energies

MDPI

Article

A Hybrid Method for Prediction of Ash Fouling on Heat Transfer Surfaces

Fangshu Cui [1], Sheng Qin [1], Jing Zhang [1], Mengwei Li [2] and Yuanhao Shi [2,3,*]

[1] School of Data Science and Technology, North University of China, Taiyuan 030051, China; fscui@nuc.edu.cn (F.C.); shengqin2016@163.com (S.Q.); izhang_jing@163.com (J.Z.)
[2] School of Electrical and Control Engineering, North University of China, Taiyuan 030051, China; lmw1997614@163.com
[3] Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China
* Correspondence: yhshi@nuc.edu.cn; Tel.: +86-136-536-42822

Abstract: Soot blowing optimization is a key, but challenging question in the health management of coal-fired power plant boiler. The monitoring and prediction of ash fouling for heat transfer surfaces is an important way to solve this problem. This study provides a hybrid data-driven model based on advanced machine-learning techniques for ash fouling prediction. First, the cleanliness factor is utilized to represent the level of ash fouling, which is the original data from the distributed control system. The wavelet threshold denoising algorithm is employed as the data preprocessing approach. Based on the empirical mode decomposition (EMD), the denoised cleanliness factor data is decoupled into a series of intrinsic mode functions (IMFs) and a residual component. Second, the support vector regression (SVR) model is used to fit the residual, and the Gaussian process regression (GPR) model is applied to estimate the IMFs. The cleanliness factor data of ash accumulation on the heat transfer surface of diverse devices are deployed to appraise the performance of the proposed SVR + GPR model in comparison with the sole SVR, sole GPR, SVR + EDM and GPR + EDM models. The illustrative results prove that the hybrid SVR + GPR model is superior to other models and can obtain satisfactory effects both in one-step- and the multistep-ahead cleanliness factor predictions.

Keywords: coal-fired power plant boiler; ash fouling prediction; cleanliness factor; machine-learning technique

check for updates

Citation: Cui, F.; Qin, S.; Zhang, J.; Li, M.; Shi, Y. A Hybrid Method for Prediction of Ash Fouling on Heat Transfer Surfaces. *Energies* **2022**, *15*, 4658. https://doi.org/10.3390/en15134658

Received: 10 May 2022
Accepted: 23 June 2022
Published: 25 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 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

Although new energy is developing and growing rapidly, coal-fired power generation still plays an important role in energy supply in many countries, due to the energy structure problems in the future for quite a long time [1–4], especially in China. To achieve carbon peak and neutrality targets, energy conservation and emission reduction in coal-fired power generation, as a large supplier of energy consumption and pollutants, have become particularly important [5–8].

Ash deposition has a significant impact on the heat transfer surfaces in coal-fired power plant boilers, as it reduces the heat transfer capability of the surfaces [9]. In addition, it causes higher CO_2 emissions and increases of CO and NOx emissions due to incomplete combustion, and safety accidents [7,10]. According to the research, ash deposition causes at least 1% energy loss under normal conditions [11]. Therefore, it is particularly important to reduce the influence of ash deposition on the heat transfer surfaces [12–14].

Soot blowing is the procedure of using high-temperature steam, or other mediums, to blast the deposits on the heat transfer surfaces. However, untimely soot blowing reduces the thermal efficiency of heat transfer surfaces duo to ash deposition, and too frequent soot blowing will lead to the waste of main steam, raised maintenance loss, and tube corrosion. Thus, soot blowing optimization of coal-fired power plant boilers is one of the research hotspots to which researchers and engineers have been paying attention [15,16].

Moreover, accurate fouling rate prediction strongly influences the efficiency and safety of coal-fired power plants, which is one of the important parts in the production process [17]. Therefore, reasonable monitoring and accurate prediction form the basis of the optimization algorithm's decision-making. Improving the accuracy and robustness of ash fouling rate prediction has become highly desirable and should not be delayed.

In recent decades, many ash fouling monitoring and prediction approaches have been developed, and can be divided into two types: mechanism model methods and data driven model methods. Chen et al. proposed a predicting model of ash fouling considering heat transfer resistance and the effect of flue gas velocity [18]. Shu et al. developed a dynamic mathematical model of fly ash and sticking in coal fired power plant boilers based on a direct simulation Monte Carlo method [19]. Tang et al. used a comprehensive fouling model integrated with the CFD framework, which was carried out with a discrete phase model and a dynamic mesh model in Ansys Fluent (Canonsburg, PA, USA) [20].

However, due to the complexity and nonlinear characteristics of the ash deposition mechanism and the complexity and variability of production environments, it is difficult to obtain an accurate mathematical model to match all working conditions. With the development of artificial intelligence, some machine-learning methods are being used to predict the ash fouling rate of heat transfer surfaces. Tong et al. proposed an online ash fouling prediction method based on wavelet analysis and support vector regression (SVR) [21]. Shi et al. used a model-based indirect ash monitoring method with support vector machine (SVM) and artificial neural network (ANN) to predict ash fouling [22,23]. Towards soot blowing optimization, Pena et al. proposed a soft computing-based model to predict the ash fouling status of heat transfer surfaces [15,24]. Tang et al. proposed an fouling prediction and multi-objective optimization method using neural networks and genetic algorithms (GA) [25].

Nevertheless, the prediction methods mentioned above have their own advantages and disadvantages. Moreover, most of these approaches use simple data preprocessing to improve the prediction performance. They cannot completely capture the main features of ash fouling series. Therefore, to improve the effectiveness of the data preprocessing method, empirical mode decomposition (EMD) is developed to solve the challenging task [26].

Therefore, to develop an effective ash fouling rate prediction technique for soot blowing optimization, this paper exploits machine learning techniques to derive a new data-driven hybrid method, enabling precise future ash fouling prediction and good extrapolation properties. Particularly, several pivotal contributions are made as follows:

1. By employing the EMD, the denoised cleanliness factor data is decomposed into several intrinsic mode functions (IMFs) and a residual sequence;
2. The SVR model is utilized to fit the residual signal, which can capture the long-term cleanliness factor degeneration. Meanwhile, the Gaussian process regression (GPR) model is applied to fit the IMFs separately, which can capture the local fluctuations;
3. The predictive property of several machine learning models are investigated, and the results prove that the combined SVR + GPR model outperforms other models;
4. For both the one-step- and multistep-ahead cleanliness factor predictions, the proposed hybrid method has high precision and good extrapolating performance.

The remaining sections of this paper are organized as follows. Section 2 presents the problem statement and data description. Section 3 describes the adopted techniques and the overall structure of this article. In Section 4, the data preprocessing and the accuracy indicators are given, followed by performance comparison tests of various models. Moreover, the experimental predicted results of the proposed method are analyzed. Finally, Section 5 concludes this paper.

2. Problem Statement and Data Description

In this paper, a 300 MW coal-fired power plant boiler in Guizhou Province in China is used as a case study. The schematic diagram of boiler heat transfer flow char is shown in Figure 1; the boiler is a drum-type with W-shape flame and reheating. The main

heat transfer surfaces include the single furnace, platen superheater, high temperature superheater, low temperature superheater, reheaters, economizers, and air preheaters.

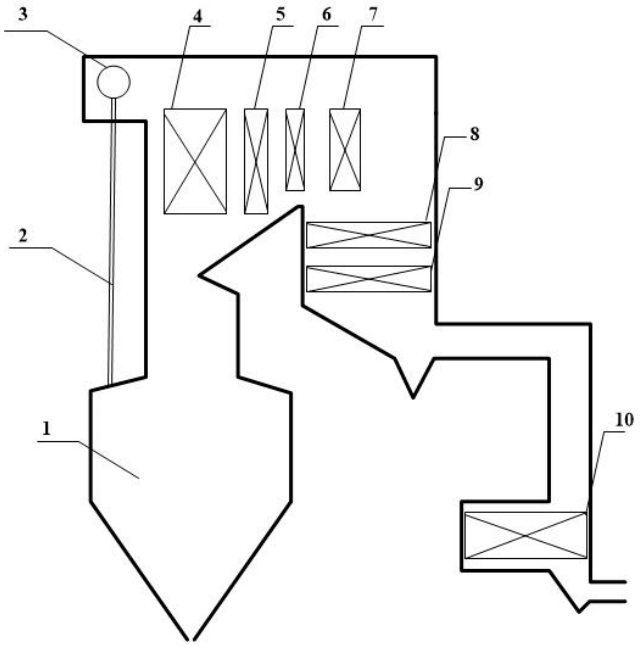


Figure 1. Flow diagram of boiler heat transfer. 1. Furnace 2. Downcomer, 3. Steam drum, 4. Platen superheater, 5. High temperature superheater, 6. High temperature reheater, 7. Low temperature superheater, 8. Low temperature reheater, 9. Economizer, 10. Air Preheater.

The heat transfer mode of the boiler mainly includes the radiation heat transfer and convective heat transfer. As shown in Figure 2, taking the convective heat transfer mode of the tail flue as an example, the fouling layer will reduce the heat transfer efficiency.

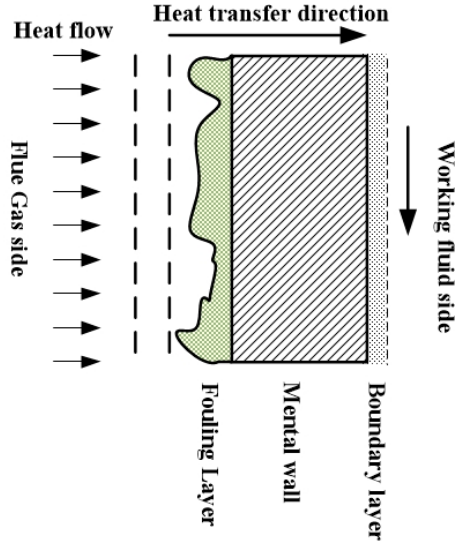


Figure 2. Schematic diagram of convectional heat transfer process.

Generally, due to the complex working conditions of the heat transfer surfaces in a boiler, the ash accumulation state on the heat transfer surfaces cannot be measured directly. The ash fouling state of the boiler, in combination with indirect parameters, can be used to reflect the fouling accumulation state of the heat transfer surfaces. In this paper, the cleanliness factor (CF) represents the ash fouling state [27].

The CF f_c of each heat transfer surface is defined as follow:

$$f_c = \frac{K_r}{K_0}, \tag{1}$$

where K_r and K_0 are the actual heat transfer coefficient of the heat transfer surface and the theoretical heat transfer coefficient, respectively. Obviously, the values of CF f_c lie in interval [0, 1], with one corresponding to the clean status of the heat transfer surface.

Theoretical heat transfer coefficient K_0 denotes the heat transfer efficiency of the heat transfer surface in the clean status, which is usually expressed as the sum of the theoretical radiation heat transfer coefficient and the theoretical convection heat transfer coefficient.

A cleanliness factor dataset in the boiler is evaluated to develop an effective ash fouling prediction system for soot blowing optimization. These include a datasets economizer, high temperature superheater (HTS), low temperature superheater (LTS), and reheater with a 1 min period. The total number of each dataset is 595, which are divided into training data and testing data.

3. Techniques

In order to acquire the reliable future cleanliness factor, the proposed hybrid method primarily applies three techniques: the EMD method to decouple the denoised ash accumulation dataset, an SVR model to capture long-term dynamics, and a GPR model to capture local fluctuations.

3.1. Empirical Mode Decomposition

EMD is a valid signal processing technique that has powerful capabilities in decomposing highly dynamic signals into a residual signal and a suite of IMFs [26]. The EMD algorithm has been widely used in numerous real-world problems, such as rotation mechanisms, ocean waves and so on. Particularly, EMD needs to satisfy the following conditions:

- In the total dataset, the number of extrema and the number of zero-crossings must be equal, or at most, different by one;
- At any dot, the envelopes that are defined by the local extrema must produce a zero mean.

The detailed process of the EMD decomposition is as follows:

Step 1: According to the upper and lower extreme points of the original signal, the upper and lower envelope lines are drawn respectively.
Step 2: Take the mean of the upper and lower envelope and plot the mean envelope.
Step 3: Subtracting the mean envelope from the original signal gives the intermediate signal.
Step 4: Judge whether the intermediate signal meets the above two conditions. If so, the signal is an IMF component. If not, perform step 1 to step 4 analysis again based on this signal. Obtaining IMF components usually requires several iterations.
Step 5: Each time you get an IMF, subtract it from the original signal and repeat the above steps, all the way to the end where the signal is just a monotone sequence or a constant sequence.

Therefore, the original signal $x(t)$ is the linear superposition of a series of IMFs, and the rest by the EMD method:

$$x(t) = \sum_{i=1}^{n} IMF_i(t) + r_n(t), \tag{2}$$

where $IMF_i(t)$ and $r_n(t)$ stand for the IMF sequence and the residual component, respectively.

3.2. Support Vector Regression

SVM is a valid machine learning technique that is used for the nonlinear problems. Generally, in the SVM, kernels are applied to transform a nonlinear issue with a lower feature space into a linear issue in a higher dimensional space [28]. It is a flexible method

that is able to model complicated problems when offered enough data. The SVM makes forecasts by using the following function [29]:

$$y(x) = \sum_{i=1}^{N} \omega_i K(x, x_i) + \varepsilon, \tag{3}$$

where ω_i represent model weights which connect the feature space to the output, $K(\cdot)$ is the kernel function and ε stands for an independent noise term.

The kernel function is of great significance for the SVM. There are many types of kernel functions, which are as follows: linear kernel, polynomial kernel, radial basis function (RBF) kernel, etc.

SVR is presented based on the SVM, and is an advantageous tool for nonlinear regression issues. The SVR copes with data in a high dimensional space by applying linear quadratic programming techniques, which give the best performance for regression results [30]. Hence, it has good and stable ability in prediction for nonlinear time series.

The differences between SVR and SVM are as follows: the former has only one kind of data, and the goal is to minimize the deviation from all data points to the optimal solution; the latter has many kinds of data, and the goal is to separate these kinds of data.

3.3. Gaussian Process Regression

Stemming from a Bayesian framework, GPR is a nonparametric, probabilistic and flexible approach that has been broadly used in prediction analyses [31–34]. The GPR is a machine learning method based on many kernel functions, and the probability distribution of GPR is described as follows:

$$F(x) \sim GPR(m(x), \kappa(x, x')), \tag{4}$$

where $m(x)$ and $\kappa(x, x')$ represent the mean function and the covariance function, respectively. In reality, there are multiple kernel functions that can be chosen for $\kappa(x, x')$.

In a regression procedure, a function plus an additive noise ($\varepsilon \sim N(0, \sigma_n^2)$) is used to model the output. The prior distribution for observations is expressed by the following function:

$$y \sim N(0, \kappa(x, x') + \sigma_n^2 I_n). \tag{5}$$

Letting the new dataset x' follow an approximate Gaussian distribution with training set x, the total joint prior distribution of known outputs y and prediction outputs y' will be denoted by the following function:

$$\begin{pmatrix} y \\ y' \end{pmatrix} \sim N\left(0, \begin{pmatrix} \kappa(x, x) + \sigma_n^2 I_n & \kappa(x, x') \\ \kappa(x, x')^T & \kappa(x', x') \end{pmatrix}\right). \tag{6}$$

Next, outputs are forecasted by computing the conditional distribution $p(y'|x, y, x')$ with the following function:

$$p(y'|x, y, x') = N(y'|x', \mathrm{cov}(y')), \tag{7}$$

where

$$\begin{cases} z' = \kappa(x, x')^T [\kappa(x, x) + \sigma_n^2 I_n]^{-1} y \\ \mathrm{cov}(y') = \kappa(x', x') - \kappa(x, x')^T [\kappa(x, x) + \sigma_n^2 I_n]^{-1} \kappa(x, x') \end{cases}. \tag{8}$$

The conditional distribution $p(y'|x, y, x')$ also follows a Gaussian distribution. z' can be regarded as the prediction value of y'. $\mathrm{cov}(y')$ is a covariance matrix on behalf of the uncertainty.

3.4. Overall Structure

The overall framework and algorithm for predicting the future cleanliness factor based on the combined SVR + GPR model are shown in Figures 3 and 4, respectively.

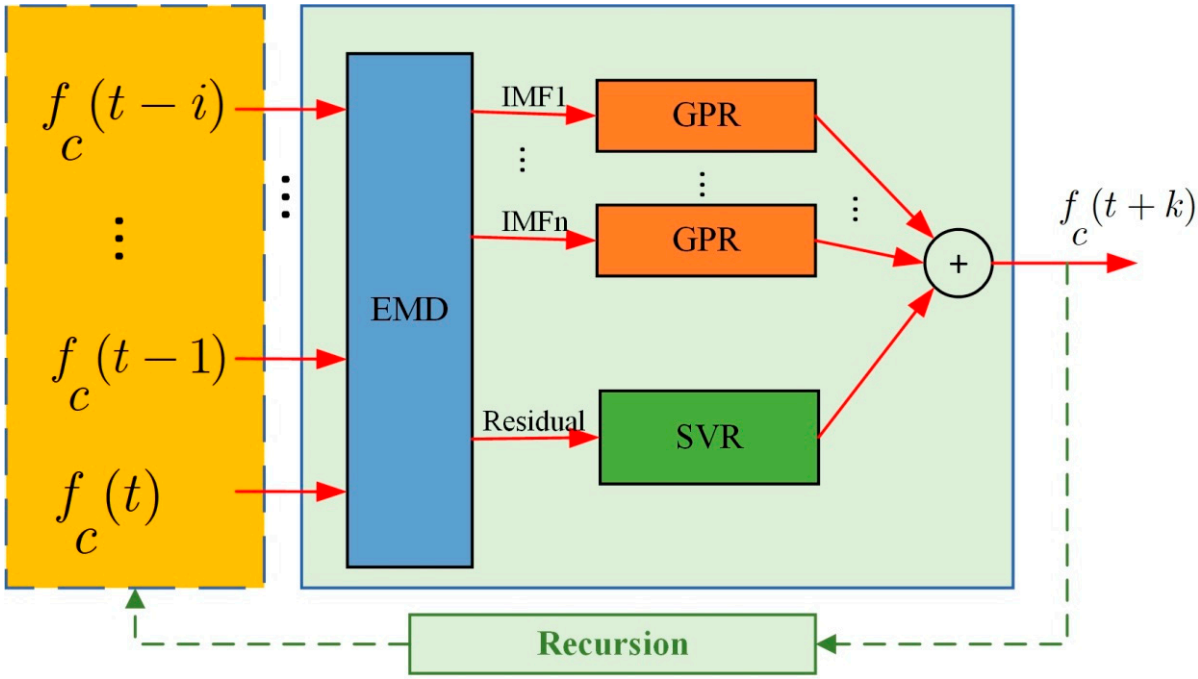


Figure 3. Framework of predicting the future cleanliness factor by combining the SVR and GPR models.

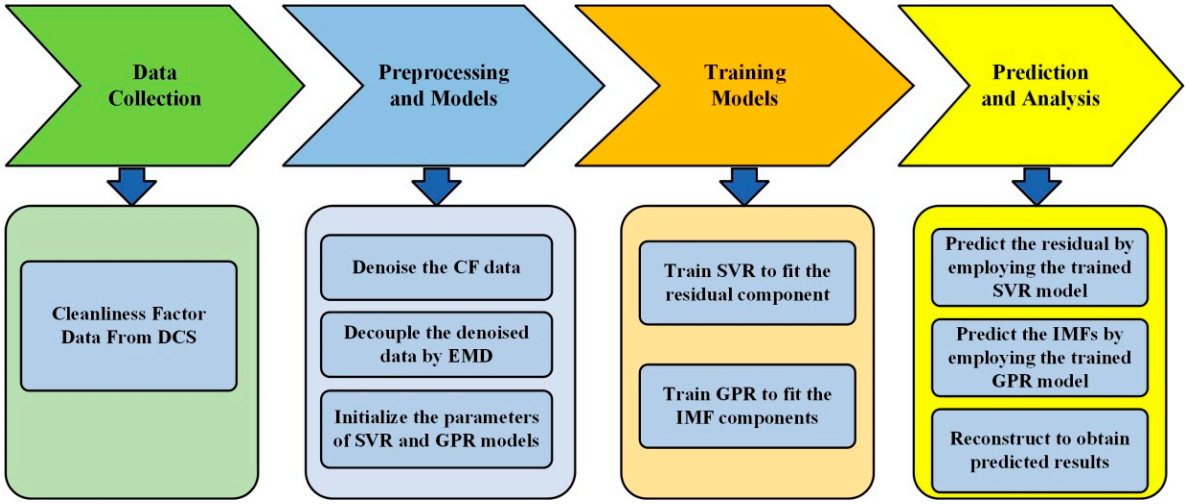


Figure 4. Workflow for predicting future cleanliness factor based on the proposed SVR + GPR model.

As shown in Figure 3, with the current and historical cleanliness factor vectors $[f_c(t-i), \ldots, f_c(t-1), f_c(t)]$ as model inputs, the output $f_c(t+k)$ can be forecasted by using the hybrid SVR + GPR model. Here, i and k denote the previous and future steps, respectively. Furthermore, the previously predicted cleanliness factor is employed as the next input of the model to further forecast the new cleanliness factor value. Hence, we should know that the prediction is guided just based on the historical cleanliness factor information. The detailed procedures of the whole algorithm framework are shown as follows:

Step 1: For data preprocessing, the original cleanliness factor degradation data is denoised by using the wavelet threshold denoising theory. Then, the EMD technique is utilized to decouple the data into several IMFs and a residual. For the model part, the suitable kernel functions are chosen for SVR and GPR, respectively. Initialize the parameters for both the SVR and GPR models.

Step 2: For the residual sequence, train the SVR model to fit the residual signal. For the IMFs, train the GPR models to fit the IMFs separately.

Step 3: Apply the well-trained SVR model to forecast the future residual value, and use the well-trained GPR models to predict the values of each IMF. Combine these results to obtain the predicted cleanliness factor value.

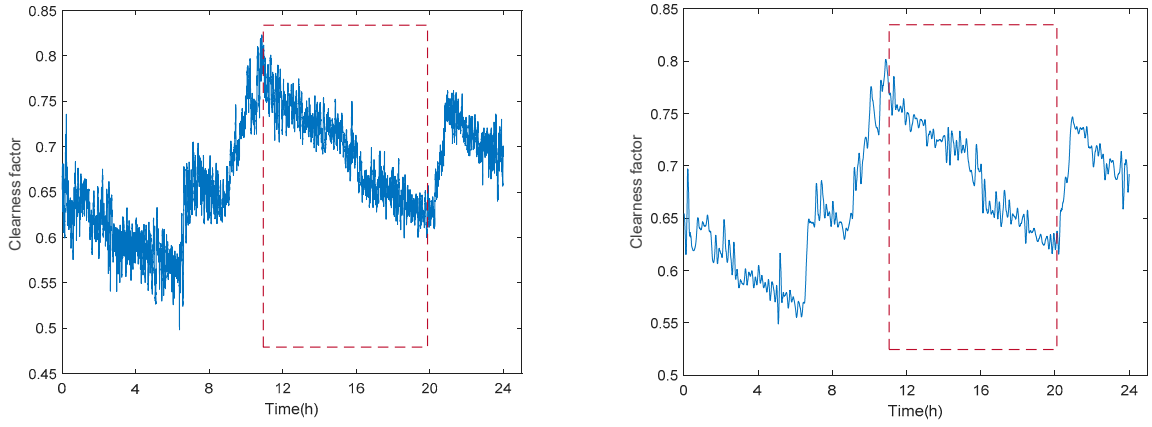
Following this program, the future cleanliness factor can be calculated, which can be used for early warning. This allows extra time to conduct preparation operations and personnel allocation for soot blowing.

4. Results and Discussions

In this section, the raw cleanliness factor dataset, which comes from the economizer, low-temperature superheater, high-temperature superheater and reheater in the coal-fired boiler of the Guizhou thermal power station, is denoised by the wavelet threshold theory and decomposed by EMD. Then, the model evaluation indicators are given to quantify the prediction performance, followed by a comparison of different models to examine their effectiveness. Finally, both one-step- and multistep-ahead cleanliness factor predictions are conducted. Here, all the tests are performed in Python 3.7 with a 3.00 GHz Intel (R) Core (TM) i7-9700 CPU.

4.1. Preprocessing

Due to the strong noise of the raw data, the original cleanliness factor data of each device is denoised using wavelet threshold theory, which is illustrated in Figure 5. Because the cleanliness factor sequences still have strong nonlinear and nonstationary features after denoising, each of them can be taken as a mixed signal with multi-scale signals. By using EMD through an iterative sifting procedure, the nonstationary dataset can be decoupled into a residual component and a series of IMFs. Figure 6 shows the corresponding decomposition results, in which three extracted IMFs and a residual are acquired with a low amount of calculation of about 0.5 s. Particularly, the local fluctuations have been eliminated by the EMD and the received residual sequence represents a whole monotonous trend to depict long-term dynamics. Simultaneously, all the local undulations are captured by the IMFs. Generally, IMF1 fluctuates strongly, while the trend of IMF2 is comparatively gentle, and the IMF3 is the gentlest.



(**a**) (**b**)

Figure 5. Comparison diagram of cleanliness factor curve of economizer before (**a**) and after (**b**) denoising by using wavelet threshold theory.

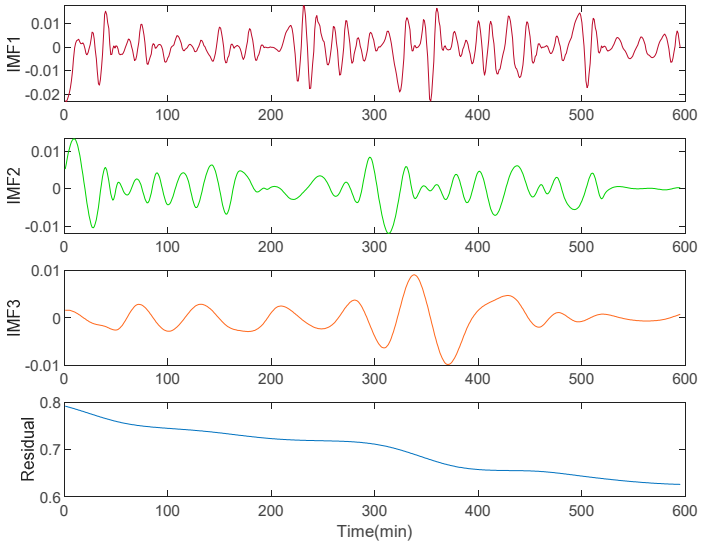


Figure 6. IMFs and residual signals for the ash accumulation curve of the economizer after EMD decomposition.

4.2. Accuracy Indicators

In order to measure the quality of the proposed models, we use the following evaluation indicators:

1. Root mean square error (*RMSE*):

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

2. Mean absolute percentage error (*MAPE*):

$$MAPE = \frac{100\%}{m}\sum_{i=1}^{m}\left|\frac{y_i - \hat{y}_i}{y_i}\right|; \tag{10}$$

3. Mean absolute error (*MAE*):

$$MAE = \frac{1}{m}\sum_{i=1}^{m}|y_i - \hat{y}_i| \tag{11}$$

where y_i and $\hat{y}_i$ stand for the true value and the predicted value of the cleanliness factor, respectively.

4.3. Performance Comparison of Multifarious Models

To emphasize the effectiveness of our proposed SVR + GPR model, the sole SVR model, the sole GPR model, the sole SVR + EMD model, and the sole GPR + EMD model are compared. In the first two models, the denoised cleanliness factor data is handled by using the sole SVR model and the sole GPR model, respectively. In the last two models, the SVRs and GPRs are utilized to deal with all signals acquired after EMD dissolution, separately.

Figures 7 and 8 and Table 1 illuminate the one-step-ahead predicted results and evaluation criteria through applying diverse models. Here, every model is trained by employing the preceding 350 data from the cleanliness factor degenerated curve of the economizer. Then, each trained model is verified with the rest of data. For the SVR and GPR models, choosing a proper kernel function is a key point. In the sole SVR, we use the radial basis function (RBF) as the kernel function. The RBF and rational quadratic (RQ) function are chosen for the kernel function of the sole GPR model. In the SVR + EMD model, the RBF is used as the kernel function for the IMFs, and the linear function is applied for the

residual component. We employ the RBF and RQ as the kernel function for the IMFs, and use the RBF for the residual signal in the GPR + EMD model. In our proposed SVR + GPR model, the linear function is selected as the kernel function for the SVR model, and the RBF and RQ functions are chosen as the kernel functions for the GPR model. Moreover, the number of inputs (i.e., the value of $i + 1$ in Figure 3) is also an important factor affecting the performance, particularly for the IMF1, which owns more fluctuations. Generally, increasing the number of input capacity terms will enhance the precision, but too large a number may result in an overfitting problem. In order to estimate the performance and restrain the model from overfitting, the experiments employing different input numbers are compared [35]. Hence, we select the input number as 10 to assure accuracy and prevent overfitting in all the models.

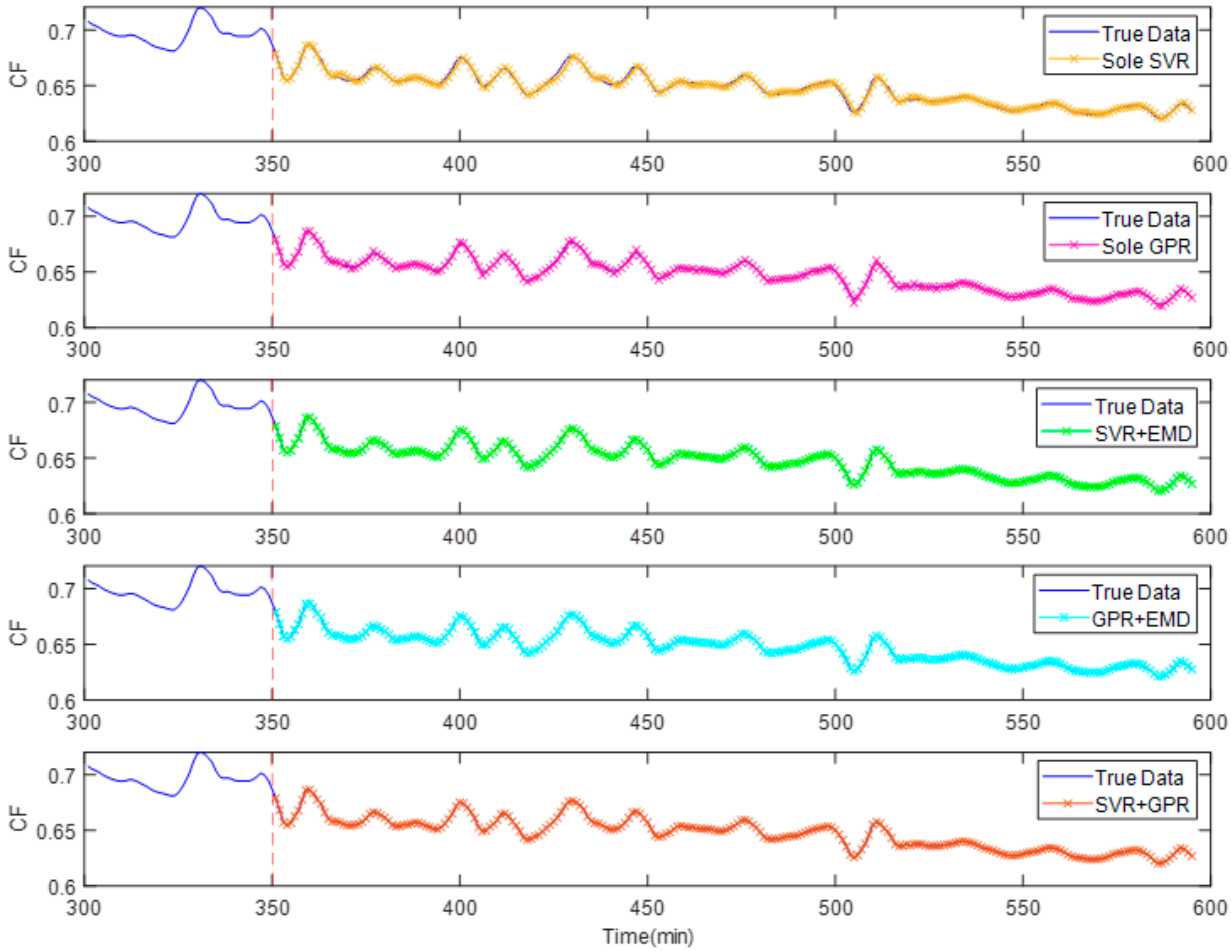


Figure 7. Prediction results of cleanliness factor for the economizer by using diverse models: sole SVR, sole GPR, SVR + EMD, GPR + EMD and SVR + GPR, from top to bottom.

From Figures 7 and 8, in the sole SVR model and sole GPR model, it is observed that the prediction curve is not smooth and several mismatches appear in the peaks of the curve, which may be due to the stochasticity and instability of the prediction. In comparison, after EMD dissolution, it is evident that all the predictions become better. Meanwhile, we find that there is a bit of a mismatch between the prediction results and the true data in the SVR + EMD model and GPR + EMD model. Finally, it is obvious that our proposed SVR + GPR method shows more efficient performance in predicting the cleanliness factor for the economizer. From Table 1, the MAPEs of the sole SVR model and sole GPR model are similar. The RMSE and MAE of the sole GPR model are better than the sole SVR model. Moreover, by using the EMD decomposition, we discover that all the accuracy indicators become better. The RMSE, MAPE and MAE of the hybrid SVR + GPR model are 0.00006,

0.00008 and 0.00005, which are the smallest compared with the other four models. In consequence, it can be concluded that our proposed SVR + GPR model presents the best the prediction of the cleanliness factor for the economizer.

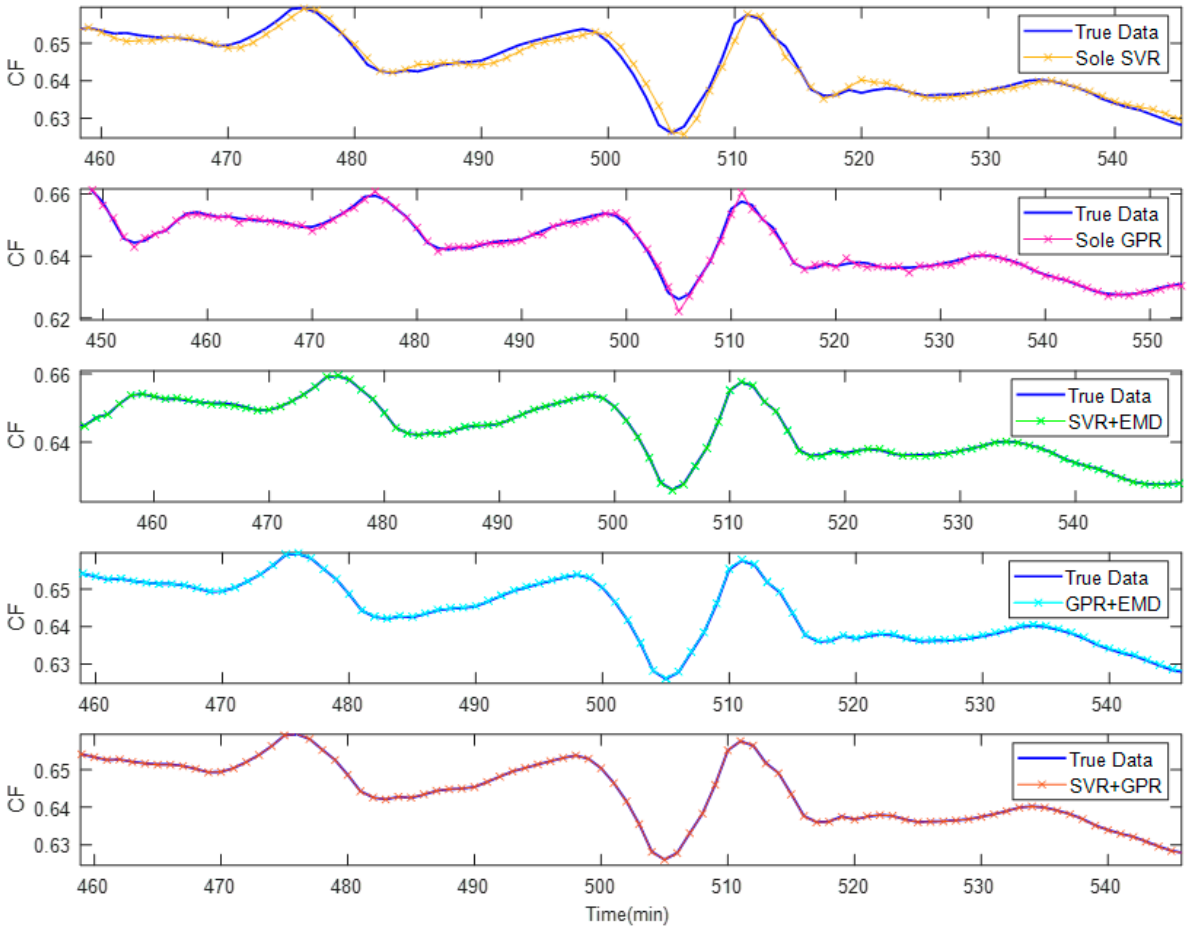


Figure 8. Prediction results of cleanliness factor for the economizer by using diverse models: sole SVR, sole GPR, SVR + EMD, GPR + EMD and SVR + GPR, from top to bottom (partially enlarged figure).

Table 1. Evaluation indicators using different models.

Methods	SVR	GPR	SVR + EMD	GPR + EMD	SVR + GPR
RMSE	0.00163	0.00092	0.00015	0.00020	0.00006
MAPE	0.02503	0.02537	0.00019	0.00022	0.00008
MAE	0.00117	0.00067	0.00012	0.00014	0.00005

4.4. Prediction Results

In order to examine the extrapolation performance of our proposed SVR + GPR model, both one-step- and multistep-ahead predictions are conducted. For the one-step test, the prediction results of the cleanliness factor for the reheater, the low-temperature superheater and the high-temperature superheater are plotted, and the evaluation indicators are also illustrated. For the multistep test, because of page limitations, the three-, six- and nine-steps-ahead predictions for the economizer and the reheater are carried out. In these tests, the input number is set as 10, and the prediction is implemented at the circle k-step ahead of the current circle. For the purpose of having enough information involved in the training process, the first 350 data points from the cleanliness factor degenerated curve of each device are applied as the training sets. Next, the well-trained model is employed to predict the future one-step or multistep cleanliness factor in the remaining circles without any

retraining procedure. In our iterative forecast, the new prediction result will be used for the next forecast, which can lead to cumulative uncertainty.

4.4.1. One-Step Prediction

Figure 9 illuminates the one-step-ahead cleanliness factor predicted results of the combined SVR + GPR model for the reheater, low-temperature superheater and high-temperature superheater. In Figure 9, it is visible that the trained model captures both the long-term dynamics and the local undulations of the ash accumulation curve for each device, as shown by the satisfying match between the prediction results and the true data for the rest of the points. In accordance with the evaluation indicators in Table 2, the RMSE, MAPE and MAE of the one-step case are all within 0.00006, 0.00009 and 0.00005, respectively, demonstrating the high precision of our proposed hybrid model.

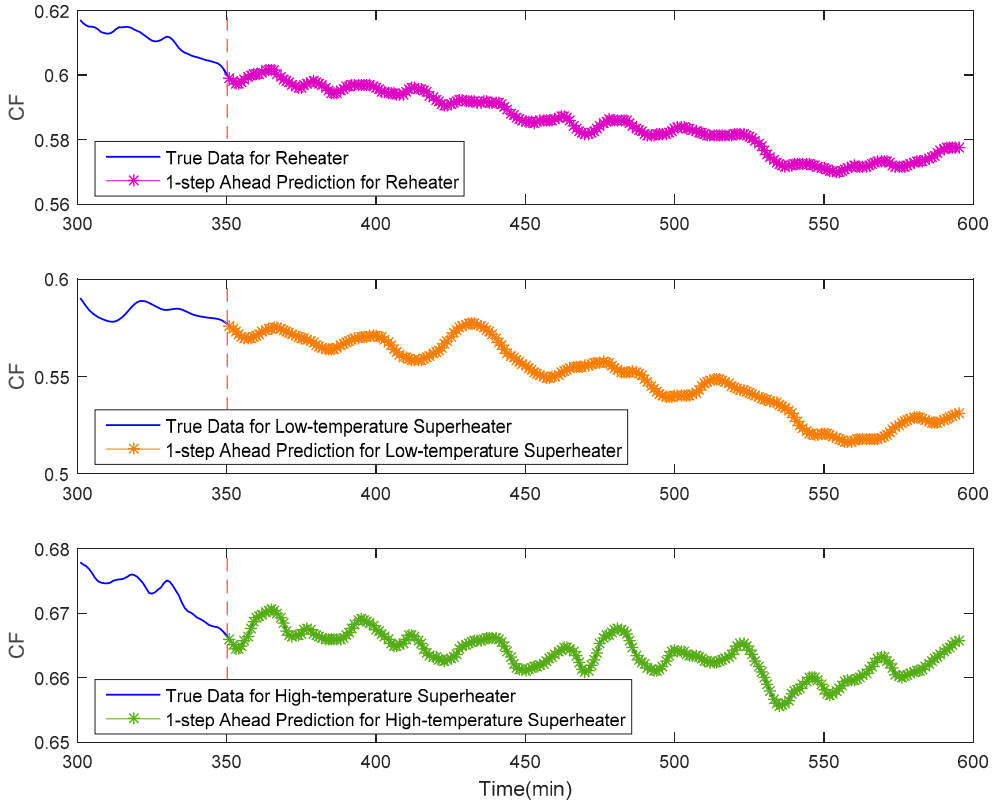


Figure 9. One-step-ahead prediction results of cleanliness factor for reheater, low-temperature superheater and high-temperature superheater.

Table 2. Accuracy indicators of one-step-ahead prediction.

Devices	Reheater	LTS	HTS
RMSE	0.00004	0.00006	0.00002
MAPE	0.00006	0.00009	0.00002
MAE	0.00003	0.00005	0.00001

4.4.2. Multistep Prediction

Figures 10 and 11 shows the k-step-ahead predicted values of cleanliness factor for the economizer and the reheater using the proposed model. It is evident that there are a few short-period mismatches in the multistep predictions, which is largely because of the shortage of a priori information. Nevertheless, the prediction results gradually rematch the actual data again as a result of the valid decoupling and the powerful long-time capture capacity of the hybrid SVR + GPR model. As the number of predicted steps increases,

there will be more mismatches, which is because the long-step forecast involves more indetermination. From Tables 3 and 4, the RMSE, MAPE and MAE for the economizer and the reheater become larger with the increase in the number of prediction steps. However, all these evaluation criteria are less than 0.003, which means that the forecast performance is satisfactory for the multistep cases. Hence, our proposed SVR + GPR model also has good quality of extrapolation for cleanliness factor multistep-ahead prediction.

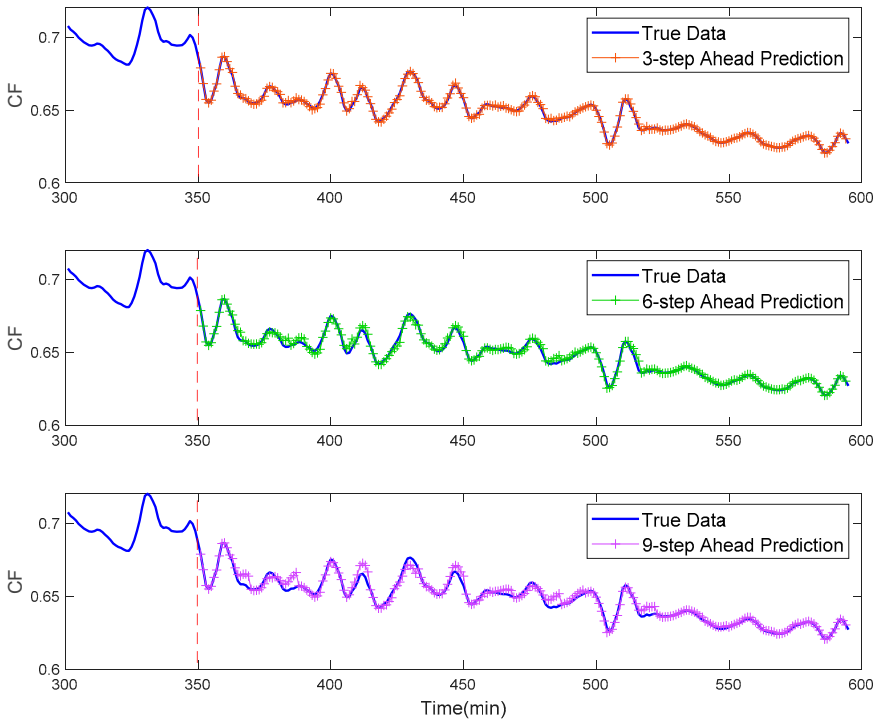


Figure 10. K-step-ahead prediction results of cleanliness factor for the economizer.

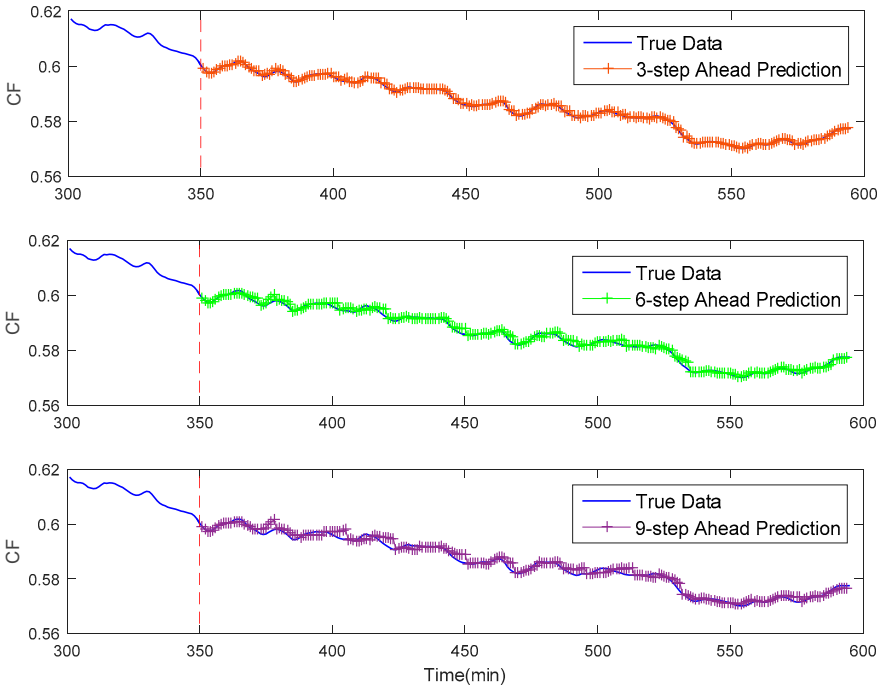


Figure 11. K-step-ahead prediction results of cleanliness factor for the reheater.

Table 3. Evaluation indicators of multistep-ahead prediction for the economizer.

Steps	Three-Step	Six-Step	Nine-Step
RMSE	0.00074	0.00174	0.00273
MAPE	0.00067	0.00167	0.00258
MAE	0.00044	0.00109	0.00169

Table 4. Evaluation indicators of multistep-ahead prediction for the reheater.

Steps	Three-Step	Six-Step	Nine-Step
RMSE	0.00043	0.00085	0.00123
MAPE	0.00048	0.00099	0.00146
MAE	0.00028	0.00058	0.00086

4.5. Discussions

This paper focuses on the exploitation of a hybrid data-driven method to achieve accurate future ash fouling prediction and good extrapolation properties. Due to the complex working conditions of the heat transfer surfaces in a boiler, it is challenging to model this physical process. The data-driven based prognostics method solves this by employing machine-learning techniques to link the cleanliness factor to the ash fouling accumulation state of the heat transfer surfaces. In reality, exploiting an adequately precise and robust data-driven model is an open research question. Despite the high quality of prediction and good extrapolating property obtained by our proposed method, the relevant performance greatly depends on the quality of data and test experiments, which is a general problem for applications of pure data-driven methods.

5. Conclusions

In this paper, an innovative hybrid method was presented to enable precise prediction of ash fouling for the heat transfer surfaces. This was the first known adhibition of integrating the SVR model and the GPR model for the future ash fouling prediction. By the way of specified result analyses and contrasts, some conclusions were received as follows:

- The long-term cleanliness factor degeneration was exactly captured by the SVR, while the local undulations were well expressed by means of GPR, further leading to advanced predicted results;
- Compared with the sole SVR, sole GPR, SVR + EMD and GPR + EMD models, our proposed SVR + GPR model is better than other models;
- In both the one-step- and multistep-ahead cleanliness factor forecasts, the combined SVR + GPR model obtained satisfactory extrapolation capability.

In future works, other machine learning techniques may be combined for future ash fouling prediction to achieve the benefits of both these methods, such as recurrent neural networks and convolutional neural networks. Furthermore, dynamical selection methods can be applied to choose the most appropriate combination model.

Author Contributions: F.C. and S.Q. are co-first authors of the paper. F.C., S.Q. and J.Z. conceived the theme; F.C. and S.Q. designed and performed the numerical simulations; F.C., S.Q., J.Z. and Y.S. analyzed the results and wrote the original manuscript; F.C., S.Q., M.L. and Y.S. did the work of data curation; F.C. and Y.S were responsible for funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China (No. 72071183), China University industry university research innovation fund project (No. 2019ITA0cxy0023) and Natural Science Foundation of Shanxi Province (No. 201901D211237, 20210302123025, 20210302124192).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Mangi, H.N.; Chi, R.; DeTian, Y.; Sindhu, L.; Lijin; He, D.; Ashraf, U.; Fu, H.; Zixuan, L.; Zhou, W.; et al. The ungrind and grinded effects on the pore geometry and adsorption mechanism of the coal particles. *J. Nat. Gas Sci. Eng.* **2022**, *100*, 104463. [CrossRef]
2. Clark, R.; Zucker, N.; Urpelainen, J. The Future of Coal-Fired Power Generation in Southeast Asia. *Renew. Sustain. Energy Rev.* **2019**, *121*, 109650. [CrossRef]
3. Ashraf, U.; Zhang, H.; Anees, A.; Mangi, H.N.; Ali, M.; Zhang, X.; Imraz, M.; Abbasi, S.S.; Abbas, A.; Ullah, Z.; et al. A Core Logging, Machine Learning and Geostatistical Modeling Interactive Approach for Subsurface Imaging of Lenticular Geobodies in a Clastic Depositional System, SE Pakistan. *Nat. Resour. Res.* **2021**, *30*, 2807–2830. [CrossRef]
4. Vo Thanh, H.; Yasin, Q.; Al-Mudhafar, W.J.; Lee, K.-K. Knowledge-based machine learning techniques for accurate prediction of CO_2 storage performance in underground saline aquifers. *Appl. Energy* **2022**, *314*, 118985. [CrossRef]
5. Zhang, W.H. Suggestions and Measures for Promoting the Energy Conservation and Emission Reduction of Coal-fired Industrial Boilers. *Energy Conserv.* **2014**, *9*, 67–69.
6. Zhou, K.; Yang, S.; Shen, C.; Ding, S.; Sun, C. Energy conservation and emission reduction of China's electric power industry. *Renew. Sustain. Energy Rev.* **2015**, *45*, 10–19. [CrossRef]
7. Yan, Q.; Wang, Y.; Baležentis, T.; Sun, Y.; Streimikiene, D. Energy-Related CO_2 Emission in China's Provincial Thermal Electricity Generation: Driving Factors and Possibilities for Abatement. *Energies* **2018**, *11*, 1096. [CrossRef]
8. Li, X. Design of energy-conservation and emission-reduction plans of China's industry: Evidence from three typical industries. *Energy* **2020**, *209*, 118358. [CrossRef]
9. Harding, N.S.; O'Connor, D.C. Ash deposition impacts in the power industry. *Fuel Process. Technol.* **2007**, *88*, 1082–1093. [CrossRef]
10. Shao, Y.; Wang, J.; Preto, F.; Zhu, J.; Xu, C. Ash Deposition in Biomass Combustion or Co-Firing for Power/Heat Generation. *Energies* **2012**, *5*, 5171. [CrossRef]
11. Valero, A.; Cortés, C. Ash fouling in coal-fired utility boilers. Monitoring and optimization of on-load cleaning. *Prog. Energy Combust. Sci.* **1996**, *22*, 189–200. [CrossRef]
12. Shi, Y.; Li, Q.; Wen, J.; Cui, F.; Pang, X.; Jia, J.; Zeng, J.; Wang, J. Soot Blowing Optimization for Frequency in Economizers to Improve Boiler Performance in Coal-Fired Power Plant. *Energies* **2019**, *12*, 2901. [CrossRef]
13. Zierold, K.M.; Odoh, C. A review on fly ash from coal-fired power plants: Chemical composition, regulations, and health evidence. *Rev. Environ. Health* **2020**, *35*, 401–418. [CrossRef] [PubMed]
14. Shi, Y.; Wen, J.; Cui, F.; Wang, J. An Optimization Study on Soot-Blowing of Air Preheaters in Coal-Fired Power Plant Boilers. *Energies* **2019**, *12*, 958. [CrossRef]
15. Peña, B.; Teruel, E.; Díez, L.I. Soft-computing models for soot-blowing optimization in coal-fired utility boilers. *Appl. Soft Comput.* **2011**, *11*, 1657–1668. [CrossRef]
16. Wen, J.; Shi, Y.; Pang, X.; Jia, J.; Zeng, J. Optimization of Boiler Soot Blowing Based on Hamilton-Jacobi-Bellman Equation. *IEEE Access* **2019**, *7*, 20850–20862. [CrossRef]
17. Li, Q.; Yan, P.; Liu, J.; Shi, Y.; Yu, D. Prediction of Pollution State of Heating Surface in Coal-Fired Utility Boilers. *IEEE Access* **2020**, *8*, 206132–206145. [CrossRef]
18. Chen, B.K.; Yan, W.P.; Zhu, Y.D.; Gao, Z.Y.; Liang, X.J. Prediction of ash fouling on heat transfer surface of coal fired utility boiler. *J. North China Electr. Power Univ.* **2004**, *31*, 32–35.
19. Zheng, S.; Zeng, X.; Qi, C.; Zhou, H. Mathematical modeling and experimental validation of ash deposition in a pulverized-coal boiler. *Appl. Therm. Eng. Des. Processes Equip. Econ.* **2017**, *110*, 720–729. [CrossRef]
20. Tang, S.Z.; Wang, F.L.; Ren, Q.; He, Y.L. Fouling characteristics analysis and morphology prediction of heat exchangers with a particulate fouling model considering deposition and removal mechanisms. *Fuel* **2017**, *203*, 725–738. [CrossRef]
21. Tong, S.; Zhang, X.; Tong, Z.; Wu, Y.; Tang, N.; Zhong, W. Online Ash Fouling Prediction for Boiler Heating Surfaces based on Wavelet Analysis and Support Vector Regression. *Energies* **2019**, *13*, 59. [CrossRef]
22. Shi, W.; Wang, J.; Shi, Y.; Liu, Z. Clustering LS-SVM models for the prediction of unburned carbon content in fly ash. In Proceedings of the 27th Chinese Control and Decision Conference (2015 CCDC), Qingdao, China, 23–25 May 2015; pp. 19–24.
23. Shi, Y.; Li, M.; Cui, F.; Wen, J.; Zeng, J. A Neural-Network-Based Method for Ash Fouling Prediction of Heat Transfer Surface in Coal-Fired Power Plant Boiler. *IEEE Access* **2021**, *9*, 109584–109604. [CrossRef]
24. Peña, B.; Teruel, E.; Díez, L.I. Towards soot-blowing optimization in superheaters. *Appl. Therm. Eng.* **2013**, *61*, 737–746. [CrossRef]
25. Tang, S.Z.; Li, M.J.; Wang, F.L.; He, Y.L.; Tao, W.Q. Fouling potential prediction and multi-objective optimization of a flue gas heat exchanger using neural networks and genetic algorithms. *Int. J. Heat Mass Transf.* **2020**, *152*, 119488. [CrossRef]
26. Han, J.; van der Baan, M. Empirical mode decomposition for seismic time-frequency analysis. *Geophysics* **2013**, *78*, O9–O19. [CrossRef]
27. Shi, Y.; Wang, J.; Liu, Z. On-line monitoring of ash fouling and soot-blowing optimization for convective heat exchanger in coal-fired power plant boiler. *Appl. Therm. Eng.* **2015**, *78*, 39–50. [CrossRef]

28. Li, Y.; Liu, K.; Foley, A.M.; Zülke, A.; Berecibar, M.; Nanini-Maury, E.; Van Mierlo, J.; Hoster, H.E. Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review. *Renew. Sustain. Energy Rev.* **2019**, *113*, 109254. [CrossRef]
29. Murphy, K.P. *Machine Learning: A Probabilistic Perspective*; MIT Press: Cambridge, MA, USA, 2012.
30. Pao-Tsun, L.; Shun-Feng, S.; Tsu-Tian, L. Support vector regression performance analysis and systematic parameter selection. In Proceedings of the 2005 IEEE International Joint Conference on Neural Networks, Montreal, QC, Canada, 31 July–4 August 2005; Volume 872, pp. 877–882.
31. Carl Edward, R.; Christopher, K.I.W. Subject Index. In *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2005; pp. 244–247.
32. Samo, Y.-L.K.; Roberts, S.J. p-Markov Gaussian Processes for Scalable and Expressive Online Bayesian Nonparametric Time Series Forecasting. *arXiv* **2015**, arXiv:1510.02830.
33. Richardson, R.R.; Osborne, M.A.; Howey, D.A. Gaussian process regression for forecasting battery state of health. *J. Power Sources* **2017**, *357*, 209–219. [CrossRef]
34. Datong, L.; Jingyue, P.; Jianbao, Z.; Yu, P. Data-driven prognostics for lithium-ion battery based on Gaussian Process Regression. In Proceedings of the IEEE 2012 Prognostics and System Health Management Conference (PHM-2012 Beijing), Beijing, China, 23–25 May 2012; pp. 1–5.
35. Liu, K.; Shang, Y.; Ouyang, Q.; Widanage, W.D. A Data-Driven Approach With Uncertainty Quantification for Predicting Future Capacities and Remaining Useful Life of Lithium-ion Battery. *IEEE Trans. Ind. Electron.* **2021**, *68*, 3170–3180. [CrossRef]