Journal of Intelligent Systems and Control
https://www.acadlore.com/journals/JISC

Adaptive Environmental Control System for Large-Scale Poultry Houses Based on Multif-LSTM

Lin Zhang[1], Jialin Ma[2*], Ashim Khadka[3]

[1] Faculty of Management Engineering, Huaiyin Institute of Technology, 223003 Huaian, China
[2] Faculty of Computer and Software Engineering, Huaiyin Institute of Technology, 223003 Huaian, China
[3] Nepal College of Information Technology, Pokhara University, 44700 Lalitpur, Nepal

* Correspondence: Jialin Ma (majl@hyit.edu.cn)

Received: 07-13-2024 **Revised:** 09-16-2024 **Accepted:** 09-24-2024

Citation: L. Zhang, J. L. Ma, and A. Khadka, "Adaptive environmental control system for large-scale poultry houses based on Multif-LSTM," *J. Intell Syst. Control*, vol. 3, no. 3, pp. 174–185, 2024. https://doi.org/10.56578/jisc030304.

Abstract: The environmental conditions in large-scale, intensive poultry farming systems require high precision, and accurate prediction of environmental factors is critical for effective control. Existing control methods generally focus on the prediction and control of individual environmental factors without considering the interdependencies among these factors, leading to low prediction and control accuracy. To address the complex nature of the environmental system in poultry houses, characterised by multi-factor dependencies, an adaptive environmental control system based on Multi-feature Long Short-Term Memory (Multif-LSTM) was proposed. The Multif-LSTM model within the system calculates the dependencies between environmental factors using correlation coefficients and establishes a multi-input, multi-output neural network architecture. External climate factors are also incorporated during the input phase. Experimental comparisons conducted in a duck house environment, with Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) models, show that the Multif-LSTM model outperforms others in terms of prediction accuracy. For NH_3 concentration, the Root Mean Square Error (RMSE), the Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2) values are 1.34, 8.3, and 0.55, respectively; for temperature, they are 0.29, 2.83, and 0.98; and for relative humidity, they are 1.73, 2.46, and 0.95, respectively. Compared to the average performance of the RNN and LSTM models, the RMSE is reduced by 2.5, MAPE by 4.6, and R^2 increased by 0.32. The results demonstrate that the Multif-LSTM model achieves higher prediction accuracy and is suitable for high-precision adaptive environmental control in poultry houses.

Keywords: Poultry farming; Environmental control; Environmental factor prediction; LSTM; Multif-LSTM

1 Introduction

With the continuous growth in demand for poultry consumption and the increasing environmental standards in the livestock industry, large-scale modern poultry farming has become an essential development pathway for the industry. Medium- and large-scale poultry farming enterprises are transitioning from traditional farming methods to large-scale, intensive, and facility-based systems [1]. Environmental control systems have become a critical component of modern poultry farming operations. In recent years, with the rapid development of new-generation information technologies such as artificial intelligence, the Internet of Things (IoT), and big data, the goals of improving productivity, preventing disease, and reducing environmental impact have become central to smart poultry farming. The farming environment serves as the primary space for poultry activities, where various environmental factors interact and merge to form a complex system. Achieving precise control over this farming environment is one of the key tasks in realizing the intelligent farming process [2–4].

Traditionally, poultry house environments have relied on manual experience for monitoring and management, which results in delayed and inefficient adjustments. In recent years, intelligent monitoring equipment and systems for farming environments have gradually been applied in practical production [5]. However, the dynamic changes in environmental parameters are closely linked to the physiological responses, behavioral characteristics, and growth rates of the animals. Effective management of summer ventilation and heat dissipation, winter insulation and dehumidification, as well as control of harmful gases such as ammonia, hydrogen sulfide, and methane, are critical for creating a healthy and comfortable poultry house environment. This is essential for improving production yield, reducing the incidence of diseases, and minimizing environmental impact. Currently, IoT-based environmental monitoring devices are highly reliable and can accurately sense environmental information within poultry houses. The key technology for the intelligent regulation of poultry house environments lies in the precise prediction of environmental factors through further data fusion [6–8].

The environmental factor prediction models of poultry houses can generally be classified into two categories: statistical methods and machine learning algorithms. Statistical methods rely on extensive experimental data, which is subsequently analyzed to identify patterns. These methods require high precision in experimental data, and their application often involves significant human and time costs. Moreover, the predictive models derived from these methods typically exhibit poor robustness and transferability, making them difficult to generalize. The second category involves machine learning-based models for predicting poultry house environmental factors, which include regression models, grey system theory, Adaptive Neuro-Fuzzy Inference System (ANFIS) models, and neural network models, among others [9–11]. Among these, the RNN is particularly effective in learning time series data, capturing the nonlinear relationships and temporal dynamics within the data. However, the RNN struggles with long-term dependencies within the data. To address this limitation, LSTM networks were developed. The LSTM enhances the RNN by replacing the hidden layer neurons with cell states and three gate structures, allowing for the control of memory and forgetting mechanisms and improving the dynamic prediction of environmental factors within poultry houses [12–16].

Among these models, in order to reduce model input variables, simplify model complexity, and enhance prediction accuracy of the model, scholars have proposed models such as the fusion of the random forest and LSTM for optimizing ammonia concentration prediction in pigsties, as well as LSTM models integrated with attention mechanisms. These models are capable of better capturing the dependencies between poultry house environmental variables and the temporal dependencies within the environmental data sequences [1, 6, 8, 10]. Additionally, models combining Empirical Mode Decomposition (EMD) with LSTM enable the transformation of non-stationary ammonia sequences into stationary ones, fully leveraging the data processing advantages of LSTM neural networks. The combination of the one-dimensional Convolutional Neural Network (CNN) with LSTM into CNN-LSTM models has been found to effectively extract sequential features from time-series data [17–20]. To avoid local optima, optimization algorithms such as the Sparrow Search Algorithm (SSA), Particle Swarm Optimization (PSO), and Genetic Algorithm (GA) have been used to optimize the hyperparameters of neural networks, thereby improving the predictive accuracy of these models [21–24].

In recent years, machine learning, particularly neural network models, has been widely used by researchers for predicting poultry house environmental factors [25–28]. Most of these studies focus on the prediction of individual environmental factors, which limits their application to single-factor control in poultry houses. The poultry house environment is a complex system with multi-factor dependencies, and these methods cannot meet the practical demands of large-scale, intensive farming, leaving a considerable gap from fully intelligent farming systems [29–31]. To address the limitations of existing poultry house environmental control systems, such as the single control model, fixed control parameters, delayed control effects, and poor adaptability, an adaptive environmental control system for poultry farming was developed in this study. This system adopts a multi-environmental factor linkage control approach, which enables precise control based on the dynamic changes in factors such as poultry house temperature, humidity, lighting, ammonia, carbon dioxide, hydrogen sulfide, methane, particulate matter, and external climate conditions. The main contributions of this study are as follows:

a) Development of the Multif-LSTM model for poultry house environmental factor prediction. This model addresses the time-series data characteristics of poultry house environments. Building on the advantages of LSTM in capturing long-term dependencies, it employs a multi-input and multi-output configuration, overcoming the limitations of existing models that can only predict single environmental factors. The proposed model enables the prediction of multiple environmental factors, improving the effectiveness of poultry house environmental control systems.

b) Consideration of the interdependencies between internal and external environmental factors. Special attention is given to incorporating external climate conditions, and correlation coefficients are used to analyze the relationships between each environmental factor and its related factors. As a result, the feature selection for environmental factors is more scientifically efficient, and the prediction accuracy is significantly improved.

2 Model and System

The precise control of poultry house environmental factors is a key technology for achieving intelligent farming. Traditional poultry house environmental control systems suffer from several limitations, including single control modes, fixed control parameters, delayed control responses, and poor adaptability. The time-series data of poultry house environmental factors consists of both linear and nonlinear sequences. Most existing research focuses on the prediction of single environmental factors, often neglecting the interactions between multiple environmental factors, especially the influence of external climate conditions. The proposed Multif-LSTM model incorporates outdoor climate factors as subsequent inputs dependent on the system and employs multi-environmental factor inputs. This model is capable of predicting multiple environmental factors simultaneously, providing enhanced adaptability. The poultry house environmental control system based on the Multif-LSTM model is characterized by multi-factor integration, dynamic regulation, and fine-grained control, enabling more accurate environmental management within poultry houses.

2.1 Control System Framework

The adaptive environmental control system for poultry farming consists of five main components: the perception module, transmission module, computational analysis module, control module, and application module. The overall framework is shown in Figure 1.

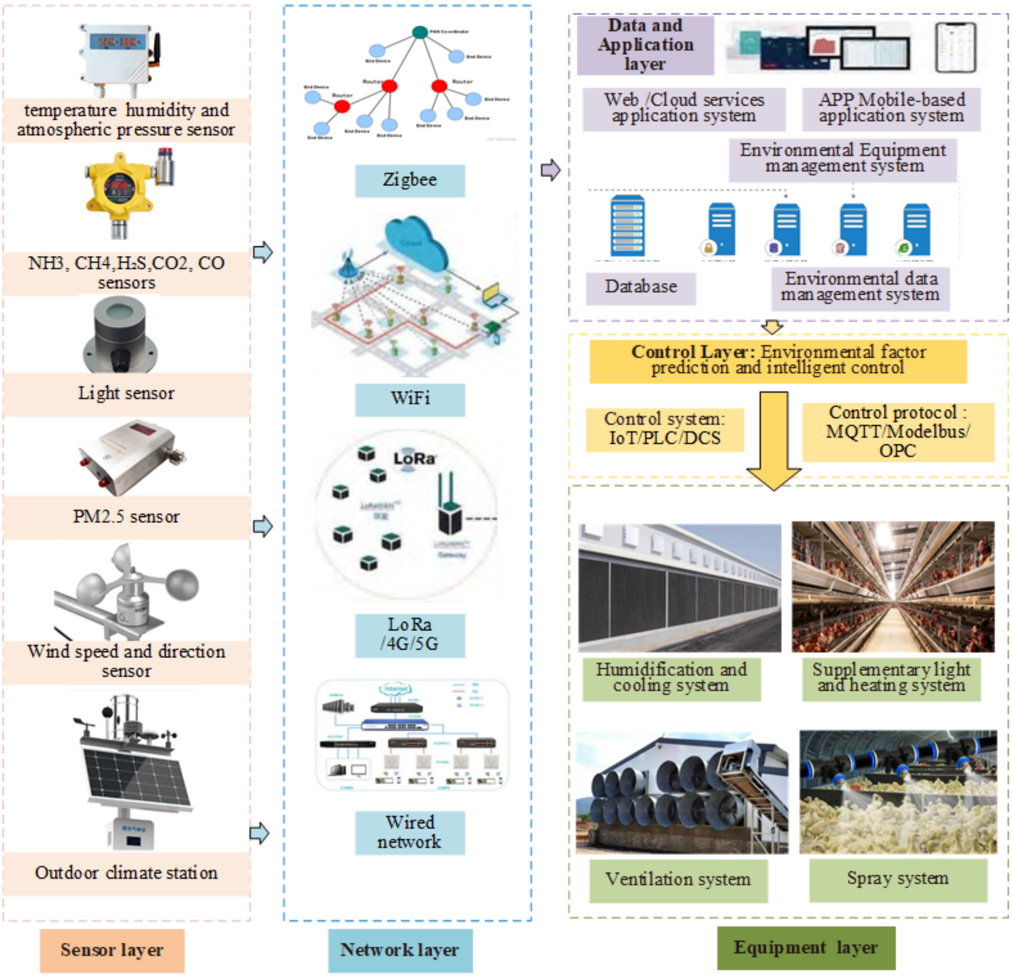

Figure 1. Framework of the adaptive environmental control system for poultry farming

a) Sensor layer

The perception module is responsible for sensing data related to the environmental indicators of the poultry house, primarily focusing on thermal conditions and air quality. The former includes factors such as temperature, relative humidity, lighting, and wind speed, all of which influence the heat and moisture balance within the poultry house and the metabolic heat production of the animals. The latter includes harmful gases, such as carbon dioxide, ammonia, methane, and hydrogen sulfide, which are produced by animal metabolism, manure, bedding, and feed decomposition, as well as dust particles, such as $\mathrm{PM}_{2.5}$, PM_{10}, and total suspended particulates (TSP), which are generated by animal activity and management within the poultry house. These particles can carry bacteria and pathogens and adsorb harmful gases. When the levels of these environmental factors exceed the physiological limits of livestock, cold or heat stress can occur, which typically leads to physiological diseases and adversely affects the health and growth of the animals. Currently, IoT sensor technology is relatively mature and can provide reliable sensors for monitoring these indicators, enabling real-time data collection. Additionally, outdoor climate data can be obtained from local public weather forecasting services or by establishing small weather stations outside the poultry house.

b) Network layer

Scholars, both domestically and internationally, have integrated various types of sensors based on the environmental monitoring requirements of livestock houses and users. Multiple wireless networking approaches have been employed to develop intelligent environmental monitoring devices with different sensor combinations, enabling information sensing and remote control under complex farming conditions [31]. Common data transmission technologies include ZigBee, WiFi, and LoRa. Indoor poultry houses, depending on factors such as area, level of centralization, and transmission distance, may use WiFi, Programmable Logic Controller (PLC), Distributed Control Systems (DCS), or other automation transmission technologies to achieve effective control [4].

c) Data and application layer

The data and application server layer in the IoT system for poultry houses is primarily responsible for the reception, storage, processing, analysis, and application of data. This can be summarized in three key functions:

• Data storage and management: Responsible for storing and managing the vast amounts of data generated within the IoT system. This data includes environmental information collected by sensors and device status, which must be stored efficiently and securely for subsequent analysis and application.

• Data processing and analysis: Involves deep analysis and mining of the stored data to extract valuable insights and information. This includes the use of cloud computing, big data, and artificial intelligence technologies to clean, transform, and aggregate the data in order to support decision-making, predictive analysis, and other functionalities.

• Management and application services: Supports various IoT applications. These applications can include services for specific functions, such as sensor management, data visualization dashboards, on-site health monitoring, and control programs. These services are presented to the end user through user-friendly interfaces in the form of web pages, forms, or mobile applications.

d) Control layer

The control layer is positioned within the data processing and analysis module. It integrates cloud computing, big data, artificial intelligence, and other technologies to conduct deep analysis and modeling of the perceived environmental data, and formulates control commands. The system adapts to the characteristics of the farming environment to achieve precise control over the poultry house environment. This layer plays a command-level role within the overall system. An adaptive poultry farming environment control system was developed in this study, which employs multi-environmental factor interlinked control. It enables precise control based on dynamic changes in poultry house temperature, humidity, light, ammonia, carbon dioxide, hydrogen sulfide, methane, particulate matter, and external climatic factors. The core of the system utilizes the Multif-LSTM model to process environmental data with time-series characteristics.

e) Equipment layer

In the current era of rapidly advancing information technology, the IoT has been effectively applied in intelligent farming, significantly enhancing farming efficiency. Common equipment in poultry houses includes humidification and cooling systems, supplementary light and heating systems, ventilation systems, and spraying systems, among others. These devices are typically controlled by the system, which issues control commands to initiate operations. Advanced poultry house environmental monitoring equipment continues to evolve towards greater intelligence and integration. These devices not only collect various types of environmental data but also possess the capabilities for data storage management, remote transmission, and intelligent control. This enables distributed monitoring and centralized management of systems such as ventilation, heat dissipation, and supplementary lighting.

2.2 Data Processing and Prediction

a) Data processing

In the model, consideration was given to various environmental factors within the poultry house, including temperature, humidity, light, ammonia, carbon dioxide, hydrogen sulfide, methane, particulate matter, and external climate factors. Since the environmental factors collected by sensors are measured in different dimensions and units, normalization is first required. Moreover, not all environmental factors interact with one another. Typically, only a subset of environmental factors are influenced by others. Therefore, the correlation between factors was analyzed using the correlation coefficient. After constructing the correlation coefficient matrix, the top k factors with the highest correlations were selected as input variables for prediction in the model. The correlation coefficient matrix is as follows:

$$C = \begin{bmatrix} \rho(x_1, x_1) & \cdots & \rho(x_1, x_n) \\ \vdots & \ddots & \vdots \\ \rho(x_n, x_1) & \cdots & \rho(x_n, x_n) \end{bmatrix}. \tag{1}$$

The correlation coefficient, also known as the Pearson correlation coefficient, is used to measure the degree of correlation between two poultry house environmental factors, X and Y. This is denoted as ρ_{XY} and can be calculated as follows:

$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{DXDY}} \tag{2}$$

where, $\mathrm{Cov}(X,Y)$ represents the covariance between X and Y, which is used to measure the degree of similarity between the two environmental factors. Specifically, $\mathrm{Cov}(X,Y)$ is defined as $E[(X - EX).(Y - EY)]$, E denotes the mathematical expectation, and D represents the variance.

b) Prediction model

The time series of poultry house environmental factors consists of both linear and nonlinear sequences. Traditional time series forecasting models typically employ single models, such as the Gated Recurrent Unit (GRU) model, CNN model, and Autoregressive Integrated Moving Average (ARIMA) model [18]. These models generally predict only a single environmental factor, with predictions based on the historical data of that specific factor. This approach overlooks the interdependencies and interactions between environmental factors, leading to lower prediction accuracy. Additionally, separate models are required to predict different environmental factors, which results in low computational efficiency.

To address the limitations of traditional poultry farming environment control systems, such as the use of a single control mode, fixed control parameters, delayed control effects, and poor adaptability, particularly the failure to consider the relationships between multiple factors, the Multif-LSTM model with multi-factor inputs and multi-outputs was designed in this study. This model not only accounts for indoor environmental factors but also incorporates the dependency of indoor conditions on external climate factors, thereby enhancing the adaptability of the prediction model.

The Multif-LSTM is a type of RNN. Compared to a conventional LSTM [14], the internal structure of the neural network is similar, with both models featuring a gating mechanism to capture dependencies. The main difference is that the Multif-LSTM adds a multi-feature selection module at the input layer, which selects the relevant dependencies for the environmental factor currently being predicted. Additionally, the overall structure of the Multif-LSTM contains independent processing units for each of the n environmental factors, allowing parallel predictions for different factors. Since the environmental factors at the next time step depend on the values of multiple previous time steps, the model processes the sequence from the past to the present, utilising effective information from later time steps. At each time step in the sequence, the input is passed to n LSTM modules for processing. The predicted values for each environmental factor are then output individually. The structure of the LSTM neural network layer and the Multif-LSTM structure are shown in Figure 2 and Figure 3, respectively.

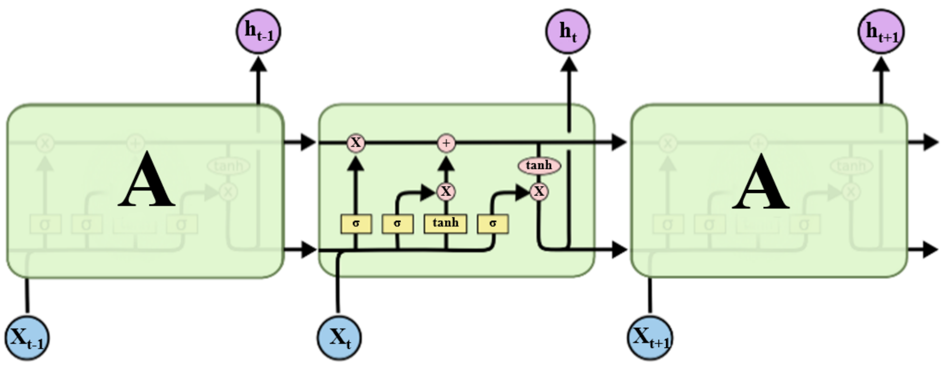

Figure 2. Structure of the LSTM neural network layer

In the Multif-LSTM structure, x_{nt} represents the input values of the n environmental factors at time t. These environmental factor values are processed through the multiple feature selection module, where the correlation coefficients are computed to select the next model input values. After selection, X_{nt} represents the dependent environmental factors of x_{nt}, where $X_{nt} = \{x_{it}, i \in \{0, \mathrm{n}\}\}$. In addition, $y_{nt} + 1$ refers to the predicted output value of the n-th environmental factor for the next time step in the Multif-LSTM model. Within the neural network layer of the Multif-LSTM structure, the forget gate determines which information should be discarded or retained. The information from the previous hidden state and the current input are simultaneously passed through the Sigmoid function. The output value ranges between 0 and 1 , with values closer to 0 indicating that the information should be forgotten, and values closer to 1 indicating that it should be retained. The f function controls which data is discarded or retained.

The input gate is used to update the cell state. The information from the previous hidden state and the current input are passed through the Sigmoid function to adjust the output values between 0 and 1, determining which information should be updated. A value of 0 indicates unimportant information, while 1 indicates important information. Additionally, the hidden state and current input are passed through the Tanh function, which compresses the values between -1 and 1 to regulate the network. The Tanh output and the Sigmoid output are then multiplied together, with the Sigmoid output determining which information in the Tanh output is important and should be retained.

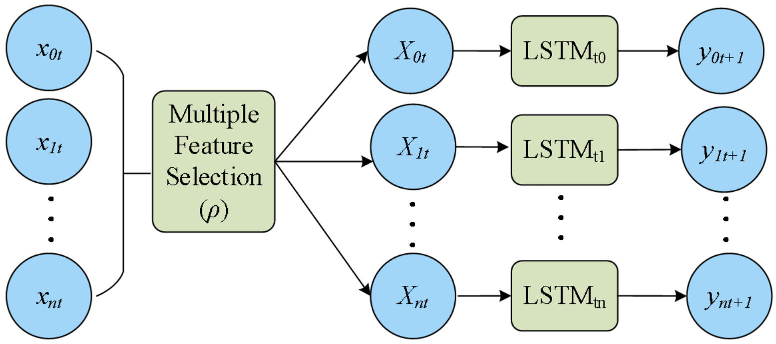

Figure 3. Structure of the Multif-LSTM

Once the current state has been determined, the output gate, denoted as O_t, controls the output of C_t. The output gate decides the value of the next hidden state, which contains relevant information from the previous inputs. The hidden state is also used for prediction purposes. Initially, the previous hidden state and the current input are passed through the Sigmoid function. Next, the newly obtained cell state is passed through the Tanh function. The output of the Tanh function and the Sigmoid function are then multiplied to determine which information should be carried by the hidden state. Finally, the hidden state is output as the current cell output, and both the new cell state and new hidden state are passed on to the next time step. Specifically, the equations for the input gate, the forget gate, the cell state update, the cell state, the output gate, and the hidden state in the LSTM module are as follows:

$$i_t = \sigma\left(W_{ix}x_t + W_{ih}h_{t-1} + b_i\right) \tag{3}$$

$$f_t = \sigma\left(W_{fx}x_t + W_{fh}h_{t-1} + b_f\right) \tag{4}$$

$$\tilde{C}_t = \tanh\left(W_{cx}x_t + W_{ch}h_{t-1} + b_c\right) \tag{5}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{6}$$

$$o_t = \sigma\left(W_{ox}x_t + W_{oh}h_{t-1} + b_o\right) \tag{7}$$

$$h_t = o_t \odot \tanh\left(C_t\right) \tag{8}$$

where, x_t represents the input sequence at time step t, h_t is the hidden state at time step t in the forward LSTM module; C_t is the cell state at time step t; W and b are the model parameters; σ is the Sigmoid function; and $\odot$ denotes element-wise multiplication.

3 Experiment and Analysis

3.1 Experimental Data

The data used in this study was obtained from a large-scale duck farming facility in a city in Jiangsu Province, China. The duck house has an area of approximately 2,000 m^2 and employs an indoor networked floor farming system. The breed used in the facility is the sheldrake, a species of terrestrial duck. The duck house is of a rectangular frame structure, with ventilation windows on both sides, a wet curtain at the front, and a fan at the rear. The facility was also equipped with heating and supplementary lighting systems, with around 40 sensor points installed throughout the house. In addition, a small weather station was located outside the duck house to gather outdoor environmental data. The data sampling interval was set at 30 minutes. A dataset spanning six months of historical data was selected for the experiment, with 90% of the data used for training and 10% reserved for testing. The environmental factors monitored in the duck house and their corresponding data points are outlined in Table 1.

Table 1. Environmental factors in the duck house

Location	Factor
Indoor	Temperature, humidity, atmospheric pressure, wind speed, wind direction, $\mathrm{PM}_{2.5}$, light intensity, ammonia, carbon dioxide, hydrogen sulfide, and methane
Outdoor	Temperature, humidity, atmospheric pressure, wind speed, wind direction, and $\mathrm{PM}_{2.5}$

3.2 Experimental Setup and Evaluation

To further validate the prediction performance of the model, comparisons were made between the proposed Multif-LSTM model and both RNN and LSTM models [14]. The hardware environment for the model experiments included an Intel Core i7 14700k processor, an NVIDIA GeForce RTX 4090 graphics card, and a 64-bit Windows 11 operating system. The model algorithms were developed using Python 3.8. The main parameter settings for training the three models were as follows: input_size=17, hidden_size=128, num_layers=2, output_size =11, learning_rate=0.001, batch_size=32, epochs=10, and dropout=0.5. The models' performance was comprehensively evaluated using RMSE, R^2, and MAPE. The corresponding formulas are given in Eqs. (9)-(11).

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - y_i')^2} \tag{9}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(y_i - y_i')^2}{\sum_{i=1}^{m}(y_i - \bar{y})^2} \tag{10}$$

$$\mathrm{MAPE} = \sum_{i=1}^{m}\left|\frac{y_i - y_i'}{y_i}\right| \times \frac{100}{m} \tag{11}$$

where, y represents the actual value, y' represents the predicted value, m is the number of samples, and $\bar{y}$ is the average value.

To provide a more intuitive assessment of the model's effectiveness, the average error between the predicted results for all environmental factors and the actual values was used to visually reflect the prediction accuracy of the model. This is represented by the following Eq. (12):

$$\eta = \frac{1}{\mathrm{n}}\sum_{i=1}^{\mathrm{n}}(\mathrm{y}_i - \bar{y}) \tag{12}$$

where, η represents the evaluation error, y_i denotes the predicted value for the i-th factor, and $\bar{y}$ is the average value of the actual values.

Furthermore, due to the different dimensions of the environmental factors, they cannot be directly input into the model for training. Therefore, normalization is required for preprocessing the data before the experiment. In this study, the Z-score normalization method was adopted. Z-score normalization, also known as zero-mean normalization or standardization, transforms the data into a distribution with a mean of 0 and a standard deviation of 1. The Z-score formula is given by: $y_{norm} = (y - \mu)/\delta$, where y represents the original data, μ is the mean of the original data, δ is the standard deviation of the data, and y_{norm} is the normalized data. This method is particularly suitable for data distributions that approximate a normal distribution.

3.3 Experimental Results and Analysis

The typical environmental factors of ammonia (NH_3), temperature, and humidity were selected from Duck House No. 3 at the research facility for the 10^{th} of December 2023. These factors were predicted using three models: RNN, LSTM, and Multif-LSTM. The predicted values were then compared with the actual measurements, and the results were plotted to observe the prediction performance of the four models, as shown in Figure 4.

From the experimental results shown in Figure 4, it can be observed that the predictions of NH_3 concentration by the three models fluctuate between 10 and 20 ppm, while the actual NH_3 concentration ranges from a minimum of 12 ppm to a maximum of 20 ppm. Over the course of the day, the actual NH_3 concentration exhibits considerable fluctuations, with the Multif-LSTM model (represented by the red curve) showing the closest match to the actual values (represented by the black curve). The LSTM model also yields relatively good predictions during certain periods, although there are instances where the predicted values deviate significantly. In contrast, the RNN model exhibits large fluctuations, reflecting poorer predictive performance. Additionally, the NH_3 concentration in the duck house is influenced by factors such as ventilation, temperature, and duck activity, leading to substantial variations in NH_3 levels.

From the temperature prediction results shown in Figure 5, it can be observed that the three models predict the temperature in the duck house to fluctuate between 4°C and 15°C, with the temperature during the night being significantly lower than during the day. All three models demonstrate good prediction accuracy, with values fluctuating closely around the actual measurements. The RNN model exhibits larger fluctuations, reflecting poorer prediction performance. The red curve for Multif-LSTM is the closest to the actual values, represented by the black curve. The LSTM model also closely approximates the actual temperature values, suggesting that the temperature in the duck house is primarily influenced by external climatic factors, resulting in relatively stable temperature values.

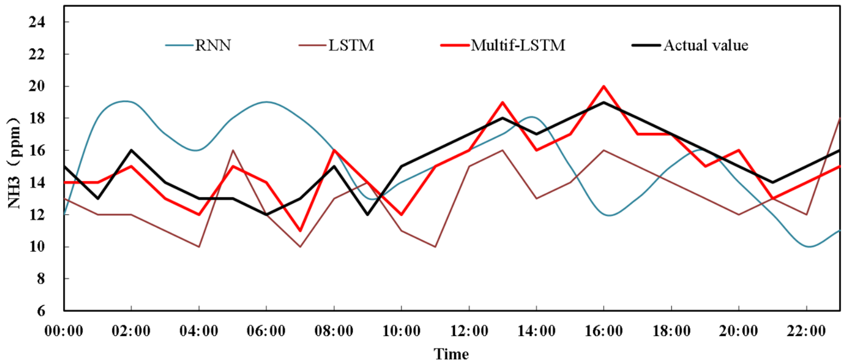

Figure 4. Curves for NH_3 predicted and actual values using the three models

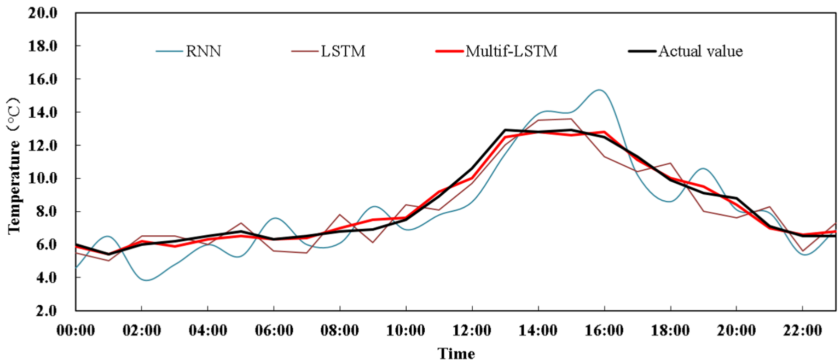

Figure 5. Curves for temperature predicted and actual values using the three models

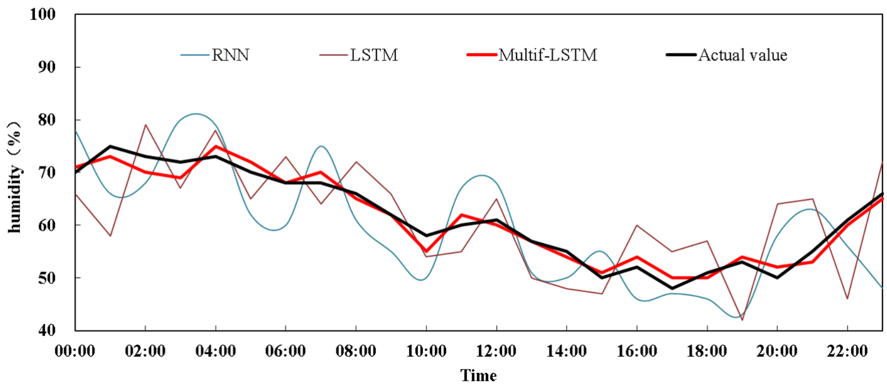

Figure 6. Curves for relative humidity predicted and actual values using the three models

From the experimental results shown in Figure 6, it can be seen that the three models predict the relative humidity in the duck house to range between 40% and 80%, while the actual values vary between 50% and 75%. The fluctuations in the actual values over the course of the day are relatively small, indicating that the relative humidity in the duck house remains stable. This is primarily due to seasonal factors. During winter, when temperatures are lower, the use of wet curtains for cooling in the duck house is minimal, and therefore, the relative humidity is mainly influenced by external climatic conditions. Among the three models, the red curve for Multif-LSTM is the closest to the actual values, represented by the black curve. The predictions from the other two models deviate from the actual values during certain periods, with the fluctuations being approximately 10%. The RNN model shows the largest deviation from the actual values compared to the other models, while the LSTM model exhibits varying degrees of phase-specific deviations.

As shown in Table 2, the Multif-LSTM model provides the best prediction results. For NH_3 concentration in the duck house, the RMSE, MAPE, and R^2 are 1.34, 8.3, and 0.55, respectively. For temperature predictions, the RMSE, MAPE, and R^2 are 0.29, 2.83, and 0.98, respectively. For relative humidity, the RMSE, MAPE, and R^2 are 1.73, 2.46, and 0.95, respectively. In contrast, the RNN model has the poorest prediction accuracy, with RMSE, MAPE, and R^2 for NH_3 concentration being 3.59, 20.3, and 0.23, respectively. The temperature predictions yield RMSE, MAPE, and R^2 of 1.31, 14.04, and 0.72, while for relative humidity, these values are 7.66, 11.5, and 0.16, respectively. Compared to the traditional RNN model, the Multif-LSTM model demonstrates an average decrease in RMSE and MAPE of 73% and 70%, respectively, across the three environmental factors. Additionally, the R^2 increases by 55%. The poor performance of the RNN model can be attributed to the structural and functional differences in the model. The RNN only processes information from the previous node, neglecting longer-range dependencies.

Table 2. Statistical metrics of prediction results for the three models

Models	NH_3(ppm)			Temperature (°C)			Humidity (%)		
	$RMSE$	$MAPE$	R^2	$RMSE$	$MAPE$	R^2	$RMSE$	$MAPE$	R^2
RNN	3.59	20.3	0.23	1.31	14.04	0.72	7.66	11.5	0.16
$LSTM$	3.00	17.83	0.35	0.85	10.02	0.88	7.91	11.7	0.10
$Multif-LSTM$	**1.34**	**8.30**	**0.55**	**0.29**	**2.83**	**0.98**	**1.73**	**2.46**	**0.95**

Compared to the LSTM model, the Multif-LSTM model demonstrates superior performance in predicting NH_3, temperature, and relative humidity. Specifically, for NH_3, the RMSE and MAPE are reduced by 55% and 59%, respectively, while R^2 improves by 59%. For temperature, the RMSE and MAPE decrease by 40% and 47%, respectively, and R^2 increases by 2%. For relative humidity, the RMSE and MAPE are reduced by 79% and 79%, respectively, while R^2 rises by 89%. These experimental results show that the Multif-LSTM model outperforms the LSTM model in all three environmental factors. The Multif-LSTM model, built upon the LSTM framework, incorporates a multi-feature neural network structure, which enables more comprehensive information extraction. Furthermore, it integrates additional environmental factors and external climatic data, leading to improved prediction accuracy.

The Multif-LSTM and LSTM models share a similar core architecture, both being feedforward neural network variants, with Multif-LSTM representing an extension of the LSTM model. In the experiment, LSTM did not account for relevant environmental factors and did not incorporate external environmental data from the duck house, which resulted in a lower prediction performance compared to Multif-LSTM. Additionally, due to experimental constraints, the dataset used in this study covered only one month, which may have limited the model's performance. Moreover, seasonal variations had a significant impact on the results. It is anticipated that with the accumulation of long-term, multi-year data, both Multif-LSTM and LSTM will exhibit better performance, with Multif-LSTM outperforming LSTM.

To further validate the model's effectiveness, the average error percentage between the predicted and actual values of all environmental factors over six months of experimental data was also used to visually reflect the performance of the three models, as shown in Figure 7.

It can be observed from Figure 7 that the RNN model exhibits the highest evaluation error rate at 12.6%, which is approximately 5% higher than that of the LSTM model, making it the model with the highest error rate among the three. The RNN model, which employs self-feedback neurons, is capable of processing time series data of arbitrary length. Its basic structure is relatively simple, storing the network output in a memory unit, which is then combined with the next input to enter the neural network. However, the RNN is a type of RNN that only has short-term memory, retaining information from the previous time step. In the poultry house, environmental factor changes are influenced by preceding time intervals and are not limited to the most recent moment. Additionally, the extent to which current environmental factor values are influenced by historical data varies across different factors. As a result, the short-term memory capacity of the RNN leads to the poorest prediction performance. The LSTM model produces an evaluation error rate of 7.61%, which is approximately 2% higher than that of the proposed Multif-LSTM model. LSTM has long-term memory capabilities and shares a similar core architecture with the Multif-LSTM model. The Multif-LSTM model, however, achieves the lowest evaluation error rate at 5.57%, outperforming the other two models. Compared to LSTM, Multif-LSTM not only retains the advantage of long-term memory but also incorporates multi-feature information, particularly considering the impact of external climatic conditions on indoor environmental factors. This integration results in the lowest prediction error rate.

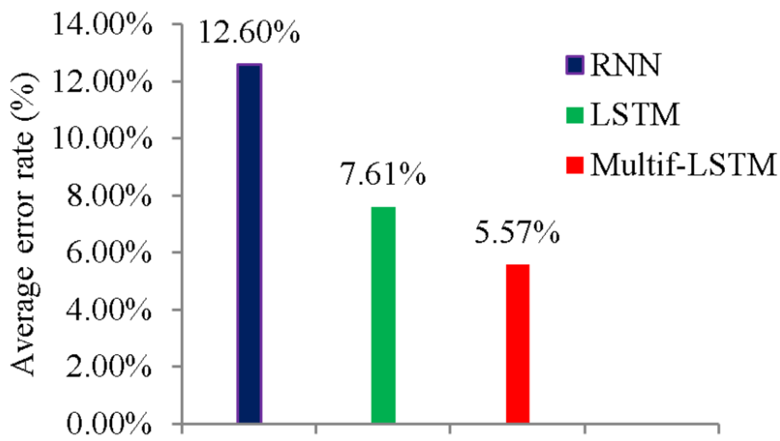

Figure 7. Average error rate of the predictions for the three models

4 Conclusion

Large-scale, intensive poultry farming systems require high-precision environmental control systems. The use of machine learning methods to predict environmental factors for preemptive control has become a new research focus. Most of the methods proposed in current studies address the control of individual environmental factors, without considering the interdependencies between multiple environmental factors in poultry houses. As a result, the accuracy of these predictions remains limited, and there is still a significant gap to achieving fully intelligent farming systems. In this study, an adaptive poultry farming environment control system was developed. This system employs multi-factor interactive control, enabling precise management based on dynamic changes in factors such as temperature, humidity, light, ammonia, carbon dioxide, hydrogen sulfide, methane, particulate matter, and external climatic conditions. The system is based on the Multif-LSTM model, which uses correlation analysis to assess the relationships between environmental factors. A neural network structure was constructed to capture the dependencies in time-series data. During the input phase, external climatic factors were also incorporated into the model. The predictive results obtained from experimental data outperform those from the RNN and LSTM models, exhibiting multi-factor integration, dynamic adjustment, fine-tuning, and adaptability. This model provides an effective approach for achieving intelligent control of poultry farming environments.

The Multif-LSTM model proposed in this study serves as the core of the intelligent poultry farming environment control system. While it demonstrates strong predictive performance, this research did not conduct an in-depth analysis of the model's hyperparameters, nor was the model's time efficiency scientifically validated. These areas will be the focus of future work.

Funding

This research was supported by the Major Natural Science Research Project of Higher Education Institutions in Jiangsu Province (Grant No.: 21KJA210006).

Data Availability

The data used to support the research findings are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Z. D. Guo, Z. Yin, Y. C. Lyu, Y. Z. Wang, S. Chen, Y. Y. Li, W. P. Zhang, and P. Gao, "Research on indoor environment prediction of pig house based on OTDBO–TCN–GRU algorithm," *Animals*, vol. 14, no. 6, p. 863, 2024. https://doi.org/10.3390/ani14060863

[2] E. Arulmozhi, J. K. Basak, T. Sihalath, J. Park, H. T. Kim, and B. E. Moon, "Machine learning-based microclimate model for indoor air temperature and relative humidity prediction in a swine building," *Animals*, vol. 11, no. 1, p. 222, 2021. https://doi.org/10.3390/ani11010222

[3] X. R. Zhang, Z. Q. Gong, G. Allinson, M. Xiao, X. J. Li, C. Y. Jia, and Z. J. Ni, "Environmental risks caused by livestock and poultry farms to the soils: Comparison of swine, chicken, and cattle farms," *J. Environ. Manage.*, vol. 317, p. 115320, 2022. https://doi.org/10.1016/j.jenvman.2022.115320

[4] M. Singh, R. Kumar, D. Tandon, P. Sood, and M. Sharma, "Artificial intelligence and IoT based monitoring of poultry health: A review," in *2020 IEEE International Conference on Communication, Networks and Satellite (Comnetsat)*, Batam, Indonesia, 2020, pp. 50–54. https://doi.org/10.1109/Comnetsat50391.2020.9328930

[5] D. H. Wu, D. Cui, M. C. Zhou, and Y. B. Ying, "Information perception in modern poultry farming: A review," *Comput. Electron. Agric.*, vol. 199, p. 107131, 2022. https://doi.org/10.1016/j.compag.2022.107131

[6] M. Rodríguez, G. Montalvo, J. Morales, M. Jiménez-Martín, M. Aparicio, and C. Piñeiro, "Environmental monitoring and disease prediction," in *Advancements and Technologies in Pig and Poultry Bacterial Disease Control*. Academic Press, 2021, pp. 145–169. https://doi.org/10.1016/B978-0-12-818030-3.00011-8

[7] M. L. Liu, H. X. Chen, Z. Y. Zhou, X. D. Du, Y. X. Zhao, H. Y. Ji, and G. H. Teng, "Development of an intelligent service platform for a poultry house facility environment based on the Internet of Things," *Agriculture*, vol. 14, no. 8, p. 1277, 2024. https://doi.org/10.3390/agriculture14081277

[8] S. Y. Peng, J. M. Zhu, Z. H. Liu, B. Hu, M. Wang, and S. H. Pu, "Prediction of ammonia concentration in a pig house based on machine learning models and environmental parameters," *Animals*, vol. 13, no. 1, p. 165, 2022. https://doi.org/10.3390/ani13010165

[9] F. Qi, X. D. Zhao, Z. X. Shi, H. Li, and W. Y. Zhao, "Environmental factor detection and analysis technologies in livestock and poultry houses: A review," *Agriculture*, vol. 13, no. 8, p. 1489, 2023. https://doi.org/10.3390/agriculture13081489

[10] Q. J. Xie, J. Q. Ni, J. Bao, and Z. B. Su, "A thermal environmental model for indoor air temperature prediction and energy consumption in pig building," *Build. Environ.*, vol. 161, p. 106238, 2019. https://doi.org/10.1016/j.buildenv.2019.106238

[11] E. Küçüktopçu, B. Cemek, and H. Simsek, "Modeling environmental conditions in poultry production: Computational fluid dynamics approach," *Animals*, vol. 14, no. 3, p. 501, 2024. https://doi.org/10.3390/ani14030501

[12] A. A. G. Martinez, I. D. A. Nääs, T. M. R. de Carvalho-Curi, J. M. Abe, and N. D. da Silva Lima, "A heuristic and data mining model for predicting broiler house environment suitability," *Animals*, vol. 11, no. 10, p. 2780, 2021. https://doi.org/10.3390/ani11102780

[13] G. Van Houdt, C. Mosquera, and G. Nápoles, "A review on the long short-term memory model," *Artif. Intell. Rev.*, vol. 53, pp. 5929–5955, 2020. https://doi.org/10.1007/s10462-020-09838-1

[14] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[15] Z. H. Huang, W. Xu, and K. Yu, "Bidirectional LSTM-CRF models for sequence tagging," *arXiv preprint arXiv:1508.01991*, 2015. https://doi.org/10.48550/arXiv.1508.01991

[16] E. Küçüktopçu and B. Cemek, "Modelling indoor environmental conditions in a commercial broiler house," *J. Agric. Sci.*, vol. 25, no. 4, pp. 440–448, 2019. https://doi.org/10.15832/ankutbd.436170

[17] F. Rojano, P. E. Bournet, M. Hassouna, P. Robin, M. Kacira, and C. Y. Choi, "Computational modelling of thermal and humidity gradients for a naturally ventilated poultry house," *Biosyst. Eng.*, vol. 151, pp. 273–285, 2016. https://doi.org/10.1016/j.biosystemseng.2016.09.012

[18] G. Corkery, S. Ward, C. Kenny, and P. Hemmingway, "Monitoring environmental parameters in poultry production facilities," in *Computer Aided Process Engineering, CAPE Forum 2013*, Graz University of Technology, Austria, 2013.

[19] X. Y. Chen, L. J. Yang, H. Xue, L. H. Li, and Y. Yu, "A machine learning model based on GRU and LSTM to predict the environmental parameters in a layer house, taking CO2 concentration as an example," *Sensors*, vol. 24, no. 1, p. 244, 2023. https://doi.org/10.3390/s24010244

[20] M. D. Dumas, S. W. Polson, D. Ritter, J. Ravel, J. Gelb Jr, R. Morgan, and K. E. Wommack, "Impacts of poultry house environment on poultry litter bacterial community composition," *PLoS One*, vol. 6, no. 9, p. e24785, 2011. https://doi.org/10.1371/journal.pone.0024785

[21] S. Y. Lee, I. B. Lee, U. H. Yeo, J. G. Kim, and R. W. Kim, "Machine learning approach to predict air temperature and relative humidity inside mechanically and naturally ventilated duck houses: Application of recurrent neural network," *Agriculture*, vol. 12, no. 3, p. 318, 2022. https://doi.org/10.3390/agriculture12030318

[22] C. Clunie, G. Batista-Mendoza, D. Cedeño-Moreno, H. Calderon-Gomez, L. Mendoza-Pittí, C. M. de la Cruz, and M. Vargas-Lombardo, "Use of data mining strategies in environmental parameters in poultry farms, a case study," in *International Conference on Technologies and Innovation*. Springer Nature Switzerland, Cham,

2023, pp. 81–94. https://doi.org/10.1007/978-3-031-45682-4_7

[23] Y. Wang, W. C. Zheng, and B. M. Li, “A modified discrete grey model with improved prediction performance for indoor air temperatures in laying hen houses,” *Biosyst. Eng.*, vol. 223, pp. 138–148, 2022. https://doi.org/10.1016/j.biosystemseng.2022.08.011

[24] A. Costantino, E. Fabrizio, A. Ghiggini, and M. Bariani, “Climate control in broiler houses: A thermal model for the calculation of the energy use and indoor environmental conditions,” *Energy Build.*, vol. 169, pp. 110–126, 2018. https://doi.org/10.1016/j.enbuild.2018.03.056

[25] N. D. S. Lima, R. G. Garcia, I. A. Nääs, F. R. Caldara, and R. Ponso, “Model-predicted ammonia emission from two broiler houses with different rearing systems,” *Sci. Agric.*, vol. 72, no. 5, pp. 393–399, 2015. https://doi.org/10.1590/0103-9016-2014-0217

[26] D. Hwang, M. J. Rothrock Jr, H. Pang, M. Guo, and A. Mishra, “Predicting Salmonella prevalence associated with meteorological factors in pastured poultry farms in southeastern United States,” *Sci. Total Environ.*, vol. 713, p. 136359, 2020. https://doi.org/10.1016/j.scitotenv.2019.136359

[27] D. K. Al Assaad, M. S. Orabi, N. K. Ghaddar, K. F. Ghali, D. A. Salam, D. Ouahrani, M. T. Farran, and R. R. Habib, “A sustainable localised air distribution system for enhancing thermal environment and indoor air quality of poultry house for semiarid region,” *Biosyst. Eng.*, vol. 203, pp. 70–92, 2021. https://doi.org/10.1016/j.biosystemseng.2021.01.002

[28] E. Küçüktopcu and B. Cemek, “Comparison of neuro-fuzzy and neural networks techniques for estimating ammonia concentration in poultry farms,” *J. Environ. Chem. Eng.*, vol. 9, no. 4, p. 105699, 2021. https://doi.org/10.1016/j.jece.2021.105699

[29] T. G. Omomule, O. O. Ajayi, and A. O. Orogun, “Fuzzy prediction and pattern analysis of poultry egg production,” *Comput. Electron. Agric.*, vol. 171, p. 105301, 2020. https://doi.org/10.1016/j.compag.2020.105301

[30] G. Ahmed, R. A. S. Malick, A. Akhunzada, S. Zahid, M. R. Sagri, and A. Gani, “An approach towards IoT-based predictive service for early detection of diseases in poultry chickens,” *Sustainability*, vol. 13, no. 23, p. 13396, 2021. https://doi.org/10.3390/su132313396

[31] W. F. Pereira, L. da Silva Fonseca, F. F. Putti, B. C. Góes, and L. de Paula Naves, “Environmental monitoring in a poultry farm using an instrument developed with the Internet of Things concept,” *Comput. Electron. Agric.*, vol. 170, p. 105257, 2020. https://doi.org/10.1016/j.compag.2020.105257