RESEARCH ON THE DOA-BP-BASED TEMPERATURE AND HUMIDITY PREDICTION MODEL FOR COMMERCIAL CULTIVATION OF AGARICUS BISPORUS

基于 DOA-BP 的双孢菇工厂化生产温湿度预测模型研究

Tianhua LI [1)], Yinhang DONG [1)], Guoying SHI [1*)], Guanshan ZHANG[1)], Chao CHEN[2)], JianchangSU[3)]
[1)] Shandong Agricultural University, College of Mechanical and Electrical Engineering/ China;
[2)] Shandong Century Smart Agricultural Technology Co., LTD
[3)] Shandong Qihe Biotech Co., LTD
E-mail: sgy509@sdau.edu.cn
DOI: https://doi.org/10.35633/inmateh-73-13

Keywords: *Mushroom house environment, Temperature and humidity prediction, BP neural network, DOA optimization algorithm, Algorithm fusion.*

ABSTRACT

Accurate prediction of environmental changes in Agaricus bisporus cultivation is essential for better managing climatic conditions within mushroom houses, ultimately enhancing the yield and quality of Agaricus bisporus. However, traditional control systems for Agaricus bisporus production environments can only monitor the current conditions and lack the ability to predict environmental changes, leading to issues such as delayed feedback on environmental data and the effectiveness of control measures. In response to these challenges, this study establishes a temperature and humidity prediction model based on the DOA-BP algorithm. Experimental results demonstrate that the DOA optimization algorithm exhibits strong global search capabilities. By rapidly searching for optimal weights and biases, it overcomes the drawback of the BP neural network getting stuck in local minima, accelerates network convergence, and improves the performance of the BP neural network. The MAE values for temperature and humidity prediction inside the mushroom house are 0.021 and 0.013, respectively. The RMSE values are 0.044 and 0.038, respectively, and the R2 values are 0.976 and 0.968, respectively. Through validation, the DOA-BP temperature and humidity prediction model proposed in this study accurately predicts the temperature and humidity inside mushroom houses. This model can enhance environmental control for cultivation, optimize resource utilization, and reduce production costs effectively.

摘要

精准预测双孢菇生产环境变化有助于更好的管理菇房内的气候条件，提高双孢菇产量与质量。但传统的双孢菇生产环境控制系统只能对当前环境状况进行监测，无法对环境变化做出预判，导致环境数据的反馈和调控措施的生效都存在滞后性等问题。针对以上问题，本文建立了基于 DOA-BP 的温湿度预测模型，实验结果表明，DOA 优化算法具有较强全局搜索能力，通过快速搜索最优权值和偏置，克服了 BP 神经网络陷入局部极小值的缺点，加快网络收敛速度，提高了 BP 神经网络的性能。该预测模型对菇房内温湿度预测的 MAE 值分别为 0.021、0.013，RMSE 值分别为 0.044、0.038，R2 值为 0.976、0.968。通过验证，本研究提出的 DOA-BP 温湿度预测模型能够精准预测菇房温湿度，可以更好的控制栽培环境，还可以合理安排资源利用，降低生产成本。

INTRODUCTION

To achieve precision, standardization, and year-round production of *Agaricus bisporus*, the rapid development of environmentally controlled closed-system factory production is evident *(Mao et al., 2018)*. Temperature and humidity, as crucial environmental factors in the growth process of *Agaricus bisporus*, play a significant role throughout the entire cultivation cycle, directly influencing the quality and yield of *Agaricus bisporus*. Closed mushroom house environmental control systems can analyse and make regulatory decisions on monitored environmental factors. However, they lack the capability to predict future environmental changes, resulting in a certain lag in the feedback of abrupt environmental data and the effectiveness of control measures *(Zhao et al., 2020)*.

Tianhua Li, Professor; Yinhang Dong, Postgraduates; Guoying Shi, Senior experimentalist; Guanshan Zhang, Doctoral students; Chao Chen, Undergraduate; Jianchang Su, Researcher.

Therefore, establishing a high-precision prediction model with temperature and humidity as output variables is crucial. This model aims to anticipate the changing trends of temperature and humidity within the mushroom house over a future period, allowing timely adjustments to maintain optimal conditions for stable *Agaricus bisporus* production.

Greenhouse microclimate prediction models can be categorized into mechanistic models and data models *(Chen et al., 2017)*. Mechanistic models simulate and predict greenhouse environmental conditions based on fluid dynamics and heat and mass transfer mechanisms *(Zhou et al., 2014)*. However, due to the dynamic interdependence of various environmental factors within the greenhouse, these models require extensive physical parameters and environmental data, leading to challenges such as complex modelling and poor adaptability *(Saberian et al., 2019)*.

Data models, also known as black boxes, employ data fitting algorithms to ensure the predicted values align with actual values *(Wang et al., 2009)*. Previous studies, such as that by *Zong et al., (2022)*, utilized multiple linear regression, support vector machine regression, and random forest regression to construct nighttime hourly temperature prediction models under three different weather conditions in the greenhouse. Zou Weidong *et al., (2015),* improved the extreme learning machine for predicting temperature and humidity in a sunlight greenhouse.

Mao Xiao Juan et al., (2023), proposed a grey wolf optimization algorithm-based long short-term memory network model for greenhouse temperature prediction. However, these models were limited by a finite number of samples, and some were only studied under specific seasons or meteorological conditions, restricting their generalization capabilities.

Tian et al., (2020), proposed a combination method of MA-ARIMA-GASVR based on mining historical temperature data time-series information to establish a temperature prediction model. However, this model is relatively complex, demanding substantial computational resources and time, limiting its practical application.

Zu et al., (2023), introduced a sparrow search algorithm (SSA)-optimized long short-term memory network (LSTM) greenhouse environment prediction model. Yet, the limited local search capabilities of this optimization algorithm may lead to local optimal solutions, impacting the model's overall search capabilities and optimization effectiveness.

Zhang et al., (2021), established an Elman neural network prediction model based on collected environmental historical data. However, this model may encounter challenges such as gradient vanishing or exploding when predicting long time series, making it difficult to stabilize training networks and affecting predictive performance.

The aforementioned studies share common issues, including limited coverage of sample data, biased model selection, complexity, and susceptibility to locally optimal solutions *(Johnstone et al., 2021)*. In this paper, a DOA-BP-based industrial production environment prediction model for *Agaricus bisporus* was proposed. With the advantages of the DOA optimization algorithm, such as strong global search ability, parallelism and simplicity, the weight and bias of the BP network were accelerated to optimize, and environmental data of different regions of mushroom houses were collected through multiple groups of sensors. The 3sigma criterion and linear interpolation method were used to process the original data, and the input matrix of the processed environmental data was constructed according to the time series and input into the DOA-BP model for training, to achieve the accurate prediction of the production environment of the *Agaricus bisporus*.

This study compares the total energy consumption and environmental data between Mushroom House 28, which uses temperature and humidity data predicted by the DOA-BP model for climate control, and Mushroom House 29, which uses real-time sensor data for climate control. The results show that Mushroom House 28's temperature control equipment consumed 10.8% less total energy, and its humidification equipment consumed 15.4% less total energy compared to Mushroom House 29. Overall, Mushroom House 28's total energy consumption was 12.6% lower than that of Mushroom House 29. Additionally, the temperature and humidity in Mushroom House 28 remained stable and within optimal ranges, while Mushroom House 29 also maintained optimal temperature and humidity ranges, but with greater fluctuations.

The remaining sections of this article are organized as follows. In the second section, the DOA and BP algorithms are individually elaborated, and the fusion process of the two algorithms is summarized. The third section discusses the source and processing methods of the required data. In the fourth section, validation of the fusion algorithm is conducted, and its practical application in the field is demonstrated. The conclusion of this article is presented in the fifth section.

MATERIALS AND METHODS

DOA-BP ENVIRONMENTAL PREDICTION MODEL

DOA optimization algorithm

The Dingo Optimization Algorithm (DOA), proposed by Hernán Peraza-Vázquez and colleagues in 2021, is a novel bio-inspired algorithm *(Hernán et al., 2021)*. This algorithm draws inspiration from the hunting strategies of wild dingoes in Australia, including group attacks, grouping strategies, scavenging behaviour, and the incorporation of survival tactics *(Cai et al., 2023)*. These additions enhance the overall efficiency and performance of the method, endowing it with strong optimization capabilities and rapid convergence speed.

The fundamental mathematical model of the algorithm is as follows.

1. Population Initialization

$$\vec{x}_i(t) = lb_i(t) + rand(ub_i(t) - lb_i(t)) \quad (1)$$

where $\vec{x}_i(t)$ represents the position of the current optimizing individual, whose value varies between the lower search boundary lb_i and the upper search boundary ub_i of the Australian wild dog individual; t is the iteration number, rand is a randomly generated number uniformly distributed between [0,1].

2. Hunting strategy

1) Group Attack: Australian dingoes engage in coordinated group attacks when hunting large prey. They work together to locate the prey's position and surround it for a group attack. As illustrated in Figure 1, this behaviour can be described by (2).

$$\vec{x}_i(t+1) = \beta_1 \sum_{k=1}^{na} \frac{\left[\vec{\varphi}_k(t) - \vec{x}_i(t)\right]}{na} - \vec{x}_*(t) \quad (2)$$

When *random* < *P* and *random* < *Q* are satisfied. Where random is a uniformly generated random number in the range *[0,1]*, and *P* and *Q* are the probability values for the dingo population to execute their respective strategies. In this context, $\vec{x}_i(t+1)$ represents the new position of an individual, na is a randomly generated integer inversely proportional to *[2, sizepop/2]*, where sizepop is the total population size. Additionally, $\vec{\varphi}_k(t)$ denotes the set of individuals that will attack, with $\varphi \in X$, where *X* is a randomly generated dingo population. $\vec{x}_i(t)$ is the current optimizing individual, $\vec{x}_*(t)$ is the best individual found in the previous iteration, and β_1 is a uniformly generated random number in the range *[-2,2]*, serving as a scaling factor for the magnitude of individual movement trajectories.

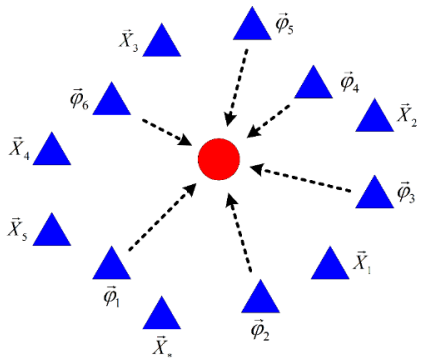


Fig. 1 - Group Attack Strategy

2) Persecution Strategy: Australian dingoes, when hunting small animals such as rabbits, employ a strategy where individual dingoes continuously pursue and approach the prey until it is captured individually. This behaviour can be described by (3).

$$\vec{x}_i(t+1) = \vec{x}_*(t) + \beta_1 * e^{\beta_2} * (\vec{x}_{r1}(t) - \vec{x}_i(t)) \quad (3)$$

When *random* < *P* is satisfied the following actions are executed.$\vec{x}_i(t+1)$ represents the new position of the individual, where $\vec{x}_i(t)$ is the current optimizing individual, and $\vec{x}_*(t)$ is the best individual found in the previous iteration. β_1 is the same as in (2), and β_2 is a uniformly generated random number in the range *[-1,1]*. r_1 is a random number generated from the range *[1, sizepop]*, and$\vec{x}_{r_1}(t)$ is the position of the randomly selected individual at index r_1, where $i \neq r_1$.

3) Scavenging: Australian dingoes exhibit scavenging behaviour when randomly wandering in their habitat and come across carrion as a food source. This behaviour can be described by (4).

$$\vec{x}_i(t+1) = \frac{1}{2}\left[e^{\beta_2} * \vec{x}_{r_1}(t) - (-1)^{\sigma} * \vec{x}_i(t) \right] \quad (4)$$

When *random* > *P* is satisfied. the following actions are executed. $\vec{x}_i(t+1)$ represents the new position of the individual, where $\vec{x}_i(t)$ is the current optimizing individual. $\vec{x}_{r_1}(t)$ is the position of the randomly selected individual at indexr_1, where$i \neq r_1$. $\beta_2$is the same as in (3), and σ is a uniformly generated random number in the range *[0,1].*

4) Survival Strategy: Due to reasons such as illegal hunting, Australian dingoes are facing the risk of extinction. The survival rate of Australian dingoes is described by (5). (6) is applicable when the survival rate is less than or equal to 0.3.

$$survival(i) = \frac{fitness_{\max} - fitness(i)}{fitness_{\max} - fitness_{\min}} \quad (5)$$

$$\vec{x}_i(t) = \vec{x}_*(t) + \frac{1}{2}\left[\vec{x}_{r_1}(t) - (-1)^{\sigma} * \vec{x}_{r_2}(t) \right] \quad (6)$$

Where $fitness_{max}$ and $fitness_{min}$ are the fitness values of the best and worst individuals in the current generation, and $fitness(i)$ is the current fitness value of the ith individual. The training data error absolute value is used as the individual fitness value, where $\vec{x}_i(t)$ represents the dingo individual with a lower survival rate that will be updated. r_1 and $r_2$are random numbers generated within the range *[1, sizepop]*, and $r_1 \neq r_2$. $\vec{x}_*(t)$ is the best individual found in the previous iteration.

Bp neural network

BP neural network is a concept proposed by *Rumelhart et al.* and other scientists in 1986. Each layer of neurons is fully connected with the adjacent layer of neurons, and there is no connection between neurons in the same layer and no feedback connection between neurons in each layer. It is a feedforward neural network system with hierarchical structure trained according to the error reverse propagation algorithm (*Xu et al., 2017*). Its training algorithm is simple, can handle various types of data, and has good adaptability to nonlinear problems (*Huang et al., 2020*).

This paper adopts a three-layer BP neural network, consisting of the input layer, hidden layer, and output layer. The number of nodes in the input and output layers is set to 6 and 2, respectively. The number of nodes in the hidden layer is determined based on the empirical formula *(Ding et al., 2023)* (7).

$$j = \sqrt{m+n} + a \quad (7)$$

In the formula, n represents the number of nodes in the input layer, m is the number of nodes in the output layer, j is the number of nodes in the hidden layer, based on experience, choosing a between 1 and 10 often helps to find a balance between computational efficiency and model accuracy. According to (7), the preferred range for j is 4 to 13.

Figure 2 shows the topology of BP neural network, where *h* is the output of the hidden layer, *Y* is the output of the output layer, W_{ij} is the weight between the j_{th} neuron of the hidden layer and the *i* neuron of the input layer, b_j is the bias of the j^{th} neuron of the hidden layer, and W_{ij} is the weight between the j^{th} neuron of the hidden layer and the l neuron of the output layer. b_1 and b_2 are the bias of neurons in the output layer.

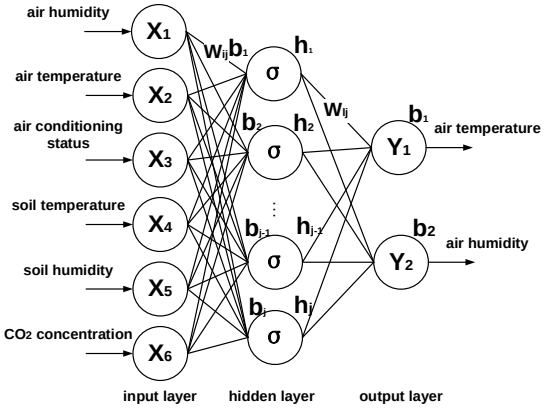


Fig. 2 - Structure of BPNN network

DOA-BP network fusion model

In the BPNN model, weights and biases play a crucial role in the model's fitting capability. In traditional BPNN models, weights and biases are typically set randomly based on empirical experience, which may lead the model to get stuck in local minima during network training, making it challenging to achieve optimal performance. Therefore, the DOA optimization algorithm is employed, known for its robust global search capabilities. Through hunting and survival strategies, optimal parameters are obtained for assigning and training the weights and biases of the BP neural network. This approach helps overcome the drawback of susceptibility to local minima, accelerates network convergence, and enhances the performance of the neural network. The algorithmic process is illustrated in Figure 3, and the specific steps are outlined below.

(1) DOA Initialization: Initialize individual positions by randomly generating a certain number of dingo individuals within the boundary conditions of the search space. Set their positions as random combinations of weights and biases.

(2) Fitness Evaluation: Input historical environmental data and perform forward propagation through the BP neural network to generate predicted values. Calculate the prediction error of each individual corresponding to the BP neural network, and evaluate the fitness of each individual using the loss function.

(3) Selection and Replication: Select individuals with better performance based on their fitness values. Retain these individuals for generating the next generation, and replicate the selected individuals to maintain the population size.

(4) Update Individual Positions: Update individual positions through group attacks, grouping strategies, and scavenging behaviour, ensuring that the updated positions do not exceed the allowable range of weights and biases.

(5) Convert the updated individual positions into combinations of weights and biases, and apply them to the BPNN model. Evaluate the positions through the fitness function to identify the individual with the optimal fitness.

(6) Update the Best Individual: Select the dingo individual with the best fitness value in the population as the current best individual. Record and update the global optimum solution.

(7) Check whether the set error or maximum iteration conditions are met. If satisfied, set the global optimal weights and biases as the parameters for the BPNN. If not satisfied, return to step (2).

(8) Model Prediction: Input new environmental data into the optimized BP neural network. Generate prediction results through forward propagation. Output the predicted temperature and humidity values.

(9) Model Evaluation and Validation: Evaluate the predictive performance of the model using a test dataset and calculate error metrics.

Through these steps, the weights and biases of the BP neural network will be set to search and optimize in the solution space of the DOA. By iteratively updating individual positions, the algorithm eventually finds the individual with the minimum fitness value, and its corresponding weights and biases are considered as the optimal parameters for the BPNN model.

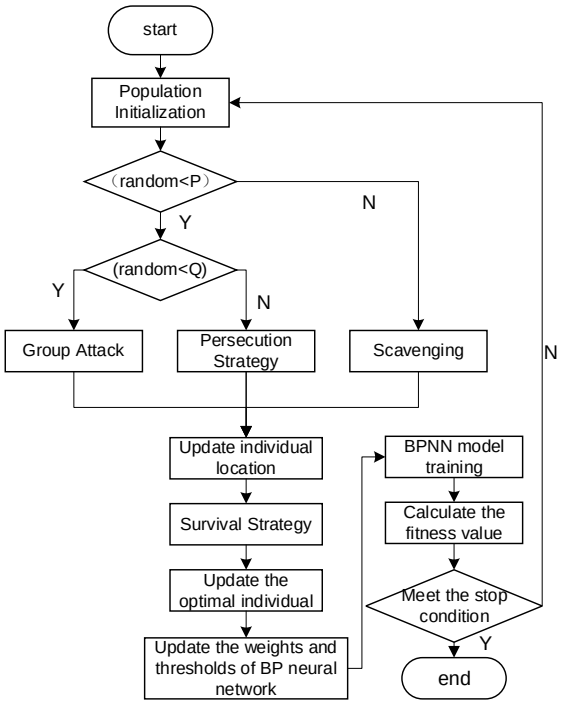


Fig. 3 - DOA-BPNN Flowchart

Model evaluation metrics

To provide a visual representation of the predictive performance and accuracy of the model, the mean absolute error *(MAE)*, root mean square error *(RMSE)*, and coefficient of determination *(R^2)* are selected as performance evaluation metrics for the model (*Liang et al., 2023*). Here, *N* represents the number of test samples, y_i is the actual value, $\hat{y}_i$ is the predicted value, and $\bar{y}_i$ is the mean value.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - y_i\right| \quad (8)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - y_i)^2} \quad (9)$$

$$R^2 = \frac{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2}{\sum_{I=1}^{N}(y_i - \bar{y}_i)^2} \quad (10)$$

DATA SOURCE AND PROCESSING

Overview of the experimental area

The experimental site is located at the Shandong Century Smart Agriculture Technology Co., Ltd. *Agaricus bisporus* Cultivation Base in Jining City, Shandong Province, China, specifically in Mushroom House No. 28 (35.224°N, 116.929°E). The mushroom house measures 18 m in length, 6.5 m in width, and 4.6 m in height. It is constructed using 100 mm thick polyurethane insulation boards, with a ground surface of 180 mm thick cement hardening. Each mushroom house has an area of approximately 117 m^2 and is equipped with two sets of aluminium alloy planting racks. These racks are 15 m long, and 1.4 m wide in planting face, with a layer height of 600 mm, totalling 6 layers.

The mushroom house is equipped with an intelligent temperature control system, humidification system, and ventilation system. When the temperature changes, the temperature control system activates the air conditioning through an S-shaped temperature control pipeline. In case of an increase in carbon dioxide concentration, the ventilation system extracts air from the mushroom house through the return air vent and introduces fresh air through the fresh air vent. When humidity decreases, the humidification system is activated, and the solenoid valve is opened for spray humidification.

Internet of Things data acquisition system

Build a greenhouse environment IoT data acquisition system, real-time and accurate. Obtaining environmental information is the premise and guarantee of realizing greenhouse environmental prediction. The IoT data-acquisition system mainly consists of a CPU (*Teik et al., 2021*), perception module and the transmission module is composed of the sensing module to complete the air temperature and relative humidity. Measurement of degree, carbon dioxide concentration and soil temperature and humidity, complete the collection of air conditioner on/off status. The relevant parameters of the sensor are shown in Table 1. The overall system architecture is shown in Figure 4.

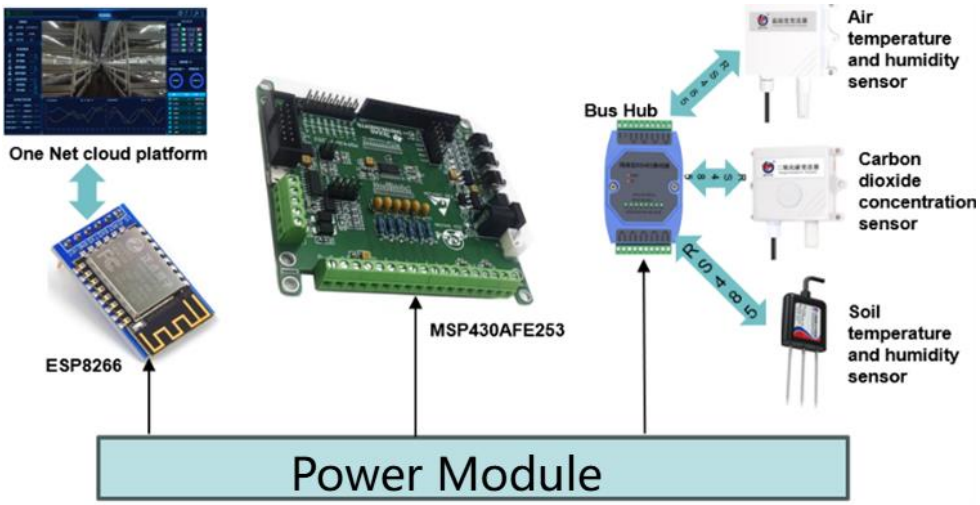


Fig. 4 - Lot system architecture diagram

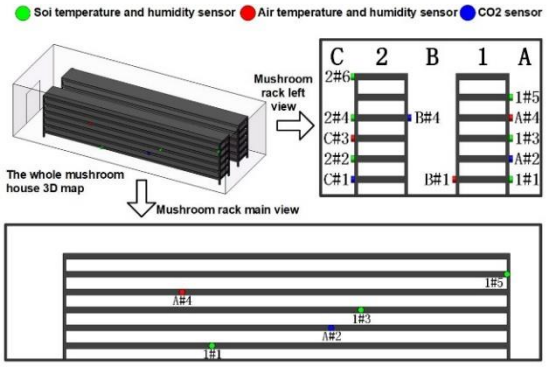


Fig. 5 - Three-dimensional map of sensor layout position

To ensure the accuracy of collecting environmental data in the mushroom house, the following sensor layout scheme is employed: Two sets of planting racks (15 m, 6 layers each), sensors are placed at 5 m intervals on 1st layer of Planting Rack #1 (1#1), 3rd layer of Planting Rack #1 (1#3), 5th layer of Planting Rack #1 (1#5), 2nd layer of Planting Rack #2 (2#2), 4th layer of Planting Rack #2 (2#4), and 6th layer of Planting Rack #2 (2#6). Air temperature and humidity sensors are placed on the west side of A#4 (1st layer, 4th layer of Planting Rack #1), east side of B#1 (1st layer of Planting Rack #1), and east side of C#3 (3rd layer of Planting Rack #2). CO2 sensors are placed on the west side of A#2 (2nd layer of Planting Rack #1), east side of B#4 (4th layer of Planting Rack #2), and east side of C#1 (1st layer of Planting Rack #2).

This optimized sensor layout maximizes the sensing range, reduces blind spots and overlapping areas, avoids resource waste and redundancy, and enhances detection accuracy and reliability.

The schematic diagram of the sensor deployment positions is shown in Figure 5, where the top left corner represents the overall three-dimensional view of the mushroom house, the top right corner is the main view of the mushroom rack, and the bottom shows the side view of Planting Rack #1. During the experiment, two sets of planting racks inside the mushroom house were used for the cultivation of *Agaricus bisporus* using the covered soil cultivation method, and the mushrooms were in the fruiting body growth stage (mushrooming stage and mushrooming period). The experimental site is illustrated in Figure 6.

a) b) c)

Fig. 6 - Interior environment of mushroom house

a) East aisle; b) Centre aisle; c) West aisle

Table 1

Sensor parameters

Sensor type	Model	Range	Precision
Air temperature and humidity sensor	RS-WS-N1-2-*	Temperature: -40℃~+80℃ Humidity:0%RH~100%RH	Temperature: ±0.4℃ Humidity: ±2%RH
CO2 sensor	RS-CO2*-*-2	0-5000ppm	±30ppm+3%F.S
Soil temperature and humidity sensor	RS-WS-*-TR-1	Temperature: -40℃~+80℃ Humidity:0%RH~100%RH	Temperature: ±0.5℃ Humidity: ±3%RH

Data collection and preprocessing

Data collection

Through the intelligent data collection system, automatically collect six types of data: air temperature, air humidity, soil temperature and humidity, CO_2 concentration, and air conditioner on/off status (0/1), was conducted from May 3, 2023, to June 7, 2023. This period corresponds to the fruiting body growth stage of *Agaricus bisporus* (fruiting stage and mushrooming period). The sampling interval was 10 minutes, resulting in 5040 data points for a single environmental parameter and a total of 30,240 data points.

The environmental data from May 3, 2023, to June 1, 2023, were used as the training set, while the data from June 2, 2023, to June 7, 2023, were used as the validation set for model verification. Some raw data, as illustrated in Figure 7 (taking the temperature data from May 3, 2023, to May 18, 2023, as an example), indicate the presence of missing data and instances of abrupt temperature changes to zero.

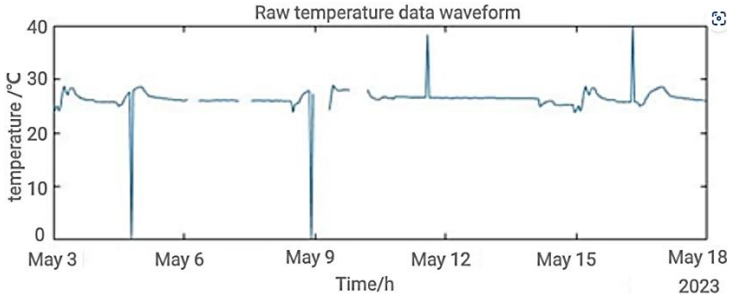


Fig. 7 - Partial raw data

Data preprocessing

Exposing sensors to prolonged high-humidity environments in the mushroom house can affect their accuracy. Additionally, issues such as network transmission quality and control equipment malfunctions may lead to anomalies and data gaps, impacting the accuracy of the prediction model. In this study, the 3-sigma criterion is employed to eliminate outliers from the collected raw data, and linear interpolation is used to impute missing data. As shown in Figure 8 (taking the temperature data from May 3, 2023, to May 18, 2023, as an example), the upper part displays temperature data after removing outliers, while the lower part shows temperature data after interpolation.

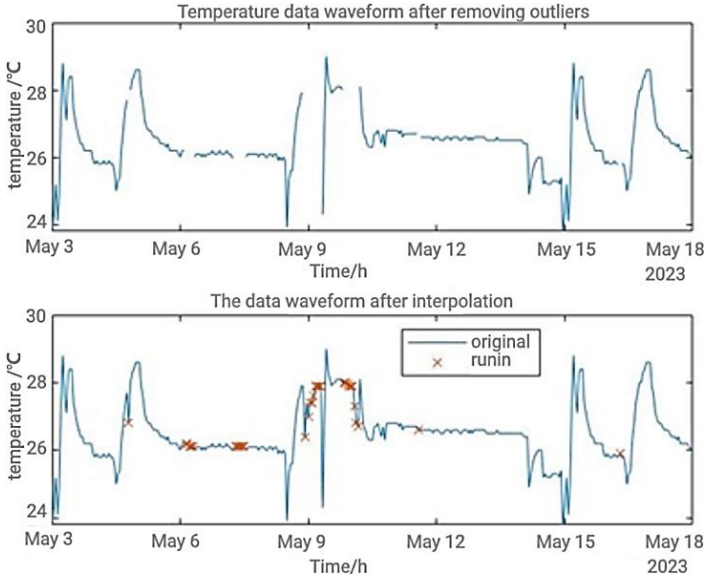


Fig. 8 - Data preprocessing

Data normalization

Normalization of data eliminates differences in range and dimensionality among various environmental parameters within the mushroom house, enhancing the stability, accuracy, and efficiency of the model. This ensures consistent weights for features, thereby better adapting to the modelling process (*Liang et al., 2023*).

$$X_i = \frac{x_i - x_{i\min}}{x_{i\max} - x_i} \tag{11}$$

Here, X_i represents the normalized value, and x_i is the environmental data within the sample dataset X. x_{imin} and x_{imax} represent the minimum and maximum values of the data belonging to the sample before normalization.

TEST

Test platform

The platform configuration used for training is as follows: operating system Windows 10 Professional 64-bit, processor 12th Gen Intel Core i9-12900KF, graphics card NVIDIA GeForce RTX 3090, and memory 128GB. The programming software used is Matlab R2023a.

Model parameter settings

The parameter settings are as follows: In the DOA optimization model, a larger population size can expand the search space and improve global search capabilities, but it increases computational complexity. Set the total number of dingoes to 200. The probability of hunting or scavenging strategy, P, is set to 0.5, and the probability of group attack or persecution, Q, is set to 0.7. The search boundaries define the range of individual searches for new solutions. A larger search boundary can expand the search space, aiding global search but may slow down the search speed. The upper and lower search boundaries of the population individuals, lb_i and ub_i, are set to -10 and 10, respectively.

The selected three-layer BP neural network model has an input layer with 6 nodes and an output layer with 2 nodes. Through program execution, under the same sample set and number of training iterations, the root mean square error (RMSE) for different numbers of hidden layer nodes is calculated, as shown in Figure 9. When the number of hidden layer nodes is 14, the RMSE is minimized.

Therefore, the BP neural network model is configured with 14 hidden layer nodes, and the activation function for each neuron is the sigmoid function:$\sigma(x) = 1/(1 + e\text{^}(-x))$. The loss function used is the mean square error $E = \frac{1}{2}\sum_{l=1}^{k}(d_l - y_l)^2$, with a learning rate ε set to 0.01. The training target for the minimum error is set to 0.0001, and the time step, determining the utilization of historical information and model prediction performance, is set to 15. The maximum number of iterations is set to 1000. The parameter settings for the DOA-BP prediction model are summarized in Table 2.

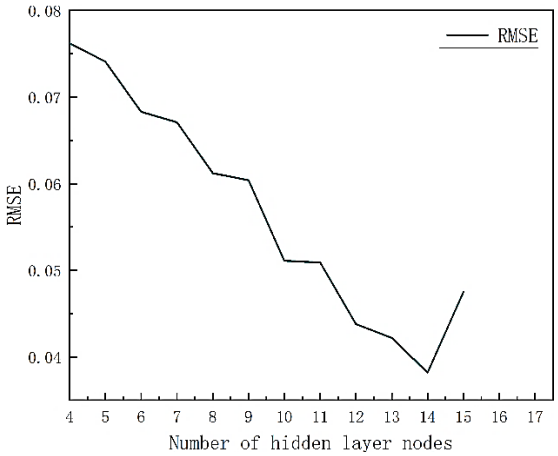


Fig. 9 - Root mean square error corresponding to node number of different hidden layers

Table 2

Model parameter Settings

Name of parameter	Parameter values
Population quantity	200
Probability of P	0.5
Probability of Q	0.7
Learning rate	0.01
Network layer number	3
Input layer node	5
Hidden layer node	13
Output layer node	2
Time step	15
Search boundary	-10~10
Target minimum error	0.001
Maximum number of iterations	1000

RESULT AND ANALYSIS

Data preprocessing impact analysis

Table 3 presents a comparative analysis of the DOA-BP prediction model's temperature forecasting performance before and after data preprocessing. From the table, it can be observed that, after data preprocessing, the DOA-BP model's Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) for temperature prediction decreased by 0.005 and 0.057, respectively, compared to before preprocessing. The coefficient of determination (R^2) increased by 0.075. Similarly, for humidity prediction, the MAE and RMSE decreased by 0.009 and 0.0047, respectively, and R^2 increased by 0.057. This indicates that data missing and anomalies have a negative impact on the prediction model, demonstrating the effectiveness of using the 3-sigma criterion and linear interpolation for data preprocessing.

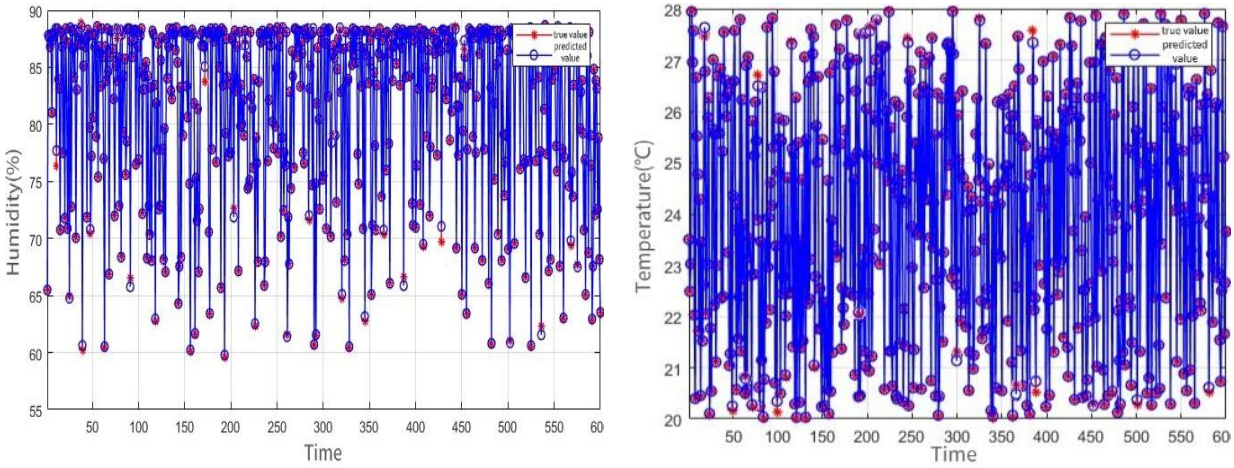
Table 3

Comparative analysis of DOA-BP prediction model's effect on temperature prediction before and after data preprocessing

Temperature	**MAE**	**RMSE**	**R^2**
Before repair	0.026	0.101	0.901
After restoration	0.021	0.044	0.976
Humidity	MAE	RMSE	R2
Before repair	0.022	0.085	0.911
After restoration	0.013	0.038	0.968

Prediction results and comparative analysis

The obtained DOA-BP prediction model was used to forecast the temperature and humidity inside the mushroom cultivation facility for the next 5 days. The results are depicted in Figure 10.



a humidity prediction *b temperature prediction*

Fig. 10 - Prediction results of temperature and humidity in mushroom house based on DOA-BP

The DOA-BP model achieved R^2, MAE, and RMSE values of 0.976, 0.021, and 0.044 for temperature prediction, respectively. For humidity prediction, the model achieved R^2, MAE, and RMSE values of 0.968, 0.013, and 0.038, respectively.

To validate the scientific and superior performance of the proposed DOA-BP neural network prediction model, this study conducted a comparative analysis of predictive performance with BP neural network, genetic algorithm optimized BP neural network (GA-BP), and particle swarm optimization optimized BP neural network (PSO-BP) under the same input parameters and computational environment.

As seen in Table 4, For humidity prediction, the R^2 of DOA-BP model increased by 8.3%, 5.9%, and 5.4% compared to BP model, GA-BP model, and PSO-BP model, respectively. Moreover, the MAE and RMSE decreased by 12.5%, 54.2% compared to BP model, 30%, 48.2% compared to GA-BP model, and 8.7%, 46.9% compared to PSO-BP model. As seen in Table 5, for temperature prediction in the mushroom cultivation facility, the R^2 of DOA-BP model increased by 6%, 4.6%, and 4.4% compared to BP model, GA-BP model, and PSO-BP model, respectively. Additionally, the MAE and RMSE decreased by 35%, 56.3% compared to BP model, 40.9%, 51.9% compared to GA-BP model, and 38.1%, 51.3% compared to PSO-BP model.

The number of iterations in which the best performance of each model for temperature and humidity prediction appears respectively were 132 and 174 for DOA-BP, 352 and 326 for BP, 205 and 241 for GA-BP, and 197 and 225 for PSO-BP. In conclusion, the DOA-BP model demonstrated superior predictive performance, accurately forecasting temperature and humidity in the mushroom cultivation facility, while also improving model efficiency.

Table 4

Comparison of humidity evaluation indexes predicted by models

Model(humidity)	MAE	RMSE	R^2	Iterations
BP	0.020	0.087	0.913	352
GA-BP	0.023	0.079	0.925	205
PSO-BP	0.021	0.078	0.927	197
DOA-BP	0.013	0.038	0.968	132

Table 5

Comparison of temperature evaluation indexes predicted by models

Model(temperature)	MAE	RMSE	R^2	Iterations
BP	0.024	0.096	0.912	316
GA-BP	0.030	0.085	0.922	241
PSO-BP	0.023	0.083	0.926	225
DOA-BP	0.021	0.044	0.976	174

Experimental validation

Mushroom house No. 28 and Mushroom house No. 29 of the mushroom culture base of Shandong Century Smart Agricultural Technology Co., LTD., Jining City, Shandong Province, were selected. The two mushroom houses were in the fruiting stage, and the suitable temperature range was 20-24℃ and the suitable humidity range was 80-85%. The automatic environmental control system of mushroom room No. 28 uses the temperature and humidity data predicted by the DOA-BP model to adjust temperature and humidification, while the automatic environmental control system of mushroom room No. 29 directly adjusts temperature and humidification according to the real-time sensor data. The observation time is from July 1, 2023 to July 4, 2023, and the environmental data is collected every 10 minutes. Record the total power consumed by the device. The total electric energy consumption and environmental data of the equipment in Room No. 28 and Room No. 29 were compared and analysed. As shown in the table 5, the total electric energy consumed by the temperature regulating equipment in Room 28 is reduced by 10.8% compared with room 29, the total electric energy consumed by the humidifying equipment is reduced by 15.4% compared with room 29, and the total electric energy is reduced by 12.6% compared with room 29.

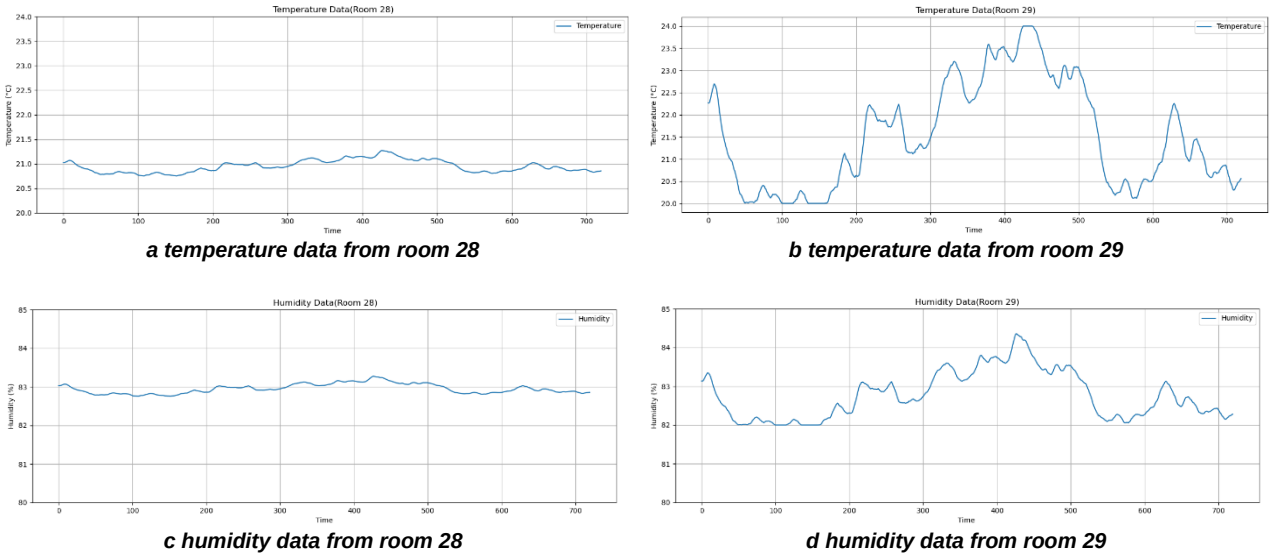
Figure 11 shows the temperature data and humidity data of mushroom room No. 28 and No. 29 within 5 days. The overall temperature and humidity of mushroom room No. 28 changed steadily, and always kept within the appropriate temperature and humidity range; Room No. 29 kept the overall temperature and humidity within the appropriate range. But the temperature and humidity change range was large.

Based on the above analysis, the DOA-BP temperature and humidity prediction model proposed in this study can accurately predict the temperature and humidity of mushroom houses, which can better control the cultivation environment, rationally arrange the utilization of resources, and reduce the production cost.

Table 6

Power consumption of equipment

Room	thermostat consumes electricity（KW*h）	humidifier consumes electrical energy（KW*h）	Total electric energy（KW*h）
Mushroom house No. 28	371	231	602
Mushroom house No. 29	416	273	689



a temperature data from room 28 *b temperature data from room 29*

c humidity data from room 28 *d humidity data from room 29*

Fig. 11 - The change of the overall temperature and humidity of the mushroom house

CONCLUSIONS

(1) The temperature and humidity prediction model for the industrialized production of *Agaricus bisporus* based on the DOA-BP proposed in this study utilizes the strong global search capability of the DOA optimization algorithm to perform rapid search and optimization of the weights and thresholds of the BP neural network. This approach addresses issues such as low model prediction accuracy resulting from manually setting weights and biases through random empirical methods. It helps overcome the drawbacks of easily falling into local minima, accelerates network convergence, and improves the overall performance of the neural network.

(2) The greenhouse environment Internet of Things data acquisition system is used to obtain environmental information in real-time and accurately. At the same time, the abnormal environmental data collected is eliminated based on the 3sigma criterion, and the missing data is interpolated by the linear interpolation method. The input matrix of the processed environmental data is constructed according to the time series and input into the DOA-BP model for training. The accuracy of temperature and humidity prediction is significantly improved by the model, which provides a guarantee for realizing the accurate prediction of the greenhouse environment.

(3) Experimental results indicate that the model achieves determination coefficients (R^2) of 0.976 and 0.968 for temperature and humidity predictions, respectively. The performance surpasses that of BP neural networks, genetic algorithm-optimized BP neural networks, and particle swarm optimization-optimized BP neural networks. While enhancing prediction accuracy, the model also considers operational efficiency, thus improving the overall predictive performance. It exhibits notable superiority in predicting the temperature and humidity of the *Agaricus bisporus* production environment.

(4) Through validation, the DOA-BP temperature and humidity prediction model proposed in this study accurately forecasts the conditions in mushroom houses. It enables better control of cultivation environments and facilitates efficient resource allocation, thereby reducing production costs.

ACKNOWLEDGEMENT

This work was supported by the Major Science and Technology Innovation Project of Shandong Province: Research and Development and Industrialization of key technologies for intelligent Factory Production of edible fungi (2022CXGC010609); Shandong Vegetable Industry Technology System Project (SDAIT-05-11)

REFERENCES

[1] Chen L., Du S., Li J., He Y., Liang M. (2017). Online identification method of parameters for greenhouse temperature prediction self-adapting mechanism model (温室环境温度预测自适应机理模型参数在线识别方法). *Transactions of the Chinese Society for Agricultural Engineering*, Vol. 33, pp. 315-321.

[2] Cai X., Zhu J., Liu M., Liu J., Meng Z., Yu Y. (2023). Peak shaving strategy of electric vehicles based on an improved dingo optimization algorithm (基于改进野狗优化算法的电动汽车调峰策略). *Energy Storage Science and Technology*, Vol. 12, pp. 1913-1919.

[3] Ding G., Shi X., Hu J. (2023). Prediction model of the aeration oxygen supply for aerobic composting using CGA-BP neural network (基于 CGA-BP 神经网络的好氧堆肥曝气供氧量预测模型). *Transactions of the Chinese Society for Agricultural Engineering*, Vol. 39, pp. 211-217.
[4] Johnstone C., Sulungu ED. (2021). Application of neural network in prediction of temperature: a review. *Neural computing and applications*, Vol. 33, pp. 11487-11498.
[5] Huang Y., Xiang Y., Zhao R., Cheng Z. (2020). Air quality prediction using Improved PSO-BP neural network. *IEEE Access*, vol. 8, pp. 346-353.
[6] Hernán P., Adrian F., Gustavo E., Ana B., Jonás V., Fernando R. (2021). A bio-inspired method for engineering design optimization inspired by dingoes hunting strategies. *Mathematical Problems in Engineering*, Vol. 138, pp. 1-19.
[7] Liang W., Xue W., Ma J., Cheng L., Xu H. (2023). Soft measurement of main steam flow in municipal solid waste incinerator based on WOA-Elman neural network (基于 WOA-Elman 神经网络的城市固废焚烧炉主蒸汽流量软测量). *Control Engineering*, Vol. 31, pp. 1-7.
[8] Mao H., Jin C., Chen Y. (2018). Research progress and prospect on control methods of greenhouse environment (温室环境控制方法研究进展分析与展望). *Transactions of the Chinese Society for Agricultural Machinery*, Vol. 49, pp. 1-13.
[9] Mao X., Bao T., Xun G., Li D., Wang B., Ren N. (2023). Prediction of temperature in the greenhouse of vegetable growing based on GWO-LSTM (基于 GWO-LSTM 的设施蔬菜温室温度预测). *Journal of Chinese Agricultural Mechanization*, vol. 44, pp. 116-123.
[10] Rumelhart D., Hinton G., Williams J. (1986). Learning representations by back-propagating errors. *Nature*, Vol. 323, pp. 533-536.
[11] Saberian A., Sajadiye S. (2019). The effect of dynamic solar heat load on the greenhouse microclimate using CFD simulation. *Renewable Energy*, Vol. 138, pp. 722-737.
[12] Tian D., Wei X., Wang Y., Zhao A., Mu W., Feng J. (2020). Prediction of temperature in edible fungi greenhouse based on MA-ARIMA-GASVR (基于 MA-ARIMA-GASVR 的食用菌温室温度预测). *Transactions of the Chinese Society for Agricultural Engineering*, Vol. 36, pp. 190-197.
[13] Teik TS., Krishnen G., Khulidin KA., Tahir M., Hashim M., Khairudin S. (2021) Automated controlled environment mushroom house. *Advances in Agricultural and Food Research Journal*, Vol. 2.
[14] Wang D., Wang M., Qiao X. (2009). Support vector machines regression and modeling of greenhouse environment. *Computers and Electronics in Agriculture*, Vol. 66, pp. 46-52.
[15] Xu B., Dan H., Li L. (2017). Temperature prediction model of asphalt pavement in cold regions based on an improved BP neural network. *Applied Thermal Engineering*, Vol. 120, pp. 568-580.
[16] Zhou W., Li Y., Wang X. (2014). Model predictive control of air temperature in greenhouse based on CFD unsteady model (基于 CFD 非稳态模型的温室温度预测控制). *Transactions of the Chinese Society for Agricultural Machinery*, Vol. 45, pp. 335-340.
[17] Zou W., Zhang B., Yao F., He C. (2015). Verification and forecasting of temperature and humidity in solar greenhouse based on improved extreme learning machine algorithm (基于改进型极限学习机的日光温室温湿度预测与验证). *Transactions of the Chinese Society for Agricultural Engineering*, 31, 194-200.
[18] Zhao Q., Song Z., Li Q., Zheng W., Liu Y., Zhang Z. (2020). Multi-point prediction of temperature and humidity of mushroom based on CNN-GRU (基于 CNN-GRU 的菇房多点温湿度预测方法研究). *Transactions of the Chinese Society for Agricultural Machinery*, Vol. 51, pp. 294-303.
[19] Zhang J., Shan H., Jing X., Li C., Zhang C., Liu H. (2021). Prediction method of greenhouse environmental factors based on Elman neural network (基于 Elman 神经网络的温室环境因子预测方法). *Journal of Chinese Agricultural Mechanization*, vol. 42, pp. 203-208.
[20] Zong C., Wang J., Song W., Gen R., Liu P., Xu D. (2022). Construction and validation of hourly air temperature prediction model in solar greenhouse at night (基于天气预报的日光温室夜间逐时气温预测模型构建). *Transactions of the Chinese Society for Agricultural Engineering*, Vol. 38, pp. 218-225.
[21] Zu L., Liu P., Zhao Y., Li T., Li H. (2023). Solar greenhouse environment prediction model based on SSA-LSTM (基于 SSA-LSTM 的日光温室环境预测模型研究). *Transactions of the Chinese Society for Agricultural Machinery*, Vol. 54, pp. 351-358.