

Article

NO_x Concentration Prediction in Cement Denitrification Process Based on EEMD-MImRMR-BASBP

Xuanzhi Liu¹, Yanxin Chen^{1,*} , Ning He² and Yanfei Yao¹ ¹ College of Materials Science and Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China² College of Electrical and Mechanical Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China

* Correspondence: chenyanxin@xauat.edu.cn; Tel.: +86-1309-697-3109

Abstract: NO_x concentration is an important indicator of the response to ammonia dosage and nitrogen emissions, and its accurate prediction allows for efficient and rational optimal control of ammonia dosage. Due to the large external noise, time lag and non-linearity of the cement denitrification process, it is difficult to derive accurate mathematical prediction models. Therefore, a new machine learning model, namely EEMD-MImRMR-BASBP, is developed. Firstly, Ensemble Empirical Mode Decomposition (EEMD) and median-averaged filtering is used to process the data and remove the noise. In order to handle the large time lags, non-linearity and non-smoothness among the variables, mutual information (MI) based on the entropy principle is proposed to calculate the lag time of the non-linear system; furthermore, according to the feature variable selection method of Max-Relevance and Min-Redundancy (mRMR), the factors with strong influence are selected as the input variables of the prediction model in combination with the results of the mechanism analysis. Then, the EEMD-MImRMR-BASBP model to predict NO_x concentration is constructed, in which the initialization parameters of the Back Propagation Neural Network (BP) are searched by Beetle Antennae Search (BAS) to effectively overcome the parameter selection problem of traditional BP prediction models. Finally, the model was applied for the NO_x concentration prediction of a real cement plant in Jiang xi and Fu ping and compared with the classical BP-based prediction model, BASBP model, the root means square error (RMSE) and mean absolute error (MAE) of the EEMD-MImRMR-BASBP model for the two production lines are only 0.2927, 0.3513 and 0.1795, and 0.2383, which have better prediction performance compared with the current model.



Citation: Liu, X.; Chen, Y.; He, N.; Yao, Y. NO_x Concentration Prediction in Cement Denitrification Process Based on EEMD-MImRMR-BASBP. *Processes* **2023**, *11*, 317. <https://doi.org/10.3390/pr11020317>

Academic Editor: Jacopo Donnini

Received: 20 December 2022

Revised: 5 January 2023

Accepted: 12 January 2023

Published: 18 January 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/>).

Keywords: NO_x emissions; mutual information; Ensemble Empirical Mode Decomposition; Max-Relevance and Min-Redundancy; Beetle Antennae Search

1. Introduction

NO_x is a key indicator for environmental monitoring during the calcination of cement kilns. This indicator determines the production efficiency, and higher NO_x values can lead to substandard emissions, which leads to shutdown and rectification in serious cases. Current denitrification technologies commonly used in cement production are Selective Non-Catalytic Reduction (SNCR) [1] and Selective Catalytic Reduction (SCR) [2], both of which are based on the reaction of ammonia and NO_x into the system to produce N₂ and H₂O. However, the ammonia flow rate in the existing denitrification process is mostly controlled manually or by PID based on the NO_x emission value after denitrification, which is influenced by the large inertia, delay and disturbance of the cement production. There are many problems, such as large and untimely ammonia flow rate adjustment and ammonia escape. Therefore, the establishment of an accurate NO_x concentration prediction model can efficiently and reasonably provide the basis for the optimal control of ammonia dosage, which is of great significance to improve the control quality of current denitrification systems and reduce denitrification costs in the cement industry.

During cement production, sensor problems and complex gas–solid reactions can lead to noisy process data and time lags, and high temperatures and variable external environments can trigger non-linearities in the working conditions, all of which affect the data-driven predictive capability, so the model-building process needs to be a special case. For the characteristics of large noise, the empirical mode decomposition (EMD) method [3] is a commonly used method for non-linear signal pre-processing compared to wavelet transform [4] and conventional filtering [5]; however, when the signal polar points are not uniformly distributed, modal aliasing can occur. In order to address these problems, the Ensemble Empirical Mode Decomposition (EEMD) method [6] is proposed, which effectively suppresses the generation of modal aliasing by adding Noise Assisted Analysis (NADA) to change the polar point characteristics of the signal. Moreover, median-averaged filtering [7] can effectively segment large, noisy data. In addition, the simplest way to deal with time lags is to calculate the correlation coefficients between input and output variables, but in real industrial processes, variables are mostly dynamic and non-linear, and the correlation coefficients cannot be handled. Mutual information (MI) based on the entropy principle [7] is an effective method to deal with non-linear time lags. If all process variables are used as input variables, it leads to redundancy of data and increases the model computation. Data analysis methods such as maximum information coefficient [8] and gray correlation analysis [9] focus on the correlation between input and output variables and ignore the redundancy between input variables, which interferes with the prediction results. The Max-Relevance and Min-Redundancy (mRMR) algorithm [10] can be a good solution to the above variable selection problem.

In addition to the input sample data, the modeling method is also an important factor affecting prediction accuracy. In recent years, with the development of machine learning, many modeling methods have been proposed and applied to the prediction of NO_x concentration. Tang et al. [11] and Tan et al. [12] implemented the prediction of NO_x based on deep confidence network and long short-term memory (LSTM) methods, respectively. Neural networks are particularly suitable in the field of multi-factor and uncertain information processing. However, there are problems of difficult structure determination and the tendency to fall into local optimality by introducing the Beetle Antennae Search (BAS) algorithm [13], which simulates the foraging principle of the aspen to search for the optimal solution in the global context [14].

A hybrid strategy NO_x concentration prediction model is proposed for the characteristics of large noise, large time lag and non-linearity in the cement production process. On the one hand, EEMD and median average filtering are used to deal with large noise. Moreover, MI and mRMR with entropy principle are used for time lag analysis and variable selection to solve the large time lag problem. On the other hand, BAS improves the BP neural network prediction ability and solves the non-linear working condition problem. Finally, a prediction model of NO_x concentration based on EEMD-MImRMR-BASBP was established using the data of the 6500t/d production line in Fu ping and verified in the 2500t/d production line of Jiang xi.

2. Analysis of NO_x Generation Mechanism

In the cement, suspension preheat decomposition system, preheater, decomposer and rotary kiln are the three main pieces of equipment used to complete the calcination of clinker. According to its principle, NO_x during the calcination of cement clinker can be divided into fuel-based and thermal-based (Figure 1). The fuel type refers to the NO_x generated by the oxidation of nitrogen-containing functional groups in pulverized coal with O_2 in the combustion air during the combustion process, and the fuel type NO_x can be generated in both decomposer and rotary kiln. Thermal NO_x refers to NO_x generated by the oxidation of N_2 in the air with O_2 at high temperatures, and the generation location is mainly in the high-temperature section of the rotary kiln.

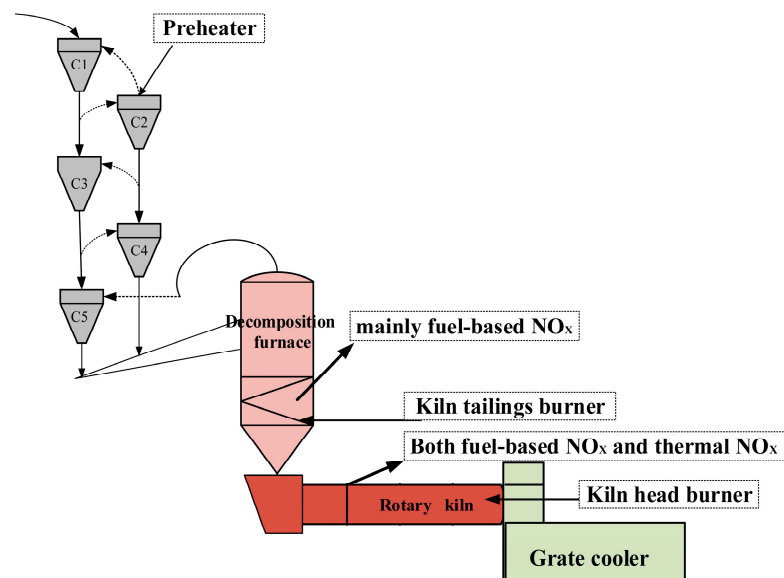


Figure 1. Diagram of NO_x generation location in cement firing process.

NO_x concentration can be detected using a flue gas analyzer. The current Fu ping Eco-cement 2# production line is equipped with two sets of flue gas analyzers, namely C1A and C1B outlets (medium-temperature gas analyzer) and kiln tail flue chamber side wall (high-temperature gas analyzer). In order to better investigate the NO_x generation trajectory, a portable gas analyzer (Testo 350) was used to measure the NO_x concentration in each section of the flue chamber and decomposer along the kiln tail preheat decomposition system from the bottom up, and based on the results, the NO_x footprint in the system was analyzed as follows:

(1) The amount of NO_x generation in the rotary kiln can be characterized according to the gas composition at the flue chamber. On the day of testing, NO_x gas was not detected in the smoke chamber, while a large amount of CO gas was present, with 0.73% O₂ content. It is thus determined that the current ventilation in the rotary kiln is low, and the reduction zone is formed in the back section of the rotary kiln, which makes the fuel type and thermal type NO_x generated in the high-temperature section be reduced.

(2) The decomposer body, decomposer gooseneck tube and decomposer outlet were tested separately, and the decomposition rate of temperature and material at the decomposer body, gooseneck tube and outlet were measured at the same time, and the NO_x, O₂ and CO gas composition data at each location were summarized, as shown in Table 1.

Table 1. Summary of the main parts of the decomposition furnace parameters.

	Gas Composition						Temperature	Measurement Uncertainty	Decomposition Rate	
	NO _x	Measurement Uncertainty	CO	Measurement Uncertainty	O ₂	Measurement Uncertainty				
	ppm	%	ppm	%	%	%	°C	°C	%	
Main Body	600.5	±5	206	±5%	0	±8	878	±1	73.2	Column A
Gooseneck	623.9	±5	136	±5%	0.01	±8	884	±1	86.8	
Export	609.7	±5	9	±2 ppm	0.09	±8	870	±1	91.1	
Main Body	614.8	±5	180	±5%	0.02	±8	878	±1	73.0	Column B
Gooseneck	645.5	±5	111	±5%	0.01	±8	886	±1	85.2	
Export	637.5	±5	0	±2 ppm	0.07	±8	871	±1	91.4	

According to the measurement, for both sides of A and B, the NO_x content along the decomposer from bottom to the top showed a trend of increasing first and then decreasing, i.e., the NO_x concentration reached the highest at the gooseneck tube. The NO_x content at the gooseneck is higher than that at the main body of the decomposer because of

the continuous generation of fuel-based NO_x in the process from the main body of the decomposer to the gooseneck. The decomposition of calcium carbonate absorbed heat in the process of gooseneck tube value decomposer exit caused the gas temperature at the exit of decomposer to be lower than that at the gooseneck tube, and under the weak reducing atmosphere, CO reacted with NO_x in the furnace, resulting in a small reduction in NO_x content.

According to the measured data, the NO_x generation footprint within the 2# production line is shown in Figure 2.

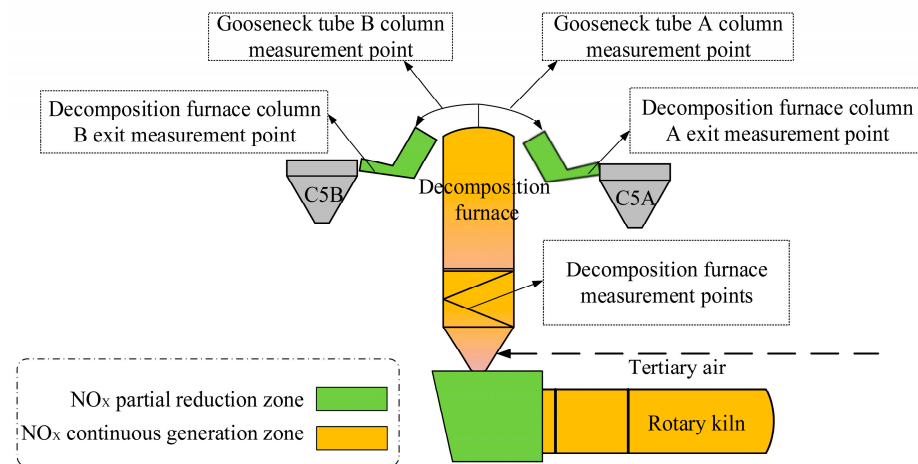


Figure 2. NO_x footprint of 2# production line.

According to the process theory and field measurements, the characterization factors of combustion temperature, air volume, calcination conditions and denitrification efficiency in the cement calcination process were selected. The above-selected characterization factors are related in Table 2.

Table 2. Relevant input variables.

Characterization Factors	Definition	Variable Name	Unit	Measurement Uncertainty
Ignite Grill Warm Degree Table Levy	T_D	Decomposition furnace temperature	$^{\circ}\text{C}$	$\pm 1^{\circ}\text{C}$
	T_{C5}	C5 temperature	$^{\circ}\text{C}$	$\pm 1^{\circ}\text{C}$
	T_S	Kiln tail flue chamber temperature	$^{\circ}\text{C}$	$\pm 1^{\circ}\text{C}$
	T_K	Kiln head hood temperature	$^{\circ}\text{C}$	$\pm 1^{\circ}\text{C}$
	G_T	Head of coal	t/h	± 0.1 t/h
	G_W	Tail coal volume	t/h	± 0.1 t/h
Denitrification efficiency characterization	L_N	Ammonia flow	l/min	0.5 L/min
Calcine Grill Work Situation Table Levy	I_{KP}	Kiln current	A	$\pm 0.5\%$
	C_{NO}	C1 Export NO_x	ppm	$\pm 5\%$
	C_{CO}	C1 export carbon monoxide	ppm	$\pm 5\% (+40 \sim +500 \text{ ppm}) / \pm 2 \text{ ppm } (0 \sim +39.9 \text{ ppm})$
	C_{O2}	C1 export oxygen	%	$\pm 8\%$
Empty Gas quantity Table Levy	P_K	Kiln head hood pressure	Pa	$\pm 1.5\% (+3 \sim +40 \text{ hPa})$
	R_H	High-temperature fan speed	rpm	0.0074 rpm
	I_{KC}	Kiln head exhaust fan current	A	$\pm 0.5\%$

3. Introduction to the Construction Idea and Method of Prediction Model

Based on the above analysis of the NO_x process and its characteristics, and the fact that the cement calcination process with strong non-linearity, large noise, large lag and coupling characteristics. Therefore, the EEMD-MImRMR-BASBP model is proposed in this paper to predict the concentration of NO_x , and the model process is shown in Figure 3.

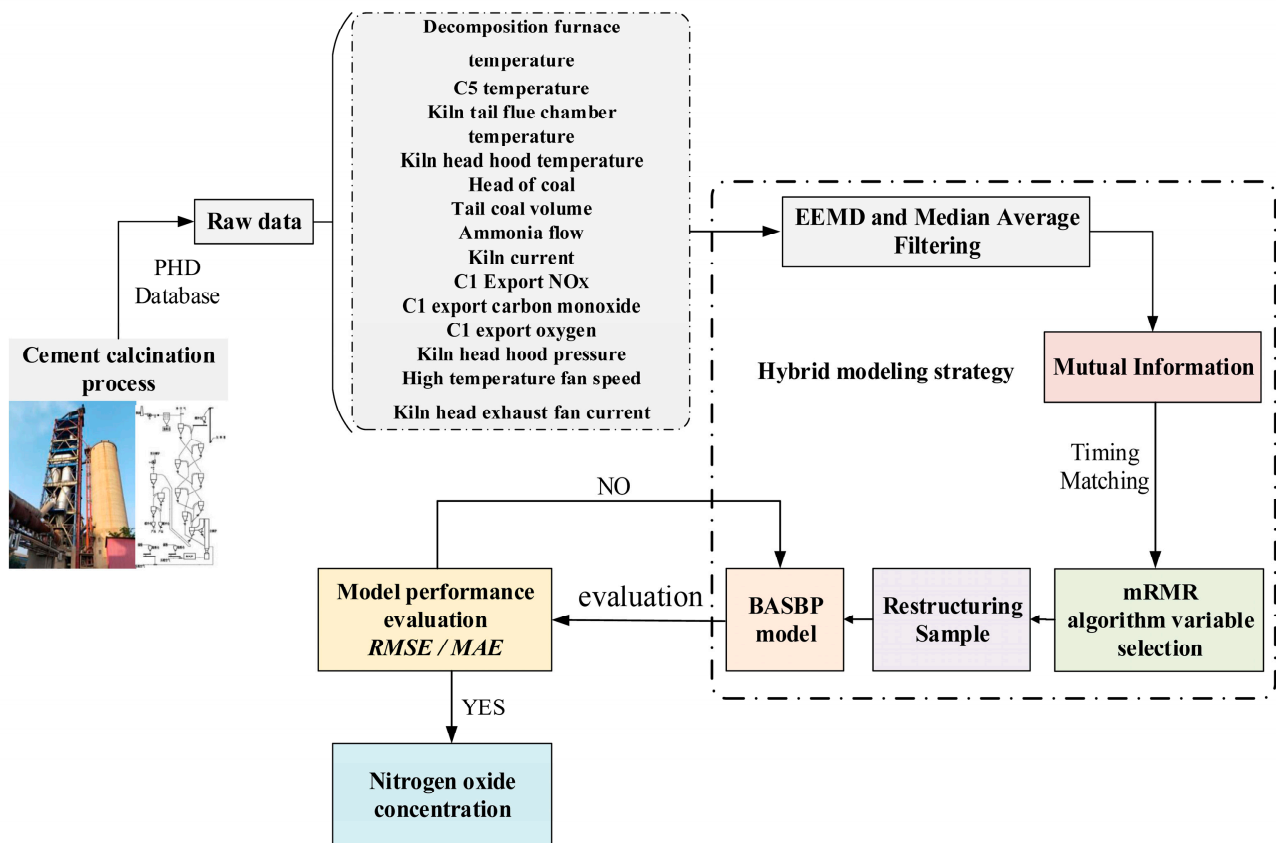


Figure 3. Modeling process of NO_x .

3.1. EEMD and Median-Averaged Filtering Data Processing Modules

Compared to EMD, EEMD can better avoid problems such as modal aliasing and enable better filtering of non-linear or non-stationary signals in industrial sites. The core idea of the EEMD method is to add i times the white noise of the standard normal distribution $n_i(t)$ to the original signal $x(t)$ and to add the white noise to the signal as a whole. The signal with the white noise added as a whole $x_i(t)$ is EMD decomposed into a series of IMFs and a residual term, which is repeated according to the amount of data, and the IMFs obtained from j times are integrated and averaged as the final result $c_j(t)$, as in Equations (1) and (2) [6].

$$x(t) + n_i(t) = x_i(t) = \sum_{j=1}^J \text{imf}_{i,j}(t) + r_{i,j}(t) \quad (1)$$

$$c_j(t) = \frac{1}{M} \sum_{i=1}^M \text{imf}_{i,j}(t) \quad (2)$$

The different $c_j(t)$ signals represent different frequencies; from $1 \sim j$, the signal frequency gradually decreases. Therefore, the deeper the layer, the lower the correlation with the noise is, while the correlation with the real signal increases, so the available correlation relationship can be used as a criterion to reject high-frequency signals. The normalized

correlation coefficient between the eigenmode component $c_j(t)$ and the original signal $x_i(t)$ is shown in Equation (3) [3].

$$r_j(x_i(t), c_j(t)) = \frac{\sum_{t=1}^N (x_i(t) - \bar{x}_i)(c_j(t) - \bar{c}_j)}{\sqrt{\sum_{t=1}^N (x_i(t) - \bar{x}_i)^2 \bullet \sum_{t=1}^N (c_j(t) - \bar{c}_j)^2}} \quad (3)$$

where N indicates the length of the original signal sampled, and a larger correlation r_j indicates a stronger correlation between the different frequency signals and the original signal. When r_j reaches a local minimum, the first j $c_j(t)$ are selected as the high-frequency signal layer, and these $c_j(t)$ are filtered using median averaging for signal processing.

The median average filter, also known as the anti-pulse interference average filter, is utilized to remove a maximum value and a minimum value from the collected data at N and then average the remaining data.

3.2. MI Timing Matching Module for the Entropy Principle

Entropy as a measure of uncertainty in information theory uses the information entropy $H(x)$ to estimate the degree of uncertainty in the input variables and the joint entropy $H(x, y)$ to represent the degree of uncertainty in the input variables x and output variables y of correlation $H(x)$ and $H(x, y)$, mathematically described as shown in (4) [15] and (5) [16].

$$H(x) = -\sum_i p_x(x_i) \log(p_x(x_i)) \quad (4)$$

$$H(x, y) = -\sum_i p_{xy}(x_i, y_i) \log(p_{xy}(x_i, y_i)) \quad (5)$$

where p_x denotes the probability distribution of x_i and $p_{x,y}$ is the joint probability distribution of x_i, y_i . Normalizing the information entropy and joint entropy yields mutual information $MI(x, y)$ [17].

$$MI(x, y) = \frac{H(x) + H(y) - H(x, y)}{\sqrt{H(x)H(y)}} \quad (6)$$

3.3. mRMR Variable Selection Module

mRMR is a filtered feature selection algorithm that finds the set of features in the original set of features that are most relevant (Max-Relevance) to the final output but least relevant (Min-Redundancy) to each other by using the non-linear relationship-mutual information as a correlation measure for the features.

The mRMR algorithm finds the feature set G by calculating the mutual information $I(x, y)$ between the random variables X and Y and calculates the relevant input variable eigenvalues $c_{x,j}(t)$ and the target eigenvalues $c_{y,j}(t)$, thereby determining the maximum correlation $\max D[G, c_{y,j}(t)]$ and the minimum redundancy $\min R(G)$ and finally integrating the maximum correlation minimum redundancy $\max \phi(D, R)$ by addition. The relevant mathematical descriptions are shown in Equations (7) [18], (8) [18,19], (9) [18,19] and (10) [18,20].

$$I[c_{x,j}(t), c_{y,j}(t)] = \iint \log \frac{p[c_{x,j}(t), c_{y,j}(t)]}{p[c_{x,j}(t)]p[c_{y,j}(t)]} dc_{x,j}(t)dc_{y,j}(t) \quad (7)$$

$$\max D[G, c_{y,j}(t)], D = \frac{1}{|G|} \sum_{c_{x_i,j}(t) \in G} I[c_{x_i,j}(t); c_{y,j}(t)] \quad (8)$$

$$\min R(G), R = \frac{1}{|G|^2} \sum_{c_{x_i,j}(t), c_{x_{i'},j}(t) \in G} I[c_{x_i,j}(t); c_{x_{i'},j}(t)] \quad (9)$$

$$\max \phi(D, R), \phi = D - R \quad (10)$$

3.4. EEMD-MImRMR-BASBP Prediction Model

BP [21] is a multilayer feedforward neural network, the main feature of this network is a multilayer feedforward network with the forward transmission of signals and backward propagation of errors. BP are highly self-learning and adaptive and are particularly suitable in the field of multi-factor, uncertain and fuzzy information processing. However, the network structure of BP is not easily determined, and the back propagation process uses a gradient descent-based algorithm. In practice, the convergence speed is slow, and the network is easily trapped in local optimum by the influence of weights and thresholds, which affects the prediction effect.

For the problem of falling into the local optimum, it has been shown that using relevant, intelligent algorithms to optimize the initial weight threshold of the BP and then training the network twice can largely improve the network performance and greatly avoid the local optimum problem caused by random initialization of the weight threshold. For example, a genetic algorithm or a particle swarm algorithm can be used to find a suitable initial weight threshold for the BP [22,23]. However, both algorithms are based on population algorithms, and the algorithms themselves require more parameters to be set. Different parameter settings have a greater impact on the final, and in actual use, the parameters need to be tested several times according to different situations, and the algorithms are difficult to use.

In order to solve the above problem, Beetle Antennae Search (BAS) was utilized, which simulates the principle of aspen foraging and thus searches for the optimal solution in the global context [24]. It is possible to achieve automatic optimal search without knowing the specific form of the function, and its individual is only one, which is significantly smaller in computation and faster in searching for the optimal solution compared to genetic algorithms and particle swarm algorithms. The principal steps are as follows.

(1) For the k dimensional spatial optimization problem, x_{lk} and x_{rk} are the coordinates of the positions of the left and right whiskers at the k iteration, respectively; x^k are the coordinates of the center of mass, d_0 is the distance between the two whiskers, and the position of the aspen facing b . The expressions for the left and right whiskers are (11) and (12).

$$x_{lk} = x^k - d_0 * b / 2 \quad (11)$$

$$x_{rk} = x^k + d_0 * b / 2 \quad (12)$$

(2) For optimization problem f , define the anterior progress length of the aspen δ^k is defined, a is the adjustment factor, and the left and right whisker values are found, as in (13) and (14) [25].

$$f_{lk} = f(x_{lk}) \quad (13)$$

$$f_{rk} = f(x_{rk}) \quad (14)$$

Then the position of the aspen shifts with the left and right whisker values, as in (15) [25,26].

$$x^{k+1} = x^k - \delta^k * b * \text{sign}(f_{rk} - f_{lk}) + a \quad (15)$$

This model adopts the idea of cooperative research to integrate the BAS algorithm and BP neural network into one system, the former is used for system initial weight optimization, and the latter is used for system weight training, both of which are mutually reinforcing. The flow chart of BAS optimized BP network is shown in Figure 4.

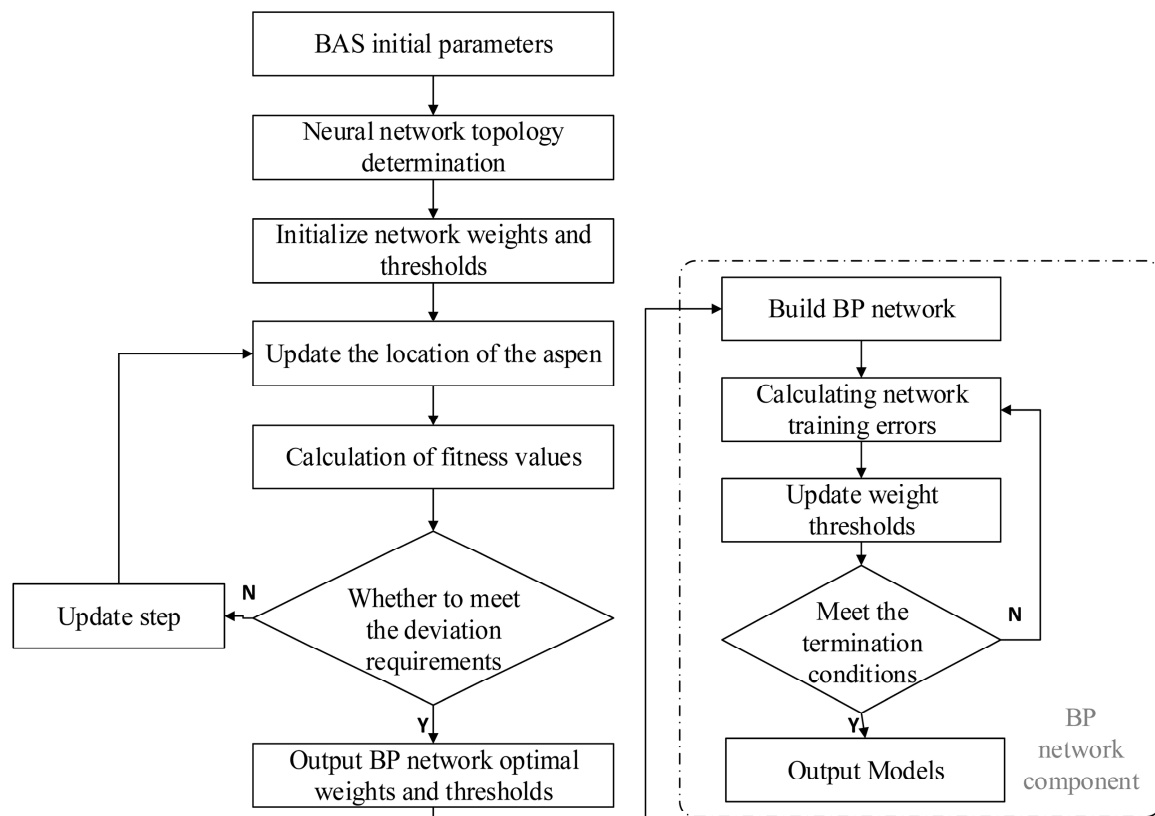


Figure 4. The BAS algorithm optimizes the BP network flowchart.

In summary, the pseudo-code of the EEMD-MImRMR-BASBP model algorithm is shown in Algorithm 1.

Algorithm1. EEMD-MImRMR-BASBP

Input: data set D of relevant variables

Output: The prediction results of the completed training

1: Data pre-processing: E = function EEMD (D , N_{std} , NE)

2: F = function Median Average Filter (E , T)

3: Timing Matching: G = function $MI(\text{Analysis Scope}, F)$

4: The data set after time-series matching through the time lag of G forms a new sample H

5: Variable Selection: I = function $mrmr_miq_d(H, f, k)$

6: Model Building: for BP repeat for I

7: Initialize BAS parameters

8: Input Data I

9: Determine the network topology

10: Initialize BP weights and thresholds

11: for net. train Param epochs = 300

12: Update the position of the beetle's left and right whiskers

13: Update beetle location

14: Calculation fitness = MSE

15: Update target network step

16: end for

17: Obtain optimal weights and thresholds

18: Find the unit output of the implicit and output layers

Deviation of the expected value from the actual value

Deviation to iteration requirements

19: end for

4. Experimental Analyses

The data from a cement company in Jiangxi province and Fuping city are used in this paper to test the proposed method. More specifically, the testing data set in this paper is obtained and configured through the Honeywell “URT Explorer” (Unified Real Time Explorer) data communication interface, which realizes bi-directional data batch transfer between DCS (Distributed Control System) and the system.

Then, the raw data were collected from a cement plant site, which totaled 25,900 sets and a sampling frequency of 10 s. For better validation of the model, the raw data from the two plants were divided into subsets A and B. The first 70% of each subset was used as the prediction training data set, and the last 30% was used as the prediction test data set. In order to verify the accuracy of the EEMD-MImRMR-BASBP model, a comparative analysis with other models is required. In this paper, four other models were constructed as comparisons: the BP model; the BASBP model; the PSOBP model; and the EEMD-MImRMR-PSOBP model. Among them, BP is a shallow neural network that has been successfully and widely used in the emission modeling of NO_x .

In order to better compare the model performance, the same model structure is used for each model, and two evaluation metrics, mean absolute error (MAE) and root mean square error (RMSE), are used to evaluate the model prediction ability separately. Specifically, RMSE contains a squared term and is more sensitive to residual outliers, i.e., it penalizes residuals with absolute values greater than 1 and rewards residuals with absolute values less than 1. The closer its value is to 0 represents a better fit, and MAE indicates the average degree of deviation of the predicted values from the true values. The MAE is a largely unbiased indicator, with no reward or penalty preference for residuals and little sensitivity to outliers. The mathematical expressions of the evaluation indicators are shown in Equations (16) and (17) [27].

$$MAE = \frac{1}{N} \sum_{i=1}^N |Y_i - \hat{Y}_i| \quad (16)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{N}} \quad (17)$$

where N is the number of test samples, Y_i is the actual measured value and $\hat{Y}_i$ is the model prediction.

(1) Data pre-processing

The raw data of the relevant input variables that are partially affected by noise in group A and the processed data based on EEMD and median average filtering are shown in Figure 5. The three variables from top to bottom are P_k , L_n and I_k . The pre-processing curve shows that the process variables are noisy, especially the kiln current, and therefore the direct use of raw data has a significant impact on the model. A comparison of the before and after plots show that the data pre-processing strategy has achieved good results for the process variables such as pressure and current and can effectively cope with the large noise characteristics of the cement calcination process data, laying the foundation for subsequent time lag estimation, variable selection and modeling.

(2) Time lag estimation

For large time lag processes, the time series matching of different input and output variables directly affects the validity of the model. The time lag interval is divided based on the analysis of the cement calcination process, and the magnitude of the mutual information between different input variables and output variables within the interval is calculated. By using the NO_x concentration time as the benchmark, the maximum time lags between the relevant input variables and output variables are confirmed, and the relevant time series are matched. The results of the time lags of the relevant input variables for their group A are shown in Table 3, respectively. From Table 3, it can be seen that the difference between the mRMR coefficients of T_S , T_D and I_{KC} without timing matching and the coefficient

values after timing matching is large, indicating that timing matching can find the relevant variables more accurately and improve the accuracy of subsequent variable selection.

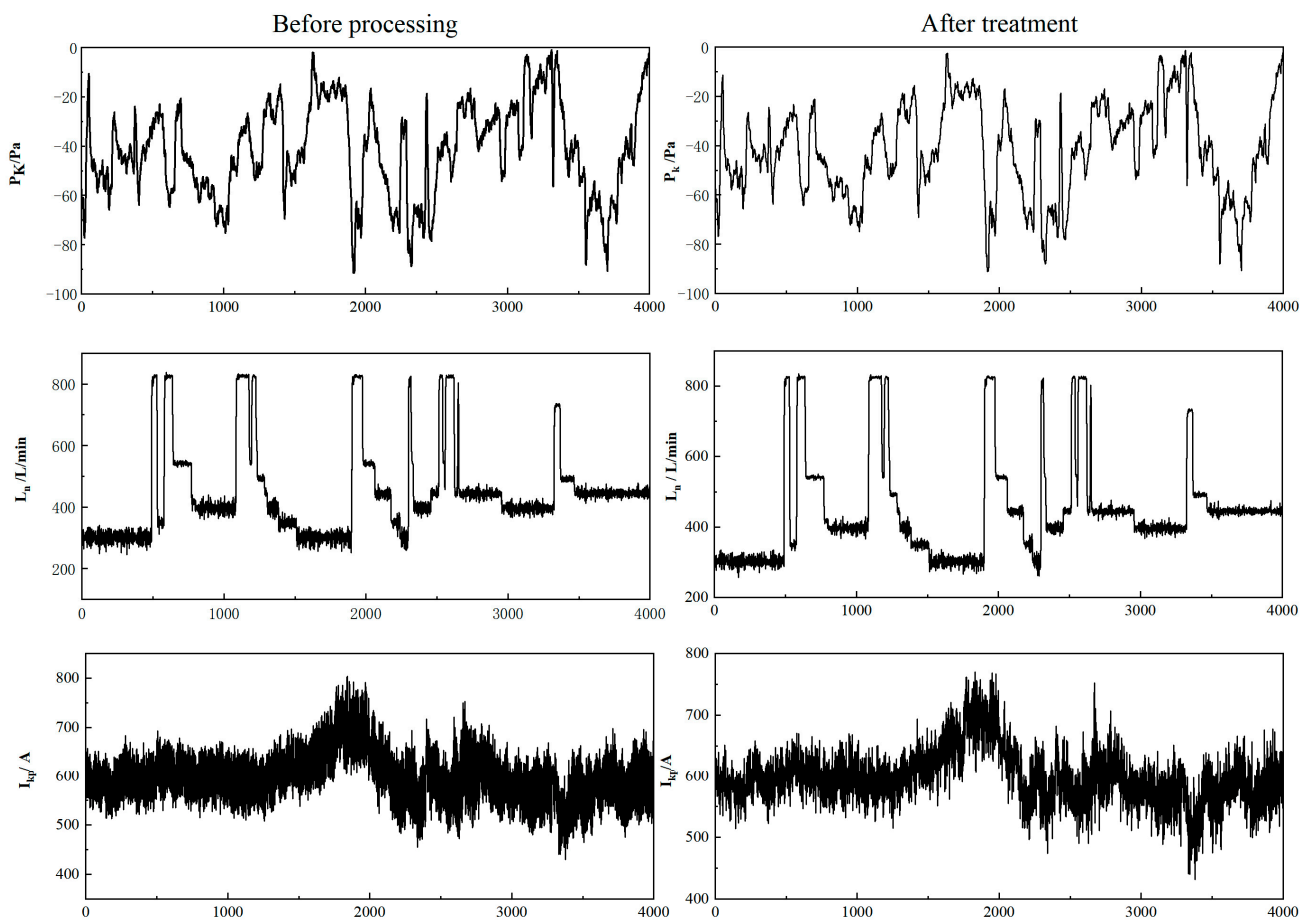


Figure 5. Comparison chart of relevant input variables before and after processing.

Table 3. Correlated input variables delay and correlated mRMR coefficient.

Definition	mRMR Coefficients after Processing of EEMD Data without MI Matching	mRMR Coefficients after Median-Averaged Filtering without MI Matching	Time Lag/s	MI-Matched mRMR Coefficients after EEMD Data Processing	MI-Matched Median-Averaged Filtered mRMR Coefficients
T_D	0.7231	0.6825	−10.4	0.8644	0.8278
T_{C5}	0.6022	0.5934	−10.75	0.7735	0.7649
T_S	0.5123	0.4957	−42	0.8782	0.8704
T_K	0.6921	0.8132	5	0.7833	0.8636
G_T	0.7835	0.8022	−4.5	0.8028	0.8544
G_W	0.8123	0.7953	−4	0.8015	0.8135
L_N	0.7568	0.8023	8.5	0.9198	0.9128
I_{KP}	0.6034	0.5984	7.5	0.7338	0.7048
C_{NO}	0.9273	0.9432	0.5	0.9368	0.9365
C_{CO}	0.7356	0.7562	−0.75	0.8172	0.8175
C_{O2}	0.5742	0.5234	17	0.7943	0.7625
P_K	0.8135	0.7975	−1	0.8094	0.8016
R_H	0.7214	0.7548	2.5	0.7878	0.7338
I_{KC}	0.7135	0.6979	7.8	0.8354	0.8679

(3) Selection of variables

On the basis of data processing based on EEMD combined with median-averaged filtering or median-averaged filtering and temporal matching of entropy principles for the relevant input and output variables, the characteristic variable selection method of mRMR is applied in order to reduce the number of variables. A larger mRMR coefficient represents a stronger correlation. mRMR coefficients were used to select the input variables for the NO_x prediction model, and the results are shown in Table 3. In this paper, the principle of maximum correlation was adhered to when selecting the input variables. The first six variables were selected in descending order based on the results in Table 3, which were C_{NO} , L_N , T_D , T_S , I_{Kc} and C_{CO} . The binding mechanism can be explained as follows: C_{NO} , representing the NO_x concentration at the C1 outlet, is closely related to the predicted NO_x concentration; L_N represents the denitrification efficiency, which can visually reflect the correctness of the NO_x prediction; T_D and T_S represent the temperature component of the cement calcination process, which affects the NO_x generation; I_{Kc} represents the amount of air in the calcination process, which has a great influence on the NO_x redox reaction; C_{CO} represents the working condition of the calcination process. This shows that the selection of variables covers the whole calcination process.

(4) Prediction results and evaluation of model performance

In order to further discuss the effectiveness of the mRMR algorithm, we conducted variable selection experiments to compare the non-linear variable selection methods listed in this paper (MIC and MI methods) using the same data set after processing and performed variable selection feasibility analysis based on the results in conjunction with the mechanism. The results are shown in Figure 6.

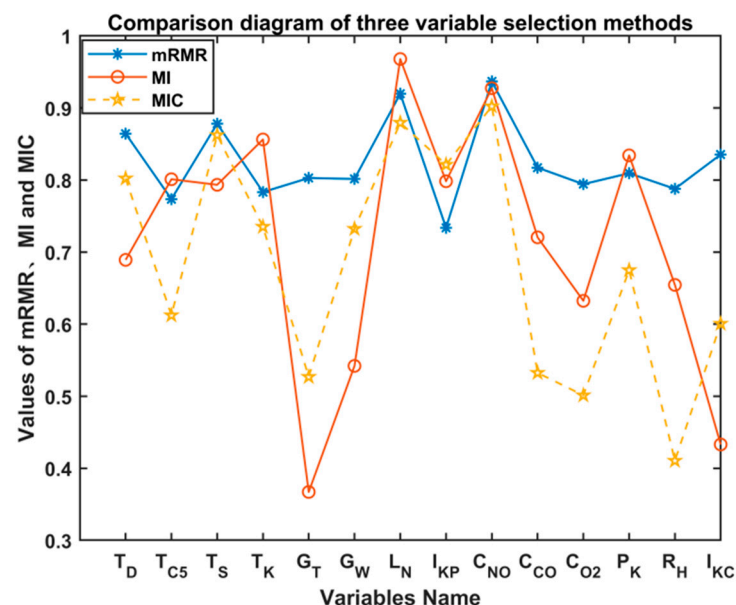


Figure 6. Comparison of variable selection methods.

As can be seen from Figure 6, the variables chosen by the three methods are not identical. If the first six variables are selected as input variables and analyzed with the mechanism of the first subsection, it can be concluded that the MI and MIC methods select too many temperature characterization variables and concentrate on the rotary kiln temperature characterization T_S and T_K . MI selects T_{C5} at the decomposer instead of T_D , which is not reasonable and makes the input variable assessment incomplete. Additionally, the mRMR algorithm can avoid the redundant relationship between temperature quantities well and thus select the best input variables.

After EEMD joint filtering of the test set A data and time series matching, the above six variables of C_{NO} , L_N , T_D , T_S , I_{Kc} and C_{CO} were selected by mRMR and substituted

into the established soft measurement model to predict the NO_x emission values, and the predicted and true values were obtained as shown in Figure 7, with $\text{RMSE} = 0.2927$, $\text{MAE} = 0.1795$, which is a good fit.

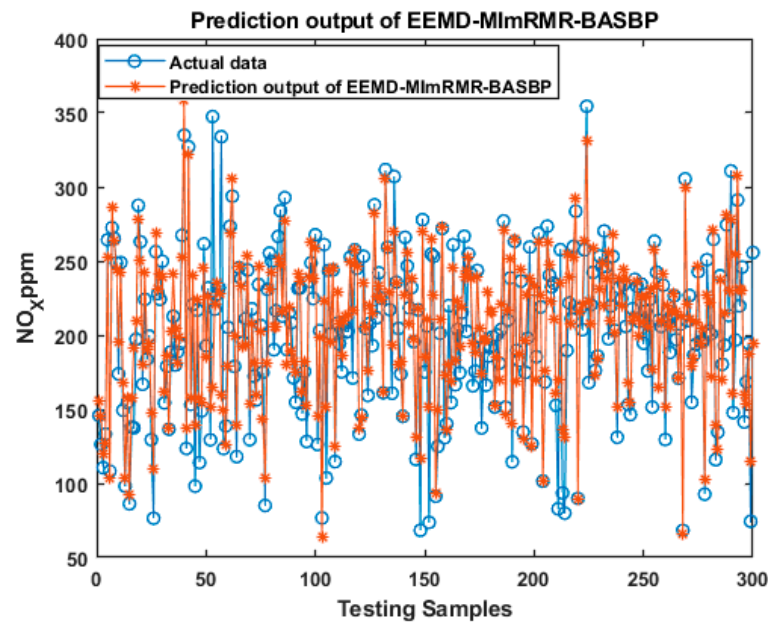


Figure 7. Model prediction results based on EEMD-MImRMR-BASBP strategy.

In order to further compare the performance of the models, the same data set was used, and the different prediction models were compared and analyzed by the two soft measurement model performance evaluation methods proposed in Section 3. where BP, BASBP and PSOBP denote NO_x based modeling without data processing and variable selection, and the rest of the parameters were kept the same during the experimental comparison. The results are shown in Table 4.

Table 4. Comparison of indicators for different modeling strategies.

Modeling Strategies	Input Variables	RMSE for Group A	MAE for Group A	RMSE for Group B	MAE for Group B
BP	$T_D, T_{C5}, T_S, T_K, G_T, G_W, L_N, I_{KP}, C_{NO}, C_{CO}, C_{O2}, P_K, R_H, I_{KC}$	0.9012	0.8643	0.9134	0.8973
PSOBP	$T_D, T_{C5}, T_S, T_K, G_T, G_W, L_N, I_{KP}, C_{NO}, C_{CO}, C_{O2}, P_K, R_H, I_{KC}$	0.8025	0.7534	0.8315	0.7954
BASBP	$T_D, T_{C5}, T_S, T_K, G_T, G_W, L_N, I_{KP}, C_{NO}, C_{CO}, C_{O2}, P_K, R_H, I_{KC}$	0.7028	0.6956	0.8012	0.7834
EEMD-MImRMR-PSOBP	$C_{NO}, L_N, T_S, T_D, I_{KC}, C_{CO}$	0.6023	0.5347	0.6953	0.6026
EEMD-MImRMR-BASBP	$C_{NO}, L_N, T_S, T_D, I_{KC}, C_{CO}$	0.2927	0.1795	0.3513	0.2383

The model performance index values for each of the five NO_x prediction models for the two production lines are given in Table 4. As can be seen from the table, the EEMD-MImRMR-BASBP model has better prediction results for both production lines. First, in terms of RMSE and MAE, the EEMD-MImRMR-BASBP model improved the mean values of 70.4%, 66.8%, 64.6% and 56.9%, respectively, over the other four models. It has higher accuracy type and stability compared with the other four models. Both PSO and BASBP predictions were significantly improved after using EEMD filtering combined with the data processing method and mRMR method to select auxiliary variables.

From the perspective of data processing and variable selection, the EEMD-MImRMR-BASBP model is superior to the BASBP model, and EEMD-MImRMR-PSOBP is superior to PSOBP. It is shown that EEMD combined with digital filtering effectively handles large, noisy process data, provides high-quality modeling data and improves the model accuracy; additionally, it verifies the effectiveness of the feature variable selection strategy of the mRMR algorithm.

From the perspective of the prediction algorithm models, the BASBP model outperforms the traditional BP model and PSOBP model, and the EEMD-MImRMR-BASBP model outperforms the EEMD-MImRMR-PSOBP model, verifying that the BAS search algorithm improved BP method has higher prediction accuracy and more accurate NO_x prediction compared with the traditional BP and PSO improved BP methods.

The NO_x concentration predicted in advance can provide a reference for SNCR ammonia flow control, change the existing lag control to feedforward control, improve the control accuracy, and can effectively guide the actual NO_x index optimization, empowering large cement production enterprises to reduce emissions and environmental protection and green manufacturing.

5. Conclusions

In this paper, a new NO_x concentration prediction model EEMD-MImRMR-BASBP is proposed. The validity of the proposed method is also verified by actual production line experiments, and the main conclusions are as follows:

(1) The proposed data processing module combining EEMD and median average filtering can effectively perform noise removal from non-linear, non-smooth signals and provide a solid foundation for subsequent timing matching and variable selection.

(2) After the MI of the proposed entropy principle is matched in the corresponding time series, six feature variables are then selected as inputs to the prediction model in combination with the mRMR algorithm, reducing the coupling between variables and eliminating the impact of information redundancy on prediction accuracy. Additionally, on the basis of building the model, the BP is improved using the BAS algorithm. For different production lines, the prediction errors of BASBP are 0.7028, 0.6956 and 0.8012, and 0.7834, which are lower than the errors of BP and PSOBP algorithms in the same period. It indicates that the improved BP has better estimation accuracy and prediction results.

(3) The proposed hybrid model strategy of the EEMD-MImRMR-BASBP model can accurately represent the predicted concentration of NO_x with smaller errors, with RMSE and MAE of 0.2927, 0.3513 and 0.1795, and 0.2383 for the two production lines, respectively, which are much better than other models, indicating that this model has a better predictive ability.

(4) Accurate NO_x prediction model results can guide cement denitrification operators to spray ammonia accurately to reduce NO_x emissions and reduce ammonia usage and ammonia escapes. In terms of subsequent cement calcination emission reduction, it is important to reduce NO_x generation through the optimization of cement calcination parameters based on NO_x prediction values without changing the actual process, which can create significant economic and environmental benefits for enterprises.

Author Contributions: Conceptualization, X.L. and Y.C.; methodology, X.L.; software, X.L.; validation, X.L.; formal analysis, X.L.; investigation, X.L.; resources, Y.C.; data curation, X.L.; writing—original draft preparation, X.L.; writing—review and editing, Y.Y., N.H. and Y.C.; visualization, Y.C.; funding acquisition, Y.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the NATIONAL KEY R&D PROGRAM OF CHINA, grant number 2016YFB0303402, SHAANXI PROVINCE KEY SCIENCE AND TECHNOLOGY INNOVATION TEAM, grant number 2021TD-53 and SHAANXI PROVINCIAL NATURAL SCIENCE BASIC RESEARCH PROGRAM, grant number 2019JLZ-05. The APC was funded by the NATIONAL KEY R&D PROGRAM OF CHINA.

Data Availability Statement: Data are available on request.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

- Li, S.; Ge, Y.; Wei, X. Experiment on NO_x reduction by advanced reburning in cement precalciner. *Fuel* **2018**, *224*, 235–240. [\[CrossRef\]](#)
- Yoo, K.S.; Park, S.W. Improvement of De NO_x efficiency of SNCR process with chemical additives in urea solution. *J. Korea Acad.-Ind. Coop. Soc.* **2017**, *18*, 663–668.
- Huang, N.E. The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis. *Proc. Math. Phys. Eng. Sci.* **1998**, *454*, 903–995. [\[CrossRef\]](#)
- Kaneko, H.; Funatsu, K. Smoothing-combined soft sensors for noise reduction and improvement of predictive ability. *Ind. Eng. Chem. Res.* **2015**, *54*, 12630–12638. [\[CrossRef\]](#)
- Donoho, D.L. *De-Noising by Soft-Thresholding*; IEEE Press: Piscataway, NJ, USA, 1995.
- Fu, R.; Yang, Y.; Yu, B.; Liu, C.; Zhang, C. High-order tensor feature Extraction from eeg Signals based on ensemble empirical Mode Decomposition. *Acta Metrol. Sin.* **2021**, *42*, 1680–1686.
- Kravchonok, A.I.; Zalesky, B.A.; Lukashevich, P.V. An algorithm for median filtering on the basis of merging of ordered columns[]]. Pattern recognition and image analysis: Advances in mathematical theory and applications in the USSR. *Pattern Recognit. Image Anal.* **2007**, *17*, 402–407. [\[CrossRef\]](#)
- Reshef, D.N.; Reshef, Y.A.; Finucane, H.K.; Grossman, S.R.; McVean, G.; Turnbaugh, P.J.; Lander, E.S.; Mitzenmacher, M.; Sabeti, P.C. Detecting novel associations in large data sets. *Science* **2011**, *334*, 1518–1524. [\[CrossRef\]](#) [\[PubMed\]](#)
- Xin, C.; Xu, J. Research on energy hub configuration optimization problem based on analytic hierarchy and grey correlation analysis. *J. Phys. Conf. Ser.* **2022**, *2354*, 012012. [\[CrossRef\]](#)
- Peng, H.; Long, F.; Ding, C. Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 1226–1238. [\[CrossRef\]](#)
- Tang, Z.H.; Zhang, Z.J. The Multi-objective optimization of combustion system operations based on deep data-driven models. *Energy* **2019**, *182*, 37–47. [\[CrossRef\]](#)
- Tan, P.; He, B.; Zhang, C.; Rao, D.; Li, S.; Fang, Q.; Chen, G. Dynamic modeling of NO_x emission in a 660MW coal-fired boiler with long short-term memory. *Energy* **2019**, *176*, 429–436. [\[CrossRef\]](#)
- Huang, Q.; Xie, L.; Yin, G.; Ran, M.; Liu, X.; Zheng, J. Acoustic signal analysis for detecting defects inside an arc magnet using a combination of variational mode decomposition and beetle antennae search. *ISA Trans.* **2020**, *102*, 347–364. [\[CrossRef\]](#) [\[PubMed\]](#)
- Jiang, X.; Li, S. BAS: Beetle antennae search algorithm for optimization problems. *arXiv* **2017**, arXiv:1710.10724. [\[CrossRef\]](#)
- Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [\[CrossRef\]](#)
- Kozachenko, L.F.; Leonenko, N.N. Sample Estimate of the Entropy of a Random Vector. *Probl. Peredachi Inf.* **1987**, *23*, 9–16.
- Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; John Wiley & Sons: Hoboken, NJ, USA, 2006.
- Özyurt, F. A fused CNN model for WBC detection with MRMR feature selection and extreme learning machine. *Soft Comput.* **2020**, *24*, 163–172. [\[CrossRef\]](#)
- Huang, L.S.; Ruan, C.; Huang, W.J.; Shi, Y.; Peng, D.L.; Ding, W.J. Wheat Powdery mildew monitoring based on GF-1 remote sensing image and relief-mRMR-GASVM model. *Trans. Chin. Soc. Agric. Eng.* **2018**, *34*, 167–175, 314.
- Cheng, X.M.; Shen, Z.F.; Xing, T.Y.; Xia, L.G.; Wu, T.J. Efficiency and accuracy analysis of multi-spectral remote sensing image classification based on mRMR feature optimization algorithm. *J. Geo-Inf. Sci.* **2016**, *18*, 815–823.
- Osowski, S. Signal flow graphs and neural networks. *Biol. Cybern.* **1994**, *70*, 387–395. [\[CrossRef\]](#)
- Jin, G.; Feng, W.; Meng, Q. Prediction of Waterway Cargo Transportation Volume to Support Maritime Transportation Systems Based on GA-BP Neural Network Optimization. *Sustainability* **2022**, *14*, 13872. [\[CrossRef\]](#)
- Zhang, M. Classification Prediction of Rockburst in Railway Tunnel Based on Hybrid PSO-BP Neural Network. *Geofluids* **2022**, *2022*, 4673073. [\[CrossRef\]](#)
- Sun, F.; Wang, X.; Han, P.; He, B. Combined EIS and BAS-BP neural network analysis of electrochemical corrosion on pipeline steel in silty soil in a Salt-Temperature coupling environment. *Int. J. Press. Vessel. Pip.* **2022**, *200*, 104807. [\[CrossRef\]](#)
- Silva, D.A.; Alves, G.I.; de Mattos Neto, P.S.; Ferreira, T.A. Measurement of fitness function efficiency using data envelopment analysis. *Expert Syst. Appl.* **2014**, *41*, 7147–7160. [\[CrossRef\]](#)
- Lima Junior, A.R.; Silva, D.A.; Mattos Neto, P.S.; Ferreira, T.A. An Experimental Study of Fitness Function and Time Series Forecasting Using Artificial Neural Networks. In Proceedings of the 12th Annual Conference Companion on Genetic and Evolutionary Computation, Portland, OR, USA, 7–11 July 2010; pp. 2015–2018.
- Zhang, F.; Gong, T.; Lee, V.E.; Zhao, G.; Rong, C.; Qu, G. Fast algorithms to evaluate collaborative filtering recommender systems. *Knowledge-Based Syst.* **2016**, *96*, 96–103. [\[CrossRef\]](#)

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.