Available at www.sciencedirect.com

INFORMATION PROCESSING IN AGRICULTURE 5 (2018) 253–268

journal homepage: www.elsevier.com/locate/inpa

Applied machine learning in greenhouse simulation; new application and analysis

Morteza Taki [a,*], *Saman Abdanan Mehdizadeh* [a], *Abbas Rohani* [b], *Majid Rahnama* [a], *Mostafa Rahmati-Joneidabad* [c]

[a] *Department of Agricultural Machinery and Mechanization, Ramin Agriculture and Natural Resources University of Khuzestan, Mollasani, Iran*

[b] *Department of Biosystems Engineering, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran*

[c] *Department of Horticultural Science, Faculty of Agriculture, Ramin Agriculture and Natural Resources University of Khuzestan, Mollasani, Iran*

ARTICLE INFO

Article history:
Received 3 November 2017
Received in revised form
15 January 2018
Accepted 16 January 2018
Available online 2 February 2018

Keywords:
Black box method
Energy lost
Environmental situation
Energy exchange

ABSTRACT

Prediction the inside environment variables in greenhouses is very important because they play a vital role in greenhouse cultivation and energy lost especially in cold and hot regions. The greenhouse environment is an uncertain nonlinear system which classical modeling methods have some problems to solve it. So the main goal of this study is to select the best method between Artificial Neural Network (ANN) and Support Vector Machine (SVM) to estimate three different variables include inside air, soil and plant temperatures (T_a, T_s, T_p) and also energy exchange in a polyethylene greenhouse in Shahreza city, Isfahan province, Iran. The environmental factors which influencing all the inside temperatures such as outside air temperature, wind speed and outside solar radiation were collected as data samples. In this research, 13 different training algorithms were used for ANN models (MLP-RBF). Based on K-fold cross validation and Randomized Complete Block (RCB) methodology, the best model was selected. The results showed that the type of training algorithm and kernel function are very important factors in ANN (RBF and MLP) and SVM models performance, respectively. Comparing RBF, MLP and SVM models showed that the performance of RBF to predict T_a, T_p and T_s variables is better according to small values of RMSE and MAPE and large value of R^2 indices. The range of RMSE and MAPE factors for RBF model to predict T_a, T_p and T_s were between 0.07 and 0.12 °C and 0.28–0.50%, respectively. Generalizability and stability of the RBF model with 5-*fold* cross validation analysis showed that this method can use with small size of data groups. The performance of best model (RBF) to estimate the energy lost and exchange in the greenhouse with heat transfer models showed that this method can estimate the real data in greenhouse and then predict the energy lost and exchange with high accuracy.

* Corresponding author.
E-mail addresses: mortezataaki@gmail.com, mtaki@ramin.ac.ir (M. Taki).

Peer review under responsibility of China Agricultural University.
https://doi.org/10.1016/j.inpa.2018.01.003

Nomenclature

Q	Heat transfer (W)
α	Heat transfer coefficient (W/m^2 K)
v	Wind speed (m/s)
I	Solar radiation (W/m^2)
T	Temperature (K)
F	View factor (-)
E	Emission coefficient (-)
A	Surface area (m^2)
c_p	Specific heat capacity (J/kg K)
d	Thickness (m)
V	Volume (m^3)
R_{b-heat}	Leaf boundary layer resistance (s/m)
l_f	Mean leaf width (m)
RBF	Radial Bias Function
SVM	Support Vector Machine
MLP	Multilayer Perceptron
BP	Back Propagation

Subscripts

$a-p$	Inside air to plant
$s-ss$	Upper to lower soil
$s-p$	Soil to plant
$ri-p$	Inside roof to plant
a	Inside air
s	Inside soil
ri	Inside roof
$s-ri$	Soil to inside roof
ss	Lower layer of inside soil
p	Plant
$a-ri$	Inside air too inside roof
$a-s$	Inside air to soil

Greek symbols

ρ	Density (kg/m^3)
σ	Stefan-Boltzmann constant (W/m^2 K^4)
λ_s	Soil thermal conductivity (W/m K)

1. Introduction

Agricultural greenhouses are source-conserving, socially supportive, commercially competitive, environmentally sound and rely on planting techniques, equipment management and constructive materials. The aim of these structures are to decrease agro-chemicals, energy and water consumption as well as waste production [1]. These goals can be obtained by some methods such as: efficient management of climatic variables, use of renewable and sustainable energy sources instead of fossil fuels, apply the suitable greenhouse covering materials and perfect physical properties with low generation of after-use waste and optimization of all applied materials for the plants. The level of inside environment control changes from the basic simple type of greenhouse to the fully closed conditioned greenhouse [2].

Typical conventional greenhouses in Iran have a plastic film covering the slanted front roof without using any blanket or thermal screen during the night. In winter, outside solar radiation is low and the air temperature even falling to below −20C°. The growers in such greenhouses must provide suitable environment conditions for plant growth and try to maintain good and economic temperature inside the greenhouse. On the other hand, facing temperature stresses (high or low) may lead to dead or burst the diseases or fungal in plants. These maintain should satisfy the grower and make a good market. Therefore, having an accurate predictive model of temperature in such greenhouses to inform the farmers about future conditions to reduce financial losses has a great importance [3].

Many modeling approaches have been utilized to help the farmers to have a good condition in greenhouse; such as mechanism, transfer function and black-box modeling. The mechanism model provides a clear physical explanation of the greenhouse environment, like the early static and dynamic model presented by Bot [4] or improved models presented by Van Henten [5] and De Zwart [6]. Static and dynamic models are used for this purpose as a function of the metrological conditions and the parameters of the greenhouse components [7,8]. The transfer function model has a simple structure [9,10], but its application is limited to linear systems. The black-box model is based on input and output data and is suitable for both linear and nonlinear modeling [11–13]. So researches try to develop these models and apply them in greenhouses to satisfy the farmers and to reduce the final cost [14–16]. He and Ma [17] applied BP neural network with PCA for modeling the internal greenhouse humidity in China. The results showed that stepwise regression method was less accurate than the BPNN based on PCA. Taki et al. [18], presented a paper to compare some mathematical models (include dynamic and Multiple Linear Regression (MLR)) with innovative method (Artificial Neural Network) and select the best one to predict inside air and roof and energy lost in a semi-solar greenhouse in Iran. Results showed that the performance of MLP model was better based on the small RMSE and MAPE and large value of EF parameters.

In the past few decades, many new training algorithms have been proposed to overcome the drawbacks of traditional neural networks and to increase their reliability. In this paradigm, one of the significant developments is a class of kernel based neural networks called Support Vector Machines (SVMs), the principle of which is rooted in the statistical learning theory and method of structural risk minimization [19]. Support Vector Machines (SVMs) have found wide application in several areas including pattern recognition, regression, multimedia, bio-informatics and artificial intelligence [20–21]. Yu et al. [2], presented a novel temperature prediction model based on a least squares support vector machine

(LS-SVM) model with parameters optimized by improved particle swarm optimization (IPSO). The results showed that IPSO can predict the maximum and minimum temperature with more accuracy than BPNN.

Because of the huge fossil energy sources in Iran, most greenhouses use high level of energy in all seasons. Also, there is no precise control on the inside environment factors of greenhouses in Iran. Therefore, this research was conducted for the first time in a conventional greenhouse in Iran, and its main objective is the feasibility of using artificial intelligence (MLP, RBF and SVM models with *k*-fold cross-validation) to control climate conditions as well as energy consumption. Based on the above comprehensive literature, there is no research in Iran on this subject. Also in this research, Randomized Complete Block (RCB) methodology was used as a new idea for analysis the type of training algorithm and kernel function for ANN and SVM models, respectively. Similar to this type of analysis, has ever been seen in researches.

2. Methodology

2.1. Polyethylene greenhouse structure

In this study, a conventional greenhouse with 5 spans and 1200 m^3 was selected. The orientation of this greenhouse is East–West (Fig. 1). Experiments were down in the greenhouse which cucumber was grown. The greenhouse was only heated overnight, using gas heating systems and so we could have the total energy consumption.

2.2. Artificial neural network

Prior to any ANN training process with the trend free data, the data must be normalized over the range of [0, 1]. This is necessary for the neurons' transfer functions, because a sigmoid function is calculated and consequently these can only be performed over a limited range of values. If the data used with an ANN are not scaled to an appropriate range, the network will not converge on training or it will not produce meaningful results. The most commonly employed method of normalization involves mapping the data linearly over a specified range, whereby each value of a variable x is transformed as follows [22]:

$$x_n = \frac{x - x_{min}}{x_{max} - x_{min}} \times (r_{max} - r_{min}) + r_{min} \quad (1)$$

where x is the original data, x_n the normalized input or output values, x_{max} and x_{min} are the maximum and minimum values of the concerned variable, respectively. r_{max} and r_{min} an correspond to the desired values of the transformed variable range. A range of 0.1–0.9 is appropriate for the transformation of the variable onto the sensitive range of the sigmoid transfer function.

Fig. 1 – Type of selected greenhouse at Shahreza city, Isfahan province.

2.2.1. Multilayer Perceptron (MLP) algorithm

MLP is a feed-forward layered network with one input layer, one output layer, and some hidden layers. Every node computes a weighted sum of its inputs and passes the sum through a soft nonlinearity. The soft nonlinearity or activity function of neurons should be non-decreasing and differentiable. The most popular function is unipolar sigmoid [22]:

$$f(\theta) = \frac{1}{1 + e^{-\theta}} \quad (2)$$

The network is in charge of vector mapping, i.e. by inserting the input vector, x^q the network will answer through the vector z^q in its output (for $q = 1, 2, \ldots, Q$). The aim is to adapt the parameters of the network in order to bring the actual output z^q close to corresponding desired output d^q (for $q = 1, 2, \ldots, Q$). The most popular method of MLP training is the Back-Propagation (BP) algorithm, and in literatures there exist many variants of this algorithm. This algorithm is based on minimization of a suitable error cost function [23].

In this study, Basic Back-propagation (BB) algorithm was employed. In this work, the learning rules of Gradient Descent Momentum (GDM) and Levenberg-Marquardt (LM) were considered. No transfer function for the first layer was used. For the hidden layers the sigmoid functions were used, and for the output layer a linear transfer function was applied as desired for estimating problems. We used an N-fold cross validation method that in this method data are randomly divided into two sets; training set and cross validation set [23]. Schematic diagram of MLP algorithm are illustrate in.

2.2.2. Radial basis function (RBF) algorithm

One type of ANN is the radial basis function (RBF) neural network which uses radial basis functions as activation functions. This ANN is a linear combination of radial basis functions. Radial Basis Functions (RBF) networks form a special architecture of neural networks that present important advantages compared to other neural network types, including simpler structure and faster learning algorithms [24]. RBF is a feed-forward neural network model with good performance and it has already proved its universal approximation ability with no local minima problem [24]. An RBF has a single hidden layer. Each node of the hidden layer has a parameter vector called center. This center is used to compare with the network input vector to produce a radially symmetrical response. Responses of the hidden layer are scaled by the connection weights of the output layer and then combined

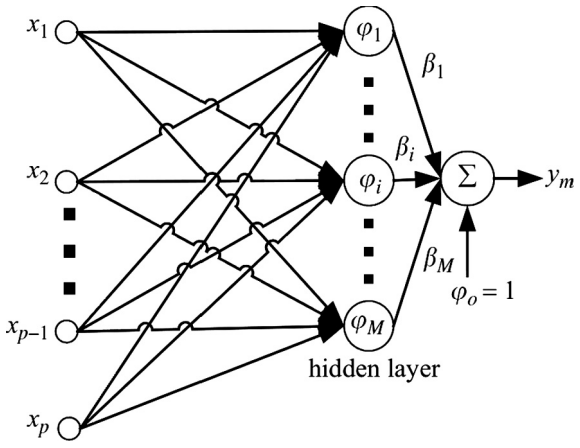

Fig. 2 – A typical RBF network configuration [24].

to produce the network output. A typical RBF network configuration is shown in Fig. 2 for a single output, where the outputs of the nonlinear activation are combined linearly with the weight vector β of the output layer to produce the network output y_m:

$$y_m = \sum_{i=0}^{M} \beta_i \varphi_i \tag{3}$$

In which β_i is the joint weighted value of the *i*th basis function. The most commonly used radial base is the Gaussian function given as:

$$\varphi_i(x) = \exp\left(-\frac{\|x - c_i\|^2}{\sigma_i^2}\right) \tag{4}$$

where c_i and σ_i are center and spread of the *i*th RBF node.

Training the radial basis function neural network is done in two steps. In the first step centers are selected from the training data (without training) or constructed by clustering the training data. The second step is basically a linear estimation of one weighting vector using ordinary least squares. RBF is an interpolating network. It can be built using all the available training points, or it can be built using reduced number of points. The selection of centers can then be performed by clustering the training data. There are three types of learning strategies used in selecting RBF centers; fixed randomly selected centers, self-organized center selection, and supervised selection of centers. A number of training algorithms has been developed for training of RBF networks [25]. Orthogonal least square (OLS) techniques are self-organized technique, and have been employed to select centers so that adequate and efficient RBF network can be obtained. The OLS uses gram-Schmidt algorithm for center selection and center updating of RBF network, while adaptive gradient descent procedure, described in Haykin [19], was used to adapt the weights. The network parameters are found such that they minimize a cost function:

$$\min J = \sum_{i=1}^{Q} (|y_{mi} - y_{di}|^2) \tag{5}$$

where Q is the number of training pattern, while y_m and y_d are the network output and desire target output, respectively. This type of learning strategy has been adopted in this work.

In this research, for MLP and RBF models, some of different training algorithms were applied and finally based on the statistical analysis, the best one was selected. Table 1 shows the list of applied training algorithms for ANN models.

2.3. *Support vector Machine (SVM)*

The Support Vector Machine (SVM), which has been used in this study, provides a computational advantage over standard SVM by converting quadratic optimization problem into a system of linear equations (Fig. 3). SVM intelligent approach considers the problem of approximating a given dataset $\{(x_1, y_1), (x_2, y_2, \ldots, (x_n, y_n)\}$ with the following nonlinear function [26]:

$$f(x) = \langle w, \Phi(x) \rangle + b \tag{6}$$

where $\langle .., . \rangle$ is dot product, $\Phi(x)$ is the nonlinear function that performs regression, b and w are bias terms and weight vector, respectively.

The optimization problem of SVM is formulated as [27]:

$$\left\{ \begin{array}{l} \min_{w,b,e} J(w,e) = \frac{1}{2}\|w\|^2 + \frac{1}{2}\gamma \sum_{k=1}^{N} e_k^2 \\ s.t. y_k = \langle w, \Phi(x_k) \rangle + b + e_k, \;\; k = 1, 2, ..., k \end{array} \right\} \tag{7}$$

where γ is a regularization parameter (also called penalty parameter, $\gamma \geqslant 0$) and e_k is the regression error for N training objects. To solve the optimization problem, Lagrange function constructed as follows [28]:

$$L(w, b, e, \alpha) = \frac{1}{2}\|w\|^2 + \frac{1}{2}\gamma \sum_{k=1}^{N} e_k^2 - \sum_{k=1}^{N} \alpha_k \{\langle w, \Phi(x_k) \rangle + b + e_k - y_k\} \tag{8}$$

where α_k is the Lagrange multiplier and the sample whose Lagrange multiplier is not equal 0, is the support vector.

The solution of Eq. (8) is determined by partially differentiating with respect to w, b, e_k and α_k [27].

2.4. *K-fold cross validation*

Cross-validation is a measurement of assessing the performance of a predictive model, and statistical analysis will generalize to an independent dataset. There are many types of cross-validation, such as repeated random sub-sampling validation, *K*-fold cross-validation, $K \times 2$ cross-validation, leave-one-out cross-validation and so on. In this study, we pick up *K*-fold cross- validation for selecting parameters of model [30–32].

The *K*-fold cross-validation is a technique of dividing the original sample randomly into *K* sub-samples. It includes 3 steps [32]:

Step 1: Divide the data *K* roughly into equal parts;

Step 2: For each $i = 1, 2, 3, \ldots, K$ fit the model with parameter γ or other $K-1$ parts, giving $\alpha^{-k}(\gamma)$ and compute its error in predicting the *k*th part:

$$E_k(\gamma) = \sum_{i \in kth\ part} [y_i - x_i \alpha^{-k}(\gamma)]^2 \tag{9}$$

Step 3: Do this for many values of γ and choose the value of γ that makes smallest error.

Table 1 – Syntax of various training algorithms.

Training algorithm	Function	Symbol	Class
Levenberg–Marquardt back propagation	Trainlm	T1	Quasi-Newton (QD)
Bayesian regularization	Trainbr	T2	Bayesian regulation back propagation
Scaled conjugate gradient back propagation	Trainscg	T3	Conjugate gradient back propagation
Resilient back propagation (Rprop)	Trainrp	T4	Resilient back propagation
Variable learning rate back propagation	Traingdx	T5	Self-adaptive learning rate
Gradient descent with momentum back propagation	Traingdm	T6	Additive Momentum
gradient descent with adaptive learning rate back propagation	Traingda	T7	Self-adaptive learning rate
Gradient descent back propagation	Traingd	T8	Gradient descent back propagation
BFGS quasi-Newton back propagation	Trainbfg	T9	Quasi-Newton (QD)
Powell–Beale conjugate gradient back propagation	Traincgb	T10	Conjugate gradient back propagation
Fletcher–Powell conjugate gradient back propagation	Traincgf	T11	
Polak–Ribiere conjugate gradient back propagation	Traincgp	T12	
One step secant back propagation	Trainoss	T13	Quasi-Newton (QD)

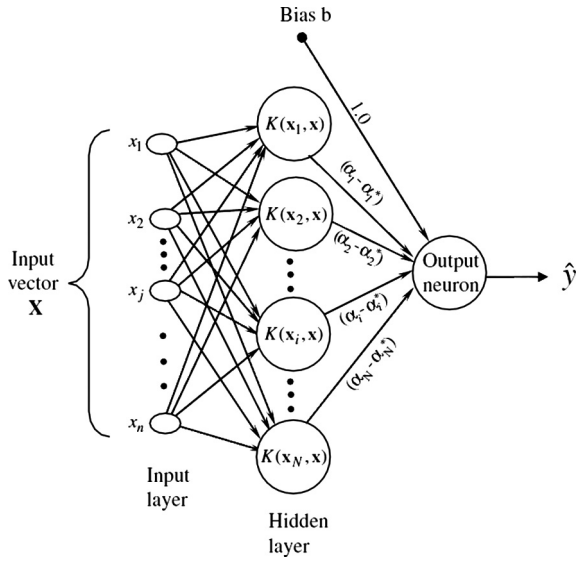

Fig. 3 – Architecture of SVM model [29].

2.5. Internal and external climate data

In this research, SHT11 sensors were used to measure both temperature and the relative humidity inside and outside the greenhouse. The accuracy of this device is ±0.4% at 20 °C and the precision measurement of the moisture is ±3% for a clear sky. Also, the TES1333R pyranometre was used for measure the global solar radiation. It is a measure of global radiation of the spectral band solar in the 400–1110 nm. Its measurement accuracy is approximately ±5%.

2.6. Energy exchange in greenhouse

In this section some of heat exchange equations in greenhouse were discussed. The goal of this part is to show the ability of the best soft computing model to adopt the energy exchange in greenhouse for environmental controlling. Some of main energy transferred between greenhouse elements by convection and conduction is expressed [33,34]:

$$Q_{a-p} = A_p \times \alpha_{a-p}(T_a - T_p) \quad (10)$$

$$Q_{a-s} = A_s \times \alpha_{a-s}(T_a - T_s) \quad (11)$$

$$Q_{s-ss} = A_s \times \lambda_s/ds(T_s - T_{ss}) \quad (12)$$

$$Q_{a-ri} = A_{ri} \times \alpha_{a-ri}(T_a - T_{ri}) \quad (13)$$

$$Q_{s-ri} = A_s \times E_s \times E_{ri} \times F_{s-ri} \times \sigma(T_s^4 - T_{ri}^4) \quad (14)$$

$$Q_{s-p} = A_s \times E_s \times E_p \times F_{s-p} \times \sigma(T_s^4 - T_p^4) \quad (15)$$

$$Q_{ri-p} = A_{ri} \times E_{ri} \times E_p \times F_{ri-p} \times \sigma(T_{ri}^4 - T_p^4) \quad (16)$$

Empirical relations reported in the literatures to estimate the heat transfer coefficients between the different surfaces in a greenhouse, are as follows [33,34]:

$$\alpha_{a-p} = \rho_a \times c_{pa}/R_{b-heat} \quad (17)$$

$$\begin{aligned} \alpha_{a-s} &= 1.7|T_a - T_s|^{\frac{1}{3}} \quad \text{if } T_a < T_s \\ \alpha_{a-s} &= 1.3|T_a - T_s|^{0.25} \quad \text{if } T_a \geqslant T_s \end{aligned} \quad (18)$$

$$\alpha_{a-ri} = 3|T_a - T_{ri}|^{1/3} \quad (19)$$

In Eq. (17), (R_{b-heat}) is the boundary layer resistance to convective heat transfer and can calculate by [33–34]:

$$R_{b-heat} = \frac{1174\sqrt{l_f}}{(l_f \times |T_c - T_a| + 207v_a^2)^{\frac{1}{4}}} \quad (20)$$

The input data for solution are given in Table 2. In order to control all effective inside and outside parameters on temperature prediction, simulation was done between 9:00 am to 16:00 pm in a conventional polyethylene greenhouse located Shahreza city.

2.7. Performance evaluation criteria

Different performance criteria have been used in literature to assess model's predictive ability. The mean absolute percentage error (MAPE), root mean square error (RMSE) and coefficient of determination (R^2) are selected to evaluate the forecast accuracy of the models in this study [34]:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{dv - pv}{dv}\right| \times 100 \quad (21)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(dv - pv)^2}{n}} \quad (22)$$

Table 2 – Input parameters used for calculation [34].

Parameter	Value	Parameter	Value
ρ_a	$1.29\frac{T_0}{T_a}$	A_s	200
E_s	0.7	λ_s	0.86
c_{p-a}	1000	ds	0.65
σ	5.67051×10^{-8}	E_p	$1 - \tau_{c-Il}$
F_{s-ri}	1.00	F_{s-c}	$F_{s-c} = 1 - \tau_{c-Il}$
E_{ri}	0.90	A_{ri}	300
A_p	$2LAI \times A_s$	k_{c-Il}	0.64
l_f	0.05	v_a	0.09
LAI	1	F_{ri-p}	$\tau_{c-Il} = e^{-k_{c-Il} \times LAI}$

$$R^2 = \left[\frac{\sum_{i=1}^{n}(dv - \overline{dv}) \times (pv - \overline{pv})}{\sum_{i=1}^{n}(dv - \overline{dv}) \times \sum_{i=1}^{n}(pv - \overline{pv})}\right]^2 \quad (23)$$

where n is number of data, dv is the desired value and pv is the predicted value. The best model is achieved when the RMSE and MAPE are minimized and R^2 is maximized. The linear regression line between dv and pv was also used to assess the models. From this perspective, the best model is the one that has the linear regression line with slope close to one, intercept close to zero and coefficient of determination close to 1 (pv = 1.00 dv + 0.00, R^2 = 0.99).

3. Results and discussion

3.1. Correlation coefficient between all variables

The procedure applied in this research started by distinguishes the correlation coefficients between all variables (Fig. 4). Studied the correlation coefficient among different characters makes it possible to decide more precisely about selected indirect selection indices and removing ineffective characters. As we can see, the inputs have a significant correlation to each other. This fact is because of these inputs have a very complex dynamic relation to each other. Every changes in one of them, can change the others and it is one of the problems to modeling the inside greenhouse environment by physical methods. In the view of classical feedback control, such a system is poorly controlled if disturbance monitors and model based feed forward control is not applied. The outside wind speed did not have a significant relation with outside solar radiation (because of unique climate in this region). Other variables have a strong relation with themselves, so we used all the variables for three future models and for each one try to decrease the predictive error.

3.2. Results of ANN (MLP and RBF) and SVM models

In this research, application of ANN (MLP and RBF) and SVM models was evaluated for prediction the inside environment variables in a polyethylene greenhouse. Because there is a significant correlation between inside and outside variables (Fig. 4), it is possible to use predictive modes to evaluate the changes in inside parameters based on outside variables. This research is focus on energy and monitoring in a conventional greenhouse and based on the literatures, such as this method has not been seen for this region. This section includes the parts:

I. Initially, the optimal parameters of ANN (MLP and RBF) and SVM models are found.
II. Then their performance is compared with each other.
III. Finally, the sensitivity analysis of the optimal model will evaluate.

3.2.1. Selection the training algorithm for ANN (MLP) model by DOE

The results of neural networks optimization models with different networks, dependent on the initial random values of the synaptic weights. So, the results in general will not be the same in two different trials even if the same training data have been used [35]. So in this research *K-fold* cross validation (K = 5 with 4 replications) was used and finally 20 different data samples were made for train and test of ANN models.

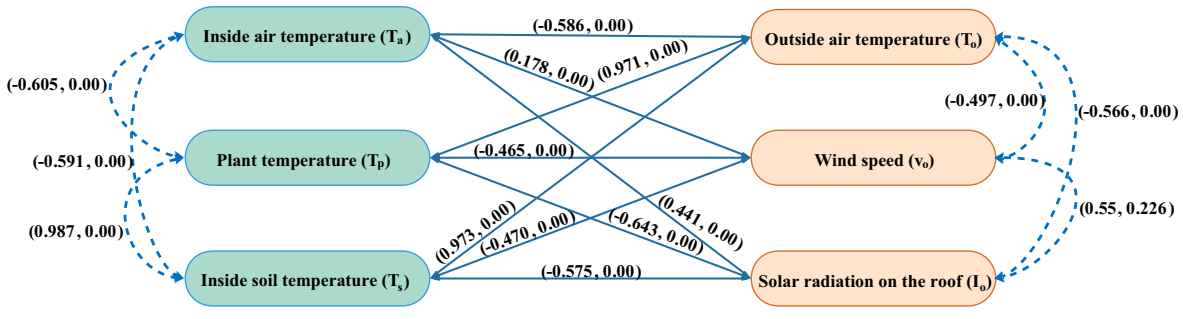

Fig. 4 – Correlation coefficient between all inputs and outputs (The first number in brackets shows the correlation value and the second number shows the value of p-value).

Table 3 – The results of the R^2 variance analysis for MLP-ANN.

Source	DF	Train			Test			Total		
		T_a	T_p	T_s	T_a	T_p	T_s	T_a	T_p	T_s
Train Alghorithm	12	0.41**	0.14**	0.18**	0.01**	0.14**	0.19**	0.41**	0.14**	0.19**
Block (Data set)	19	0.005**	0.01*	0.01*	0.383*	0.01*	0.01*	0.006**	0.01*	0.01*
Error	228	–	–	–	–	–	–	–	–	–
Total	259	–	–	–	–	–	–	–	–	–

* Significant at 5% (The values of the numbers inside the table represent the mean squares of the R^2 criterion).
** Significant at 1% (The values of the numbers inside the table represent the mean squares of the R^2 criterion).

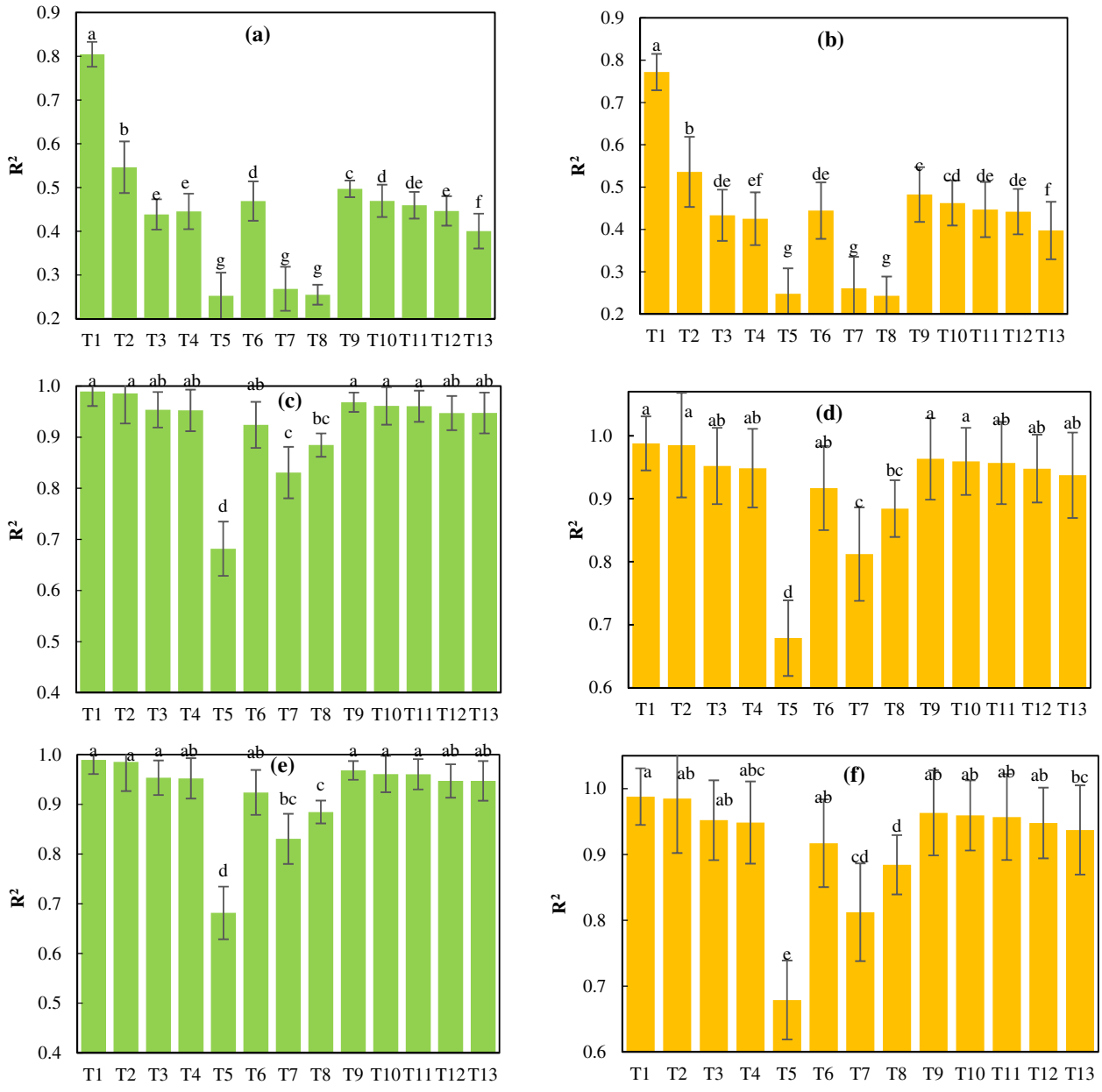

Fig. 5 – The results of the R^2 average comparison (13 algorithms) for MLP model in train and test phases (a, c and e for train; b, d and f for test) of T_a (a and b), T_p (c and d) and T_s (e and f) variables.

Also, the type of training algorithm can greatly affect the performance of models [36]. So, 13 types of training algorithms (Table 1) were used to train the ANN models. In this research, a novel statistical analysis was used. Randomized Complete Block (RCB) design was used to compare the performance of training algorithms as an innovative idea. So, the

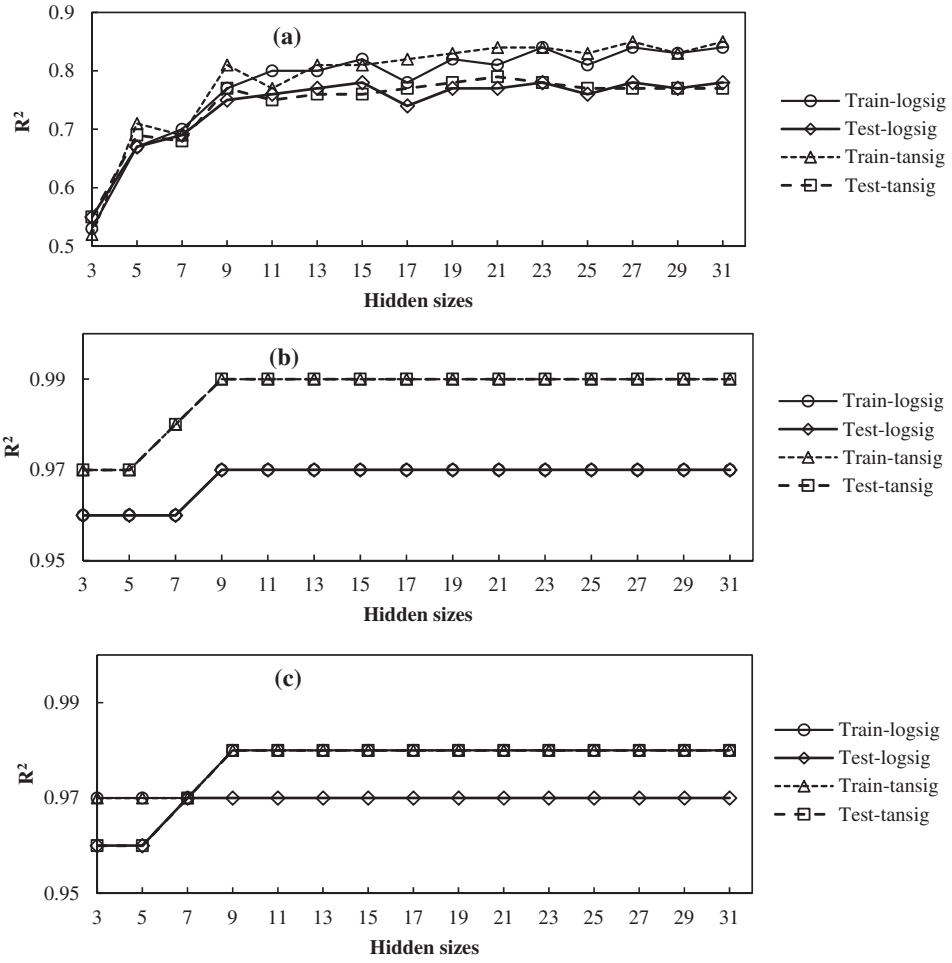

Fig. 6 – The mean value of R^2 changes at training and testing of MLP model by logsig and tansig transfer functions and different number of neurons in hidden layer (a, b and c for T_a, T_p and T_s).

Table 4 – The results of the R^2 variance analysis for RBF-ANN.

Source	DF	Train			Test			Total		
		T_a	T_p	T_s	T_a	T_p	T_s	T_a	T_p	T_s
Train Alghorithm	12	0.47**	0.04**	0.02**	0.47**	0.04**	0.02**	0.47**	0.04**	0.02**
Block (Data set)	19	0.23**	0.03**	0.01**	0.29**	0.04**	0.01**	0.24**	0.03**	0.01**
Error	228									
Total	259									

*Significant at 5% (The values of the numbers inside the table represent the mean squares of the R^2 criterion and DF is the degree of freedom).
** Significant at 1% (The values of the numbers inside the table represent the mean squares of the R^2 criterion and DF is the degree of freedom).

type of training algorithm was considered as the experimental treatments. The MLP model was also trained by each of training algorithm with 20 different data set and 20 times. The results of replications were considered as experimental blocks in RCB design. The results of R^2 variance analysis for MLP neural network model in three stages of training, testing and total for the three variables (T_a, T_p and T_s) are shown in Table 3. As the results show, the effect of training algorithm type on the correlation between the actual and predicted values (R^2) in all stages, is significant (level of 1%). The same result is obtained for the type of data set (test blocks). With understanding the effect of training algorithms on MLP performance, the R^2 average comparison of the MLP model for the three variables T_a, T_p and T_s in the two phases of training and testing with the Least Square Difference (LSD) at the 5% level was analyzed (Fig. 5).

As Fig. 5 shows, there are significant differences between training algorithms. The trainlm algorithm for the three output variables has a significantly higher R^2 than other algorithms. The performance of trainbr is also compatible with the trainlm algorithm. The worst performance for the MLP neural network in this research is traingdx. In general, the training algorithms were clustered into 7 categories for T_a estimation. In contrast, this parameter for T_p and T_s

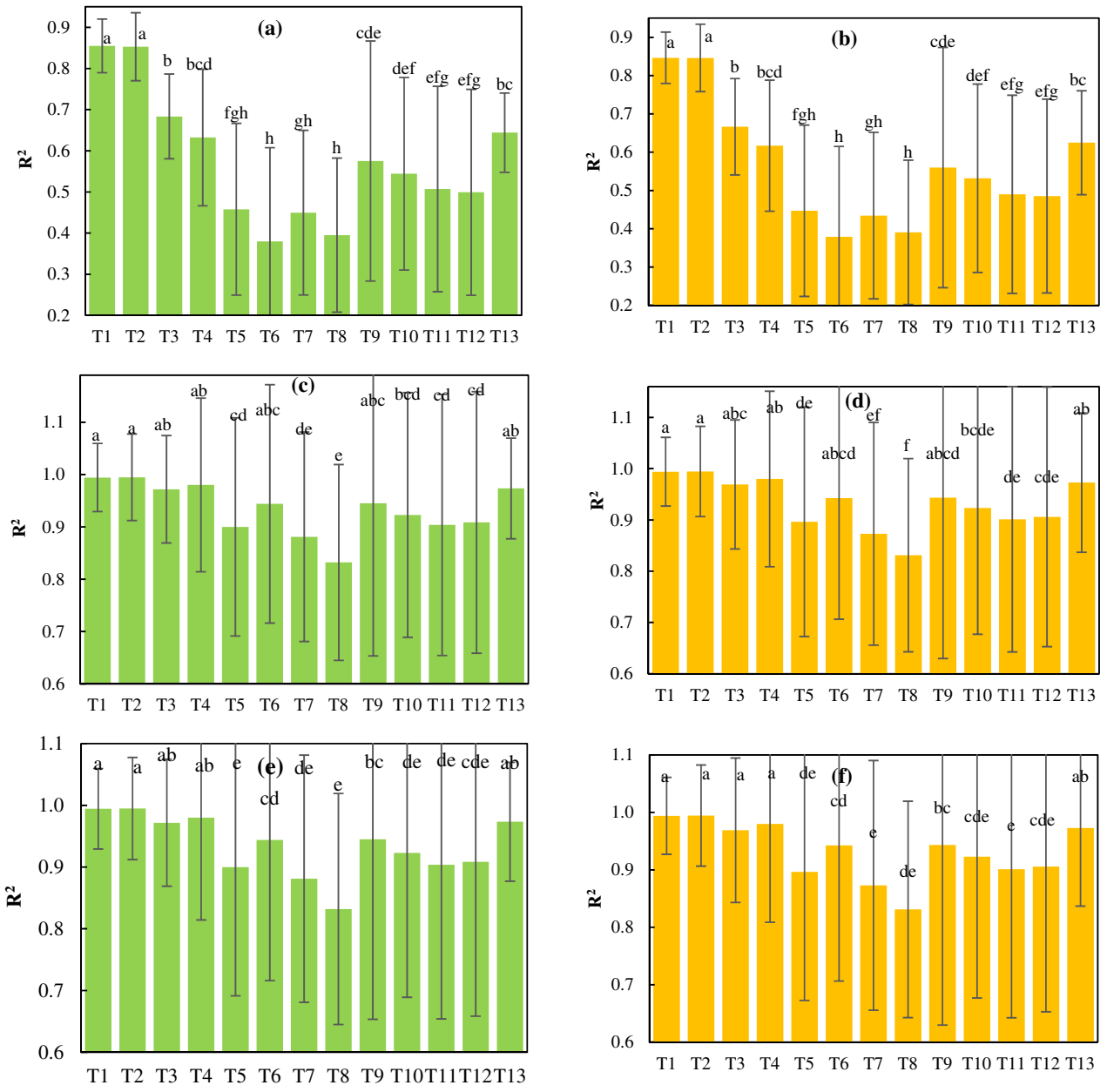

Fig. 7 – The results of training algorithm (13 algorithms) on RBF model performance in train and test phases (a, c and e for train; b, d and f for test) of T_a (a and b), T_p (c and d) and T_s (e and f).

algorithms were clustered in 4 groups. This is because of the complexity and variations of T_a than T_p and T_s. Other parameters that can affect the prediction function of the MLP model are the number of neurons in the hidden layer and the type of activation function. Fig. 6 shows the mean value of R^2 changes for logsig and tansig transfer functions and the number of different neurons in the hidden layer for the trainlm algorithm in training and testing.

As the results indicate, increasing number of neurons in the hidden layer will uptrend the R^2 value. In this research, 21, 9, and 9 neurons in the hidden layer were selected as the best number of neurons for T_a, T_p and T_s modeling by MLP network, respectively. Adding a number of neurons more than these values will not change much in the R^2 for MLP model. Also, the results showed that the tansig as an activation function in the hidden layer was better than logsig.

3.2.2. *Selection the training algorithm for ANN (RBF) model*

As Fig. 6 showed, the type of data set and training algorithm can mainly affect the predictive function of ANN models. Similar to the MLP model, RCB design was used to compare the statistical function for RBF model. Training algorithms and dataset determine the treatment and block of the RCBD design, respectively. Table 4 shows the ANOVA analysis for three output variables (T_a, T_p and T_s) at three stages of the network. The values of the numbers inside the table represent the mean squares (MS) of the R^2 criterion for the RBF model. As the results show, training algorithm and type of data set

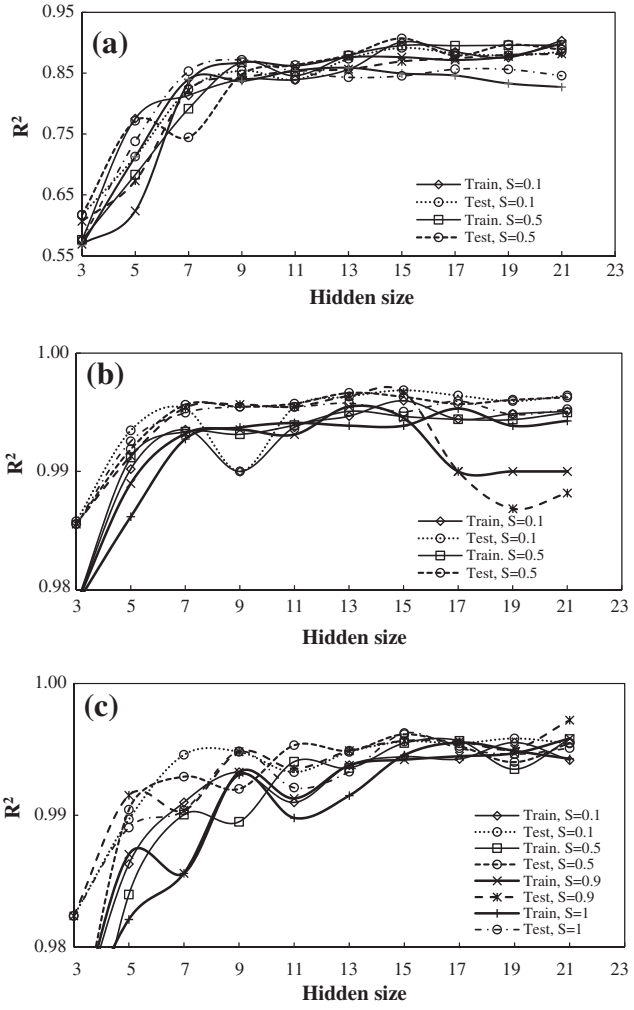

Fig. 8 – The mean value of R^2 at training and testing of RBF model for different number of neurons in hidden layer and different values of spread parameter (a, b and c for T_a, T_p and T_s).

has a significant effect on RBF network prediction performance (in all cases, p-value <0.05).

After analysis of variance, using LSD method at 5% level, the average of R^2 for different training algorithms was compared with each other (Fig. 7). As it can show, training algorithms for all the output variables were clustered in 8, 6 and 5 groups. Fig. 7 shows that the results of the statistical comparisons are similar to each other in two phases of training and testing. This means that if the learning algorithm has a good performance to find the relationship between variables in training step, the model almost can be successful at the test phase. The difference between mean values of R^2 for trainlm and trainbr algorithms in all output variables at both stages of training and testing was not significant. Also, the performance of these algorithms has a significant difference with other algorithms. But based on experience, we chose trainbr algorithm as the best training algorithm for RBF. Here, unlike MLP, the T8 algorithm had the worst performance compared to others.

In RBF model, the spread parameter and hidden size can affect the performance in all cases. Fig. 8 shows the effect of these parameters on the R^2 value for the RBF model with trainbr algorithm. As it can show, with increasing the hidden size, the R^2 value also increase. The process of increasing and changing of R^2 is also dependent on the spread parameter values. For T_a, S = 0.5 with 15 neurons, for T_p, S = 0.1 with 15 neurons and for T_s, S = 1 with 15 neurons in RBF network design was selected.

3.2.3. *Selection the best kernel function for SVM model*

The model's prediction performance can be affected by the type of kernel function [36]. In this research, RCB design was used to compare the type of kernel function. For SVM model, 4 kernel functions including linear (Poly 1), second-order polynomial (Poly 2), third-order polynomial (Poly 3) and radial bias (RBF) function were used. Training algorithms and dataset, determine the treatment and block of RCB design, respectively. Table 5 shows the results of ANOVA for SVM model. As the results show, training algorithm and type of data set have significant effect on SVM model prediction performance (in all cases, p-value <0.05). After analysis of variance, using LSD method at 5% level, the average of R^2 for different training algorithms was compared with each other (Fig. 9). The results show that liner function is the best training algorithm for SVM model.

Recently, some of researchers used SVM model in greenhouse. Yu et al. [2], presented a novel temperature prediction model based on support vector machine and improved particle swarm optimization (IPSO). The IPSO with probability of mutation was employed to optimize the required hyper parameters of the SVM model. The performance of the IPSO–SVM model compared with traditional modeling approaches by applying it to predict solar greenhouse temperatures. The results showed that its predictions of the maximum and minimum temperature were more accurate than

Table 5 – The results of the R^2 variance analysis for SVM model.

Source	DF	Train			Test			Total		
		T_a	T_p	T_s	T_a	T_p	T_s	T_a	T_p	T_s
Kernel function	3	0.89**	0.00**	0.00**	0.60**	0.00**	0.00**	0.82**	0.00**	0.00**
Block (Data set)	19	0.00*	0.00**	0.00**	0.01**	0.00**	0.01*	0.00*	0.00**	0.00*
Error	57									
Total	79									

* Significant at 5% (The values of the numbers inside the table represent the mean squares of the R^2 criterion and DF is the degree of freedom).
** Significant at 1% (The values of the numbers inside the table represent the mean squares of the R^2 criterion and DF is the degree of freedom).

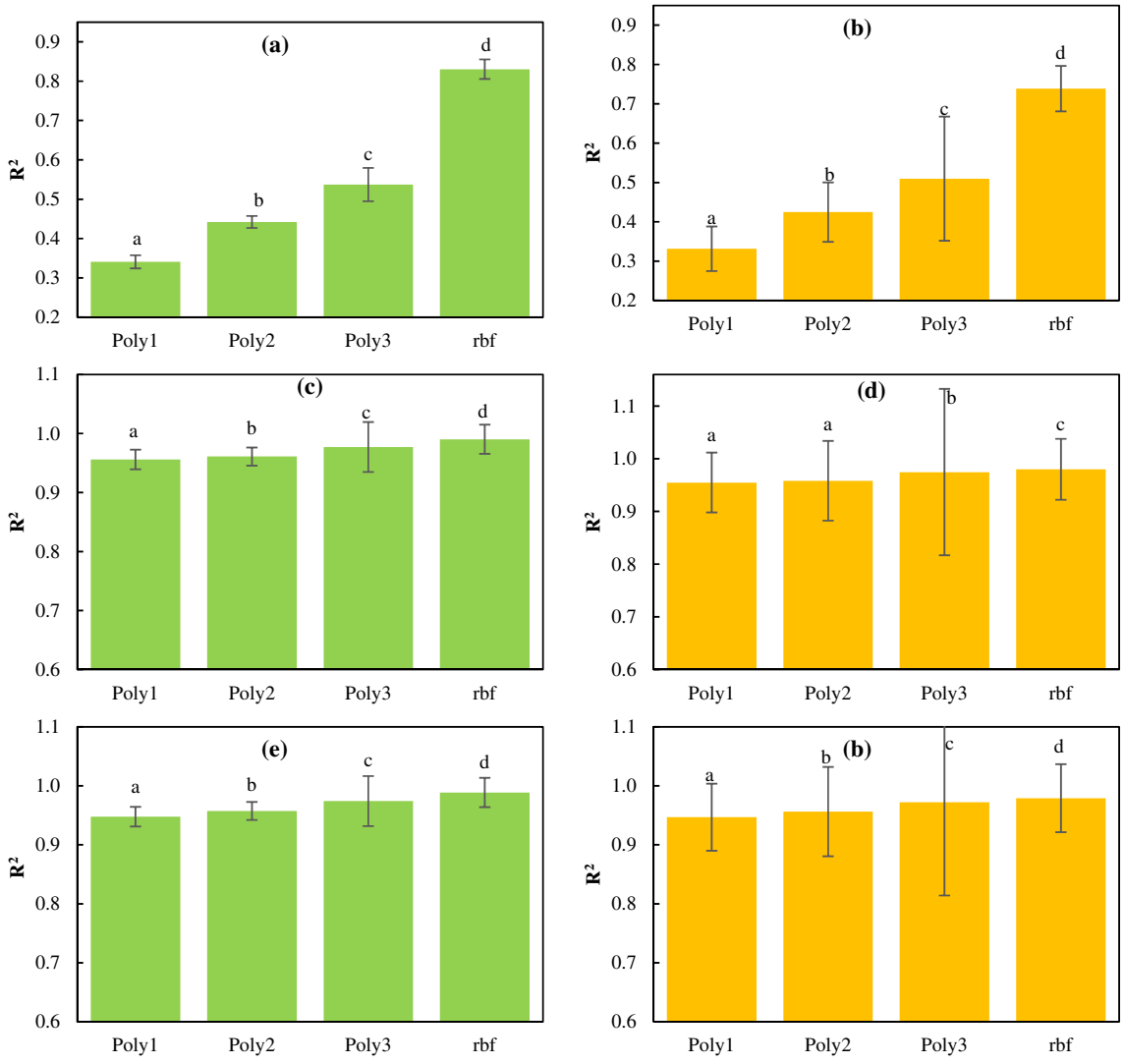

Fig. 9 – The result of the R^2 (4 algorithms) for SVM model in train and test phases (a, c and e for train; b, d and f for test) of T_a, T_p and T_s.

those of the standard support vector machine and back propagation neural network.

3.3. *Selection the best model between MLP, RBF and SVM*

In the previous sections, the best training algorithm and kernel function for MLP, RBF and SVM models were selected. Here, our goal is to evaluate, compare and select the best predictive model. As the results showed, the type of data set has a significant effect on the performance of the models. So, 100 different data sets were used to compare the results of the models with each other. In Table 6, the results of RMSE, MAPE and R^2 for MLP, RBF and SVM networks are presented based on the percentage of the training data set. In this table, by reducing the size of training data sets, the generalizability characteristics of the MLP, RBF and SVM models can evaluate. The results show that the standard deviation values for RBF errors are less than the MLP and SVM, so the RBF model is less sensitive to changing the training dataset than other models. The results indicate that by decreasing the number of training set, the training errors will decrease but the test phase errors will increase. Although the error values of the three models are acceptable in all sizes of the training set. The comparison of the results shows that the RBF model has less predictive errors.

Based on the results of Table 6, the RBF model was selected to predict inside environment variables in this research. Fig. 10 shows the results of the RBF model evaluation for 50% of the total data for training. As the results show, there is a very good agreement between the actual and predicted data by this model.

3.4. *The results of sensitivity analysis*

In this section, the sensitivity analysis was carried out. The current well-established sensitivity analysis methods can be classified into two categories: local and global techniques [37]. In this research, local sensitivity technique was selected and used. Local approaches estimate the effect of a single factor on the outputs while keeping all the others fixed at their

Table 6 – Evaluation the MLP, RBF and SVM methods versus the size of training data sets.

TS*		Model	Train			Test		
			RMSE	MAPE	R^2	RMSE	MAPE	R^2
80	T_a	MLP	0.17 ± 0.01	0.80 ± 0.09	0.82 ± 0.03	0.20 ± 0.02	0.93 ± 0.10	0.76 ± 0.05
		RBF	0.13 ± 0.00	0.58 ± 0.03	0.89 ± 0.01	0.13 ± 0.01	0.59 ± 0.07	0.89 ± 0.03
		SVM	0.17 ± 0.01	0.69 ± 0.03	0.83 ± 0.02	0.21 ± 0.03	0.89 ± 0.08	0.73 ± 0.05
	T_p	MLP	0.19 ± 0.00	0.74 ± 0.23	0.99 ± 0.00	0.20 ± 0.05	0.75 ± 0.23	0.99 ± 0.00
		RBF	0.12 ± 0.01	0.44 ± 0.03	0.99 ± 0.00	0.12 ± 0.01	0.44 ± 0.06	0.99 ± 0.00
		SVM	0.19 ± 0.00	0.85 ± 0.01	0.99 ± 0.00	0.26 ± 0.04	1.03 ± 0.10	0.97 ± 0.00
	T_s	MLP	0.10 ± 0.00	0.44 ± 0.03	0.98 ± 0.00	0.12 ± 0.01	0.51 ± 0.05	0.98 ± 0.00
		RBF	0.07 ± 0.01	0.28 ± 0.07	0.99 ± 0.00	0.07 ± 0.01	0.28 ± 0.08	0.99 ± 0.00
		SVM	0.11 ± 0.00	0.54 ± 0.01	0.98 ± 0.00	0.15 ± 0.02	0.64 ± 0.06	0.97 ± 0.00
60	T_a	MLP	0.17 ± 0.02	0.75 ± 0.11	0.83 ± 0.04	0.21 ± 0.03	0.94 ± 0.12	0.74 ± 0.07
		RBF	0.13 ± 0.01	0.57 ± 0.05	0.89 ± 0.01	0.13 ± 0.01	0.57 ± 0.06	0.89 ± 0.02
		SVM	0.17 ± 0.01	0.68 ± 0.04	0.82 ± 0.02	0.23 ± 0.03	0.94 ± 0.05	0.70 ± 0.07
	T_p	MLP	0.16 ± 0.01	0.64 ± 0.04	0.99 ± 0.00	0.19 ± 0.01	0.76 ± 0.05	0.99 ± 0.00
		RBF	0.11 ± 0.01	0.42 ± 0.04	0.99 ± 0.00	0.12 ± 0.01	0.45 ± 0.03	0.99 ± 0.00
		SVM	0.19 ± 0.00	0.88 ± 0.02	0.99 ± 0.00	0.30 ± 0.04	1.16 ± 0.07	0.97 ± 0.00
	T_s	MLP	0.10 ± 0.02	0.43 ± 0.09	0.98 ± 0.00	0.13 ± 0.02	0.54 ± 0.08	0.98 ± 0.00
		RBF	0.06 ± 0.00	0.26 ± 0.02	0.99 ± 0.00	0.06 ± 0.00	0.26 ± 0.02	0.99 ± 0.00
		SVM	0.11 ± 0.00	0.56 ± 0.01	0.98 ± 0.00	0.17 ± 0.02	0.72 ± 0.04	0.97 ± 0.00
40	T_a	MLP	0.17 ± 0.01	0.78 ± 0.08	0.82 ± 0.03	0.24 ± 0.03	1.05 ± 0.08	0.67 ± 0.09
		RBF	0.14 ± 0.00	0.60 ± 0.03	0.89 ± 0.01	0.14 ± 0.00	0.60 ± 0.04	0.89 ± 0.01
		SVM	0.18 ± 0.01	0.67 ± 0.05	0.81 ± 0.03	0.25 ± 0.03	1.01 ± 0.06	0.65 ± 0.07
	T_p	MLP	0.17 ± 0.02	0.65 ± 0.12	0.99 ± 0.00	0.22 ± 0.03	0.82 ± 0.09	0.99 ± 0.00
		RBF	0.12 ± 0.01	0.42 ± 0.06	0.99 ± 0.00	0.12 ± 0.01	0.45 ± 0.03	0.99 ± 0.00
		SVM	0.20 ± 0.01	0.91 ± 0.05	0.98 ± 0.00	0.37 ± 0.07	1.31 ± 0.14	0.96 ± 0.01
	T_s	MLP	0.11 ± 0.01	0.47 ± 0.06	0.98 ± 0.00	0.15 ± 0.02	0.65 ± 0.10	0.97 ± 0.00
		RBF	0.06 ± 0.00	0.26 ± 0.02	0.99 ± 0.00	0.06 ± 0.00	0.26 ± 0.02	0.99 ± 0.00
		SVM	0.12 ± 0.00	0.57 ± 0.02	0.98 ± 0.00	0.21 ± 0.03	0.81 ± 0.08	0.95 ± 0.00

* TS: Training Size.

nominal values. Table 7 shows the results of sensitivity analysis.

The results of Table 7 show that all the outputs (T_a, T_p and T_s) definitely depended on outside air temperature (T_o). It is true based on heat transfer equations. Some of researchers applied inside and outside variables for estimation the greenhouse environment, but for more accurate, we used only outside variables as the inputs. A lot of researches used dynamic models since 1960 to 2016 for simulation and modeling the greenhouse environment. Joudi and Farhan [38], presented a dynamic model to predict the inside air and soil temperature in an innovative greenhouse in Iraq. The input parameters of this model collected from measured meteorological conditions and the thermo-physical properties of the greenhouse components which include the cover, inside air, and soil. Comparisons between the predicted and measured results showed good agreement. The absolute error in this dynamic model was lower that 10% for inside air and soil temperature.

Du et al. [39], applied the simulation method to predict the inside air and soil in a greenhouse with heat pipe system. The model validated with experimental data and found to be in close agreement. The absolute error between predicted and desired data was about ±20%. The results showed that for estimation the inside variables in all types of greenhouses, soft computing models can be very successful than physical and mathematical models. The results of some researches are the proofs of conclusion in this research [40–44]. For the future studies, some of other soft computing models such as Adaptive Network Based Fuzzy Inference System (ANFIS) and Gaussian Process Regression (GPR) should be evaluate.

3.5. *Energy exchange in greenhouse*

The final part of this paper is to comparison the results of RBF, MLP and SVM models to estimate the total energy exchange inside the greenhouse by using mathematical Eqs. (10)–(20). Energy exchange and lost through the greenhouse include $Q_{ri\text{-}p}$, $Q_{s\text{-}p}$, $Q_{a\text{-}ri}$, $Q_{s\text{-}ss}$, $Q_{a\text{-}p}$, $Q_{a\text{-}s}$ and $Q_{s\text{-}ri}$ that cover the conduction, convection and radiation phenomena between some parts of greenhouse. The inside air, soil and plant temperature effect these parameters. Figs. 11 and 12 show the comparison of RBF, MLP and SVM models to predict energy lost and exchange. As Fig. 11 shows, the direction of heat fluxes between inside air and soil until noon is to air because the soil is warmer than air (incoming solar radiation can warm the inside soil of greenhouse), but afternoon the direction is to inside air. The heat exchange direction between inside air and crop at morning is to air (almost for evapotranspiration) and before night it will be decrease. One of the main energy lost corridor at greenhouse is by conduction between top and underground soil of greenhouse. Fig. 11 shows the direction of heat lost whole the day (from top to underground). Finally, the convection heat exchange between inside air and inside roof of greenhouse especially in winter is the

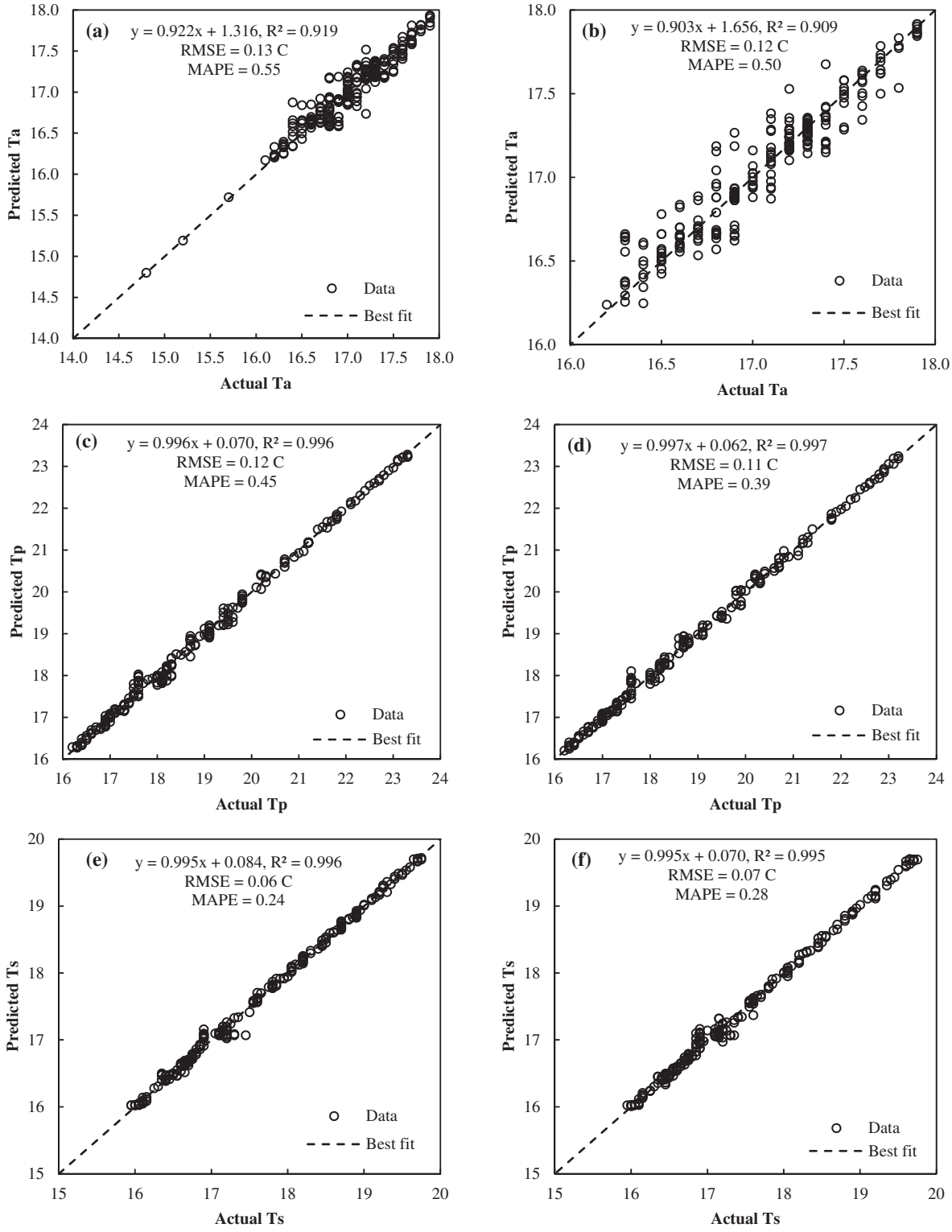

Fig. 10 – The agreement between actual and predicted values of T_a, T_p and T_s with RBF model at training and test phases (a and b; c and d; e and f are the train and test of T_a, T_p and T_s, respectively).

important factor for energy monitoring in a greenhouse and can be seen in Fig. 11. The results show that the RBF model has the best performance in all cases. As it can show, the difference between the results of RBF and actual data is not significant but other models (MLP and SVM) had a significant difference with RBF model.

Fig. 12 shows the results of radiation heat exchange modeling with all three models between inside roof, soil and plant. It shows the direction of heat flux from plant to soil and inside roof. Actually in greenhouse, crop temperature is higher than soil and roof temperature. As it can show, the result of RBF model is very similar to actual data.

Table 7 – The results of sensitivity analysis for remove the ineffective variables.

	Varaible	RMSE	MAPE	R^2
T_a	All	0.12	0.52	0.92
	All exclude T_o	0.44	2.07	0.20
	All exclude v_o	0.13	0.51	0.91
	All exclude I_o	0.28	1.25	0.58
T_p	All	0.11	0.42	0.99
	All exclude T_o	0.79	3.06	0.82
	All exclude v_o	0.13	0.42	0.99
	All exclude I_o	0.38	1.43	0.96
T_s	All	0.07	0.26	0.99
	All exclude T_o	0.50	2.06	0.77
	All exclude v_o	0.09	0.33	0.99
	All exclude I_o	0.19	0.75	0.97

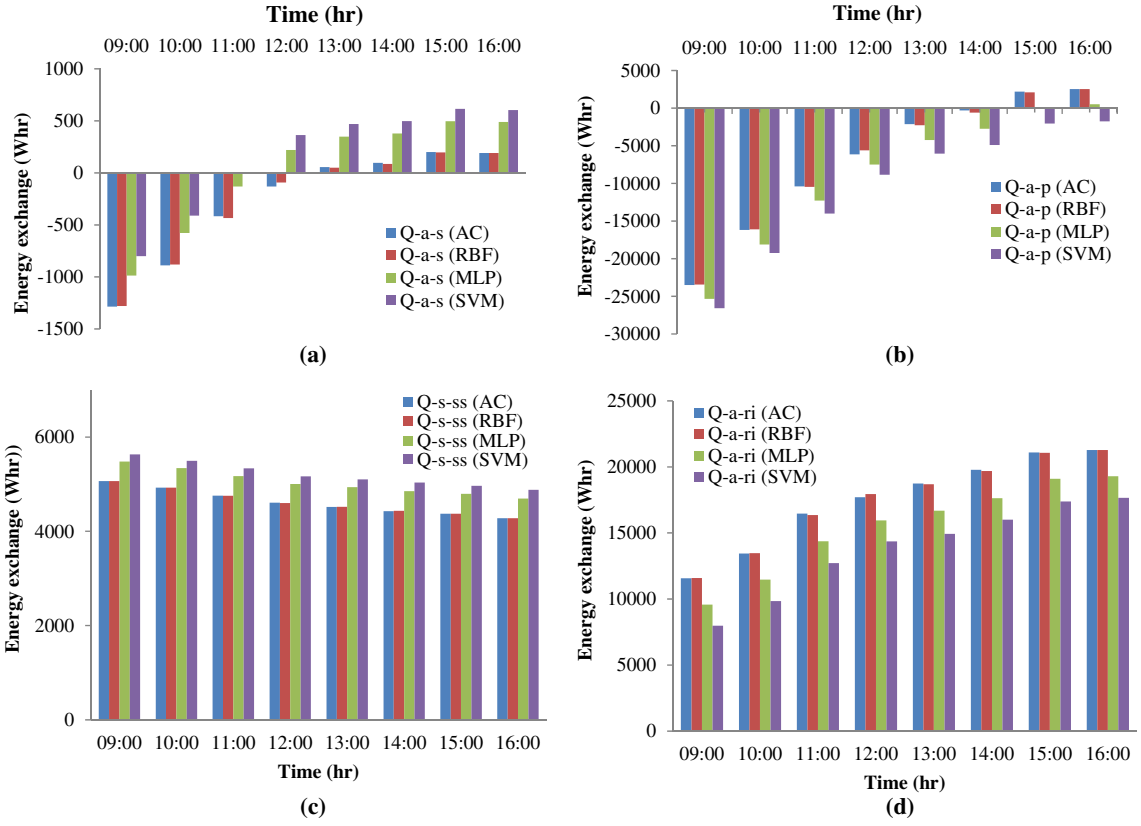

Fig. 11 – Comparison of hourly energy lost and exchange predicted by RBF, MLP and SVM models (a, b and c convection and d conduction heat exchange).

In this paper, we tried to show the fact that innovative methods are simple and more accurate than physical heat and mass transfer method to predict the environment changes. Furthermore, this method can use to predict other changes in greenhouse such as final yield, evapotranspiration, humidity, cracking on the fruit, CO_2 emission and so on. So the future research will focus on the other soft computing models such as ANFIS, GPR, Time Series and ... to select the best one for modeling and finally online control of greenhouse in all climate and different environment.

4. Conclusion

The paper presents a comparison between Artificial Neural Network (ANN) and Support Vector Machine (SVM) models to predict three point temperatures (T_a, T_p and T_s) and energy exchange in a conventional greenhouse at Shahreza city in Isfahan province, Iran. For this purpose, some inside and outside variables used as input and the relation between them was examined. The *K-fold* cross validation (K = 5 with 4 replications) was used to show the ability of models to pre-

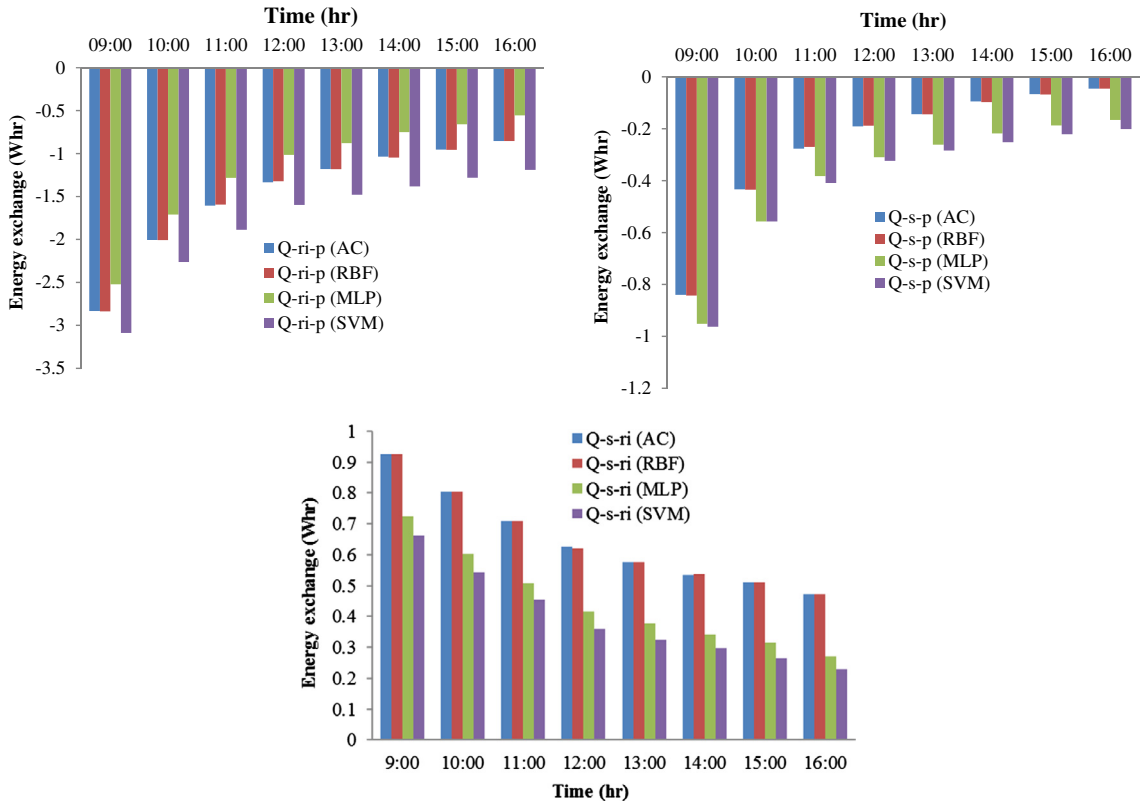

Fig. 12 – Comparison of hourly energy lost and exchange predicted by RBF, MLP and SVM models (long wave radiation).

dict inside parameters. 13 different training algorithms were used for ANN models. In this research, randomized complete block methodology was used to compare the performance of training algorithms for all models as an innovative idea. All three models were trained by each of training algorithm with 20 different data set (K = 5 with 4 replications). The results of these 20 replications were considered as experimental blocks in randomized complete block design. The results showed that type of training algorithm is very important in all three models. Comparison of the models showed that RBF has lowest error between other models. The range of RMSE and MAPE factors for RBF model to predict T_a, T_p and T_s were between 0.07 and 0.12 °C and 0.28–0.50%, respectively. Also the results showed that RBF model can estimate the energy exchange and lost in greenhouse with high accuracy. Such forecasts can be used by farmers as an appropriate advanced notice for changes in temperatures. So they can apply preventative measures to avoid damage caused by extreme temperatures. More specifically, predicting a greenhouse temperature can not only provide a basis for greenhouse environment management decisions that can reduce the planting risks, but also can be as a basic research for the feedback-feed-forward type of climate control strategy.

Conflict of interest

The authors declare that there is no conflicts of interest.

Acknowledgments

The authors would like to thank the editor in chief and the anonymous referees for their valuable suggestions and useful comments that improved the paper content substantially. This study was supported by a grant (961/06) from Ramin Agriculture and Natural Resources University of Khuzestan, Iran. The authors are grateful for the support provided by this University.

REFERENCES

[1] Taki M, Rohani A, Rahmati-Joneidabad M. Solar thermal simulation and applications in greenhouse. Info Proc Agri 2018;5:83–113.
[2] Yu H, Chen Y, Gul Hassan S, Li D. Prediction of the temperature in a Chinese solar greenhouse based on LS-SVM optimized by improved PSO. Comput Electron Agric 2016;122:94–102.
[2] Yu H, Chen Y, Gul Hassan S, Li D. Prediction of the temperature in a Chinese solar greenhouse based on LS-SVM optimized by improved PSO. Comput Electron Agric 2016;122:94–102.
[3] Taki M, Abdi R, Akbarpour M, Ghasemi-Mobtaker H. Energy inputs-yield relationship and sensitivity analysis for tomato greenhouse production in Iran. Agric Eng Int: CIGR J 2013;15 (1):59–67.
[4] Bot, G. Greenhouse climate: from physical process to a dynamical model. PhD Dissertation, Agric. Univ. Wageningen, The Netherlands, 1983.

[5] Van Henten EJ. Greenhouse climate management: an optimal control approach. PhD Dissertation. the Netherlands: Agric. Univ. Wageningen; 1994.
[6] De Zwart HF. Analyzing energy-saving options in greenhouse cultivation using a simulation model. PhD Dissertation. the Netherlands: Agric. Univ. Wageningen; 1996.
[7] Bot GPA. Physical modelling of greenhouse climate. In: Proceeding of the IFAC/ISHS Workshop; 1991. pp. 7–12.
[8] Abdel Ghany AM, Helal IM. Solar energy utilization by a greenhouse: general relations. Renewable Energy 2011;36:189–96.
[9] Udinkten C. Modeling and simulation in greenhouse climate control. Acta Hort 1985;174:461–7.
[10] Nielsen H, Madsen P. Identification of transfer functions for control of greenhouse air temperature. J Agric Eng Res 1995;60(1):25–34.
[11] Wang D, Wang M, Qiao X. Support vector machines regression and modeling of greenhouse environment. Comput Electron Agric 2009;66:46–52.
[12] Dariouchy E, Aassif K, Lekouch L, Bouirden G. Maze Prediction of the intern parameters tomato greenhouse in a semi-arid area using a time-series model of artificial neural networks. Measurement 2009;42:456–63.
[13] Abdi R, Taki M, Akbarpour M. An analysis of energy input-output and emissions of greenhouse gases from agricultural productions. Int J Nat Eng Sci 2012;6(3):73–9.
[14] Ruano AE, Crispim EM, Conceição EZE, Lúcio M. Prediction of building's temperature using neural networks models. Energy Build 2006;38(6):682–94.
[15] Abdi R, Taki M, Jalali A. Study on energy use pattern, optimization of energy consumption and CO_2 emission for greenhouse tomato production. Int J Nat Eng Sci 2013;7 (1):01–4.
[16] Zou W, Yao F, Zhang B, He C, Guan Z. Verification and predicting temperature and humidity in a solar greenhouse based on convex bidirectional extreme learning machine algorithm. Neurocomputing 2017. https://doi.org/10.1016/j.neucom.2017.03.023.
[17] He F, Ma C. Modeling greenhouse air humidity by means of artificial neural network and principal component analysis. Comput Electron Agric 2010;71S:S19–23.
[18] Taki M, Ajabshirchi Y, Ranjbar SF, Rohani A, Matloobi M. Heat transfer and MLP neural network models to predict inside environment variables and energy lost in a semi-solar greenhouse. Energy Build 2016;110:314–29.
[19] Haykin S. Neural networks and learning machines. 3rd ed., Prentice Hall; 2009.
[20] Wang G, Qiu Y, Li H. Temperature forecast based on SVM optimized by PSO algorithm. Int. Conf. Intell. Comput. Cogn. Inform 2010:259–62.
[21] Lins ID, Araujo M, Moura MDC, Silva MA, Droguett EL. Prediction of sea surface temperature in the tropical Atlantic by support vector machines. Comput Stat Data Anal 2013;61:187–98.
[22] Rohani A, Abbaspour-Fard MH, Abdolahpour S. Prediction of tractor repair and maintenance costs using Artificial Neural Network. Expert Syst Appl 2011;38:8999–9007.
[23] Taki M, Mahmoudi A, Ghasemi-Mobtaker H, Rahbari H. Energy consumption and modeling of output energy with multilayer feed-forward neural network for corn silage in Iran. Agric Eng Int: CIGR J 2012;14(4):93–101.
[24] Iliyas SA, Elshafei M, Habib MA, Adeniran AA. RBF neural network inferential sensor for process emission monitoring. Control Eng Pract 2013;21:962–70.
[25] Chen S, Cowan CFN, Grant PM. Orthogonal least square learning algorithm for radial basis function network. IEEE Trans Neural Networks 1991;2:302–9.
[26] Jung HC, Kim JS, Heo H. Prediction of building energy consumption using an improved real coded genetic algorithm based least squares support vector machine approach. Energy Build 2015;90:76–84.
[27] Cao H, Xin Y, Yuan Q. Prediction of biochar yield from cattle manure pyrolysis via least squares support vector machine intelligent approach. Bioresour Technol 2016;202:158–64.
[28] Suykens JAK, Vandewalle J. Least squares support vector machine classifiers. Neural Process Lett 1999;9:293–300.
[29] Tripathi A, Srinivas VV, Nanjundiah RS. Downscaling of precipitation for climate change scenarios: a support vector machine approach. J Hydrol 2006;330:621–40.
[30] Tvedskov TF, Meretoja TJ, Jensen MB, Leidenius M, Kroman N. Cross-validation of three predictive tools for non-sentinel node metastases in breast cancer patients with micrometastases or isolated tumor cells in the sentinel node. Eur J Surg Oncol 2013;40:435–41.
[31] Shaoa C, Paynabarb K, Kima TH, Jinc JH, Hu SJ, Spicerd JP, et al. Feature selection for manufacturing process monitoring using cross-validation. J Manuf Syst 2013;32:550–5.
[32] Jiang P, Chen J. Displacement prediction of landslide based on generalized regression neural networks with K-fold cross-validation. Neurocomputing 2016. https://doi.org/10.1016/j.neucom.2015.08.118.
[33] Taki M, Ajabshirchi Y, Ranjbar SF, Rohani A, Matloobi M. Modeling and experimental validation of heat transfer and energy consumption in an innovative greenhouse structure. Inform Process Agric 2016;3:157–74.
[34] Van Straten G, Van Willigenburg E, Van Henten R, Van Oothghem R. Optimal control of greenhouse cultivation. Taylor and Francis, New York: CRC Press; 2011.
[35] Pandey S, Hindoliya DA, Mod R. Artificial neural networks for predicting indoor temperature using roof passive cooling techniques in buildings in different climatic conditions. Appl Soft Comput 2012;12:1214–26.
[36] Rohani A, Taki M, Abdollahpour M. A novel soft computing model (Gaussian process regression with K-fold cross validation) for daily and monthly solar radiation forecasting (Part: I). Renew Energy, doi: 10.1016/j.renene.2017.08.061.
[37] Chen L, Yan G, Tianxing W, Ren H, Calbo J, Zhao J, et al. Estimation of surface shortwave radiation components under all sky conditions: modeling and sensitivity analysis. Remote Sens Environ 2012;123:457–69.
[38] Joudi K, Farhan A. A dynamic model and an experimental study for the internal air and soil temperatures in an innovative greenhouse. Energy Convers Manage 2015;91:76–82.
[39] Du J, Bansal P, Huang B. Simulation model of a greenhouse with a heat-pipe heating system. Appl Energy 2012;93:268–76.
[40] Seginer I. Some artificial neural network applications to greenhouse environmental control. Comput Electron Agric 1997;18:167–86.
[41] Linker R, Seginer I. Greenhouse temperature modeling: a comparison between sigmoid neural networks and hybrid models. Math Comput Simul 2004;65:19–29.
[42] Ferreira PM, Faria EA, Ruano AE. Neural network models in greenhouse air temperature prediction. Neurocomputing 2002;43:51–75.
[43] Ehret DL, Hill BD, Helmer T, Edwards DR. Neural network modeling of greenhouse tomato yield, growth and water use from automated crop monitoring data. Comput Electron Agric 2011;79:82–9.
[44] Taki M, Rohani A, Soheili-Fard F, Abdeshahi A. Assessment of energy consumption and modeling of output energy for wheat production by neural network (MLP and RBF) and Gaussian process regression (GPR) models. J Clean Prod, doi: 10.1016/j.jclepro.2017b.11.107.