Research Article

An Adaptive Control Combination Forecasting Method for Time Series Data

Hongyan Jiang,1,2 Dianjun Fang,1,3 and Xinyan Zhang1

1School of Mechanical Engineering, Tongji University, Shanghai 201804, China
2School of Mechatronic and Power Engineering, Jiangsu University of Science and Technology, Zhenjiang 212003, China
3Sino-German Institute of Intelligent Technologies, Qingdao 266000, China

Correspondence should be addressed to Dianjun Fang: 1733316@tongji.edu.cn

Received 19 January 2021; Revised 6 May 2021; Accepted 19 May 2021; Published 3 June 2021

Academic Editor: Samuel N. Jator

Copyright © 2021 Hongyan Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

According to the individual forecasting methods, an adaptive control combination forecasting (ACCF) method with adaptive weighting coefficients was proposed for short-term prediction of the time series data. The US population dataset, the American electric power dataset, and the vibration signal dataset in a hydraulic test rig were separately tested by using ACCF method, and then, the accuracy analysis of ACCF method was carried out in the study. The results showed that, in contrast to individual methods or combination methods, the proposed ACCF method was adaptive to adopt one or some of prediction methods and showed satisfactory forecasting results due to flexible adaptability and a high accuracy. It was also concluded that the higher the noise ratio of the tested datasets, the lower the prediction accuracy of the ACCF method; the ACCF method demonstrated a better prediction trend with good volatility and following quality under noisy data, as compared with other methods.

1. Introduction

A time series is a set of statistics and usually collected at regular intervals. Time series data occur naturally in many application areas, such as economics, medicine, weather data, ocean engineering, finance, and engineering control. Time series data are obtained by the sensors, and they refer to the large, diverse datasets of information that cannot be easily processed by using standard computers. Based on the past performance, time series forecasting is an analysis used to forecast future value, which is still a challenging research topic nowadays [1–3].

The three common methods for time series forecasting included physical, statistical, and artificial intelligence. In physical methods, effective forecasting results rely on physical information [4, 5], but it was not proficient in dealing with short-term series with complex calculation process. Statistical models, including Period-Sequential Index (PSI) [6], moving average (MA) [7], autoregressive integrated moving average (ARIMA) [8], exponential smoothing [9], Kalman filter [10], and grey forecasting [11], effectively tackled linear features but gave larger error for a fluctuant, seasonal one [12], noise, or instability [13]. Artificial intelligence models, subsuming BP neural network (BP-NET) [14, 15], support vector machines (SVM) [16], fuzzy logic models [17], and least square support vector machine (LSSVM) [18], have exhibited significant advantages in dealing with nonlinear problems. These artificial intelligence models offered higher forecasting accuracy than physical or statistical models, but their prediction was mostly relying on training datasets, and they are easy to get stuck or suffer from overfitting in the local optima [19, 20].

Because of the inherent disadvantages of each model, nowadays, the effective information of multiple models has been used to predict time series, and weight problem of combination model is becoming the research focus. Weights could be allocated to the various forecasts produced by individual models, so as to achieve a combined forecast [21, 22]. For example, Clark et al. [23] derived mean square error-minimizing weights for combining the restricted and unrestricted forecasts and assigned more weights on the
restricted model and less weights on the unrestricted model. However, Clark’s method supported the conventional wisdom that simple averages were hard to beat and applied only to averaging to nested models. Hao et al. [24] introduced entropy weight method into the combination prediction model. Xie et al. [25] combined linear regression prediction model and grey model to get a various weight combination model. The two methods have better forecasting accuracy in only Dam’s settlement or landslip, but weights were assigned to all participating single models, and forecast accuracy in other time series cases was unwarrantable or unknown. Gao et al. [26] established five kinds of combination forecasting models, including suboptimal weight, optimal weight, grey comprehensive correlation degree weight, entropy weight, and neural network. In Gao’s study, the weighted constraint criterion was given particular attention; nevertheless, negative weight of single item prediction model might occur in the combination model. Song and Fu [27] have also found out that some models failed to combine the advantages of single models, such as the combination of the autoregressive integrated moving average model and neural network model, or the combination of neural network and other forecasting models.

In my opinion, the main problem of the above combination methods is that the statistical distribution information of the forecasting errors with the historical time is not paid more attention or is ignored, leading to unreasonable weight distribution and even negative weights. Therefore, existing combined forecast models were still lacking the predicted reliability, particularly under a condition of noise. In the study, weighting coefficient for each model was adaptively determined based on their own statistical forecasting performance for historical data. The rest of the paper is organized as follows. Section 2 contains the methodology of combination forecasting method. Section 3 contains the steps of computation. Section 4 contains the results and discussion of short-term prediction cases. Section 5 contains the conclusions.

2. Methodology

2.1. The Forecasting Methods

2.1.1. Individual Methods

(1) Period-Sequential Index Method

The Period-Sequential Index (PSI) method [6] had been studied by the author in the previous work. The PSI model introduced the period index (PI) and sequential index (SI) to describe the dataset structure information in vertical and horizontal dimensions, respectively. Figure 1 shows the schematic diagram of the PSI algorithm. \( H_{2\cdot1} \), \( H_{1\cdot1} \) denote reference historical periods, i.e., the year before last year and last year. \( H_0 \) represents the forecasting period. The period for \( H_{1\cdot1} \) and \( H_0 \) is uniform, defined as \( T \). When time \( t_i \) is upcoming, the forecasting value at time of \( t_i \) based on PSI method is mainly dependent on the PI and the SI. The PSI method is described as follows:

\[
Y_1(t_i) = aPI(t_i) \cdot K_0 + (1 - a)SI(t_{i-1}) \cdot y(t_{i-1}),
\]

where \( t_i \) (independent variable) is a forecasting time, \( Y_1(t_i) \) (dependent variable) is a forecasting value at time of \( t_i \), \( K_0 \) is the reference coefficient for period index, \( y(t_{i-1}) \) is the observed value at historical time of \( t_{i-1} \), \( a \) is the optimized weighing factor of PSI method, \( PI(t_i) \) is the forecasting period index (dependent variable) at time of \( t_i \) and \( SI(t_{i-1}) \) is the forecasting sequential index (dependent variable) at time of \( t_{i-1} \).

\[
\begin{align*}
PI(t_i) &= 0.5 \cdot \frac{y(t_i - 2T)}{K_{-2}} + 0.5 \cdot \frac{y(t_i - T)}{K_{-1}} \\
SI(t_{i-1}) &= 0.5 \cdot \frac{y(t_{i-2} - 2T)}{y(t_{i-1-2} - 2T)} + 0.5 \cdot \frac{y(t_{i-1} - T)}{y(t_{i-1-1} - T)},
\end{align*}
\]

where \( y(t_{i-2} - 2T) \) and \( y(t_i - T) \) describe the reference historical data at time of \( t_{i-2} - 2T \) and \( t_i - T \), respectively. \( K_{-2} \) and \( K_{-1} \) are reference functions of period index. A standard period average is originally set to be a reference function of period index, where it is defined as a constant. The more detailed derivation can be found in Ref. [6].

(2) Exponential Smoothing Method

The Exponential Smoothing (ES) method [9] is often used in practice to forecast time series. Suppose that the observed values for time series are \( y(t_1), y(t_2), \ldots, y(t_{i-1}) \) at time of \( t_1, t_2, \ldots, t_{i-1} \), respectively. For ES method, the forecast value at time of \( t_i \) is dependent on the observed value at time of \( t_{i-1} \) and the forecasting value at time of \( t_{i-1} \). The ES method is defined as

\[
Y_2(t_i) = \beta y(t_{i-1}) + (1 - \beta)Y_2(t_{i-1}),
\]

where \( Y_2(t_1) \) and \( Y_2(t_{i-1}) \) represent the forecasting values at time of \( t_1 \) and \( t_{i-1} \) by using ES method, \( y(t_{i-1}) \) is the observed value at historical time of \( t_{i-1} \), and \( \beta \) is the smoothing parameter, which can be adjusted between 0 and 1. Higher \( \beta \) will produce a forecast, which is more responsive to recent changes in the data, whilst also being less robust to any errors that could occur.

(3) Moving Average Method

The moving average (MA) method [7] is simple and widely used, and it performs well in forecasting competitions against more sophisticated approaches. A simple Moving Average is a common average of the previous \( n \) data points in time series data, and each point in the time series data is equally weighted. The MA method can be described as follows:

\[
Y_3(t_i) = \frac{y(t_{i-1}) + y(t_{i-2}) + \ldots + y(t_{i-n})}{n}
\]

(4)
where $Y_d(t_i)$ represents the forecasting values at time of $t_i$ by using MA method, $n$ is the number of data points used in the calculation, and $y(t_{i-n})$ is the observed data point value at time of $t_{i-n}$.

(4) Autoregressive Integrated Moving Average Method

Autoregressive integrated moving average (ARIMA) [8] is one of the most popular statistical linear models for forecasting time series data. It is a combination of autoregression AR($p$) (an additive linear function of $p$ past observations), moving average MA($q$) ($q$ random errors), and $d$ which is an integer making a series to be stationary. The general form of the forecast equation for ARIMA ($p$, $q$, $d$) model can be written as follows:

$$Y_4(t_i) = c + \sum_{j=1}^{p} \phi_j y(t_{i-j}) + \sum_{k=1}^{q} \theta_k \epsilon(t_{i-k}) + \epsilon(t_i), \quad (5)$$

where $Y_d(t_i)$ is a forecasting value at time of $t_i$ by using ARIMA method, $c$ is the constant representing the intercept, $\phi_j$ and $y(t_{i-j})$ are the parameters and regressors for AR part of the model, respectively; $\theta_k$ and $\epsilon(t_{i-k})$ are the parameters and regressors of the MA part of the model, respectively; and $\epsilon(t_i)$ is the white noise at time of $t_i$.

(5) BP Neural Network Method

BP neural network (BP-NET) method [14, 15] can realize self-learning and memory functions of machine. Figure 2 gives the simple structure of BP-NET. It can be seen that BP-NET is composed of input layer, output layer, and hidden layer. Using this three-layer structure, BP can simulate any complex nonlinear relationship through nonlinear elements. The basic calculation principle is divided into three steps: forward calculation (calculate the output of each node in turn based on the input); error backpropagation (calculate the gradient of each node according to the loss function); weight update. Because this method has excellent data processing and relationship building capabilities, it has been widely used in forecasting.

(6) Grey Forecasting Method

Grey forecasting method (GM) [11] indicates one variable and one-order grey forecasting model. This grey differential equation is formed by an original time series $y(t_i)$ using accumulated generating operation (AGO) technique. It is denoted as follows:

$$Y_6(t_i) = \left[ y(t_i) - \left( \frac{b}{a} \right) e^{-a(i-1)}(1 - e^a) \right], \quad (6)$$

where $Y_d(t_i)$ is a forecasting value at time of $t_i$ by using GM method, $a$ is a developing coefficient, and $b$ is a control variable. $a$ and $b$ are denoted as

$$[a \ b]^T \left( B^T B \right)^{-1} B^T C,$$

where

$$B = \begin{bmatrix} -\frac{1}{2} \left( \sum_{k=1}^{2} y(t_k) + y(t_i) \right) & -\frac{1}{2} \left( \sum_{k=1}^{2} y(t_k) + \sum_{k=1}^{3} y(t_k) \right) & \cdots & -\frac{1}{2} \left( \sum_{k=1}^{i-1} y(t_k) + \sum_{k=1}^{i-2} y(t_k) \right) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & 1 \end{bmatrix},$$

$$C = [ y(t_3) \ y(t_4) \ \cdots \ y(t_{i-1})]^T.$$
2.1.2. Combination Forecasting Methods. Because it is too risky to rely on the forecasts produced by an individual method, the combination forecasting method was widely used in the study. There is one historical piece of data \( y(t_1), y(t_2), \ldots, y(t_m) \) occurring at the corresponding time \( t_1, t_2, \ldots, t_m \). The forecasting value of the \( j \)-th model (\( j = 1, 2, \ldots, n \)) and the corresponding weight coefficient are \( \lambda_j(t_0) \) and \( \omega_j \) respectively, at time \( t_0 \). The linear combination is generally calculated according to [28]

\[
F(t_i) = \sum_{j=1}^{n} w_j Y_j(t_i) \quad \text{and} \quad \sum_{j=1}^{n} w_j = 1.
\]

(9)

Considering the actual situation and the calculation complexity, three common weight methods [27], including inverse variance (IV) method, mean square error inverse (MSEI) method, and simple weighted average (SWA) method, were used to compute weight coefficients in the study. The methods are based on the sum of squared errors.

The weight coefficients of the IV method were computed in equation (10). In a case of the larger sum of squared errors in a single method, this method is assigned a smaller weight. On the contrary, a larger weight is assigned to the smaller squared errors in a single model.

\[
\omega_j = \frac{e_j^{-1}}{\sum_{j=1}^{n} e_j^{-1}},
\]

(10)

where \( n \) indicates the number of single models, and \( e_j \) is the \( j \)-th single model.

\[
e_j = \sum_{k=1}^{j-1} (y(t_k) - Y_j(t_k))^2 \quad \text{and} \quad j = 1, 2, \ldots, n.
\]

(11)

Similarly, the weight coefficients of the MISEI method were computed as follows:

\[
\omega_j = \frac{e_j^{-0.5}}{\sum_{j=1}^{n} e_j^{-0.5}}.
\]

(12)

For the SWA method, the sum of squared errors of each model was ranked by a descending ranking order. Then, a new array of \( j_1, j_2, \ldots, j_n \) could be defined and represent a ranking order of single model. This meant that one individual model with a higher value of \( j \) would have a lower forecasting error. The weight coefficients of the SWA method were further given in equation (13). It is seen that the weight coefficient is larger in a condition of higher \( j \) in order to minimize the sum of squared errors.

2.2. The Adaptive Control Combination Forecasting Method. In the following work, one total tested data set occurred at time \( t_1, t_2, \ldots, t_{N-1}, t_N \), and is defined as

\[
y(t_1), y(t_2), \ldots, y(t_{t-1}), y(t_t), (t_{t+1}), \ldots, y(t_{N-1}), y(t_N).
\]

This total tested data set is a time series based on an equal interval of time \( (dt) \) and a fixed time period \( (T) \). For example, while \( dt \) is month, day, hour, minute, or second, the corresponding \( T \) of the tested data are year, month, day, and minute, respectively. Now, the time \( t \) (\( 3T < t < t_s \)) is assumed to be upcoming, so the tested data set during three time periods \( (3T) \) before \( t \) and \( y(t_t) \) is extracted from the total tested data set. It is defined that the extracted tested data set contains \( s + 1 (s = 3T/dt) \) data in number and is given in detail as follows: \( y(t_{t-s}), y(t_{t-s-1}), y(t_{t-s-2}), \ldots, y(t_{t-2}), y(t_{t-1}), y(t_t) \). Using this extracted tested data set, each individual method can give forecasts at time of \( t_t \) based on that mentioned in Section 2.1.1 and get the modeling data set: \( Y_1(t_t), Y_2(t_t), \ldots, Y_{n-1}(t_t), Y_n(t_t) \).

Similarly, as the forecasting time \( t_t \) moves back to \( t_0, t_{t-1}, t_{t-2}, \ldots, t_{t-s-2}, t_{t-s-1}, \) and \( t_{t-s} \) in sequence, the all extracted tested datasets can be generated, and they are listed in Table 1.

Using each individual method, corresponding forecasts at time of \( t_0, t_{t-1}, t_{t-2}, \ldots, t_{t-s-1}, \) and \( t_{t-s} \) are defined as the first modeling datasets, which are listed in Table 2.

In fact, the first modeling datasets are \( n \) forecasting datasets corresponding to every individual forecasting model, respectively. Then, the mean absolute percentage errors (MAPE) [29, 30] are computed to obtain the datasets \( \text{MAPE}_1, \text{MAPE}_2, \ldots, \) and \( \text{MAPE}_n \), given in equation (14). The mean values and the standard deviations for \( \text{MAPE}_1, \text{MAPE}_2, \ldots, \) and \( \text{MAPE}_n \) can be further given in equation (14). They are defined as the second modeling datasets, which are listed in Table 3.

![Figure 2: The simple structure of BP-NET.](image-url)
Based on the previous preparation work of equation (14), our present work is to produce a forecast $F(t_i)$ for the upcoming time $t_i$ ($3s < i \leq N$). Consequently, an adaptive control combination forecasting (ACCF) method is given in equation (14).

$$F(t_i) = w_1(t_i) \cdot Y_1(t_i) + w_2(t_i) \cdot Y_2(t_i) + \ldots + w_{n-1}(t_i) \cdot Y_{n-1}(t_i) + w_n(t_i) \cdot Y_n(t_i), \quad (15)$$

where $w_1, w_2, \ldots, w_{n-1}$, and $w_n$ are the weighting coefficients, which are dependent variables with the change of the upcoming time $t_i$. Based on the performance-based approach [31], an adaptive weight for each model is determined based on their own forecasting performance and can be defined as follows:
deviation are defined as $\text{MAPE}_m$ and $\sigma_m$, respectively. Thus, the smallest MAPE range is defined between $\text{MAPE}_m - 3\sigma_m$ and $\text{MAPE}_m + 3\sigma_m$. When statistical arrays range from $\text{MAPE}_j - 3\sigma_j$ to $\text{MAPE}_j + 3\sigma_j$ produced by models that overlap partially or fully with the smallest MAPE range (from $\text{MAPE}_m - 3\sigma_m$ to $\text{MAPE}_m + 3\sigma_m$), where $k_j$ is 1; otherwise, $k_j$ is 0.

In the ACCF model, the modeling datasets, including $\text{MAPE}_j$, $\sigma_j$, $w_j$; and $k_j$ ($j = 1, 2, \ldots, n$) must be corrected on every upcoming time $t_i$ according to the solution of equations (14), (16), and (17). Finally, key parameters $w_j$ are finally taken into equation (15) to make a prediction on time $t_i$. Because $w_{ij}$ in equation (16) must be updated and revised on every upcoming time $t_i$ so as to ensure prediction accuracy with high robustness, the new developed model is propitious to short-term prediction. In the condition of long-term prediction, however, this new model cannot update the value of the weights in time and may be result in a large percentage forecast error.

2.3. Evaluation Index. The MAPE during time period $(T)$ is further measured to evaluate the obtained results. Because there are $s/3$ data in number every time period, the error $\text{MAPE}_k (k = 1, 2, \ldots, (N-3s)/(s/3))$ during the $k_{th}$ time period can be calculated by equation (17). For total tested data set, the number of predicting time periods is equal to $(N-3s)/(s/3)$, and calculation equations of total error indicators $\text{MAPE}_{\text{all}}$ are given in the below formula (18). To verify the superiority of the ACCF approach, statistical count lower than $\text{MAPE}_{\text{all}}$ (defined Count) and percentage of Count in the sum (defined Per) are also calculated by equation (18) for this new array $\text{MAPE}_k$.

\[
\text{MAPE}_k = \frac{1}{(s/3)} \sum_{i=3s+1}^{3s+k_{th}} \left( \frac{y(t_i) - F(t_i)}{y(t_i)} \right) \cdot 100\%,
\]

\[
\text{MAPE}_{\text{all}} = \frac{1}{(N-3s)/(s/3)} \sum_{k=1}^{(N-3s)/(s/3)} \text{MAPE}_k,
\]

\[
\text{Count} = \text{count}\{\text{MAPE}_k < \text{MAPE}_{\text{all}}\},
\]

\[
\text{Per} = \frac{\text{Count}}{(N-3s)/(s/3)} \cdot 100\%.
\]
where \( y(t_i) \) is the measured value at time of \( t_i \); \( F(t_i) \) is the predicting value at time of \( t_i \).

After obtaining the predicting error of \( \text{MAPE}_{\text{all}} \) the forecasting accuracy (\( FA_{\text{all}} \)) can be calculated by using equation (19). Thus, the forecasting accuracy of the model will be better when the \( FA \) is approaching 100%.

\[
FA_{\text{all}} = 100\% - \text{MAPE}_{\text{all}}.
\] (19)

### 3. Steps of Computation

In the study, six individual forecasting methods, including PSI, ES, MA, ARIMA, GM, and BP-NET, are used to construct ACCF method described in Section 2.2. The flow chart of the proposed ACCF model for short-term prediction is summarized in Figure 3. It gives a rolling forecast process, and the detailed steps are as follows:

Step 1: The total tested data set \((y(t_1), y(t_2), \ldots, y(t_{n-1}))\) is initialized at time of \( t_{n-1} \).

Step 2: The all tested datasets in Table 1 are extracted from total tested data set.

Step 3: By solving equations (1) and (3)–(6), ..., the first modeling datasets in Table 2 are predicted by using each individual model, such as PSI method, ES method, ARIMA method, MA method, GM method, and BP-NET method.

Step 4: Equation (14) is solved, and the second modeling datasets in Table 3 are got;

Step 5: Equations (16) and (17) are solved, and the third modeling datasets in Table 4 are further obtained;

Step 6: Using \( Y_1(t_i), Y_2(t_i), \ldots, Y_6(t_i) \) (in Table 2) and \( w_1, w_2, \ldots, w_6 \) (in Table 4), equation (15) is calculated to give a new combination forecast \( F(t_i) \) at time step of \( t_i \).

Step 7: If the time steps of the stop condition \((t_i \geq t_N)\) are satisfied, the search stops, as well as output parameters of \( \text{MAPE}_{\text{all}}, \text{Count}, \text{Per}, \) and \( FA_{\text{all}} \) by solving equations (18) and (19); Otherwise, the time step is added, new generated data \( y(t_i) \) is added into the total tested data set, and then the procedure returns to step 2.

### 4. Results and Discussion

Table 5 shows three groups of tested time series datasets. As shown in Table 5, the first dataset (named USP), with total samples of 492, is US population between January 1979 and December 2019, which is from the US Census Bureau hosted by the Federal Reserve Economic Database (FRED) [32]. FRED has a data platform found US population data and updated population information of every month. The second dataset (named AEP), with total samples of 3432, is the American hourly electric power consumption data between March 13, 0:00, and August 2, 23:00, in 2018, which comes from PJM’s website and is in megawatts (MW) [33]. PJM is a regional transmission organization (RTO) in the United States, and part of the Eastern Interconnection grid operating an electric transmission system. The third dataset (named VS) was experimentally obtained with a hydraulic test rig. This test rig consists of a primary working, a test system, and a secondary cooling-filtration circuit, which are connected via the oil tank [34, 35]. The system cyclically repeats constant load cycles (duration 60 seconds). The test system is equipped with several sensors measuring process values such as vibration, with standard industrial 20 mA current loop interfaces connected to a data acquisition system. In the study, these vibration signals in hydraulic test rig with a sampling frequency of 1 Hz during 8580 seconds [34] were measured and used as the third dataset.

By using the developed ACCF method, a comparison analysis between the real value and the forecasting value was implemented and showed a direct observation of the prediction, so as to evaluate the confidence of the ACCF method.

#### 4.1. Periodic Recognition and Prediction on USP Dataset

The USP datasets were used as a tested dataset to show periodic detection and prediction results. In order to calculate second modeling datasets \( \text{MAPE}_1, \text{MAPE}_2, \text{MAPE}_3, \text{MAPE}_4, \text{MAPE}_5, \text{MAPE}_6 \) produced by six individual models, respectively in equation (14), monthly population data of USP during contiguous 36 months are in turn trained to forecast the population data during next month. Figure 4 shows evolution of the mean values of MAPE with year when predicting by using the PSI, ES, ARIMA, MA, BP-NET, and GM methods. It can be seen that the ARIMA method gives a much better prediction accuracy.

Based on the ACCF method, the weighting coefficients can be given by solving equations (16) and (17): \( w_1 = 0 \) for PSI, \( w_2 = 0 \) for ES, \( w_3 = 1 \) for ARIMA, \( w_4 = 0 \) for MA, \( w_5 = 0 \) for BP-NET and \( w_6 = 0 \) for GM. It also shows that the ACCF method can adaptively seek the prediction methods with much higher accuracy and abandon other prediction methods with poor accuracy. Then, the USP dataset from January 1988 to December 2019 is circularly predicted by solving equation (15), as shown in Figure 5. It demonstrates a good prediction trend with good volatility and following quality by using the ACCF method.

#### 4.2. Periodic Recognition and Prediction on AEP Dataset

The AEP datasets were used as a tested data set to show periodic detection and prediction results. Figure 6 shows evolution of the mean values of MAPE with time when predicting by using the PSI, ES, ARIMA, MA, BP-NET, and GM methods. By solving equations (16) and (17), the weighting coefficients of models at different forecasting times are shown in Table 6. It can be seen that the sum of six weight coefficients is always equal to 1; \( w_4 \) and \( w_5 \) are equal to zero; the values of \( w_2 \) are always less than those of \( w_1 \) and \( w_3 \). Thus, because of the worst performance of MA and GM method, their weights are not assigned. But PSI method and ARIMA method play more important role than other individual methods in the prediction of AEP dataset. Then, the AEP dataset from
March 22, 0:00, to August 2, 23:00, in 2018 performed a rolling prediction by solving equation (15), as shown in Figure 7. It demonstrates that the ACCF method shows a better prediction trend with good volatility and following quality.

### 4.3. Periodic Recognition and Prediction on VS Dataset

The VS datasets were used as a tested dataset to show periodic detection and prediction results. Figure 8 shows evolution of the mean values of MAPE with time when predicting by using the PSI, ES, ARIMA, MA, BP-NET, and
The mean values of MAPE (%)

**MAPE\textsubscript{1}** for PSI

**MAPE\textsubscript{2}** for ES

**MAPE\textsubscript{3}** for ARIMA

**MAPE\textsubscript{4}** for MA

**MAPE\textsubscript{5}** for BP-NET

**MAPE\textsubscript{6}** for GM

Figure 4: Evolution of the MAPE mean values with time on USP dataset.

---

**Figure 5: Time series pattern and prediction on USP dataset.**

---

**Figure 6: Evolution of the MAPE mean values with time on AEP dataset.**
GM methods. Based on the ACCF method, the weighting coefficients by each base model can be computed by solving equations (16) and (17). Table 7 shows the weighting coefficients of models at different forecasting times. It can be seen that the sum of six weight coefficients is equal to 1, which is similar to the result in Section 4.2, but only \( w_4 \) and \( w_6 \) are equal to zero. PSI method, ES method, ARIMA method, and BP-NET method all play important role in the prediction of VS dataset. Then, the VS dataset from the 541th second to the 8580th second is further predicted by using the

| Forecasting time/seconds | PSI weight \( w_1 \) | ES weight \( w_2 \) | ARIMA weight \( w_3 \) | MA weight \( w_4 \) | BP-NET weight \( w_5 \) | GM weight \( w_6 \) |
|-------------------------|---------------------|-----------------|------------------|--------------|-----------------|-----------------|
| 217                     | 0.4978              | 0.0000          | 0.5122           | 0.0000       | 0.0000          | 0.0000          |
| 432                     | 0.5859              | 0.0000          | 0.4141           | 0.0000       | 0.0000          | 0.0000          |
| 648                     | 0.3875              | 0.0000          | 0.3929           | 0.0000       | 0.2196          | 0.0000          |
| 864                     | 0.4762              | 0.0000          | 0.5238           | 0.0000       | 0.0000          | 0.0000          |
| 1080                    | 0.2967              | 0.2414          | 0.2566           | 0.0000       | 0.2052          | 0.0000          |
| 1296                    | 0.4296              | 0.0000          | 0.3370           | 0.0000       | 0.2334          | 0.0000          |
| 1512                    | 0.5832              | 0.0000          | 0.4168           | 0.0000       | 0.0000          | 0.0000          |
| 1728                    | 0.3642              | 0.0000          | 0.3795           | 0.0000       | 0.2563          | 0.0000          |
| 1944                    | 0.3324              | 0.0000          | 0.3977           | 0.0000       | 0.2699          | 0.0000          |
| 2160                    | 0.3998              | 0.0000          | 0.3525           | 0.0000       | 0.2477          | 0.0000          |
| 2376                    | 0.3449              | 0.0000          | 0.3589           | 0.0000       | 0.2962          | 0.0000          |
| 2592                    | 0.3369              | 0.0000          | 0.3623           | 0.0000       | 0.3008          | 0.0000          |
| 2808                    | 0.4346              | 0.0000          | 0.5654           | 0.0000       | 0.0000          | 0.0000          |
| 3024                    | 0.3791              | 0.0000          | 0.3557           | 0.0000       | 0.2652          | 0.0000          |
| 3240                    | 0.3900              | 0.0000          | 0.3844           | 0.0000       | 0.2257          | 0.0000          |
| 3432                    | 0.4782              | 0.0000          | 0.5218           | 0.0000       | 0.0000          | 0.0000          |

Figure 7: Time series pattern and prediction on AEP dataset. (a) Forecasts from the 216th hour to the 1728th hour. (b) Forecasts from the 1729th hour to the 3432th hour.
ACCF method, as shown in Figure 9. Thus, once again, the ACCF method demonstrates a better prediction trend with good volatility and following quality.

4.4. Accuracy Analysis of ACCF Method

4.4.1. Comparison with Other Forecasting Models. Based on each prediction model, MAPEₖ (k = 1, 2, …, (N–3s)/s(3)), MAPEₖ for MA, MAPEₖ for BP-NET, and MAPEₖ for GM are calculated by equation (18). Then, the statistical test of MAPEₖ is shown in Table 8 on USP dataset, in Table 9 on AEP dataset, and in Table 10 on VS dataset. Max value of MAPEₖ, standard deviation of MAPEₖ can also be got and listed in Tables 8–10. It can be seen from Tables 8 and 9 that MAPEₖ, max value of MAPEₖ and standard deviation of MAPEₖ are smaller by ACCF method, as compared to those by other methods. Meanwhile, greater values of Count and Per can be got by ACCF method than other methods. Furthermore, in Table 10, max value and standard deviation of MAPEₖ by ACCF method are slightly greater than that of IV method and MSEI method, but ACCF method has the smaller value of MAPEₖ and the greater values of Count and Per over all methods. A possible reason is that statistical distribution law of historical forecasting errors was delved deeper by using ACCF method, and weighting coefficient for each model was modified more reasonably, leading to smaller MAPEₖ as well as greater statistical Count. Therefore, the ACCF method has the highest incidence of delivering the best predictions over all compared forecasting methods on USP dataset, AEP dataset, and VS dataset.
Table 11 presents a more visual view of prediction accuracy of each prediction model. It can be noticed that most of $FA_{all}$ values for individual methods are all more than 95%, and the four combination forecasting methods have higher $FA_{all}$ values of more than 97%. Obviously, the combination forecasting effect is most desirable and superior to the individual methods. It is interesting to know that the ACCF method is observed to be the best for the prediction of the USP AEP and VS datasets, due to higher forecasting accuracy. When judging by the $FA_{all}$ values of the four combination methods, ACCF method and IV method are superior to MSEI method and SWA method. In general, the developed ACCF algorithm is adaptive to adopt one (for USP) or some (for AEP and VS) among individual prediction methods, achieving satisfactory accuracy ($FA_{all} > 98.8\%$) in time series prediction, and it is suitable to the prediction of three datasets used in this study due to higher values of $FA_{all}$.

Figure 9: Time series pattern and prediction on VS dataset. (a) Forecasts from the 541th second to the 4320th second. (b) Forecasts from the 4321th second to the 8580th second.

| Models | $MAPE_k$ (%) | $MAPE_{all}$ (%) | $Count$ (MAPE$_k$ < 0.006) | Per (=Count/32) (%) |
|--------|--------------|-----------------|-----------------------------|-------------------|
| PSI    | 0.112        | 0.061           | 0                           | 0                 |
| ES     | 0.112        | 0.079           | 0                           | 0                 |
| ARIMA  | 0.033        | 0.006           | 21                          | 66                |
| MA     | 0.168        | 0.119           | 0                           | 0                 |
| BP-NET | 0.126        | 0.048           | 0                           | 0                 |
| GM     | 6.437        | 4.688           | 0                           | 0                 |
| IV     | 0.033        | 0.007           | 19                          | 59                |
| MSEI   | 0.054        | 0.021           | 0                           | 0                 |
| SWA    | 1.003        | 0.726           | 0                           | 0                 |
| ACCF   | 0.033        | 0.006           | 22                          | 69                |
4.4.2. Impact of Noise Ratio. In order to test robustness of the ACCF algorithm, the noisy data were further added to the AEP data set and the VS data set. The ratio of the standard deviation (STD) of added noise to the STD of original dataset is in a range from 0.00 to 0.50 in the study. The prediction accuracy of the ACCF method under noisy data was computed and compared with that of other forecasting methods, as shown in Tables 12 and 13.

As can be observed from Tables 12 and 13, when the proportion of noisy data increases, the FA all value of each algorithm decreases in all cases. Considering the different natural periodicity of each time series dataset, comparison methods can obtain different accuracy on different datasets. The FA all values of forecasting methods are between 76.25% and 98.18% for the AEP dataset and are in a range of 94.85%–98.77% for the VS dataset. In addition, for AEP dataset or VS dataset with noise ratio of 0.00, the ACCF method is superior to other methods, and the IV method is in the second place. With the increasing noise ratios, however, the ACCF method almost keeps the highest FA all value in two cases against other comparison algorithms (including not only individual methods, but also combination methods). For example, as the noise ratio changes from 0.0 to 0.5, the FA all of IV method decreases from

| Models | Max value of MAPE_k (%) | Standard deviation of MAPE_k (%) | MAPE all (%) | Count [MAPE_k < 1.094] | Per (=Count/134) (%) |
|--------|-------------------------|---------------------------------|--------------|-------------------------|----------------------|
| PSI    | 2.390                   | 0.401                           | 1.270        | 46                      | 34                   |
| ES     | 4.519                   | 0.697                           | 3.127        | 0                       | 0                    |
| ARIMA  | 3.018                   | 0.358                           | 1.314        | 35                      | 26                   |
| MA     | 6.416                   | 0.937                           | 4.617        | 0                       | 0                    |
| BP-NET | 4.589                   | 0.689                           | 1.959        | 10                      | 7                    |
| GM     | 66.269                  | 10.719                          | 22.116       | 0                       | 0                    |
| IV     | 2.891                   | 0.319                           | 1.141        | 67                      | 50                   |
| MSEI   | 3.130                   | 0.392                           | 1.406        | 24                      | 18                   |
| SWA    | 9.842                   | 1.464                           | 2.892        | 0                       | 0                    |
| ACCF   | 2.270                   | 0.279                           | 1.094        | 76                      | 57                   |

| Models | Max value of MAPE_k (%) | Standard deviation of MAPE_k (%) | MAPE all (%) | Count [MAPE_k < 1.189] | Per (=Count/134) (%) |
|--------|-------------------------|---------------------------------|--------------|-------------------------|----------------------|
| PSI    | 1.669                   | 0.140                           | 1.237        | 50                      | 37                   |
| ES     | 1.746                   | 0.139                           | 1.376        | 12                      | 9                    |
| ARIMA  | 1.933                   | 0.169                           | 1.367        | 22                      | 16                   |
| MA     | 2.574                   | 0.209                           | 1.964        | 0                       | 0                    |
| BP-NET | 1.913                   | 0.169                           | 1.454        | 7                       | 5                    |
| GM     | 8.304                   | 1.013                           | 4.705        | 0                       | 0                    |
| IV     | 1.593                   | 0.136                           | 1.220        | 56                      | 42                   |
| MSEI   | 1.571                   | 0.129                           | 1.274        | 35                      | 26                   |
| SWA    | 1.953                   | 0.161                           | 1.266        | 44                      | 33                   |
| ACCF   | 1.618                   | 0.138                           | 1.189        | 68                      | 51                   |

| Models | FA all (%) | FA all (%) | FA all (%) |
|--------|------------|------------|------------|
| PSI    | 99.939     | 99.730     | 98.763     |
| ES     | 99.921     | 96.873     | 98.624     |
| ARIMA  | 99.994     | 98.686     | 98.633     |
| MA     | 99.881     | 95.383     | 98.036     |
| BP-NET | 99.952     | 98.041     | 98.546     |
| GM     | 95.312     | 77.884     | 95.295     |
| IV     | 99.993     | 98.859     | 98.780     |
| MSEI   | 99.979     | 98.594     | 98.726     |
| SWA    | 99.274     | 97.108     | 98.734     |
| ACCF   | 99.994     | 98.906     | 98.811     |
98.120% to 90.882% for AEP dataset, and from 98.742% to 97.443% for VS dataset, while the FA\textsubscript{all} of ACCF method decreases only from 98.184% to 90.955% for AEP dataset, and from 98.770% to 97.452% for VS dataset. This might be due to the high stochasticity of the tested dataset under high noise ratio, and thus the comparison models could not capture the actual trend of historical forecasting errors. Especially for IV, MSEI, and SWA models, the statistical distribution information of the forecasting errors with the historical time is not considered, so they show lower robustness than developed ACCF model. On the contrary, because the statistical forecasting errors are used to correct weights in real time, the robustness of ACCF method is better than other comparison methods for noisy data.

Therefore, it is concluded that the proposed ACCF algorithm obtains higher prediction accuracy on time series datasets and is more robust to noisy data than other individual methods, as well as combination methods.

5. Conclusions

(i) According to the individual forecasting methods, such as PSI, ES, ARIMA, MA, and BP-NET methods, an ACCF method with adaptive weighting coefficients is proposed for short-term prediction of the time-series data.

(ii) The combination forecasting methods are most desirable and superior to the individual methods. In contrast to other forecasting methods, the proposed ACCF method is adaptive to adopt one or some of prediction methods and shows satisfactory forecasting quality due to its flexible adaptability and high forecasting accuracy. The ACCF method is extremely suitable for short-term prediction of time series datasets.

(iii) The higher the noise ratio of the tested datasets, the lower the prediction accuracy of the ACCF method. But the proposed ACCF methods can still achieve significant advantages compared with other forecasting methods in terms of forecasting accuracy. The ACCF method demonstrates a better prediction trend with good volatility and following quality.

**Abbreviations**

- $dt$: Equal interval of time
- $ej$: The sum of squared errors of $j$-th single model
- $FA$: Forecasting accuracy
- $kj$: Either 0 or 1 ($j = 1, 2, \ldots, n$)
- $K0$: Correction coefficient for period index
- MAPE: Mean absolute percentage error

**Table 12:** Prediction accuracy of comparison algorithms on AEP dataset.

| Models | $FA_{all}$ (%) | 0.05 | 0.10 | 0.15 | 0.20 | 0.25 | 0.30 | 0.35 | 0.40 | 0.45 | 0.50 |
|--------|----------------|------|------|------|------|------|------|------|------|------|------|
| PSI    | 98.073         | 97.252 | 96.324 | 95.348 | 94.381 | 93.466 | 92.587 | 91.716 | 90.824 | 89.881 |
| ES     | 96.725         | 96.346 | 95.792 | 95.115 | 94.370 | 93.603 | 92.827 | 92.050 | 91.278 | 90.517 |
| ARIMA  | 97.853         | 96.997 | 96.126 | 95.244 | 94.358 | 93.475 | 92.603 | 91.753 | 90.933 | 90.154 |
| MA     | 95.304         | 95.047 | 94.632 | 94.080 | 93.409 | 92.642 | 91.816 | 90.967 | 90.135 | 89.357 |
| BP-NET | 96.986         | 95.825 | 94.608 | 93.386 | 92.209 | 91.118 | 90.106 | 89.156 | 88.252 | 87.377 |
| GM     | 77.871         | 77.785 | 77.649 | 77.489 | 77.332 | 77.192 | 77.047 | 76.865 | 76.612 | 76.257 |
| IV     | 98.120         | 97.344 | 96.542 | 95.724 | 94.900 | 94.080 | 93.266 | 92.461 | 91.666 | 90.882 |
| MSEI   | 97.905         | 97.164 | 96.386 | 95.583 | 94.771 | 93.960 | 93.155 | 92.359 | 91.573 | 90.801 |
| SWA    | 97.048         | 96.689 | 96.110 | 95.390 | 94.610 | 93.835 | 93.078 | 92.336 | 91.608 | 90.892 |
| ACCF   | 98.184         | 97.399 | 96.576 | 95.741 | 94.918 | 94.127 | 93.354 | 92.581 | 91.788 | 90.955 |

**Table 13:** Prediction accuracy of comparison algorithms on VS dataset.

| Models | $FA_{all}$ (%) | 0.05 | 0.10 | 0.15 | 0.20 | 0.25 | 0.30 | 0.35 | 0.40 | 0.45 | 0.50 |
|--------|----------------|------|------|------|------|------|------|------|------|------|------|
| PSI    | 98.725         | 98.637 | 98.510 | 98.354 | 98.180 | 97.996 | 97.805 | 97.609 | 97.409 | 97.205 |
| ES     | 98.600         | 98.521 | 98.402 | 98.255 | 98.095 | 97.933 | 97.768 | 97.599 | 97.425 | 97.245 |
| ARIMA  | 98.596         | 98.506 | 98.377 | 98.223 | 98.055 | 97.886 | 97.716 | 97.542 | 97.365 | 97.181 |
| MA     | 98.022         | 97.963 | 97.867 | 97.746 | 97.609 | 97.464 | 97.312 | 97.151 | 96.977 | 96.789 |
| BP-NET | 98.503         | 98.418 | 98.301 | 98.161 | 98.005 | 97.841 | 97.671 | 97.495 | 97.316 | 97.132 |
| GM     | 95.290         | 95.281 | 95.267 | 95.242 | 95.203 | 95.149 | 95.082 | 94.931 | 94.856 | 94.856 |
| IV     | 98.742         | 98.663 | 98.552 | 98.416 | 98.266 | 98.108 | 97.945 | 97.779 | 97.611 | 97.443 |
| MSEI   | 98.693         | 98.619 | 98.513 | 98.382 | 98.236 | 98.081 | 97.921 | 97.757 | 97.590 | 97.423 |
| SWA    | 98.702         | 98.632 | 98.530 | 98.405 | 98.263 | 98.111 | 97.932 | 97.758 | 97.521 | 97.417 |
| ACCF   | 98.770         | 98.686 | 98.570 | 98.431 | 98.276 | 98.115 | 97.949 | 97.782 | 97.616 | 97.452 |

Max values: 98.184, 97.399, 96.576, 95.741, 94.918, 94.127, 93.354, 92.581, 91.788, 90.955
MAPE_1: New historical arrays given in equation (14) by PSI method
MAPE_2: New historical arrays given in equation (14) by ES method
MAPE_3: New historical arrays given in equation (14) by ARIMA method
MAPE_4: New historical arrays given in equation (14) by MA method
MAPE_5: New historical arrays given in equation (14) by BP-NET method
MAPE_6: New historical arrays given in equation (14) by GM method

N: Number of forecasting methods
Pi(t_i): Period index at time of t_i
Si(t_i): Sequential index at time of t_i
t_i: Time (t_1 = t_2, ..., t_N)
T: Fixed time period
y(t_i): Observed value at time of t_i
Y_j(t_i): Forecasting value at time of t_i by PSI method
Y_j(t_i): Forecasting value at time of t_i by ES method
Y_j(t_i): Forecasting value at time of t_i by ARIMA method
Y_j(t_i): Forecasting value at time of t_i by MA method
Y_j(t_i): Forecasting value at time of t_i by BP-NET method
Y_j(t_i): Forecasting value at time of t_i by GM method
w_j: Weighting coefficients by ACCF method (j = 1, 2, ..., n)

MAPE_j: Mean values for historical arrays MAPE_j (j = 1, 2, ..., n)
MAPE_m: The smallest values in MAPE_j (j = 1, 2, ..., n)
σ_j: Standard deviations for historical arrays MAPE_j (j = 1, 2, ..., n)
σ_m: Standard deviations for historical arrays MAPE_m
F(t_i): Forecasting value at time of t_i by ACCF method
A: Optimized weighing factor of PSI method
β: Smoothing parameter.

Data Availability

The data presented in this study are available upon request from the corresponding author.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors’ Contributions

Conceptualization was performed by H. J., D. F.; formal analysis was performed by H. J., D. F.; investigation was performed by H. J., D. F.; original draft was written by H. J., X. Z.; reviewing and editing were performed by H. J., D. F.

Acknowledgments

This study was funded by the Key Projects for International Cooperation in Scientific and Technological Innovation between Governments (grant no. 2017YFE0101600). The authors would like to express their sincere thanks to technical support from Sino–German Institute for Intelligent Technologies.

References

[1] G. Shi, J. Guo, W. Huang, and B. M. Williams, “Modeling seasonal heteroscedasticity in vehicular traffic condition series using a seasonal adjustment approach,” Journal of Transportation Engineering, vol. 140, no. 5, p. 04014012, 2014.
[2] C. Stefanakos, “Fuzzy time series forecasting of nonstationary wind and wave data,” Ocean Engineering, vol. 121, pp. 1–12, 2016.
[3] N. D. Bokde, Z. M. Yaseen, and G. B. Andersen, “ForecastTB an R package as a test-bench for time series forecasting-application of wind speed and solar radiation modeling,” Energies, vol. 13, no. 10, p. 2578, 2020.
[4] M. D. Felice, A. Alessandri, and P. M. Ruti, “Electricity demand forecasting over Italy: potential benefits using numerical weather prediction models,” Electric Power Systems Research, vol. 104, pp. 71–79, 2013.
[5] Y. Qian, R. Yan, and R. X. Gao, “A multi–time scale approach to remaining useful life prediction in rolling bearing,” Mechanical Systems and Signal Processing, vol. 83, pp. 549–567, 2017.
[6] H. Y. Jiang, D. J. Fang, K. Spicher, F. Cheng, and B. X. Li, “A new period-sequential index forecasting algorithm for time series data,” Applied Sciences, vol. 9, no. 20, pp. 1–12, 2019.
[7] D. K. Barrow, “Forecasting intraday call arrivals using the seasonal moving average method,” Journal of Business Research, vol. 69, no. 12, pp. 6088–6096, 2016.
[8] V. Š. Ediger and S. Akar, “ARIMA forecasting of primary energy demand by fuel in Turkey,” Energy Policy, vol. 35, no. 3, pp. 1701–1708, 2007.
[9] Z. Dong, D. Yang, T. Reindl, and W. M. Walsh, “Short-term solar irradiance forecasting using exponential smoothing state space model,” Energy, vol. 55, pp. 1104–1113, 2013.
[10] C. D. Zuluaga, M. A. Álvarez, and E. Giráldos, “Short-term wind speed prediction based on robust Kalman filtering: an experimental comparison,” Applied Energy, vol. 156, pp. 321–330, 2015.
[11] C.-H. Wang and L.-C. Hsu, “Using genetic algorithms grey theory to forecast high technology industrial output,” Applied Mathematics and Computation, vol. 195, no. 1, pp. 256–263, 2008.
[12] U. Beyaztas, S. Q. Salih, K.-W. Chau, N. Al-Ansari, and Z. M. Yaseen, “Construction of functional data analysis modeling strategy for global solar radiation prediction: application of cross-station paradigm,” Engineering Applications of Computational Fluid Mechanics, vol. 13, no. 1, pp. 1165–1181, 2019.
[13] T. Niu, J. Wang, K. Zhang, and P. Du, “Multi-step-ahead wind speed forecasting based on optimal feature selection and a modified bat algorithm with the cognition strategy,” Renewable Energy, vol. 118, pp. 213–229, 2018.
[14] Z.-H. Guo, J. Wu, H.-Y. Lu, and J.-Z. Wang, “A case study on a hybrid wind speed forecasting method using BP neural network,” Knowledge-Based Systems, vol. 24, no. 7, pp. 1048–1056, 2011.
[15] C. Voyant, M. Muselli, C. Paoli, and M.-L. Nivet, “Optimization of an artificial neural network dedicated to the multivariate forecasting of daily global radiation,” *Energy*, vol. 36, no. 1, pp. 348–359, 2011.

[16] S. K. Aggarwal, A. Goel, and V. P. Singh, “Stage and discharge forecasting by SVM and ANN techniques,” *Water Resources Management*, vol. 26, no. 13, pp. 3705–3724, 2012.

[17] T. Korol and A. Korodi, “An evaluation of effectiveness of fuzzy logic model in predicting the business bankruptcy,” *Romanian Journal of Economic Forecasting*, vol. 14, pp. 92–107, 2011.

[18] J. Duan, X. Qiu, W. Ma, X. Tian, and D. Shang, “Electricity consumption forecasting scheme via improved LSSVM with maximum correntropy criterion,” *Entropy*, vol. 20, no. 2, p. 112, 2018.

[19] J. Liu, X. Wang, and Y. Lu, “A novel hybrid methodology for short-term wind power forecasting based on adaptive neuro-fuzzy inference system,” *Renewable Energy*, vol. 103, pp. 620–629, 2017.

[20] S. Zhang, Y. W. Liu, Y. Liu, J. Wang, and C. Wang, “Research on combined model based on multi-objective optimization and application in wind speed forecast,” *Applied Sciences*, vol. 9, no. 3, p. 423, 2019.

[21] Y. Zhang, Y. Wei, Y. Zhang, and D. Jin, “Forecasting oil price volatility: forecast combination versus shrinkage method,” *Energy Economics*, vol. 80, pp. 423–433, 2019.

[22] W. Chen, H. Xu, Z. Chen, and M. Jiang, “A novel method for time series prediction based on error decomposition and nonlinear combination of forecasters,” *Neurocomputing*, vol. 426, pp. 85–103, 2021.

[23] T. E. Clark and M. W. Mcraken, “Combining forecasts from nested models,” *Oxford Bulletin of Economics and Statistics*, vol. 71, no. 3, pp. 303–329, 2009.

[24] S. F. Hao, Y. M. Fang, and J. W. Yang, “Application of combination model of entropy method to landslide deformation prediction,” *Journal of Surveying and Mapping Engineering*, vol. 23, pp. 62–64, 2014.

[25] P. P. Xie, T. Huang, and Y. Liu, “Application of changeable weight combination forecasting model to dam settlement predicting,” *Journal of Surveying and Mapping Engineering*, vol. 24, pp. 74–76, 2015.

[26] C. Gao, X. Cui, and N. Gao, “Study on deformation parallel combination prediction under different restriction criterion,” *Geodesy and Geodynamics*, vol. 34, pp. 91–94, 2014.

[27] C. Song and X. Fu, “Research on different weight combination in air quality forecasting models,” *Journal of Cleaner Production*, vol. 261, p. 121169, 2020.

[28] S. Liu, L. Wang, H. Qu, and Z. G. Wang, “Optimal forecast combination based on neural networks for time series forecasting,” *Applied Soft Computing*, vol. 66, pp. 1–17, 2018.

[29] N. Kourentzes, “Intermittent demand forecasts with neural networks,” *International Journal of Production Economics*, vol. 143, no. 1, pp. 198–206, 2013.

[30] E. Bas, C. Grosan, E. Egrioglu, and U. Yolcu, “High order fuzzy time series method based on pi-sigma neural network,” *Engineering Applications of Artificial Intelligence*, vol. 72, pp. 350–356, 2018.

[31] J. Nowotarski, E. Raviv, S. Trück, and R. Weron, “An empirical comparison of alternative schemes for combining electricity spot price forecasts,” *Energy Economics*, vol. 46, pp. 395–412, 2014.

[32] “Kaggle - dataset - United States population time series data,” 2019, https://www.kaggle.com/census/population-time-series-data.

[33] “Kaggle - dataset - condition monitoring of hydraulic systems, hourly power consumption data,” 2018, https://www.kaggle.com/robikscube/hourly-energy-consumption.

[34] “Kaggle - dataset - condition monitoring of hydraulic systems, condition assessment of a hydraulic test rig based on multi sensor data,” 2019, https://www.kaggle.com/jacostupa/condition-monitoring-of-hydraulic-systems.

[35] N. Helwig, E. Pignanelli, and A. Schütze, “Condition monitoring of a complex hydraulic system using multivariate statistics,” in *Proceedings of the Conference Record - IEEE Instrumentation and Measurement Technology Conference*, pp. 210–215, Pisa, Italy, May 2015.