

Modeling and simulation of temperature control system in plant factory using energy balance

Mingqiu Zhang^{1,2}, Wei Zhang^{1*}, Xiaoyu Chen³, Fei Wang³, Hui Wang², Jisheng Zhang², Linhui Liu²

(1. College of Engineering, Heilongjiang Bayi Agricultural University, Daqing 163319, Heilongjiang, China;

2. College of Modern Manufacturing Engineering, Heilongjiang University of Technology, Jixi 158100, Heilongjiang, China;

3. Beidahuang Kenfeng Seed Co., Ltd., Harbin 150431, China)

Abstract: Closed production systems, such as plant factories and vertical farms, have emerged to ensure a sustainable supply of fresh food, to cope with the increasing consumption of natural resource for the growing population. In a plant factory, a microclimate model is one of the direct control components of a whole system. In order to better realize the dynamic regulation for the microclimate model, energy-saving and consumption reduction, it is necessary to optimize the environmental parameters in the plant factory, and thereby to determine the influencing factors of atmosphere control systems. Therefore, this study aims to identify accurate microclimate models, and further to predict temperature change based on the experimental data, using the classification and regression trees (CART) algorithm. A random forest theory was used to represent the temperature control system. A mechanism model of the temperature control system was proposed to improve the performance of the plant factories. In terms of energy efficiency, the main influencing factors on temperature change in the plant factories were obtained, including the temperature and air volume flow of the temperature control device, as well as the internal relative humidity. The generalization error of the prediction model can reach 0.0907. The results demonstrated that the proposed model can present the quantitative relationship and prediction function. This study can provide a reference for the design of high-precision environmental control systems in plant factories.

Keywords: plant factory, temperature control system, mechanism simulation, random forest, cart model, generalization error

DOI: 10.25165/ijabe.20211403.6114

Citation: Zhang M Q, Zhang W, Chen X Y, Wang F, Wang H, Zhang J S, et al. Modeling and simulation of temperature control system in plant factory using energy balance. *Int J Agric & Biol Eng*, 2021; 14(3): 66–75.

1 Introduction

Plant factories, one type of closed system, are designed to maximize production density, productivity, and use the efficiency of natural resource to alleviate the local scarcity of urban food supply. This high-level agricultural production approach can efficiently adjust the internal microclimate to a more suitable state for plant growth, thereby improving the quality and yield of crops^[1]. There are some incorporating features of plant factories, such as artificial lighting and active cooling, particularly on the excellent hermetic seal, tunable light intensity, forced circulation channels of internal air, and high CO₂ absorption rate^[2]. Currently, the switching control system is commonly used to manipulate the artificial lighting, temperature, humidity, CO₂ concentration, and nutrient solution in most facilities of plant factories^[3-6]. However, the input and output signals of various parameters always vary in constant impulse^[7-9], easily resulting in power loss and annual

maintenance cost of the equipment^[10]. Therefore, it is necessary to select a continuous variable control system for the environmental devices to obtain the desired values of outputs. Then, the main component for this continuous control system can be a microclimate environment model, generally related to temperature, humidity, CO₂ in plant factories. Establish a highly accurate microclimate model becomes a great challenge for the subsequent procedures in a plant factory, particularly on the precise control of crop growth, quantitative management, environmental atmosphere optimization, saving energy, and reducing resource consumption^[11-13].

Temperature is one of the important climate variables affecting crop growth. The physical laws of this continuous variable can be represented in a continuous dynamic mode^[11,14]. The control system is designed, usually based on a logic rule algorithm, in most devices, such as artificial light sources, temperature control equipment. Thus, the specific procedure can be driven by discrete events, concurrently varying in the continuous dynamic changes^[15,16]. The interaction between discrete events and continuous dynamic changes can make the temperature a complex hybrid dynamic system^[17,18].

Most efforts have been made on the theories and methodologies for the prediction of a temperature control system in a plant factory over the last decades^[19-22]. The conventional prediction methods were mostly based on the principle of sequential control systems, representing by autoregressive moving average model (ARMA), and non-stationary autoregressive moving average model (ARIMA), with emphasis on ARIMA model with a linear approximation system and disturbance characteristics^[13,19,20,23]. These widely used methods can be

Received date: 2020-08-24 **Accepted date:** 2021-04-01

Biographies: **Mingqiu Zhang**, PhD candidate, Associate Professor, research interests: agricultural mechanization engineering, Email: zmq9521@126.com; **Xiaoyu Chen**, Master, research interests: phenotypic identification of greenhouse automatic platform crops, Email: 592095502@qq.com; **Fei Wang**, Master, research interests: novel organic binding phosphate fertilizer, Email: 672360226@qq.com; **Hui Wang**, Master, research interests: mathematical mechanization, Email: huihui_666@163.com; **Jisheng Zhang**, Master, research interests: mechanical design and theory, Email: 921142278@qq.com; **Linhui Liu**, Master, research interests: data analysis, Email: 375977508@qq.com.

***Corresponding author:** **Wei Zhang**, PhD Supervisor, Professor, research interests: agricultural mechanization engineering. College of Engineering, Heilongjiang Bayi Agricultural University, Daqing 163319, Heilongjiang, China. Tel: +86-15246110999, Email: zhang66wei@163.com.

attributed to the dominant features, such as the fewer input data to deal with, and simple models^[13]. However, the high stability of the original time series has become highly demanded, while the probability distribution of white noise is also required to know ahead of time. The method based on a priori probability can cause some deviation when ordering in the time series during data processing. Specifically, the large prediction error can occur with non-normal distribution data and multi-collinearity.

This study aims to construct a model of internal temperature variation in a plant factory based on the energy balance theory^[24,25], and further to explore the input factors that influence the internal temperature, such as volume flow, supply air temperature, internal relative humidity. Using the classification and regression trees (CART) algorithm in random forest (RF) theory^[26-28], a model system with in-out measurements was established to predict the behavior of the temperature control system, with emphasis on economically feasible application.

2 Theoretical background for mechanism model of temperature control system

This section presents the theoretical models of energy balance in closed plant production, further to deduce the mechanism model of a temperature control system, particularly for energy loads balance, crop transpiration, and artificial lighting in plant factories. In a building structure, there are non-uniform spatial distributions of temperature in the horizontal and vertical directions^[29,30], partly due to the combined influences of many factors, such as the internal environment, control mode, planting structure, and types of crops^[31,32]. Nevertheless, the internal temperature in the plant factory can be assumed as a uniform spatial distribution in a specific research period, and then to establish the relationship between temperature and time.

2.1 Energy balance in a conventional greenhouse

A typical light-transmitting greenhouse with natural ventilation consists of some structural materials (glass, plastic), surface soil, architectural components, heating and vents^[33]. Taking the air and crops in the greenhouse as the research object, two forms of energy can be divided, including the sensible heat and latent heat. The sensible heat is the exchange of energy between the system and the surrounding, including the solar radiation energy (Q_R), conduction and convection passing through greenhouse materials (Q_F), the energy transferred by different architectural components and production systems (Q_C), soil heat transfer energy (Q_{Soil}), the energy transferred by natural ventilation (Q_{V1}), the energy transferred by mechanical ventilation (Q_{V2}), energy supply by the heating system (Q_S), heat energy from the artificial lighting system, and the energy consumed by photosynthesis. The latent heat is the transfer of energy to a system during a phase transition of matter, covering the energy transferred by crop transpiration (Q_{tran}), the heat absorption and release during water evaporation and condensation (Q_H). To simplify the model, the heat energy from artificial lighting, and energy consumed by photosynthesis can be neglected due to their relatively small values, compared with the the sensible heat. Therefore, a simplified model can be established to illustrate the energy balance in a typical greenhouse, as shown in Figure 1. The energy balance equation can be represented by Equation (1).

$$Q_R + Q_F + Q_C + Q_{Soil} + Q_S + Q_H - Q_{V1} - Q_{V2} - Q_{tran} = 0 \quad (1)$$

2.2 Energy balance in a plant factory

An artificial lighting system is generally installed to provide the related energetic fluxes in a plant factory. The opaque thermal

insulation materials can be used to build a highly insulated closed architecture, like a warehouse in shape, thereby minimizing the heat transfer from the structural wall to the outside. Inside a plant factory, forced air circulation is usually utilized to transfer all the energy passing through the internal system^[34,35]. Therefore, the artificial lighting radiation energy Q_{rad} can be selected in a plant factory, rather than the solar energy Q_R , in a typical greenhouse, where Q_{diss} is the thermal energy of artificial lighting. In the case of completely artificial production and opaque facades, the architecture can be considered as adiabatic, whereas the Q_F , Q_{soil} , Q_C can also be neglected. When the temperature control device is used to monitor the heating and cooling, the energies Q_{V1} , Q_{V2} can become Q_0 (the energy flux flowing out during forced ventilation), and Q_S (the energy flux transferred into the indoor via the treated air). The soil moisture evaporation can be omitted, due to the excellent sealing performance under the thin film covering soil surface, or nutrients cultivation pond. The water vapor condensation can also be omitted with the relatively small value, compared with that of the total crop transpiration. Thus, a simplified model can be established to illustrate the energy balance of the plant factory, as shown in Figure 2. The energy balance equation of the plant factory can be represented by Equation (2).

$$Q_S + Q_{rad} + Q_{diss} + Q_H - Q_{tran} - Q_0 = 0 \quad (2)$$

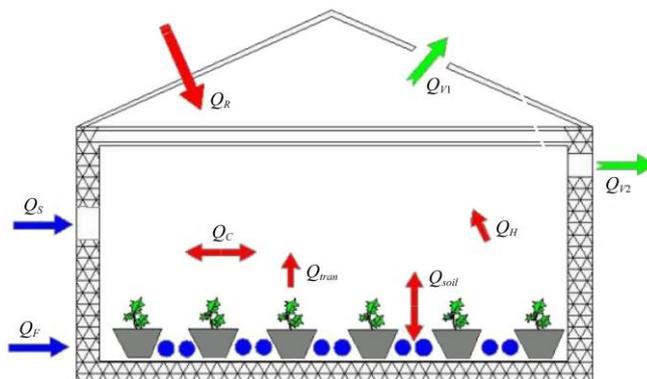


Figure 1 Illustration for the energy balance model in a typical greenhouse

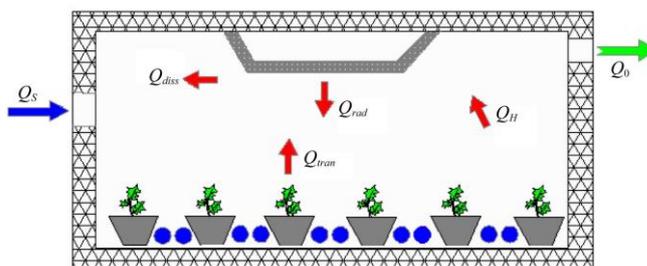


Figure 2 Illustration for the energy balance model in a plant factory

2.3 Temperature control system mechanism model

2.3.1 Load balance model in a plant factory

Generally, the required temperature and relative humidity can be gained for the internal air atmosphere in the closed production system of a plant factory. However, there are some inevitable disturbing factors, which can remarkably change the temperature and relative humidity indoors. These influencing parameters can be called load. According to the first law of thermodynamics, loads of heat and humidity in a plant factory can be expressed as^[36]:

$$Q_1 - Q_0 = \frac{dQ}{dt} \quad (3)$$

where, Q_1 represents the energy in a plant factory per unit time, kJ; $\frac{dQ}{dt}$ is the change rate of stored energy in a plant factory, kJ/s.

Then, the energy change relationship of indoor air can be written as

$$\frac{dQ}{dt} = m \cdot c_{ap} \cdot \Delta t = \rho_a \cdot V \cdot c_{ap} \cdot \frac{dt_a}{dt} \quad (4)$$

where, m is the internal air mass, kg; $m = \rho_a \cdot V$; V is the indoor volume, m^3 ; ρ_a is the indoor air density in kg/m^3 , and $\rho_a = 1.199 kg/m^3$ for the air in a plant factory; c_{ap} is the specific heat of air at constant pressure, and the $c_{ap} = 1.009 kJ/(kg \cdot ^\circ C)$; Δt is the change rate of internal temperature, $^\circ C/s$; t_a is the indoor temperature, $^\circ C$.

The energy in a plant factory per unit time generally includes the heat transferred into the indoor via the treated air and the load energy from the indoor heat source. There is a great influence of exterior climate on the heating load in greenhouses, but in a plant factory, the heating load can be omitted, due to a much smaller effect than the internal loads caused by lighting and plants, particularly when the system is running smoothly, and no human activity disturbing. Thus, Equation (5) can be gained in this case.

$$Q_1 = Q_{in} + Q_S = Q_S + Q_{rad} + Q_{diss} + Q_H - Q_{tran} \quad (5)$$

where, Q_{in} denotes the heating load caused by the indoor heat source.

In the typical plant factory, the energy flux transferred into the indoor via the treated air Q_S can be written as

$$Q_S = k_s m_s (t_s - t_a) = k_s \rho_a q_s (t_s - t_a) \quad (6)$$

where, k_s is the capacity coefficient of the supply air, $k_s = 1.21 kJ/(m^3 \cdot ^\circ C)$; m_s is the air flow quality of the supply air, kg/s ; q_s is the volume flow rate, m^3/h ; t_s is the temperature of supply air, $^\circ C$.

In a plant factory, the energy flux flowing out during forced ventilation can be written as

$$Q_0 = m c_{ap} (t_a - t_r) \quad (7)$$

where, t_r is the temperature of returning air, $^\circ C$. The rest of the parameters are the same as mentioned above. Because $t_r = t_a$ in a plant factory, then $Q_0 = 0 kJ$, finally, the governing equation can be obtained,

$$\frac{dt_a}{dt} = \frac{Q_m + k_s \cdot \rho_a \cdot q_s \cdot (t_s - t_a)}{\rho_a \cdot V \cdot c_{ap}} \quad (8)$$

2.3.2 Energy consumption of crop transpiration

The crop transpiration directly determines the internal energy balance in a closed system. In practice, the crop transpiration coefficient can be defined as the fraction of the radiation load that dissipated by the crop as the latent heat. An accurate estimate of the crop transpiration coefficient becomes necessary in recent years. The Penman-Monteith crop transpiration method^[37] can be used to evaluate the energy produced by crop transpiration in fully controlled plant factories.

1) Crop transpiration model

In a plant factory, the internal evapotranspiration can be equal to the crop transpiration, as the evaporation of the surface soil is relatively small. The Bowen ratio energy balance method (BREB) can usually be selected to investigate the crop transpiration^[18], then the transfer can be represented by the following equations:

$$\lambda E = \frac{(1 - e^{-kLAI})}{1 + \beta} R_n \quad (9)$$

$$\beta = \gamma \frac{t_1 - t_2}{e_1 - e_2} \quad (10)$$

where, λE is the latent heat consumed by crop canopy transpiration,

W/m^2 ; The leaf area index (LAI) shows an important role to determine the energy balance. The LAI is defined as the ratio of leaf area divided by cultivation panel area, $LAI = A_p/A$, A_p is the area of leaves, m^2 ; A is the area of surface soil, m^2 ; R_n is the total net radiation intensity of the crop canopy, W/m^2 ; k is the extinction coefficient, dimensionless; β is the Bowen ratio, dimensionless; λ is the thermometer constant; $\gamma = 0.646 kPa/^\circ C$; t_1 , t_2 are the air temperature at the high position of 1 and 2, respectively, K ; e_1 and e_2 are the water pressure at the high position of 1 and 2, respectively, kPa .

According to Gaudriaan's study^[6], the equation can be written as

$$e_{in} = RH_{in} \cdot e_{in}^* = RH_{in} \cdot e_0 \cdot e^{\left(\frac{17.4t_a}{239+t_a}\right)} \quad (11)$$

where, e_{in}^* is the saturated vapor pressure of the internal air, kPa ; e_{in} is the actual vapor pressure of the internal air, kPa ; RH_{in} is the internal relative humidity, %; e_0 is the saturated water pressure of the air at the temperature of $0^\circ C$, and $e_0 = 0.6107 kPa$.

The actual water pressure of the air at the position of crop canopy e_{inp} (kPa) can be written as

$$e_{inp} = RH_{in} e_{inp}^* = e_0 RH_{in} e^{\left(\frac{17.4t_p}{239+t_p}\right)} \quad (12)$$

where, e_{inp}^* is the saturated water pressure of the air at the crop canopy, kPa ; t_p is the temperature of the leaves at the crop canopy, $^\circ C$. Then, the Bowen ratio β can be written as

$$\beta = \gamma \frac{t_p - t_a}{e_0 RH_{in} \left(e^{\left(\frac{17.4t_p}{239+t_p}\right)} - e^{\left(\frac{17.4t_a}{239+t_a}\right)} \right)} \quad (13)$$

2) Energy consumption of crop transpiration

The crop absorbs and emits radiation, while exchanges heat with air and transpire. Hence, the energy consumed by crop transpiration in a plant factory can be obtained as

$$Q_{tran} = A_p \lambda E = A_p \frac{(1 - e^{-kLAI})}{1 + \beta} R_n \quad (14)$$

2.3.3 Transfer of sensible heat between crop canopy and air

According to Fick's first diffusion law^[24], Equation (15) can be written for a sensible heat model between crop canopy and air.

$$H = LAI \cdot \rho_a \cdot c_{ap} \frac{(t_p - t_a)}{r_a} \quad (15)$$

where, H is the sensible heat exchange between the crop canopy and the air, W/m^2 ; r_a is the single aerodynamic resistance to the boundary layer of crop leaf, s/m . The value of aerodynamic resistance is taken as $100 s/m$ when the forced air circulation is working, whereas the value is set as $200 s/m$ when the forced air circulation is off^[24]. Thus, the expression of sensible heat transfer between the crop canopy and air can be written as

$$Q_H = A \cdot H = A_p \cdot \rho_a \cdot c_{ap} \frac{(t_p - t_a)}{r_a} \quad (16)$$

2.3.4 Energy balance equation of artificial lighting source

Artificial lighting sources in plant factories mainly include agricultural high-pressure sodium lamps (HPS) and light-emitting diodes (LED). The HPS was used in this study.

Two parts can be allocated for the energy generated by artificial lighting sources. One is dissipated in the form of heating, and another is absorbed by the plant canopy in the form of radiation, while converted into the heat when contacting various surfaces, and then into the internal air via the convection in a plant factory.

The energy balance equation of artificial lighting source can be written as

$$Q_L = Q_{diss} + Q_{rad} = \tau \sum_{i=1}^n p_i t_{W_i} \quad (17)$$

where, Q_L is the energy output when the artificial light source works, W; τ is the power factor of the artificial lighting source, dimensionless; p is the power when a single artificial lighting source is used, W; n is the number of artificial lighting sources; t_W is the operating time of artificial lighting source, h.

The mechanism model of the internal temperature control system in the plant factory can be established as Equation (18).

$$\frac{dt_a}{dt} + \left(\frac{k_s \cdot q_s}{V \cdot c_{ap}} - \frac{A_p}{r_a \cdot V} \right) t_a = \frac{k_s \cdot q_s \cdot t_s}{V \cdot c_{ap}} + \frac{\tau \sum_{i=1}^n p_i t_{W_i}}{\rho_a \cdot V \cdot c_{ap}} + \frac{A_p}{r_a \cdot V} t_p - \frac{A_p \cdot R_n \cdot (1 - e^{-kLAI})}{\rho_a \cdot V \cdot c_{ap} + \frac{\rho_a \cdot V \cdot c_{ap} \cdot \gamma (t_p - t_a)}{e_0 RH_{in} \left(e^{\frac{17.4t_p}{239+t_p}} - e^{\frac{17.4t_a}{239+t_a}} \right)}} \quad (18)$$

In Equation (18), the following explanation can be clarified.

1) The temperature control system in the plant factory presents some typical features, including strong nonlinearity, strong coupling, and large time lag^[19,20]. Particularly, the strong interference and time-varying control are also the dominant characteristics in a plant factory, due to the variation in the air volume and temperature of temperature device, the type and working time of artificial lighting source, the types of crops grown in the greenhouse, and the planting pattern of the crops.

2) For the internal temperature, the main influencing factors are, the volume flow of supply air (q_s), the temperature of supply air (t_s), the power of artificial lighting source (p), the running time of artificial lighting sources (t_W), and the internal relative humidity (RH_{in}). Reversely, the internal relative humidity can also be affected by parameters q_s , t_s , p , t_p .

3) There are obvious delays in the influence of the internal relative humidity on the internal temperature. The reason can be that the volume flow of supply air (q_s), and the temperature of supply air (t_s), first imposed on the internal temperature (t_a), and then the varied internal temperature can impose on the internal relative humidity.

3 Modeling and analysis of temperature control system

3.1 Determination of model type

For modeling a temperature control system, it is necessary to comprehensively evaluate the influence of various environmental factors on temperature^[19,38], and thereby to select the dominant influence factors for the optimization of the model. There is a complicated mechanism in the greenhouse system, with emphasis on some environmental factors that change relatively slowly in a short-term evaluation. According to the above-mentioned mechanism model, the relationships between the variables q_s , t_s , p , RH_{in} and t_a , and these variables' influences level on t_a in temperature model were identified. Through taking these environmental factors including indoor humidity, supply air temperature, air volume flow as well as light into consideration, the changes of indoor temperature were predicted. Normally, the temperature control system can be assumed as a multiple-input single-output (MISO) system. The output parameter of the temperature control system can be set as the measured value of the internal temperature, whereas the input can be the volume flow of supply air (q_s), the temperature of supply air (t_s), the power of the

artificial lighting source (p), and the internal relative humidity (RH_{in}). Based on the controllability of artificial lighting sources and the characteristics of forced circulation of the internal air within the closed system of the plant factory, the controlling equipment can be continually controlled and adjusted, and the input and output signals of various parameters always vary in constant impulse.

In this study, a system identification method based on the in-out observation was selected to accurately predict the temperature behavior in the plant factory. Specifically, the run data of a system was first used to identify the structure and parameters of the model, and then to establish a dynamic analysis model, finally to predict the temperature variation from its past trajectory in a plant factory^[39]. Based on environmental data collected from the test, the correlation analysis is usually utilized to verify the feasibility of these selected variables for the temperature control system model. Consequently, the prediction of temperature control systems at a certain time in the future can be gained, mostly starting from the data of influencing factors using various mathematical methods.

Random Forest Regression (RFR) can be seen as a kind of integrated machine learning method, which applied random resampling technology bootstrap and random splitting of nodes technology to construct multiple decision trees for obtaining the final classification results by voting. RFR has some advantages such as higher accuracy, the controlled generalization error, fast convergence speed as well as a few adjustable parameters to effectively avoid overfitting. It is suitable for all kinds of operation of data sets, especially for ultra-high dimensional feature vector space. Meanwhile, RF has the ability to analyze complex interaction classification features, and it has good robustness for noise data and data with missing values. Hence, this paper applies Random Forest Algorithm for model identification and regression prediction.

3.2 Random forest theory

3.2.1 CART model

The random forest can be defined as^[40-43],

$$Forest = \{h_1(X), h_2(X), \dots, h_k(X)\} \quad (19)$$

where, $h_k(X) = h(x, \theta_k)$, θ is the random vector of each decision tree, $h(X, \theta)$ is the output value, k is the number of decision trees.

Only two parameters can be adjusted in a random forest approach, the number of trees in the forest, and the selected number of key split features for each tree. Under the law of large numbers (LLN), the random forest algorithm can represent high classification accuracy, while preventing overfitting during the training process.

In random forest, the classification and regression trees (CART) algorithm is a powerful and popular predictive machine learning technique. This learning method can be used to output the conditional probability distribution of the random variable Y , under the condition of the given input random variable X .

Suppose that X is the input vector, Y is the output vector, particularly, Y is a continuous variable, then a given training set can be written as

$$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\} \quad (20)$$

Assuming that the input space was divided into M elements, denoted by R_1, R_2, \dots, R_M , each of which has a fixed output value, the regression tree model can be expressed as

$$f(x) = \sum_{m=1}^M c_m I(x \in R_m) \quad (21)$$

The squared error $\sum_{x_i \in R_m} (y_i - f(x_i))^2$ can be used to represent the prediction error of the regression tree on the training set.

The optimal value of c_m on the element R_M can be denoted by \hat{c}_m which is the mean value of the output y_i , corresponding to all input instances x_i , on the above element R_M , namely,

$$\hat{c}_m = \text{ave}(y_i | x_i \in R_m) \quad (22)$$

Using following equation to find the optimal segmentation variable j , and the optimal segmentation point s ,

$$\min_{j,s} \left[\min_{c_1} \sum_{x_i \in R_1(j,s)} (y - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(j,s)} (y - c_2)^2 \right] \quad (23)$$

3.2.2 Analysis of random forest model

1) Convergence of random forest

Given a set of classifiers, $h_1(x), h_2(x), \dots, h_k(x)$, the randomly selected training set for the distribution of vector, X can be obtained using the random vector, Y , where X is the input vector, while, Y is the output vector^[44,45]. Consequently, the margin function for the defining sample points (x,y) can be written as,

$$mg(X,Y) = \text{ave}_k I(h_k(X) = Y) - \max_{j \neq Y} \text{ave}_k I(h_k(X) = j) \quad (24)$$

where, $I()$ is the indicator function; ave_k is the operator of average value.

The margin of a data point is defined as the proportion of votes for the correct class minus the maximum proportion of votes for the other classes. Thus, under majority votes, the positive margin means correct classification and vice versa.

2) Generalization error of random forest

In random forest, the generalization error, also called the generalization gap, can be defined as

$$PE^*(forest) = P_{X,Y}(mg(X,Y) < 0) \quad (25)$$

In random forest $h_k(X) = h(x, \theta_k)$. When the size of a decision tree is large enough, the law of large numbers can be followed. In all sequences, $\theta_1, \theta_2, \dots, \theta_k$, $PE^*(forest)$ almost converges everywhere to

$$P_{X,Y} \left\{ P_\theta(h(X, \theta) = Y) - \max_{j \neq Y} P_\theta(h(X, \theta) = j) < 0 \right\} \quad (26)$$

It infers that the random forest cannot produce the over-fitting problem as the increase of decision trees, but it may produce the generalization error within a certain limit.

This can be expressed formally as a boundary (the maximum) on the generalization error as follows:

$$PE^*(forest) \leq \frac{\bar{p}(1 - S^2)}{S^2} \quad (27)$$

where, S is the intensity of generalization ability; \bar{p} is the average correlation coefficient.

4 Experiment

4.1 Test conditions

A field experiment was carried out continuously in a fully enclosed plant greenhouse with artificial lighting from August 20, 2018 to September 10, 2019 in the Beidahuang Kenfeng Seed Industry Co., Ltd., Harbin province, China. The structure size of the greenhouse is 3.6 m×2.4 m×3.0 m. Four high-pressure sodium lamps (MASTER son-T plus 400W E40 SLU) were used for the artificial lighting. A compressor condensing unit served as the temperature-control device where the power of the condensing unit is 5.5 kW, the design pressure is 3.2 MPa (high-pressure side), and 1.9 MPa (low-pressure side), and the total nominal cooling capacity is 11.7 kW. A CO₂ gas tank was set to ensure the normal

photosynthesis of crops. During the test period, the environmental control facilities work continuously.

A type of cold field maize, *Demeiya No.1*, was set as research materials, produced by Beidahuang Kenfeng Seed Industry Co., Ltd. Harbin City, Heilongjiang province, China. Before sowing, the treatment can be made on the maize seeds, including screening, disinfecting, and coating. The seedling time of maize was set from BBCH-scale 00~13-15, i.e., from sowing to 3-6 leaf stage in the growing phase of maize^[46,47]. A small-scale cultivation platform 0.2 m above the ground was designed, on which some pots with the black calcium soil were placed in the form of a single layer planting. In each time, 16 pots were used for sowing, while two maize seeds in each pot were sown to ensure the seedling emergence rate. After the emergence of maize, the weak seedling was manually removed, whereas the strong one was retained.

The specific environment settings in the greenhouse were that the indoor temperature of 18 °C-26 °C, the CO₂ concentration remained constant at 440 μmol/mol, the internal relative humidity (RH_{in}) of 60%-80%, the power of artificial lighting source of 400 W (single lighting source). The working time was 12 h/d, where the opening time of lighting source was from 20:00 to 8:00 the next day, whereas, the closed time was daily from 8:00 to 20:00. A wireless sensor networks (WSN) monitoring system was used to automatically collect data, while real-time regulate the microclimate in the greenhouse. The collected environmental data were sent to the monitoring computer, sampling and recording once per minute.

4.2 Data acquisition

4.2.1 Measurement of canopy temperature and radiation intensity

The vertical distance between the artificial lighting source and crop canopy was 1.5 m, and remained the same during the whole measurement. A handheld spectrometer (UPRTEK MK-350N) was used to measure the illuminance and photosynthetic photon flux density (PPFD) of the crop canopy. PPFD is defined as the number of micromoles of light that hit any given square meter of the crop each second^[48]. When the artificial lighting source was running continuously to a smooth state, the illuminance at the canopy was 34542 lx, and the PPFD (400-700 nm) was 520 μmol/(m² s), as shown in Figure 3. A sensor transmitter (RSTONG GT20) for temperature and humidity was installed at the height of the crop canopy to measure the canopy temperature.



Figure 3 Measurement data for a) illumination b) PPFD of crop canopy

4.2.2 Measurement of indoor temperature and humidity

At the center of the indoor ground, a temperature and humidity sensor transmitter (RSTONG GT20) was installed 1.5 above the ground level, to measure the indoor temperature and humidity in a greenhouse.

4.2.3 Measurement of volume and temperature of air supply

There were two square vents in the size of 400 mm×400 mm, symmetrically arranged in the wall 2 m above the ground level. An air flow detector (FLUKE 922) was placed in the center of each vent, to measure the volume flow and temperature of the air at the outlet in the greenhouse.

4.2.4 Determination of leaf area

A high-throughput plant imaging system (Scanalyzer HTS) was used to measure the leaf area of canopy crops when the maize seedlings grew to the stage of three leaves and one heart. This image system was developed by Beidahuang Kenfeng Seed

Industry Co., Ltd., Heilongjiang Province, China^[49,50]. Two parts can be divided for the measuring procedure, one was the image recognition for the leaf area of each selected seedling, another is to sum the leaf area of all measured seedlings. Thus, the leaf area of crop canopy can be obtained, as shown in Figure 4.

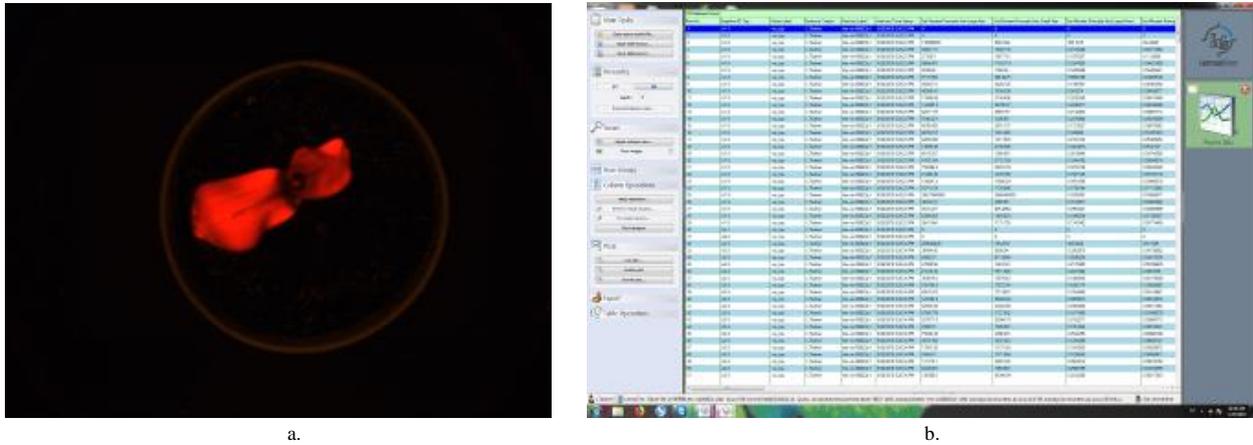


Figure 4 Scanalyzer HTS a) near-infrared imaging and b) software interface of data acquisitionment

5 Results and discussion

5.1 Preprocessing of acquired system data

Table 1 lists small parts of acquired system data extracted from a total of 30 000 sets of measured sampling data. Each set captured in each 60 s from 9:21 on August 20, 2018 to 5:25 on September

10, 2018. The data processing of the test was completed through Python software. The prediction of Random Forest Regression was achieved through the sklearn module. 75% of all data was randomly selected as the training set to train a prediction model using the random forest algorithm (CART). The rest of the measured data was set as the test set.

Table1 Small parts of acquired system sample data

System time	Volume flow of supply air q_s /m ³ h ⁻¹	Temperature of supply air t_s / °C	State of artificial lighting P (open $p=1$, close $p=0$)	Internal relative humidity $RH_{in}/\%$	Internal temperature t_a / °C
2018/8/20 9:21	5.25	23.8	1	80.95	23.51
2018/8/20 9:22	5.25	23.8	1	79.94	23.72
2018/8/20 9:23	5.27	23.9	1	79.84	23.77
2018/8/20 9:24	5.27	23.9	1	78.32	24.08
2018/8/20 9:25	5.29	24.0	1	76.75	24.49
2018/8/20 9:26	5.31	24.1	1	75.03	24.88
2018/8/20 9:27	5.31	24.1	1	74.27	25.03
2018/8/20 9:28	4.78	21.7	1	76.04	24.13
2018/8/20 9:29	4.81	21.8	1	79.88	22.60
2018/8/20 9:30	5.05	22.9	1	72.17	23.92
2018/8/20 9:31	5.16	23.4	1	71.35	23.35
2018/8/20 9:32	5.16	23.4	1	70.13	23.04
2018/8/20 9:33	5.14	23.3	1	66.64	23.35
2018/8/20 9:34	5.14	23.3	1	65.39	23.17
2018/8/20 9:35	5.11	23.2	1	63.86	23.17
2018/8/20 9:36	5.09	23.1	1	62.08	23.28
2018/8/20 9:37	5.09	23.1	1	61.61	23.15
2018/8/20 9:31	5.07	23.0	1	60.09	23.30
2018/8/20 9:39	5.07	23.0	1	59.97	23.28

5.2 Standard error analysis

The mean square error (MSE) was selected to evaluate the proposed prediction model. The statistical parameters can be set as the sum of squares, and the mean for the standard error at the corresponding point of prediction and the raw data. The equation can be written as

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (28)$$

where, MSE is the mean square error; y_i is the raw data; \hat{y}_i is the predicted data; n is the number of prediction points. Generally, the smaller the mean square error is, the more accurate the prediction of the model is.

5.3 Prediction results

5.3.1 Selection of parameters

Based on the acquired data sets, four parameters, $M=4$, were selected to predict the indoor temperature in the greenhouse, including the volume flow of supply air q_s , the temperature of

supply air t_s , the internal relative humidity RH_{in} , and the state of artificial lighting (open, $p=1$, close, $p=0$). When the number of decision trees in the random forest is N , the mean square error of the random forest can achieve the minimum, as shown in Table 2.

Table 2 Mean square error (MSE) of model at various number of decision trees (N)

N	100	200	300	500	1000
MSE	0.0909	0.0911	0.0907	0.0907	0.0907

5.3.2 Evaluation of prediction results

(1) Generalization error of model

To verify the prediction accuracy of the proposed model, 20 samples were randomly selected from all test samples, to compare the actual values of parameters with the predicted values. The selection of the verification approach was because, in the proposed prediction system, the variation of time was not involved in the predicted temperature data, with emphasis on the processing parameters, including the internal relative humidity, the temperature of supply air, the volume flow of supply air, and with/without artificial lighting.

In the test data set in the whole day, the sampling data at 20 moments were randomly selected for numbering, and thereby obtaining the actual and predicted values of the indoor temperature, as shown in Table 3. The comparison results of the actual and predicted internal temperature values as shown in Figure 5. In the randomly selected training data, the smaller distance between the real and predicted values will make higher accuracy for the model prediction, as compared with Table 3 and Figure 5.

Table 3 Actual and predicted values of internal temperature in the whole day

Sample number	Actual internal temperature/ °C	Predicted internal temperature/ °C
1	26.4000	26.3937
2	26.4800	26.2450
3	23.5600	23.3348
4	23.5600	23.5791
5	23.8500	23.7326
6	23.2200	23.0415
7	22.7900	23.2718
8	24.8000	24.8374
9	22.7900	22.872
10	23.3500	23.2117
11	22.7900	22.8421
12	23.5100	23.5153
13	23.5900	23.6861
14	26.5300	26.2154
15	26.1200	26.0854
16	23.0200	23.3149
17	23.2800	23.9107
18	23.6400	23.9947
19	23.5100	24.6729
20	23.5600	23.8855

Artificial lighting sources generally display a periodical variation trend in plant factories. In this study, an individual investigation has been made on the sample data with/without the lighting conditions ($p=1$ or $p=0$), and a continuously lighting state under the whole day. In three cases, 75% of the total data under the different conditions was used to establish the model, whereas, the rest of the data were used to generate the prediction for the evaluation of accuracy. Consequently, the generalization error of

the three models can be achieved, as shown in Table 4.

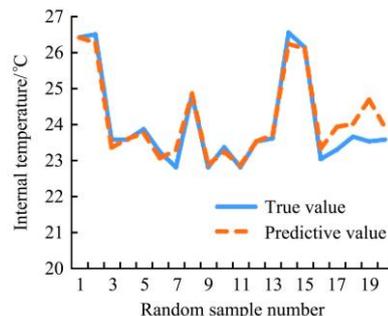


Figure 5 Comparison of actual and predicted internal temperature in the whole day

Table 4 Generalization errors of three model for the artificial lighting

State of artificial lighting	Generalization error
With artificial lighting ($p=1$)	0.1135
Without artificial lighting ($p=0$)	0.0761
Continuously lighting in a whole day	0.0907

The generalization errors of the model were 0.1135, 0.0761, and 0.0907, respectively, for the case with/without artificial lighting ($p=1/p=0$), and continuously lighting in a whole day, indicating that the predictions of the three models were accurate. Specifically, the prediction accuracy of the case without artificial lighting ($p=1$) was the highest, whereas, that of the case with artificial lighting was the lowest. It can be inferred that the usage of artificial lighting source can increase the power consumption of the corresponding control equipment in the greenhouse, and thereby impose the change of indoor temperature. The reason is that the lighting source can inevitably radiate the heat energy to the surroundings, although it has only a little effect on the model prediction.

(2) Weight of parameters

Based on the accurate model prediction, it is necessary to determine the weight value of various parameters influencing the indoor temperature, in order to avoid the effects of the multicollinearity between different attributes on the prediction processing. Thus, the weight value of different parameters can be obtained using the above-mentioned prediction model and the parameters of training samples, in the three cases, with/without artificial lighting ($p=1/p=0$), and continuously lighting in a whole day, respectively. Figures 6-8 show the weight values of four parameters, including the volume flow of supply air (q_s), the temperature of supply air (t_s), and the internal relative humidity (RH_{in}) under the three conditions.

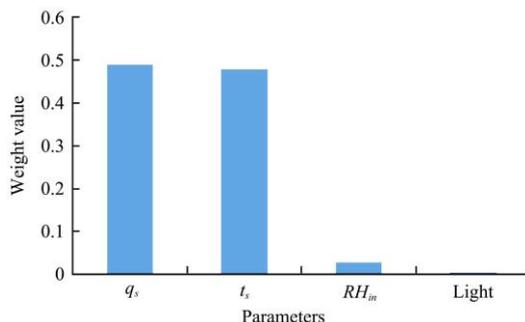


Figure 6 Weight value of the volume flow of supply air q_s , the temperature of supply air t_s , the internal relative humidity RH_{in} , and lighting, under the continuously lighting for a whole day

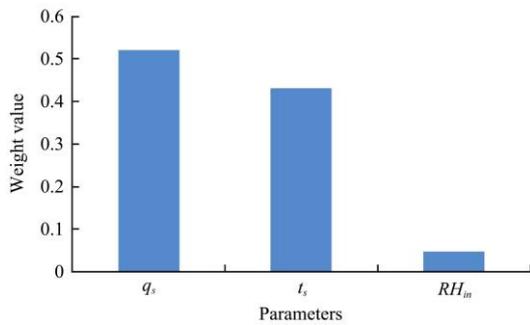


Figure 7 Weight value of the volume flow of supply air q_s , the temperature of supply air t_s , the internal relative humidity RH_{in} , with artificial lighting ($p=1$)

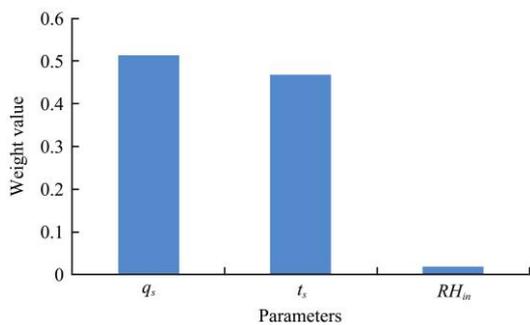


Figure 8 Weight value of the volume flow of supply air q_s , the temperature of supply air t_s , the internal relative humidity RH_{in} , without artificial lighting ($p=0$)

In Figure 6, the weight values for t_s , q_s , RH_{in} , and the lighting source were 0.4889, 0.4778, 0.0283, and 0.005, respectively, under the continuously lighting for a whole day in the greenhouse.

In Figure 7, the weight values for t_s , q_s , and RH_{in} were 0.4306, 0.5214, and 0.0480, respectively, with the artificial lighting ($p=1$) in the greenhouse.

In Figure 8, the weight values for t_s , q_s , and RH_{in} were 0.4669, 0.5142, and 0.0189, respectively, without the artificial lighting ($p=0$) in the greenhouse.

It can be inferred that the temperature of supply air t_s , and the volume flow of supply air q_s , demonstrated large effects on the prediction of internal temperature, whether with or without the artificial lighting in the greenhouse, whereas, the internal relative humidity RH_{in} , can also affect the predicted value, although the weight value was relatively small.

Compared with the observation in the experiments, the main parameters influencing the internal temperature can be the temperature of supply air t_s , the volume flow of supply air q_s , the internal relative humidity RH_{in} , indicating the basically consistent with the analysis of the mechanism model. Although the artificial lighting sources have only little effect on model predictions, the radiated heat energy can significantly increase the power consumption of control equipment in the greenhouse, thereby affecting the changes of the indoor temperature. The value of leaf area index LAI was relatively small, indicating the latent heat of transpiration was relatively lower, compared with the overall power consumption in the greenhouse. This trend can be possibly due to the small number of maize seedlings were planted in this experiment. Furthermore, there was a relatively small effect of the internal relative humidity RH_{in} , on the predicted temperature, and an obvious lag also occurred.

From the analysis above, the temperature mechanism model in plant factories (PFLAs) was constructed based on the modeling of a typical greenhouse system, using the energy balance theory, in

order to obtain the disturbance input affecting the internal temperature in the system. The feasibility of the proposed model was verified by the test method.

Similar results can be found in the research from Meng et al.^[14] in which they proposed to construct the simulation model of thermal environment in solar greenhouses, based on the law of mass-energy transfer in a greenhouse, in order to comprehensively consider the effects of several parameters, including the meteorological conditions outside the greenhouse, covering materials of greenhouses, surface soil, the crops and ventilation inside the greenhouse, crop evaporation and phase change of water vapor, on the internal temperature in the greenhouse. Alternatively, the study from Luuk et al.^[24] also showed that the energy distribution of sensible and latent heat, and crop transpiration model described by a mechanism model of plant factories. The research of Qin et al.^[19] also showed that the correlation between perturbation input. The temperature/humidity in the greenhouse directly determined the identification and simulation of parameters in the system model. Jiang et al.^[20] investigated various factors influencing the dynamic system of the humidity in a greenhouse, and thereby obtaining the microclimate model via the experiment simulation, indicating that the testing method can be used to ensure the reliability of the numerical model. Compared with the selection of the linear static approach in the local domain to approximate the non-linear dynamic model, proposed by Li et al.^[10], the random forest algorithm in this study can be well used to identify and predict the regression of the dynamic model, indicating higher prediction accuracy, while better tracking the model change, and the strong applicability. The method in two previous studies can be consistent with this study. Wu et al.^[43] adopted the tuning features of the generalization error in random forest algorithm to improve the accuracy of prediction. Wang et al.^[27] employed the random forest algorithm to sort the weight values of four basic characteristic variables. Previous experiments also found that the heat dissipation of high-pressure sodium lamps can be greater, thereby increasing the power consumption of environmental control equipment in plant factories. Therefore, the rational selection for the type of artificial lighting source is benefit to reduce the power consumption of plant factories during the growth period of crops.

6 Conclusions

Currently, various greenhouse models are mostly based on a typical greenhouse system model. A high-performance system model is lacking for the optimal control of plant factories. This makes it difficult to be applied in the actual plant factories. In this paper, the mechanism modeling of temperature factors affecting the microclimate of a plant factory has been carried out, to obtain a complex model with perfect mechanism and high accuracy, with emphasis on including the dynamics of the actuator. The necessary simplification and simulation prediction of the complex model have been made to provide a sound reference for the system analysis and controller design for a plant factory. The following conclusions were drawn.

1) Based on the energy balance theory, a temperature mechanism model was established including the dynamic operation of the actuator for a plant factory, and further, the simplified model can ensure the real-time control process.

2) The random forest algorithm was used for the model identification, simulation, and regression prediction. The mechanism model analysis and the prediction model can achieve

basically consistent data, indicating that the reliability of the model was verified by the test method. The main factors that affect the internal temperature can be the volume flow, the temperature of supply air, and the internal relative humidity in a plant factory. Specifically, the weight values of the three parameters were 0.4889, 0.4778, and 0.0283, respectively, whereas, the generalization error of the prediction model was 0.0907.

3) The practical significance of this study is to clarify the modeling method for temperature mechanism in a control-oriented plant factory. This finding can offer a promising insightful approach to establish a model suitable for the system control of plant factories based on the previous models, and further to guide the subsequent research direction in the near future, particularly for the accurate prediction of the internal temperature in a plant factory in modern agriculture.

Acknowledgements

The authors acknowledge Beidahuang Kenfeng Seed Industry Co., Ltd., Heilongjiang Province, China for providing the test platform for this study.

References

- [1] Xu L H, Su Y P, Liang Y M. Requirement and current situation of control-oriented microclimate environmental model in greenhouse system. *Transactions of the CSAE*, 2013; 29(19): 1–15. (in Chinese)
- [2] Kozai T. Resource use efficiency of closed plant production system with artificial light: concept, estimation and application to plant factory. *Proceedings of the Japan Academy. Series B, Physical and Biological Sciences*, 2013; 89(10): 447–460.
- [3] Takakura T. Research exploring greenhouse environment control over the last 50 years. *Int J Agric & Biol Eng*, 2019; 12(5): 1–7.
- [4] Ren Y Z, Wang M J, Saeeda I A, Chen X R, Gao W L. Progress, problems and prospects for standardization of greenhouse-related technologies. *Int J Agric & Biol Eng*, 2018; 11(1): 40–48.
- [5] Song J W. Grow light for plant factory using quantum dot LED. *Journal of International Council on Electrical Engineering*, 2016; 6(1): 13–16.
- [6] Ma X, Lin C H, Qi L, Jiang L K, Tan Y X, Liang Z W, et al. Effect of different lighting quality and intensities on quality of rice seedling by greenhouse stereoscopic nursing. *Transactions of the CSAE*, 2015; 31(11): 228–235. (in Chinese)
- [7] Wu Y, Li L, Li S S, Wang H K. Optimal control algorithm of fertigation system in greenhouse based on EC model. *Int J Agric & Biol Eng*, 2019; 12(3): 118–125.
- [8] Qi F, Zhou X Q, Ding X M, Wei X M. Discussion on classification method of protected agricultural engineering technology. *Transactions of the CSAE*, 2012; 28(10): 1–7. (in Chinese)
- [9] Stefański P, Siedlarz P, Matysik P, Rybka K. Usefulness of LED lightings in cereal breeding on example of wheat, barley and oat seedlings. *Int J of Agric & Biol Eng*, 2019; 12(6): 32–37.
- [10] Dai J F, Luo W H, Xu G B, Li Y X, Wang X C, Ye J, et al. Simulation of greenhouse air temperature, humidity and canopy transpiration in Yangtze River Delta. *Transactions of the CSAE*, 2005; 5: 107–112. (in Chinese)
- [11] Acquah S J, Yan H F, Zhang C, Wang G Q, Zhao B S, Wu H M, et al. Application and evaluation of Stanghellini model in the determination of crop evapotranspiration in a naturally ventilated greenhouse. *Int J Agric & Biol Eng*, 2018; 11(6): 95–103.
- [12] Al Mamun Hossain S A, Wang L X, Liu H S. Improved greenhouse cucumber production under deficit water and fertilization in Northern China. *Int J Agric & Biol Eng*, 2018; 11(4): 58–64.
- [13] Li J, Qin L L, Yue D Z, Wu G, Xue M S, Chen W, et al. Experiment greenhouse temperature system modeling and simulation. *Journal of System Simulation*, 2008; 7: 1869–1875. (in Chinese)
- [14] Meng L L, Yang Q C, Wen J, Zhang Y, Fang H. Visual simulation model for thermal environment in Chinese solar greenhouse. *Transactions of the CSAE*, 2009; 25(1): 164–170. (in Chinese)
- [15] Kozai T. Towards sustainable plant factories with artificial lighting (PFALs) for achieving SDGs. *Int J Agric & Biol Eng*, 2019; 12(5): 28–37.
- [16] Liu N, Ji F, Xu L J, He D X. Effects of LED light quality on the growth of pepper seedling in plant factory. *Int J Agric & Biol Eng*, 2019; 12(5): 44–50.
- [17] Lin D L. Hybrid modeling and analysis of greenhouse control systems. Doctoral dissertation. Shanghai: Shanghai University, 2010; 171p. (in Chinese)
- [18] Wang Xiaochan. Research on microclimate simulation and energy consumption prediction of modern greenhouse in subtropical area. Nanjing: Nanjing Agriculture University, 2003. (in Chinese)
- [19] Qin L L, Ma G Q, Chu Z D, Wu G. Molding and control of greenhouse temperature-humidity system based on grey prediction model. *Transactions of the CSAE*, 2016; 32(S1): 233–241. (in Chinese)
- [20] Jiang Y X, Qin L L, Shi C, Wu G. Modern greenhouse humidity system modeling based on mechanism model. *Journal of Jiangnan University (Natural Science Edition)*, 2013; 12(5): 535–540. (in Chinese)
- [21] Zhang X, He D X, Niu G H, Yan Z N, Song J X. Effects of environment lighting on the growth, photosynthesis, and quality of hydroponic lettuce in a plant factory. *Int J Agric & Biol Eng*, 2018; 11(2): 33–40.
- [22] Wu C F, Deng J S, Wang K, Ma L G, Tahmassebi A R S. Object-based classification approach for greenhouse mapping using Landsat-8 imagery. *Int J Agric & Biol Eng*, 2016; 9(1): 79–88.
- [23] Wang L S, Hou T, Jiang M. Improved multi-objective evolutionary algorithm for optimization control in greenhouse environment. *Transactions of the CSAE*, 2014; 30(5): 131–137. (in Chinese)
- [24] Graamans L, van den Dobbelsteen A, Meinen E, Stanghellini C. Plant factories; crop transpiration and energy balance. *Agricultural Systems*, 2017, 153: 138–147.
- [25] Dong C, Shao L Z, Fu Y M, Wang M J, Xie B Z, Yu J, et al. Evaluation of wheat growth, morphological characteristics, biomass yield and quality in Lunar Palace-1, plant factory, green house and field systems. *Acta Astronautica*, 2015; 111: 102–109.
- [26] Ye H C, Huang W J, Huang S Y, Cui B. Identification of banana fusarium wilt using supervised classification algorithms with AV-based multi-spectral imagery. *Int J Agric & Biol Eng*, 2020; 13(3): 136–142.
- [27] Wang L J, Kong Y R, Yang X D, Xu Y, Liang L, Wang S G. Classification of land use in farming areas based on feature optimization random forest algorithm. *Transactions of the CSAE*, 2020; 36(4): 244–250. (in Chinese)
- [28] Ping R, Zhou S S, Li D. Cost sensitive random forces classification algorithm for highly unbalanced data. *Pattern Recognition and Artificial Intelligence*, 2020; 33(3): 249–257. (in Chinese)
- [29] Cheng Y S, He D X. A photosynthesis continuous monitoring system for CAM plants. *Int J Agric & Biol Eng*, 2019; 12(3): 141–146.
- [30] Wang L N, Wang B R. Greenhouse microclimate environment adaptive control based on a wireless sensor network. *Int J Agric & Biol Eng*, 2020; 13(3): 64–69.
- [31] Yan Z N, He D X, Niu G H, Zhou Q. Growth, nutritional quality, and energy use efficiency in two lettuce cultivars as influenced by white plus red versus red plus blue LEDs. *Int J Agric & Biol Eng*, 2020; 13(2): 33–40.
- [32] Zheng J F, He D X, Ji F. Effects of light intensity and photoperiod on runner plant propagation of hydroponic strawberry transplants under LED lighting. *Int J Agric & Biol Eng*, 2019; 12(6): 26–31.
- [33] Xu F, Shang C, Li H L, Xue X Z. Comparison of thermal and light performance in two typical Chinese solar greenhouses in Beijing. *Int J Agric & Biol Eng*, 2019; 12(1): 24–32.
- [34] Song J X, Xu L, He D X, Tuskagoshi S. Estimating EC and ionic EC contribution percentage of nutrient solution based on ionic activity. *Int J Agric & Biol Eng*, 2019; 12(2): 42–48.
- [35] Wang W S, Wang C, Pan D Y, Zhang Y K, Luo B, Ji J W. Effects of drought stress on photosynthesis and chlorophyll fluorescence images of soybean (*Glycine max*) seedlings. *Int J Agric & Biol Eng*, 2018; 11(2): 196–201.
- [36] Han B Q, Li S L. Theory and application of refrigeration and air conditioning. Beijing: China Machine Press, 2002; 569p. (in Chinese)
- [37] R ós Salazar J D, Candelo-Becerra J E, Hoyos Velasco F E. Growing arugula plants using aeroponic culture with an automated irrigation system. *Int J Agric & Biol Eng*, 2020; 13(3): 52–56.
- [38] Wang Y, Qi X N, Shao J F, Liu Y Z, Li Z X. Effects of light intensity at full growing stage on the growth and yield of different maize varieties. *Journal of Jilin Agricultural University*, 2008; 30(6): 769–773. (in Chinese)
- [39] Wang F L, Dong Z G, Wu Z H, Fang K. Optimization of maize planting density and fertilizer application rate based on BP neural network.

- Transactions of the CSAE, 2017; 33(6): 92–99. (in Chinese)
- [40] Cao Z F. Study on optimization of random forests algorithm. Doctoral dissertation. Beijing: Capital University of Economics and Business, 2014; 134p. (in Chinese)
- [41] Zhang H W, Wang M W, Gan L X. Automatic text classification model based on random forest. *Journal of Shandong University (Natural Science)*, 2006; 3: 139–143. (in Chinese)
- [42] Hu Y R, Li J Q, Liu H G, Fan M P, Wang Y Z. Infrared spectral study on the origin identification of *Boletus Tomentipes* based on the random forest algorithm and data fusion strategy. *Spectroscopy and Spectral Analysis*, 2020; 40(5): 1495–1502. (in Chinese)
- [43] Wu X Y, He J H, Zhang P, Hu J. Power system short-term load forecasting based on improved random forest with grey relation projection. *Automation of Electric Power Systems*, 2015; 39(12): 50–55. (in Chinese)
- [44] Zheng L L, Xu J Y, Wang X L, Liu B G. Study on relationships between vegetation community distribution and topsoil factors based on random forests in shoaly wetlands of Poyang Lake. *Soils*, 2020; 52(2): 378–385. (in Chinese)
- [45] Song X F, Jiao L C. Classification of hyperspectral remote sensing image based on sparse representation and spectral information. *Journal of Electronics & Information Technology*, 2012; 34(2): 268–272. (in Chinese)
- [46] Lancashire P D, Bleiholder H, Den Boom T V, Langelüdde P. A uniform decimal code for growth stages of crops and weeds. *Annals of Applied Biology*, 2008; 119(3): 561–601.
- [47] Weber E, Bleiholder H. Erläuterungen zu den BBCH-dezimal-codes für die Entwicklungsstadien von Mais, Raps, Faba-bohne, Sonnenblume und Erbse-mit Abbildungen. *Gesunde Pflanzen*, 1990; 42(9): 308–321. (in German)
- [48] Yin J, Liu X Y, Miao Y L, Gao Y, Qiu R C, Zhang M, et al. Measurement and prediction of tomato canopy apparent photosynthetic rate. *Int J Agric & Biol Eng*, 2019; 12(5): 156–161.
- [49] Li H, Yin J, Zhang M, Sigrimis N, Gao Y, Zheng W G. Automatic diagnosis of strawberry water stress status based on machine vision. *Int J Agric & Biol Eng*, 2019; 12(1): 159–164.
- [50] Wen D M, Ren A X, Ji T, Flores-Parra I M, Yang X T, Li M. Segmentation of thermal infrared images of cucumber leaves using K-means clustering for estimating leaf wetness duration. *Int J Agric & Biol Eng*, 2020; 13(3): 161–167.