International Journal of Research in Science & Technology
Volume 2 | Issue 4 | April 2015 | ISSN: 2349-0845

Support Vector Machine for Wind Speed Prediction

Mrs. Sangita B. Patil[1] and Mr. Bapu K. Patil[2]

[1]Electrical Engineering Department, GHRIET, Wagholi, Pune, Maharashtra, India
[2]Maintenance Department, Maxion Wheels India Ltd., Pune, Maharashtra, India

Article Info

Article history:

Received on 10th April 2015
Accepted on 14th April 2015.
Published on 16th April 2015

Keyword:

Artificial Neural Network,
Backpropagation Algorithm,
Support Vector Machine (SVM),
Forecasting,
Kernels

ABSTRACT

The energy is a vital input for the social and economic development of any nation. With increasing agricultural and industrial activities in the country, the demand for energy is also increasing. The increasing use of natural and renewable energy sources is needed to take the burden of our current dependency on fossil fuels. Development and analysis of renewable energy models helps utility in energy forecasting, planning, research and policy making. The wind power is a clean, inexhaustible, and almost a free source of energy. But the integration of wind parks with the power grid has resulted in many challenges for the utility in terms of commitment and control of power plants. As wind speed and wind direction fluctuate frequently, the accurate long-term and short-term forecasting of wind speed is important for ascertaining the wind power generation availability. To deal with wind speed forecasting, many methods have been developed such as physical method, which use lots of physical considerations to reach the best forecasting precision and other is the statistical method, which specializes in finding the relationship of the measured power data. Wind speed can be predicted by using time series analysis, artificial neural network, Kalman Filter method, linear prediction method, spatial correlation models and wavelet, also by using the support vector machines. In this paper, the SVM is used for day ahead prediction of wind speed using historical data of wind park. In this paper Support Vector Machine (SVM) results are compared with feedforward Backpropagation neural network. It is observed that the Mean Absolute Percentage Error (MAPE) by SVM method is around 7% and correlation coefficient is close to 1. This justifies the ability of SVM for wind speed prediction task than Backpropagation algorithm.

Corresponding Author:

Mrs. Sangita B. Patil
Electrical Engineering Department
G. H. Raisoni Institute of Engineering, Wagholi
Pune, India
sangita.patil@raisoni.net

I. INTRODUCTION

Highlight As the energy is a vital input for the social and economic development of any nation, the increasing use of natural and renewable energy sources is needed to take the burden of our current dependency on fossil fuels. The significantly high level of fossil fuel products burnt each and every day is polluting of the air and surrounding environments and may also be contributing to climate change. The proper allocation of widely available renewable energy sources as solar, wind, bio-energy and hydropower in meeting future energy demand has become today's need. Development and analysis of renewable energy models helps utility in energy forecasting, planning, research and policy making. Wind speed prediction is needed as wind energy is discontinuous and it is also one of the main sources of alternative energy to the exhaustible non-renewable resources.

Many of the studies proposed in the literature are based on artificial intelligence (AI) techniques such as Artificial Neural Network (ANN), Fuzzy Logic (FL), Expert System (ES), etc. ANN has attracted a great deal of attention among these techniques because of their ability to handle noisy data and their learning capabilities. Support Vector Machine (SVM) which is quite a new method and used in this work which provide efficient and powerful classification and regression algorithms that are capable of dealing with high-dimensional input features and with theoretical bounds on the generalization error and sparseness of the solution provided by statistical learning theory.

The wind speed is uncertain and random hence its prior estimation is of great importance in meeting future energy needs. Support vector regression model is developed using historical data of three years to estimate hourly wind speed. In wind speed prediction, the actual values of wind speed and estimated values are compared and it is seen that Support Vector Regression (SVR) can predict wind speed with acceptable accuracy as compared to Feedforward Backpropagation Neural Network.

II. NEURAL NETWORKS

An A feed-forward neural network has layers of processing elements, which makes independent computations on data that it receives and passes the results to another layer and finally, a subgroup of one or more processing elements determine the output from the network. Each processing element makes its computation based upon a weighted sum of its inputs. The first layer is always the input layer and the last layer is always the output layer. The layers placed between the first and the last layers are the hidden layers [1]. The processing elements are seen as units that are similar to the neurons in a human brain, and hence, they are referred to as cells or artificial neurons.

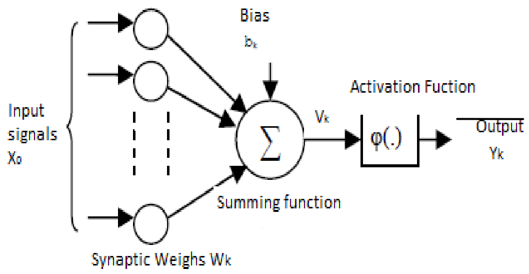

Figure 1. Model of an Artificial Neural Network

The output of any neuron is the result of thresholding, if any, of its internal activation, which, in turn, is the weighted sum of the neuron's inputs. The Activation function is denoted by φ (v), defines the output of a neuron in terms of the induced local field v. Here, Sigmoid function is used as the activation function. It is defined by the equation,

$$\varphi(v) = 1/(1 + \exp(-av)) \qquad (1)$$

Where 'a' is the slope parameter of the sigmoid function. For the prediction of wind speed, historical data (i.e., wind speed, wind gust, temperature, humidity, wind direction and barometric pressure) of October month for year 2006, 2007 and 2008 is collected. For wind speed prediction, historical data (i.e., wind speed, wind gust, temperature, humidity, wind direction and barometric pressure) of October month for year 2006, 2007 and 2008 is used. For predicting average wind speed of a hour of today, average wind speed of an hour of previous two days and average temperature, humidity, wind gust, wind direction and barometric pressure of same hour of previous one day are considered as inputs to the Feedforward Backpropagation neural network. The output is the predicted wind speed of day'd'. Thus the network is provided with seven inputs and Feedforward Backpropagation neural network will return one predicted value of wind speed of the next day. Samples of 22 days of October 2006, 2007 & 2008 are used for training and validation. Out of total 1584 samples 70% samples are used for training and 30% samples are used for testing. Samples of 23rd October, 2008 to 30th October, 2008 are used for evaluating the performance of the Feedforward Backpropagation neural network by comparing predicted wind speed values with actual wind speed values. ANN toolbox in MATLAB environment is used and a program is written for training the Feedforward Backpropagation neural network.

III. SUPPORT VECTOR MACHINE

SVM stands for Support Vector Machines. It is an elegant and highly principled learning method for a feed forward network design with a single hidden layer of units that are non linear. Its deviation follows the principle of structural risk minimization which is based on the fact that the error rate of the learning machine on test data is bounded by sum of training error rate. As the name suggests, the design of the machine hinges on the extraction of a subset of the training data that serves as support vectors and hence, represents a stable data characteristic [9]. There is a common factor which is to use a technique known as "kernel trick" which is to enable operations to be performed in the input space rather than the potentially high dimensional feature space [9]. SVMs are based on the concept of decision planes that define decision boundaries. A decision plane is a boundary between a set of objects having different class memberships. SVM includes the polynomial learning machine, radial-basis function network and two layer perceptron as special instances. These methods provide different representations of intrinsic statistical regularities contained in the training set.

A. Architecture of SVM

The original feature space can always be mapped to some higher-dimensional feature space where the training set is separable using Regression. The basic idea behind support vector machines (SVM) for regression is to map the data x into a high dimensional feature space through a nonlinear mapping. Once mapping is done then SVMs perform a linear regression in this feature space. The architecture of Support Vector Machine is shown in Figure 2.

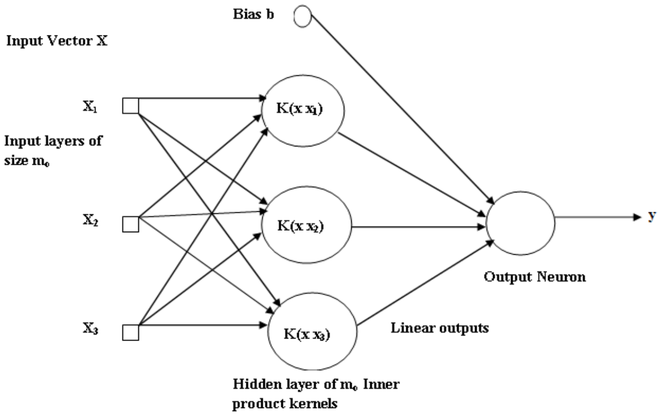

Figure 2. Architecture of Support Vector Machine

B. Support Vector Regression

Support Vector Machines can also be used for function approximation tasks, e.g., to discover the functional relationship between a dependent variable y and independent variables x in presence of noise, expressed as $y = f(x) + N$. Where N denotes a random noise component whose distributional properties are unknown. The unknown functional form f(x) must be estimated from the training set $\{(x_i, y_i): i=1,\ldots\ldots,n\}$ by minimizing the estimated risk or cost function (e.g., the mean squared error) [5].

The advantages of the SVM method are as follows:

- Highly accurate compared to other methods of forecasting.
- Ability to model complex nonlinear decision boundaries
- Less prone to over fitting than other forecasting models
- Provides a compact description of the learned model
- Can be used for prediction as well as classification.

Support vector machine method is established on the basis of the Vapnik-Chervonenks dimension of the statistic studying theory and the principle of structure risk minimization, and better solves the actual problems of little samples, nonlinear, high dimension, local minimum etc. At present it has been applied to models recognition, text classification and electrical power load forecasting [12]. The paper uses SVM algorithm to estimate the regression function and thus to forecast the hourly wind speed. The thought of the algorithm is as follows: Suppose the samples $\{(x_i, y_i)\}(i = 1,2,\ldots,m)$,where m is the sample amount, x_i the input vector y_i the target output data is given. Giving the samples have a nonlinear relation in most conditions, the estimated function *f* can be defined in the following way: use a nonlinear function φ to put a sample point to be reflected to the high dimension space, and then make a linear regression in the high dimension characteristic space, and finally get the nonlinear regression effect in the original space.

$$f(x,w) = w\varphi(x) + b = (w,\varphi(x)) + b \qquad (2)$$

where *w* is the weighting vector, *b* the bias, $(w,\varphi(x))$ the inner product. In order to train *w* and *b*, the following functional need to be minimized.

$$R(c,\varepsilon) = \frac{1}{m} c \sum_{i=1}^{m} L_\varepsilon(y_i - f(x_i), x_i) + \frac{1}{2} w.w \qquad (3)$$

where ε is used to weight risk through ε-insensitivity loss function, and the first item is the empirical error item, the second the normalization item, *c* the normalization constant which determines some balance relation between the empirical error item and the normalization item. In order to solve the function (3), the kernel function $k(x,x_i)$ is introduced in (4).

$$f(x,\alpha_i,\alpha_i^*) = \sum_{i=1}^{m} (\alpha_i - \alpha_i^*) k(x,x_i) + b \qquad (4)$$

where α_i, α_i^* are the Langrange multipliers. The problem described in (4) is a convex programming problem each solution of which is the globally optimal solution, so these are no existence of local extremum.

IV. Prediction Of Wind Speed Using Support Vector Machine

Algorithm:

Collection of data: For prediction of wind speed, historical data (i.e., wind speed, wind gust, temperature, humidity, wind direction and barometric pressure) of October month for year 2006, 2007 and 2008 is collected.

Selection of input & output: For wind speed prediction, historical data (i.e., wind speed, wind gust, temperature, humidity, wind direction and barometric pressure) of October month for year 2006, 2007 and 2008 is used. For predicting average wind speed of a hour of today, average wind speed of an hour of previous two days and average temperature, humidity, wind gust, wind direction and barometric pressure of same hour of previous one day are considered as inputs to the support vector machine. The parameters that are considered as inputs to the model are listed in Table I. The output is the predicted wind speed of day ‘d’. Thus the network is provided with seven attributes and support vector machine will return one predicted value of wind speed of the next day.

TABLE I. List Of Input Parameters To The Network

PARAMETERS	UNITS
Temperature	Deg.C
Humidity	%RH
Wind Gust	m/s
Wind Direction	Deg.M
Barometric Pressure	Mb
Wind Speed of (d-1)	m/s
Wind Speed of (d-2)	m/s

Training & validation data: Samples of 22 days of October 2006, 2007 & 2008 are used for training and testing. Out of total 1584 samples 70% samples are used for training and 30% samples are used for testing. Samples of 23rd October, 2008 to 30th October, 2008 are used for evaluating the performance of the network by comparing predicted wind speed values with actual wind speed values.

LIBSVM tool [15] is used in MATLAB environment and a program is written for training the SVR model. Training of SVM regression model require selecting the cost function C and Kernel functions parameters, which influence the ensuring model performance. In this simulation, RBF kernel is considered as it is advantageous in complex non-separable data [16].

The network is trained, starting with some kernels and some values of parameters C and γ. Observing the performance of minimization of error; one can try another kernels and C and γ values.

If the performance of the SVM model is not satisfactory then it is retrained with different values of C and γ till mean squared error is within tolerance [6].

The optimum value of C and γ is selected through iterative process. The values of C and γ are finalized for which mean square error is minimum. The time required for getting best parameters was 30 minutes. The mean squared error during training was 2.52e-005. The best C=200 and γ=8. The numbers of support vectors for this trained model are 38.

V. Evaluation Of Performance Of Support Vector Machine For Regression

Once the SVM model is trained then samples which were neither used during training and testing are used for the evaluating the performance of the SVM model. Hourly wind speed values are predicted for seven days of October 2008 and then compared with actual wind speed values of these samples.

The performance of SVM model is evaluated in terms of Absolute percentage error (APE) and Mean Absolute Percentage Error (MAPE).

Let W_a be the actual wind speed in m/s and W_f be the forecasted wind speed in m/s. Then Absolute percentage error is defined as

$$APE=\left|\frac{W_a-W_f}{W_a}\right|\times 100 \qquad (5)$$

and MAPE is given by

$$MAPE=\frac{1}{N}\sum_{i=1}^{n} APE \qquad (6)$$

Where N = time block

The wind speed values are predicted for eight days of October month of 2008 and compared with actual wind speeds. The performance for these days is evaluated as discussed by calculating APE and MAPE.

Figure 3 shows comparison of actual and predicted values of wind speed for 23rd October 2008. It is observed that the minimum absolute percentage error obtained by SVM is 1.206% and maximum APE is 17.65%. The mean absolute percentage error is found as 7.19%. The results obtained by SVM are compared with the results of Feedforword Backpropagation Neural Network. Promising results have been obtained using SVM compared to Feedforword Backpropagation Neural Network.

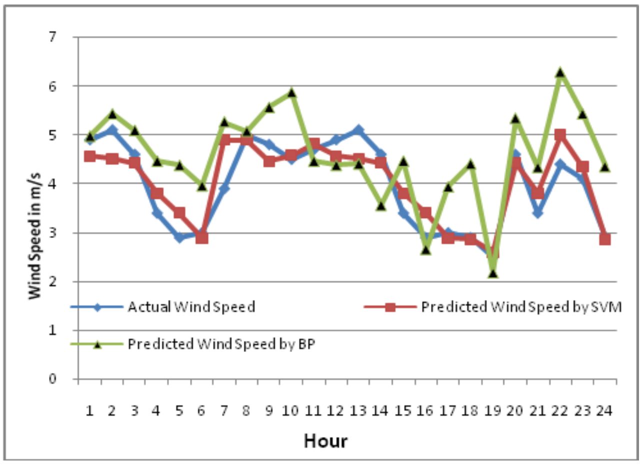

Figure 3. Actual and estimated wind speed for 23rd October 2008

Figure 4 shows comparison between predicted and actual values of the wind speed for 24th October 2008. It is observed that minimum absolute percentage error (APE) obtained by SVM is found as 1.87% and maximum APE is 26.13%. The mean APE found is 7.22%. The values are compared with the values obtained from the Backpropagation Feedforword Network. It is seen that the results of SVM are better than the Feedforword Backpropagated neural Network.

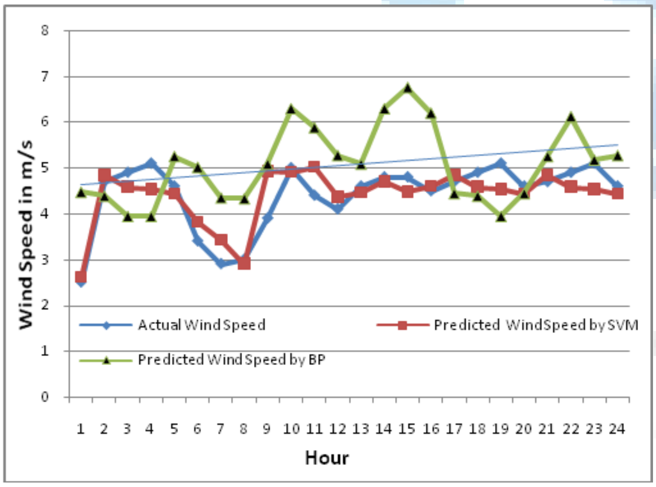

Figure 4. Actual and estimated wind speed for 24th October 2008

Figure 5 shows comparison between actual and predicted values of wind speed of 25th October 2008. Here minimum APE obtained by SVM is found as 1.20% and maximum APE found is 24.13%. The mean APE found is 7.19%. From above Fig. It is seen that the predicted wind speed values obtained by SVM are nearer to the actual values than the Backpropagated Neural Network. Graph shows the superiority of SVM over the other method.

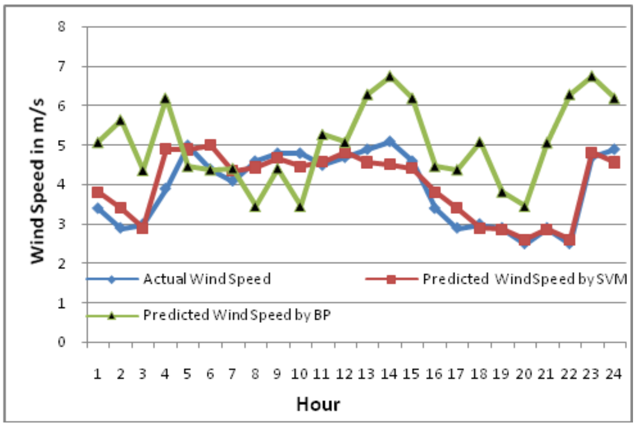

Figure 5. Actual and Estimated wind speed for 25th October 2008

Comparison between actual and predicted values of wind speed of 26th October 2008 is shown in Figure 6. Here minimum APE obtained by SVM is found as 1.12% and maximum APE found is 20.13%. The mean APE obtained by SVM found is 7.67%. Wind speed predicted by SVM is more accurate compared to Feedforword Backpropagation Neural Network

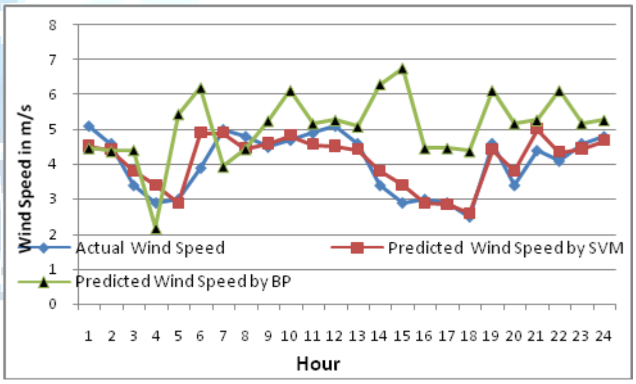

Figure 6. Actual and Estimated wind speed for 26th October 2008

Comparison between actual and predicted wind speed of 27th October 2008 is shown in Figure 7 Here minimum APE obtained by SVM is found as 1.87% and maximum APE found is 17.65%. The mean APE obtained by SVM found is 6.98%

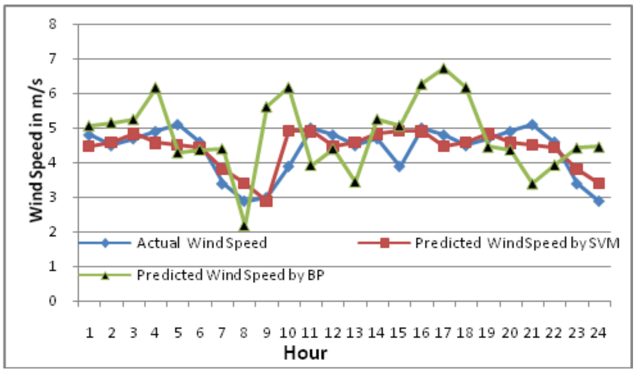

Figure 7. Actual and estimated wind speed for 27th October 2008

Comparison between actual and predicted wind speed of 28th October 2008 is shown in Figure 8. Here minimum APE obtained by SVM is found as 1.05% and maximum APE found is 26.13%. The mean APE obtained by SVM found is 7.14%

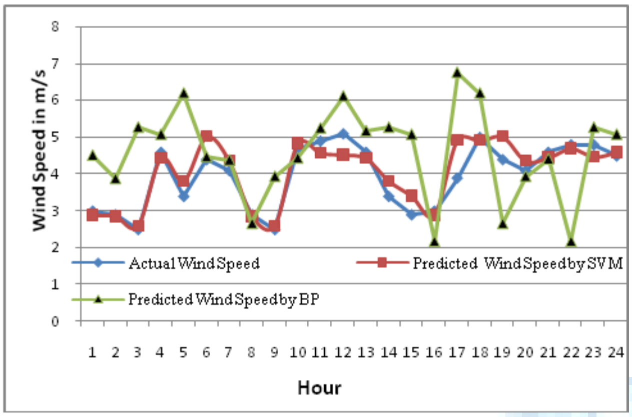

Figure 8. Actual and Estimated wind speed for 28th October 2008

Figure 9 shows comparison between actual and predicted wind speed of 29th October 2008. Here minimum APE obtained by SVM is found as 0.0020% and maximum APE found is 27.90%. The mean APE obtained by SVM found is 7.66%

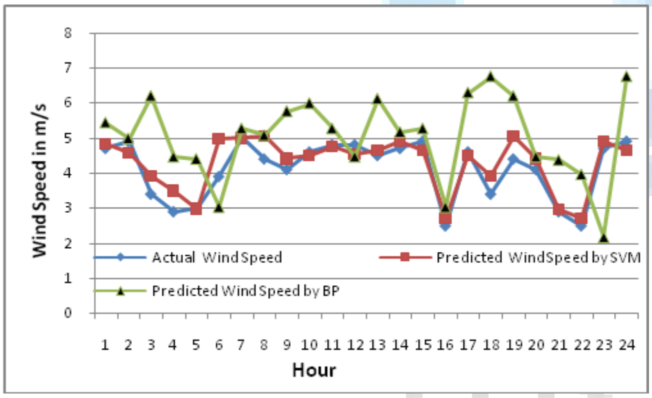

Figure 9. Actual and estimated wind speed for 29th October 2008

Comparison between actual and predicted wind speed of 30th October 2008 is shown in Figure 10. Here minimum APE obtained by SVM is found as 1.020% and maximum APE found is 20.40%. The mean APE obtained by SVM found is 7.87. The comparative graph shows that the wind speed values predicted by SVM is satisfactory as compared to the values obtained by Feedforward Backpropagation neural network.

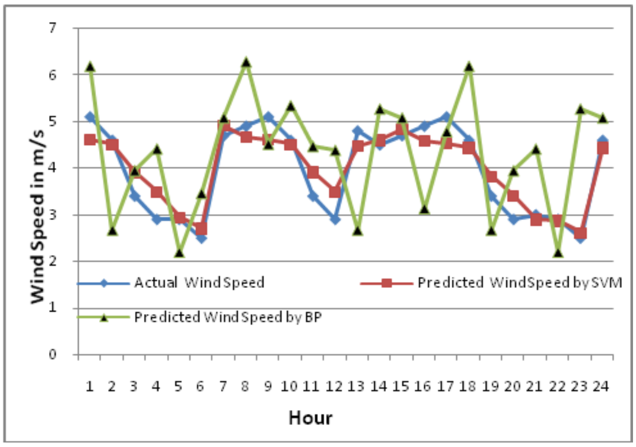

Figure 10. Actual and estimated wind speed for 30th October 2008

Calculation of Correlation Co-efficient between Actual Wind Speed and predicted Wind Speed

Correlation Co-efficient: - A measure of the strength of linear association between two variables. Correlation will always between -1.0 and +1.0. When the value of correlation coefficient is zero, it means there is no correlation. If the correlation is positive, there is a positive relationship (it means that the occurrence of one implies the occurrence of other). If it is negative, the relationship is negative (It means occurrence of one is negatively correlated with occurrence of other).

Formula: Correlation(r) =

$$\left[N\sum XY-\left(\sum X\right)\left(\sum Y\right)/Sqrt\left(\left[N\sum X^2-\left(\sum X\right)^2\right]\left[N\sum Y^2-\left(\sum Y\right)^2\right]\right)\right]$$

(7)

Where

N=Number of values or elements

X=Actual wind speed

Y=Predicted wind speed

$\sum$ XY=Sum of the product of actual wind speed and predicted wind speed

ΣX = Sum of actual wind speeds

ΣY = Sum of predicted wind speeds

ΣX^2= Sum of square actual wind speeds

ΣY^2 = Sum of square predicted wind speeds

Correlation coefficient is calculated for each day. Minimum correlation coefficient is 0.80 and maximum correlation coefficient is 0.94.

VI. Conclusion

There are many method of forecasting wind speed such as numeric weather prediction (NWP), statistical and ANN methods, hybrid methods etc. In this work Support Vector Machine model and Backpropagation model are developed using three years data of wind speed, temperature, barometric pressure, wind gust, wind direction, humidity on hourly average basis to predict hourly wind speed. It is observed that SVM is better than Backpropagation. For the SVM model, Mean Absolute Percentage Error (MAPE) is around 7% and correlation coefficient is close to 1. This justifies the ability of SVM for wind speed prediction task. Results obtained from Support Vector Machine model are satisfactory and promising.

REFERENCES

[1] K. Sreelakshmi, P. Ramakanthkumar, "Short Term Wind Speed Prediction Using Support Vector Machine Model", WSEAS Transactions on Computer Science, Issue 11, Volume 7, pp-1828-1837, November 2008.

[2] K. Sreelakshmi, P. Ramakanthkumar, "Performance Evaluation of Short Term Wind Speed Prediction Techniques", IJCSNS International Journal of Computer science and Network Security, Volume 8 No.8, pp-162-169, August 2008.

[3] L. S. Moulin, A. P. Alves da Silva, M. A. El-Sharkawi, and R. J. Marks II, "Support Vector Machines for Transient Stability Analysis of Large-Scale Power Systems", IEEE Transactions on Power Systems, Volume 19, No. 2, pp 818-825, May 2004.

[4] Yongli Wang, Dongxiao NIU, "Optimization of Support Vector Machines Based on Rough Sets in Seasonal Load Forecasting", Journal of Computational Information Systems 7:5 (2011) 1479-1487.

[5] D. Thukaram, and Rimjhim Agrawal, "Discrimination of Faulted Transmission Lines Using Multi Class Support Vector Machines", 16th National Power Systems Conference, pp. 497 – 502,15th-17th December, 2010.

[6] B. Ravikumar, D. Thukaram and H. P. Khincha, "Application Of Support Vector Machines For Fault Diagnosis In Power Transmission System," Generation, Transmission & Distribution, IET, pp. 119-130, 2008.

[7] S. R. Samantaray and P. K. Dash, "High impedance fault detection in distribution feeders using extended kalman filters and support vector machine", Euro. Transaction on Electrical Power (2009), Published online in Wiley Inter Science (www.interscience.wiley.com) DOI: 10.1002/etep.321.

[8] K.R.Sudha, Y. Butchi, Raju Prasad Reddy.P.V.G.D, "Adaptive Power System Stabilizer Using Support Vector Machine", International Journal of Engineering Science and Technology Volume 2(3), 2010, pp. 442-447.

[9] V. Malathi and N.S. Marimuthu, "Support Vector Machine for Discrimination Between Fault and Magnetizing Inrush Current in Power Transformer", Journal of Computer Science 3 (11): 894-897, 2007.

[10] Robert Lukomski, Kazimierz Wilkosz, "Utilization of Support Vector Machine Classifiers to Power System Topology Verification", Modern Electric Power Systems 2010, Wroclaw, Poland MEPS'10 - paper P48.

[11] Ravikumar B., Thukaram D. and H. P. Khincha, "Knowledge Based Approach for Transmission line Distance Relay Coordination", Fifteenth National Power Systems Conference (NPSC), IIT Bombay, pp. 397-402, December 2008.

[12] P. K. Dash, S. R. Samantaray, and Ganapati Panda, "Fault Classification and Section Identification of an Advanced Series-Compensated Transmission Line Using Support Vector Machine", IEEE Transactions On Power Delivery, Volume 22, No. 1, pp. 67-73, January 2007.

[13] Samantaray, S.R. Dash, P.K. Panda, G., "Fault Classification and Ground detection using Support Vector Machine", National Institute of Technology, Rourkela; TENCON 2006. 2006 IEEE Region 10 Conference, pp. - 1-3, 14-17 Nov. 2006.

[14] Muhammad Nizam, Azah Mohamed, Majid Al-Dabbagh, and Aini Hussain, "Using Support Vector Machine for Prediction Dynamic Voltage Collapse in an Actual Power System", International Journal of Electrical and Electronics Engineering 3:10 2009 pp. – 601-606.

[15] Chih-Chung Chang and Chih-Jen Lin, "LIBSVM: a Library for Support Vector Machines".

[16] Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin, "A Practical Guide to Support Vector Classification".

Mrs. Patil Sangita B has completed her post graduation in Electrical Engineering from Pune Vidyarthi Griha's College of Engineering, Pune and Graduation from AISSM'S Institute of Information & Technology, Pune. Presently She is working as Assistant Professor In G. H. Raisoni Institute of Engineering & Technology. Her area of interest is Artificial Neural Network, Support Vector Machines and Wind Power Generation.

Mr. Bapu K Patil has completed his Masters In Business Studies from Neville Wadia College of Management & Research, Pune-1 and Graduation in Mechanical Engineering from Cusrow Wadia Institute of Technology Pune. Presently he is working as Dy. Manager In Maxion Wheels Ltd., Pune. His area of interest is Hydraulics, Maintenance in Mechanical Engineering.