



OPEN Temperature prediction of solar greenhouse based on NARX regression neural network

Maosheng Gao^{1,2}, Qingli Wu³, Jianke Li^{1,2}✉, Bailing Wang^{1,2}, Zhongyu Zhou^{1,2}, Chunming Liu¹ & Dong Wang⁴

Temperature has an important influence on plant growth and development. In protected agriculture production, accurate prediction of temperature environment is of great significance. However, due to the time series, nonlinear and multi coupling characteristics of temperature, it is difficult to achieve accurate prediction. We proposed a method for building a solar greenhouse temperature prediction model based on a timeseries analysis, that considers the time series characteristics and dynamic temperature changes in the greenhouse system. The method would predict the temperature of greenhouse, and provide reference for the temperature change law in protected agriculture. A parameter analysis was performed on the nonlinear autoregressive exogenous (NARX) neural network, and a solar greenhouse temperature time series prediction model was established using the NARX regression neural network. The results showed that the proposed model depicted a maximum absolute error of 0.67 °C, and model correlation coefficient of 0.9996. Compared with the wavelet and BP neural networks, the NARX regression neural network accurately predicted and significantly outperformed in the solar greenhouse temperature prediction model. Moreover, the prediction model can accurately predict temperature trends within the solar greenhouse and is pivotal to obtaining precise control of solar greenhouse temperature.

Solar greenhouses are designed to insulate heat and increase production^{1,2}. In recent years, the area of greenhouses in China has continued to expand^{3,4}. As of 2020, greenhouses covered 340 hectares, 90% of which are solar greenhouses located in northern China⁵. Solar green houses are operated by simple, manual processes. In winter, the manager will close the greenhouse based on experience to ensure the indoor temperature⁶. However, because these temperature control decisions are made in the absence of scientific assessment, they are often erroneous or tardy, which leads to delays in shed opening time, negatively impacts the greenhouse's sunlight exposure time in winter, and contributes to a decline ineffective crop photosynthesis^{7,8}. Therefore, to achieve effective greenhouse temperature regulation, it is particularly important to predict the greenhouse temperature in advance⁹.

Solar greenhouses are nonlinear, large inertia, strong coupling, and time-varying complex systems¹⁰. As such, implemented control measures and changes in external climate will impact the greenhouse environment and thus, facilitate changes in the solar greenhouse crops' physiological characteristics¹¹. The modeling methods of greenhouse temperature can be divided into mechanism models based on energy balance and non-mechanism models based on data drive. The mechanism model has the advantages of strong interpretability, but it is difficult to apply to the actual agricultural production due to its low accuracy, poor versatility and difficulties in obtaining parameters. With the development of sensor technology, data acquisition becomes easier and easier. Data driven model is widely used in agriculture because of its high precision and strong generalization ability. Holthuijzen¹² has used five statistical methods to predict temperature. Although statistical methods can predict time series data, their ability to deal with nonlinear problems is limited, and their fitting accuracy is generally low.

To circumvent this issue, many scholars use intelligent algorithms to solve the challenges associated with greenhouse modeling. Examples include but are not limited to: fuzzy control, neural network, genetic, and nonlinear regression algorithms^{13–16}. Among them, the application of neural networks in greenhouse modeling is particularly prominent¹⁷, as modeling process difficulties are greatly reduced due to the neural networks' black box nature¹⁸. Ferreira et al.¹⁹ used radial basis function neural network to predict the greenhouse temperature. The parameters of the model were optimized by Levenberg–Marquardt method. However, the model belongs

¹Meteorological Administration of Yangling City, Yangling 712100, China. ²Yangling High-Tech Agricultural Meteorological Technology Combined Research Center, Yangling 712100, China. ³Shaanxi Meteorological Cadre Training College, Xi'an 710016, China. ⁴College of Mechanical and Electronic Engineering, Northwest Agriculture and Forestry University, Yangling 712100, China. ✉email: 250610861@qq.com

to static neural network and the temperature time series change information is not considered. The long-term prediction accuracy is low.

Considering the time series change characteristics of temperature, we designed a closed loop neural network Nonlinear Auto Regressive Model with Exogenic Inputs (NARX). Compared with the traditional neural network, this network takes the predicted temperature as the feedback input, which belongs to a dynamic model. The closed loop network can predict the temperature according to the previous temperature to improve the accuracy of the model. In order to train the model easily, the error back propagation algorithm was used to train the parameters. In researches of temperature prediction, NARX has dynamic time series characteristics and is convenient for training.

In this work, a dynamic neural network was employed to develop a model that enables solar greenhouse temperature series prediction with time series characteristics. Herein, the time series concept, which is based on traditional neural network modeling, is introduced; and a NARX recurrent neural network (NARXNN) prediction model with two primary external input influencing factors—light and temperature difference—is proposed. Results show that the prediction accuracy and network performance of a NARXNN prediction model based on time series analysis is significantly better than that of other algorithm models, which has practical significance for accurate regulation of greenhouse temperature.

Materials and methods

Testing platform structures. In this work, greenhouse temperature and temperature-related environmental data were experimentally obtained, from December 12–17, 2021. Testing was conducted at Northwest A&F University of Agriculture and Forestry Science and Technology in Wutun Town, Yanliang, Xi'an City (latitude 34° 35' 11" N, longitude 109° 08' 54" E). The experimental greenhouse was 46 m long and 3.8 m high, with 1 m thick east, west, and rear walls. Experimental apparatus consisted of a db130 air temperature and humidity sensor, with a measurement accuracy of ± 0.5 °C at ambient temperature (25 °C); and a DAVIS6450 solar radiation sensor, with a $\pm 5\%$ measurement accuracy at full scale.

Using ZigBee wireless communication and general packet radio service (GPRS) technology, the monitoring node collected the greenhouse environmental factor information, and uploaded the data to the website through the root node and DTU. The test data acquisition system is shown in Fig. 1.

The monitoring node is composed of a power supply module, core processing module, sensor module, and debugging module. All the monitoring nodes are networked to collect and transmit the environmental data from each node—which consists of indoor/outdoor air temperature, humidity, and light radiation data that has been collected every 5 min. The root node is composed of the serial port, power, and core processing modules. Using ZigBee wireless sensor network technology, the latter receives the data collected by each monitoring node and

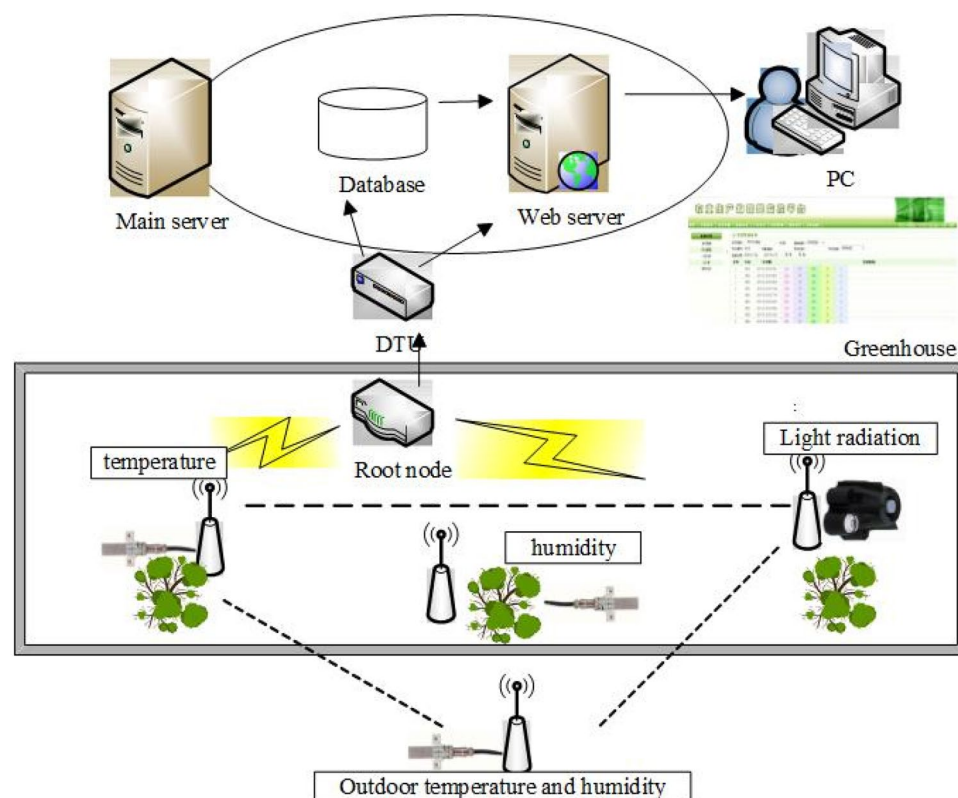


Figure 1. Experimental data acquisition system.

sends it to the DTUs through the RS232 serial port. The central server then uses GPRS technology to have the DTU upload the environmental factor information to the agricultural production Internet of Things monitoring platform. Users can then access the platform online to obtain test data.

Experiment design. The greenhouse remained vacant during the experiment to avoid the influence of crops on indoor temperature changes. The temperature detection values differed as a function of height and distance from the rear wall, so the placement position of each sensor was determined as shown in Fig. 2. Nodes 1, 2, 7, and 8 were 1 m equidistant from the ground; Nodes 2, 3, 8, and 9 were 3.5 m equidistant from the back wall; Nodes 4, 5, and 6 were 0.8 m equidistant from the ground within the same vertical plane; and Nodes 1, 2, and 3 were symmetrical with nodes 7, 8, and 9 relative to the vertical planes of nodes 4, 5, and 6. In addition, Nodes 1–9 were outfitted with air temperature and humidity sensors, while nodes 4, 5, and 6 were outfitted with light radiation sensors. Node 0 was located outdoors, at a minimum distance of 1.5 m from the greenhouse, and a height of 2 m. This node was also outfitted with an air temperature/humidity sensor and light radiation sensor.

The test was conducted during clear weather conditions, such that the low temperatures and absence of light associated with actual extreme weather could be simulated. The greenhouse remained in an opened state from Dec. 24th–25th, during which time the greenhouse quilt remained open at night. The greenhouse was closed before sunset on the evening of the 25th and open after sunrise on the morning of the 27th, during which time the quilt remained closed during the day. The shed was then closed again at sunset on the evening of the 28th. The specific time associated with opening and closing the shed in the morning and evening was managed by the greenhouse manager and performed in conjunction with the greenhouses of other crops. The five-day data collection time spanned from 00:00 on January 24, 2018 to 00:00 on January 30, 2018. After the data was obtained, the average temperature value of each node in the greenhouse was calculated and taken as the indoor temperature data, while the indoor/outdoor temperature difference was calculated and used for indoor/outdoor temperature difference data. After the outliers in the test data were excluded, the sample set consisted of 1440 sets of environmental factors.

Pretreatment of test data. Because temperature, temperature difference, and radiation intensity are characterized using different dimensions, the data could not be directly compared. As such, if the NARXNN model was established using the raw data, the model accuracy would be negatively impacted. Thus, the sample set was first normalized using the normalization mapping formula in Formula (1):

$$y = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

where $x, y \in \mathbb{R}^n$; $x_{\min} = \min(x)$; $x_{\max} = \max(x)$, x is the data to be normalized, and y is the normalized data. Using this formula, the original data was normalized to the range of $[0, 1]$.

To reserve sufficient opening action time and temperature response time, and to ensure prediction accuracy, 20 min was selected as the prediction time. Equal spacing started from the first data point in the first 1440 sets of data. Next, 361 sets of data i.e., every 4th set, was selected to comprise the new sample set. Of the reduced sample set, the first 300 data sets were used as the training set and the last 61 sets as the prediction set.

Greenhouse temperature prediction model based on NARXNN. In this work, a NARXNN time series prediction model with external input was used to establish a corresponding greenhouse temperature prediction model with the aim of resolving the temperature prediction challenges associated with solar greenhouse. The model was developed using static neurons and network feedback, and combined the linear autoregression (ARX) model's nonlinear processing with a neural network, giving it strong nonlinear identification ability.

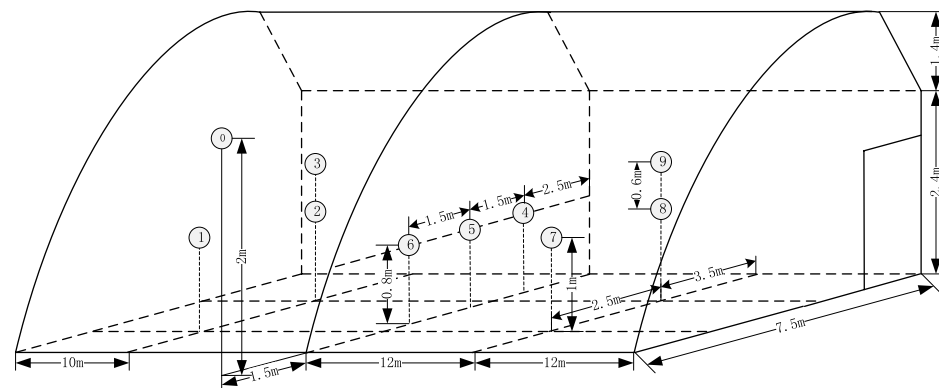


Figure 2. Schematic diagram of the solar greenhouse architecture and sensor distribution.

NARXNN based greenhouse temperature prediction model methodology. Using the sample set described above, a NARX regression neural network prediction model was developed. The greenhouse's historical and current light radiation time and inside/outside temperature difference were used as inputs, while the greenhouse's future temperature at a given time served as the output. A flowchart of the NARXNN algorithm is shown in Fig. 3. First, the appropriate prediction model structure was selected based on the network's input and output characteristics, a nonlinear autoregressive neural network model with external input was constructed, and the network mode was simultaneously established. Next, the time series data was pre-prepared, and the 361 training sets were divided into new training sets, verification sets, and test sets in a 70:15:15ratio. Then, the network parameters were determined and the algorithm was taught to achieve network training. Finally, predictions were made using the trained network.

Network pattern selection. The standard NARXNN operates in Parallel neural network mode²⁰. Given the time series continuity, the prediction output is fed back as one-dimensional input of the next moment to add the historical series influence on the future moment's predicted value. In practice, temperature data is collected and stored in real time, and the indoor temperature of the current and historical moments are known when making predictions. Therefore, a static network was built with the actual historical temperature value as one of the dimensional inputs. This resulted in a Series-Parallel neural network mode that turns the NARXNN into a pure forward neural network. Compared with the traditional BP neural network, the network structure adds a delay order, so the static neural network modeling function can be directly used, which reduces modeling difficulty and improves the network's prediction accuracy to a certain extent. The network pattern is shown in Fig. 4.

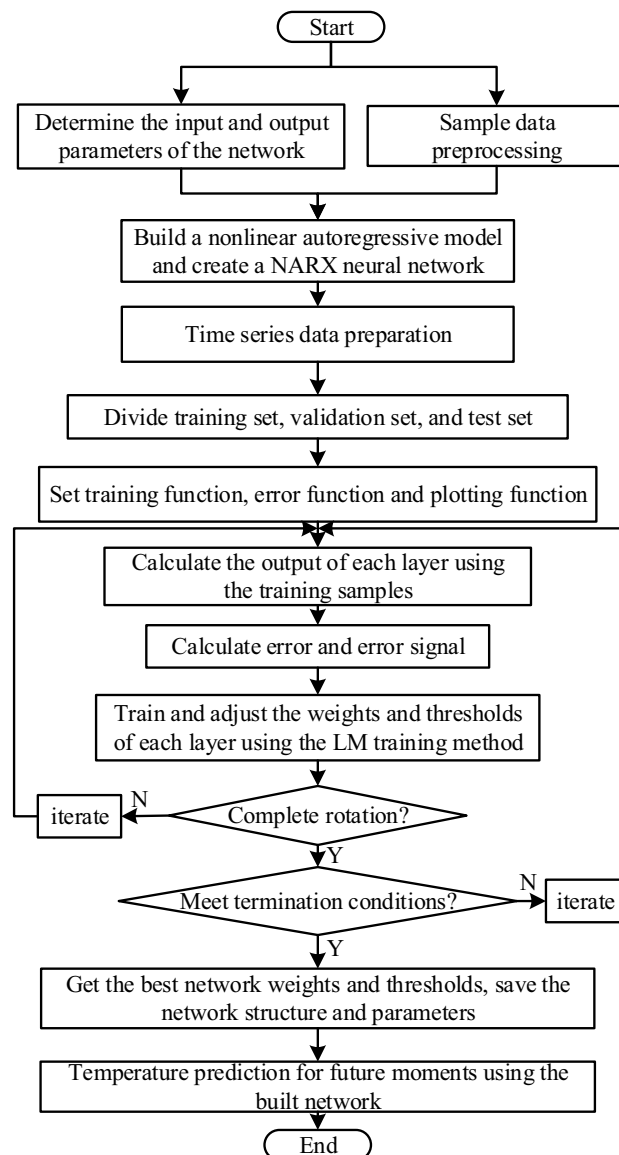


Figure 3. Specific flowchart of the NARXNN algorithm.

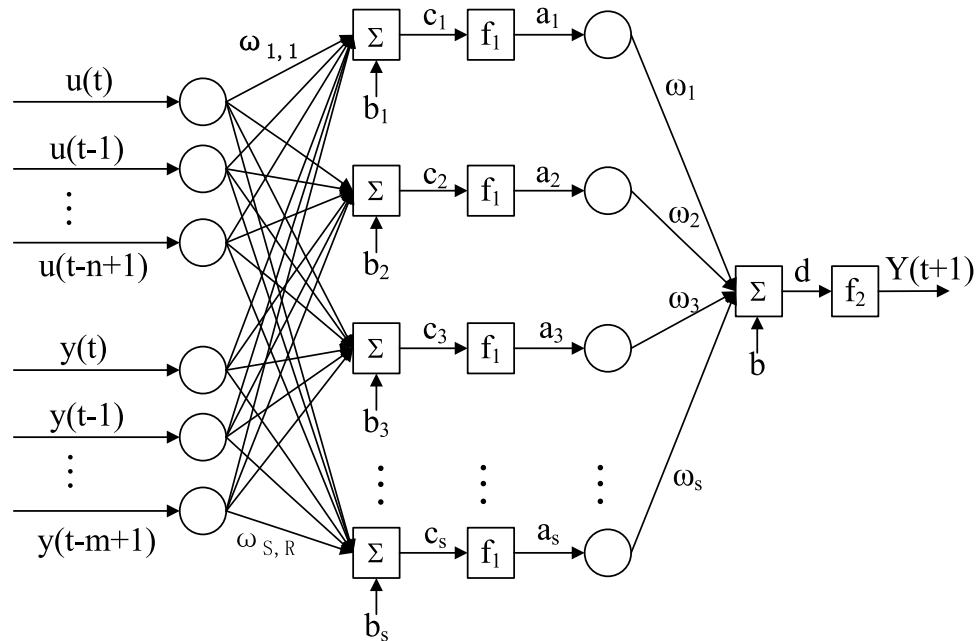


Figure 4. Series parallel neural network model.

In Fig. 4, n and m represents the system's input and output delay orders, respectively; $R = n + m$ and indicates the current network's number of input nodes; ω_{ij} is the connection weight between the i th neuron in the network's hidden layer and the j th element in the input vector; S is the number of hidden nodes; b_i is the i th neuron's bias value in the network's hidden layer; c_i is the net input of the i th neuron in the hidden layer; f_1 is the transfer function of a crypto layer neuron; a_i is the output of the i th neuron in the hidden layer; ω_i is the i th connection weight between the network output layer neurons and the hidden layer neurons; b is the neurons' bias value in the network's output layer; d is the net input of neurons in the output layer; f_2 is the transfer function of neurons at the network's output layer; and $Y(t+1)$ is the network's predicted output.

If the current moment is t , the input–output relationship between the model and the $t+1$ moment temperature value prediction is as follows²¹:

$$Y_{t+1} = \sum_{i=1}^s \omega_i f_1[c_i(t)] + b \quad (2)$$

The vector form is shown in Formula (3):

$$c_i(t) = \sum_{j=1}^n \omega_{i,j} u(t-j+1) + \sum_{j=1}^m \omega_{i,j+n} y(t-j+1) + b_i \quad (3)$$

Formula (2) and Formula (3) are combined to obtain:

$$Y(t+1) = f(u(t), u(t-1), \dots, u(t-n+1), y(t), y(t-1), \dots, y(t-m+1)) \quad (4)$$

where $u(t)$ is the temperature influencing factor term at the current moment, $u(t-n+1)$ is the temperature influencing factor term at the moment of $t-1+1$, $y(t)$ is the indoor temperature value at the current moment, $y(t-m+1)$ is the indoor temperature value at the $t-m+1$ moment, and $Y(t+1)$ is the predicted indoor temperature at the next moment.

NARX regression neural network design. The selected input and output delay orders were $n=3$ and $m=3$ —i.e., there were six input layer nodes.

The number of neurons in the implicit layer is often based on experience, and in turn, has resulted in large network errors. To circumvent this issue, different numbers of neurons were tested for network training; and the number of neurons that produced the best training results was selected (Fig. 5).

The performance function is defined as the training result detection function and is used as the basis for determining whether the network training result is good or bad. As shown in Fig. 5, the NARXNN performance function value depicts an inverse relationship with the number of hidden layer nodes until the latter value reaches 10. At this point, the trend reverses and the two variables depict a positive correlation. While increasing the number of hidden layer nodes makes the network more accurate, it also results in a discordant fit, which affects the network's performance. In this case, the performance value function is maximized when the number

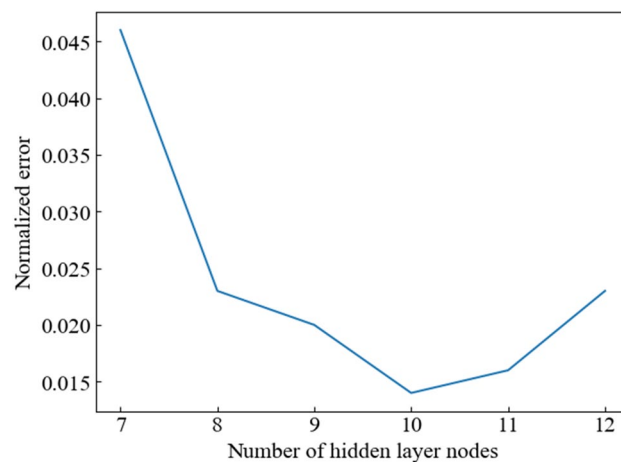


Figure 5. Performance function curve with the number of hidden layer nodes.

of nodes in the hidden layer is 10; thus, 10 was selected as the value for this variable. The prediction result was the temperature value of the next moment, and the number of nodes in the output layer was 1.

In summary, the number of network nodes was as follows: 6 input nodes, 10 hidden layer nodes, and 1 output node.

Neuronal activation functions are often divided into global functions—such as S-type functions, and local functions—such as RBF functions²². Since NARXNN is a global approximation network, its neuronal excitation function should be a global function. Therefore, the influence of different S-type transfer functions on network performance was investigated. Four models training step parameters were analyzed to determine the optimal transfer function combination. In addition, the training set's root mean square error, deterministic coefficient, and training time were tested under different combinations of transfer functions. The results are shown in Table 1.

Table 1 shows that different combinations of transfer functions have an important impact on model training. The model rms error and deterministic coefficient obtained in the sequence numbers 3, 4, 7, 8, 11, 12, and 16 depict superior performance. Among them, group 12 has the smallest root mean square error and the decision coefficient closest to 1, as well as a small training step and short training time. Thus 12 was selected as the optimal transfer function combination. Accordingly, the implicit layer's transfer function was tansig, and the output layer's transfer function was purelin.

There are numerous types of neural network learning algorithms. In this study, the error function and network computation time of the pseudo-Newtonian backpropagation algorithm, gradient descent backpropagation algorithm, Powell-Beale reset algorithm, gradient descent backpropagation algorithm for adaptive adjustment of

Order number	Transfer function		Training step	RMS error	Coefficient of determination	Training time
	Implication level	Output level				
1	radbas	radbas	8	113.0576	0.7486	18.049
2	radbas	logsig	13	137.0880	0.1512	21.977
3	radbas	tansig	13	0.2326	0.9978	22.603
4	radbas	purelin	13	0.0224	0.9998	16.719
5	logsig	radbas	14	142.0744	0.06037	22.967
6	logsig	logsig	10	112.9237	0.7537	23.047
7	logsig	tansig	21	0.9222	0.991	20.620
8	logsig	purelin	7	0.0166	0.9998	18.409
9	tansig	radbas	7	116.1256	0.6604	22.727
10	tansig	logsig	25	112.8936	0.7565	24.207
11	tansig	tansig	19	0.0611	0.9995	23.625
12	tansig	purelin	15	0.0133	0.9999	19.9925
13	purelin	radbas	6	113.5387	0.7458	27.386
14	purelin	logsig	28	113.5672	0.7348	40.981
15	purelin	tansig	5	1.3032	0.9878	23.410
16	purelin	purelin	22	0.0172	0.9998	18.203

Table 1. The transfer function's effect on the mode's prediction accuracy.

learning rate, and Levenberg–Marquardt backpropagation algorithm were assessed and compared. The results are shown in Table 2.

According to Table 2, the gradient descent backpropagation algorithm, which is known by the root mean square error value and the decision coefficient, is not suitable for the data characteristics and was therefore eliminated. Furthermore, the gradient descent backpropagation algorithm with the adaptive adjustment learning rate shows better adaptability, but still lags behind the remaining three algorithms. Comparing the LM backpropagation algorithm, BFGS learning algorithm, and Plwell-Beale reset algorithm, the LM backpropagation algorithm performed best—depicting a coefficient of determination = 0.9999, root mean square error = 0.0133, and requiring only 15 training steps. Considering the indicators that were compared, the LM backpropagation algorithm offered the optimal value, indicating that it can solve the actual problem faster and more accurately than the others. As such, the LM backpropagation algorithm was selected for this study.

In summary, the model used herein depicts the following parameters: network structure (3–10–1); hidden layer activation function (tagsig function); output layer activation function (purelin linear transfer function); training function (trainlm function); and the error performance function (mse function).

Model evaluation. In this work, the measured value and the predicted value were compared and used as the model evaluation benchmark. Moreover, the maximum absolute error, mean absolute error (MAE), root mean square error (MSE), decision coefficient, and calculation time were used as the specific evaluation indicators to determine the model's advantages and disadvantages. The model's feasibility was assessed by determining whether the indicators were within the acceptable range, while the model's quality was evaluated by measuring the distance of each indicator from the best standard.

Results and discussion

Training effectiveness of the NARXNN prediction model. The already established NARXNN prediction model was used to train the study model, and the training effectiveness is shown in Fig. 6. Note that both the training set and test set reached the target error value in the 8th training. In contrast, the verification set reached the target error value after the 15th training, at which time the verification set error was 0.021865. Notably, the verification set error continued to rise, so the network required only 15 iterations to complete the training. These results indicate that the model prediction is characterized by a fast convergence speed and high accuracy.

Feasibility of the NARXNN prediction model. A predictive model's feasibility is usually determined by autocorrelation and cross-correlation tests. When the prediction error is obviously autocorrelated, and the input and error have obvious cross-correlation, then the prediction error is dependent on the time series, as well

Predicted models	Training step	R ²	RMSE	Operation time
trainlm	15	0.9999	0.0133	19.9925
trainbfg	29	0.9995	0.0487	20.271
traingd	0	27.3047	625.6171	14.239
traingda	170	0.9825	1.18318	22.806
traincgb	44	0.9976	0.2494	25.438

Table 2. Evaluation index of different learning algorithms.

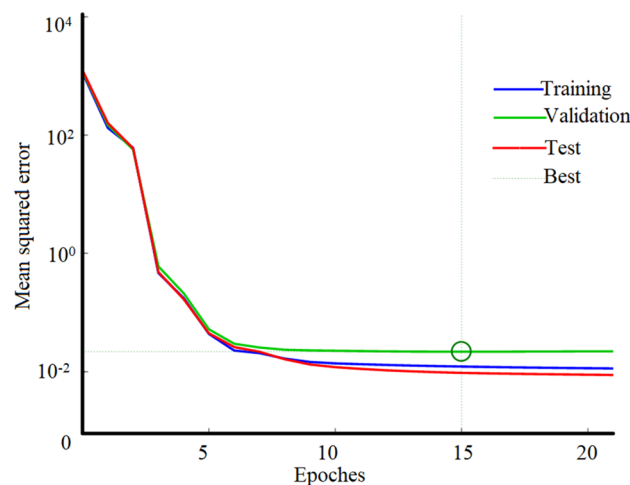


Figure 6. NARXNN training chart.

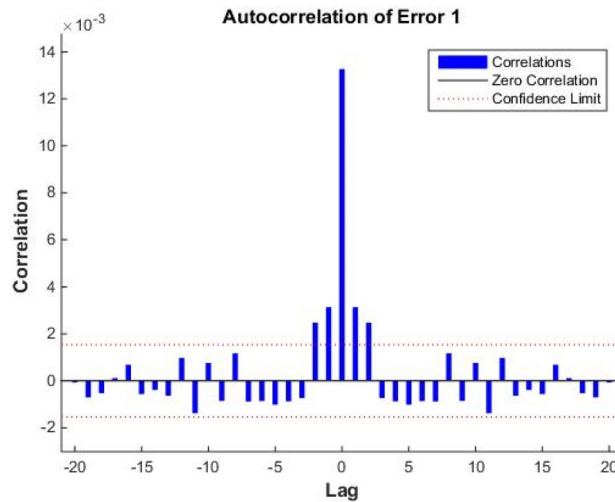


Figure 7. Error autocorrelation diagram.

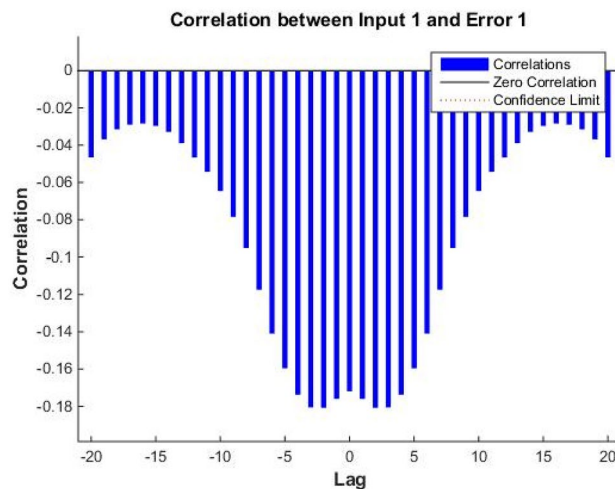


Figure 8. Input and output error correlation diagram.

as the input. Under these conditions, the model's general performance significantly decreases and the prediction function will fail. In the NARXNN prediction algorithm, the error autocorrelation graph and the input and error correlation graph were visualized, as shown in Figs. 7 and 8, respectively.

In practice, it is impossible to guarantee complete disparity. When calculating the degree of autocorrelation, the confidence interval should be set to 95%²³. The error should be maximized when lag is 0, and in all other cases, below the confidence interval. The lower the correlation between the input and the error, the better the results, and the closer its value is to zero in the plot. According to Fig. 7, the autocorrelation coefficient of the temperature prediction error sequence is 10^{-3} , it is only maximized when lag is 0, and in other cases, it does not exceed the confidence interval from the third group. These results demonstrate that the temperature prediction error sequence is a random process and does not have any correlation. Figure 8 shows that the input and error correlation is very low, the maximum value of the correlation coefficient's absolute value does not exceed 0.2, and that all values are within the confidence interval. The combined results from the error autocorrelation analysis and the input and error correlation analysis comprehensively demonstrate that it is feasible to use the model to predict the solar greenhouse temperature data.

NARXNN prediction model results analysis. To verify the accuracy and adaptability of the model prediction results, the 61 groups of prediction sample sets that were not used to train the model, were verified and analyzed via the heterometric calibration method. Due to the 1:3delay ratio, this process began with the fourth moment; thus, a total of 58 groups were analyzed. Specifically, the sample sets' measured and predicted values were compared and analyzed, the correlation between these values was calculated, and a correlation analysis plot was generated in which x is the indoor temperature measured value and f is the room temperature predicted value.

After fitting a correlation curve between the measured and predicted values (see Fig. 9), it was determined that the coefficient of determination = 0.9976, slope of the correlation curve = 1.015, and the ordinate intercept = - 0.056. These results demonstrate that the measured value is highly correlated with the predicted value. In addition, an error analysis of the prediction results from the same 58 sets was simultaneously performed. Results showed that the maximum absolute error = 0.67 °C. Thus, the solar greenhouse temperature time series prediction model established in this work can achieve high-precision greenhouse temperature prediction. As such, it can successfully be used to provide a scientific basis for forecasting greenhouse conditions, and therefore enable accurate greenhouse temperature regulation. To further display the temperature prediction results, the predicted temperature is displayed in the form of time series as shown in Fig. 10. The trend of predicted results is completely consistent with the real results. The peaks and troughs of daily temperature could be almost perfectly captured. Although there is error in the numerical value, the error is very small and within an acceptable range.

In the research of temperature prediction, neural network algorithm has strong adaptability because of its data-driven characteristics. Wang et al.²⁴ adopted BP neural network to obtain a good temperature prediction model for solar greenhouse. Wang²⁵ also obtained a temperature prediction model with excellent performance by using the Wavelet neural network. Compared with them, the proposed method adds a temperature feedback mechanism. To further evaluate the performance of the NARXNN based greenhouse temperature prediction model proposed herein, the same training and prediction sets were used in conjunction with the wavelet neural network time series prediction algorithm and the BP neural network to establish a model. Subsequently, the prediction performance of all three algorithms was evaluated. The evaluation results are shown in Fig. 11.

As shown in Fig. 11, the model based on the NARXNN algorithm depicts the best performance, with the R^2 of 0.9996, the MAE of 0.19, the maximum absolute error of 0.67, and the MSE of 0.12. The wavelet neural network timing analysis prediction algorithm and the BP network both use the historical temperature of the three historical moments as the input and the current moment temperature as the output to train the network. The NARXNN prediction algorithm and the wavelet neural network time series prediction algorithm enhance the sequence learning ability more so than the BP neural network prediction algorithm. As such, when solving the timing prediction problem, the former two algorithms show better prediction performance than the latter. Comparing the wavelet neural network timing analysis prediction with the NARXNN prediction algorithm, the latter adds a delay factor to consider the environmental factor timing characteristics. Subsequently,

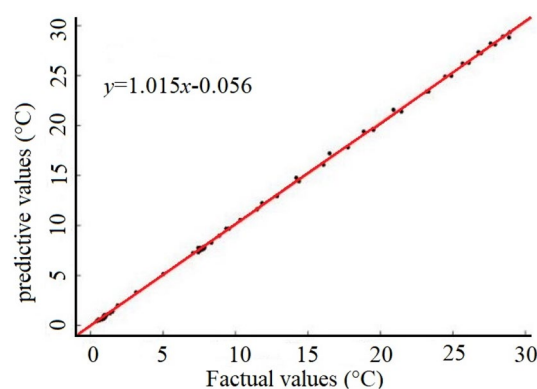


Figure 9. Temperature prediction of NARXNN.

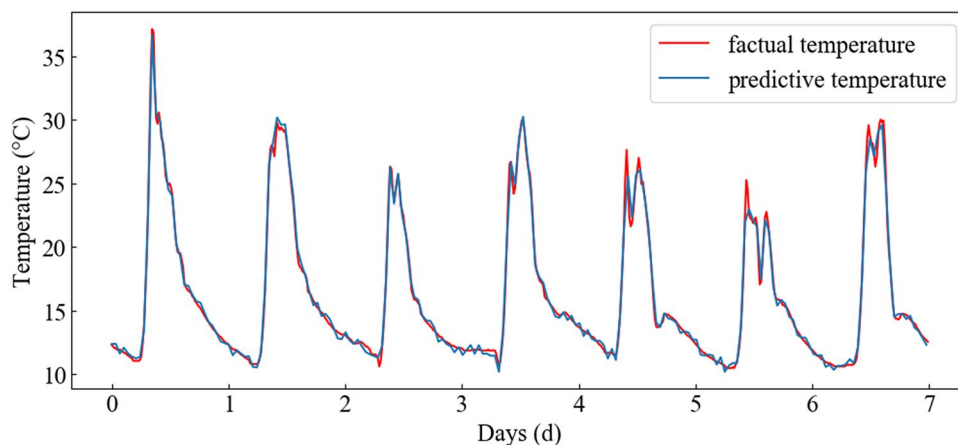


Figure 10. Temperature prediction results.

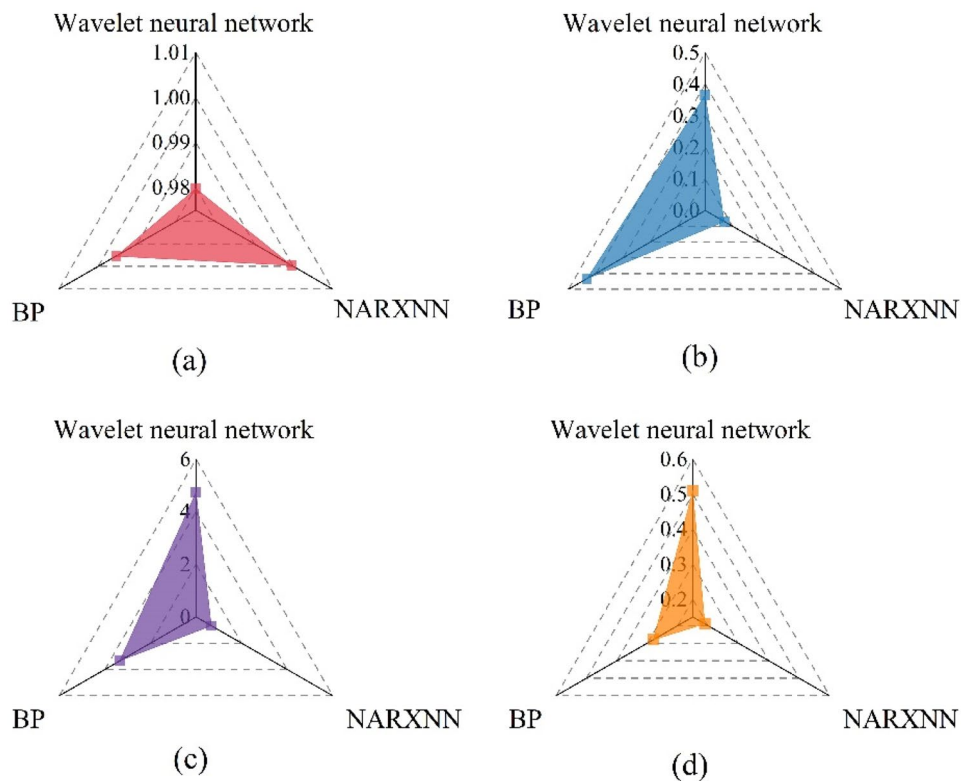


Figure 11. Evaluation index of greenhouse temperature prediction model. (a: R^2 , b: MSE, c: Maximum absolute error, d: MAE).

it uses the main environmental impact factor's optical radiation value and the indoor and outdoor temperature difference value as the network input to correct the network and predict the temperature at the next moment.

Conclusions

In this work, the performance of a NARXNN time series prediction model was compared with that of models using the wavelet neural network time series prediction algorithm and BP algorithm. Results showed that the NARXNN time series prediction model depicted both accurate results and the highest performance, and thus, is proposed for solar greenhouse ambient temperature prediction.

Compared with traditional greenhouse temperature prediction models, the model established herein is based on a time series and introduces two main environmental factors as model inputs—temperature difference and light radiation. Moreover, the proposed model depicted a maximum absolute error of 0.67 °C, and model correlation coefficient of 0.9996. Compared with traditional temperature prediction models, the temperature prediction accuracy offered by the NARXNN time series prediction model is greatly improved. Due to the greenhouse environment lag, greenhouse environment temperature regulation needs to be predicted and activated in advance. The model can achieve accurate temperature prediction for the next 20 min, which increases the advance decision-making time, and subsequently offsets the response time of the indoor temperature to the roller shutter quilt action, in effect, enabling proactive control.

The temperature is not only related to the current environmental conditions, but also to the accumulated heat in the past. The proposed model does not introduce accumulated heat, which limits the long-term temperature prediction performance. In the future research, the parameters related to heat accumulation, such as light duration, can be used as model inputs.

Data availability

The datasets used and/or analyzed during the current study available from the corresponding author on reasonable request.

Received: 5 September 2022; Accepted: 9 November 2022

Published online: 28 January 2023

References

1. Naseer, M. *et al.* Bioeconomic evaluation of extended season and year-round tomato production in Norway using supplemental light. *Agric. Syst.* **198**, 103391 (2022).
2. Guo, Y., Zhao, H. & Zhang, S. Modeling and optimization of environment in agricultural greenhouses for improving cleaner and sustainable crop production. *J. Clean. Prod.* **285**, 124843 (2021).

3. Kalkhajeh, Y. K., Huang, B. & Hu, W. Environmental soil quality and vegetable safety under current greenhouse vegetable production management in China. *Agric. Ecosyst. Environ.* **307**, 107230 (2021).
4. Zhang, M. *et al.* Energy-saving design and control strategy towards modern sustainable greenhouse: A review. *Renew. Sustain. Energy Rev.* **164**, 112602 (2022).
5. Gao, L. *et al.* Structure, function, application, and ecological benefit of a single-slope, energy-efficient solar greenhouse in china. *HortTechnology* **20**(3), 626–631 (2010).
6. Moon, A., & Lee, S. Design of the Farmer's experience based greenhouse operating systems. In *IEEE International Conference on Trust* (2013).
7. Silke, H. *et al.* Remote control of greenhouse vegetable production with artificial intelligence-greenhouse climate, irrigation, and crop production. *Sensors* **19**, 1807 (2019).
8. Xin, P. *et al.* Optimization and control of the light environment for greenhouse crop production. *Sci. Rep.* **9**, 8650 (2019).
9. Zhang, J. Climate change prediction using artificial neural network. *Adv. Comput. Sci. Technol.* **11**(4), 285–296 (2018).
10. Ma, L., Ji, J., & He, C. The study of greenhouse temperature modeling based on fuzzy neural network. In *International Conference on E-Product E-Service and E-Entertainment* 1–4 (2010).
11. Liu, R. *et al.* A fast and practical one-dimensional transient model for greenhouse temperature and humidity. *Comput. Electron. Agric.* **186**, 106186 (2021).
12. Holthuijzen, M. F. A comparison of five statistical methods for predicting stream temperature across stream networks. *Utah State Univ.* 6535 (2017).
13. Yau, J. *et al.* Modelling of ventilation rate and heating rate using multi-module fuzzy control system for a greenhouse. *Eur. J. Eng. Res. Sci.* **5**(7), 800–806 (2020).
14. Xin, P. *et al.* An improved photosynthesis prediction model based on artificial neural networks intended for cucumber growth control. *Appl. Eng. Agric.* **34**(5), 769–787 (2018).
15. Hu, J. *et al.* Model for tomato photosynthetic rate based on neural network with genetic algorithm. *Int. J. Agric. Biol. Eng.* **12**, 179–185 (2019).
16. Wang, D., Wang, M. & Qiao, X. Support vector machines regression and modeling of greenhouse environment. *Comput. Electron. Agric.* **66**(1), 46–52 (2009).
17. Jung, D. *et al.* Timeserial analysis of deep neural network models for prediction of climatic conditions inside a greenhouse. *Comput. Electron. Agric.* **173**, 105402 (2020).
18. Jini, E. R., & Sunny, S. Neural networks in data mining. *Int. J. Eng. Res. Technol.* **4**(6) (2018).
19. Ferreira, P. M. & Ruano, A. E. Predicting the greenhouse inside air temperature with RBF neural networks. *IFAC Proc. Vol.* **34**(11), 67–72 (2001).
20. Oliveira, A. G. *et al.* Simulations of the interplanetary magnetic field conditions with NARX networks. In *33rd International Cosmic Ray Conference* (2013).
21. Ouyang, H. T. Nonlinear autoregressive neural networks with external inputs for forecasting of typhoon inundation level. *Environ. Monit. Assess.* **189**(8), 376 (2017).
22. Mehmet, E. Novel neuronal activation functions for feedforward neural networks. *Kluwer Acad. Publ.* **28**(2), 63–79 (2008).
23. Afshar, P. J., Bahrampour, A. & Shahesmaeli, A. Determination of the trend of incidence of cutaneous Leishmaniasis in Kerman province 2014–2020 and forecasting until 2023. A time series study. *PLOS Negl. Trop. Dis.* **16**(4), e0010250 (2022).
24. Wang, H. *et al.* Temperature prediction model for sunlight greenhouse based on Bayesian regularization BP neural network. *Hub. Agric. Sci.* **54**(17), 4300–4303 (2015).
25. Wang, G. Research on short-term prediction of temperature based on “compact” wavelet neural network. *J. Comput. Sci. Appl.* **8**(1), 5–14 (2020).

Acknowledgements

This research was supported by Natural Science Basic Research Program of Shaanxi (Grant No. 2017JQ3025).

Author contributions

M.G.: Conceptualization, Methodology, Resources, Validation, Data curation, Writing—original draft; Q.W.: Investigation, Writing—review & editing; J.L.: Conceptualization, Writing—review & editing, Supervision, Funding acquisition; B.W.: Conceptualization, Methodology; Z.Z.: Writing—review & editing, Supervision; C.L.: Writing—review & editing; D.W.: review & editing. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023