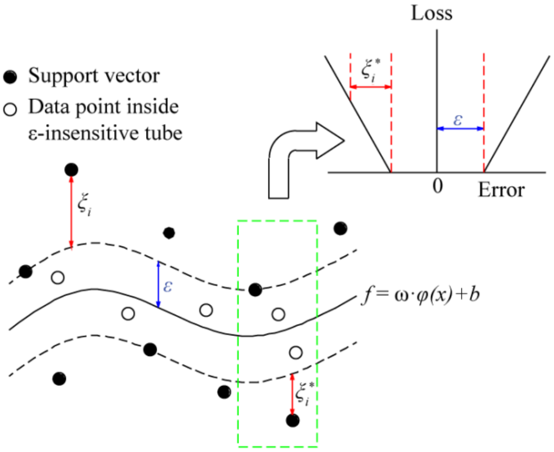

Figure 1. The non-linear Support Vector Regression

By the use of kernels, all necessary computations can be performed directly in input space, without having to compute the mapping. Some commonly used kernels are: linear, polynomial, sigmoid and radial basis function (RBF) kernel. The RBF kernel nonlinearly maps samples into a higher dimensional space and it can handle the case when the relationship between dependent and independent variables is non-linear. Therefore it is one of the most widely used for various regression problems and it can be written as:

$$K(x_i, x_j) = \exp\left(-\gamma\|x_i - x_j\|^2\right) \tag{4}$$

where γ is the kernel parameter. Support vector algorithm for nonlinear case can be written in its dual form:

$$\begin{aligned}
&\max \quad -\frac{1}{2}\sum_{i,j=1}^{l}\left(\alpha_i-\alpha_i^*\right)\left(\alpha_j-\alpha_j^*\right)K\left(x_i,x_j\right)-\varepsilon\sum_{i=1}^{l}\left(\alpha_i+\alpha_i^*\right)+\sum_{i=1}^{l}y_i\left(\alpha_i-\alpha_i^*\right)\\
&\text{subject to} \quad \begin{cases}\sum_{i=1}^{l}\left(\alpha_i-\alpha_i^*\right)=0\\ 0\le\alpha_i,\alpha_i^*\le C\end{cases}
\end{aligned}\tag{5}$$

Only a number of coefficients α_i and α_i^* will be different from zero, and the data points associated to them are called the support vectors. Finally the kernel function allows the decision function of nonlinear SVR to be expressed as follows:

$$f(x)=\sum_{i=1}^{l}\left(\alpha_i-\alpha_i^*\right)K\left(x_i,x_j\right)+b \tag{6}$$

The parameters that define the nonlinear SVR are the cost constant C, the radius of the insensitive tube ε, and the parameter related to kernel (γ for the RBF kernel). These parameters are mutually dependent, so changing the value of one parameter influences other parameters. The parameter C controls the smoothness or flatness of the approximation function. A greater C corresponds to greater penalty of errors and makes the learning machine more complex [25] . Smaller C may cause the errors to be excessively tolerated, which can lead to learning machine with poor approximation. The parameter ε also affects smoothness and dominates the number of support vectors (smaller ε leads to more support vectors and more complex machine). Determining optimal parameters for nonlinear SVR is often a heuristic trial-and-error process, while there are attempts in literature to define optimal range of values depending of the size and noise of the training data [26].

3. Case study

The Norwegian University of Science and Technology (NTNU) campus Gløshaugen is a typical representative of the group of “mixed use” buildings. It consists of 35 objects with very wide purposes and significant energy use, such as classrooms, sport facilities, laboratories, kitchen, etc [27]. The buildings are usually multi-functional, and most of them have laboratories, which might indicate possible high energy use [28]. The total campus area is approximately 300,000 m^2. Building and Energy Management System (BEMS) and web-based Energy Monitoring System (Energy Remote Monitoring – ERM) are available for operation and energy monitoring at NTNU. Hourly heat and electricity consumption from all installed meters and submeters can be collected on ERM. The main meter for the entire campus is installed by the district heating supplier, so the daily heating energy use measured by this meter are taken as relevant for creating the model.

4. SVM model development

The input variables for the developed model are all the available measured meteorological parameters gathered at the local weather station. Additionally categorical values that define day of the week (values 1 to 7) and month of the year (1 to 12) are taken into account as inputs. For the modeling of daily heating energy use (output variable), the following input variables were used: the mean daily outdoor temperature [°C], the mean daily wind speed [m/s], total daily solar radiation [Wh/m^2], the minimum daily temperature [°C], the maximum daily temperature [°C], relative humidity [%], day of the week and the month of the year. Various studies showed that the accuracy can be improved by introducing previous target values (in this case heating energy use of the previous day) as additional input value. Therefore, partial autocorrelation, which shows how a variable (daily heating energy use) is correlated with itself for different lags (in this case, days) is investigated.

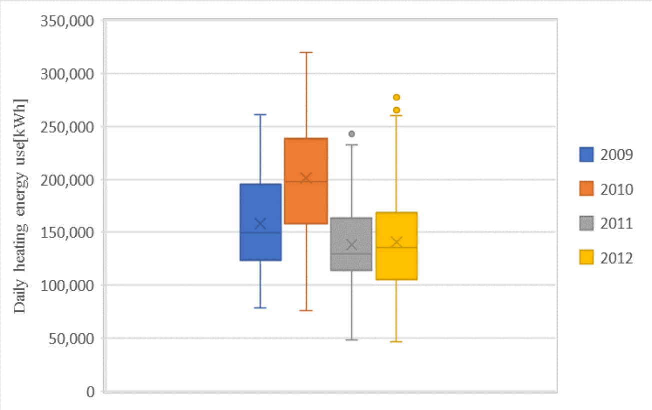

Figure 2. Boxplot of daily heating use

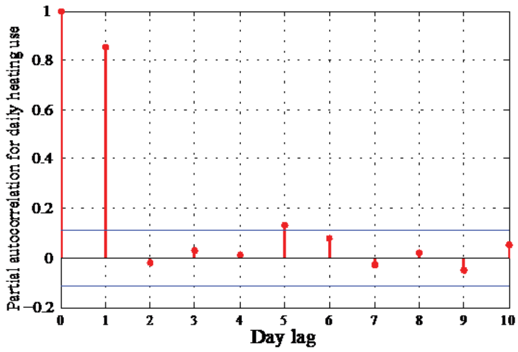

Figure 3. Partial autocorrelation function

Figure 2 represents boxplot of the daily heating energy use for the chosen database. It can be seen that there are some outliers in the year 2011 and 2012. In Figure 3 the results of the partial autocorrelation function for daily district heating use of the campus are presented. It indicates that the daily heating use of the observed day is the most correlated with the use of the previous day (value 0.85). This information implies that another potentially significant input variable is the heating use of the previous day, without necessity to take into account further previous use (heating use for two or more days ago). The first model, SVM1, was developed using available eight variables as input. For the model SVM2 the heating use of the previous data was added as additional input variable in order to investigate potential improvement of the prediction accuracy. Considering that there were obviously different patterns of energy use and behavior occurring during working days and weekends, it was decided to develop heating use prediction models only for the working days. For training the models, data for the working days in the coldest 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), while the year 2012 (100 samples) was used for testing. Data with obvious errors and heat meter malfunctions were removed from the dataset. In order to overcome the numerical difficulties and to avoid that the variables with greater numerical values dominates the ones with smaller range all input and output variables in the training and test set were normalized using min-max normalization, where their values are scaled to range [0,1] using the linear scaling function. The prediction accuracy was measured by the coefficient of determination (R^2), root mean square error (RMSE) and the mean absolute percentage error (MAPE). For building the Support Vector Machine (SVM) model, library for Support Vector- libsvm for Matlab was used [29]. It has been indicated in the recent literature that the centralized feature of the RBF enables it to effectively model the regression process, therefore it was chosen for kernel in this study. The parameters that define the selected nonlinear SVR are the cost constant C (Equation (3)), the radius of the insensitive tube ε (Equation (3) and Figure 1) and the RBF kernel parameter γ (Equation (4)). Training of the SVM model comprises of finding the optimal combination of these parameters. One of the advantages of SVM is the possibility to use the grid search method in order to overcome potential shortcomings of the trial and error method. When using neural networks for prediction, more parameters needs to be known a priori, such as training algorithms, number of layers, number of neurons in hidden layers, etc. The optimal parameters are usually chosen by trial-and-error method. In libsvm library, the “gridregression.py”, which conducts grid search over the combination of the parameters, in defined range, with specified step is also available. It calculates the training error for all combinations of parameters and picks out the combination achieving the smallest error. The similar search can be performed directly in Matlab, by calculating prediction error for all the combinations of parameter. The first step is to define wider range of parameters with bigger step in order to find the optimal „area“ of combinations. After that the defined step can be lowered down, in order to perform „finer“ grid search within the optimal area. To improve the generalization ability, the adopted grid search technique usually uses a cross-validation procedure. In the paper, all combinations of parameters (C, ε, γ) are tried using 5-fold cross validation and the one with the best cross-validation MSE is selected. The optimal combination of parameters was found as (C, ε, γ) = (16, 0.0078, 0.0641).

5. Results

The SVM model using eight input variables (available meteorological parameters, day of the week and month of the year) achieved the MAPE of 3.5498% for training and 5.3084% for testing data. Unlike for the neural networks, in the case of SVM model introducing additional input variable, heating use of the previous day (HCp) did not show the significant increase in the prediction accuracy. MAPE for the testing period for model SVM2 was 5.2888%. In order to be able to adequately compare the achieved results with the previously published work, heating energy use of the previous day was taken into account as additional input variable. Figure 4 and Figure 5 show the comparison between the measured daily district heating use and the prediction of the model SVM2 for the training and test period, respectively.

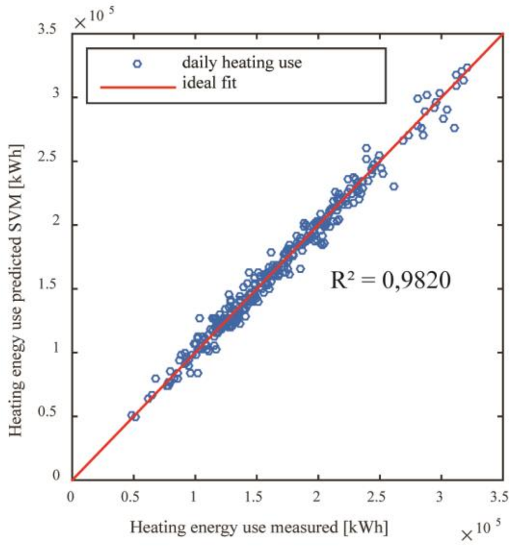

Figure 4. Comparison between measured heating use and prediction of SVM2 model for training period

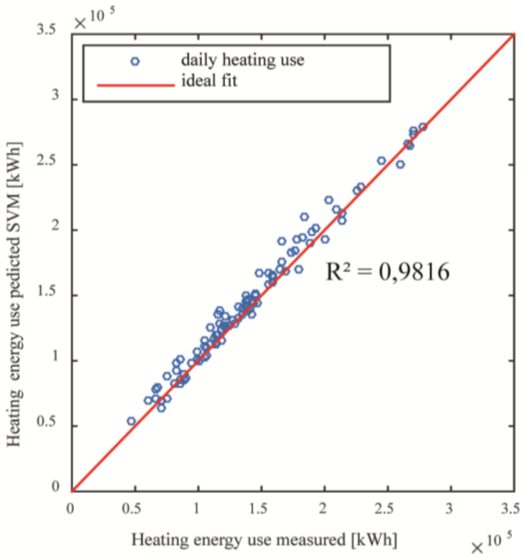

Figure 5. Comparison between measured heating use and prediction of SVM2 model for test period

Figure 6 and Figure 7 show the prediction results for the training and test period respectively. Regardless of the big daily energy use variation, from 40,000 kWh to 320,000 kWh, the model was capable to predict daily energy use with sufficient accuracy. A brief summary of the results and model quality comparison are given in Table 1

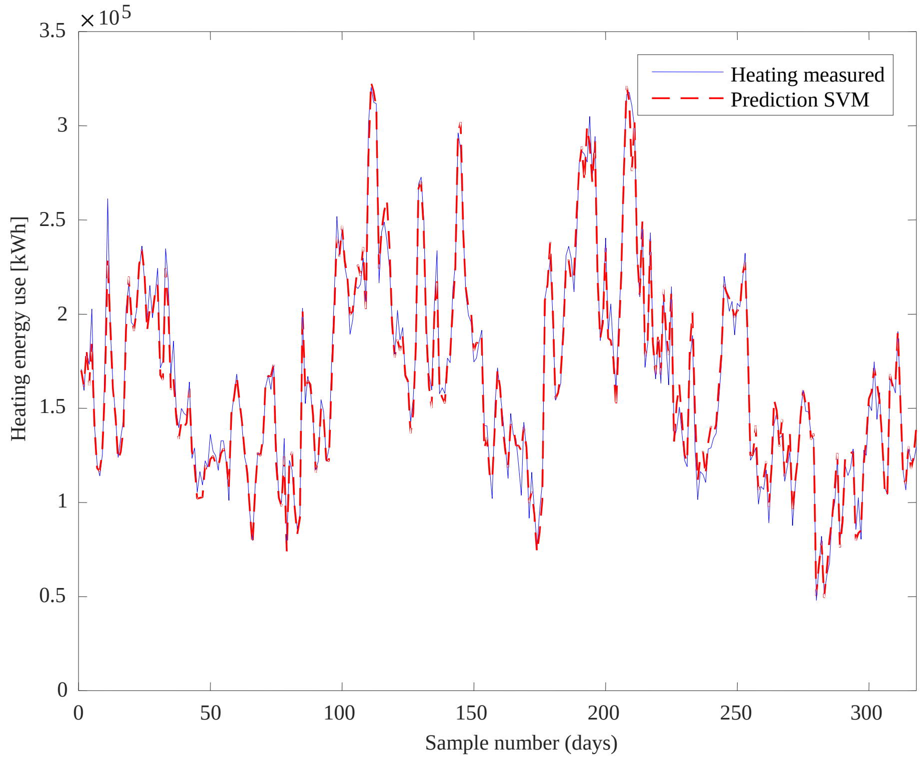

Figure 6. Prediction results of the SVM2 model for the training period

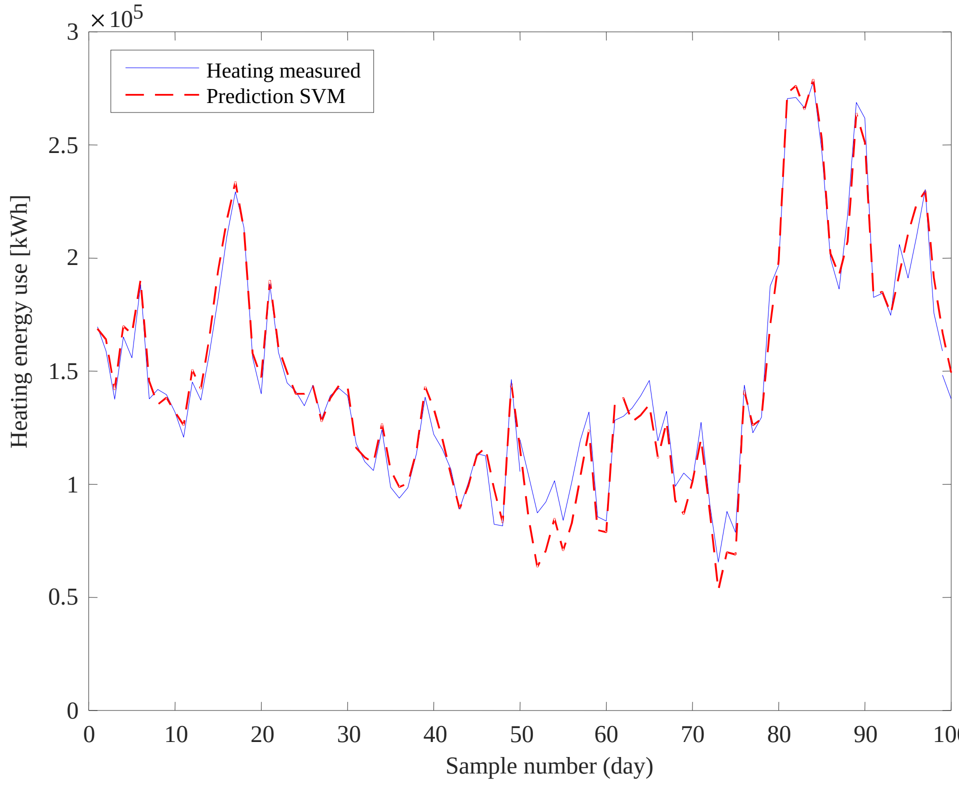

Figure 7. Prediction results of the SVM2 model for the test period

In Table 1, the prediction indices for training and testing of all the compared models are presented. Here, only the brief overview of the previously published models is presented. In paper [11] three different neural networks (FFNN, RBFN and ANFIS) were developed. Possible improvement of accuracy by creating ensemble of neural networks was also investigated in [11]. The results showed that the simple combination of the outputs of the individual networks (simple-SAV, weighted-WAV or median based averaging-MAV) can improve the prediction quality.

Table 1. Prediction indices for various models

model	R^2 [-]		RMSE [kWh]		MAPE [%]	
	training	test	training	test	training	test
SVM1 (8 input variables)	0.9819	0.9816	7.423	8.616	3.5498	5.3084
SVM2 (with HCp)	0.9820	0.9816	7.410	8.602	3.5446	5.2888
FFNN [11]	0.9836	0.9814	7.152	8.496	3.4402	5.6283
RBFN [11]	0.9764	0.9816	8.474	8.849	4.4663	5.6682
ANFIS [11]	0.9826	0.9783	7.200	9.115	3.8737	5.5778
Multistage RBFN ensemble with 8 clusters [10]	0.9877	0.9821	6.128	8.548	3.0870	5.4934
Ensemble MAV with 8 clusters [10]	0.9867	0.9818	6.379	8.582	3.1070	5.4982
MS ANFIS-FCM 5 clusters [12]	0.9880	0.9814	6.029	8.221	2.9960	5.3810
MS ANFIS-gaussmf with 6 clusters [12]	0.9904	0.9814	5.394	8.334	2.7766	5.3887
Ensemble SAV with 2 clusters [13]	-	0.9828	-	8.640	-	5,2691

The next idea was to train 50 different FFNNs and use k-means clustering to select the ensemble members. In order to improve the ensemble efficiency, it was necessary to ensure both accuracy and diversity between individuals. First, clustering had been used to divide networks in groups, and then the most accurate individual network was selected for the ensemble. In paper [12] FFNN and ANFIS networks in the second level were used to create the multistage ensemble. Also, different ANFIS models were constructed: using different membership functions (trimf, gbellmf, gaussmf), fuzzy C-means clustering (FCM) and subtractive clustering. In [10] RBFN was proposed for the second stage. In [13] k-means clustering was used for creating subsets used to train individual RBFN. Due to the resampling of training dataset, the prediction results for the training period are not presented. The second step was to aggregate the outputs of the individual networks separately trained on different training dataset. Table 1 shows the overview of the best results from each of the mentioned studies. All of the proposed algorithms showed improvement compared to the single neural networks.

6. Discussion

In Table 1 it is possible to compare the results of the SVM model with various single NN models, and also with other previously published improvements. Introducing additional input variable (heating energy use of the previous day) – model SVM2, improved the accuracy of the prediction (MAPE is 5.2888% comparing to MAPE for SVM1 which is 5.3084%). The accuracy is not significantly increased, so unless the prediction is done for just one day ahead (when the heating use of the previous day is known and measured), that additional input variable could increase the overall error of model (if the energy use of previous day is also predicted). Both SVM models achieved better accuracy comparing to the various single NNs. Only model SVM1 does take into account previous day, while all other models have nine input variables. Even the SVM without using the heating use of the previous day as additional input had better results than neural networks with nine input variables. Also, both SVM models achieved better prediction accuracy than the developed multistage models. Although all of the proposed algorithms showed improvement compared to the single neural networks, the single SMV model had better results even than the more complex models. It can be seen that the SVM with MAPE for the test period of 5,2888% is the model that can stand side by side with all the other innovative methodologies. In paper [25], the authors introduced the SVM model to predict hourly cooling load in the building, while in [30] is shown that SVM outperforms various NN models. For the prediction of daily maximum temperature [31] SVM model developed using real data measured in meteorological stations obtains accurate prediction, while it performed better than multi-layer perceptron (MLP) and extreme learning machine (ELM). The same conclusion is found for the wind speed prediction [32], where the SVM model showed better results comparing to MLP with different number of hidden neurons. The database consisting of the measured daily heating use was used as case study for several previously published studies. The comparison of the results presented in this paper with those studies shows that SVM outperforms not just various single NNs, but also more complex ensemble models. Achieving MAPE of 5.3084% and R^2=0.9816 in testing period, shows that the SVM model can be used for the prediction of the daily heating energy use with high accuracy. These results are encouraging, in a way that they show that SVM can be used for building energy use estimation, which is clearly topic of great interest nowadays. This kind of models can be used to indicate the meter malfunctioning, if the readings are significantly different from the results of the previously adequately trained model. They can help in identifying higher consumption than normal, so the management can pay more attention and try to find the source. The significant part of the NTNU campus is being rented to other users, which is often for this type of building, so the prediction model can help in accurately calculating users heating bills. In order to adequately carry out energy planning, it is necessary to have tools that can provide clear picture on expected energy use for different building types. In this case study, the SVM models have proven to be very successful technique. For industrial plants, the energy supplier requests that the expected consumption in following period is estimated. These values significantly affect the bills, with high penalties for overpassing the maximum. Therefore, the accurate prediction is not only useful, but often highly required.

7. Conclusion

This study aimed to develop the SVM model for the prediction of the daily heating energy use of the NTNU campus Gløshaugen for the working days in the coldest period. The model was trained and tested on the real, measured data. For development of the SVM1 model eight input variables were used (meteorological parameters, day of the week and month of the year), while for the SVM2 model, daily heating use of the previous day was used as additional input variable. The SVM model results were compared with the previously published results using the same database and same input variables. Both SVM models showed better accuracy than any of the individual neural network models (FFNN, RBFN and ANFIS). The improvement of prediction accuracy by adding the heating use of the previous day as input variable was not so significant like in case of the neural networks. The SVM prediction is comparable, and in most of the cases even better than the innovative methodologies that involve creating conventional and multistage ensembles. This study showed that the SVM is often wrongfully neglected in solving the prediction problems. The optimal combination of the parameters during training of the SVM model can be found using grid search method, which makes the modeling easier. These results may be used to recover the lost data on daily energy use for the purpose of issuing correct energy bills, since the significant part of the campus is leased to other users. It is expected that this theoretically well known algorithm finds wider application, especially in the prediction problems. Future studies should be faced towards possible improvements of the proposed algorithm.

ACKNOWLEDGEMENT

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] Council., E. P. a. 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] Zhao, H.-x. and Magoulès, F. A review on the prediction of building energy consumption. *Renewable and Sustainable Energy Reviews, 16* (2012), 6, pp. 3586-3592. doi:10.1016/j.rser.2012.02.049

[3] Foucquier, A., Robert, S., Suard, F., Stéphan, L. and Jay, A. State of the art in building modelling and energy performances prediction: A review. *Renewable and Sustainable Energy Reviews*, *23* (2013), pp. 272-288. doi:10.1016/j.rser.2013.03.004

[4] Kusiak, A., Li, M. and Zhang, Z. A data-driven approach for steam load prediction in buildings. *Applied Energy*, *87* (2010), 3, pp. 925-933. doi:10.1016/j.apenergy.2009.09.004

[5] Vujić, B., Vukmirović, S., Vujić, G., Jovičić, N., Jovičić, G. and Babić, M. Experimental and artificial neural network approach for forecasting of traffic air pollution in urban areas: the case of subotica. *Thermal Science*, *14* (2010), suppl., pp. 79-87. doi:10.2298/TSCI100507032V

[6] Ćirić, I. T., Ćojbašić, Ž. M., Nikolić, V. D., Živković, P. M. and Tomić, M. A. Air quality estimation by computational intelligence methodologies. *Thermal Science*, *16* (2012), suppl. 2, pp. 493-504. doi:10.2298/TSCI120503186C

[7] Özener, O., Yüksek, L. and Özkan, M. 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. doi:10.2298/TSCI120321220O

[8] Esen, H., et al., Artificial neural networks and adaptive neuro-fuzzy assessments for ground-coupled heat pump system. *Energy and Buildings*, *40* (2008), 6, pp. 1074-1083. doi:10.1016/j.enbuild.2007.10.002

[9] Kumar, R., Aggarwal, R. and Sharma, J. Energy analysis of a building using artificial neural network: A review. *Energy and Buildings*, *65* (2013), pp. 352-358.

[10] Jovanović, R. Ž., Sretenović, A. A. and Živković, B. D. Multistage ensemble of feedforward neural networks for prediction of heating energy consumption. *Thermal Science* (2015) , online first. doi:10.2298/TSCI150122140J

[11] Jovanović, R. Ž., Sretenović, A. A. and Živković, B. D. Ensemble of various neural networks for prediction of heating energy consumption. *Energy and Buildings*, *94* (2015), pp. 189-199. doi:10.1016/j.enbuild.2015.02.052

[12] Jovanovic, R. and Sretenovic, A. Various multistage ensembles for prediction of heating energy consumption, *Modeling, Identification and Control, 36* (2015), 2, pp. 119-132. doi: 10.4173/mic.2015.2.4

[13] Jovanovic, R. and Sretenovic, A. Ensemble of radial basis neural networks with k-means clustering for heating energy consumption prediction. *FME Transactions*, *44* (2016), 3, pp. 217-223.
[14] Ćirić, I. T., Ćojbašić, Ž. M., Nikolić, V. D., Igić, T. S. and Turšnek, B. A. Intelligent optimal control of thermal vision-based Person-Following Robot Platform. *Thermal Science*, *18* (2014), 3, pp. 957-966. doi:10.2298/TSCI1403957C
[15] Ekonomou, L. Greek long-term energy consumption prediction using artificial neural networks. *Energy*, *35* (2010), 2, pp. 512-517. doi:10.1016/j.energy.2009.10.018
[16] Vapnik, V., The nature of statistical learning theory. 2013: Springer Science & Business Media.
[17] Naradasu, K. R., Jyothirmai, S. and Ramesh, R. Towards artificial intelligence based diesel engine performance control under varying operating conditions using support vector regression. *Thermal Science*, *17* (2013), 1, pp. 167-178. doi:10.2298/TSCI120413218N
[18] Dong, B., Cao C., and Lee S.E., Applying support vector machines to predict building energy consumption in tropical region. *Energy and Buildings*, *37*, (2005). pp. 545-553. doi:10.1016/j.enbuild.2004.09.009
[19] Esen, H., Esen, M., and Ozsolak O., Modelling and experimental performance analysis of solar-assisted ground source heat pump system. *Journal of Experimental & Theoretical Artificial Intelligence 29*, (2017), pp. 1-17. doi:10.1080/0952813X.2015.1056242
[20] Esen, H., et al., Modeling a ground-coupled heat pump system by a support vector machine. *Renewable Energy*, *33* (2008), 8, pp. 1814-1823. http://dx.doi.org/10.1016/j.renene.2007.09.025
[21] Esen, H., et al., Modelling of a new solar air heater through least-squares support vector machines, *Expert Systems with Applications*, *36* (2009), 7, pp. 10673-10682. http://dx.doi.org/10.1016/j.eswa.2009.02.045
[22] Wang, L., *Support Vector Machines: theory and applications*, Springer Science & Business Media, 2005
[23] Gunn, S. R. Support vector machines for classification and regression. *ISIS technical report*, *14* (1998).
[24] Smola, A. J. and Schölkopf, B. A tutorial on support vector regression. *Statistics and computing*, *14* (2004), 3, pp. 199-222. doi: 10.1023/B:STCO.0000035301.49549.88
[25] Li, Q., Meng, Q., Cai, J., Yoshino, H. and Mochida, A. Applying support vector machine to predict hourly cooling load in the building. *Applied Energy*, *86* (2009), 10, pp. 2249-2256. doi:10.1016/j.apenergy.2008.11.035
[26] Cherkassky, V. and Ma, Y. Practical selection of SVM parameters and noise estimation for SVM regression. *Neural networks*, *17* (2004), 1, pp. 113-126. doi:10.1016/j.apenergy.2008.11.035
[27] Sretenovic, A. *Analysis of energy use at university campus*. Master, NTNU, Trondheim, 2013.
[28] Guan J., Nord N., and Chen S., Energy planning of university campus building complex: Energy usage and coincidental analysis of individual buildings with a case study. *Energy and Buildings*, *124*, (2016), pp. 99-111. doi:10.1016/j.enbuild.2016.04.051
[29] Chang, C.-C. and Lin, C.-J. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *2* (2011), 3, pp. 27. doi:10.1145/1961189.1961199
[30] Li, Q., et al., Predicting hourly cooling load in the building: a comparison of support vector machine and different artificial neural networks, *Energy Conversion and Management*, *50* (2009), 1, pp. 90-96. doi:10.1016/j.enconman.2008.08.033
[31] Paniagua-Tineo, A., et al., Prediction of daily maximum temperature using a support vector regression algorith,. *Renewable energy*, *36*, (2011), 11, pp. 3054-3060. doi:10.1016/j.renene.2011.03.030
[32] Mohandes, M.A., et al., Support vector machines for wind speed prediction, *Renewable energy*, *29*, (2004), 6, pp. 939-947. doi:10.1016/j.renene.2003.11.009