agriculture

MDPI

Article

Temperature Prediction of Mushrooms Based on a Data—Physics Hybrid Approach

Mingfei Wang [1,2], Xiangshu Kong [2], Feifei Shan [2], Wengang Zheng [2], Pengfei Ren [3], Jiaoling Wang [4], Chunling Chen [1], Xin Zhang [2,*] and Chunjiang Zhao [1,2,*]

[1] School of Information and Electrical Engineering, Shenyang Agricultural University, Shenyang 110866, China; wangmf@nercita.org.cn (M.W.)
[2] Research Center of Intelligent Equipment, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China; shanff@nercita.org.cn (F.S.)
[3] Institute of Agricultural Resources and Environment, Shangdong Academy of Agricultural Sciences, Jinan 250100, China; zhaoli@city-office.com.cn
[4] Nanjing Institute of Agricultural Mechanization, Ministry of Agriculture and Rural Affais, Nanjing 210014, China; zhaojx@nercita.org.cn
* Correspondence: zhangx@nercita.org.cn (X.Z.); zhaocj@nercita.org.cn (C.Z.)

Abstract: Temperature has a significant impact on the production of edible mushrooms. The industrial production of edible mushrooms is committed to accurately maintaining the temperature inside the mushroom room within a certain range to achieve quality and efficiency improvement. However, current environmental regulation methods have problems such as lagging regulation and a large range of temperature fluctuations. There is an urgent need to accurately predict the temperature of mushroom houses in the future period to take measures in advance. Therefore, this article proposes a temperature prediction model for mushroom houses using a data–physical hybrid method. Firstly, the Boruta-SHAP algorithm was used to screen out the key influencing factors on the temperature of the mushroom room. Subsequently, the indoor temperature was decomposed using the optimized variational modal decomposition. Then, the gated recurrent unit neural network and attention mechanism were used to predict each modal component, and the mushroom house heat balance equation was incorporated into the model's loss function. Finally, the predicted values of each component were accumulated to obtain the final result. The results demonstrated that integrating a simplified physical model into the predictive model based on data decomposition led to a 12.50% reduction in the RMSE of the model's predictions compared to a purely data-driven model. The model proposed in this article exhibited good predictive performance in small datasets, reducing the time required for data collection in modeling.

Keywords: data-physics hybrid; VMD; prediction; mushroom room

Citation: Wang, M.; Kong, X.; Shan, F.; Zheng, W.; Ren, P.; Wang, J.; Chen, C.; Zhang, X.; Zhao, C. Temperature Prediction of Mushrooms Based on a Data—Physics Hybrid Approach. *Agriculture* **2024**, *14*, 145. https://doi.org/10.3390/agriculture14010145

Received: 15 December 2023
Revised: 17 January 2024
Accepted: 18 January 2024
Published: 19 January 2024

1. Introduction

The temperature of the mushroom room is very important for the cultivation process of edible mushrooms [1]. The life activities of edible mushrooms depend on a series of functional proteins, and protein activity is closely related to temperature changes [2]. When the temperature inside the mushroom room exceeds the appropriate upper limit temperature for edible mushrooms, the mycelium grows weakly and is susceptible to invasion by diseases and pests. Once the tolerance limit is exceeded, the mycelium will be irreversibly damaged. When the temperature inside the mushroom room is below the appropriate lower limit, the mycelium slows down or stops. There are also some varieties, such as Lentinula edodes, that only form fruiting bodies when suitable for temperature changes [3]. Therefore, precise temperature control in mushroom houses is an important part of the edible mushroom cultivation process.

Traditional temperature control methods for mushroom houses mainly include the classic on/off method and PID method, which have problems such as low energy utilization and temperature overshoot in practical applications [4]. The model predictive control (MPC) method has been extensively studied and shown to have advantages such as strong tracking ability, energy conservation, and consumption reduction in air conditioning control [5–7]. The three elements of MPC are predictive model, rolling optimization, and feedback correction [8,9], which combine to achieve closed-loop control of the control system. The predictive model is a prerequisite for achieving MPC.

Temperature prediction methods are mainly divided into white box models [10], black box models [11], and gray box models [12] in terms of prediction mechanisms. White box models need accurate distinguishment of thermal parameters for each part in the heat balance equation. These parameters include the design parameters of the main components of the system (including compressor, condenser, and evaporator), as well as the heat carried away by air exchange. This process is a challenging task. With the development of the Internet of Things and deep learning technology, the cost of obtaining a large amount of real-time data has been reduced. In this context, data-driven mushroom house temperature prediction methods, represented by machine learning, have been widely applied, especially deep learning prediction methods represented by LSTM [13], GRU [14], BiLSTM [15], etc. There are still many issues that need to be explored when applying such methods.

The primary issue for accurate prediction is the inclusion of complex nonlinear and non-stationary features in raw data. Some scholars believe that data decomposition methods are effective as they overcome the inherent limitations of a single model [16]. Various time series decomposition methods, such as wavelet transform (WT) and empirical mode decomposition (EMD), have been proposed. The performance of WT largely depends on the function and threshold [17]. EMD also has issues [18] such as modal mixing and boundary effects. Therefore, various improved EMD methods have been proposed, such as Empirical Mode Decomposition (EEMD) [19], Complementary Ensemble Empirical Mode Decomposition (CEEMD) [20], Complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) [21], and Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (ICEEMDAN) [22]. While these improved EMD methods lack a mathematical theoretical foundation on one hand [23], on the other hand, they also struggle with noise not being eliminated in EEMD and CEEMDAN. In contrast, Variational Mode Decomposition (VMD) can reduce data volatility and effectively suppress noise. It is currently considered the most effective data decomposition technique [24,25]. Li et al. [26] conducted a comprehensive comparison of EMD, EEMD, VMD, and Empirical Wavelet Transform (EWT) techniques for building heat consumption prediction modeling, and found that VMD achieved the best effect.

However, pure data-driven methods still face the following challenges [27]:

1. Poor generalization performance: The strong dependence of most deep learning methods on data may lead to poor transferability performance of prediction models. There are constraints on the thermal equilibrium equation for the temperature data of mushroom houses.
2. Poor interpretability: Non-interpretable artificial intelligence models, such as neural networks, hide their internal logic from users, posing inherent risks. It is crucial to clarify the interactions between input features and output labels. The SHapley Additive exPlanations (SHAP) method facilitates a more appropriate interpretation of the outputs of machine learning models [28]. The SHAP method has been extensively explored in the literature across various engineering and scientific disciplines [29–31].

The advantage of the gray box method lies in incorporating physical knowledge into the construction model, which is more reliable than the black box model. However, with the increase in available data, the performance of such models has bottlenecks. Simplified resistance–capacitance (RC) models have been widely used in MPC applications, but RC models cannot guarantee improved predictive performance over longer training cycles [32].

Therefore, there is an urgent need to explore a modeling method that uses existing knowledge rules as prior knowledge for deep learning models. The aim is to improve the model's generalization performance and interpretability, while taking advantage of the massive data brought by the Internet of Things technology to overcome issues such as low prediction accuracy caused by complex features in the raw data. This article combines physical rules, data decomposition, and deep learning to form a comprehensive method. It aims to achieve high-precision temperature prediction of mushroom houses in extremely short periods, providing technical support for subsequent model predictive control methods. The novelty of this article lies in:

1. To reduce the computational complexity of the prediction model, the Boruta method was used to select input data related to indoor temperature in mushroom houses. SHAP values for each feature were further calculated to enhance the interpretability of the model.
2. To reduce the noise in the original data and improve the performance of the prediction model, VMD was used to decompose historical temperature data into sub-modal components with different center frequencies.
3. To predict the temperature, each modal component was individually utilized as input features for networks. Subsequently, the predicted temperatures from each network were summed to obtain the final temperature prediction.
4. To ensure the physical boundaries of learning dynamics, the heat balance equation was incorporated into the loss functions of prediction model.

2. Material and Methods

The structure of the proposed model, based on a hybrid method of data–physics, is illustrated in Figure 1. The model consisted of four components: data collection and preprocessing stage, key influencing factor selection, data decomposition, and data prediction.

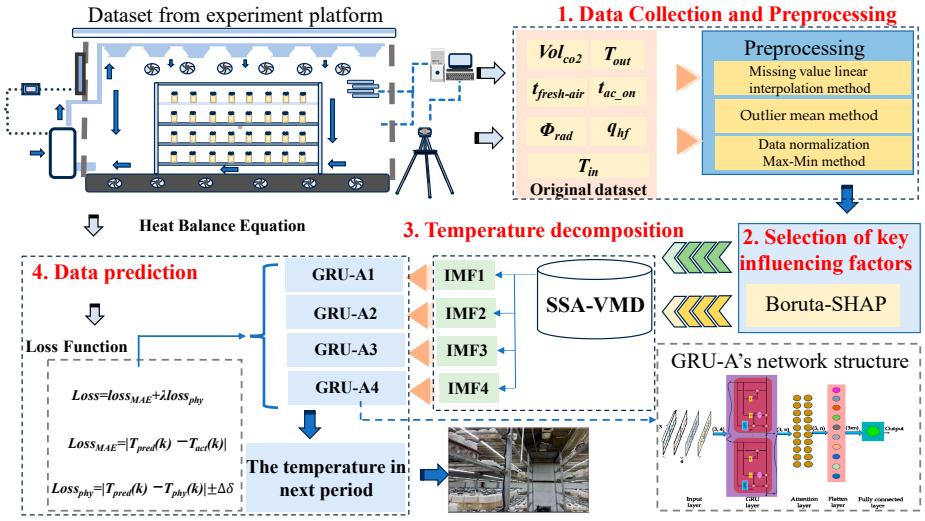


Figure 1. Prediction process of mushroom temperature based on the data–physics hybrid model.

2.1. Data Collection and Preprocessing

The experiment was conducted in a mushroom factory in Beijing, China. The collected data included outdoor temperature T_{out}, solar radiation Φ_{rad}, air conditioning on duration t_{ac_on}, fresh air on duration $t_{fresh-air}$, indoor CO_2 concentration Vol_{CO2}, heat flux from outside to indoor q_{hf}, and indoor temperature T_{in}. The data were collected from 12 May to 12 September 2022, with a 1 min interval. The data collection process can be referred to in ref. [33]. During the experiment, the upper and lower limits of the temperature in the mushroom room were set between 14 °C and 15 °C. For the convenience of display, data were selected from 15 August to 24 August. The duration of air conditioning operation

was displayed in days, while the average indoor temperature was displayed in hours. The collection of all environmental data was illustrated in Figure 2.

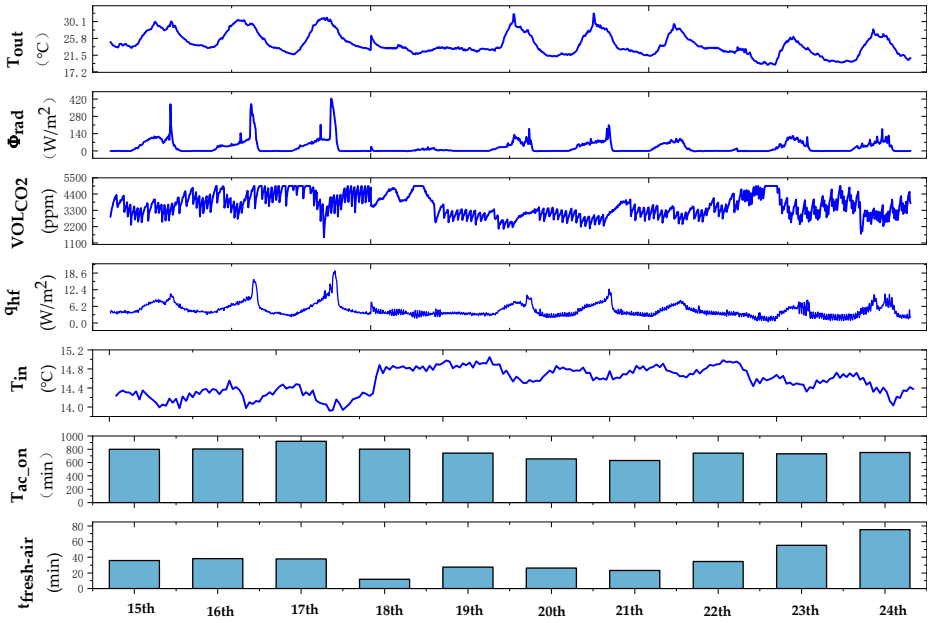


Figure 2. The raw dataset of the mushroom house environment.

After collecting data using sensors, it underwent preprocessing using the missing value linear interpolation method, outlier mean method, and Max–Min data normalization method. All data were divided into a training set, validation set, and test set in an 8:1:1 ratio.

To handle missing data, linear interpolation was employed to fill in the gaps using the data preceding and following the missing position, as depicted in Equation (1):

$$x(k+n) = x(k) + n \times \frac{x(k+m) - x(k)}{m} (0 < n < m) \quad (1)$$

where $x(k+n)$ represents the missing data for $(k+n)$ period, $x(k)$ is the original data for k period, $x(k+m)$ is the original data for $(k+m)$ period.

To address abnormal data, a smoothing filter was applied using the mean method, as illustrated in Equation (2):

$$x(k) = \frac{x(k-1) + x(k+1)}{2} \quad (2)$$

where $x(k)$ is the substituted abnormal data, and $x(k-1)$ and $x(k+1)$ represent adjacent valid data.

To normalize the data, the calculation method is presented in Equation (3):

$$x_n = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (3)$$

where x_n is the normalized value, x_{max} is the maximum value, and x_{min} is the minimum value.

As demonstrated in ref. [33], this article still studied the predicting the temperature of mushroom houses in the next 10 min based on historical 30 min environmental data [33].

The actual production cycle of the factory's Hypsizygus marmoreus is 23 days. The mushroom room needs to be disinfected for a few days without the need for air conditioning during this period. Therefore, the period from 1 August to 24 August was selected as the unit, with a fixed total dataset length of 3312.

2.2. Selection of Key Influencing Factors

This section employed the Boruta–SHAP method to identify the key factors influencing the temperature of mushroom room.

Boruta is a heuristic algorithm based on random forest learners. Its core idea involves constructing shadow features, aggregating the original features and shadow features into a new feature matrix, generating new feature scores, and ultimately selecting the optimal feature set related to the dependent variable. In addition to generating feature sorting, Boruta categorizes features into Confirmed, Tentative, and Rejected types for qualitative evaluation of their importance.

SHAP [28] is a method based on game theory that provides optimal SHAP values to explain the contribution of each feature. The SHAP value represents the level of contribution of each attribute.

Boruta–SHAP feature selection is a method that combines the Boruta and SHAP feature selection algorithms. The SHAP value reflects the importance of each feature in the model's prediction.

The feature selection process was briefly outlined as follows:

1. The Boruta algorithm was employed to determine the feature importance, confirming all features except for the ventilation duration as "Confirmed" and labeling the ventilation duration as "Tentative".

2. The SHAP library was utilized to calculate the SHAP values for each feature falling within the "Confirmed" and "Tentative" range, and subsequently ranked them.

After the above process, the ranking of feature importance is shown in Figure 3.

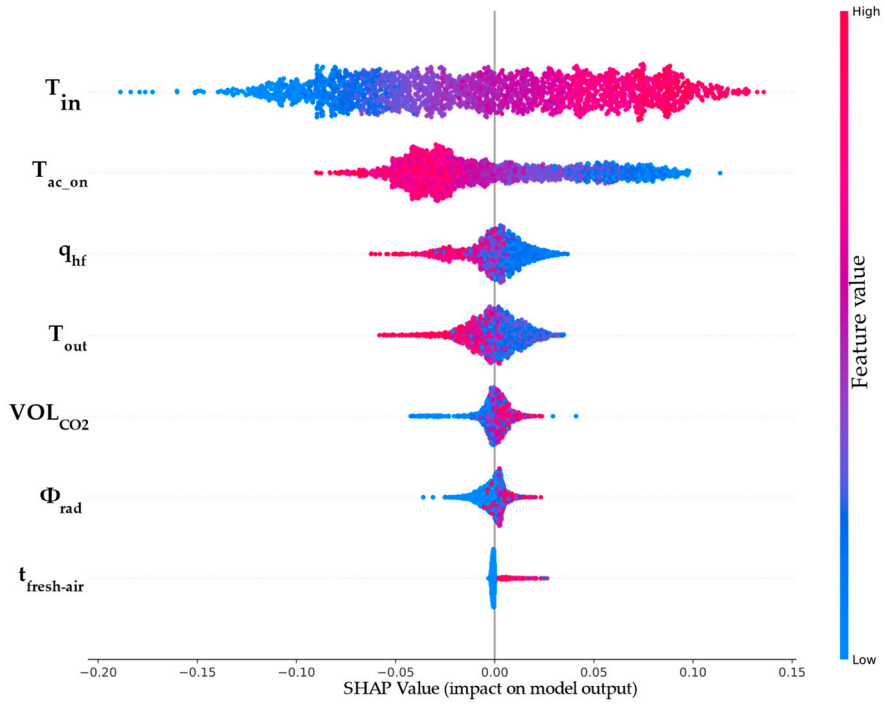


Figure 3. The impact of each feature on the model output.

The previous temperature of the mushroom room had the most significant impact on the results among all the characteristics, followed by the duration of air conditioning. This is illustrated by the vertical sorting of the graph, which represents the ranking of the impact of a certain feature on the predicted temperature. The horizontal axis represents the magnitude of the impact of samples from different periods in a single feature on the predicted temperature. The color represents the size of feature values.

The first row showed the indoor temperature of the mushroom room. The extended length indicates that the primary factor influencing the future temperature of the mush-

room room is historical temperature. The height of this row is used to represent the similarity in data values collected at different time periods. To facilitate visualization, the row height was used to represent this situation. The red dot on the right, representing a high value of mushroom room temperature, indicates a positive impact on future mushroom room temperatures. The blue dot on the left suggests a negative correlation with future temperatures.

The red dot on the second line indicates that the longer the air conditioning is turned on, the lower the temperature in the mushroom room in the future. The blue dot is located on the left side of this line, indicating that the shorter the air conditioning on time, the higher the temperature in the future due to external heat transfer.

In the heat flux section of the third line, the positions of the red and blue dots are roughly similar to factor 2. The magnitude of the heat flux value is also negatively correlated with the future temperature of the mushroom room. The row height of the red dots is smaller than that of the blue dots, indicating that the heat transferred from the outside to the inside through the enclosure structure mainly occurs slowly.

The outdoor temperature section in the fourth line exhibits a distribution similar to that of factor 3 in the graph. This is due to the fact that the heat transferred from solar radiation and external temperature to the mushroom room must pass through the wall, with the heat flux accurately reflecting this process. The CO_2 concentration section in line 5 represents the heat generated by mushroom respiration. The higher the concentration, the higher the temperature in the mushroom room in the future period. Similarly, solar radiation in the sixth line is positively correlated with the future temperature of the mushroom house. In the fresh air section of the seventh line, there are many samples with SHAP values of 0, indicating that the duration of fresh air opening is not significantly related to the future temperature. However, a small number of red dots on the left also indicates that prolonged ventilation can cause an increase in temperature.

It is worth noting that for a factor with a mixture of red and blue, the correlation with the results should be determined by density distribution. For example, there are two blue dots on the far right of CO_2 concentration. However, the overall red color on the right indicates a positive correlation between CO_2 concentration and future temperature.

2.3. Mushroom Temperature Decomposition

VMD has the capability to decompose signals into a series of subsequences with finite bandwidth [34]. When using VMD to process signals, it is necessary to pre-determine the number of modes for IMF decomposition, denoted as k, as well as the penalty factor α. The value of k determines the number of IMFs through decomposition. A large k value can lead to over-decomposition and increased computational complexity. Conversely, a small value of k may result in loss or aliasing of the decomposed IMFs. α determines the width of the IMF bandwidth. The smaller the value of α, the larger the bandwidth of each IMF component obtained. On the contrary, a larger α value will result in a smaller bandwidth for the IMF component. It is evident that the choice of k and α significantly influences the final outcome of the VMD decomposition [35]. Hence, it is crucial to identify the optimal combination of the number of decompositions (k) and penalty factors (α) when employing VMD.

To mitigate bias stemming from subjective experiences in selecting K and α for VMD, this paper adopted the Sparrow Search Algorithm (SSA) to optimize the VMD decomposition process.

The SSA optimization algorithm, proposed by Xue, J et al. [36], exhibits fewer iterations, faster convergence speed, and stronger search ability, thereby substantially decreasing the training time of the model. In terms of optimization ability, search accuracy, stability, and convergence performance, SSA surpasses biological heuristic algorithms such as GA, PSO, and GWO.

This article took envelope entropy, reconstruction error, and the weighted sum of neural network training time as the objective function, as shown in Equation (4):

$$f = min(aS + be_{VMD} + cK) \quad (4)$$

where f is the objective function value; S is the envelope entropy value; e_{VMD} is the reconstruction error of data after VMD decomposition; and a, b, and c are weighting coefficients.

Set the population size of the sparrow search algorithm to 50, the number of iterations to 100, the proportion of vigilantes to 10% of the total population, and the upper and lower limits of K to [1, 10]. The upper and lower limits of the constraint on α were [1000, 3000]. The final optimization result was K = 4 and α = 1000. The detailed solution process could be found in Algorithm 1.

Algorithm 1: Pseudocode for VMD parameter optimization

Input: Number of modal components K, Penalty coefficient α, original temperature data x;
Output: Optimal number of modal components K_{best} and Penalty coefficient α_{best};
1 Set SSA parameters: Population number, Iteration number N_{iter};
2 K and α were obtained according to the individual position of the sparrow, and the VMD decomposition result $IMF_1 \ldots IMF_K$ is calculated;
3 Calculated $IMF_1 \ldots IMF_K$ local minimum envelope entropy, reconstruction error, neural network training time;
4 The fitness value was calculated according to Equation (4), and sorted to find the best and worst individual;
5 For the iteration variable i, perform the following steps in N_{iter};
6 Updated the location of discoverers, followers, and watchmen;
7 The fitness value was calculated according to Equation (4), and the sparrow position was updated to form a new population;
8 Check whether the maximum number of iterations was reached;
9 $i = i + 1$
10 Obtain K_{best}, α_{best}.

2.4. Data Prediction

After SSA–VMD decomposition, four modal components were obtained. The Gate Recurrent Unit and Attention (GRU–A) model were used to predict each component. Finally, the final prediction was obtained by adding up the modal components.

To clarify the optimal hyperparameters in GRU–A, this paper still used the SSA method to optimize the number of neurons and learning rate hyperparameters of the four GRU–A models. The objective function was to minimize the loss function. Based on experience, the constraint condition was that the number of GRU neurons was between [5,30], and the learning rate was one of 0.1, 0.01, or 0.001. Set the population size in the Sparrow Search Algorithm to 100 and the number of iterations to 10. The calculated results were shown in Table 1.

Table 1. Optimization results of neural network hyperparameters.

Parameter	GRU-1	GRU2-2	GRU-3	GRU-4	Learn Rate
Result	15	15	27	28	0.001

The loss function consisted of a data model and a physical model. The data model adopted the default MAE model. The physical model can be obtained from the thermal equilibrium equation.

The pathways that cause heat changes inside the mushroom house include: (1) heat exchange between the internal environment of the mushroom house and the external environment; (2) the heat produced by mushroom growth; (3) cooling/heating of air

conditioning; (4) the heat exchanged through door gaps, lighting, fresh air, etc. The heat balance equation of the mushroom house is shown in Formula (5):

$$\Delta Q_{indoor} = \Delta Q_{envelope} + \Delta Q_{fungus} + \Delta Q_{AC} + \Delta Q_{other} \tag{5}$$

where ΔQ_{indoor} is the change in overall indoor heat, J;

$\Delta Q_{envelope}$ is the heat transferred from indoors to outdoors through the enclosure structure during a single prediction period, J;

ΔQ_{fungus} is the heat generated by indoor mushroom respiration during a single prediction period, J;

ΔQ_{AC} is the amount of heating provided by the air conditioner to regulate room temperature during a single prediction period, J;

ΔQ_{other} is the indoor heat generated through door gaps, lighting, and fresh air during a single prediction period, J.

$\Delta Q_{envelope}$ can be obtained from Formula (6):

$$\Delta Q_{envelope} = \Delta hf \times D \tag{6}$$

where Δhf represents the heat flux transmitted from indoors to outdoors through the enclosure structure during a single prediction period, which can be directly measured by sensors; D represents the heated surface area of the mushroom room.

ΔQ_{fungus} is the heat generated by indoor mushroom respiration during a single prediction period. During the aerobic decomposition of glucose in the process of mushroom respiration, the energy utilization rate is approximately 40% [37], while the remaining energy is released in the form of thermal energy.

ΔQ_{AC} can be obtained from Formula (7):

$$\Delta Q_{AC} = \Delta t_{AC} \times W_{rp} \tag{7}$$

where Δt_{AC} is the duration of air conditioning operation within a single prediction period, min, W_{rp} is the rated power of the air conditioning system, W.

Due to the relatively small size of ΔQ_{other}, it can be ignored.

$$T_{phy}k - T_{act}(k-1) = (\Delta Q_{envelope} + \Delta Q_{fungus} + \Delta Q_{AC} + \Delta Q_{other})cm \tag{8}$$

where $T_{phy}k$ is the average temperature obtained from the physical model during period k, °C;

$T_{act}(k-1)$ is the average temperature measured during the $k-1$ period, °C;

c is the specific heat capacity of indoor air at constant pressure, kJ/kg °C;

m is is the total weight of indoor air, kg.

The constructed thermal equilibrium model, although interpretable, had limited accuracy. Its mathematical expression was:

$$LOSS_{phy} = \left| T_{pre-IMF1}(k) + T_{pre-IMF2}(k) + T_{pre-IMF3}(k) + T_{pre-IMF4}(k) - T_{phy}(k) \right| \pm \Delta\delta \tag{9}$$

where $T_{pre-IMF1}(k)$, $T_{pre-IMF2}(k)$, $T_{pre-IMF3}(k)$, and $T_{pre-IMF4}(k)$ are the predicted values of each sub-model, respectively;

$\Delta\delta$ is the set error term and takes a value of 0.5.

Using the above formula as a constraint for optimizing the data loss function, we introduce Lagrange multipliers λ to transform the entire loss function into an unconditional constraint form:

$$LOSS_{total} = LOSS_{MAE} + \lambda LOSS_{phy} \tag{10}$$

Based on the principle of having a similar scale for the previous and subsequent items, λ was taken as 0.01.

3. Results and Discussion

3.1. Evaluation of Model Performance by Decomposition Methods

To evaluate the impact of EMD, EEMD, EWT, and VMD on the performance of prediction models, Python library functions such as EMD signal 1.5.1, ewtpy0.2, and vmdpy0.2 were utilized, respectively. EMD and EEMD automatically find the optimal number of modal components, while EWT and VMD employ SSA optimization methods to optimize key parameters. The input features of the model encompassed the historical temperature of the mushroom room, duration of air conditioning operation, indoor CO_2 concentration, and heat flux of the inner walls. The original data were decomposed into 9, 9, 3, and 4 modal components, and each modal component was predicted individually. The decomposition details were presented in Table 2. The predicted R^2, RMSE, and training time were depicted in Figures 4 and 5, respectively.

Table 2. Performance evaluation of different decomposition methods.

Decomposition	Reconstruction Error	Envelope Entropy	Number of Modal Components
EMD	1.93×10^{-16}	10.993	9
EEMD	1.92×10^{-14}	11.059	9
EWT	7×10^{-3}	11.030	3
VMD	1.9×10^{-2}	10.886	4

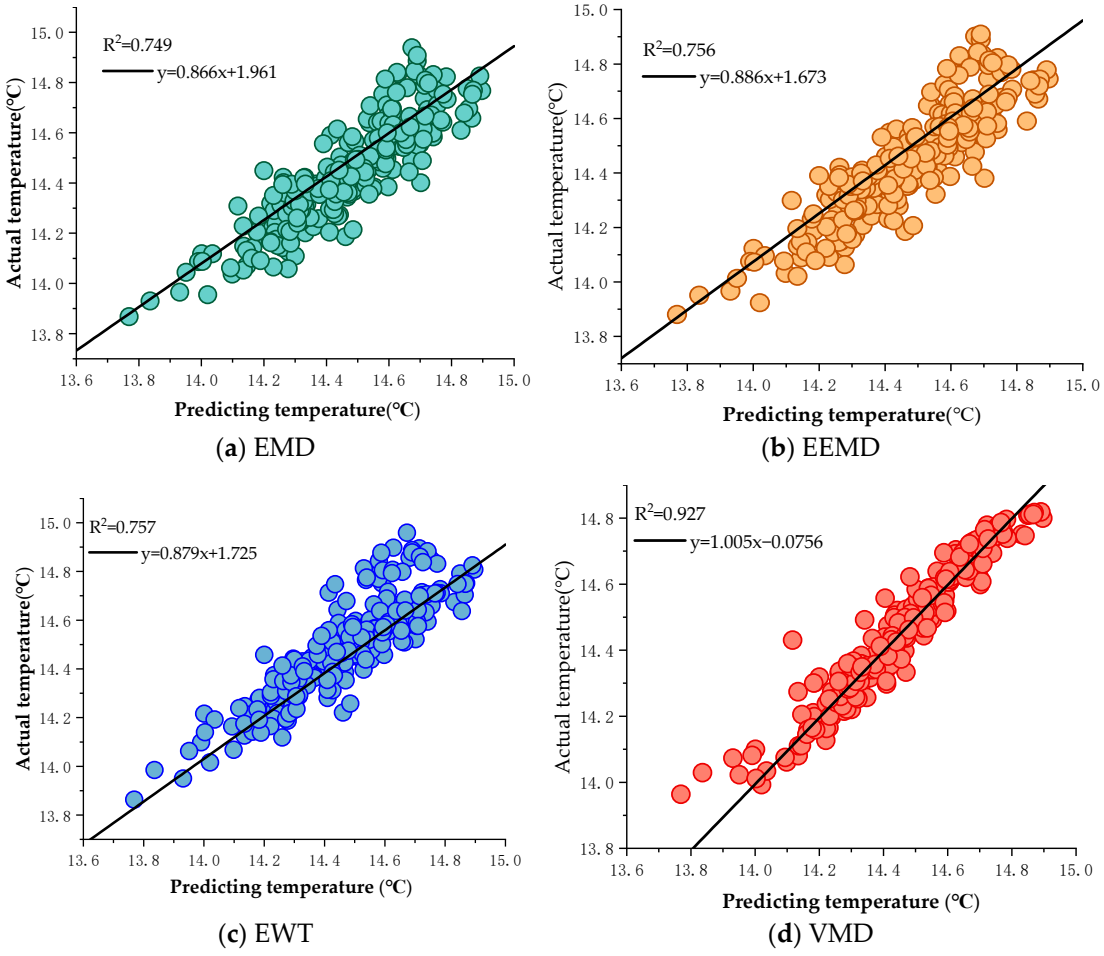
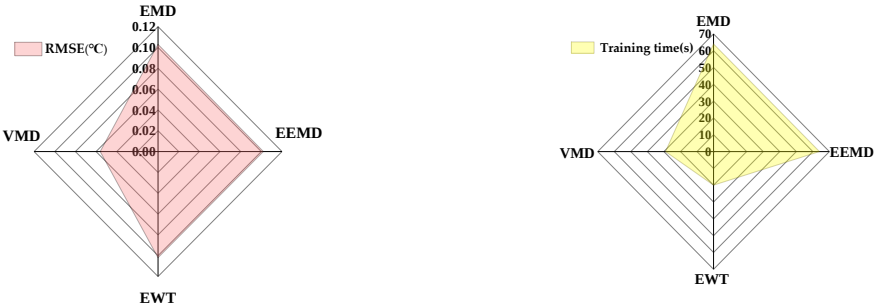


Figure 4. Comparison of R^2 values based on different decomposition modes: (**a**) model based on EMD decomposition; (**b**) model based on EEMD decomposition; (**c**) model based on EWT decomposition; (**d**) model based on VMD decomposition.



(**a**) Comparison of RMSE values (**b**) Comparison of training time

Figure 5. Comparison of predictive performance of models based on different decomposition modes: (**a**) comparison of RMSE values predicted by models under four decomposition modes; (**b**) comparison of training time for model prediction under four decomposition modes.

According to Table 2, it can be observed that the EMD decomposition method exhibited the highest reconstruction accuracy for this particular feature. On the other hand, the VMD decomposition method demonstrated the best envelope entropy. It is worth noting that envelope entropy was computed by analyzing the signal's envelope, which effectively captured the complexity of the signal envelope and, to some extent, reflected the patterns of data changes.

Figures 4 and 5 demonstrated that the models utilizing EMD, EEMD, and EWT decompositions yielded similar results in terms of predicting R^2 and RMSE. The training time of the prediction model employing EWT decomposition was reduced by 68% compared to other methods. Specifically, when compared to the model using VMD decomposition, the training time was reduced by nearly 10 s. This reduction can be attributed to the correlation between the number of modal components decomposed by EMD, EEMD, and VMD.

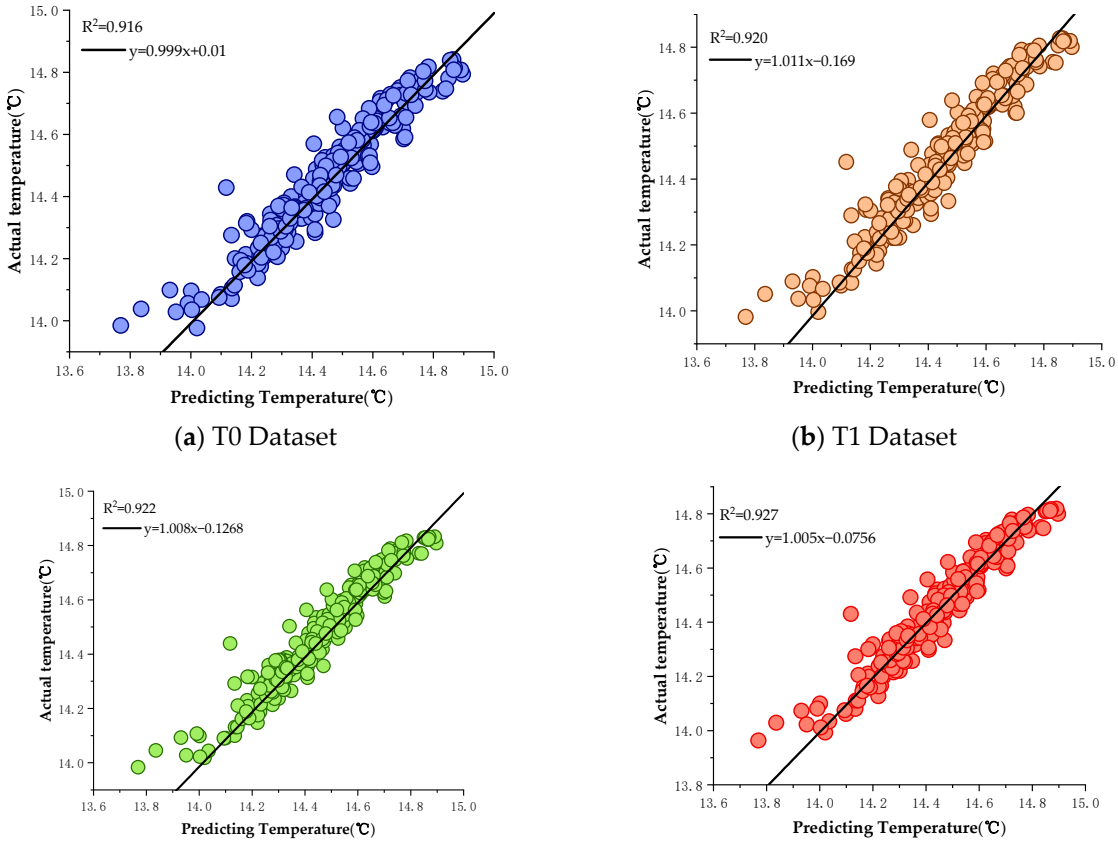
Furthermore, compared with the prediction model decomposed by EMD and other methods, the RMSE of the model decomposed by VMD decreased by 45.10%, and the R^2 increased by 22.46%, respectively. This improvement can be attributed to the fact that the VMD decomposition method, optimized by SSA, yielded the smallest envelope entropy. The method effectively reduced the volatility of the original data, suppressed noise, and enhanced prediction accuracy. These findings suggest that the purpose of decomposition should be to maximize the representation of features rather than to restore the original data losslessly.

3.2. Evaluation of Model Performance on Vary Feature

To assess the predictive performance using the Boruta–SHAP algorithm, various combinations of factors were selected and input into the SSA-VMD-GRU-A-PHY model. The complete set of collected features was used as dataset T0. A significant number of samples in the duration of fresh air opening exhibited SHAP values of 0. Additionally, the duration of fresh air opening was considered tentative. In reality, the duration of fresh air opening is approximately 2 min per hour, and it had no noticeable impact on indoor temperature changes. Therefore, this feature was excluded, resulting in the formation of dataset T1. The two factors of outdoor temperature and solar radiation contributed to heat transfer through walls, thereby affecting the indoor environment. The heat flux flowing through the inner wall in the data measured reflected this heat. Therefore, the outdoor temperature and solar radiation were filtered out. The dataset T2 was formed by the new air opening time, indoor temperature, heat flux flowing through the inner wall, air conditioning opening time, and indoor CO_2 concentration. Based on dataset T2, the duration of fresh air opening was excluded to form dataset T3.

By comparing Figure 6a–d, it showed that excluding the duration of fresh air opening, R^2 increased by 0.44% and 0.54%, respectively. Figure 7 showed that RMSE decreased by 1.67% and 3.45%, respectively, and the time was shortened by 0.3 and 0.5 s, respectively. This phenomenon arose due to the network's acquisition of redundant features throughout

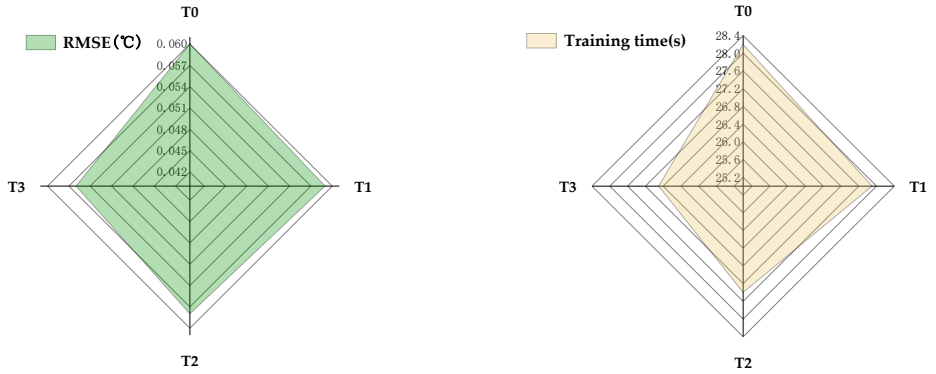
the training process, leading to a certain level of overfitting. Consequently, this overfitting hampered prediction accuracy and escalated the computational complexity of the neural network.



(**a**) T0 Dataset (**b**) T1 Dataset

(**c**) T2 Dataset (**d**) T3 Dataset

Figure 6. Comparison of R^2 values based on different decomposition modes: (**a**) model based on T0 Dataset; (**b**) model based on T1 Dataset; (**c**) model based on T2 Dataset; (**d**) model based on T3 Dataset.



(**a**) Comparison of RMSE values (**b**) Comparison of training time

Figure 7. Comparison of predictive performance of models based on different datasets: (**a**) comparison of RMSE values predicted by models under four datasets; (**b**) comparison of training time for model prediction under four datasets.

By comparing Figure 6b,c, it was evident that the exclusion of outdoor temperature and solar radiation, which exhibited a strong correlation with heat flux, resulted in an increase of 0.76% and 0.66% in R^2, respectively. Furthermore, Figure 7 demonstrated a decrease of

5.08% and 3.33% in RMSE, accompanied by a reduction in time by 1 and 0.8 s, respectively. These findings highlighted the possibility of the model acquiring redundant information when incorporating features that were highly correlated with other variables. Consequently, it necessitated the implementation of feature selection and processing techniques.

3.3. Evaluation of Model Performance on Fixed Dataset

To further validate the performance of the proposed data physical hybrid method for predicting mushroom room temperature (SSA-VMD-GRU-A-PHY) in practical applications, GRU, GRU–A, GRU-PHY, GRU-A-PHY, and SSA-VMD-GRU-A were selected for comparison in predicting mushroom room temperatures.

As shown in Figures 8 and 9, the GRU–A model exhibited a 6.89% decrease in *RMSE* and a 6.4% increase in R^2 compared to the GRU model. This improvement can be attributed to the effective enhancement of important feature weights through attention mechanisms, resulting in improved prediction accuracy. Compared with the GRU–A model, the SSA-VMD-GRU-A model predicts a 40.74% decrease in RMSE and a 24.34% increase in R^2. These results highlighted the significant enhancement in prediction accuracy achieved by employing different hyperparameters for different modal components in deep learning prediction. To illustrate the advantages of introducing physical rules into the loss function of neural networks, the following performance comparisons were conducted. Compared with the GRU model, the GRU-PHY model predicted a 26.72% decrease in RMSE and a 21.37% increase in R^2. Compared with the GRU–A model, the RMSE predicted by the GRU-A-PHY model decreased by 28.70% and the R^2 increased by 18.43%. Compared with the GRU–A model, the GRU-A-PHY model predicted a 28.70% decrease in RMSE and an 18.43% increase in R^2. Compared with the SSA-VMD-GRU-A model, the SSA-VMD-GRU-A-PHY model proposed in this paper predicted a 12.50% decrease in *RMSE* and a 2.5% increase in R^2. These findings indicate that introducing physical rules based on thermal equilibrium into the loss function of neural networks can significantly enhance the predictive performance of the model.

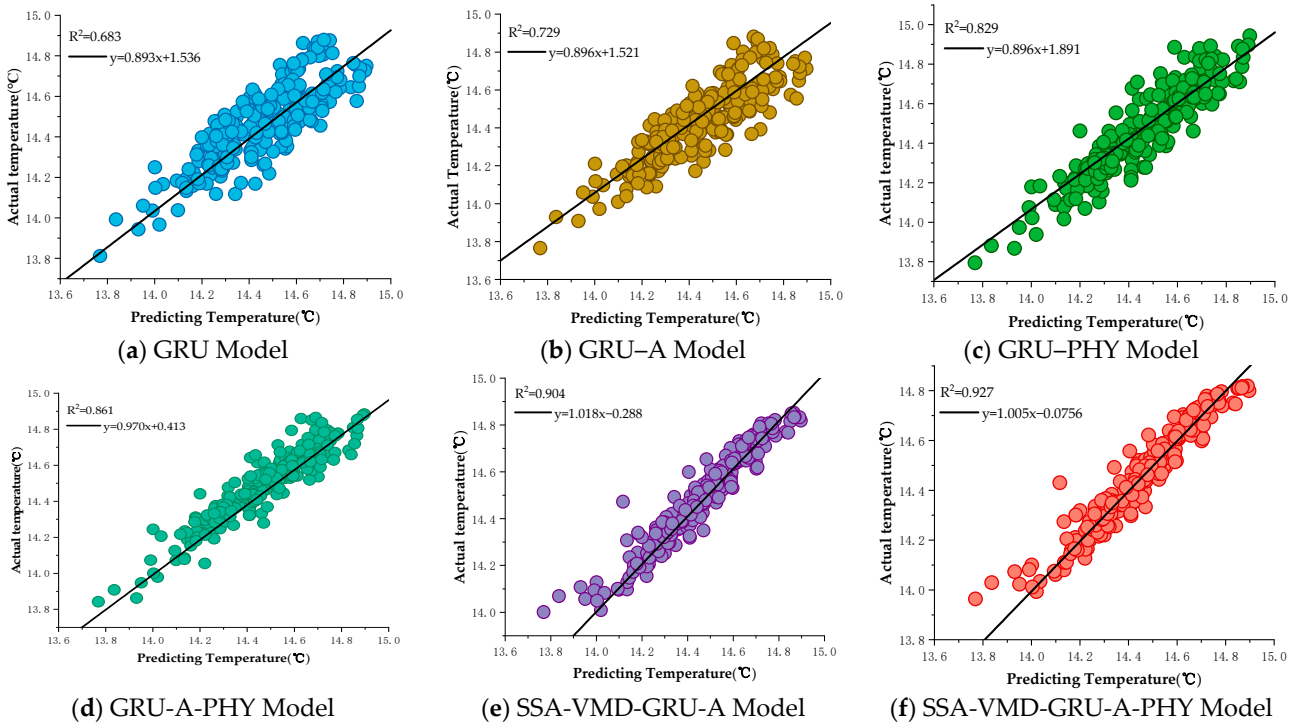


Figure 8. Comparison of R^2 values based on different prediction modes: (**a**) GRU Model; (**b**) GRU–A Model; (**c**) GRU–PHY Model; (**d**) GRU-A-PHY Model; (**e**) SSA-VMD-GRU-A Model; (**f**) SSA-VMD-GRU-A-PHY Model.

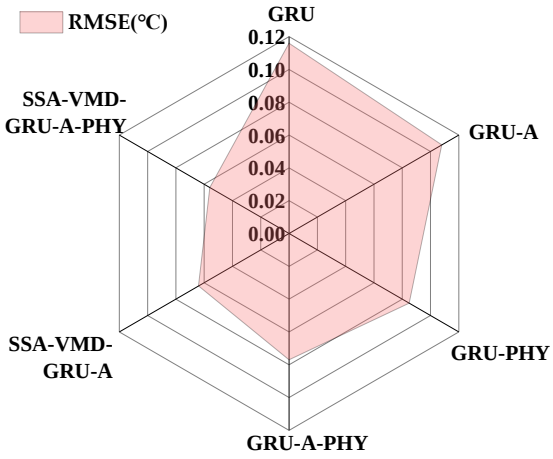


Figure 9. Comparison of *RMSE* values based on different prediction modes.

3.4. Evaluation of Model Performance on Different Datasets

According to ref. [38], neural network prediction models that incorporate physical rules require less training datasets compared to fully data-driven neural networks [35]. To assess the influence of the training dataset size on prediction accuracy of the proposed SSA-VMD-GRU-A-PHY model in this article, different durations were considered. Specifically, the first 20 days, 15 days, 10 days, 5 days, 3 days, and 1 day of 23 August were used as the model training sets. 23 August was designated as the model validation set, while 24 August was used as the model testing set.

As shown in Table 3, when the training dataset was collected for 20 days, 15 days, and 10 days, there was no significant change in the model's prediction accuracy. The RMSE values were recorded as 0.055 °C, 0.054 °C, 0.058 °C, while the corresponding R^2 values were 0.944, 0.944, and 0.937, respectively.

Table 3. Comparison of model performance under different training sets.

Prediction Model	Training Dataset (Days)	RMSE (°C)	R^2
SSA-VMD-GRU-A-PHY	1	0.162	0.509
	3	0.120	0.730
	5	0.078	0.884
	10	0.058	0.937
	15	0.054	0.944
	20	0.055	0.944
SSA-VMD-GRU-A	1	0.177	0.411
	3	0.154	0.555
	5	0.128	0.692
	10	0.094	0.832
	15	0.064	0.923
	20	0.060	0.533
GRU–A	1	0.185	0.359
	3	0.170	0.457
	5	0.162	0.509
	10	0.113	0.762
	15	0.105	0.793
	20	0.105	0.793

However, a noticeable change in prediction accuracy was observed when the training dataset was reduced to 5 days. In this case, the RMSE increased by 34.5% and the R^2 decreased by 5.7% compared to the 10-day training data. Furthermore, when the training data were further shortened to just one day, the RMSE increased by 179% and the R^2

decreased by 45.7%. These results can be attributed to the reduction of training data, resulting in overfitting of the model and poor performance in the test set. Therefore, it is recommended to ensure a training dataset of more than 5 days when utilizing the SSA-VMD-GRU-A-PHY model for temperature prediction.

Furthermore, a comparative analysis was conducted among the VMD-GRU-A-PHY, VMD-GRU-A, and GRU-A models using different datasets. As illustrated in the table, the proposed VMD-GRU-A-PHY model achieved an RMSE of 0.120 °C and an R^2 of 0.730 when trained for 3 days. On the other hand, the VMD-GRU-A and GRU–A models required 5 and 10 days of training data, respectively, to achieve a similar level of accuracy. These results can be attributed to the presence of physical constraints that impose limitations on the parameter updates within the neural network.

3.5. Discussion on Model Applications

The model constructed in this article was applicable to the factory production of Hypsizygus marmoreus. The temperature prediction model constructed in this article is suitable for the industrial production of Flammulina velutipes, which is similar to the cultivation of Hypsizygus marmoreus. The only difference is the target temperature setting, and the suitable growth temperature for Flammulina velutipes is 12–13 °C. The common characteristic among these edible mushroom types is the requirement for high concentrations of CO_2 during the mushroom emergence period to support the proper differentiation of the stipe. Ventilation has a relatively small impact on the temperature of the mushroom room.

For Agaricus bisporus, it is necessary to control the CO_2 concentration while adjusting the target temperature to prevent yield reduction. In addition, increasing ventilation appropriately can also prevent the occurrence of diseases. Therefore, when predicting the temperature of the mushroom room, the influence of fresh air on the temperature of the mushroom room cannot be ignored, and this part needs to be substituted into the construction of the heat balance equation.

Compared with pure data-driven methods, the model constructed in this article significantly shortened the training set. However, when applying it, it is still recommended that the shortest time period should not be less than one mushroom cultivation cycle, as mentioned in this article, which is 23 days for *Hypsizygus marmoreus*.

The model is suitable for the industrial production of edible mushrooms with air conditioning and other regulatory equipment. For the widely existing solar greenhouses and plastic greenhouses in China, these facilities have relatively weak temperature control conditions and are easily affected by external conditions. Therefore, the factors considered will significantly increase.

4. Conclusions

In contrast to conventional commercial buildings, factory mushroom houses possess distinct characteristics, including high internal planting density, substantial indoor thermal inertia, and a noticeable heat generation cycle associated with mushroom sticks. This study presents the GRU–A prediction model that integrates Boruta–SHAP feature selection, SSA-VMD feature decomposition, and physical model fusion. The following conclusions were summarized:

Compared to the GRU–A prediction model, the RMSE of the model decreased by 40.74%, and the R^2 increased by 24.34% when the data was decomposed by VMD before prediction. These results demonstrate that the model's performance significantly improved after decomposing the feature data.

In comparison to other decomposition methods such as EMD, EEMD, and EWT, the RMSE of the model decomposed by VMD decreased by 45.10% and the R^2 increased by 22.78%. The results show that the VMD decomposition method, optimized by SSA, effectively suppresses noise and enhances key features.

When compared to a simple data-driven model, the proposed SSA-VMD-GRU-A-PHY model predicted a decrease of 12.50% in RMSE, an increase of 2.5% in R^2, and a 40% reduction in the training dataset data. These results suggest that incorporating physical rules based on thermal equilibrium into the loss function of neural networks significantly enhances the model's predictive performance and reduces the data collection time required for training.

Future work will continue to focus on improving the model's prediction efficiency and interpretability. Non-deep learning methods such as the Autoregressive Integrated Moving Average model (ARIMA) and eXtreme Gradient Boosting (XGBoost) will be employed to predict different modal components after data decomposition. In addition, the optimization method for selecting weight values of multiple loss functions will also be a worthwhile research direction.

Author Contributions: Conceptualization, M.W. and C.Z.; methodology, M.W.; software, X.K.; validation, M.W., F.S. and J.W.; formal analysis, C.C.; resources, P.R.; data curation, X.K.; writing—original draft preparation, M.W.; writing—review and editing, X.Z. and W.Z.; visualization, M.W.; supervision, C.Z. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by the Beijing Academy of Agricultural and Forestry Sciences Science and Technology Innovation Capacity Construction Project (KJCX20230410), the China Agriculture Research System of MOF and MARA (CARS-20), and the Beijing edible fungi Innovation Team (BAIC03-2023).

Institutional Review Board Statement: The study did not require ethical approval.

Data Availability Statement: Data are contained within the article.

Acknowledgments: The authors would like to thank the anonymous reviewers for their constructive comments, which helped improve the quality of this paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wang, H.; Sun, G.; Pang, J.; Xie, Y.; Li, Y. Research progress on the influence of environmental stress on the growth and development of edible fungi. *Edible Med. Mushrooms* **2017**, *25*, 110–112.
2. Tianyu, H.; Tingting, Z.; Ming, G. Advances in low temperature stimulation for the fruiting body formation of edible fungi. *Microbiol. China* **2023**, *50*, 5518–5533. [CrossRef]
3. Pang, J.; Sun, G.; Yu, J.; Meng, H.; Wang, H.; Liu, Y.; Yang, Z. Cluster analysis based on the growth temperature of 35 edible fungi. *Edible Fungi China* **2015**, *34*, 34–37+41. [CrossRef]
4. Kathirgamanathan, A.; De Rosa, M.; Mangina, E.; Finn, D.P. Data-driven predictive control for unlocking building energy flexibility: A review. *Renew. Sustain. Energy Rev.* **2021**, *135*, 110120. [CrossRef]
5. Shao, J.; Huang, Z.; Chen, Y.; Li, D.; Xu, X. A practical application-oriented model predictive control algorithm for direct expansion (DX) air-conditioning (A/C) systems that balances thermal comfort and energy consumption. *Energy* **2023**, *269*, 126748. [CrossRef]
6. Godoy, J.L.; Schierloh, R.M. Predictive management of the hybrid generation dispatch and the dispatchable demand response in microgrids with heating, ventilation, and air-conditioning (HVAC) systems. *Sustain. Energy Grids Netw.* **2022**, *32*, 100857. [CrossRef]
7. Yang, S.; Wan, M.P.; Ng, B.F.; Dubey, S.; Henze, G.P.; Chen, W.; Baskaran, K. Experimental study of model predictive control for an air-conditioning system with dedicated outdoor air system. *Appl. Energy* **2020**, *257*, 113920. [CrossRef]
8. Wang, H.; Bo, S.; Zhu, C.; Hua, P.; Xie, Z.; Xu, C.; Wang, T.; Li, X.; Wang, H.; Lahdelma, R.; et al. A zoned group control of indoor temperature based on MPC for a space heating building. *Energy Convers. Manag.* **2023**, *290*, 117196. [CrossRef]
9. Drgoňa, J.; Arroyo, J.; Cupeiro Figueroa, I.; Blum, D.; Arendt, K.; Kim, D.; Ollé, E.P.; Oravec, J.; Wetter, M.; Vrabie, D.L.; et al. All you need to know about model predictive control for buildings. *Annu. Rev. Control* **2020**, *50*, 190–232. [CrossRef]
10. Tol, H.İ.; Madessa, H.B. Development of a white-box dynamic building thermal model integrated with a heating system. *J. Build. Eng.* **2023**, *68*, 106038. [CrossRef]
11. Hu, Y.; Qin, L.; Li, S.; Li, X.; Zhou, R.; Li, Y.; Sheng, W. Adaptive corrected parameters algorithm applied in cooling load prediction based on black-box model: A case study for subway station. *Energy Build.* **2023**, *297*, 113429. [CrossRef]
12. Talib, A.; Park, S.; Im, P.; Joe, J. Grey-box and ANN-based building models for multistep-ahead prediction of indoor temperature to implement model predictive control. *Eng. Appl. Artif. Intell.* **2023**, *126*, 107115. [CrossRef]
13. Lu, H.; Wu, J.; Ruan, Y.; Qian, F.; Meng, H.; Gao, Y.; Xu, T. A multi-source transfer learning model based on LSTM and domain adaptation for building energy prediction. *Int. J. Electr. Power Energy Syst.* **2023**, *149*, 109024. [CrossRef]

14. Zhang, J.; Poon, K.H.; Kwok, H.H.L.; Hou, F.; Cheng, J.C.P. Predictive control of HVAC by multiple output GRU-CFD integration approach to manage multiple IAQ for commercial heritage building preservation. *Build. Environ.* **2023**, *245*, 110802. [CrossRef]
15. Zuo, H.; Zheng, W.; Wang, M.; Zhang, X. Prediction of Heat and Cold Loads of Factory Mushroom Houses Based on EWT Decomposition. *Sustainability* **2023**, *15*, 15270. [CrossRef]
16. Peng, T.; Li, Y.; Song, Z.; Fu, Y.; Nazir, M.S.; Zhang, C. Hybrid intelligent deep learning model for solar radiation forecasting using optimal variational mode decomposition and evolutionary deep belief network—Online sequential extreme learning machine. *J. Build. Eng.* **2023**, *76*, 107227. [CrossRef]
17. Zhao, Z.; Yun, S.; Jia, L.; Guo, J.; Meng, Y.; He, N.; Li, X.; Shi, J.; Yang, L. Hybrid VMD-CNN-GRU-based model for short-term forecasting of wind power considering spatio-temporal features. *Eng. Appl. Artif. Intell.* **2023**, *121*, 105982. [CrossRef]
18. Chen, Z.; Liu, B.; Yan, X.; Yang, H. An Improved Signal Processing Approach Based on Analysis Mode Decomposition and Empirical Mode Decomposition. *Energies.* **2019**, *12*, 3077. [CrossRef]
19. Lu, Y.; Sheng, B.; Fu, G.; Luo, R.; Chen, G.; Huang, Y. Prophet-EEMD-LSTM based method for predicting energy consumption in the paint workshop. *Appl. Soft Comput.* **2023**, *143*, 110447. [CrossRef]
20. Shi, J.; Teh, J. Load forecasting for regional integrated energy system based on complementary ensemble empirical mode decomposition and multi-model fusion. *Appl. Energy* **2024**, *353*, 122146. [CrossRef]
21. Gao, Y.; Hang, Y.; Yang, M. A cooling load prediction method using improved CEEMDAN and Markov Chains correction. *J. Build. Eng.* **2021**, *42*, 103041. [CrossRef]
22. He, Y.; Tsang, K.F. Universities power energy management: A novel hybrid model based on iCEEMDAN and Bayesian optimized LSTM. *Energy Rep.* **2021**, *7*, 6473–6488. [CrossRef]
23. Lu, P.; Ye, L.; Pei, M.; Zhao, Y.; Dai, B.; Li, Z. Short-term wind power forecasting based on meteorological feature extraction and optimization strategy. *Renew. Energy* **2022**, *184*, 642–661. [CrossRef]
24. Han, L.; Jing, H.; Zhang, R.; Gao, Z. Wind power forecast based on improved Long Short Term Memory network. *Energy* **2019**, *189*, 116300. [CrossRef]
25. Sareen, K.; Panigrahi, B.K.; Shikhola, T.; Chawla, A. A robust De-Noising Autoencoder imputation and VMD algorithm based deep learning technique for short-term wind speed prediction ensuring cyber resilience. *Energy* **2023**, *283*, 129080. [CrossRef]
26. Li, Y.; Zhu, N.; Hou, Y. Comparison of empirical modal decomposition class techniques applied in noise cancellation for building heating consumption prediction based on time-frequency analysis. *Energy Build.* **2023**, *284*, 112853. [CrossRef]
27. Yu, Y.; Si, X.S.; Hu, C.H.; Zhang, J.X. A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Neural Comput.* **2019**, *31*, 1235–1270. [CrossRef]
28. Lundberg, S.; Lee, S.I. A Unified Approach to Interpreting Model Predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Volume 30. [CrossRef]
29. Sushanth, K.; Mishra, A.; Mukhopadhyay, P.; Singh, R. Real-time streamflow forecasting in a reservoir-regulated river basin using explainable machine learning and conceptual reservoir module. *Sci. Total Environ.* **2023**, *861*, 160680. [CrossRef]
30. El Bilali, A.; Abdeslam, T.; Ayoub, N.; Lamane, H.; Ezzaouini, M.A.; Elbeltagi, A. An interpretable machine learning approach based on DNN, SVR, Extra Tree, and XGBoost models for predicting daily pan evaporation. *J. Environ. Manag.* **2023**, *327*, 116890. [CrossRef]
31. Kellner, L.; Stender, M.; und Polach, F.V.B.; Ehlers, S. Predicting compressive strength and behavior of ice and analyzing feature importance with explainable machine learning models. *Ocean Eng.* **2022**, *255*, 111396. [CrossRef]
32. Lee, H.; Heo, Y. Simplified data-driven models for model predictive control of residential buildings. *Energy Build.* **2022**, *265*, 11267. [CrossRef]
33. Wang, M.; Zheng, W.; Zhao, C.; Chen, Y.; Chen, C.; Zhang, X. Energy-Saving Control Method for Factory Mushroom Room Air Conditioning Based on MPC. *Energies* **2023**, *16*, 7623. [CrossRef]
34. Bazghandi, R.; Hoseintabar Marzebali, M.; Abolghasemi, V.; Hedayati Kia, S. A Novel Mode Un-Mixing Approach in Variational Mode Decomposition for Fault Detection in Wound Rotor Induction Machines. *Energies* **2023**, *16*, 5551. [CrossRef]
35. Yang, J.; Zhou, C.; Li, X. Research on Fault Feature Extraction Method Based on Parameter Optimized Variational Mode Decomposition and Robust Independent Component Analysis. *Coatings* **2022**, *12*, 419. [CrossRef]
36. Xue, J.; Shen, B. A novel swarm intelligence optimization approach: Sparrow search algorithm. *Syst. Sci. Control Eng.* **2020**, *8*, 22–34. [CrossRef]
37. Yinhui, C. CFD Analysis of Environment of Edible Mushroom House and Optimization of Ventilation Structure. Master's Thesis, Jiangsu University, Zhenjiang, China, 2020. [CrossRef]
38. Gokhale, G.; Claessens, B.; Develder, C. Physics informed neural networks for control oriented thermal modeling of buildings. *Appl. Energy* **2022**, *314*, 118852. [CrossRef]