Article

Application of Artificial Intelligence for Modeling the Internal Environment Condition of Polyethylene Greenhouses

Elham Bolandnazar [1], Hassan Sadrnia [1], Abbas Rohani [1], Francesco Marinello [2] and Morteza Taki [3,*]

1 Department of Biosystems Engineering, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad 91779-48978, Iran; arohani@um.ac.ir (A.R.)
2 Department of Land, Environment, Agriculture and Forestry, University of Padova, 36100 Vicenza, Italy; francesco.marinello@unipd.it
3 Department of Agricultural Machinery and Mechanization, Faculty of Agricultural Engineering and Rural Development, Agricultural Sciences and Natural Resources University of Khuzestan, Mollasani 63417-73637, Iran
* Correspondence: mtaki@asnrukh.ac.ir

Abstract: Accurate temperature prediction and modeling are critical for effective management of agricultural greenhouses. By optimizing control and minimizing energy waste, farmers can maintain optimal environmental conditions, leading to improved crop yields and reduced financial losses. In this study, multiple models, including Multiple Linear Regression (MLR), Radial Basis Function (RBF), and Support Vector Machine (SVM), were compared to predict greenhouse air temperature. External parameters, such as air temperature (T_{out}), relative humidity (H_{out}), wind speed (W), and solar radiation (S), were used as inputs for these models, and the output was the inside temperature. The results showed that the RBF model with the LM (Levenberg–Marquardt) learning algorithm outperformed the other models, achieving the lowest error and the highest coefficient of determination (R^2) value. The RBF model produced RMSE, MAPE, and R^2 values of 1.32 °C, 3.23%, and 0.931, respectively. These results demonstrate that the RBF model with the LM learning algorithm can reliably predict greenhouse air temperatures for the next two hours. The ANN model can be applied to optimize time management and reduce energy losses, improving the overall efficiency of greenhouse operations.

Keywords: greenhouse temperature; polyethylene cover; prediction; soft-computing; reliability

1. Introduction

Greenhouse air temperature is a critical factor that directly impacts crop yield. Adverse temperature conditions can lead to fungal diseases and even plant death [1]. To overcome these issues, it is essential to employ efficient, low-cost, and reliable heating or cooling technologies. During the summer, optimal internal conditions can be achieved by using the shade of covering materials, natural or hybrid ventilation with roof or side openings, advanced cooling systems (such as direct or indirect evaporative coolers), and better utilization of alternative energy sources [2]. These measures can help to maintain optimal environmental conditions within the greenhouse, leading to improved crop yields and reduced losses due to adverse temperature conditions. By adopting these alternative heating or cooling technologies, farmers can achieve a more sustainable and efficient greenhouse operation [3].

Solar radiation and the crops inside the greenhouse are two critical factors that significantly affect the energy balance of the greenhouse. The energy balance of the greenhouse is a crucial aspect of maintaining optimal growing conditions for plants and ensuring efficient energy use [4]. Solar radiation is the primary source of energy for greenhouses. The amount of solar radiation that enters the greenhouse depends on various factors, such as the orientation and design of the greenhouse, the latitude of the location, and weather

conditions [5]. The solar radiation that enters the greenhouse is absorbed by the crops, soil, and other surfaces inside the greenhouse, leading to an increase in the internal temperature. The absorption of solar radiation can be affected by the transmissivity of the greenhouse cover material, the shading provided by the crops, and the orientation of the greenhouse [6].

The crops inside the greenhouse also play a crucial role in the energy balance of the greenhouse. The crops absorb solar radiation and convert it into energy through photosynthesis. This energy is then used for growth, respiration, and other metabolic processes. The crops also release water vapor through transpiration, which affects the humidity and temperature inside the greenhouse. The crop canopy can also affect the distribution of solar radiation inside the greenhouse, leading to variations in the temperature and growth of different parts of the crop [7]. To maintain optimal growing conditions for plants and ensure efficient energy use, it is essential to balance the amount of solar radiation entering the greenhouse with the energy consumed by the crops and the greenhouse itself. This balance can be achieved through various methods, such as adjusting the orientation and design of the greenhouse, using shading materials, and optimizing the use of heating and cooling systems [3].

Real-time prediction of greenhouse temperatures a few hours ahead can provide farmers with enough time to adopt measures to maintain optimal environmental conditions for their plants, even using low-cost technologies [1]. Therefore, an accurate and fast temperature forecasting method is essential for greenhouse temperature control. Prediction models for controlling the greenhouse environment are mainly categorized into two groups: physical and black box models [8]. Physical models rely on mathematical methods, while black box models use modern computational techniques to make predictions [9]. Both models are complex and have many parameters, but physical models require parameter values to be specified and can be beneficial for climate control purposes [10], while minimizing the digitization footprint [11]. By utilizing these models, farmers can improve their greenhouse management practices and maintain optimal environmental conditions for their plants, leading to improved crop yields and reduced financial losses.

Several researchers have applied artificial intelligence methods to predict internal environmental variables in greenhouses. Eddine et al. [12] utilized an Artificial Neural Network (ANN) and an adaptive Neuro-fuzzy Inference System (ANFIS) for temperature prediction, CO_2 concentration, humidity, and internal radiation. Yu et al. [1] used the Least Squares Support Vector Machine (LSSVM) for inside-environment prediction. Li et al. [13] employed Partial Least-Square Regression (PLSR) and Backpropagation Neural Network (BPNN) for inside temperature and humidity prediction. Singh and Tiwari [14] utilized ANN for relative humidity and air temperature prediction. Reyes-Rosas et al. [15] developed dynamic models for plant, air, greenhouse cover, and soil temperature. Yue et al. [16] used the ANN (RBF model) for the foresight of air temperature and humidity. Özden et al. [17] utilized ANN for the anticipation of energy consumption in temperature control. Chen et al. [18], employed a PSO-based robust MPC (Model Predictive Control) scheme for the prediction of air temperature. Shojaei et al. [19] used the ANN (MLR model) for air temperature prediction. Finally, Taki et al. [20] developed a dynamic and ANN model for air and roof temperature prediction, while Taki et al. [3] employed ANN and SVM for soil, air, and plant temperature prediction. These studies demonstrate the potential of artificial intelligence methods for predicting internal environmental variables in greenhouses, which can lead to improved crop yields and reduced financial losses.

Jiroft city in Kerman province, Iran, is known for its exceptional climatic diversity and fertile agricultural regions, including a thriving greenhouse production industry. The southeast of Kerman Province is one of the most fertile regions in Iran and has a significant number of greenhouse cucumber farmers, producing 280,105 tons within a 14,075 ha area. Due to the high variability of climate in the area, it is essential to evaluate and control greenhouse internal variables. However, there is a lack of research on temperature forecasting based on upcoming hours in the region. Thus, this study aims to investigate the prediction of greenhouse indoor temperature using MLR, RBF, and SVM models and select
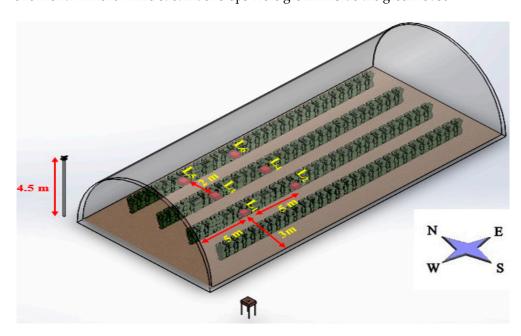
the best one for implementation. Furthermore, the study aims to evaluate the performance of the selected model in forecasting temperatures for future hours, particularly during seasons when plant stress due to warming temperatures is a concern. By addressing these research gaps, this study can contribute to improved greenhouse management practices in Jiroft City, leading to improved crop yields and reduced financial losses.

2. Materials and Methods

2.1. Experimental Data

The data for this study were collected from a greenhouse located in Jiroft, Kerman province, Iran, which is situated in the Southeast of Iran at 28°40′ latitude and 57°44′ longitude. The greenhouse was covered with polyethylene plastic and utilized a natural ventilation system by opening and closing roof and wall windows during hot and dry conditions. An evaporative cooling system (fan-pad) was used to reduce the temperature and increase humidity within the greenhouse when outside air temperatures exceeded acceptable levels for the crops. During colder seasons, gas and diesel heaters were used to increase the internal greenhouse temperature. The study collected internal and external climate parameters, including air temperature, relative humidity, wind speed, and solar radiation, using sensors and related devices. The AM2303 digital temperature and humidity sensors were used to capture temperature and humidity data with a precision of 0.1 °C and 1.2%, respectively. Radiation data were collected using the TES132 solar power meter, which captures radiation in the wavelength range of 400 to 1100 nm with a precision of 5%. Wind speed data were captured using the DT186 anemometer with a storage capacity of 32,000 data points in the range of 1.1 to 20 ms^{-1}. The wind speed outside the greenhouse was also measured using an anemometer at a height of 1.5 m above the ground. Figure 1 shows the sensor locations within the greenhouse. The data were collected from 8 am to 4 pm with a 5-min interval for one month in 2020 while cucumber crops were grown inside the greenhouse.

Figure 1. Situation of temperature and humidity sensors, a solar meter, and an anemometer in the greenhouse.

2.2. Multiple Linear Regressions (MLR)

This study utilized the MLR model to predict greenhouse temperature based on the independent variables, including outside air temperature (T_{out}), relative humidity (H_{out}), wind speed (W), and solar radiation (S), as inputs, and the inside greenhouse air temperature as the output. The regression models employed to predict indoor temperature

included linear, two-factor interaction (2FI), quadratic, and reduced quadratic models, as demonstrated in previous research [21]:

$$Y = \beta_0 + \sum_{i=1}^{m}\beta_i X_i + \varepsilon \tag{1}$$

$$Y = \beta_0 + \sum_{i=1}^{m}\beta_i X_i + \sum_{i=1}^{m}\sum_{j=2}^{l}\beta_{ij} X_i X_j + \varepsilon \tag{2}$$

$$Y = \beta_0 + \sum_{i=1}^{m}\beta_i X_i + \sum_{i=1}^{m}\sum_{j=2}^{l}\beta_{ij} X_i X_j + \sum_{i=1}^{m}\beta_{ii} X_i^2 + \varepsilon \tag{3}$$

$$Y = \beta_0 + \sum_{i=1}^{m}\beta_i X_i + \sum_{i=1}^{m}\beta_{ii} X_i^2 + \varepsilon \tag{4}$$

where Y represents the dependent variable (greenhouse inside temperature), X_i is the independent variable for i = 1, 2, 3, 4 ($X_1 = W$, $X_2 = S$, $X_3 = T_{out}$, and $X_4 = H_{out}$), β_0 is a constant coefficient, β_i is the linear effect, β_{ii} is the quadratic effect, β_{ij} is the interaction, and ε is the model error.

2.3. Radial Basis Function (RBF) Model

The ANN model is a widely used prediction method that typically consists of at least three layers: input, hidden, and output. The input layer size is equal to the number of model inputs, with an equivalent weight assigned to each input. The hidden layer comprises several neurons that enhance the ANN model's performance by selecting an appropriate number of neurons in this layer. The output layer size is equal to the network output, with only one neuron in this study since the goal is to predict greenhouse temperature [21].

RBF models are advantageous for predicting indoor air temperature in greenhouses because they can capture the nonlinear relationships that exist between input variables and the output temperature. For example, the relationship between temperature and humidity is nonlinear, and RBF models can accurately capture this relationship, which is essential for predicting indoor air temperature in greenhouses. The RBF method is a type of neural network in which each neuron in the hidden layer utilizes a nonlinear activation function. The bias factor is employed to converge the RBF neural network during the training phase and reach the global minimum. The output of the RBF neural network for each input (x) is calculated as per previous research [22]:

$$Y = W^T \Phi = \sum_{j=1}^{L_2} w_{ij}\phi(\|x - c_i\|) \tag{5}$$

where w_{ij} is the weight vector of the connections between the hidden layer and the output, L_2 represents the number of neurons in the hidden layer, c_i is the center of the neurons in the hidden layer, and ϕ is the Gaussian function that is calculated based on [22]:

$$\phi_i(x) = exp\left(-\frac{\|x - c_i\|^2}{\sigma_i^2}\right) \tag{6}$$

where σ_i is the spread parameter. In this study, to select the optimal network structure, the spread parameter changed from 0.4 to 0.7 (SP_1 = 0.4, SP_2 = 0.5, SP_3 = 0.6, and SP_4 = 0.7). Figure 2 shows the structure of the radial basis function model [23].
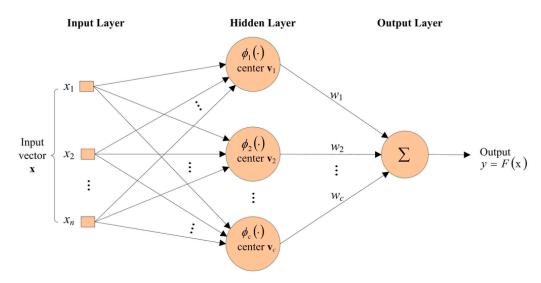
Figure 2. The structure of the RBF model [23].

Training a neural network is a crucial step in achieving significant error in the RBF model. One of the most widely used methods for this purpose is the error back-propagation algorithm, as demonstrated in previous research [20]. In this study, thirteen training functions were employed to train the RBF model using back-propagation training algorithms, as illustrated in Figure 3. These functions were utilized to optimize the weights in the RBF model and improve the accuracy of greenhouse temperature prediction. By utilizing multiple training functions, the study aimed to identify the most effective method for training the RBF model and contribute to improved greenhouse management practices.
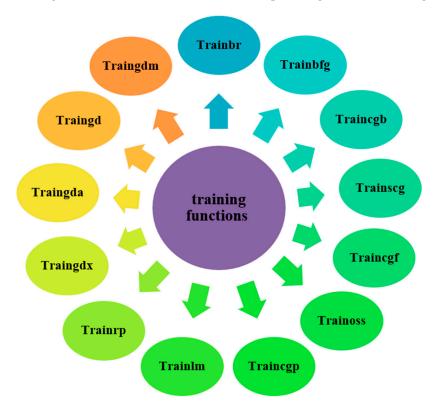

Figure 3. Types of training functions applied in the RBF model.

According to previous research, each optimal continuous function in the ANN model can be approximated with a hidden layer containing a sufficient number of neurons [24]. Therefore, this study utilized a single hidden layer for the RBF model and evaluated its

performance by altering the number of neurons between 3 and 25 in the hidden layer, selecting the optimal structure to estimate greenhouse temperature. Additionally, the literature suggests that linear transfer functions for the output layer in the neural network method can be effectively employed to approximate any complex function [25]. Consequently, this study took advantage of transfer functions for the RBF method in the output layer to improve the accuracy of greenhouse temperature prediction.

2.4. Support Vector Machine (SVM)

The SVM model acts as a proper computational method because it can solve quadratic optimization problems. The SVM method considers the problem of estimating a given dataset $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ based on [26]:

$$f(x) = \langle w, \Phi(x) \rangle + b \tag{7}$$

where, $\langle w, \Phi(x) \rangle$ represents the dot product, $\Phi(x)$ is the nonlinear function that generates regression, and b and w are the bias term and weight vector, respectively. In this study, Equation (8) is used to optimize the problems in the SVM method [3]:

$$\left\{ \begin{array}{c} \min_{w,b,e} J(w,e) = \frac{1}{2}\|w\|^2 + \frac{1}{2}\gamma \sum_{k=1}^{N} e_k^2 \\ s.t. y_k = \langle w, \Phi(x_k) \rangle + b + e_k, k = 1, 2, \ldots, k \end{array} \right\} \tag{8}$$

where γ is an adjusted parameter (also called a penalty parameter, $\gamma \geq 0$) and e_k is the regression error for N training goals. In this method, the Lagrange function was used to solve the optimization problem [3]:

$$L(w, b, e, \alpha) = \frac{1}{2}\|w\|^2 + \frac{1}{2}\gamma \sum_{k=1}^{N} e_k^2 - \sum_{k=1}^{N} \alpha_k \{\langle w, \Phi(x_k) \rangle + b + e_k - y_k\} \tag{9}$$

where α_k represents the Lagrange multiplier, and the sample whose Lagrange multiplier is not equal to *0* is the support vector. Having w, b, e_k, and α_k, we can partially specify the solution of Equation (9) [3]. One of the main advantages of SVM models for predicting indoor air temperature in greenhouses is their ability to handle high-dimensional data. Greenhouse environments can be highly complex, with many variables affecting indoor air temperature. SVM models can handle this complexity by using kernel functions to transform the input data into a higher-dimensional space, where it can be separated into different classes more easily. In this study, MATLAB software version 2015b was used to analyze different methods.

2.5. Normalization and K-Fold Cross Validation Method

As per previous studies, it is recommended to normalize data over the range [0, 1] before conducting any training process with data in various domains or ranges [3]. One of the challenges of prediction models is producing different results in each iteration and randomly selecting data. So, proper selection of training data can enhance the neural network's results [26]. In this study, the K-fold cross-validation method was employed to improve prediction and evaluate the stability and generalizability of the models. For this purpose, the K-fold method was implemented using 100 sets, and the best set of data was chosen for predicting greenhouse temperature. Figure 4 illustrates the steps for applying the K-fold method in this study. By utilizing these methods, the study aimed to improve the accuracy of greenhouse temperature prediction and contribute to improved greenhouse management practices.

Full data

Partition

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	
Model copy # 1	Validation	Training				Predictions for Fold 1
Model copy #2	Training	Validation	Training			Predictions for Fold 2
Model copy #3	Training	Validation	Training			Predictions for Fold 3
Model copy #4	Training	Validation	Training			Predictions for Fold 4
Model copy #5	Training			Validation		Predictions for Fold 5

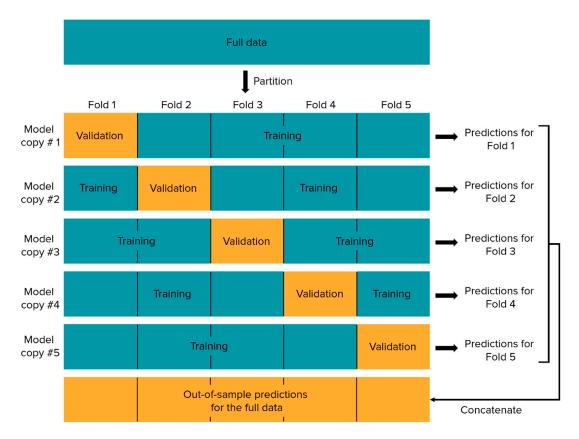Out-of-sample predictions for the full data

Concatenate

Figure 4. The steps for applying the K-fold method.

2.6. Performance Evaluation Criteria

In this study, various performance criteria, including Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and coefficient of determination (R^2), were employed to evaluate the performance of MLR, RBF, and SVM methods in predicting indoor temperature in a greenhouse, as per previous research [27]:

$$MAPE = \frac{1}{n}\sum_{j=1}^{n}\left|\frac{Td_j - Tp_j}{Td_j}\right| \times 100 \tag{10}$$

$$RMSE = \sqrt{\frac{\sum_{j=1}^{n}\left(Td_j - Tp_j\right)^2}{n}} \tag{11}$$

$$R^2 = \left[\frac{\sum_{j=1}^{n}\left(Td_j - \overline{Td}\right)\left(Tp_j - \overline{Tp}\right)}{\sum_{j=1}^{n}\left(Td_j - \overline{Td}\right) \times \sum_{j=1}^{n}\left(Tp_j - \overline{Tp}\right)}\right]^2 \tag{12}$$

where n is the number of observational data, Tp_j is the temperature value predicted by the models used for the prediction, Td_j is the temperature value measured in the greenhouse, and $\overline{Td}$ and $\overline{Tp}$ are the average values of the whole desired and predicted temperatures (Taki et al., 2016b). In this study, when the MAPE and RMSE are minimum and R^2 is maximum, the model gives the highest accuracy.

3. Results and Discussion

3.1. Sensitivity Analysis: Selecting the Effective Inputs and Best Data Set for Training

Table 1 presents the correlation coefficients between the output and all inputs, indicating a complex dynamic relationship between all inputs except wind speed. The results suggest that all inputs, except wind speed, are significantly correlated. The impact of

changes in the quantities of each input on the variation of others is also demonstrated in Table 1, with positive and negative correlations indicating direct and inverse effects, respectively. Notably, relative humidity inside the greenhouse exhibits a negative correlation with other inputs, meaning that an increase in air temperature leads to a decrease in relative humidity. Additionally, changes in solar radiation can directly and inversely affect temperature and humidity, respectively.

Table 1. Correlation coefficient between all inputs and outputs.

	W	**S**	$\mathbf{T_{out}}$	$\mathbf{H_{out}}$
S	−0.130 **			
T_{out}	0.025 [ns]	0.643 **		
H_{out}	−0.097 *	−0.487 **	−0.723 **	
T_{in}	0.001 [ns]	0.685 **	0.799 **	−0.633 **

** Correlation is significant at the 0.01 level; * Correlation is significant at the 0.05 level; [ns] is not significant.

Due to the high cost of sensors and the need for applications with fewer inputs, the sensitivity analysis method was utilized for the MLR model in this study. The best set of input variables for temperature estimation by the MLR model was selected, and the results were employed as the basis for input variable selection in other methods used in this research, such as the RBF and SVM models, to predict greenhouse temperature. As illustrated in Table 2, wind speed (W) produced the poorest results, with the highest RMSE and MAPE observed when wind speed was considered as an input variable. This outcome suggests that the airflow was calm during the data collection period and there was no risk of convection. Therefore, the best combination of data from sensors, including the solar meter, temperature and humidity sensors, and humidity sensor, was utilized as inputs for all prediction models. Based on the results presented in Table 2, the input variables S, T_{out}, and H_{out} were utilized for temperature estimation by the MLR, RBF, and SVM models.

Table 2. Selection of the best group of variables by the MLR model.

Parameter Model Criteria	**RMSE**	**MAPE**
W	3.92	11.34
S	3.72	10.82
T_{out}	2.96	8.25
H_{out}	3.85	11.20
S, T_{out}	2.65	7.12
S, H_{out}	3.24	9.39
T_{out}, H_{out}	3.04	8.44
S, T_{out}, H_{out}	2.55	6.87

To enhance the evaluation and generalizability of the MLR and RBF models and prevent over-fitting, the dataset was partitioned into several strategies. Partitioning strategies, such as 80–20%, 70–30%, and 60–40%, were employed to train and test the models, with the best strategy selected based on the lowest RMSE and MAPE and the highest R^2, as shown in Table 3. It reveals that the forecast error of the MLR and RBF models was affected by the percentage of data allocated for training and testing. Additionally, reducing the amount of data for training resulted in increased error rates for both models. Therefore, the optimal choice for training and testing the datasets was found to be 80–20%, which resulted in the lowest RMSE and MAPE values. All subsequent sections of this study are based on this selection, with 80% of the data used for training and 20% for testing. It is worth noting that the number of epochs was fixed at 500 and may be adjusted as needed.

Table 3. The results of selecting the best data set for training and testing the MLR and RBF models.

Partition Strategy	MLR		RBF	
	RMSE	MAPE	RMSE	MAPE
80%, 20%	2.55	6.87	1.32	3.23
70%, 30%	2.56	6.87	1.40	3.50
60%, 40%	2.57	6.86	1.52	3.69

3.2. MLR Model Results

Once the best variable was identified and the optimal percentage of data for training and testing was determined for the MLR model, different forms of the MLR method were assessed, including linear, interaction, pure quadratic, and quadratic, to identify the most effective model for temperature prediction. Table 4 summarizes the RMSE and MAPE values for each model, with the quadratic model demonstrating the best performance, with RMSE and MAPE values of 2.55 and 6.87, respectively. The regression equations for the selected variables (S, T_{out}, and H_{out}) using the MLR model are presented in Table 5. In a related study, Taki et al. [20] used the MLR model to forecast air and roof temperatures in a semi-solar greenhouse, with RMSE values of 1.14 °C and 1.18 °C, respectively, for temperature prediction.

Table 4. The result of evaluating various MLR models.

Parameter Model Criteria	RMSE	MAPE
Linear	2.82	7.93
2FI	2.66	7.17
Red. Quad	2.65	7.12
Quad	2.55	6.87

Table 5. Regression model equation for temperature prediction.

Parameter	$Y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{23} x_2 x_3$				
	α	β_1	β_2	β_3	β_{23}
T_{in}	18.22	−0.02	0.30	−0.17	0.01

The ANOVA results for the MLR quadratic model are displayed in Table 6. All model factors were significant at the 1% level. The results also indicate that R^2 and R^2_{adj} were nearly equal. This indicates that the calibrated regression model was able to explain temperature variations and that the model inputs were sufficient to estimate greenhouse temperatures, with the model being highly accurate.

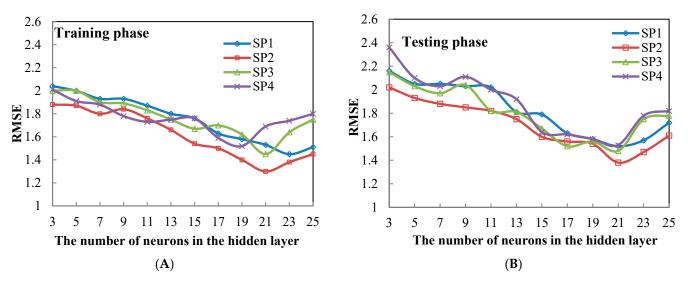**Table 6.** Analysis of variance (ANOVA) for the MLR regression model.

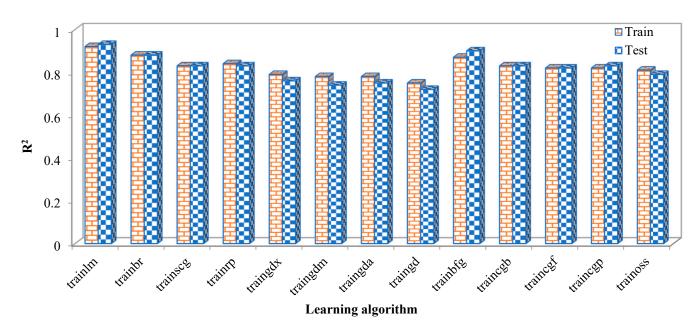|	DF	SS	PC	MS	F
Model	5	2419.38	80.99	483.88	80.95 **
x_1	1	1652.93	55.33	1652.93	276.53 **
x_2	1	581.88	19.48	581.88	97.34 **
x_3	1	14.55	0.49	14.55	2.43 **
$x_2 x_3$	1	24.77	0.83	24.77	4.14 **
x_1^2	1	145.24	4.86	145.24	24.3 **
Error	95	567.86	19.01	5.98	
Total	100	2987.24	100	29.87	
$R^2_{adj} = 0.72$					

** Correlation is significant at the 0.01 level.

3.3. The Results of RBF Model

The spread parameter and number of neurons in the hidden layer are two important factors in the RBF model that must be optimally adjusted for optimal prediction [28]. Given the importance of these two factors, this study selected the optimum network structure in the RBF model by changing the two parameters and their effect on the prediction error in two training and testing phases. For selecting the optimal network structure, the spread parameter was changed from 0.4 to 0.7 (SP_1 = 0.4, SP_2 = 0.5, SP_3 = 0.6, and SP_4 = 0.7), and the number of neurons increased from 3 to 25. Then, the effect of changing the two parameters on the forecast error was investigated (Figure 5). As it can be seen, with the spread parameter equal to 0.5 and 21 neurons in the hidden layer, the least error was obtained for the prediction of greenhouse temperature in both the training and testing phases (RMSE values equal to 1.3 °C for training and 1.38 °C for testing phases). Thus, the most optimal network structure in the RBF model for the prediction of greenhouse temperature was selected as 3-21-1. In this study, 13 different algorithms were employed to train the network and adjust weights until achieving an acceptable error, as depicted in Figure 6. Among the 13 learning algorithms used for the RBF model, the LM (Levenberg-Marquardt) algorithm demonstrated superior performance for temperature modeling, with a coefficient of determination (R^2) value of approximately 0.93. Similarly, in a related study by Taki et al. [21], various training algorithms were utilized to train the RBF model for predicting air, plant, and soil temperatures, with the Trainbr algorithm yielding the lowest MAPE.

The spread parameter and number of neurons in the hidden layer are crucial factors in the RBF model that must be optimally adjusted for accurate prediction [28]. Given the importance of these two parameters, this study selected the optimal network structure in the RBF model by altering the two parameters and examining their impact on prediction error in two training and testing phases. So, the spread parameter was changed from 0.4 to 0.7 (SP$_1$ = 0.4, SP$_2$ = 0.5, SP$_3$ = 0.6, and SP$_4$ = 0.7) and the number of neurons increased from 3 to 25 to determine the optimal network structure. Then, the effect of changing the two parameters on the forecast error was evaluated, as shown in Figure 6. The results revealed that the least error was obtained for predicting greenhouse temperature in both the training and testing phases when the spread parameter was set to 0.5 and the number of neurons was 21 (RMSE value of 1.3 °C for training and 1.38 °C for testing phases). As a result, the optimal network structure in the RBF model for predicting greenhouse temperatures was determined to be 3-21-1.

Figure 5. The relationship between the number of neurons in the hidden layer based on RMSE in the RBF model at the training (**A**) and testing (**B**) phases.

Figure 6. The relationship between all types of learning algorithms based on the R^2 values in the RBF model.
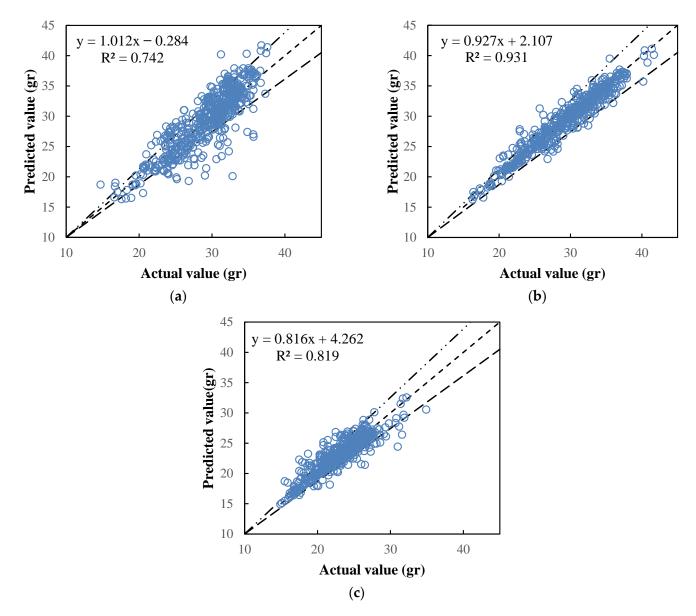
3.4. SVM Model Results

In this study, the SVM model was utilized for temperature prediction. The performance of the SVM model was found to be dependent on the selection of an appropriate kernel function. We evaluated the performance of four types of kernel functions, including linear, Gaussian, RBF, and polynomial, as shown in Table 7. The RBF kernel function demonstrated superior performance compared to the other functions. Similarly, in a related study by Taki et al. [3], various types of kernel functions were employed to predict air, plant, and soil temperatures, with the RBF kernel function yielding the best results for temperature prediction.

Table 7. The result of evaluating the four types of Kernel functions in the SVM model.

The Types of Kernel Function	RMSE (°C)	MAPE (%)
linear	2.81	7.69
Gaussian	2.51	6.54
rbf	2.20	5.36
polynomial	2.51	6.54

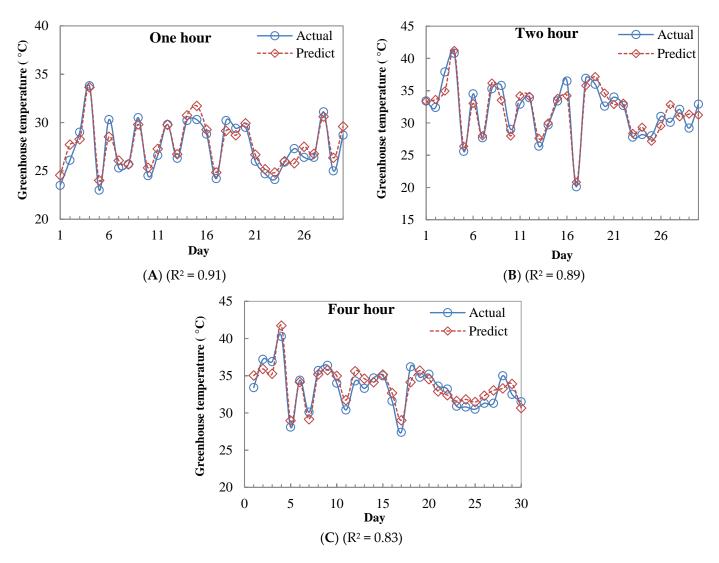### 3.5. The Performance Comparison of the Soft Computing Models

In this section, the criterion of agreement between predicted and experimental values was used to select the best prediction model. The evaluation was based on the results of the optimum parameters in the MLR, RBF, and SVM models. Figure 7 illustrates the distribution of actual and predicted data for these models, indicating that the majority of data points fell within two lines of ±10%. The best result was achieved when the regression line's slope and width from the origin were close to one and zero, respectively. Additionally, a coefficient of determination closer to one indicated that a higher percentage of the overall variations in the predicted temperature could be explained by the linear relationship between predicted and actual temperature (as described by the regression equation). Based on the results, the RBF model demonstrated acceptable accuracy for temperature prediction, with a determination coefficient (R^2) of 0.93. In a related study by Taki et al. [29], MLP, RBF, SVM, ANFIS, and MLR models with K-fold cross-validation were developed to predict global solar radiation. The comparison between predicted and actual results revealed that the RBF methodology performed with higher accuracy in estimating solar radiation.

(**a**)

(**b**)

(**c**)

Figure 7. Comparison between the desired and predicted values by the MLR (**a**), RBF (**b**), and SVM (**c**) models.

3.6. Evaluation of the Best Topology for RBF Model Performance

In this study, in addition to choosing the best model for temperature prediction, early estimation of internal greenhouse temperature was investigated. Such an approach could provide farmers with sufficient time to prevent temperature rises by implementing temperature control measures and ensuring favorable conditions for plant growth. After identifying the best model for temperature estimation (RBF), we evaluated its performance for 30 days with predictions of one, two, and four hours, respectively (model outputs). The RBF model's structure used to predict greenhouse temperature at different time intervals is displayed in Figure 8. The results indicated an increase in the error between predicted and measured values with increasing time intervals. Therefore, based on the findings of this research, predicting greenhouse temperatures for the next two hours could facilitate proper and timely management, reduce energy loss, and mitigate environmental problems.

Figure 8. The results of the RBF model to predict the greenhouse temperature in the next hour (**A**), next two hours (**B**), and next four hours (**C**).

4. Conclusions

The objective of this study was to predict indoor greenhouse temperatures using MLR, RBF, and SVM models. The key findings of this study are presented as follows:

1. A sensitivity analysis (MLR model) was conducted to identify the variable that contributed the most to the prediction error. The results indicated that using wind speed (W) as an input variable could increase the error in temperature prediction. Therefore, this variable was omitted from all models used for temperature prediction;

2. Model comparison revealed that the RBF model was the most suitable model for temperature prediction in conventional greenhouses, with RMSE values of 1.3 °C (training phase) and 1.38 °C (testing phase);

3. Performance evaluation of the RBF model showed that it can predict greenhouse temperature for the next two hours with acceptable accuracy;

4. The findings of this research demonstrated that the application of ANN models can aid farmers in building smart greenhouses, managing time, reducing plant stress and energy consumption, and increasing crop yield in Jiroft City in the future;

5. This study aimed to investigate the effectiveness of using artificial intelligence for estimating internal variables in greenhouses. However, to develop a more robust model for smart greenhouses, it is important to collect data over several years. Additionally, future studies should consider the potential impact of climate change on

weather conditions and how it may affect the internal environment of the greenhouse. By considering long-term weather patterns and climate change, more accurate models can be developed to effectively predict internal variables in smart greenhouses. Ultimately, this will lead to more efficient resource use and improved plant growth.

Nomenclature

H_{out}	Relative humidity of the air outside the greenhouse (%)
T_{out}	The temperature of the air outside the greenhouse (°C)
Tin	The temperature of the air inside the greenhouse (°C)
W	Wind speed outside the greenhouse (ms^{-1})
S	Solar radiation (Wm^{-2})
MSE	Mean Square Error
RMSE	Root Mean Square Error
MAPE	Mean Absolute Percentage Error
R^2	Coefficient of determination
ANN	Artificial Neural Network
MLR	Multiple Linear Regression
RBF	Radial Basis Function
SVM	Support Vector Machine
ANFIS	An adaptive neuro-fuzzy inference system
SP	Spread Parameter
Trainbr	Bayesian regularization back-propagation
Trainbfg	BFGS quasi-Newton back-propagation
Traincgb	Powell-Beale conjugate gradient back-propagation
Trainscg	Scaled conjugate gradient back-propagation
Traincgf	Fletcher-Powell conjugate gradient back-propagation
Trainoss	One step secant back-propagation
Traincgp	Polak-Ribiere conjugate gradient back-propagation
Trainlm	Levenberg-Marquardt back-propagation
Trainrp	Resilient back-propagation
Traingdx	Gradient descent w/momentum and adaptive back-propagation
Traingda	Gradient descent with adaptive back-propagation
Traingdm	Gradient descent with momentum back-propagation
Traingd	Gradient descent back-propagation

References

1. Yu, H.; Chen, Y.; Hassan, S.G.; Li, D. Prediction of the temperature in a Chinese solar greenhouse based on LSSVM optimized by improved PSO. *Comput. Electron. Agric.* **2016**, *122*, 94–102. [CrossRef]
2. Gadhesaria, G.; Desai, C.; Bhatt, R.; Salah, B. Thermal analysis and experimental validation of environmental condition inside greenhouse in tropicalwet and dry climate. *Sustainability* **2020**, *12*, 8171. [CrossRef]
3. Taki, M.; Abdanan Mehdizadeh, S.; Rohani, A.; Rahnama, M.; Rahmati-Joneidabad, M. Applied machine learning in greenhouse simulation; new application and analysis. *Inf. Process. Agric.* **2018**, *5*, 253–268. [CrossRef]
4. Cossu, M.; Sirigu, A.; Deligios, P.A.; Farci, R.; Carboni, G.; Urracci, G.; Ledda, L. Yield Response and Physiological Adaptation of Green Bean to Photovoltaic Greenhouses. *Front. Plant Sci.* **2021**, *12*, 655851. [CrossRef] [PubMed]
5. Nauta, A.; Tasnim, S.H.; Lubitz, W.D. Simulation of an Earth-Air Heat Exchanger in a Commercial Greenhouse to Improve Energy Efficiency. *J. Biosyst. Eng.* **2023**, 1–18. [CrossRef]
6. Cossu, M.; Tiloca, M.T.; Cossu, A.; Deligios, P.A.; Pala, T.; Ledda, L. Increasing the agricultural sustainability of closed agrivoltaic systems with the integration of vertical farming: A case study on baby-leaf lettuce. *Appl. Energy* **2023**, *344*, 121278. [CrossRef]
7. Nauta, A.; Han, J.; Tasnim, S.H.; Lubitz, W.D. Performance Evaluation of a Commercial Greenhouse in Canada Using Dehumidification Technologies and LED Lighting: A Modeling Study. *Energies* **2023**, *16*, 1015. [CrossRef]
8. Stanghellini, C.; de Jong, T. A model of humidity and its applications in a greenhouse. *Agric. For. Meteorol.* **1995**, *76*, 129–148. [CrossRef]
9. Van Beveren, P.J.M.; Bontsema, J.; Van Straten, G.; Van Henten, E.J. Minimal heating and cooling in a modern rose greenhouse. *Appl. Energy* **2015**, *137*, 97–109. [CrossRef]
10. He, F.; Ma, C. Modeling greenhouse air humidity by means of artificial neural network and principal component analysis. *Comput. Electron. Agric.* **2010**, *71*, 129–148. [CrossRef]
11. Kayad, A.; Sozzi, M.; Paraforos, D.S.; Rodrigues, F.A.; Cohen, Y.; Fountas, S.; Medel-Jimenez, F.; Pezzuolo, A.; Grigolato, S.; Marinello, F. How many gigabytes per hectare are available in the digital agriculture era? A digitization footprint estimation. *Comput. Electron. Agric.* **2022**, *198*, 107080. [CrossRef]
12. Eddine, C.; Mansouri, K.; Mourad, M.; Belmeguenai, A. Adaptive Neuro-Fuzzy Inference Systems for Modeling Greenhouse Climate. *Int. J. Adv. Comput. Sci. Appl.* **2016**, *7*, 129–148. [CrossRef]
13. Li, X.F.; Qin, L.L.; Ma, G.Q.; Wu, G. Modeling greenhouse temperature by means of PLSR and BPNN. In Proceedings of the 2016 35th Chinese Control Conference, CCC, Chengdu, China, 27–29 July 2016; pp. 2196–2200. [CrossRef]
14. Singh, V.K.; Tiwari, K.N. Prediction of greenhouse micro-climate using artificial neural network. *Appl. Ecol. Environ. Res.* **2017**, *15*, 767–778. [CrossRef]
15. Reyes-Rosas, A.; Molina-Aiz, F.D.; Valera, D.L.; López, A.; Khamkure, S. Development of a single energy balance model for prediction of temperatures inside a naturally ventilated greenhouse with polypropylene soil mulch. *Comput. Electron. Agric.* **2017**, *142*, 9–28. [CrossRef]
16. Yue, Y.; Quan, J.; Zhao, H.; Wang, H. The prediction of greenhouse temperature and humidity based on LM-RBF network. In Proceedings of the 2018 IEEE International Conference on Mechatronics and Automation, Changchun, China, 5–8 August 2018. [CrossRef]
17. Özden, S.; Dursun, M.; Aksöz, A.; Saygın, A. Prediction and Modelling of Energy Consumption on Temperature Control for Greenhouses. *J. Polytech.* **2018**, *76*, 129–148. [CrossRef]
18. Chen, L.; Du, S.; He, Y.; Liang, M.; Xu, D. Robust model predictive control for greenhouse temperature based on particle swarm optimization. *Inf. Process. Agric.* **2018**, *5*, 329–338. [CrossRef]
19. Shojaei, M.H.; Mortezapour, H.; Jafari Naeimi, K.; Maharlooei, M.M. Temperature Prediction of a Greenhouse Equipped with Evaporative Cooling System Using Regression Models and Artificial Neural Network (Case Study in Kerman City). *Iran. J. Biosyst. Eng.* **2019**, *49*, 567–576.
20. Taki, M.; Ajabshirchi, Y.; Ranjbar, S.F.; 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*, 314–329. [CrossRef]
21. Taki, M.; Ajabshirchi, Y.; Ranjbar, S.F.; Rohani, A.; Matloobi, M. Modeling and experimental validation of heat transfer and energy consumption in an innovative greenhouse structure. *Inf. Process. Agric.* **2016**, *3*, 157–174. [CrossRef]
22. 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* **2018**, *115*, 411–422. [CrossRef]
23. Fan, Z.-C.; Hwang, W.-J. Efficient VLSI Architecture for Training Radial Basis Function Networks. *Sensors* **2013**, *13*, 3848–3877. [CrossRef] [PubMed]
24. Haykin, S. *Neural Networks: A Comprehensive Foundation*, 2nd ed.; Prentice Hall: Saddle River, NJ, USA, 1998.
25. Hagan, M.T.; Demuth, H.B. *Neural Network Design*; PWS: Boston, MA, USA, 1997.
26. Bolandnazar, E.; Rohani, A.; Taki, M. Energy consumption forecasting in agriculture by artificial intelligence and mathematical models. *Energy Sources Part A Recovery Util. Environ. Eff.* **2020**, *42*, 1618–1632. [CrossRef]
27. Khodabakhshikoulaei, A.; Sadrnia, H.; Tabasizadeh, M.; Ghobadian, B.; Raghavan, V. Bioethanol fuel quality assessment using dielectric spectroscopy. *Biofuels* **2021**, *76*, 1–9. [CrossRef]

28. Tafarroj, M.M.; Kolahan, F. Using an optimized RBF neural network to predict the out-of-plane welding distortions based on the 3-2-1 locating scheme. *Sci. Iran.* **2019**, *26*, 869–878. [CrossRef]

29. Taki, M.; Rohani, A.; Yildizhan, H. Application of machine learning for solar radiation modeling. *Theor. Appl. Climatol.* **2021**, *143*, 1599–1613. [CrossRef]