http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 13 (9): 1656-1661, 2014
ISSN 1812-5638 / DOI: 10.3923/itj.2014.1656.1661

Tank Temperature and Humidity Prediction in Earthworm Treatment Using Support Vector Regression with Tuning-based on TLBO Optimization

[1]Yongming Cai, [1]Xuenong Zhang, [1]Lingyan Gu and [2]Lidan Fan
[1]College of Medical Information Engineering, GuangDong Pharmacy University, China
[2]Department of Computer Science, University of Texas at Dallas, USA

Abstract: By combining the use of support vector regression and predictive control techniques, we find that it is possible to control tank air temperature and humidity in earthworm treatment. In this study, we use Support Vector Regression (SVR) to predict temperature and humidity in the earthworm biological treatment reactor used to treat the municipal sludge, where the temperature and humidity belong to the nonlinear dynamic model system. Base on this model, we design a nonlinear model predictive controller. Furthermore, we propose an optimization algorithm to generate online control signals under the control constraints. To build and generalize an earth worm treatment model, the Teaching Learning Based Optimization (TLBO) technique is adopted to adjust the hyper-parameters of SVR. Experimental results show that the tuned SVR model by TLBO has good regression precision and is generalizable.

Key words: Teaching-learning based optimization, support vector machine, earthworms, sludge

INTRODUCTION

A large amount of wastewater and sludge have been produced in the urbanization (Suthar, 2009), the municipal wastewater treatment plants are used to handle them. However, it is not easy to process this huge volume of sludge which contains plenty of rich organic matter minerals, heavy metals and other toxic substances. In the past ten years, some researchers suggested that the earthworm can reduce the harmful matter in the sludge. They showed that earthworms can breakdown a wide range of organic materials, indicating that it can change the sludge into useful fertilizer. Therefore, the so called earthworm composting sludge treatment technology can be used to treat municipal sewage sludge, reducing the pollution.

Earth worms need proper living conditions including temperature and humidity. So, it is very important to control the temperature and humidity condition of earthworm treatment tank air. Many researchers are working on modeling auto control air condition. Since for the auto control, one of the most important things is to predict the air temperature and humidity. Thus, a lot of work focuses on temperature and humidity prediction. However, to get better result, all traditional methods including the artificial neural networks are based on large scale of sample data. In practice, it is not always easy to get large amount of sample data. To do prediction, Support Vector Machine (SVM) is a method that performs well on small scale of sample data. SVM bases on the theory of statistical learning. Therefore, the Support Vector Regression (SVR), which is one categoryof SVM, is adopted to model and predict earthworm treatment tank air temperature and humidity. The SVR is a reformulation of the standard Support Vector Machine (SVM), its implifies the standard SVM model in a great extent by applying linear least squares criteria to the loss function, which replaces traditional quadratic programming method. Similarly, SVR also has the character of simplicity and excellent generalization ability. SVR has been applied in many pattern recognition problems. When SVR is used to predict the temperature and humidity condition of earthworm treatment tank air, two parameters, the regularization parameter C and the kernel parameter σ^2, need to be determined. They impact the regression accuracy and generalization ability of SVR. To obtain excellent generalization ability of SVR, it is very important to choose appropriate values for the two parameters. How to choose the two parameters of SVR could be thought as an essential optimization task. This calls for the use of advanced meta-heuristic approaches, such as evolutionary method, population-based method, etc.

In recent years, Rao *et al.* (2012) proposed a new and efficient meta-heuristic optimization method, namely Teaching-Learning-Based Optimization (TLBO) algorithm and it is based on the philosophy of teaching and learning

Corresponding Author: Yongming Cai, College of Medical Information Engineering, GuangDong Pharmacy University, China

and is similar to other population-based optimization techniques, such as Particle Swarm Optimization (PSO), evolutionary optimization (DE), Artificial Bee Colony (ABC), Gravitational Search Algorithm (GSA) and Coupled Simulated Annealing (CSA) and so on. The TLBO is also a population-based optimization method and adopts a population of solutions to proceed to the global solution. Some researchers have applied the TLBO to some complex computational problems, such as data clustering, mechanical design, electrochemical discharge machining, design of planar steel frames and so forth. The results show that the performance of TLBO is better than existing best search optimization techniques including ABC, Genetic Algorithm (GA) and Grenade Explosion Method (GEM). Moreover, in this study we run experiment and the experimental results show that the TLBO could find better solutions and have much faster convergence speed. Additionally, to obtain a well generalized model of earthworm treatment tank air temperature and humidity, we use the TLBO to adjust the two parameters of SVR. The experimental results show that the tuned SVR by TLBO has well regression precision and generalization ability.

REVIEW OF RELATED WORKS

Teaching-learning-based optimization: The teaching-learning-based optimization algorithm is inspired by the phenomenon of the influence of a teacher on the output of learners in a class. A good teacher motivates the learners in a class and thereby helps in improving the average performance of all learners in the class. The reason is that every learner in a class is likely to follow their teacher and improves their own performance. Simultaneously, all learners in the class also try to interact with each other to improve their own performance. The whole procedure of the TLBO contains two phases, the Teacher phase and the Learner phase.

Teacher phase: In the teacher phase, the most knowledgeable person in the society is considered as a teacher. Learners would learn from the teacher and the teacher would bring the learners up to his or her level in terms of knowledge. The teaching role is assigned to the best individual ($X_{teacher}$). It means that the teacher would make efforts to move the mean of all the learners in a class up to his or her level depending on his or her ability. The algorithm attempts to improve individuals (X_i) by moving their positions towards the position of the $X_{teacher}$ by taking into account of the current mean of the individuals (X_{mean}). This is constructed by using the mean-value for each parameter within the problem space and represents the qualities of all learners in current generation. Equation 1 simulates how a learner's improvement is influenced by the difference between the teacher's knowledge and the qualities of all learners. For stochastic purpose, two randomly-generated parameters are introduced to the equation, where r is a random number between 0 and 1 and T_F is a teaching factor with the value being either 1 or 2.

$$X_{new} = X_i + r\ (X_{teacher} - (T_F . X_{mean})) \tag{1}$$

Learner phase: In learner phase, the learners could increase their knowledge by interactions among themselves. A learner can communicate randomly with other learners in order to improve his or her knowledge through group discussions, formal communications, presentations, etc. A learner learns something new from the learners who have more knowledge than him or her. The learning process can be described as follows Eq. 2 and 3.

For a learner X_i, randomly select another learner X_j where $i \neq j$:

$$X_{new} = X_i + r.\ (X_j - X_i),\ \text{If}\ X_j > X_i \tag{2}$$

$$X_{new} = X_i - r.\ (X_i - X_j),\ \text{If}\ X_j < X_i \tag{3}$$

Additionally, infeasible individuals must be appropriately handled to determine whether one individual is better than another, when applied to constrained optimization problem.

Support vector regression: Structural Risk Minimization (SRM) adopted by SVM could provide better generation ability than the empirical risk minimization used by traditional methods. While for to do prediction, it is proper to use the Support Vector Regression algorithms (SVR) (Smola *et al*., 1998), which is the deformation of SVM. The SVR are appealing algorithms for a large variety of regression problems (Mohandes *et al*., 2004; Chen *et al*., 2011), since they not only consider the error approximation to the data, but also the generalization of the model, i.e., their capability to improve the prediction of the model when a new dataset needs to be evaluated.

The SVR method consists of training a model with input as a set of samples $\{(x_1,y_1),(x_2,y_2),\ldots,(x_l,y_l)\} \subset R_n \times R$, where x_l is the lth input vector and y_l the corresponding objective value. The classifier parameters in the optimization problem of the SVR are found by resolving the following optimization problem, which expresses the maximization of the margin $2/\|w\|$ and the minimization of the training error (Eq. 4):

$$\min \frac{1}{2}\| w\|^2 + c\sum_{i=1}^{m} l(y_i, f(\varnothing(x_i)))$$
$$\text{s.t. } y_i\ (w.\varnothing(x_i)+b) \geq 1 - \xi_i,\ \xi_i \geq 0,\ i = 1,\dots, l \quad (4)$$

where, w controls the smoothness of the model, $\varnothing(x_i)$ is a function of projection of the input space to the feature space, b is aparameter of bias, x_i is a feature vector of the input space with dimension N, y_i is the output value to be estimated and $l(y_i, f(x))$ is the loss function selected. ξ_i is the slack variable. The dual form of this optimization problem is usually obtained through the minimization of the Lagterm function, constructed from the objective function and the problem constraints. In this case, the dual form of the optimization problem is as follows (Eq. 5):

$$\max\left(-\frac{1}{2}\sum_{i,j=1}^{l}(a_i - a_i^*)K(x_i, x_j) - \varepsilon\sum_{i=1}^{l}(a_i + a_i^*) + \sum_{i=1}^{l} y_i(a_i + a_i^*)\right) \quad (5)$$
$$\text{Subject to} \sum_{i=1}^{l} y_i(a_i - a_i^*) = 0 \quad a_i, a_i^* \in [0, C]$$

In addition to these constraints, the Karush-Kuhn-Tucker conditions must be fulfilled and also the bias variable b, must be obtained. In the dual formulation of the problem, function $K(x_i, x_j)$ is the kernel matrix, which is formed by the evaluation of a kernel function, equivalent to the dotproduct $\varnothing(x_i)$, $\varnothing(x_j)$. A usual selection for this kernel functionis a Gaussian function, being described as follows (Eq. 6):

$$K(x_i, x_j) = \exp\left(-\frac{\| x_i - x_j\|^2}{2\sigma^2}\right) \quad (6)$$

where, γ is the regularization parameter, the final form of the function f (x) depends on the Lagrange multipliers a_i, a_j^* and the decision function of the SVR is as follows (Eq. 7):

$$f(x) = \sum_{i=1}^{l}(a_i - a_i^*)K(x_i, x) \quad (7)$$

MATERIALS AND METHODS

Digital earthworm treatment tank: The digital earthworm treatment tank is a cuboid whose size is $2\times1\times0.6$ m^3. There are intake and exhaust ports on the top of the equipment. The air from the exhaust port flush through the three-way valve being cooled by the radiator and come into the tank for recycle. The fresh and warmed mixed air will come into the tank through the intake pipe. The switch of the cool and warm cycle fan is controlled by the auto control system which will keep the air temperature and humidity of the tank stable. The sketch map of the earthworm treatment tank air temperature and humidity control system is show as Fig. 1.

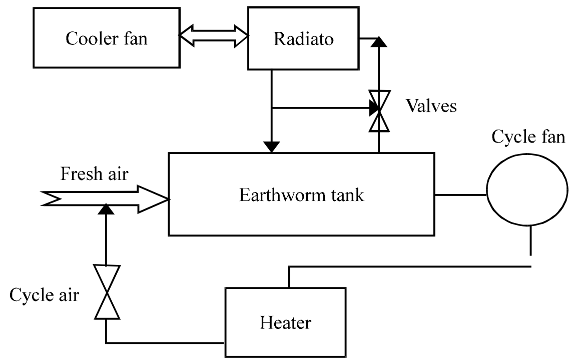

Fig. 1: Earthworm treatment tank control system

The water temperature in radiators will continue fluctuate because of the shift of the ON-OFF switch. The air in the earthworm treatment tank is forced to flow by the three-way valves. Heat energy exchanges when the air flows through the radiator. The three-way valves can change the air flow speed and direction. When air temperature and humidity of the earthworm treatment tank is lower than the threshold, it needs to reduce the cool water and control the fan to make sure that more air comes into the heater.

The max air flow speed of the fan in our experiment is up to 18.3 m^3 every minute. But in order to protect the fan, the speed is limited to 75% of the max calibration. The fan speed and the radiator threshold volume are system control variables. There are six cycle fans in total for the good air cycle, of which four fans' rated power is 0.63 kW and the other two is 0.75 kW. The heat send out by the six working fans can be the energy source of the heater. The frequency of sampling the six variables is once every 60 sec. The sample characters include air temperature y_1, air humidity y_2, fresh air temperature w_1, humidity w_2, fan speed u_1 and radiator u_2.

Earthworm treatment tank air temperature and humidity model: There is coupling phenomenon during the process of controlling earthworm treatment temperature and humidity. If the air temperature inside the tank is lower than the threshold, then it needs to heat the air. During the process of increasing the air temperature, the condensate will be evaporated and which in turn increases the air humidity. If the air temperature is too high, then it needs to be cooled down and which subsequently can lead the water in the air to condense and the humidity to become lower. The coupling degree will change according to the tank volume size.

There needs change for the compressor parameters according to the earthworm treatment tank size. Accordingly, some parameters of the air temperature and

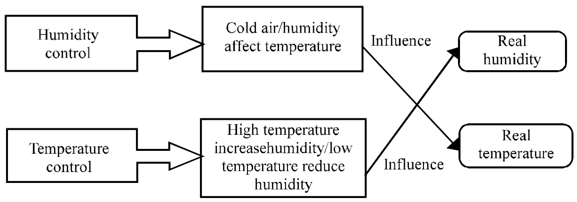

Fig. 2: Temperature and humidity model

humidity control model need to be changed too. Under the other situation, the earthworm treatment tank should change the air temperature and humidity according to the living condition of the earthworm. The earthworm treatment tank air temperature and humidity model is showed as Fig. 2.

By analyzing the mathematical model of the earthworm treatment tank air temperature and humidity, it is found to have an obvious property called time delay. Therefore, it is a complicate problem to accurately control the air temperature and humidity which have serious time delay strong coupling. The air temperature and humidity change with time can be fixed according to the character of earthworm treatment dynamic control system.

The k+1th situation ofthe system output can be calculated according to the kth, k-1th and k-2th situations. The function is as follows (Eq. 8 and 9):

$$y_1(k+1) = F_1(y_1(k), y_2(k), w_1(k), w_2(k), u_1(k), u_2(k), y_1(k-1), y_2(k-1), w_1(k-1), w_2(k-1), u_1(k-1), u_2(k-1), y_1(k-2), y_2(k-2), w_1(k-2), w_2(k-2), u_1(k-2), u_2(k-2)) \quad (8)$$

$$y_2(k+1) = F_2(y_1(k), y_2(k), w_1(k), w_2(k), u_1(k), u_2(k), y_1(k-1), y_2(k-1), w_1(k-1), w_2(k-1), u_1(k-1), u_2(k-1), y_1(k-2), y_2(k-2), w_1(k-2), w_2(k-2), u_1(k-2), u_2(k-2)) \quad (9)$$

Functions F_1 and F_2 are nonlinear functions of the support vector regression model. So, the earthworm treatment tank air temperature and humidity prediction function can be supposed as Eq. 11:

$$y_i = \sum_{i=1}^{mj} a_{ji} k(v_i, v) + b_j, j = 1, 2 \quad (10)$$

In Eq. 10, K is a kernel function. V_i, $i = 1, \ldots, m_j$ is the support vector of SVM model. m_j is the number of support vectors. a_{ji} and b_j are constants which can be obtained through training. Two SVM models need to be designed for the two output variables. Every output variable has its own SVM model, support vectors and constant coefficients. Vector v and v_i are the combination status variables y_1 and y_2 of sampling times k, k-1 and k-2th, input variables u_1, u_2 and w_1, w_2. Total 18 input variables determine the values of function F_1 and F_2 in Eq. 9 and 10.

Suppose using the gauss kernel function, Eq. 10 can be written as Eq. 11:

$$y_i = \sum_{i=1}^{mj} a_{ji} \exp\left(-\frac{\| v - v_i \|}{2\sigma_j^2}\right) + b_j, j = 1, 2 \quad (11)$$

It can be seen that functions F_1 and F_2 share the same input variable v from Eq. 8 and 9. The input vector has 18 characters. Every character will be used in the SVM kernel function. It is noticed that the different characters have their own value ranges. Infunction F_1, the first parameter (temperature) is in the range 17-30°C and the second parameter (humidity) is in the range 40-99%. Every parameter can affect the kernel function by the square error value. If the absolute values of some of the elements are bigger than others, they would be the mayor elements that affect the kernel function in the end. So the raw data needs to be preprocessed before put into the SVM model. All the characteristic values and target values are regulated to [-1,1]. Two parameters of the kernel function under the SVM model need to be initially set. For simplicity, the combined parameter $g = 1/\sigma^2$ is used to replace. The K-fold cross verification is used to verify the parameters of kernel function.

Nonlinear model is presented to predict the effect of a sequence of actions on the control target. Predicting the control target is to optimize the detail index and output the preset target value. As for the earthworm treatment tank air temperature and humidity control model, the objective function and constraint condition can be expressed as follows (Eq. 12):

$$\min_u \sum_{k=0}^{p-1} \{\sum_{i=1}^{2} \gamma i \left(\frac{y_i(k) - r_i}{\bar{y}_i - r_i}\right)^2 + \theta i u_i^2(k) + \phi_i \Delta u_i^2(k)\} + \sum_{i=1}^{2} \eta_i \left(\frac{y_i(p) - r_i}{\bar{y}_i - r_i}\right)^2$$

$$\text{subject to} \quad \begin{cases} y_i(k+1) = \sum_{j=1}^{N_i} a_{ij} \exp(-g_i \| v_j(k) - v_{ij}^* \|) + b_i, i = 1, 2 \\ 0.1 \le u_1 \le 0.75 \\ 0.1 \le u_2 \le 1 \end{cases} \quad (12)$$

In the above equations, y_i, θ_i, ϕ_i and ϕ_i are adjustment coefficients, which aim to adjust the importance of each factor. The different importance level of these factors will significantly influence the control effect.

TUNING SVR BASED ON TLBO TO MODEL EARTHWORM TREATMENT TANK

In this section, we adopt the tuned SVR by TLBO to model earthworm treatment tank air temperature and humidity. In order to demonstrate the validity of TLBO, another three optimization methods: GSA, ABC and PSO are employed to adjust the two parameters of SVR. Firstly, the earthworm treatment tank air temperature and humidity data and the parameters of optimization algorithms are given. Then the experimental results are analyzed in detail.

Earthworm composting sludge materials: We use the bright-colored Eiseniafetida which has deep longitudinal stripes between segments and likes warm and humid conditions. Healthy adult earthworms aged above 3 months, in about 350 mg and with obvious clitellum were selected. Before experiments, they were tamed in an artificial climate box for 7 days.

Artificial soil was prepared according to the standard OECD guideline and by simulating the characteristics of natural soil (OECD, 1984). The components were blended in proportion and fully mixed by adding deionized water. The artificial soil was adjusted to pH 7.0±0.2. The tanks had air holes on the bottom; a layer of cotton gauze was laid on an internal bottom. Dewatered sludge containing 66% water was added to each basin. The earthworm treating tanks air condition targets are set at 20±2°C and humidity 70±2%. After 28 days when the sludge was processed fully by the earthworms, these earthworms were collected.

We adopt the tuned SVR by optimization algorithms to build the mathematical model of earthworm treatment tank air temperature and humidity, which could state the mapping relation between air temperature and humidity and the operational conditions of the earthworm treatment tank. Here, the earthworm treatment tank air temperature and humidity would be taken as the output of the SVR model and the above the parameters are chosen as inputs of the SVR model.

Tuning parameters of SVR: In SVR, the two major hyper-parameters: regularization parameter C and kernel parameter σ^2 would directly affect the generalization ability and regression accuracy. So, the selection of appropriate hyper-parameters is a crucial step in obtaining well-adjusted SVR. In this work, the TLBO algorithm is adopted to tune the hyper-parameters of SVR. During the selection of hyper-parameters, each learner represents a feasible solution consisted of a vector $<C, \sigma^2>$. In addition, there is still another important factor for tuning hyper-parameters of SVR by TLBO. That is how to define the fitness function in order to estimate the marks of learners. Here, the fitness function is defined as leave-one-out cross-validation.

The objective is to minimize the fitness value, so the learner with the minimal fitness value will outperform others and should be reserved during the optimization process. Accordingly, the optimal hyper-parameters can be selected.

As TLBO algorithm is an algorithm specific parameter-less algorithm, only population size and number of generations need to be specified to run the algorithm. We set the population size at 20 and the maximum iteration 50. There are another three optimization methods including GSA, ABC, PSO, which are also employed to adjust the hyper-parameters of SVR.

RESULTS AND DISCUSSION

In this subsection, the predicted earthworm treatment tank air temperature and humidity by SVR are presented. The emphasis is paid on the generalization ability and regression accuracy, especially on the generalization ability.

As seen from Fig. 3, the predicted temperature correlates swith the real data. It also can be seen from Fig. 4, the predicted humidity runs well with the real data too. Because the earthworm living condition need constant temperature at 20±2°C, it is important to adjust the air temperature to fix it. The experiment begins with the initial conditions of y_1 = 28.1°C and y_2 = 56%. The initial

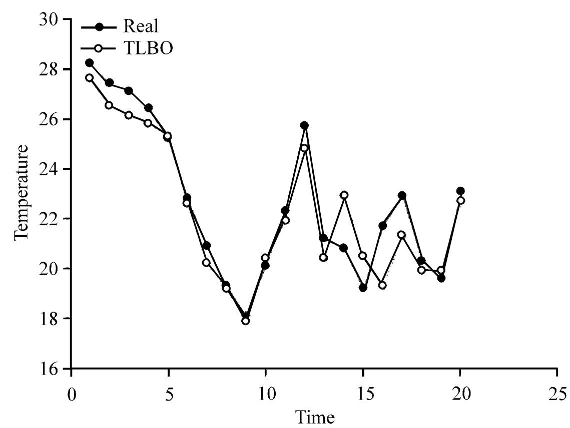

Fig. 3: Measured vs. predicted values of temperature by TLBO training set

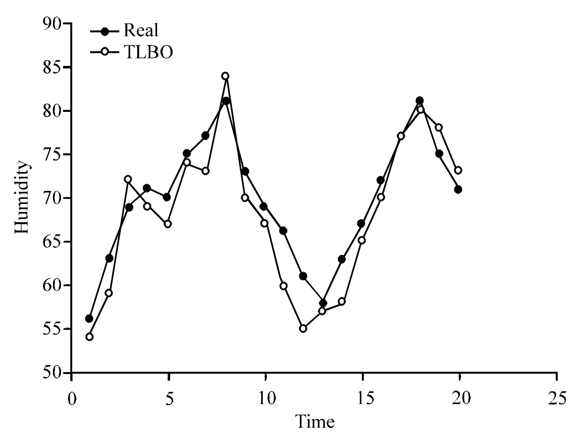

Fig. 4: Measured vs. predicted values of humidity by TLBO training set

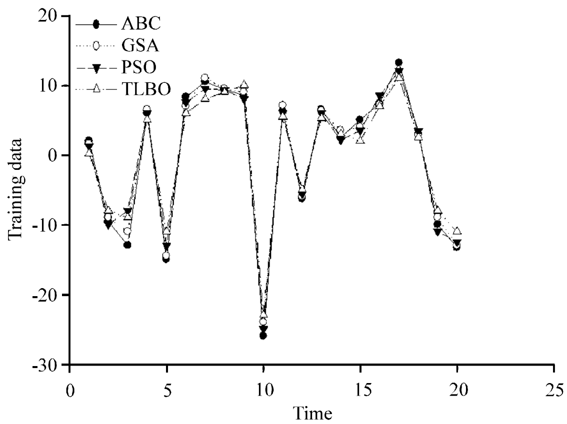

Fig. 5: Forecast errors of tuned SVR model by four methods for training set

Table 1: Values of tuning parameters

γ_1	γ_2	θ_1	θ_2	$\varnothing_1$	$\varnothing_2$	η_1	η_2
2	1	0	0	$0.6/0.75^2$	0.6	4	4

set points are y_1 = 20°C and y_2 = 70%. These are changed to y_1 = 21°C and y_2 = 72% at the 6th recording time point and to y_1 = 22°C and y_2 = 55% at the 20th recording point. The tuning parameters used are listed in Table 1.

It can be included from Fig. 5 that the tuned SVR model by TLBO has best regression accuracy, whose Mean Square Error (MSE) is the least one. So, the hyper-parameters found by using TLBO are the best ones for the training set and the regression accuracies of the tuned SVR model by the five methods are very well. So, we could see that the training effect is very good for the four optimization methods.

The training data error is considered as criteria in this experiment. Error is the average difference between the obtained best solution and the global solution, which indicates the ability of the algorithm in reaching the global optimum solution. The population size of the optimization algorithms is the same as the training and the variable interval is [0.001, 1000].

In sum, the tuned model by TLBO shows better identification and generalization abilities under various operating conditions than the ones by the other three methods. Therefore, it could be adopted by modeling earthworm treatment tank air temperature and humidity. It is verified by the experimental results that the tuned model by TLBO is very convenient, direct and accurate and can give a general and suitable way to predict the earthworm treatment tank air temperature and humidity under various operating conditions.

CONCLUSIONS

To get a well-generalized and identification model, the TLBO is employed to adjust the hyper-parameters of SVR which is used to model earthworm treatment tank air temperature and humidity. Simulation results show that the TLBO has very good regression precision and better generalization ability. TLBO algorithm performs correctly in all cases. It can be easily customized to suit the optimization of other types of thermal systems involving large number of variables and objectives. These features boost up the applicability of the proposed algorithm for the thermal systems optimization. Therefore, the TLBO can be applied to optimization problems in various fields.

ACKNOWLEDGMENTS

The authors acknowledge financial support from the Guangdong Science and Technology Plan Foundation (2012B031000018), the Natural Science Foundation of Guang Dong China (S2011010002511) is also acknowledged.

REFERENCES

Chen, J.L., H.B. Liu, W. Wu and D.T. Xie, 2011. Estimation of monthly solar radiation from measured temperatures using support vector machines-a case study. Renewable Energy, 36: 413-420.

Mohandes, M.A., T.O. Halawani, S. Rehman and A.A. Hussain, 2004. Support vector machines for wind speed prediction. Renewable Energy, 29: 939-947.

OECD, 1984. Guideline for testing of chemicals. Avian Reproduction Test, in Section 4, Number 207, Adopted on April 4, 1984.

Rao, R.V., V.J. Savsani and D.P. Vakharia, 2012. Teaching-learning-based optimization: An optimization method for continuous non-linear large scale problems. Inform. Sci., 183: 1-15.

Smola, A.J., N. Murata, B. Scholkopf and K.R. Muller, 1998. Asymptotically optimal choice of ε-loss for support vector machines. Proceedings of the 8th International Conference on Artificial Neural Networks, September 2-4, 1998, Skovde, Sweden, pp: 105-110.

Suthar, S., 2009. Vermistabilization of municipal sewage sludge amended with sugarcane trash using epigeic *Eisenia fetida* (Oligochaeta). J. Hazard. Mater., 163: 199-206.