

Daily Reference Evapotranspiration Estimation Based on Least Squares Support Vector Machines

Dachun Chen*

Hehai University, Nanjing 210098, P.R. China,
Water Conservancy and Civil Engineering College,
Xinjiang Agricultural University, Urumqi 830052, P.R. China
vision_studio@sina.com

Abstract. As a key hydrological parameter, daily reference evapotranspiration (ET_o) determines the accuracy of the hydrological number of the crop, and, consequently, the regional optimization disposition of water resources. At present, the main methods for ET_o estimation are the Penman-Monteith (PM) equation and its modified formula, both of which are based on climatic factors such as temperature, radiation, humidity, and wind velocity, among others. Unfortunately, these required data are not always available in Xinjiang Uighur Autonomous Region, China, which is a semiarid area. Hence, this paper puts forward, for the first time, a least squares support vector machine (LSSVM) model for estimating ET_o. The LSSVM model used in this study considers climatic factors as input variables and the ET_o calculated by the Penman-Monteith equation as an output variable. Compared with the artificial neural network (ANN) model, which was developed with the same data, LSSVM prediction shows higher accuracy, efficiency, and generalization performance. Therefore, it can be used as a complementary ET_o estimation method.

Keywords: Daily reference evapotranspiration, LSSVM, ANN, Semiarid area, Xinjiang.

1 Introduction

Accurate evapotranspiration (ET_o) estimates are necessary to determine the water requirement of crops for irrigation scheduling. The only recommended method for ET_o estimation is the physically-based complex Penman-Monteith (PM) equation using complete meteorological data [1]. The ANN model has been successfully applied to evapotranspiration estimation. Kumar [2] built an ANN model for ET_o estimation and concluded that it could predict ET_o better than the conventional method, that is, the PM equation. Trajkovic [3] applied a sequentially adaptive radial basis function (RBF) network for ET_o forecasting. As well, many studies have used neural networks to model ET_o as a function of climatic variables [4–12].

* Corresponding author.

The support vector machine (SVM) approach was recently recognized for its ability to capture nonlinear regression relationships between variables [13]. Eslamian and Kisi [13,14] investigated the accuracy of SVM in modeling ETo, and their comparison results revealed that the SVM could be successfully used in modeling the ETo process.

This paper discusses ETo estimation by using a least squares support vector machine (LSSVM). In comparison with ANN and SVM, it shows higher accuracy, efficiency, and generalization performance.

2 Materials and Methods

2.1 Investigation Areas

The 145th Regiment of Shihez Reclamation Area, located south of the XinJiang Uighur Autonomous Region, was explored in the present study. The data sets were gathered from meteorological stations located at east longitude 86.05°, north latitude 44.32°, and altitude of 443.7 m. Data for the following weather variables used for this study were gathered from 1998 to 2009: precipitation (mm), relative humidity (%), wind speed (m/s), and maximum and minimum temperature (°C).

2.2 Estimation of Reference Evapotranspiration

The Penman-Monteith equation for calculation of ETo proposed by Allen [1] is expressed as:

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T + 273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (1)$$

where ETo is the reference evapotranspiration [mm day⁻¹], R_n is the net radiation at the crop surface [MJ m⁻² day⁻¹], G is the soil heat flux density [MJ m⁻² day⁻¹], T is the mean daily air temperature at 2 m height [°C], u_2 is the wind speed at 2 m height [m s⁻¹], e_s is the saturation vapor pressure [kPa], e_a is the actual vapor pressure [kPa], $e_s - e_a$ is the saturation vapor pressure deficit [kPa], Δ is the slope vapour pressure curve [kPa °C⁻¹], and γ is the psychrometric constant [kPa °C⁻¹].

2.3 LSSVM and Models Evaluation [15, 16]

Suppose that we are given a training data set of n data points $\{x_i, y_i\}_{i=1}^n$, where $x_i \in R^n$ is the i -th input vector and $y_i \in R$ is the corresponding i -th target. For

binary classification problems, y_i takes only two possible values $\{-1, +1\}$, whereas y_i takes any real value for regression problems. The goal is to find a function $V : x \rightarrow y$ that minimizes the residuals for the given data and generalizes well with unseen data. We transform the input patterns into the reproducing kernel Hilbert space (RKHS) by a set of mapping functions $\phi(x)$ [16]. An inner product in the feature space has an equivalent kernel in input space, $K(x, x') = \langle \phi(x), \phi(x') \rangle$, provided certain conditions hold. K is a symmetric positive definite function, which satisfies Mercer's conditions. To learn the unknown function, we solve a Tikhonov functional of the special form:

$$\begin{aligned} \min_{w,b,e} P(w, e) &= \frac{1}{2} w^T w + C \sum_{i=1}^n e_i^2 \\ \text{s.t.} \quad f(x) &= w^T \phi(x_i) + b + e_i \\ & \quad i = 1, 2, \dots, n \end{aligned} \quad (2)$$

where $C > 0$ is the regularization factor, e_i is the residual between the output y_i and $f(x_i)$, w is the weight vector, and b is called the bias term.

Using the Lagrangian multiplier method, the Lagrangian for (2) is:

$$L(w, b, e; \alpha) = P(w, e) - \sum_{i=1}^n \alpha_i \{w^T \phi(x_i) + b + e_i - y_i\} \quad (3)$$

where $\alpha_i, i = 1, 2, \dots, n$ are the Lagrangian multipliers corresponding to (2). The Karush-Kuhn-Tucker (KKT) conditions (3) are:

$$\left\{ \begin{aligned} \frac{\partial L}{\partial w} = 0 &\rightarrow w = \sum_{i=1}^n \alpha_i \phi(x_i), \\ \frac{\partial L}{\partial b} = 0 &\rightarrow \sum_{i=1}^n \alpha_i = 0, \\ \frac{\partial L}{\partial e_i} = 0 &\rightarrow \alpha = C e_i, \\ \frac{\partial L}{\partial \alpha_i} = 0 &\rightarrow w^T \phi(x_i) + b + e_i - y_i = 0, \\ & \quad i = 1, 2, \dots, n \end{aligned} \right. \quad (4)$$

In the numerical solution proposed by Suykens [12], the KKT conditions of (3) are reduced to a linear system by eliminating w and e , resulting in:

$$\begin{bmatrix} 0 & \vec{1}^T \\ \vec{1} & K(x_i, x_j) + C^{-1}I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (5)$$

where $y = [y_1, y_2, \dots, y_n]^T$, $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_n]^T$, $\vec{1} = [1, 1, \dots, 1]^T$. Note that $K(x_i, x_j) + C^{-1}I$ is symmetric and positive definite since the matrix K is semi-positive definite and the diagonal term C^{-1} is positive. Solving (5) for α and b , the discriminate function can be obtained from $f(x) = \sum_{i=1}^n \alpha K(x, x_i) + b$.

Four common choices of kernels are: $K(x, z) = x^T z$ (linear kernel), $K(x, z) = (r + x^T z)^d$ (polynomial kernel of degree d), $K(x, z) = e^{-\frac{\|x-z\|_2^2}{\sigma^2}}$ (RBF kernel), and $K(x, z) = \tanh(k_1 x^T z + k_2)$ (multilayer perceptron kernel).

The performance of the different models is evaluated based on the criteria of the root mean square error (RMSE) and square value of coefficient of correlation r . These two statistical parameters used for the performance evaluation are given as follows [11]:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_p - y_e)^2}{n}} \quad (6)$$

$$r = \frac{\sum_{i=1}^n (y_p - \bar{y}_p)(y_e - \bar{y}_e)}{\sqrt{\sum_{i=1}^n (y_p - \bar{y}_p)^2 \sum_{i=1}^n (y_e - \bar{y}_e)^2}} \quad (7)$$

where y_p and y_e represent PM method and temperature-based model ETo values estimated for the i -th values, \bar{y}_p and \bar{y}_e represent the average values of the corresponding variables, and n represents the number of observations.

2.4 Data Normalization

In this study, limited climatic data gathered from 1998 to 2009 were used. Daily average data of these 12 years in the 145th Regiment, making up 2568 data points,

were divided into three parts for the purposes of training (60%), validation (20%), and testing (20%). Note that only those data in the crop growth period (March–September) were used in this study. In order to overcome the negative result associated with extreme values, input and output data sets were scaled in the range [0, 1] using the following equation:

$$y_{norm} = \frac{y_i - y_{min}}{y_{max} - y_{min}} \tag{8}$$

where y_{norm} is the normalized dimensionless variable, y_i is the observed value of variable, and y_{min} and y_{max} are the minimum and maximum values of the observed variable.

3 Discussion of Results

For similar bases of comparison, the same training and verification sets were used for the ANN and LSSVM models. RMSE and r^2 were employed to evaluate the performances of the models developed. In this study, the ANN model architecture was 4–10–1, the Levenberg-Marquardt algorithm parameter was 0.002, and the LSSVM model kernel function used was RBF with parameters $(C, \epsilon, \sigma) = (50, 0.0001, 1.5)$.

A summary of the statistical performances of the different models is presented in Table 1. Compared with the ANN model’s consumption time (119.78 s), the LSSVM model shows better performance (35.265 s). Results from the statistical criteria indicate nearly the same fitting degree r^2 (0.986–0.987), but the LSSVM model performs better than ANN model in terms of RMSE.

To illustrate the model estimation performance, Figure 1 shows ANN and LSSVM ETo estimates with 642 nodes (March 2006–April 2009) configuration and PM ETo. The figure shows that both ANN- and LSSVM-estimated ETo values agree closely with the PM-estimated value and follow the same trend. LSSVM also has less deviation around the ETo peak compared with ANN, and shows better generalization capacity due to its Vapnik-Chervonenkis (VC) theory base.

Table 1. ANN and LSSVM models and their coefficients of correlation for the training and validation periods in the 145th regiment

Model	r2			RMSE			Consumption Time(s)
	Training	Validation	Testing	Training	Validation	Testing	
ANN	0.996	0.994	0.986	4.212	4.317	3.189	119.78
LSSVM	0.996	0.995	0.987	0.112	0.115	0.117	35.265

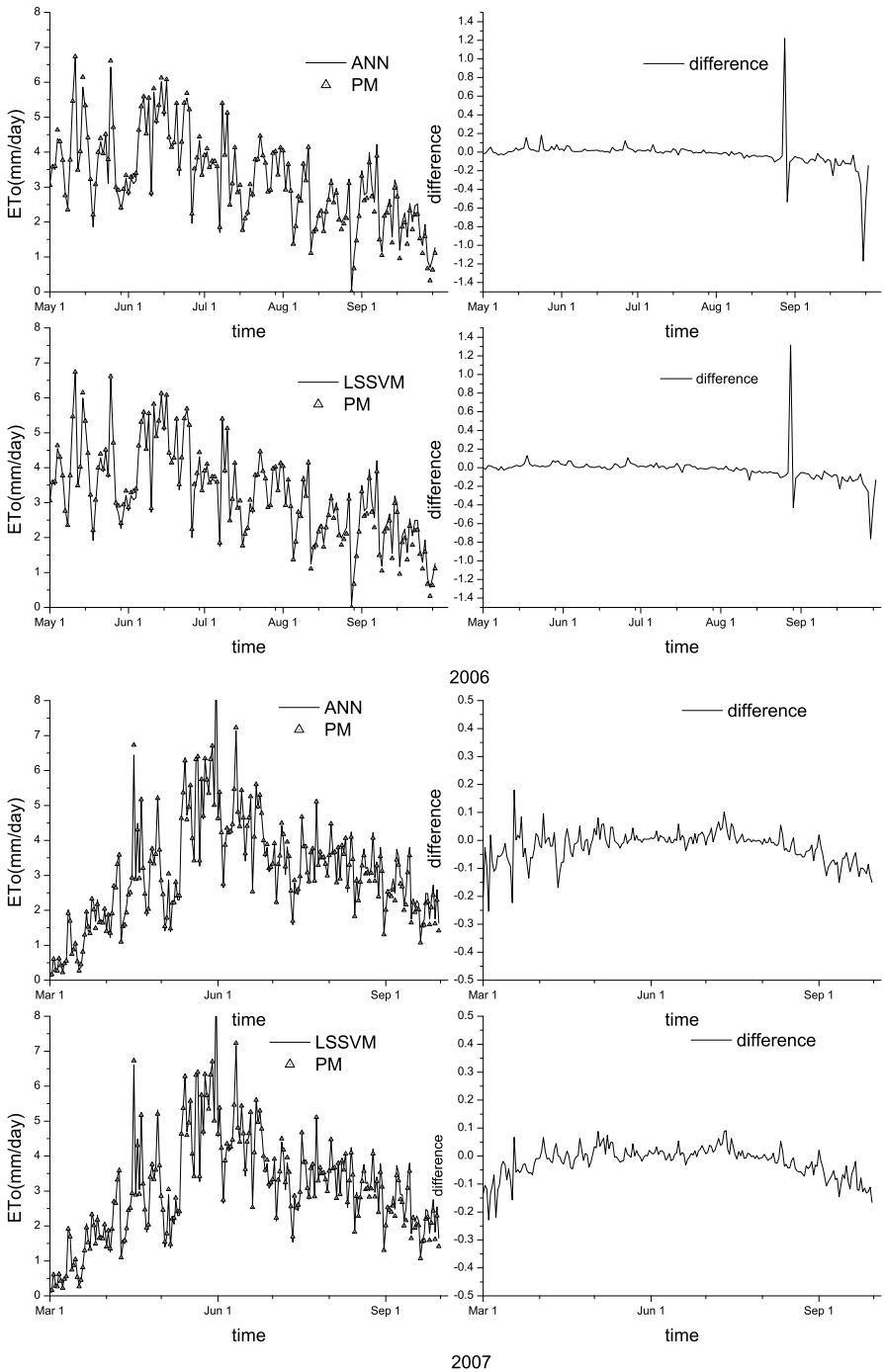


Fig. 1. Comparison of ETo values calculated by the proposed ANN, LSSVM, and PM methods

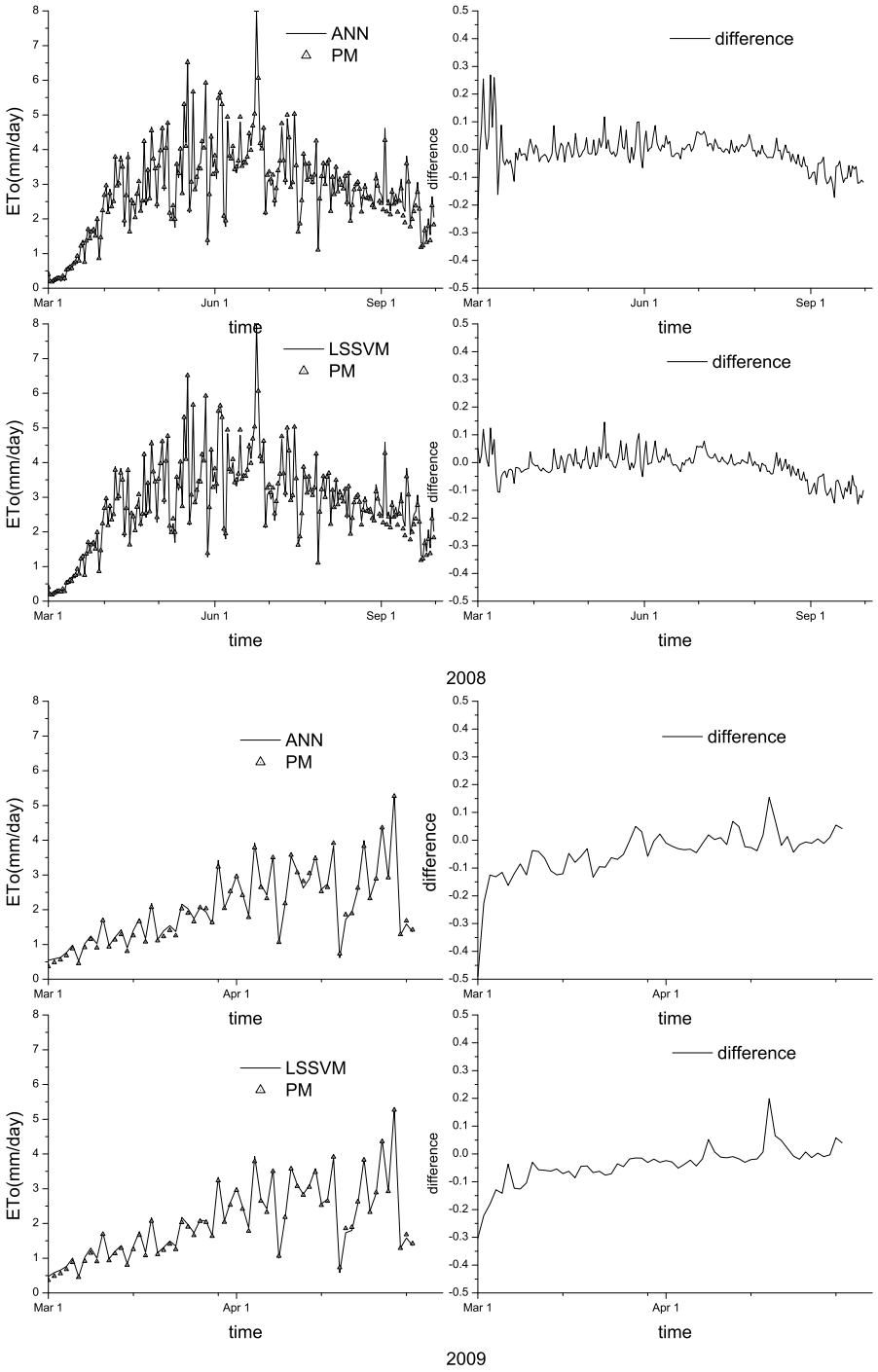


Fig. 1. (Continued)

Table 2. Comparison data in April 2009

Date	PM	ANN		LSSVM	
		Testing	Error	Testing	Error
2009-4-1	2.954	2.985	-0.010	3.025	-0.024
2009-4-2	2.413	2.465	-0.022	2.482	-0.029
2009-4-3	1.779	1.833	-0.030	1.870	-0.051
2009-4-4	3.783	3.912	-0.034	3.926	-0.038
2009-4-5	2.645	2.730	-0.032	2.704	-0.022
2009-4-6	2.320	2.425	-0.045	2.420	-0.043
2009-4-7	3.499	3.542	-0.012	3.560	-0.017
2009-4-8	1.058	1.038	0.019	1.004	0.051
2009-4-9	2.179	2.174	0.002	2.163	0.007
2009-4-10	3.572	3.537	0.010	3.612	-0.011
2009-4-11	3.070	3.118	-0.016	3.112	-0.014
2009-4-12	2.816	2.626	0.067	2.843	-0.009
2009-4-13	3.044	2.894	0.049	3.097	-0.017
2009-4-14	3.473	3.554	-0.023	3.578	-0.030
2009-4-15	2.520	2.586	-0.026	2.571	-0.020
2009-4-16	2.643	2.743	-0.038	2.690	-0.018
2009-4-17	3.913	3.840	0.019	3.884	0.007
2009-4-18	0.730	0.617	0.154	0.584	0.199
2009-4-19	1.854	1.722	0.071	1.733	0.065
2009-4-20	1.888	1.923	-0.019	1.796	0.048
2009-4-21	2.621	2.586	0.013	2.568	0.021
2009-4-22	3.826	3.991	-0.043	3.856	-0.008
2009-4-23	2.321	2.358	-0.016	2.366	-0.019
2009-4-24	2.885	2.904	-0.007	2.864	0.007
2009-4-25	4.359	4.407	-0.011	4.417	-0.013
2009-4-26	2.922	2.908	0.005	2.913	0.003
2009-4-27	5.267	5.327	-0.011	5.319	-0.010
2009-4-28	1.284	1.271	0.011	1.288	-0.003
2009-4-29	1.678	1.586	0.055	1.580	0.058
2009-4-30	1.420	1.360	0.042	1.363	0.040

4 Conclusions

In this work, LSSVM is proposed to be a novel technique for nonlinear function approximation. It is a very specific type of learning algorithm characterized by the capacity control of the decision function, the use of the kernel functions, and the sparseness of its solution. Established on the unique theory of structural risk minimization to estimate a function by minimizing the upper boundary of the generalization error, LSSVM is very resistant to the over-fitting problem, eventually achieving high generalization performance in solving various nonlinear function approximation problems. Another key property of LSSVM is that training it is equivalent to solving linear systems; its solution is always unique and globally optimal, unlike the training of other networks, which requires non-linear optimization with the danger of getting stuck with local minima.

An LSSVM model was developed to estimate ETo based on an extensive number of data sets with limited climatic information. Its superiority was demonstrated by comparison with the ANN model. All the simulation results showed that the LSSVM model is more effective and efficient in estimating ETo than the ANN model.

Although this paper showed the effectiveness of the LSSVM model, several issues require further investigation. Firstly, how to determine the hyperparameters of LSSVM is an important issue that needs to be addressed. Secondly, in this study, only the RBF kernel function was investigated. Additional research is necessary to explore more useful kernel functions to improve the performance of the LSSVM model.

References

1. Allen, R.G., Pereira, L.S., Raes, D., Smith, M.: Crop evapotranspiration: guidelines for computing crop water requirements. In: Proceedings of the Irrigation and Drainage Paper No. 56. Food and Agricultural Organization, United Nations, Rome (1998)
2. Kumar, M., Raghuvanshi, N.S., Singh, R., et al.: Estimating evapotranspiration using artificial neural network. *Journal of Irrigation and Drainage Engineering* 128, 224–233 (2002)
3. Trajkovic, S., Todorovic, B., Stankovic, M.: Forecasting of reference evapotranspiration by artificial neural networks. *Journal of Irrigation and Drainage Engineering* 129, 454–457 (2003)
4. Odhiambo, L.O., Yoder, R.E., Yoder, D.C., et al.: Optimization of fuzzy evapotranspiration model through neural training with input-output examples. *Trans. ASAE American Society of Agricultural and Biological Engineering* 44, 1625–1633 (2001)
5. Aksoy, H., Guven, A., Aytek, A., et al.: Discussion of Generalized regression neural networks for evapotranspiration modeling. *Hydrological Sciences Journal* 52, 825–831 (2007)
6. Kişi, O.: Generalized regression neural networks for evapotranspiration modeling. *Hydrological Sciences Journal* 51, 1092–1104 (2006)
7. Kişi, O., Ozturk, O.: Adaptive Neurofuzzy Computing Technique for Evapotranspiration Estimation. *Journal of Irrigation and Drainage Engineering* 133, 368–379 (2007)

8. Gorka, L., Amaia, O.B., Jose, J.L.: Comparison of artificial neural network models and empirical and semi-empirical equations for daily reference evapotranspiration estimation in the Basque Country (Northern Spain). *Agricultural Water Management* 95, 553–565 (2008)
9. Ali, R.K.: Comparative study of Hargreaves's and artificial neural network's methodologies in estimating reference evapotranspiration in a semiarid environment. *Irrigation Science* 26, 253–259 (2008)
10. Sungwon, K., Hung, S.K.: Neural networks and genetic algorithm approach for nonlinear evaporation and evapotranspiration modeling. *Journal of Hydrology* 351, 299–317 (2008)
11. Wang, Y.-M., Seydou, T., Tienfuan, K.: Neural network approach for estimating reference evapotranspiration from limited climatic data in Burkina Faso. *WSEAS Transactions on Computers* 7, 704–713 (2008)
12. Seydou, T., Wang, Y.-M., Tienfuan, K.: Artificial neural network for modeling reference evapotranspiration complex process in Sudano-Sahelian zone. *Agricultural Water Management* 97, 707–714 (2010)
13. Eslamian, S.S., Abedi-Koupai, J., Amiri, M.J., et al.: Estimation of daily reference evapotranspiration using support vector machines and artificial neural networks in greenhouse. *Res. J. Env. Sci.* 3, 439–447 (2009)
14. Kisi, O., Cimen, M.: Evapotranspiration modelling using support vector machines. *Hydrological Sciences Journal* 54, 918–928 (2009)
15. Vapnik, V.N.: *The Nature of Statistical Learning*. Springer, Berlin (1995)
16. Suykens, J.A.K., Vandewale, J.: Least squares support vector machine classifiers. *Neural Processing Letters* 9, 293–300 (1999)