Application of Support Vector Machines On Medium Term Load Forecasting

Miloš B. Stojanović[1], Miloš M. Božić[2] and Milena M. Stanković[3]

Abstract – **Medium term load forecasting using Support Vector Machines (SVM) is presented in this paper. The forecasting is performed for the electrical maximum daily load of the period of one month. The data considered for forecasting consist of quarter hours daily load for the City of Niš for one year, daily average temperature and daily average wind speed. An analysis of available data was performed and most important attributes for the formation of the SVM model are selected. It was shown that the size and structure of the training set may significantly affect the accuracy of load forecasting. The results obtained from training set with and without involved average wind speed data are presented and the results shows the approach without wind speed data is more accurate.**

Keywords – **Load forecasting, Support vector machines, Regression, Time series.**

I. Introduction

Load forecasting has always been important for economic and reliable operation of an electric utility. However, with the deregulation of the energy industries, load forecasting is even more important. With supply and demand fluctuating and the changes of weather conditions and energy prices increasing more during peak situations, load forecasting is vitally important for utilities. Load forecasting helps an electric utility to make important decisions including decisions on purchasing and generating electric power, load switching, and infrastructure development. Load forecasts can be divided into three categories based on prediction time period: long term, medium term and short term. Long term forecasts are related to time period which are longed then a year, medium forecasts are usually from a week to a year and short term forecasts which are usually from one hour to one week.

Load forecasting methods mostly use statistical techniques or artificial intelligence algorithms such as regression, neural networks, fuzzy logic, and expert systems. The use of Artificial Neural Networks (ANN) [7] has been a widely studied electric load forecasting technique. More recent powerful technique for use in electric load forecasting is SVM [4]. Unlike ANN, which try to define complex functions of the input feature space, SVM perform a nonlinear mapping (by using so-called kernel functions) of the data into a high dimensional space. SVM then use simple linear functions to create linear decision boundaries in the new space. The problem of choosing an architecture for a ANN is replaced in SVM by the problem of choosing a suitable kernel.

Medium term load forecasting using SVM are presented in this paper. The paper is organized as follows. In Section II, SVM techniques are presented. Data analysis is given in Section III. Section IV explains and describes our model in detail. Section V presents experimental results and provides model evaluation, and the conclusions are given in Section VI.

[1]Miloš B. Stojanović is with the Faculty of Electronics Engineering, A. Medvedeva 14, 18000 Niš, Serbia, E-mail: milos.stojanovic@vtsnis.edu.rs

[2]Miloš M. Božić is with the Faculty of Electronics Engineering, A. Medvedeva 14, 18000 Niš, Serbia, E-mail: milos1bozic@yahoo.com

[3]Milena M. Stanković is with the Faculty of Electronics Engineering, A. Medvedeva 14, 18000 Niš, Serbia, E-mail: milena.stankovic@elfak.ni.ac.rs

II. SVM

SVM are developed by Vapnik and assistants in 1995. to resolve the issue of data classification. Two years later, the version of SVM is proposed that can be successfully applied to the data regression problem. This method is called Support Vector Regression (SVR) and it is the most common form of SVM that is applied in practice [1]. SVM are based on the principle of structural risk minimization (SRM), which is proved to be more efficient than the empirical risk minimization (ERM), which is used in neural networks. SRM minimizes an upper bound of expected risk as opposed to ERM that minimizes the error on the training data [2].

A brief review of SVR method is given, which is successfully used to predict the data that are organized as time series. The goal of SVR is to generate a model (function) which will perform prediction of unknown output values based on the known input parameters. In the training phase the formation (training) of the model is performed based on the known training data (x_1, y_1), (x_2, y_2), …, (x_n, y_n), where x_i are input vectors, and y_i outputs associated to them. Each input vector consists of numeric attributes. In the test phase the trained model on the basis of new inputs $x_1, x_2, ..., x_n$ makes prediction of output values $y_1, y_2, ..., y_n$. SVR is an optimization problem in which is needed to determine the parameters ω and b to minimize [3]:

$$\min_{\omega,b,\xi,\xi^*} \frac{1}{2}\omega^T\omega + C\sum_{i=1}^{n}(\xi_i + \xi_i^*), \tag{1}$$

with conditions:

$$y_i - (\omega^T - \phi(x_i) + b) \le \varepsilon + \xi_i, \tag{2}$$

$$(\omega^T\phi(x_i) + b) - y_i \le \varepsilon + \xi_i^*,$$

$$\xi_i, \xi_i^* \ge 0, i = 1, 2, \dots, n,$$

where x_i is mapped in the multi-dimensional vector space with mapping ϕ, ξ_i is the upper limit of training error and

ξ_i^* the lower. Mapping ϕ is also known as kernel function. The parameters that control the quality of the regression are mapping ϕ, and the constants C and ε. C is a constant which determines the "cost of error", i.e. determines the tradeoff between the model complexity and the degree to which deviations larger than ε are tolerated. Parameter ε controls the width of ε insensitive zone, used to fit the training data [9]. The goal of SVR is to place as many input data (vectors x_i) inside the tube $|y-(\omega^T\phi(x)+b)| \leq \varepsilon$, which is shown on Fig 1. If x_i is not inside the tube, an error occurs ξ_i or ξ_i^*.

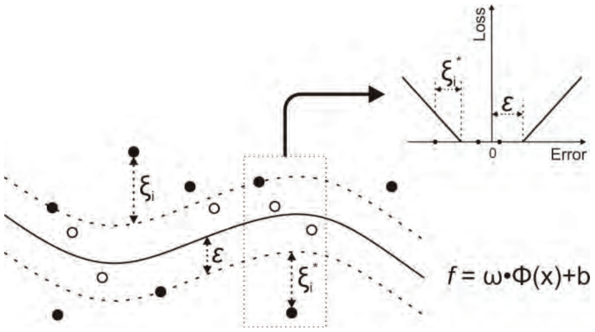

Fig. 1. ε tube of nonlinear SVR

In the least-square regression ε is always zero and the data are not mapped in a multidimensional space. Therefore, SVR provides greater flexibility and can be applied to the various sets of problems [4].

For the experiments in this paper we used a publicly available library LibSVM - A Library for Support Vector Machines [5] in C# version.

III. Data Analysis

Before we proposed the solution, some observations about properties of load demand are examined first. Some relations between load demand and other information like time and climate influence are identified.

Behavior of load demand: Load demand data given are quarter hours recorded from 2008. to 2009. for City of Niš. On Fig 2. maximum daily loads from 2008. to 2009. are presented.

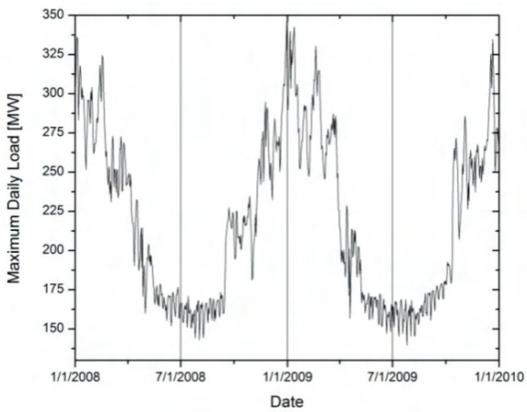

Fig. 2. Load pattern during period of two years

Relations between electricity usage and weather conditions in different seasons are observed from data on Fig 2. Load demand in winter period is higher compared to summer period. Additionally, can be observed that the seasonal load pattern is almost the same for both years. Also, load pattern periodicity exists in every week. Usage of electricity is relatively constant during week days and drops in the weekend.

Climate influence: Relation between climate and load demand also was shown in previous works on short term load forecasting [4,7]. Climate conditions influence on load demand pattern analysis may include temperature, wind speed, humidity, pressure and illumination. In this paper two climate parameters are considered, temperature and wind speed.

From Fig 3. can be easily observed negative correlation between maximum daily load and temperature. The correlation coefficient between maximum daily load and temperature is -0.942, which tells us that there is very strong connection between these two variables. Reason for that is heating use in winter period of year, higher temperature causes lower load demands.

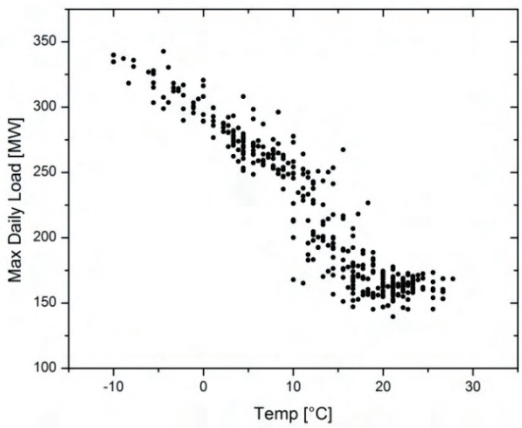

Fig. 3. Correlation between the max daily load and the temperature

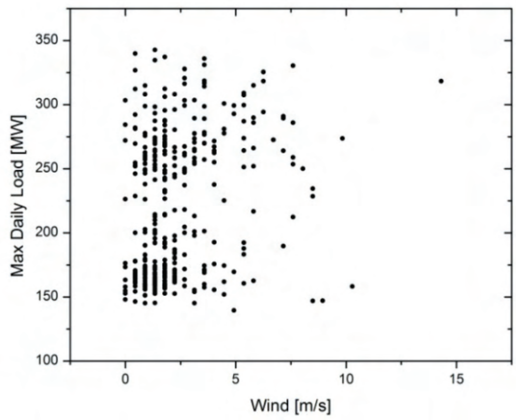

Fig. 4. Correlation between max daily load and the wind speed

Another climate parameter which may have significant influence on load demand is wind speed. Correlation between wind speed and maximum daily load is presented on Fig. 4. and this correlation coefficient is 0.263, which implies weak connection between these variables, although it was not initially expected. Because of that, two different models are built, one with wind speed included as feature and one without wind speed.

Holiday influence: Holidays and other local events also may affect the load demand. These events are usually local and their influences highly depend on the customs of the area. Major holidays such as Christmas or New Year have more influences on load demand then other holidays.

IV. Data Preparation and Model Formation

Here, we considered what kind of information should be included to prepare datasets to build SVM models. Each component of input vectors is called a feature (attribute), and it is very important to identify real features.

Calendar attributes. In Section III, weekly pattern of load demand is discussed. We pointed out earlier that, the load demand on week days is relatively constant and higher than in the weekend. Also, the load demand on holidays is lower than on non-holidays. This information (weekdays and holidays) is useful for encoding in training entries. Actually, many works [4,7,8] have used the calendar information to model the problem.

Temperature. The temperature data is another possible feature, since load demand and temperature have causal relation in between. Weather information which includes temperature, wind speed, sky cover, humidity and etc., has also been used in most load forecasting works. But there is one difficulty: for mid-term load forecasting, temperature data for several weeks away are needed.

Wind speed. Beside temperature, wind speed is another climate parameter which has important influence on load demand. As with temperature date, if we want to encode the wind speed in our training entries, we will also need forecasted wind speed data for several weeks away.

Time series style. Another information which we considered to encode as the feature is the past load demand. With this approach, concept of time-series is introduced into our models [8]. If y_i is the target value for prediction, the vector x_i includes several previous target values $y_{i-1},\ldots,y_{i-\Delta}$ as attributes. In the training phase all y_i are known but for future prediction, $y_{i-1},\ldots,y_{i-\Delta}$ can be values from previous predictions. This approach is used because we know, the past load demand could affect and imply the future load demand. In this paper, we used Δ=7, which means that seven past daily load demands are used for attributes.

The various factors that affect the load forecast are analyzed and appropriate features are chosen, from which we form vectors which will be used as inputs for SVM. Our proposed model is shown on Fig 5. Input vectors for SVM model is composed of the following features:

- Maximum daily load for past seven days (L_{i-k}), k=1,…,7,
- Average daily temperatures (T_i),
- Average daily wind speed (W_i),
- Day Of the Week (DOW).

On the basis of established vectors and the known values of the load in the selected time interval SVM generates model which forecasts the load for a period of one month (training of the model). Trained model is forecasting the load for a period of one month (the application of the model).

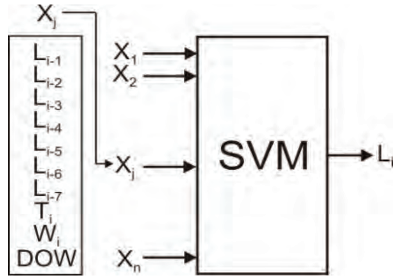

Fig. 5. Proposed architecture

In addition to the selected attributes, in order to get "good" model, we should determine the following parameters:

1. Mapping ϕ (kernel),
2. Constant C (“cost of error“),
3. Constant ε (width of tube).

For mapping ϕ RBF kernel is selected (Radial Basis Function) which is the most widespread in practical use:

$$\phi(x_i)^T\phi(x_i) = e^{-\gamma\|x_i-y_i\|^2}, \gamma > 0.$$

For constant ε it was chosen the value of 0.1. It was found that the choice of ε (tested values were in range of 0.01 to 0.5 with the step of 0.01) does not affect significantly on the accuracy of load forecasts.

It is necessary to determine the parameters C and γ. It is not known in advance which values of these parameters are best choice for a given problem. The parameters were determined using Cross - Validation procedure and Grid - Search [6].

The training set is randomly divided into training and testing parts, for example in relation 1:9. Then the learning algorithm is applied to training part and evaluation of the quality of prediction is performed on testing part. This procedure is repeated n times e.g. 10 and we selected a pair C and γ with which the best accuracy is achieved. Couples C and γ should be search exponentially, for example, $C = 2^{-5}$, 2^{-3}, ... , 2^{20} , $\gamma = 2^{-15}$, 2^{-13}, ... , 2^{3} [6]. Obtained large values for C in range 2^{18} - 2^{22} generates models that very well performs data fitting, i.e. tolerates very few errors. Values for γ in range 2^{-5} - 2^{-7} are near to default values proposed by LibSVM (γ=1/n, where n is number of attributes in vector).

V. Experimental Results

In electricity load forecasting, the prediction accuracy is generally evaluated using Mean Absolute Percentage Error (MAPE) [10]. The equation describing this error is:

$$MAPE = 100 \cdot \frac{1}{n}\sum_{i=1}^{n}\left|\frac{P_i - \hat{P}_i}{P_i}\right|, \tag{3}$$

where P_i and $\hat{P}_i$ are the real and the predicted value of maximum daily electrical load on the i^{th} day and n is the number of days in the month.

To evaluate the accuracy of the model, maximum daily load forecasting for January 2010 is done. Training model was committed with several different data encodings and segmentations.

TABLE I
MAPE USING DIFFERENT DATA PREPARATION

Segments	With wind speed	Without wind speed
All	4.25%	3.84%
Winter	4.42%	3.56%
Dec-Jan-Feb	5.49%	3.44%
Summer	15.68%	15.28%

Table I shows MAPE errors generated by different data encodings and segmentations. The first column represents the data segments used, and then the next two columns show the prediction with and without the wind speed feature. In Table I, it can be observed that model built without wind speed feature using the "Dec-Jan-Feb" data, exceeds all others. Trained models with segments formed from "Winter" and "Dec-Jan-Feb" data generally have smaller errors compared to training models with "All" and "Summer" data segments.

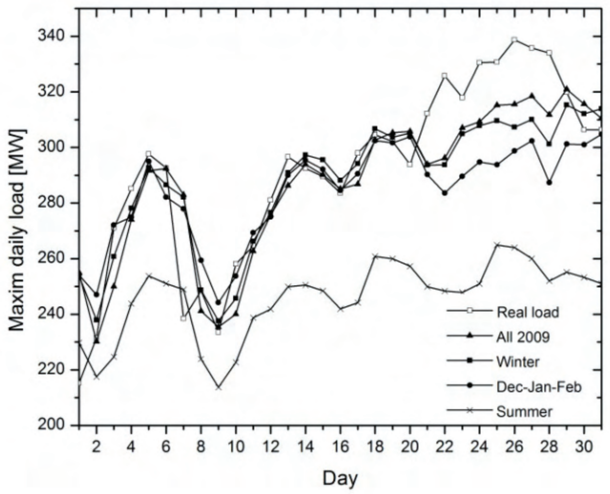

Fig. 6. Estimates and real load demand in Jan. 2010, with all features

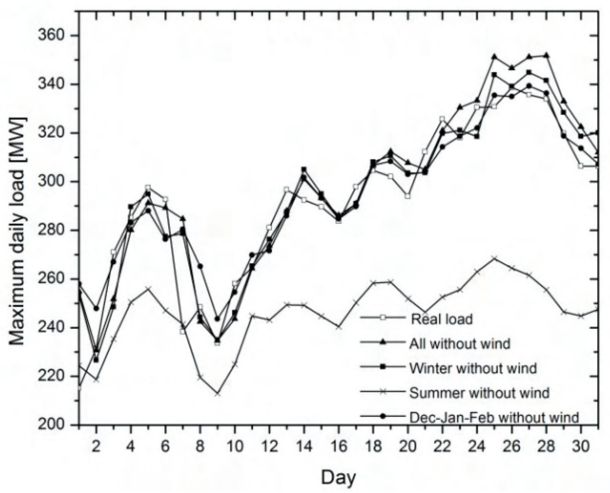

Fig. 7. Estimates and real load demand in Jan. 2010, without wind speed feature

In Fig 6. and Fig 7., the estimates and real load demand in January 2010 are shown, with all features and without wind speed feature respectively. An interesting observation can be made from the results: models built without the wind speed feature generally perform better than those built with. Also, the most deviations from real load are in days 1 and 7. Reason for that is these days are holidays, Ney Year and Christmas. As we mentioned earlier, holiday events also affect the load demand and firstly holiday information have been involved in our model. The problem is that the historical load data on holidays are not enough to train the model well, and this makes the behavior of load demand on these days hard to predict. However, later we remove holiday information from our model.

VI. CONCLUSION

In this paper several SVM models are made with different data segments. We find that choosing appropriate data segments may improve performance of models, since the load demand appears to have different distribution in different seasons. Furthermore, models build with climate information (temperature and wind speed) require future climate data for several weeks away. This difficult may lead to inaccurate prediction and because of that models without wind speed feature were build. Rejection wind speed feature from models was significantly increase models performance. However, the use of causal relation between temperature and load demand might need further consideration.

REFERENCES

[1] D. Basak, S. Pal, D.C. Patranabis, "Support Vector Regression", Neural Information Processing – Letters and Reviews, Vol. 11, No. 10, October 2007.
[2] S. R. Gunn, "Support Vector Machines for Classification and Regression", Technical Report, Image Speech and Intelligent Systems Research Group, University of Southampton, 1997.
[3] D. Basak, S. Pal, D. C.Patranabis, "Support Vector Regression", Neural Information Processing – Letters and Reviews, Vol. 11, No. 10, October 2007.
[4] A. Jain, B. Satish, "Clustering based Short Term Load Forecasting using Support Vector Machines", Bucharest Power Tech Conference June 28th - July 2nd 2009.
[5] C. C. Chang, C. J. Lin, "LibSVM: a library for support vector machines", Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm/.
[6] C. W. Hsu, C. C. Chang, C. J. Lin, "A Practical Guide to Support Vector Classification", Department of Computer Science National Taiwan University, Taipei 106, Taiwan.
[7] D. Ortiz-Arroyo, M. Skov, Q. Huynh, "Accurate Electricity Load Forecasting with Artificial Neural Networks", Proceedings of the 2005 International Conference on Computational Intelligence for Modeling, Control and Automation.
[8] N. Amjady, "Short-term hourly load forecasting using time-series modeling with peak load estimation capability", IEEE Transactions on Power Systems 16(3):498–505, August 2001.
[9] V.Cherkassky, Y.Ma, "Selection of meta-parameters for Support Vector Regression", International Conference Madrid, Spain, August 28-30, 2002 Proceedings.
[10] P. Bunnoon, K. Chalermyanont, C. Limsakul, "A Computing Model of Artificial Intelligent Approaches to Mid-Term Load Forecasting: a state-of-the-art survey for the researcher", International Journal of Engineering Technology, Vol. 2, No. 1, February 2010.