Turkish Journal of Science & Technology
Volume 11 (1), 1-7, 2016

Estimation of Monthly Mean Ambient Temperatures with Support Vector Machines

H. ESEN, A. SENGUR
Department of Energy Systems Engineering, Faculty of Technology, Firat University, 23119, Elazig, TURKEY
hikmetesen@firat.edu.tr

(Received: 11.07.2015; Accepted: 20.01.2016)

Abstract

In this study, the support vector machines (SVMs) have been used for the estimation of monthly mean ambient temperature in Elazığ (38.41˚ N, 39.14˚ E), Turkey. The model was trained and tested for four years (2002-2005) of some monthly mean meteorological values. Inputs of the network were relative humidity, local pressure, vapour pressure, and wind velocity monthly values and the output was the monthly mean outdoor temperature. The efficiency of the proposed method was demonstrated by using the 4-fold cross validation test. The proposed SVM model produced the most accurate results for partition 4 that's why the minimum root-mean squared (RMS), coefficient of variation (COV) and mean error function (MEF) and maximum coefficient of multiple determinations (R^2) values were obtained for these partitions. It is found that RMS value is 0.7691, the R^2 value is 0.9980, COV value is 5.5586, and MEF value is 1.6339 for partition 4. These results testify that the SVM can be a valuable tool for monthly ambient temperature prediction in particular and other meteorological predictions in general.

Keywords: Ambient temperature, estimation, meteorology, support vector machine, Elazığ

Aylık Ortalama Dış Hava Sıcaklığının Destek Vektör Makineleri ile Tahmini

Özet

Bu çalışmada, Elazığ ilinin aylık ortalama dış hava sıcaklığının tahmini için destek vektör makineleri (support vector machines (SVMs)) yöntemi kullanıldı. Model, dört yıllık (2002-2005) bazı aylık ortalama meteorolojik değerler için eğitildi ve test edildi. Modelin giriş değerleri bağıl nem, yerel basınç, buhar basıncı ve rüzgar hızlarının aylık değerleri iken çıkış değeri ise aylık ortalama dış hava sıcaklığıdır. Önerilen metodun verimi, 4-katlı (kısımlı) çapraz geçerlilik testi kullanılarak gösterildi. Bu kısımlar içinde minimum ortalama karekök (RMS), değişim katsayısı (COV) ve ortalama hata fonksiyonu (MEF) değerleri ile maksimum çoklu saptama katsayısı (R^2) değerlerine sahip olan 4. kısım çapraz geçerlilik için en doğru sonuçları veren SVM modeli kuruldu. 4. kısım için sırasıyla RMS, R^2, COV ve MEF değerleri 0.7691, 0.9980, 5.5586 ve 1.6339 olarak bulundu. Bu sonuçlar, destek vektör makinesinin (SVM) kısmi olarak aylık ortalama dış hava sıcaklık tahmini için genel olarak da başka meteorolojik tahminler için faydalı bir araç olabileceğini kanıtlamaktadır.

Anahtar kelimler: Dış hava sıcaklığı, tahmin, meteoroloji, destek vektör makinesi, Elazığ

1. Introduction

Estimation of meteorological data plays a very important role in the modern Heating, Ventilation and Air Conditioning (HVAC) systems for guaranteeing the thermal comfort, energy saving and reliability. The knowledge relating to the climatic parameters like monthly or hourly mean values of relative humidity, pressure, rainfall, visibility, ambient temperature and wind velocity are valuable in the thermal analysis of building, heating and cooling load calculations to decide the accurate sizing of an air-conditioning system for thermal comfort and in the performance assessment and optimum design of numerous solar energy systems [1, 2]. Dombaycı and Çivril [2] used the artificial neural network for the estimation of hourly ambient temperature in Denizli, Turkey. The training and test results show that there was a good correspondence between the predicted and measured values. Accurate estimation of monthly or hourly air temperatures has a number of vital applications in the industry, agriculture and the environment [3]. For instance, the information of variation in ambient temperature has a significant value in predicting the solar radiation [4-6], hourly energy consumption and cooling load estimation in buildings [7-8] and room air temperature prediction [2, 9].

Superior generalization performance is obtained from SVM regression and more importantly, the performance does not depend on the dimensionality of the input data. The SVM is derived from the statistical learning theory [10-11], and is a two-layer network with the inputs transformed by the kernels corresponding to a subset of the input data. The output of the SVM is a linear function of the weights and the kernels. The weights and the structure of the SVM are obtained simultaneously by constrained minimization for a given precision level of the modelling error. In the constrained minimization, kernels corresponding to data points that are within the error bounds are removed. The support vector regression (SVR) is formed by the retained kernels [11-13], and the data points associated with the retained kernels are referred to as the support vectors (SVs). Since the kernels of the SVR are similar to the basis functions of the radial basis function (RBF) network with scatter partitioning [11, 14], it is shown here that the SVR can be reformulated as a RBF network with basis functions normalized such that they form a partition of unity [15, 16]. SVM implement classifiers of an adjustable complexity, controlling the latter for optimal generalization ability i.e. the performance for the future unknown samples. Applicational aspects of SVMs had been shown in Ref. [17].

In this study, an SVM model was developed in order to use to estimate the monthly ambient temperature in Elazığ (38.41° N, 39.14° E), Turkey. The model was trained and tested with four years (2002-2005) of monthly mean temperature values obtained from Turkish State Meteorological Service. The performance of the proposed mean temperature prediction methodology was evaluated by using several statistical validation parameters. Moreover, we employed the cross-validation test for measure the robustness of the proposed methodology. We used four-fold cross validation test and we calculate the performance evaluation methods for each partition

2. Support Vector Machine (SVM)

Assume a set of training data $\{(x_1, y_1),....,(x_\ell, y_\ell)\}$, where each $x_i \subset R^n$ denotes the input space of the sample and has a corresponding target value $y_i \subset R$ for *i=1,…, l* where *l* corresponds to the size of the training data [10, 18]. The consideration of the regression problem is to determine a function that can approximate future values accurately. The generic SVM estimating function takes the form:

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

where $w \subset R^n, b \subset R$ and Φ denotes a non-linear transformation from R^n to high dimensional space. The goal is to find the value of w and b such that values of x can be determined by minimizing the regression risk:

$$R_{reg}(f) = C\sum_{i=0}^{\ell}\Gamma(f(xi) - yi) + \frac{1}{2}\|w\|^2\,, \tag{2}$$

where $\Gamma(\cdot)$ is a cost function, C is a regularization constant and vector w can be written in terms of data points as:

$$w = \sum_{i=1}^{\ell}(\alpha_i - \alpha_i^*)\Phi(x_i)\,. \tag{3}$$

By substituting Equation (3) into Equation (1), the generic Equation can be rewritten as:

$$f(x) = \sum_{i=1}^{\ell}(\alpha_i - \alpha_i^*)(\Phi(x_i)\cdot\Phi(x)) + b$$
$$= \sum_{i=1}^{\ell}(\alpha_i - \alpha_i^*)k(x_i, x) + b\,. \tag{4}$$

In Equation (4) the dot-product can be replaced with function $k(x_i, x)$, known as the kernel function. Kernel functions enable dot product to be performed in high-dimensional feature space using low dimensional space data input without knowing the transformation Φ. All kernel functions must satisfy Mercer's condition that corresponds to the inner product of some feature space. The RBF is commonly used as the kernel for regression and it is defined by [10, 18]:

$$k(x_i, x) = \exp\left\{-\gamma |x - x_i|^2\right\}. \quad (5)$$

The ε -insensitive loss function is the most widely used cost function. The function is denoted as following:

$$\Gamma(f(x)-y) = \begin{cases} |f(x)-y| - \varepsilon, & for \ |f(x)-y| \geq \varepsilon \\ 0 & otherwise \end{cases}. \quad (6)$$

By solving the quadratic optimization problem in Equation (7), the regression risk in Equation (2) and the ε -insensitive loss function Equation (6) can be minimized:

$$\frac{1}{2}\sum_{i,j=1}^{\ell}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)k(x_i, x_j) - \sum_{i=1}^{\ell}\alpha_i^*(y_i - \varepsilon) - \alpha_i(y_i + \varepsilon)$$

$$\sum_{i=1}^{\ell}\alpha_i - \alpha_i^* = 0, \quad \alpha_i, \alpha_i^* \in [0, C]. \quad (7)$$

The Lagrange multipliers α_i and α_i^*, can be used to represent solutions to the above quadratic problem that act as forces pushing predictions towards target value y_i. Only the non-zero values of the Lagrange multipliers in Equation (7) are useful in estimation the regression line and are known as support vectors. For all points inside the ε -tube, the Lagrange multipliers equal to zero do not contribute to the regression function. The constant *C* introduced in Equation (2) determines penalties to estimation errors [10, 18].

Now, the value of w in terms of the Lagrange multipliers has to be solved. For the variable b, it can be computed by applying Karush-Kuhn-Tucker (KKT) conditions which, in this case, implies that the product of the Lagrange multipliers and constrains has to equal zero:

$$\begin{aligned} &\alpha_i(\varepsilon + \zeta_i - y_i + (w, x_i) + b) = 0 \\ &\alpha_i^*(\varepsilon + \zeta_i^* + y_i - (w, x_i) - b) = 0 \end{aligned}, \quad (8)$$

and

$$\begin{aligned} &(C - \alpha_i)\zeta_i = 0 \\ &(C - \alpha_i^*)\zeta_i^* = 0 \end{aligned}, \quad (9)$$

where ζ_i and ζ_i^* are slack variables used to measure errors outside the ε -tube. Since $\alpha_i, \alpha_i^* = 0$ and $\zeta_i^* = 0$ for $\alpha_i^* \in (0, C)$, b can be computed as follows:

$$\begin{aligned} &b = y_i - (w, x_i) - \varepsilon \quad for \ \alpha_i \in (0, C) \\ &b = y_i - (w, x_i) + \varepsilon \quad for \ \alpha_i^* \in (0, C) \end{aligned}. \quad (10)$$

Putting it all together, we can use SVM and SVR without knowing the transformation [12, 18].

3. SVMs for Predicting the Mean Temperature

The dataset used in this study was the monthly mean temperature values belonging the years 2002 through 2005, measured by Turkish State Meteorological Service for Elazığ. The SVM based automatic mean temperature prediction model has 4 inputs and one output. (SVM) have been introduced within the context of statistical learning theory and structural risk minimization so it is accepted to be a powerful methodology and it is used in a wide range of applications such as classification, regression, and estimation which has also led to many other recent developments in kernel based learning methods in general [19].

The key to obtaining a highly accurate SVM estimation is to choose a proper set of regularization parameter *C*, and kernel parameters. For obtaining the optimum SVM parameters, such as regularization parameter, and optimum kernel function and parameters several test have been carried out with different kernel functions. The constant ε is used to find the target function that not only lies as close as possible to the border of the ε-tube but also is as flat as possible. The larger ε is, the flatter the function will be, and the fewer support vectors will be. On the other hand, however, a larger ε leads to larger estimation errors. Therefore, the value of ε ought to be determined in a way that it is proportional to the input noise level σ. On the other hand, a large *C* assigns higher penalties to errors so that the regression is trained to minimize error with lower generalization while a small C assigns fewer penalties to errors; this allows the minimization of margin with errors, thus higher generalization ability. If *C* goes to infinitely large, SVR would not allow the occurrence of any error and result in a complex model, whereas when C goes to zero, the result would tolerate a large amount of errors and the model would be less complex.

For obtaining the optimum parameters a grid search algorithm was employed in the parameter plane for RBF kernel function where minimum RMS value is determined. The RMS parameter is defined as following;

$$RMS = \sqrt{\frac{\sum_{m=1}^{n}(y_{pre,m} - t_{mea,m})^2}{n}} \quad (11)$$

Moreover, several statistical methods, the coefficient of multiple determinations (R^2) and the coefficient of variation (cov) were used to compare predicted and actual values for computing the model validation. The R^2, cov and mean error function (MEF) parameters are denoted at Eq. (12), Eq. (13) and Eq. (14) respectively;

$$R^2 = 1 - \frac{\sum_{m=1}^{n}(y_{pre,m} - t_{mea,m})^2}{\sum_{m=1}^{n}(t_{mea,m})^2} \quad (12)$$

$$COV = \frac{RMS}{\left|\bar{t}_{mea,m}\right|}100 \quad (13)$$

$$MEF = \frac{1}{n}\sum_{m=1}^{n}\frac{|y_{pre,m} - t_{mea,m}|}{\max(t_{mea}) - \min(t_{mea})}100 \quad (14)$$

where *n* is the number of data patterns in the independent data set, $y_{pre,m}$ indicates the predicted, $t_{mea,m}$ is the measured value of one data point *m*, and $\bar{t}_{mea,m}$ is the mean value of all measured data points.

The efficiency of the proposed method was demonstrated by using the 4-fold cross validation test. In 4-fold cross validation, the dataset is randomly split into four exclusive subsets *(X1...X4)* of equal size and the holdout method is repeated 4 times. These subsets contain 30, 30, 30 and 30 samples *(30+30+30+30=120)* respectively. At each time, one of the four subsets is used as the test set and the other three subsets are put together to form a training set. The advantage of this method is that it is not important how the data is divided. Every data point appears in a test set only once, and appears in a training set 3 times. Therefore, the verification of the efficiency of the proposed method against to the over-learning problem should be demonstrated.

4. Results and Discussions

In order to obtain the optimal model parameters of the SVM, as we mentioned earlier, a grid search algorithm was employed in the parameter space. By using the RBF-kernel, after applying the grid search algorithm, the obtained optimum C and sigma value was 155 and 20 respectively. The corresponding RMS, COV, R^2 and MEF values for these parameters are given in Table 1.

Table 1. The statistical model validation results of 4-fold cross-validation

	Statistical model validation parameters			
Partition	RMS	R^2	COV	MEF
1	1.1955	0.9935	10.0160	3.2129
2	0.8076	0.9977	5.8623	1.8049
3	0.7214	0.9979	5.7055	1.5263
4	0.7691	0.9980	5.5586	1.6339
Mean	0.8734	0.9968	6.7856	2.0445
Std. Dev.	0.2176	0.0022	2.1572	0.7873

One can observe several conclusions from Table 1. For example, the proposed SVM model produced the most accurate results for partition 4 that’s why the minimum RMS, COV and MEF and maximum R^2 values were obtained for this partition. On the other hand, the worse results were obtained for partition 1. The COV value, which was 10.016, is the highest value in all COV values. Furthermore, the COV, RMS and MEF values for partition 1 is bigger than the mean COV, RMS and MEF values. The related other actual and predicted temperature values were given in Fig.1 to Fig.6 respectively for each partition.

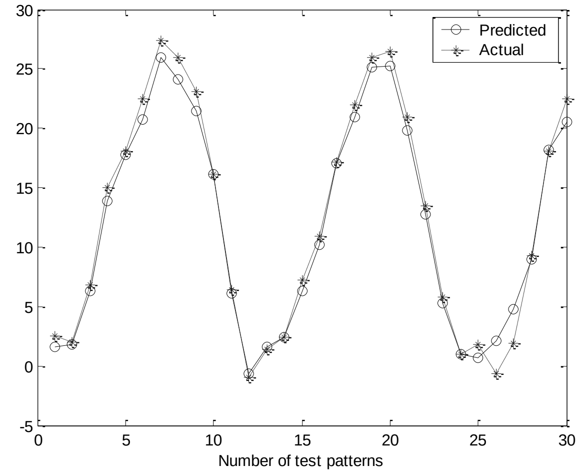

Fig. 1. The performance of the SVM on the partition 1

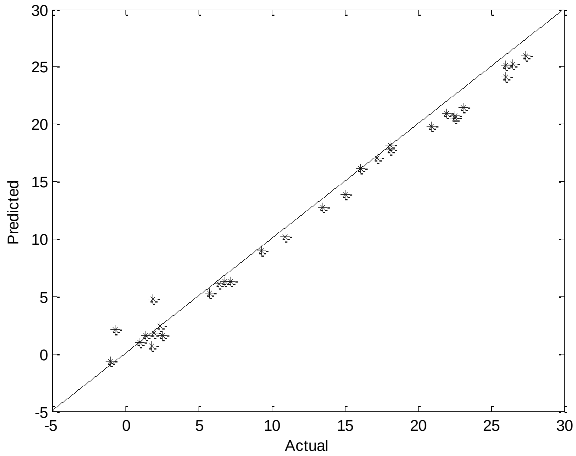

Fig. 2. Actual vs predicted for SVM model at partition 1

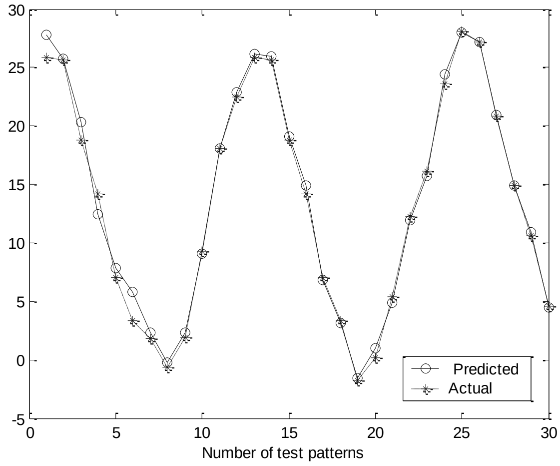

Fig. 3. The performance of the SVM on the partition 2

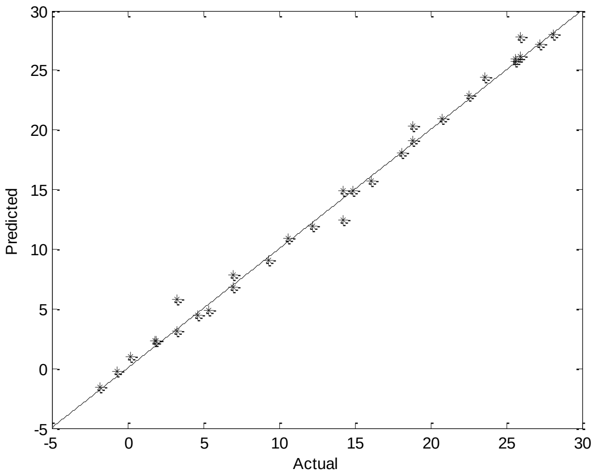

Fig. 4. Actual vs predicted for SVM model at partition 2

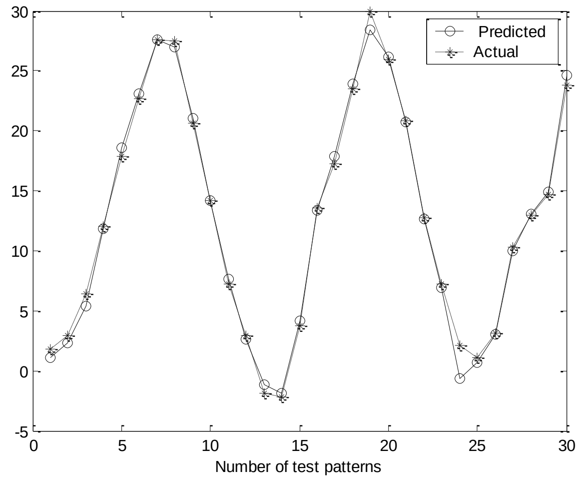

Fig. 5. The performance of the SVM on the partition 3

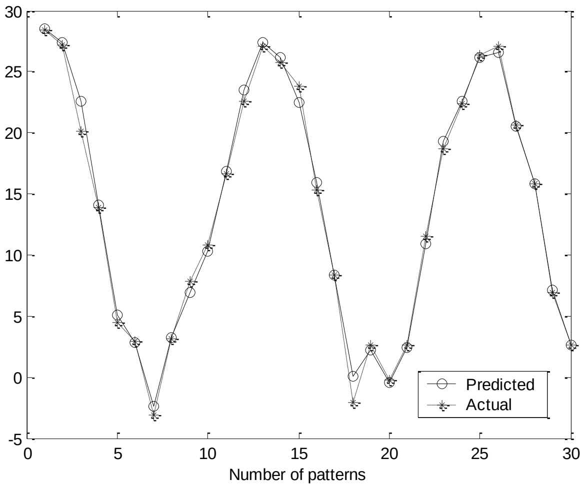

Fig. 7. The performance of the SVM on the partition 4

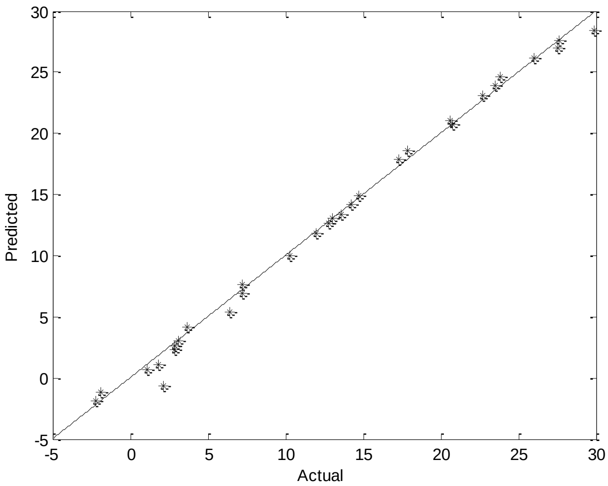

Fig. 6. Actual vs predicted for SVM model at partition 3

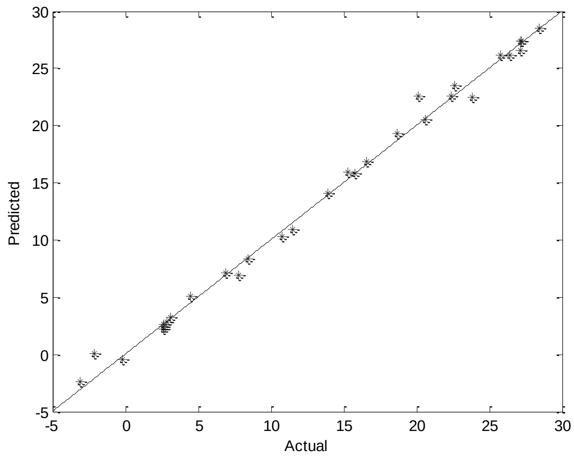

Fig. 8. Actual vs predicted for SVM model at partition 4

The actual and the predicted temperature values were given in Fig. 7 and Fig. 8, respectively for partition 4.

5. Conclusions

In this study, a SVM based methodology was intended to adopt the prediction of monthly mean temperature values belonging the years 2002 through 2005, measured by Turkish State Meteorological Service for Elazığ. The performance of the proposed mean temperature prediction methodology was evaluated by using several statistical validation parameters. Moreover, we employed the cross-validation test for measure the robustness of the proposed methodology. We used four-fold cross validation test and we calculate the performance evaluation methods for each partition and obtained the mean 0.8734, 0.9968, 6.7856 and 2.0445 RMS, R^2, COV and MEF values respectively. Future work will attempt to further improve the estimation accuracy by using dedicated seasonal models and including temperature data on a larger number of previous days.

For obtaining the optimum SVM parameters a grid search algorithm was employed on parameter plane with RBF kernel function. We also used linear-kernel function but no performance improvement was obtained so we did not give the results here. From the above results, it can be concluded that SVM is a feasible method for prediction of the temperature values. The computation of SVM model is faster compared with other machine learning techniques, because there are fewer free parameters and only support vectors are used in the generalization process.

6. References

1. Parishwad, G. V., Bhardwaj, R. K., and Nema, V.K. (1998). Prediction of monthly-mean hourly relative humidity, ambient temperature, and wind velocity for India. *Renewable Energy* **13**(3), 363-380.
2. Dombaycı, Ö. A., and Çivril, Ö. (2006). Estimation of hourly mean ambient temperatures with artificial neural networks. *Mathematical and Computational Applications*, **11**(3), 215-224.
3. Abdel-Aal, R. E. (2004). Hourly temperature forecasting using abductive networks. *Engineering Applications of Artificial Intelligence* **17**(5), 543-556.
4. Sözen, A., Arcaklıoglu, E., Özalp M., and Kanıt, E. G. (2004). Solar-energy potential in Turkey. *Applied Energy* **80**(4), 367-381.
5. Sözen, A., Arcaklıoglu, E., Özalp, M., and Kanıt, E. G. (2004). Use of artificial neural networks for mapping of solar potential in Turkey. *Applied Energy* **77**(3), 273-286.
6. Dorvlo, A. S. S., Jervase, J. A., and Lawati A. A. (2002). Solar radiation estimation using artificial neural networks. *Applied Energy* **71**(4), 307-319.
7. Mihalakakou, G., Santamoruris, M., and Tsangrassoulis, A. (2002). On the energy consumption in residential buildings. *Energy and Buildings* **34**(7), 727-736.
8. Ben-Nakhi, A. E., and Mahmoud, M. A. (2004). Cooling load prediction for buildings using general regression neural networks. *Energy Conversion & Management* **45**(13-14), 2127-2141.
9. Yang, I. H., and Kim, W. K. (2004). Prediction of the time of room air temperature descending for heating systems in buildings. *Building and Environment*, **39**(1), 19-29.
10. Vapnik, V. N. (1995). The Nature of Statistical Learning Theory. Springer, New York.
11. Zhong, Z. D., Zhu, X. J., and Cao, G. Y. (2006). Modeling a PEMFC by a support vector machine. *Journal of Power Sources* **160** (1), 293–298.
12. Drezet, P. M. L., and Harrison, R. F. (1998). Support vector machines for system identification, UKACC International Conference on CONTROL'98, 1–4 September 1998, pp. 688–692.
13. Schölkopf, B., Bartlett, P., Smola, A. J., and Williamson, R. (1999). Shrinking the tube: a new support vector regression algorithm. In: Kearns, M.S., Solla, S.A., Cohn, D.A. (Eds.), *Advances in Neural Information Processing Systems* Vol. 11. MIT Press, Cambridge, MA.
14. Schölkopf, B., Sung, K., Burges, C. J. C., Girosi, F., Niyogi, P., Poggio, T., and Vapnik, V.N. (1997). Comparing support vector machines with gaussian kernels to radial basis function classifiers. *IEEE Transactions on Signal Processing* 45, 2758–2765.
15. Harris, C. J., Wu, Z.Q. and Gan, Q. (1999). Neurofuzzy state estimators and their applications. In: Sinha, N.K., Gupta, M.M. (Eds.), Soft Computing and Intelligent Systems. Academic Press, New York, pp. 377–402.
16. Chan, W. C., Chan, C. W., Cheung, K.C., and Harris, C.J. (2001). On the modelling of nonlinear dynamic systems using support vector neural networks. *Engineering Applications of Artificial Intelligence* **14**(2), 105-113.
17. Belousov, A.I., Verzakov, S.A., and Frese, J.V. (2002). Applicational aspects of support vector machines. *Journal of Chemometrics* **16**, 482-489.
18. Wu, C-H., Wei, C-C., Chang, M-H., Su, D-C., and Ho, J-M. (2008), Travel time prediction with support vector regression, avaliable at: www.iis.sinica.edu.tw/papers/hoho/2844-F.doc.
19. Esen, H., Inalli, M., Sengur, A., and Esen, M. (2008). Modelling a ground-coupled heat pump system by a support vector machines. *Renewable Energy* **33**(8), 82-91.