

Article

Decomposition-Based Multi-Step Forecasting Model for the Environmental Variables of Rabbit Houses

Ronghua Ji ¹, Shanyi Shi ¹, Zhongying Liu ² and Zhonghong Wu ^{2,*}¹ College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China² State Key Laboratory of Animal Nutrition, College of Animal Science and Technology, China Agricultural University, Beijing 100193, China

* Correspondence: wuzhh@cau.edu.cn

Simple Summary: Forecasting rabbit house environmental variables is critical to achieving intensive rabbit breeding and rabbit house environmental regulation. As a result, this paper proposes a decomposition-based multi-step forecasting model for rabbit houses using a time series decomposition algorithm and a deep learning combinatorial model. The experimental results demonstrated that the proposed method could provide accurate decisions for rabbit house environmental regulation.

Abstract: To improve prediction accuracy and provide sufficient time to control decision-making, a decomposition-based multi-step forecasting model for rabbit house environmental variables is proposed. Traditional forecasting methods for rabbit house environmental parameters perform poorly because the coupling relationship between sequences is ignored. Using the STL algorithm, the proposed model first decomposes the non-stationary time series into trend, seasonal, and residual components and then predicts separately based on the characteristics of each component. LSTM and Informer are used to predict the trend and residual components, respectively. The aforementioned two predicted values are added together with the seasonal component to obtain the final predicted value. The most important environmental variables in a rabbit house are temperature, humidity, and carbon dioxide concentration. The experimental results show that the encoder and decoder input sequence lengths in the Informer model have a significant impact on the model's performance. The rabbit house environment's multivariate correlation time series can be effectively predicted in a multi-input and single-output mode. The temperature and humidity prediction improved significantly, but the carbon dioxide concentration did not. Because of the effective extraction of the coupling relationship among the correlated time series, the proposed model can perfectly perform multivariate multi-step prediction of non-stationary time series.

Keywords: correlated time series; multivariate multi-step prediction; deep learning algorithm; environment variables forecasting



Citation: Ji, R.; Shi, S.; Liu, Z.; Wu, Z. Decomposition-Based Multi-Step Forecasting Model for the Environmental Variables of Rabbit Houses. *Animals* **2023**, *13*, 546. <https://doi.org/10.3390/ani13030546>

Academic Editor: Cesare Castellini

Received: 25 November 2022

Revised: 29 January 2023

Accepted: 2 February 2023

Published: 3 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The future development trend of the rabbit breeding industry is toward intensification. The precise regulation of rabbit house environmental variables is the premise of intensive breeding, and the ability to accurately predict rabbit house environmental variables is the foundation for achieving environmental regulation.

Temperature, relative humidity, and carbon dioxide concentration are the most important environmental variables in a rabbit house. The prediction of rabbit house environmental variables falls under the category of time series prediction. Time series forecasting algorithms are classified into two types based on implementation theory: traditional mathematical algorithms [1,2] and machine learning algorithms [3,4]. To effectively predict the environmental parameters of livestock houses, researchers have proposed Elanco ammonia concentration prediction equation [5] for cattle houses and ammonia concentration mass

balance model [6] for swine houses using traditional mathematical forecasting algorithms. However, the prediction methods based on mathematical principles have poor generalization performance and low stability. Machine learning algorithms are classified as either machine learning or deep learning algorithms [7,8]. Researchers have proposed using models such as Support Vector Regression (SVR) [9,10] for machine learning algorithms to predict livestock house environmental variables. These models can accurately predict non-stationary single-parameter time series data. The variables in the rabbit house environment are coupled, and the machine learning model's relatively simple structure makes fully exploring the coupling relationship between the variables difficult. With the advent of deep learning, researchers have used Long Short Term Memory (LSTM) [11], Back Propagation Network (BPN) [12], and other models to predict the gas concentration of the chicken house, as well as the temperature and humidity of the pig house. Deep learning algorithms can mine the complex characteristics of variables more effectively.

According to the forecasting step, time series forecasting algorithms are classified as single-step forecasting or multi-step forecasting. Single-step time series forecasting algorithms, such as Elman neural network [13] and SVR [14], have produced relatively accurate results; however, in rabbit house scenarios, environmental regulation is usually delayed by a small amount of time [15], so single-step forecasting algorithms do not meet the time requirement. Nonetheless, multi-step forecasting algorithms such as LSTM-based deformation models [16,17], SVR-based models [18], and Echo State Network-based models [19] can be used to predict rabbit house environmental variables in multiple steps.

Based on the number of model output variables, time series forecasting algorithms are further classified as univariate forecasting algorithms and multivariate forecasting algorithms [20,21]. The series of rabbit house environmental variables exhibits both periodicity and nonlinearity [22] and mutual coupling between variables [23]. The prediction accuracy will suffer if the prediction is solely based on the law of the predicted variable and ignores the impact of other variables. Therefore, in rabbit house environmental variable time series prediction, multivariate and multi-step forecasting corresponds to the actual demand. Researchers have proposed a few prediction models based on Graph Neural Network [24], Recurrent Neural Network [25], and Echo State Network [26,27] that improved prediction accuracy by effectively extrapolating and fully utilizing the coupling relationship between multiple variables, which provides a reference for the multivariate and multi-step prediction of rabbit house environmental variables, as deep learning algorithms have rapidly developed.

The multivariate and multi-step time series forecasting model's input–output mapping is quite complex and has a low prediction accuracy. To effectively carry out long sequence prediction by fully mining the dependence relationship between time series, researchers have proposed various models based on mathematical transformations [28], Temporal Convolutional Network [29], and Ensemble Empirical Mode Decomposition [30]. Informer [31], a long series forecasting model based on the attention mechanism proposed in 2021, effectively extracts the coupling relationship between the correlated time series through the complex nonlinear mapping relationship established by the encoder and decoder, reduces computational complexity by using the sparse self-attention mechanism, and predicts the correlated time series accurately and efficiently. Using a sparse self-attention mechanism reduces computational complexity, allowing for more accurate and efficient prediction of correlated long-time series.

To ensure the algorithm's prediction accuracy, it is critical to allow sufficient time for control of decision-making in practical situations. This paper proposes a decomposition-based multi-step forecasting model for non-stationary correlated time series to address this issue.

To improve the prediction accuracy of rabbit house environmental variables and provide sufficient time for control of decision-making, this paper proposes to establish a decomposition-based multi-step forecasting model for rabbit house environmental vari-

ables based on previous research results and to provide an effective reference for rabbit house environmental supervision decision-making.

2. Materials and Methods

2.1. Time Series Forecasting Algorithms

2.1.1. Long Short Term Memory

LSTM is a time series prediction model that can effectively avoid gradient disappearance and gradient explosion. The trend component obtained by STL decomposition is the long-term change trend of this variable, with stable fluctuation and a small standard deviation, which can be accurately predicted using a simple structure time series prediction model. To that end, LSTM is used in this study to predict the trend component of decomposed time series data. Figure 1 depicts the LSTM cell structure.

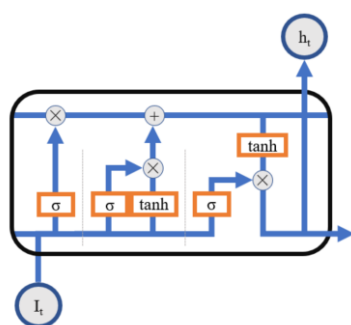


Figure 1. Structure graph of short and long time memory network unit.

As shown in Figure 1, for a given input sequence $I = \{I_1, I_2, \dots, I_n\}$, LSTM time series data prediction process is mainly divided into three steps. In the first forgetting step, the LSTM unit receives the input data h_{t-1} of the hidden layer unit at the previous time step and the input data I_t at the current time step for weighted sum calculation to control which input of the previous time step needs to be left over. The data obtained in the previous step are fed into the sigmoid function in the second step to obtain the f_t , which is the Information that must be forgotten. Furthermore, the weighted sum of h_{t-1} and x_t is calculated. The output step is the third step. The second stage data are fed into the sigmoid and tanh functions to obtain the input gate value and state information. The cell state is updated based on the most recent time data. The value of the output gate is determined to obtain the output data O_t , and the current hidden layer unit value H_t is later taken to the next unit.

2.1.2. Informer

The Informer uses the Encoder-Decoder architecture, whose overall structure is shown in Figure 2. For the time series sample pair $\{I_j, I_{j+1}, \dots, I_{j+s-1}\}$ $\{I_{j+s-l}, \dots, I_{j+s-1+p}\}$, the encoder receives a long sequence $\{I_j, I_{j+1}, \dots, I_{j+s-1}\}$ as input, through the sparse self-attention module, combined with the self-attention distillation mechanism. The feature vector V_e is obtained. $\{I_{j+s-l}, \dots, I_{j+s-1+p}\}$ through mask processing, and the feature vector V_d is obtained through the input to the sparse self-attention module of the decoder. The query vector is calculated by V_e , then the key vector and the value vectors are calculated by V_d . Finally, the full attention mechanism and the fully connected layer are used to obtain the output, where s , l , and p denote the three custom variables `seq_len`, `label_len`, and `pred_len` for data processing, respectively.

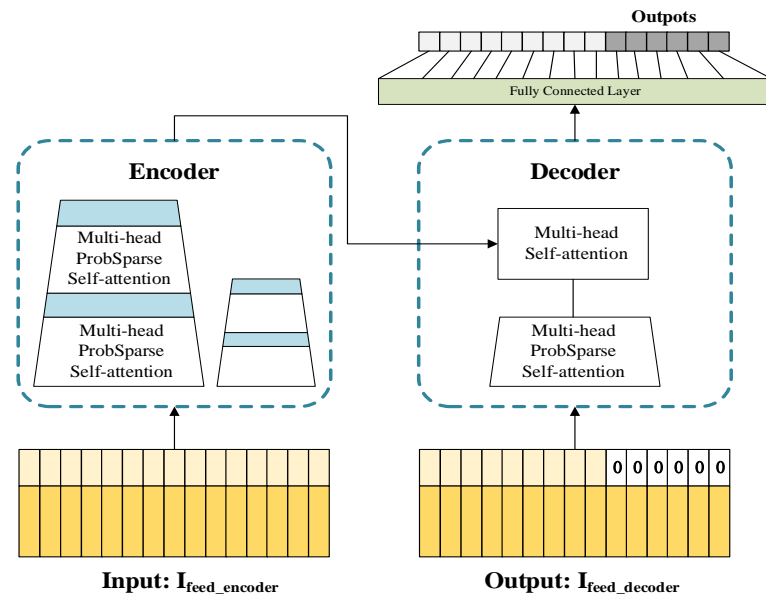


Figure 2. Informer structure diagram.

2.2. Model Structure

A non-stationary series is the rabbit house variates time series. Non-stationary series are typically divided into three components: trend, seasonal, and residual, and each data component has distinct characteristics. It was discovered that decomposing the non-stationary time series first and then predicting each part individually improve prediction operation accuracy significantly. Figure 3 depicts the basic structure of the decomposition-based multi-step forecasting model for rabbit house environmental variables.

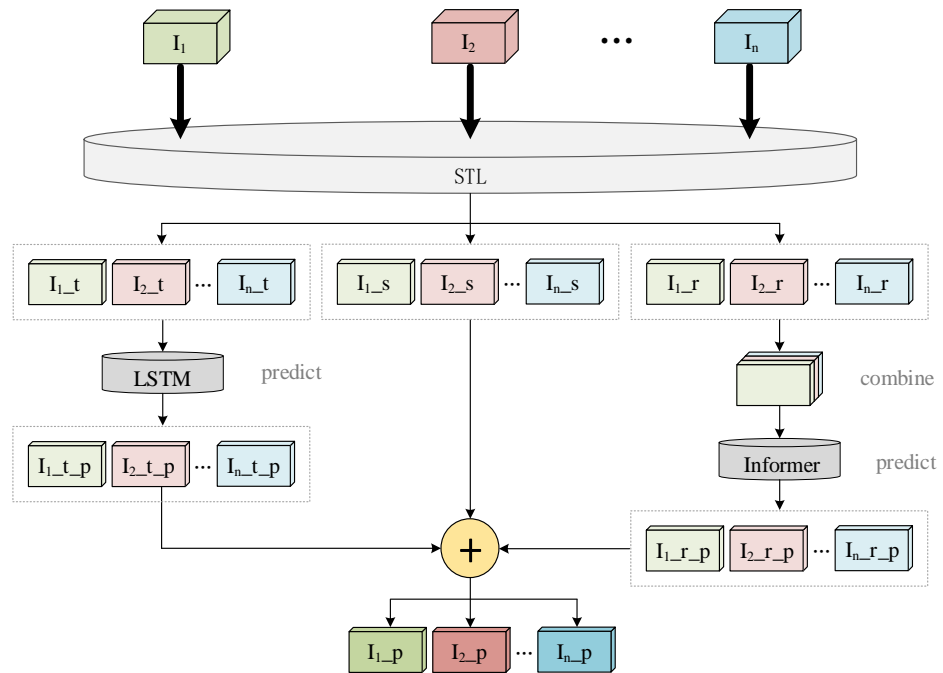


Figure 3. The fundamental structure of the decomposition-based multi-step forecasting model for the environmental variables of rabbit house.

The decomposition-based multi-step forecasting model for rabbit house environmental variables includes three primary stages, as shown in Figure 3.

The first stage is to decompose the time series into three parts. The input of model is the correlated time series $\{I_1, I_2, \dots, I_n\}$. The time series is decomposed into a trend sequence $\{I_{1_t}, I_{2_t}, \dots, I_{n_t}\}$, the seasonal sequence $\{I_{1_s}, I_{2_s}, \dots, I_{n_s}\}$, and the residual sequence $\{I_{1_r}, I_{2_r}, \dots, I_{n_r}\}$ by Seasonal and Trend decomposition using Loess (STL), which is a robust and versatile time series decomposition method and is used to decompose time series in this paper.

The second stage entails predicting each of the three components individually. Because the trend component of a time series is relatively stable, the LSTM model is used for prediction. Then the outputs of the LSTM model are the predicted trend component of each series $\{I_{1_{t_p}}, I_{2_{t_p}}, \dots, I_{n_{t_p}}\}$. The residual component is predicted by the Informer model. The outputs of the Informer model are the predicted residual component $\{I_{1_{r_p}}, I_{2_{r_p}}, \dots, I_{n_{r_p}}\}$. The seasonal component does not need to be predicted because it is a periodic constant.

The final stage is to obtain the final prediction result, which can be obtained by adding the predicted values of the trend component, residual component, and seasonal component of the correlated time series. The following equation shows how the final forecast value is calculated:

$$I_{i_p} = I_{i_{t_p}} + I_{i_s} + I_{i_{r_p}} \quad (1)$$

where I_{i_p} is the final forecast value of the time series. $I_{i_{t_p}}$ and $I_{i_{r_p}}$ are the predicted values of trend component and residual component for the i_{th} series. I_{i_s} is the seasonal component of the i_{th} series.

2.3. Data Acquisition

The environmental variables of the rabbit house were collected at Qingdao Kangda Rabbit Co. Ltd. in Shandong Province, China. The company has a large rabbit breeding base. The rabbit house is a 46-m-long enclosed structure with a span of 11.7 m and a height of 4.9 m.

Figure 4 depicts the current situation inside the rabbit house. The environmental variables inside the rabbit house were collected using twelve automated temperature and humidity recorders (Apresy, 179-TH) and three automated carbon dioxide recorders (Tian Jian Hua Yi, EZY-1). Figure 5 depicts the sensors' planar distribution.



Figure 4. The actual situation inside the rabbit house.

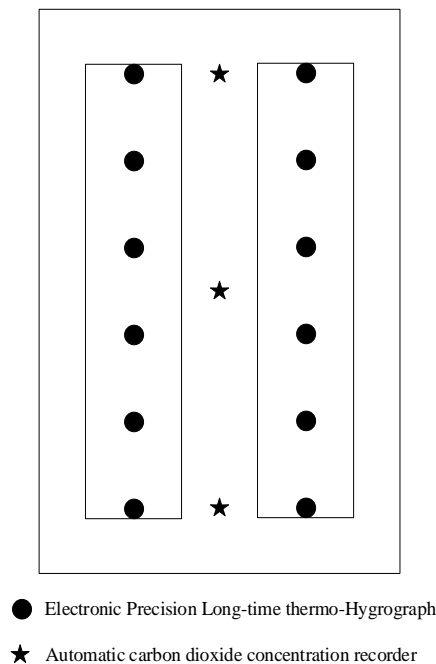


Figure 5. The planar distribution of the sensors inside the rabbit house.

The environmental variables inside the rabbit house were collected at a frequency of 10 min from 00:00 on 30 September 2020 to 23:50 on 28 November 2020. Finally, a total of 8640 rabbit house environmental variables were obtained.

2.4. Dataset Preprocessing

2.4.1. Missing Data

Due to the interference of equipment, external factors, and other factors during the acquisition data process of rabbit house environmental variables, the collected data contain some missing values, which is primarily caused by mechanical failures in data collection or storage. Based on the distribution, missing values can be classified as completely random missing or completely non-random missing. The missing data in rabbit house environmental variables are not random. Because the rabbit house's environmental variables are obtained through continuous timing acquisition every ten minutes, there will be no sudden environmental change. In this study, the mean value of the data before and after the missing value was used to fill in the missing value.

2.4.2. Normalization

The dimensions and dimensional units of the rabbit house environment variables vary. When the model is built directly, it tends to focus on variables with larger dimensions, resulting in low model prediction accuracy and slowing training speed. To eliminate the dimensional effect between the parameters, the data must be standardized so that the preprocessed data is limited to a specific range, typically between zero and one. This method's calculation formula is as follows:

$$y_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (2)$$

where x_i and y_i are the values before and after normalization, respectively. x_{min} and x_{max} are the minimum and maximum values in the same variables, respectively.

2.5. Dataset Analysis

The environmental variables measured inside the rabbit house were processed and analyzed. The results are shown in Table 1.

Table 1. Statistical information of the rabbit house environment variables.

Variable	Unit	Range	Mean	Standard Deviation
Temperature	°C	[9.3, 25.57]	19.09	2.79
Relative Humidity	%rh	[32.42, 84.3]	66.03	8.86
CO ₂ Concentration	ppm	[138.75, 1858.57]	802.23	250.82

As shown in Figure 6, three line graphs visualize the fluctuation of the temperature, the relative humidity, and the carbon dioxide concentration variables of the rabbit house within a week.

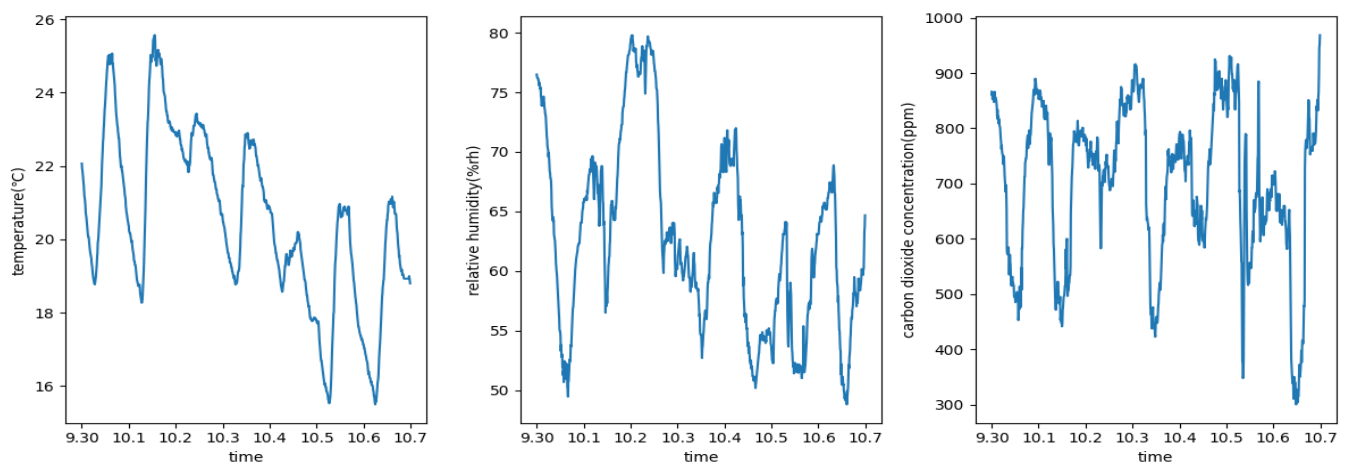
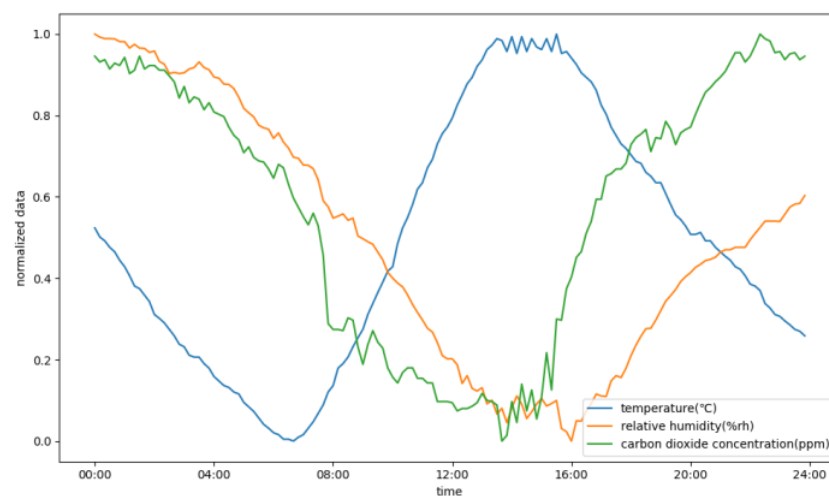
**Figure 6.** Rabbit house environmental variables within a week.

Table 1 and Figure 6 show that the environmental variables of the rabbit house have a strong periodicity.

To summarize, the three most important environmental variables within the rabbit house fluctuate on a regular basis. The temperature fluctuates in the opposite direction of the relative humidity and carbon dioxide concentration once every 24 h, while the latter two fluctuate in very similar ways.

Figure 7 depicts the normalized temperature, relative humidity, and carbon dioxide concentration in the rabbit house over one day.

**Figure 7.** The rabbit house environmental variables within a day.

The temperature in the rabbit house reaches its lowest point around 7:00 a.m. every day, gradually rises to its peak around 15:00 p.m. every afternoon, and then gradually decreases toward nightfall, as shown in Figure 7. Furthermore, the data show that when the temperature is at its highest, the relative humidity and carbon dioxide concentration are at their lowest, implying that the trend in change of the relative humidity and carbon dioxide concentration is the inverse of that of temperature. Pearson's correlation coefficient was calculated using the data from day one from the rabbit house's temperature, relative humidity, and carbon dioxide concentration environment variables. Table 2 summarizes the results of those calculations.

Table 2. Person's Correlation Coefficient for rabbit house environmental variables.

	Temperature	Relative Humidity	CO ₂ Concentration
Temperature	1	−0.8276	−0.5149
Relative Humidity	−0.8276	1	0.6505
CO ₂ Concentration	−0.5149	0.6505	1

Consequently, three conclusions can be drawn: first, the daily temperature in the rabbit house is strongly correlated with the relative humidity; second, the temperature is correlated with the carbon dioxide concentration; and third, the relative humidity is correlated with carbon dioxide concentration.

The rabbit house time series of environmental variables show periodicity and a strong coupling relationship.

2.6. Experimental Settings

The dataset contains a total of 8640 samples representing rabbit house environmental variables. Temperature, relative humidity, and carbon dioxide concentration are all included in each sample. The first 7680 samples are chosen as the training set, and the last 960 consecutive samples are chosen as the test set.

Model parameters have a significant impact on model performance. The following are the proposed model parameters. The LSTM model is made up of three layers: an input layer, a hidden layer, and an output layer. The LSTM model is fed with the temperature, relative humidity, and carbon dioxide concentration trends. The LSTM model's output is the predicted value of the trend component of temperature, relative humidity, or carbon dioxide concentration. The number of hidden layers is set to four. The Informer model is made up of an encoder and a decoder. The residual component of temperature, relative humidity, and carbon dioxide concentration are the encoder's inputs, and the predicted value of the inputs is the decoder's output. The parameters for each module of the Informer model are shown in Table 3.

Table 3. The Parameters of the Informer model.

Main Modules	Sub-Model	Parameters	Type/Value
Encoder	Sparse Self-Attention Module	Number	1
		Number of heads	8
		Number of units	5
	Self-attentive distillation module	Number of stack levels	3,2,1
Decoder	Attention Mechanisms Module	Number of stories	1
		Type	Full attention
	Fully connected layer	Number of stories	1
		Number of hidden units	512

The training parameters of Informer models are set as shown in Table 4.

Table 4. The Training parameters for the Informer model.

Parameters	Value/Type
Optimizer	Adam
Epoch_size	6
Batch_size	96
Loss Function	MSE
Initial Learning Rate	0.001
Learning rate adjustment method	Each epoch goes down by half
Iteration times	6
Drop_out	0.05

3. Results and Discussion

3.1. Effect of Model Parameter Settings on Prediction Accuracy

The input sequence length of the encoder and decoder modules has a significant impact on model performance in the Informer model. There is the encoder input sequence length, denoted by seq_len , and the decoder input sequence length, denoted by $label_len$ and $pred_len$. The predicted sequence's length is $label_len$, and the model's output sequence is the last $pred_len$ data in the predicted sequence.

Based on preliminary test results and the experience of relevant breeding experts, we discovered that it takes 20–70 min to adjust the environmental variables in the rabbit house from the initial value to the desired value. The model's output sequence length should be greater than the maximum regulation time to allow enough time for the rabbit house's environmental regulation.

The carbon dioxide concentration in the rabbit house varies greatly over time and is difficult to predict. Therefore, in the experiment of selecting the sequence length of the informer model, the carbon dioxide concentration was chosen as the model output. The Dataset_KD_minute class has been designed to load, segment, and normalize the dataset recorded to ensure that the environment variables of the rabbit house are fit for the Informer model. Dataset_KD_minute class divides the data into sequence data and time stamp data. Time stamped data are the corresponding time stamp of the sequence data, which is transformed into a vector, and extended to the same dimension as the sequence data vector through the neural network, and superimposed with the sequence data vector as the final input data.

The sequence length of the Informer model is set as follows: seq_len is set as 24 (4 h), 36 (6 h), 48 (8 h), 72 (12 h), and 84 (14 h) separately. $label_len$ is set $n-24$ and $n-12$ separately, where n is the corresponding seq_len value; $pred_len$ is set as 7 (70 min), 9 (90 min), and 12 (120 min) separately. The experiments were carried out separately using the above model parameters, which were then combined. Table 5 displays the experimental results.

It can be found that MSE decreases with an increase in $pred_len$ when seq_len and $label_len$ are the same. Considering that $pred_len$ is stable, the MSE decreases while seq_len increases. The reason for this is that as the input sequence lengthens, the correlation between the input and output sequences weakens. The best prediction is generated when seq_len is set to 24, $label_len$ is set to 12, and $pred_len$ is set to 7.

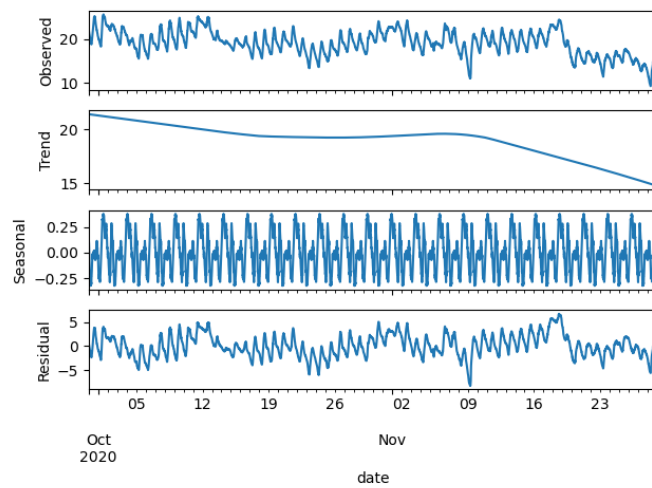
Table 5. Effect of model parameters on prediction accuracy of the model.

seq_len	label_len	pred_len	MSE (ppm)
24	0	7	0.05255134
		9	0.05477738
		12	0.082935
	12	7	0.04983531
		9	0.0659451
		12	0.08847401
36	12	7	0.0786688
		9	0.07736252
		12	0.11616067
	24	7	0.05427639
		9	0.06950006
		12	0.09658156
48	24	7	0.08289333
		9	0.0991382
		12	0.14262368
	36	7	0.07137094
		9	0.08005989
		12	0.12634541
72	48	7	0.09766254
		9	0.13091703
		12	0.15992408
	60	7	0.0786742
		9	0.13212924
		12	0.15188387
84	60	7	0.09868335
		9	0.1437092
		12	0.16232075
	72	7	0.08037249
		9	0.13032459
		12	0.15770504

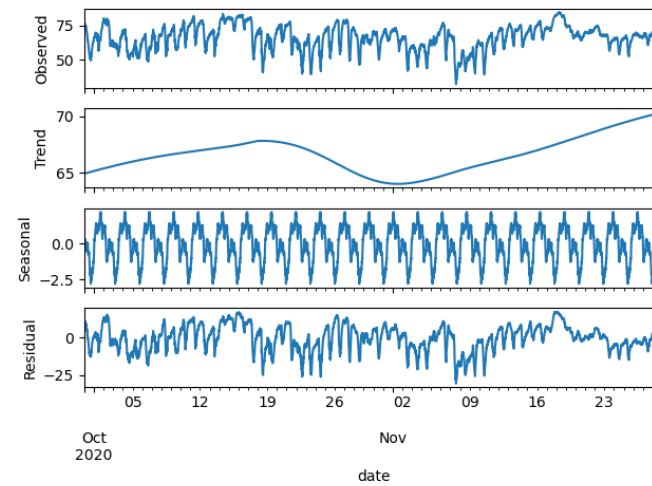
3.2. Effect of Sequence Decomposition on Model Prediction

In this section, the experiment was carried out to investigate if the model can accurately predict the environmental parameters of the rabbit house. The goal of this experiment is to evaluate if sequence decomposition improves prediction accuracy, or in other words, to investigate if the proposed model can accurately predict the environmental variables of the rabbit house.

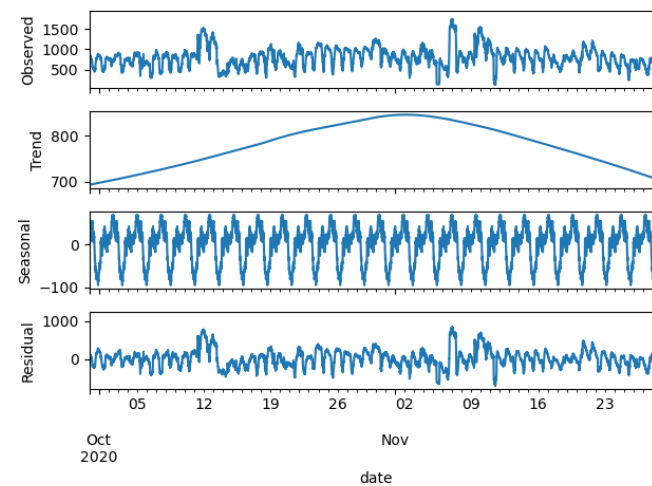
The STL algorithm is primarily used to decompose the environmental variables, and the decomposition results are shown in Figure 8.



(a)



(b)



(c)

Figure 8. The decomposition results for the rabbit house environmental parameters. (a) Temperature decomposition diagram, (b) Humidity decomposition diagram, (c) Concentrations of carbon dioxide decomposition diagram.

The change trend of the residual component environmental variables of the rabbit house is found to be similar to that of the original variables. Table 6 shows the results of calculating the correlation coefficient among the residual component environmental variables of the rabbit house.

Table 6. Correlation coefficient of the residual component environmental variables of the rabbit house.

Variable	t_r	h_r	cd_r
t_r	1	−0.8277	−0.6424
h_r	−0.8277	1	0.7198
cd_r	−0.6424	0.7198	1

t_r, h_r, and c_r in Table 6 are the residual component environmental variables of the rabbit house, which were decomposed by the STL algorithm. After decomposition, the correlation between the residual components of the rabbit house's environmental variables improves in comparison to Table 2. The coefficient of correlation between residual temperature and relative humidity has shifted slightly. Besides, the carbon dioxide concentration residual component with the temperature residual component or with the relative humidity residual component, the correlation coefficient increased by 24.76% or by 10.65%, respectively.

To test whether the proposed model can predict the correlation time series, the Informer, SVR, XGBoost, and the proposed model were used to predict the environmental variables of the rabbit house. SVR is a supervised machine learning model that seeks the best-fitting equation. It is capable of predicting time series data. SVR has high prediction accuracy as well as good robustness and generalization ability. XGBoost is a boosting-based supervised learning model. A CART decision tree is generated in each iteration to fit the difference between the sum of the predicted results of all previous trees and the true value. It has the benefits of not overtraining and quick training speed. Table 7 displays the results of the comparison experiments.

Table 7. The prediction results of two models for the prediction of correlated time series.

Predicted Variable	Model	MAE	MSE	RMSE
Temperature	Informer	0.12279760	0.02539201	0.15934873
	SVR	0.16925226	0.02864633	0.10879960
	XGBoost	0.13623399	0.02993292	0.17301134
	Proposed model	0.01167244	0.00030404	0.01743695
Humidity	Informer	0.12256409	0.02699366	0.16429749
	SVR	0.11360447	0.04006461	0.20016146
	XGBoost	0.13044998	0.02799125	0.16730587
	Proposed model	0.00261323	0.00015509	0.01245363
CO ₂ Concentration	Informer	0.15820476	0.04983531	0.22323823
	SVR	0.14900408	0.05078254	0.22534983
	XGBoost	0.16991315	0.05427113	0.23296165
	Proposed model	0.15369803	0.04048497	0.20120878

As shown above, when the proposed model predicts temperature and humidity, the model evaluation indicators MSE, MAE, and RMSE increase by more than 90% when compared to other models. Namely, when predicting temperature and humidity, the prediction effect of the rabbit house time series can be significantly improved. When used to predict carbon dioxide concentration, the proposed model's prediction accuracy is less

improved. This is because the fluctuation of carbon dioxide concentration in the rabbit house is much greater than the fluctuation of temperature and humidity, indicating that the proposed model is better suited for time series prediction with small fluctuations. Communication with breeders about the cause of the large variation in carbon dioxide concentration reveals that carbon dioxide concentration is affected not only by temperature and relative humidity but also by external factors such as rearing methods, resulting in high expectations for the model’s generalization ability. However, in terms of current breeding needs, the model’s accuracy can meet the decision-making needs of rabbit house environmental regulation.

3.3. Effect of Input Variable Types on Model Performance

We performed several experiments to evaluate how the number of input variable types affected model performance when predicting the correlation time series. The experiments were carried out as follows. The proposed model takes one type, two types, and three types as input and outputs a single variable. For example, the model inputs are set to temperature; temperature and relative humidity; temperature and CO₂ concentration; temperature, relative humidity, and CO₂ concentration; temperature, relative humidity, and CO₂ concentration. In contrast, the proposed model is used to predict temperature, and the model prediction accuracy, including MAE, MSE, and RMSE, is recorded.

The results show that the best prediction performance is obtained when the model’s input variable type is three. The percentage improvement in model prediction accuracy was calculated when the model input was three types versus one type and two types. Figure 9 depicts the outcomes. The three areas in Figure 9 represent the model output as temperature, relative humidity, and CO₂ concentration, respectively. The abscissa represents the input type, and the ordinate represents the percentage improvement in model prediction accuracy.

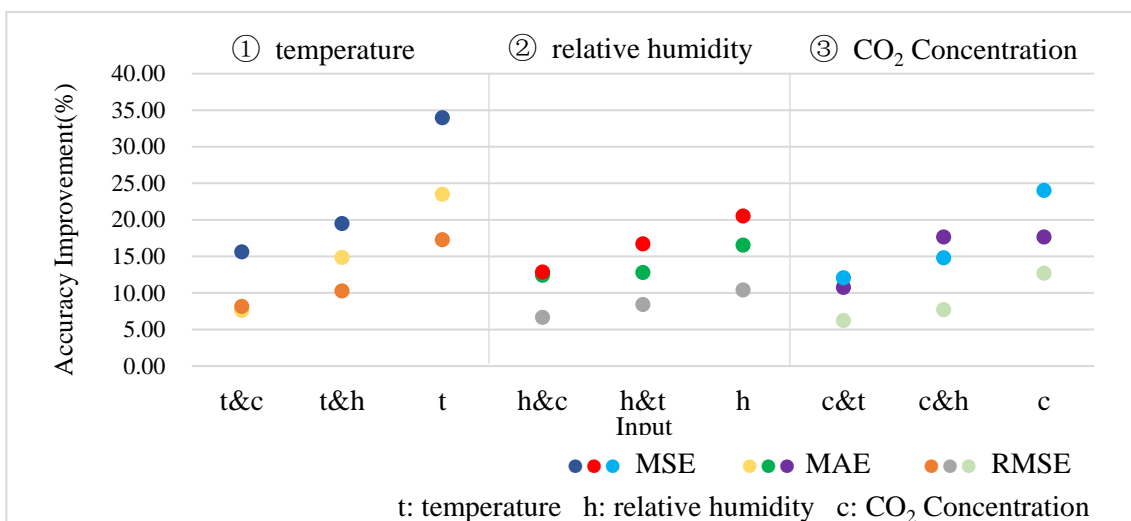


Figure 9. The prediction accuracy improvement with different input.

The results show that when temperature, relative humidity, and carbon dioxide concentration are used to predict the environmental variables of the rabbit house, the proposed model has the best prediction effect. This is because the rabbit house’s environmental variables have a strong coupling relationship. Simultaneously, the coupling relationship between the remaining terms of the three variables is highlighted after decomposing the correlation time series, which is more conducive to improving the prediction effect of the proposed model.

4. Conclusions

The model performance suffers when we perform multi-step prediction for correlated non-stationary time series of rabbit house environment because we ignore the time series' coupling relationship. We propose a decomposition-based multi-step forecasting model for rabbit house environmental variables to address this issue.

The proposed model first decomposes the non-steady time series into trend, seasonal, and remainder components using the STL algorithm and then predicts separately based on the characteristics of each component.

Because its main environmental variables (temperature, humidity, and carbon dioxide concentration) correlate with non-stationary time series, the proposed model predicted the rabbit house's environmental variables to validate its performance. The time series of the rabbit house's main environmental variables has a strong time correlation and a significant coupling relationship. A number of comparative experiments were also carried out. Accordingly, we reached the following conclusions:

- (1) By decomposing the correlation time series and then making specific predictions, the proposed model can realize multi-step prediction for correlated non-stationary time series. The experimental results demonstrated that the proposed model could perform multi-step prediction for the rabbit house's environmental variables. It can effectively improve temperature and humidity predictions in particular, but not carbon dioxide concentration predictions.
- (2) The length of the Informer model's input and output in the proposed model has a significant impact on the model's performance. When *seq_len* and *label_len* are the same, MSE decreases with increasing *pred_len* for the rabbit house's environmental variables prediction. MSE decreases as *seq_len* increases when *pred_len* remains constant. The effect of predicting environmental variables of the rabbit house is the best when *seq_len*, *label_len*, and *pred_len* are set to 24, 12, and 7, respectively.
- (3) Temperature, relative humidity, and carbon dioxide concentration were the environmental variables of the rabbit house. The absolute value of Pearson's correlation coefficient between any of the above two was greater than 0.5, indicating that the time series of environmental variables of the rabbit house not only had their own time correlation but also had a significant coupling relationship with other variables. The model's complexity and performance are heavily influenced by the types of input and output. When temperature, relative humidity, and carbon dioxide concentration are used as model inputs, and a single parameter is used as model output, the best prediction result for rabbit house environmental variables is obtained. The coupling relationship between the remaining terms of the correlation time series is highlighted further after the time series is decomposed, which improves the model's prediction performance. When temperature, relative humidity, and carbon dioxide concentration are used as inputs, the proposed model has the best prediction effect for the rabbit house's environmental variables.

Author Contributions: Conceptualization, Z.W. and R.J.; methodology, S.S. and R.J.; software, S.S.; validation, S.S. and Z.L.; formal analysis, S.S. and R.J.; investigation, S.S. and Z.L.; resources, S.S. and Z.L.; data curation, S.S. and R.J.; writing—original draft preparation, S.S. and R.J.; writing—review and editing, S.S. and R.J.; visualization, S.S. and Z.L.; supervision, Z.W., Z.L. and R.J.; project administration, Z.W. and R.J.; funding acquisition, Z.W. and R.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the earmarked fund for CARS (CARS-43-D-2) the National Key Research and Development Program of China (No. 2022YFD1301104) and Beijing Innovation Consortium of Agriculture Research System (BAIC01–2022).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: If anyone needs the data used in this study, please contact: jessic1212@cau.edu.cn.

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

References

1. Xu, W.; Peng, H.; Zeng, X.; Zhou, F.; Tian, X.; Peng, X. A Hybrid Modeling Method Based on Linear AR and Nonlinear DBN-AR Model for Time Series Forecasting. *Neur. Process. Lett.* **2022**, *54*, 1–20. [[CrossRef](#)]
2. Tian, D.; Wei, X.; Wang, Y.; Zhao, A.; Mu, W.; Feng, J. Prediction of temperature in edible fungi greenhouse based on MA-ARIMA-GASVR. *Trans. Chin. Soc. Agric. Eng.* **2020**, *36*, 190–197. [[CrossRef](#)]
3. Jiang, H.; Li, D.; Jing, W.; Xu, J.; Huang, J.; Yang, J.; Chen, S. Early Season Mapping of Sugarcane by Applying Machine Learning Algorithms to Sentinel-1A/2 Time Series Data: A Case Study in Zhanjiang City, China. *Remote Sens.* **2019**, *11*, 861. [[CrossRef](#)]
4. Chen, L.; He, Q.; Liu, K.; Li, J.; Jing, C. Downscaling of GRACE-Derived Groundwater Storage Based on the Random Forest Model. *Remote Sens.* **2019**, *11*, 2979. [[CrossRef](#)]
5. Pas, M.; Pas, N.; Gruber, S.; Kube, J.; Teeter, J.S. Modeling and prediction accuracy of ammonia gas emissions from feedlot cattle. *Appl. Anim. Behav. Sci.* **2019**, *35*, 347–356. [[CrossRef](#)]
6. Arulmozhi, E.; Basak, J.K.; Sihalath, T.; Park, J.; Kim, H.T.; Moon, B.E. Machine learning-based microclimate model for indoor air temperature and relative humidity prediction in a swine building. *Animals* **2021**, *11*, 222. [[CrossRef](#)]
7. Ma, Q.; Li, S.; Shen, L.; Wang, J.; Wei, J.; Yu, Z.; Cottrell, G.W. End-to-End Incomplete Time-Series Modeling from Linear Memory of Latent Variables. *IEEE Trans. Cybernet.* **2020**, *50*, 4908–4920. [[CrossRef](#)]
8. Li, M.; Xu, D.; Geng, J.; Hong, W. A ship motion forecasting approach based on empirical mode decomposition method hybrid deep learning network and quantum butterfly optimization algorithm. *Nonlinear. Dynam.* **2022**, *107*, 2447–2467. [[CrossRef](#)]
9. Chen, Y.; Cheng, Y.; Cheng, Q.; Yu, H.; Li, D. Short-term prediction model for ammonia nitrogen in aquaculture pond water based on optimized lssvm. *Int. Agric. Eng. J.* **2017**, *26*, 416–427.
10. Liu, S.; Huang, J.; Xu, L. Combined model for prediction of air temperature in poultry house for lion-head goose breeding based on PCA-SVR-ARMA. *Trans. Chin. Soc. Agric. Eng.* **2020**, *36*, 9. [[CrossRef](#)]
11. Xie, Q.; Zheng, P.; Bao, J. Thermal Environment Prediction and Validation Based on Deep Learning Algorithm in Closed Pig House. *Trans. Chin. Soc. Agric. Mach.* **2020**, *51*, 353–361. [[CrossRef](#)]
12. Xie, B.; Ma, Y.W.; Wan, J.Q.; Wang, Y.; Guan, Z.Y. An accuracy model for on-line prediction of effluent ammonia nitrogen in anammox treatment system based on pca-bp algorithm. In Proceedings of the 2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA), Beijing, China, 8–11 September 2017; IEEE: Piscataway, NJ, USA, 2017.
13. Zhang, Y.; Wang, X.; Tang, H. An improved Elman neural network with piecewise weighted gradient for time series prediction. *Neurocomputing* **2019**, *359*, 199–208. [[CrossRef](#)]
14. Chandra, R.; Ong, Y.; Goh, C. Co-evolutionary multi-task learning for dynamic time series prediction. *Appl. Soft. Comput.* **2018**, *70*, 576–589. [[CrossRef](#)]
15. Saleem, M.H.; Potgieter, J.; Arif, K.M. Automation in Agriculture by Machine and Deep Learning Techniques: A Review of Recent Developments. *Precis. Agric.* **2021**, *22*, 2053–2091. [[CrossRef](#)]
16. Khan, M.; Wang, H.; Riaz, A.; Elfatyany, A.; Karim, S. Bidirectional LSTM-RNN-based hybrid deep learning frameworks for univariate time series classification. *J. Supercomput.* **2021**, *77*, 7021–7045. [[CrossRef](#)]
17. Chu, X.; Jin, H.; Li, Y.; Feng, J.; Mu, W. CDA-LSTM: An evolutionary convolution-based dual-attention LSTM for univariate time series prediction. *Neur. Comput. Appl.* **2021**, *33*, 16113–16137. [[CrossRef](#)]
18. Majidpour, M.; Nazaripouya, H.; Chu, P.; Pota, H.; Gadh, R. Fast Univariate Time Series Prediction of Solar Power for Real-Time Control of Energy Storage System. *Forecasting* **2018**, *1*, 107–120. [[CrossRef](#)]
19. Yao, X.; Wang, Z.; Zhang, H. A novel photovoltaic power forecasting model based on echo state network. *Neurocomputing* **2019**, *325*, 182–189. [[CrossRef](#)]
20. Xiao, F.; Huisman, Q.E. Prediction of biopersistence of hydrocarbons using a single parameter. *Chemosphere* **2018**, *213*, 76–83. [[CrossRef](#)]
21. Yu, B.; Yin, H.; Zhu, Z. *Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting*; Cornell University Library: Ithaca, NY, USA, 2018; Reprinted.
22. Song, L.; Wang, Y.; Zhao, B.; Liu, Y.; Mei, L.; Luo, J.; Zuo, Z.; Yi, J.; Guo, X. Research on Prediction of Ammonia Concentration in QPSO-RBF Cattle House Based on KPCA Nuclear Principal Component Analysis. *Procedia. Comput. Sci.* **2021**, *188*, 103–113. [[CrossRef](#)]
23. Alfaseeh, L.; Tu, R.; Farooq, B.; Hatzopoulou, M. Greenhouse gas emission prediction on road network using deep sequence learning. *Transport. Res. D-Tr. E.* **2020**, *88*, 102593. [[CrossRef](#)]
24. Huang, F.; Yi, P.; Wang, J.; Li, M.; Peng, J.; Xiong, X. A dynamical spatial-temporal graph neural network for traffic demand prediction. *Inform. Sci.* **2022**, *594*, 286–304. [[CrossRef](#)]
25. Rahman, A.; Srikumar, V.; Smith, A.D. Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks. *Appl. Energy* **2018**, *212*, 372–385. [[CrossRef](#)]

26. Zhang, G.; Zhang, C.; Zhang, W. Evolutionary echo state network for long-term time series prediction: On the edge of chaos. *Appl. Intell.* **2020**, *50*, 893–904. [[CrossRef](#)]
27. Zhang, H.; Hu, B.; Wang, X.; Xu, J.; Wang, L.; Sun, Q.; Wang, Z. Self-organizing deep belief modular echo state network for time series prediction. *Knowl. Based Syst.* **2021**, *222*, 107007. [[CrossRef](#)]
28. Diez-Sierra, J.; Del Jesus, M. Long-term rainfall prediction using atmospheric synoptic patterns in semi-arid climates with statistical and machine learning methods. *J. Hydrol.* **2020**, *586*, 124789. [[CrossRef](#)]
29. Fu, Y.; Hu, Z.; Zhao, Y.; Huang, M. A Long-Term Water Quality Prediction Method Based on the Temporal Convolutional Network in Smart Mariculture. *Water* **2021**, *13*, 2907. [[CrossRef](#)]
30. Wang, S.; Qiu, J.; Li, F. Hybrid Decomposition-Reconfiguration Models for Long-Term Solar Radiation Prediction Only Using Historical Radiation Records. *Energies* **2018**, *11*, 1376. [[CrossRef](#)]
31. Zhou, H.; Zhang, S.; Peng, J.; Zhang, S.; Zhang, W. Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting. *arXiv* **2020**, arXiv:2012.07436. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.