SHORT TERM WIND SPEED PREDICTION USING SUPPORT VECTOR MACHINE MODEL

1. K. SREELAKSHMI
Assistant Professor, Dept of TCE
RV College of Engineering
Bangalore - 560059
Karnataka, INDIA
Email: sreelakshmik68@gmail.com
Web: http://www.rvce.ac.in

2. Dr. P. RAMAKANTH KUMAR
Professor, Dept of CSE
RV College of Engineering
Bangalore – 560059
Karnataka, INDIA
Email: pramakanth_2000@yahoo.com
Web: http://www.rvce.ac.in

Abstract – Wind speed prediction in short term is required to asses the effect of wind on different objects in action in free space, like rockets, navigating ships and planes, guided missiles satellites in launch etc. Forecasting also helps in usage of wind energy as an alternative source of energy in Electrical power generation plants. The wind speed depends on the values of other atmospheric variables, such as pressure, moisture content, humidity, rainfall etc. This paper reports a Support Vector Machine model for short term wind speed prediction. The model uses the values of these parameters, obtained from a nearest weather station, as input data. The trained model is validated using a part of data. The model is then used to predict the wind speed, using the same meteorological information.

Keywords — Short term wind speed prediction, Support Vector Machine [SVM], forecasting, hyper plane, kernels, classification.

1 Introduction

The Energy is required for the agricultural and industrial activities in any country. The demand for energy is ever increasing. An efficient energy model will help in forecasting and planning energy needs. Wind speed prediction is necessary as wind is an intermittent source of energy and it is also one of the main sources of alternative energy to the depleting petroleum resources [1]. Accurate estimation of wind speed is also required to control the capabilities of wind turbines. The wind speed is different at every point on the surface of the turbines due to three dimensional and time-dependent wind fields. Hence, effective wind speed cannot be measured by an anemometer.

In recent years, Support Vector Machines (SVM) with linear or nonlinear kernels have become one of the most promising learning algorithms for classification as well as for regression which are two fundamental tasks in data mining via the use of kernel mapping, Variants of SVMs have successfully incorporated effective and flexible nonlinear models [21].

A Support Vector Machine (SVM) is an algorithm that uses a nonlinear mapping to transform the original training data into a higher dimension. This is a promising new method for the classification of both linear and non linear data. Within this new dimension, it searches for the linear optimal separating hyper-plane that is, a "decision boundary" separating the tuples of one class from another.

The advantages of the SVM method are as follows:

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

Owing to the innumerable plus points of the SVM method, it has a wide range of applications. SVM have been applied to a number of areas, including handwritten digit recognition, object recognition, speaker identification and also, benchmark time series prediction tests.

2 Literature review

There are various models available for wind speed prediction. They are classified as Statistical, Intelligent systems, Time series, Fuzzy logic, neural network methods. These models use meteorological data as inputs, topological data as inputs, wind turbine data as inputs and are usually suited for long term predictions. They need more computational time and lack in accuracy. Also adaptation to fast changes is not inbuilt into these models. They are based on non-statistical approaches and depend on the experience of a meteorologist.

Time series models are based on historical wind data and statistical methods, Example: ARMA model, which is also called Persistence model [2]. Time series involves number of equations with many variables and hence require intensive computations.

Fuzzy models are also used to estimate wind speed. They are found to be more efficient than the conventional ARMA models [3].

Regression techniques are found to be less efficient compared to Artificial Neural Network model (ANN) models [4].

Kalman filter models are found to be 10% better than the ARMA. These models are found to be superior to ARMA [5].

Artificial Neural Network models are best suited for wind speed prediction applications as they do not require mathematical models and adapts automatically to changes in the inputs to minimize mean square errors. They have the capability to deal with large data sets. [6].

The contemporary models use Back propagation algorithm [7], Radial basis functions [8], Wavelet techniques and Support vector machines for short term forecasting [9].

SVMs [16] are discriminative classifiers based on Vapnik's structural risk minimization principle. They can implement flexible decision boundaries in high dimensional feature spaces. The implicit regularization of the classifier's complexity avoids over fitting and mostly this leads to good generalizations. Some further properties are commonly seen as reasons for the success of SVMs in real-world problems: The optimality of the training result is guaranteed, fast training algorithms exist and little a-priori knowledge is required, i.e. only a labeled training set. Further, SVM models are found to take less computational times compared to ANN models.

Support Vector Machines [SVMs] have achieved excellent recognition results in various pattern recognition applications [13, 14]. Also in off-line optical character recognition (OCR), they have been shown to be comparable or even superior to the standard techniques like Bayesian classifiers or multilayer perceptrons [15].

Support Vector Machines (SVM), the latest neural network algorithms, are introduced to wind speed prediction and their performance is compared with the multilayer perceptron (MLP) neural networks, for modeling mean daily wind speed of Madina city, Saudi Arabia. It is concluded that SVM compare favorably with the MLP model based on the root mean square errors between the actual and the predicted data [19].

Support vector machine models are shown to be more effective than Kalman filter based models for wind speed prediction [20].

3 Support vector machines

Support vector machines (SVM) are a set of related supervised learning methods used for classification and regression. Their common factor is the use of a technique known as the "kernel trick" to apply linear classification techniques to non-linear classification problems. 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.

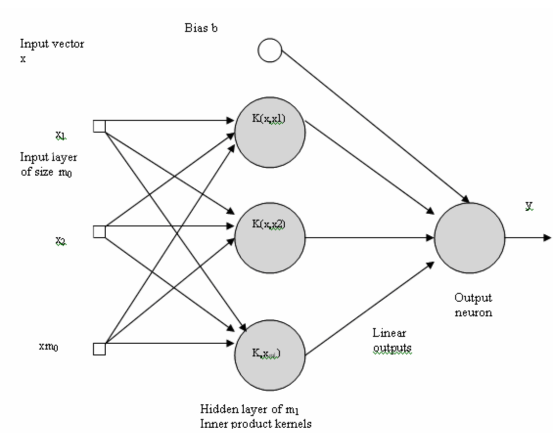

Fig.1 Architecture of support vector machine [12]

3.1 Linear Classification

This method helps in classifying some data points into two classes by a hyper-plane. The hyper plane is so chosen as to separate the data points "neatly", with maximum distance to the closest data point from both classes; this distance is called the margin. Now, if such a hyper-plane exists, it is known as the maximum-margin hyper-plane or the optimal hyper-plane, as are the vectors that are closest to this hyper-plane, which are called the support vectors.

3.1.1 Formalization of linear classifier [17]

Consider data points of the form:

$$\{(x_1, c_1), \ldots (x_n, c_n)\}$$

Where the c_i is either 1 or −1. This constant denotes the class to which the point x_i belongs. This can be viewed as training data, which denotes the correct classification. The SVM is expected to eventually classify, by means of the dividing hyper-plane, that takes the form,

$$w.x - b = 0 \tag{1}$$

The quest for maximum margin leads to support vectors and parallel hyper-planes closest to these support vectors in either class. It can be shown that these parallel hyper-planes can be described by equations

$$w.x - b = 1, \tag{2}$$

$$w.x - b = -1 \tag{3}$$

The hyper-planes are required to maximize the distance from the dividing hyper-plane and to have no data points between them. By using geometry, the distance between the hyper-planes can be found to be 2/|w|; where it is needed to minimize |w|. To exclude data points, it is required to ensure that for all i either

$$w.x_i - b \geq 1 \tag{4}$$

$$w.x_i - b \geq -1 \tag{5}$$

This can be rewritten as:

$$c_i(w.x_i - b) \geq 1 \quad 1 \leq i \leq n \tag{6}$$

The problem now is to minimize |w| subject to the constraint (6).

After the SVM has been trained, it can be used to classify unseen 'test' data. This is achieved using the following decision rule;

$$\hat{c} = \begin{cases} 1, & if \quad w.x + b \geq 0 \\ -1, & if \quad w.x + b \geq\leq 0 \end{cases}$$

Writing the classification rule in its dual form reveals that classification is only a function of the Support vectors, i.e. the training data that lie on the margin.

3.1.2 Non-linear SVM [18]

The original feature space can always be mapped to some higher-dimensional feature space where the training set is separable:

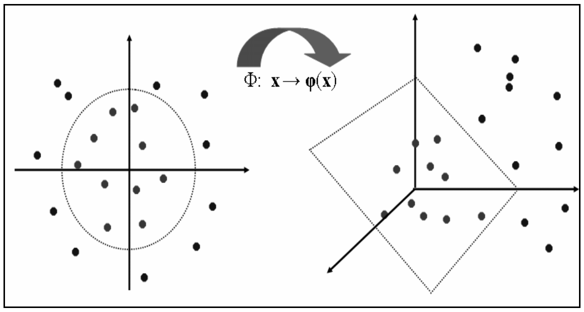

Fig.2 Feature space mapping to higher dimension

Here the main aim of regression is to approximate a function g(x) from a given set of samples:

$$G = \{(x_i, y_i)\}_{i=1}^{N} \tag{7}$$

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.

$$f(x) = \sum_{i=1}^{D} w_i \phi_i(x) + b \tag{8}$$

Where

$\{\phi_i(x)\}_{i=1}^{D}$ Are called features, b and

$\{w_i\}_{i=1}^{D}$ Are coefficients that have to be estimated from the data.

Thus, a nonlinear regression in the low dimensional input space is transferred to a linear regression in a high dimensional (feature) space. The coefficients $\{w_i\}_{i=1}^{D}$ can be found from the data by minimizing the following function:

$$R[w] = \frac{1}{N}\sum_{i=1}^{N} |f(x_i) - y_i|_\varepsilon + \lambda \|w\|^2 \tag{9}$$

where k is a regularization constant and the cost function defined by

$$|f(x_i) - y_i|_\varepsilon = \begin{cases} |f(x) - y| - \varepsilon & for |f(x_i) - y_i| \ge \varepsilon \\ 0 & otherwise \end{cases} \tag{10}$$

is called Vapnik's ε insensitive loss function. It can be shown that the minimizing function has the following form

$$f(x, \alpha, \alpha^*) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) k(x_i . x) + b \tag{11}$$

with

$$\alpha_i \alpha_i^* = 0 \quad , \quad \alpha_i \alpha_i^* \ge 0$$

i= 1,…,N and the kernel function k(xi,x) describes the inner product in the D-dimensional feature space the inner product in the D-dimensional feature space.

$$k(x, y) = \sum_{j=1}^{D} \phi_j(x) \phi_j(y) \tag{12}$$

It is important to note that the features ϕ_j need not be computed; rather what is needed is the kernel function that is very simple and has a known in analytical form. The only condition required is that the kernel function has to satisfy Mercer's condition. Some of the mostly used kernels include polynomial, Gaussian, and sigmoidal. Note also that for Vapnik's e-insensitive loss function, the Lagrange multipliers α_i, α_i^* are sparse, i.e. they result in nonzero values after the optimization only if they are on the boundary, which means that they satisfy the Karush– Kuhn–Tucker conditions.

The coefficients α_i, α_i^* are obtained by maximizing the following form.

$$R(\alpha^*, \alpha) = -\varepsilon \sum_{i=1}^{N} (\alpha_i^* + \alpha_i) + \sum y_i (\alpha_i^* - \alpha_i) - \frac{1}{2} \sum_{i,j=1}^{N} (\alpha_i^* + \alpha_i) x (\alpha_i^* - \alpha_i) k(x_i, x_j) \tag{13}$$

Subject to

$$\sum_{i=1}^{N} (\alpha_i^* - \alpha_i) = 0, \quad 0 \le \alpha_i \quad , \alpha_i^* \le C$$

The number of coefficients α_i, α_i^* are different from zero , and the data points associated with them are called support vectors. The parameters ε are C are decided by the user. Computing b requires a more direct use of the Karush–Kuhn–Tucker conditions that lead to the quadratic programming problems stated above. The key idea is to pick those values

$\alpha k, \alpha_k^*$ for a point x_k on the margin, i.e. α_k or α_k^* in the open interval (0, C). One x_k would be sufficient but for stability purposes it is recommended that one take the average over all points on the margin

4 SVM model methodology

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. 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.

The prerequisites include the three files train.txt, validation.txt and test.txt for training, validation and testing respectively for which the data set is divided into three sets (in the ratio 50%, 25%, 25%). The first set will be used for training. The second set will be used to ascertain the correct kernel and setting to use when performing prediction. The final set of data will be used in prediction.

Short term wind speed prediction involves the following steps:

a) Data Acquisition & Pre-processing
b) Data Conversion & Normalization
c) Statistical Analysis
d) Design of SVM model
e) Training
f) Validation
g) Testing

The weather report and the related parameter values are collected at a Weather Station at periodic time intervals, say every ten minutes. Our work involves the utilization of six different parameters values which are acquired from the weather station report. A historical data of 10 years is considered for the experimentation.

The inputs to the SVM model are raw data from six text files that are read, filtered and processed to obtain normalized data. Patterns are generated and statistical analysis is performed for good correlation among the input data values. A large part of the data is fed into the training network and the remaining part into the testing network. Finally, the wind speed is predicted using the model.

Table1: List of parameters for the network

PARAMETERS	UNITS
Mean temperature	Deg.C
Humidity	%RH
Wind gust	m/s
Wind direction	Deg.M
Barometric pressure	Mb
Wind speed	m/s

Tables 2 and 3 show the correlation results and sample .txt input data file respectively.

Table 2 Correlation results

Correlation	Mean temp	Humidity	Wind gust	Wind dir	Baro pres	Wind speed
Mean temp	1					
Humidity	-0.68531	1				
Wind gust	0.210195	-0.33041	1			
Wind dir	0.005796	-0.00528	-0.02107	1		
Baro pres	-0.35301	0.031238	0.030917	-0.0965	1	
Wind speed	0.193882	-0.31846	0.960779	-0.02305	0.053841	1

Table 3: Sample data input file

NORMAL.TXT - Notepad

File Edit Format View Help

DATE	TIME	MTEMP	HUM	WGUST	WDIR	BARO	WSPEED
2004-02-08	15:50	0.4729	0.5019	0.1356	0.6831	0.8278	0.2070
2004-02-08	16:00	0.4669	0.5038	0.1227	0.6399	0.8277	0.1718
2004-02-08	16:10	0.4610	0.4990	0.1302	0.4823	0.8277	0.2203
2004-02-08	16:20	0.4548	0.5048	0.1432	0.4871	0.8276	0.2291
2004-02-08	16:30	0.4495	0.4885	0.1722	0.4607	0.8276	0.2643
2004-02-08	16:40	0.4462	0.5005	0.1733	0.5055	0.8275	0.2687
2004-02-08	16:50	0.4411	0.5125	0.2034	0.5249	0.8273	0.2599
2004-02-08	17:00	0.4347	0.5279	0.1432	0.7220	0.8274	0.2335
2004-02-08	17:10	0.4275	0.5408	0.1550	0.7830	0.8274	0.1938
2004-02-08	17:20	0.4200	0.5600	0.1216	0.6885	0.8274	0.1806
2004-02-08	17:30	0.4118	0.5720	0.1249	0.6772	0.8274	0.2026
2004-02-08	17:40	0.4030	0.5836	0.1410	0.5897	0.8274	0.2115
2004-02-08	17:50	0.3944	0.5927	0.1195	0.6361	0.8274	0.2115
2004-02-08	18:00	0.3860	0.5970	0.1302	0.6232	0.8275	0.1982
2004-02-08	18:10	0.3774	0.6004	0.1421	0.5692	0.8275	0.2070
2004-02-08	18:20	0.3693	0.6042	0.1367	0.4866	0.8275	0.2159
2004-02-08	18:30	0.3613	0.6066	0.1485	0.5471	0.8274	0.2247
2004-02-08	18:40	0.3543	0.6124	0.1324	0.5589	0.8275	0.2026
2004-02-08	18:50	0.3481	0.6263	0.1098	0.4602	0.8275	0.1498
2004-02-08	19:00	0.3417	0.6335	0.0980	0.5271	0.8277	0.1278
2004-02-08	19:10	0.3353	0.6215	0.0807	0.5514	0.8277	0.1322
2004-02-08	19:20	0.3289	0.6057	0.0872	0.5390	0.8277	0.1278
2004-02-08	19:30	0.3245	0.6009	0.0797	0.4693	0.8278	0.1101

5 Test Cases and results

The following tables and graphs show the tests and results of experiments carried out in using SVM .

Table 4 Test specifications for polynomial kernel

Sl No. of test case :	1
Name of test :	SVM Test 1
C	2
Epsilon	0.1
Kernel Type	Polynomial
Degree	2

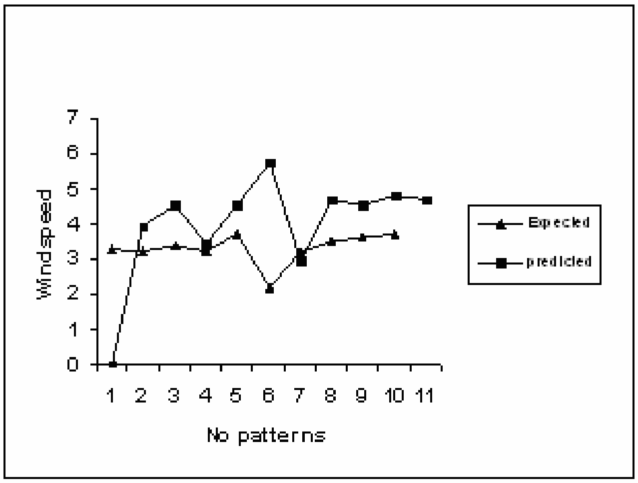

Fig 3 Outputs for polynomial kernel

Table 5 Test specifications for radial kernel

Sl No. of test case :	2
Name of test :	SVM Test 2
C	1
Epsilon	0.01
Kernel Type	Radial
gamma	0.001

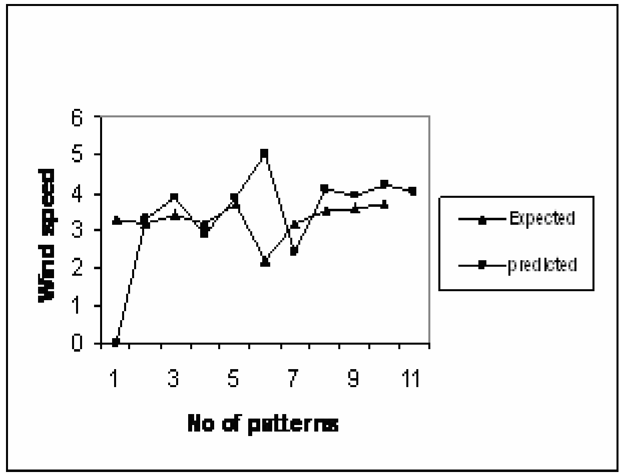

Fig 4 Outputs for radial kernel

Table 6 Test specifications for neural kernel

Sl No. of test case :	3
Name of test :	SVM Test 3
C	1
Epsilon	0.05
Kernel Type	neural
a	0.02
b	0.05

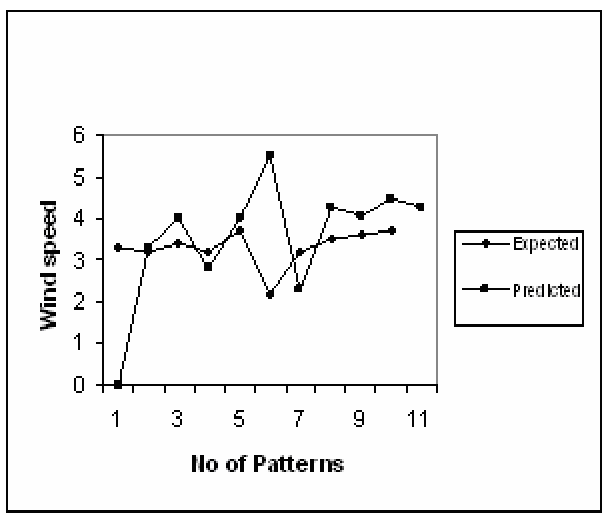

Fig 5 Outputs for Neural kernel

Table 7 Test specifications for Anova kernel

Sl No. of test case :	4
Name of test :	SVM Test 4
C	2
Epsilon	0.02
Kernel Type	Anova
gamma	0.005
degree	3

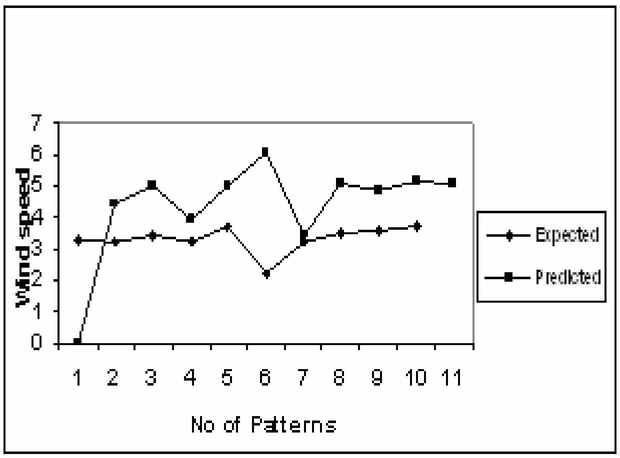

Fig 6 Outputs for Anova kernel

Table 8a Test specifications for Dot kernel

Sl No. of test case :	5
Name of test :	SVM Test 5
C	1
Epsilon	0.01
Kernel Type	Dot

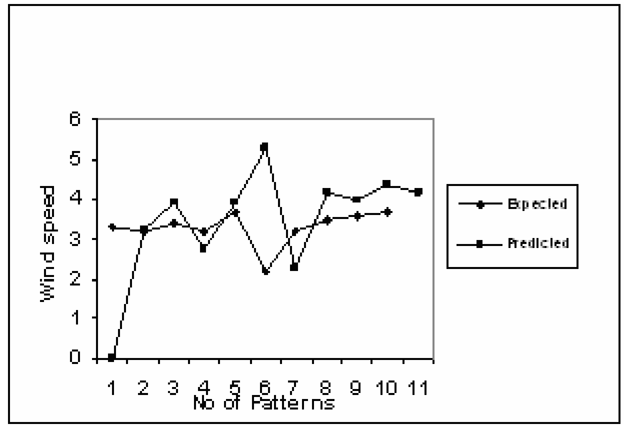

Fig 7a Outputs for Dot kernel with C1 epsilon 1

Table 8b Test specifications for Dot kernel

Sl No. of test case :	6
Name of test :	SVM Test 6
C	2
Epsilon	0.005
Kernel Type	Dot

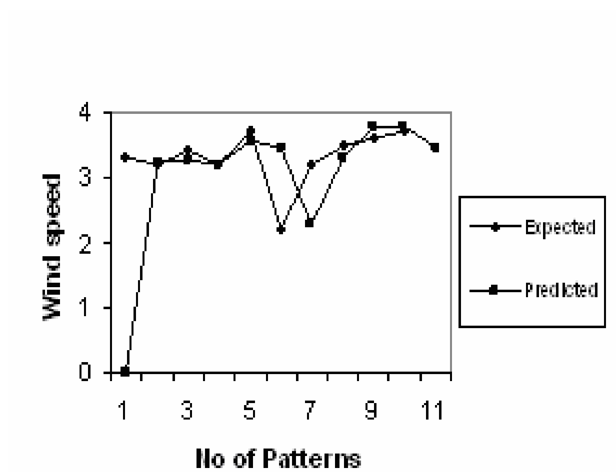

Fig 7b Outputs for Dot kernel with C2 epsilon 2

6 Conclusions

The literature available for wind speed modeling reveals that majority of the models are being utilized for electrical power demand forecasting. Wind speed prediction is carried out as part of power demand forecasting. Though many short term models are presented, the accuracy of the models still need to be improved.

In this paper, the wind speed forecasting is carried out using Support Vector Machine model. The predicted values are fairly matching the expected values. The results of SVM model predictions are almost following the expected pattern of wind speed variations. Analysis shows that the errors in predictions vary with the type of kernel, values of C, Epsilon and the No. of input patterns. Further the error percentage is lowest for 6 input patterns and highest for 5 input patterns, with all the kernels. The dot kernel with a C=2, epsilon = 0.005 seems to be the optimum as it results in a maximum error of 6.49% and minimum error of 2.94% , for 5 and 6 input patterns respectively. This leads to a computational accuracy of 93.51% at 5 input patterns and 97.06% at 6 input patterns.

SVM models are found to take less computational times compared to Artificial Neural Network models, using back propagation algorithms.

7 References

1. X.Wang, G. Sideratos, N. Hatziargyriou and L. H. Tsoukalas, " Wind speed forecasting for power system operational planning" *, 8th Intl. Conf. on Probabilistic Methods Applied to Power Systems,* Iowa State University, Iowa, September 12-16, 2004, pp 470 - 474.

2. Anthanasios Sfetsos and Costas Siriopoulos, "Time series forecasting of averaged data with efficient use of information", *IEEE Trans on Systems, Man and Cybernetics-Part A: Systems and Humans"*, Digital object identifier- 10.1109/TMSCA2005.851133.

3. Ioannis.G.Damousis, Mians C. Alexadis, John B Theocharis and Petros,S.Dokopoulos ,"A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation", *IEEE Trans. on Energy Conversions,* Vol. 19, No.2, PP 352 – 361, June 2004.

4. Shuhui Li, Donald C Wunsch, Edgar O'Hair and Michael G Giesselmann, " Comparative analysis of regression and artificial neural network models for wind turbine power curve estimation", *Journal of Solar Energy Engineering,* Vol. 123, pp 327 – 332, November 2001

5. Shuhui Li, " Wind power prediction using recurrent multilayer perceptron neural networks", *0-7803-7989-6/03/$17.00 ©2003 IEEE,* PP- 2325 – 2330

6. Henrique Steinhertz Hippert, CarlosEduardo Pedreira and Reinaldo Castro Souza, "Neural networks for short term load forecasting: A review and evaluation", *IEEE Trans. on Power Systems,* Vol 16. No.1 pp 44- 55, February 2001.

7. T.G.Barbounis and J B Theocharis, "Locally recurrent neural networks optimal filtering algorithms: application to wind speed prediction using spatial correlation", Proceedings *of The Intl. Joint Conf. on Neural Networks,* Canada, July 31 – August 4, 2005, pp 2711 - 2716.

8. G. Sideratos and N. Hatziargyriou, "Using radial basis neural networks to estimate wind power production", *1-4244-1298-6/07/$25.00 © 2007 IEEE.*

9. Enrique Romero and Daniel Toppo, "Comparing support vector machines and feed forward neural networks with similar hidden layer weights", *IEEE Trans. on Neural Networks,* Vol 18, no 3. May 2007, pp 959 – 963.

10. T.G. Barbounis, J B Theocharis, Minas. C. Alexadis and Petros. S. Dokopoulos, " Long term wind speed and power forecasting using local recurrent neural network models", *IEEE Trans. on Energy Conversion* ,Vol 21, no. 1, March 2006, pp 273 – 284.

11. Alireza Khotanzad, Reza Afkhami Rohani and Dominic Maratukulam, "ANNSTLF – artificial neural network short-term load forecaster – generation three." *IEEE Trans. on Power Systems*, Vol 13, No.4, November 1998, pp 1413-1422.

12. Simon Haykin, *"Neural networks- a comprehensive foundation*", 2 edition, Pearson Education, 2003.

13. C. Burges. "A tutorial on support vector machines for pattern recognition", *Data Mining and Knowledge Discovery,* Vol. 2, No 2, pp121–167, 1998.

14. N. Cristianini and J. Shawe-Taylor. *Support Vector Machines.* Cambridge University Press, 2000.

15. D. DeCoste and B. Schölkopf. "Training invariant support vector machines", *Machine Learning,* Vol 46, No1/3, p161, 2002.

16. Heloise Hse, A. Richard Newton. *"Sketched Symbol Recognition using Zernike Moments".*

17. Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin, *A Practical Guide to Support Vector Classification*.

18. Chang, C.-C. And C.-J. Lin *LIBSVM: a library for support vector machines*, (2001).

19. M A Mohandes, T O Halawani, S Rehman, Ahmed A Hussain, "Support Vector Machines for Wind speed prediction", *Renewable Energy,* Vol.29, Issue 6, May 2004, pp 939-947.

20. Xiyun Yang Xiaojuan Han Lingfeng Xu Yibing Liu, "Soft Sensor Based on Support Vector Machine for Effective Wind Speed in Large Variable Wind Turbines", Proc. of ICARCV '06. 9th International Conference on Control, Automation, Robotics and Vision, 5-8 Dec. 2006, pp 1-4.

21. Yuh-Jye Lee, Su-Yun Huang, "Reduced Support Vector Machines: A Statistical Theory", *IEEE Trans. On Neural Networks",* Vol. 18, No 1, January 2007.

8 Biographies

The first author is K. Sreelakshmi, Assistant Professor, Dept of Telecommunication Engg., RV College of engineering Bangalore – 560059. She obtained her BE and ME degrees from Bangalore university, with specializations in Electronics and Communication and is currently pursuing her PhD degree. Her research interests are in the field of Neural network based models and their applications, Electromagnetics and microwaves. She is a member of IETE and ISTE. She has worked in Indian Institute of Science, Bangalore as a JRF. Her work at IISc included characterization of electronic devices at cryogenic temperature, which has resulted in an international paper publication in the Journal, “Cryogenics “.

The second author is Dr.Ramakanthkumar P Bangalore 25/01/1970. He Completed his B.E in Computer Science and Engineering from S.J.College of Engineering, Mysore, and M.S in Software Systems from BITS, Pilani. He completed his Ph.D. from Mangalore University in the area of Pattern Recognition. He has authored more than five publications in reputed journals. His areas of interest are Image Processing, Pattern Recognition and Natural Language Processing. He is a member of the Computer Society of India (CSI) and a life member of Indian Society for Technical Education (ISTE). He has completed number research and consultancy projects for DRDO organizations.