

Article

# Decomposition-Based Dynamic Adaptive Combination Forecasting for Monthly Electricity Demand

Zhineng Hu <sup>1</sup>, Jing Ma <sup>1</sup>, Liangwei Yang <sup>2</sup>, Xiaoping Li <sup>1,\*</sup> and Meng Pang <sup>3</sup>

<sup>1</sup> Business School, Sichuan University, Chengdu 610064, China; huzn@scu.edu.cn (Z.H.); majing937824563@outlook.com (J.M.)

<sup>2</sup> Big Data Research Center, University of Electronic Science and Technology of China, Chengdu 611731, China; liangwei\_yang@outlook.com

<sup>3</sup> Sichuan Regulatory Commission, State Administration of Energy of China, Chengdu 641000, China; y805797986@gmail.com

\* Correspondence: lixiaoping@scu.edu.cn; Tel.: +86-28-8541-0276

Received: 23 January 2019; Accepted: 25 February 2019; Published: 28 February 2019



**Abstract:** (1) Background: Electricity consumption data are often made up of complex, unstable series that have different fluctuation characteristics in different industries. However, electricity demand forecasting is a prerequisite for the control and scheduling of power systems. (2) Methods: As most previous research has focused on prediction accuracy rather than stability, this paper developed a decomposition-based combination forecasting model using dynamic adaptive entropy-based weighting for total electricity demand forecasting at the engineering level. (3) Results: To further illustrate the prediction accuracy and stationarity of the proposed method, a comparison analysis using an analysis of variance and an orthogonal approach to solve the least squares equations was conducted using classical individual models, a combination forecasting model, and a decomposition-based combination forecasting model. The proposed method had a very satisfactory overall performance with good verification and validation compared to autoregressive integrated moving average (ARIMA) and artificial neural-networks (ANN). (4) Conclusion: As the proposed method dynamically combines various forecast models and can decompose and adapt to various characteristic data sets, it was found to have an accurate, stable forecast performance. Therefore, it could be broadly applied to forecasting electricity demand and developing electricity generation plans and related energy policies.

**Keywords:** dynamic adaptive forecast; entropy-based weighting; electricity demand forecasting

## 1. Introduction

As electricity demand forecasting is essential for energy management, maintenance scheduling, and secure modern power management [1], researchers have long been focused on developing optimum forecasting methods.

Depending on the selected time horizon, demand forecasting can be short-term from an hour to a week, medium-term from a week to a year, or long-term for over a year. Medium-term electricity demand forecasting, and especially monthly forecasting, is not only used to balance supply and demand [2] but is also an important index for many associated decisions [3] such as equipment maintenance, fuel trading, and bilateral electricity transactions, which are generally made months in advance.

Many scientific and engineering methods to improve the accuracy of electricity demand forecasting have been proposed in recent years [4,5]. Prediction accuracy and stability are vital for

electricity demand forecasting at the engineering level. However, as complex machine learning methods require staff, hardware, and software, local authorities and local electricity regulatory commissions are generally seeking effective, simple, electricity demand forecasting models. The main objective of this paper, therefore, is to propose a dynamic adaptive forecasting model to accurately and effectively forecast total electricity demand at the engineering level that could be broadly applied to electricity demand forecasting.

The remainder of the paper is organized as follows. Section 2 presents the problem and research methods, Section 3 outlines the data handling, error measurement forecasting, entropy, combination forecasting, the decomposition-based combination forecast model and its framework, Section 4 verifies and validates the proposed model through a step by step comparison analysis with six individual models, a combination forecasting model, a decomposition-based combination forecasting model and other models, and Section 5 summarizes the paper and gives guidance on future research directions.

## 2. Literature Review

Because of rising electricity demand, growing environmental/health concerns, and shrinking resource availability, the need for more accurate electricity demand forecasting models has increased, which has in turn attracted significant research attention [6]. Two main types of forecasting methods have been developed—individual models and hybrid models.

There are two main types of individual models—classical forecasting models that analyze past electricity consumption at the engineering level such as exponential smoothing [7], ARIMA [8], state space models [9], grey models [10] and linear regression [11], and machine learning models such as artificial neural-networks (ANN) [12], support vector regression [13], Gaussian processes [14] and ensemble learning methods [15].

Many studies have shown that a combination of different models can improve forecast performance and, as no one forecasting method has been found to obtain the best results for every circumstance, combination forecasting is probably the best way to improve accuracy [16]. When models are combined, they can include a wider range of electricity consumption features and overcome the defects in the individual models. There are two main types of hybrid electricity load forecasting models: in the first, the electricity demand is first predicted separately by the different models, the weight of each model is then calculated, and the final forecasting value is determined by adding each model's forecasting value multiplied by its weight [17]; and in the second, the hybrid model decomposes the electricity consumption data into several sub-series, each sub-series is forecast using a suitable model, and the final forecast results are the sum of each of the sub-series' forecasting results [18]; that is, the first is a combination of models and the second involves data decomposition.

As noted by Wang et al. [19], the more system factors that can be considered, the higher the forecasting precision. Electricity data rules and characteristics can be easily obtained using decomposition or combination [20]; for example, Zhang et al. [21] applied a decomposition approach to forecast short-term electricity demand, in which the data series were split into two new series and two different models trained to forecast these separately, Li et al. [22] proved that a random forest technique based on ensemble empirical mode decomposition was able to improve the forecasting accuracy of daily enterprise electricity consumption, Laouafi et al. [1] developed a combination methodology for electricity demand forecasting, for which six individual models were applied to the real-time load data, and the final load estimation obtained by adding each model's forecasting value multiplied by its weight, and Chen et al. [23] proposed a generalized model for wind turbine faulty condition detection using a combination prediction approach and information entropy.

However, there have been few studies that have used both decomposition and a combination forecasting model, and most studies have sought to prove method effectiveness by comparing the error means rather than the statistical significance. To simultaneously achieve adaptive, controllable processes, this paper proposes a decomposition-based combination forecast method that can dynamically adapt to various forecast models using combination, and dynamically adapt to

the various characteristics of the different data sets through decomposition. Previous research has tended to determine the combined weights for the combination model using only one measurement such as MAE [24] or absolute percentage error [1]; however, this paper applies five entropy-based error measurements to dynamically weight the individual model, and introduces the analysis of variance and an orthogonal approach to solve the least squares equations to analyze the effectiveness of the statistical significance of the proposed method.

### 3. Materials and Methods

#### 3.1. Data Handling

This study was based on electricity consumption data from January 2007 to October 2009 in Sichuan Province, China. The data sets were generated from the Electricity Saving Association, which is the major source of power data in the province and covers almost all sectors of the economy. While some of the observations were suspect (for example, almost all data from September 2009), all data were used for comparison purposes.

To better validate the proposed method, total electricity demand and all components were forecast. The national economy electricity consumption classifications divides total electricity demand into four levels, which respectively cover 2, 5, 8, and 23 types of industries for the decomposition. Table 1 lists the industries and their codes, and Figure 1 shows the decomposition details.

**Table 1.** Power consumption by industry sector.

Code	Power Industry	Code	Power Industry
s <sub>00</sub>	Total power consumption	s <sub>45</sub>	Fishery power consumption
s <sub>11</sub>	Nonresidential power consumption	s <sub>46</sub>	Coal mining power consumption
s <sub>12</sub>	Residential power consumption	s <sub>47</sub>	Production and supply of electricity, gas and water power consumption
s <sub>21</sub>	Total Primary industry power consumption	s <sub>48</sub>	Manufacturing power consumption
s <sub>22</sub>	Total Second industry power consumption	s <sub>49</sub>	Warehousing industrial power consumption
s <sub>23</sub>	Total Tertiary industry power consumption	s <sub>410</sub>	Transportation power consumption
s <sub>24</sub>	Urban residential power consumption	s <sub>411</sub>	Postal service power consumption
s <sub>25</sub>	Rural residential power consumption	s <sub>412</sub>	Real estate power consumption
s <sub>32</sub>	Industry power consumption	s <sub>413</sub>	Financial industrial power consumption
s <sub>31</sub>	Agriculture, forestry, animal husbandry and fishery power consumption	s <sub>414</sub>	Leasing, business services, resident services and other services power consumption
s <sub>33</sub>	Construction industry power consumption	s <sub>415</sub>	Catering and accommodation power consumption
s <sub>34</sub>	Information transfer, computer services and software power consumption	s <sub>416</sub>	Wholesale and retail power consumption
s <sub>35</sub>	Business, catering and accommodation power consumption	s <sub>417</sub>	Computer services and software power consumption
s <sub>36</sub>	Financial, real estate and residential property services power consumption	s <sub>418</sub>	Information transfer power consumption
s <sub>37</sub>	Public utilities and management organizations power consumption	s <sub>419</sub>	Public management, social organization, and management organization
s <sub>38</sub>	Transportation, warehousing and postal power consumption	s <sub>420</sub>	Health, social security and social welfare power consumption
s <sub>41</sub>	Forestry power consumption	s <sub>421</sub>	Education power consumption
s <sub>42</sub>	Services power consumption for agriculture, forestry, animal husbandry and fishery	s <sub>422</sub>	Scientific research, technical services, and geological prospecting power consumption
s <sub>43</sub>	Agriculture power consumption	s <sub>423</sub>	Water environment and public facilities management power consumption
s <sub>44</sub>	Animal husbandry power consumption		

$y_{ij}^1, y_{ij}^2, \dots, y_{ij}^{34}$  denote the monthly electricity data for industry  $s_{ij}$ , where  $i$  represents  $i$  level at which total electricity demand is divided and  $j$  represents  $j$ th types of industry at level  $i$ . To better elaborate the process, the data were divided into training data and test data. The procedures for each industry were as follows: first, time series  $s_{ij}^1 = \{y_{ij}^1, y_{ij}^2, \dots, y_{ij}^{24}\}$  were used as the training data set with a corresponding forecast value of  $\{y_{ij}^{25}\}$ . Then,  $s_{ij}^2 = \{y_{ij}^1, y_{ij}^2, \dots, y_{ij}^{25}\}$  were used as the training

data set to forecast  $\{\hat{y}_{ij}^{26}\}$ . Similarly,  $s_{ij}^{10} = \{y_{ij}^1, y_{ij}^2, \dots, y_{ij}^{33}\}$  were used to forecast  $\{\hat{y}_{ij}^{34}\}$ , as shown in Table 2.

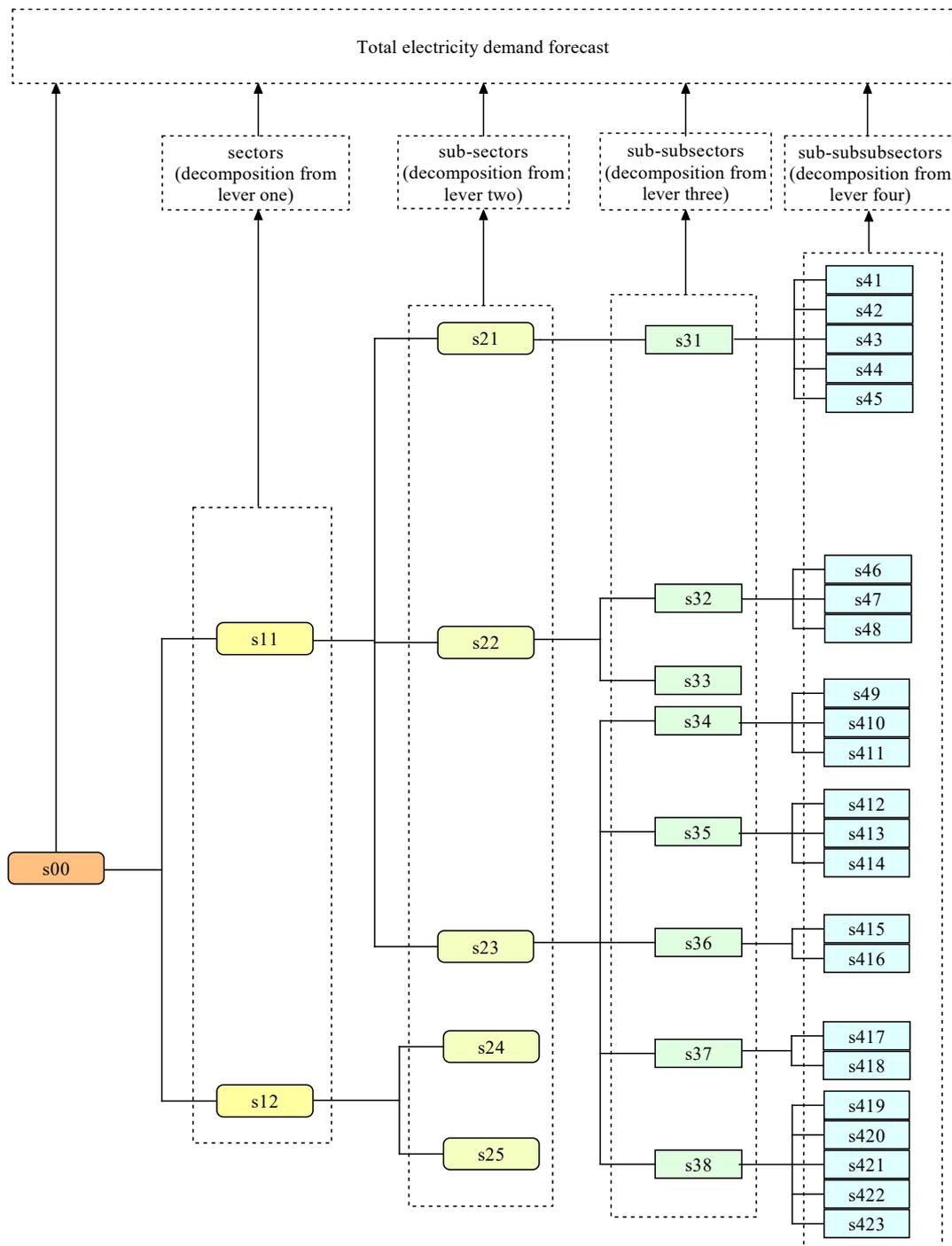


Figure 1. Electricity decomposition.

Table 2. Data division for each industry.

Training Set	Test Set
$s_{ij}^1 = \{y_{ij}^1, y_{ij}^2, \dots, y_{ij}^{24}\}$	$\{y_{ij}^{25}\}$
$s_{ij}^2 = \{y_{ij}^1, y_{ij}^2, \dots, y_{ij}^{25}\}$	$\{y_{ij}^{26}\}$
$s_{ij}^3 = \{y_{ij}^1, y_{ij}^2, \dots, y_{ij}^{26}\}$	$\{y_{ij}^{27}\}$
$s_{ij}^4 = \{y_{ij}^1, y_{ij}^2, \dots, y_{ij}^{27}\}$	$\{y_{ij}^{28}\}$
$s_{ij}^5 = \{y_{ij}^1, y_{ij}^2, \dots, y_{ij}^{28}\}$	$\{y_{ij}^{29}\}$
$s_{ij}^6 = \{y_{ij}^1, y_{ij}^2, \dots, y_{ij}^{29}\}$	$\{y_{ij}^{30}\}$
$s_{ij}^7 = \{y_{ij}^1, y_{ij}^2, \dots, y_{ij}^{30}\}$	$\{y_{ij}^{31}\}$
$s_{ij}^8 = \{y_{ij}^1, y_{ij}^2, \dots, y_{ij}^{31}\}$	$\{y_{ij}^{32}\}$
$s_{ij}^9 = \{y_{ij}^1, y_{ij}^2, \dots, y_{ij}^{32}\}$	$\{y_{ij}^{33}\}$
$s_{ij}^{10} = \{y_{ij}^1, y_{ij}^2, \dots, y_{ij}^{33}\}$	$\{y_{ij}^{34}\}$

### 3.2. Forecast Error Measurement

The Mean Absolute Deviation (MAD), Mean Squared Error (MSE), Cumulative Sum of Forecast Errors (CFE), Mean Absolute Percentage Error (MAPE), and Tracking Signal and Absolute Relative Error (ARE) were all applied to assess the forecasting capability, the definitions for which are as follows:

$$\begin{aligned}
 \text{MAD} &= \frac{\sum_{t=1}^n |A_t - F_t|}{n}, & \text{MSE} &= \frac{\sum_{t=1}^n (A_t - F_t)^2}{n}, \\
 \text{CFE} &= \frac{\sum_{t=1}^n (A_t - F_t)}{n}, & \text{MAPE} &= 100\% * \frac{\sum_{t=1}^n (|\frac{A_t - F_t}{A_t}|)}{n}, \\
 \text{Trk.Signal} &= \frac{\sum_{t=1}^n (A_t - F_t)}{\text{MAD}}, & \text{ARE} &= \left| \frac{A_t - F_t}{A_t} \right|,
 \end{aligned}$$

where  $A_t$  is the actual value in period  $t$ ,  $F_t$  is the forecast value in period  $t$ , and  $n$  is the total number of forecast periods..

The MAD, MSE, CFE, MAPE, and Trk.Signal were calculated for each training set, and the combination models selected by considering all five measurements, which were weighted using the entropy-based method outlined in Sections 3.3 and 3.4. ARE statistical analyses were used on each test set to assess the models and derive additional information, as detailed in Section 4.

### 3.3. Entropy

To weight the five measurements, the entropy [25] was revised as outlined in the following steps.

Step 1: The error matrix  $D = [x_{ab}]_{mn}$  was calculated for every time series  $s_{ij}^l$ , where  $x_{ab}$  was the  $a$ th index value ( $a$ th individual forecast method) for the  $b$ th index attribute ( $b$ th forecast error measurement), and  $m$  and  $n$  were the number of individual forecast methods and the number of forecast error measurements.

Step 2: Proportion  $p_{ab}$ , which was the  $a$ th model for the proportion of the feature or the contribution of the feature under the terms of the  $b$ th index, was calculated. To capture more information from the data, the proximity index  $r_{ab} = \sqrt{x_b^*/x_{ab}}$  was applied, where  $x_b^* = \min |x_{ab}|$  was the ideal value for the  $j$ th index attribute; then,  $p_{ab} = r_{ab} / \sum_{a=1}^m r_{ab}$ .

Step 3: The entropy for the  $b$ -term indicator  $E_b$  was defined to represent the total entropy contribution of all models to the  $b$ th indicators, and  $E_b = -d \sum_{a=1}^m p_{ab} \ln p_{ab}$ , where the constant  $d = 1 / \ln m$ .

Step 4: The index difference coefficients  $g_b$ ;  $g_b = 1 - E_b$  were calculated, in which the difference coefficient  $g_b$  indicated the extent of the inconsistent index contributions under the  $b$ th index attribute determined by  $E_b$ .

Step 5: The weighting coefficients were determined. After weighting the coefficient,  $P_b = g_b / \sum_{b=1}^n g_b$  was the normalized weight for the weight coefficient.

### 3.4. Combination Forecasting Model

As is well known, the classical forecasting methods—moving average (MA), the moving average with linear trend (MAT), single exponential smoothing with linear trends (SEST), double exponential smoothing with linear trends (DEST), the Holt-Winters additive algorithm (HWA), and the Holt-Winters multiplicative algorithm (HWM)—each have a different ability to deal with trends and seasonality [26–29].

However, compared with other complex methods and especially with modern intelligent forecasting models, these classical models are unable to completely adapt to different features in different data series. Therefore, to prove the effectiveness of the proposed method, these simpler methods were applied as individual methods in this paper rather than the modern intelligent forecasting models.

MA and exponential smoothing models are simple to use and do not require an in-depth knowledge of forecasting methods. HWM is the most complex model of the six models, so the overall computation efficiency of the proposed method would be restricted by HWM if the six models were employed in a parallel computation environment. However, the main advantages of HWM are its ease of application, speed, and reduced computational burden [30]; therefore, because its forecasting accuracy is similar to more complex methods, this method could be more easily applied by decision makers.

The combination forecast model (CF) was used to forecast the subsequent month's electricity demand for all the time series ( $s_{ij}^l$ ) in the training set. As suggested by Laouafi et al. [1], combined forecasting can improve forecast accuracy and can be employed to combine the forecasts from the primary individual models.

$$F(f_1(y), f_2(y), \dots, f_6(y)) = w_1 f_1(y) + w_2 f_2(y) + \dots + w_6 f_6(y), \quad (1)$$

where  $f_i$  is related to the individual forecasting methods that transform the real electricity consumption  $y$  of the previous months to its forecasting  $\hat{y}$ , and  $w_i$  represents the weight of the  $i$ th individual forecasting method calculated from real electricity consumption  $y$  in the preceding months.

For each time series  $s_{ij}^l$ , there was a weight vector  $w$ , which was calculated as follows:

Step 1:  $w_a^b$  was first calculated, which was the weight of the  $a$ th individual forecasting method judged by the  $b$ th forecasting error measurement. The inverse error weights ensured that the models with smaller errors were assigned a greater weight:

$$w_a^b = \frac{e_{ab}^{-1}}{\sum_{i=1}^n e_{ab}^{-1}} \quad (\forall a = 1, 2, \dots, 6, \quad \forall b = 1, 2, \dots, 5),$$

where  $e_{ab}$  is the error of the  $a$ th forecasting model judged by the  $b$ th forecasting error measurement.

Step 2: The entropy of  $p_b$  was calculated as described in Section 3.3, where  $p_b$  was aligned to the weight of the  $b$ th forecasting error measurement and  $\sum p_b = 1$ .

Step 3: The combined weight  $w_a$  was calculated

$$w_a = \sum_{j=a}^m w_a^b p_b \quad (\forall a = 1, 2, \dots, 6 \quad \forall b = 1, 2, \dots, 5).$$

### 3.5. Decomposition-Based Combination Forecasting

A decomposition-based combination forecasting model is proposed, in which  $G_{ij}^u$  denotes the electricity demand forecasting for industry  $s_{ij}$  using decomposition-based combination forecasting at level  $u$  (DCF $u$ ).

$$G_{ij}^u = \sum_v F_{uv}, \quad i = 0, 1, 2, 3, \quad u = i, \dots, 4, \quad (2)$$

where  $F_{uv}$  is the forecast value using the combination forecasting model for industry  $s_{uv}$ , and  $v$  represents the  $v$ th types of industry at level  $u$ .

For example, based on Equation (2), the total electricity demand can be forecast using the decomposition-based combination method, as follows:

$$G_{00}^1 = F_{11} + F_{12} \quad (\text{DCF1}), \quad G_{00}^2 = \sum_{v=1}^5 F_{2v} \quad (\text{DCF2}),$$

$$G_{00}^3 = \sum_{v=1}^8 F_{3v} \quad (\text{DCF3}), \quad G_{00}^4 = \sum_{v=1}^{23} F_{4v} \quad (\text{DCF4}).$$

To explain the proposed method more clearly, a flow chart was developed, as shown in Figure 2 (decomposition-based CF at level four).

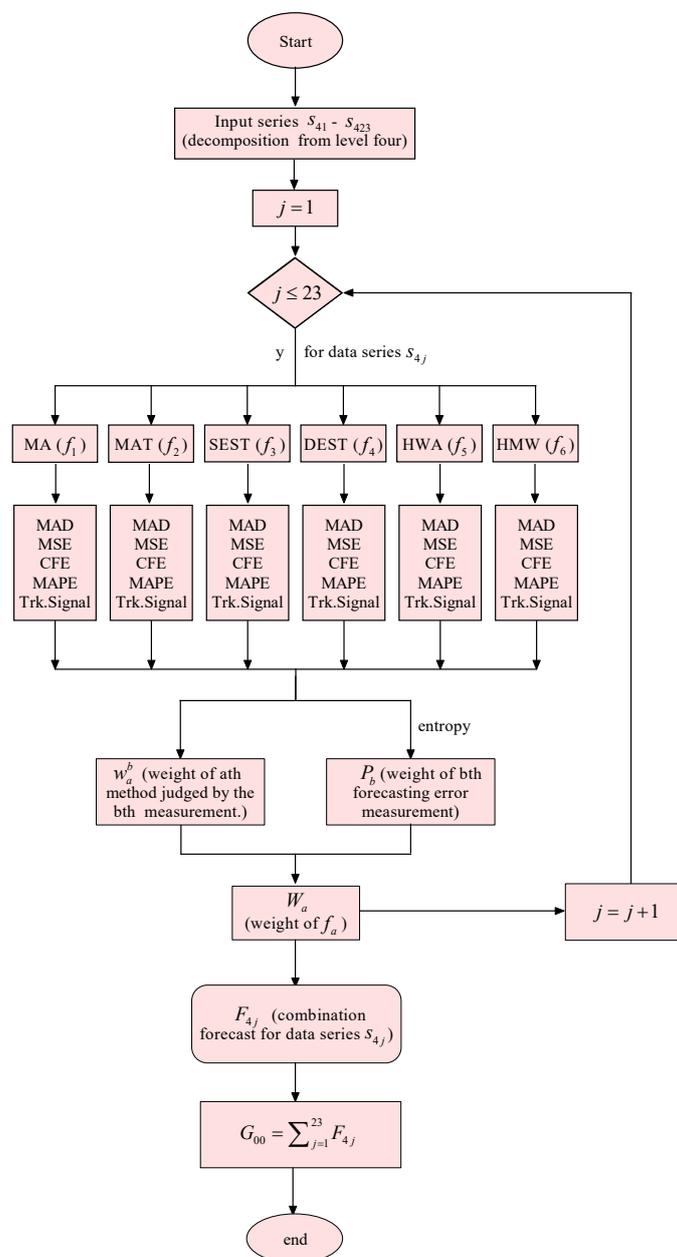


Figure 2. The Flow chart for the dynamic adaptive forecasting process.

The analysis framework for the work was as follows:

Step 1: The total electricity demand ( $s_{00}$ ) in the target society was decomposed into four industry levels (Figure 1), coded at level one to four:  $s_{1j}$ ,  $s_{2j}$ ,  $s_{3j}$ ,  $s_{4j}$  respectively (Table 1).

Step 2: For  $s_{ij}$ , ten time series groups were determined for each industry;  $s_{ij}^1 \sim s_{ij}^{10}$  (see the details in Section 3.1).

Step 3: For  $s_{4j}^l$ , the future electricity demand was forecast using the six individual models and CF. For  $s_{3j}^l$ , the future demand was forecast using the six individual models and CF and DCF4. For  $s_{2j}^l$ , the future demand was forecast using the six individual models and CF, DCF3, and DCF4. For  $s_{1j}^l$ , the future demand was forecast using the six individual models and CF, DCF2, DCF3, and DCF4. For  $s_{00}^l$ , the future demand was forecast using the six individual models and CF, DCF1, DCF2, DCF3, and DCF4.

Step 4: To compare and analyze the six individual models and the combination model for  $s_{ij}$ , an ANOVA and an orthogonal approach to solve the least squares equations were applied for which the AREs were regarded as the observation variables; therefore, there were 10 observations with 7 classification variables. For details, see Sections 4.1 and 4.2).

Step 5: The decomposition-based CF models were compared and analyzed. For  $s_{3j}$ , an ANOVA and an orthogonal approach were applied for which the AREs were regarded as the observation variables; therefore, there were 10 observations and 8 independent variables, as shown in Section 4.3.1. For  $s_{2j}$ , an analysis of variance was applied for which there were 9 independent variables, as shown in Section 4.3.2. For  $s_{1j}$ , an analysis of variance was applied for which there were 10 independent variables, as shown in Section 4.3.3. For  $s_{00}$ , an analysis of variance was applied for which there were 11 independent variables, as shown in Section 4.3.4.

Step 6: To further substantiate the forecasting results for DCF4, a comparison analysis with other forecasting methods such as ARIMA and ANN was conducted, as shown in Section 4.4.

## 4. Results and Discussion

This section gives step-by-step comparison analyses for the individual models, the combination forecasting model, and the decomposition-based combination forecasting method.

### 4.1. Analysis of the Individual Methods

As both MAPE and MSE are average values, the outliers are also averaged, and therefore it is often difficult to detect the worst prediction and the forecasting stability. Therefore, to compare the impact and significance of the forecasting accuracy for each forecast method, in this section, an ANOVA was conducted with the comparison tests on ARE.

The forecasting error distribution (shown in Figure 3) was represented using schematic box-and-whisker plots, each of which contained forecast errors from the individual and the CF methods.

With the mean, median, Q1 (first quartile), Q3 (third quartile), and the outliers (more above  $Q3 + 1.5 * IQR$ , where  $IQR = Q3 - Q1$ ) as the comparison indicators, the adaptability of the individual methods was estimated, with the Q1, Median, and Q3 columns respectively providing the lower quantile, median, and upper quantile. The best, the second best, the worst, and the second worst individual methods for each industry are shown in Table 3.

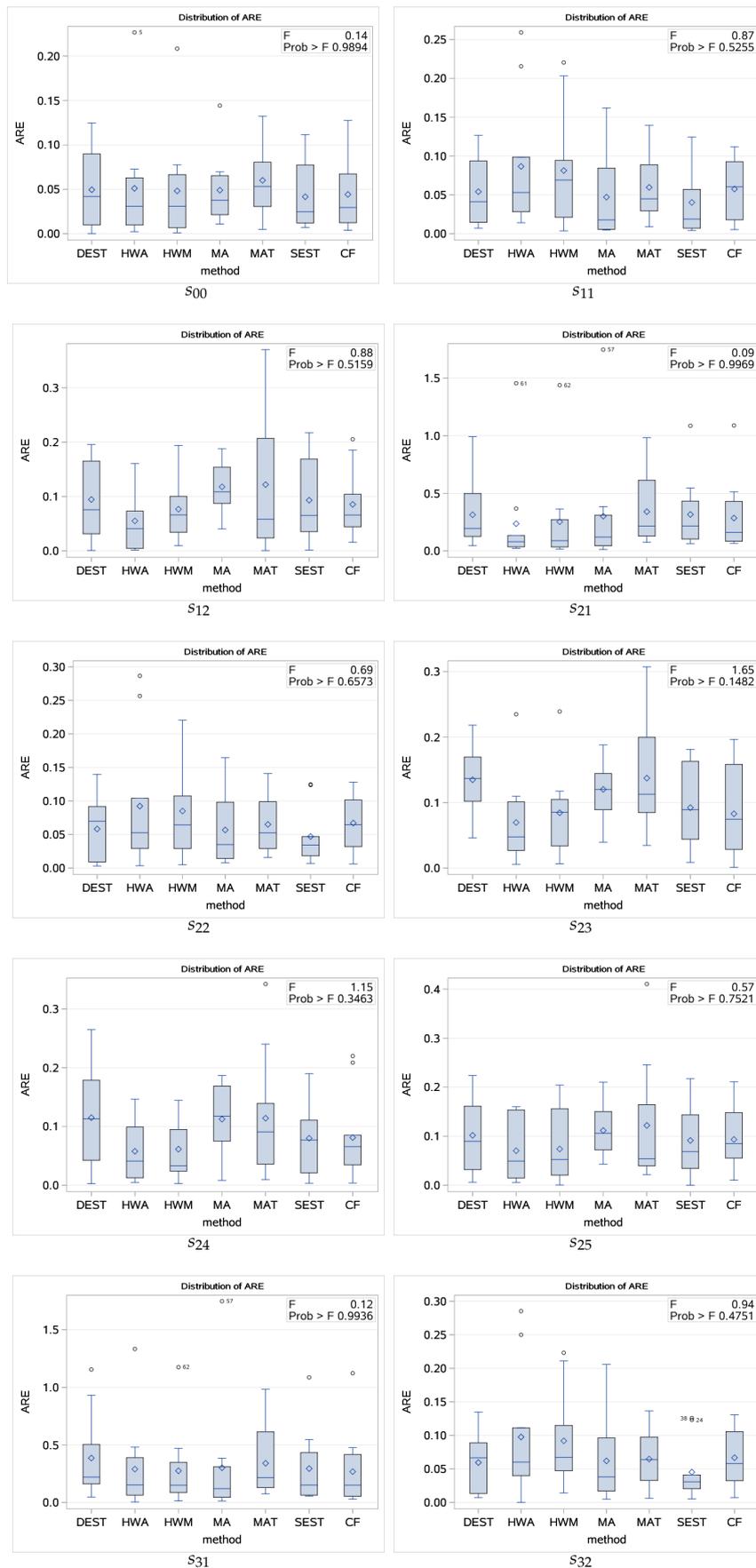


Figure 3. Cont.

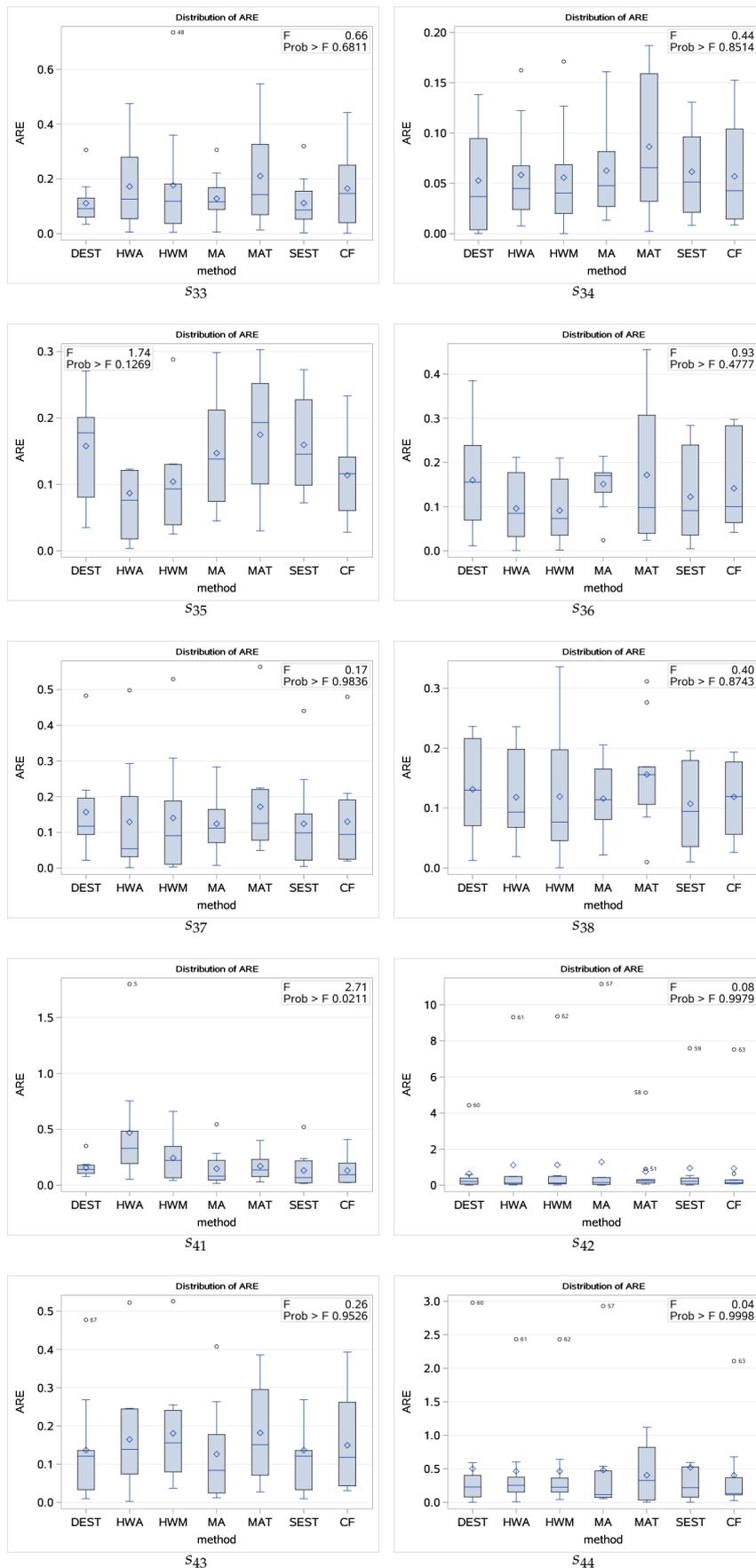


Figure 3. Cont.

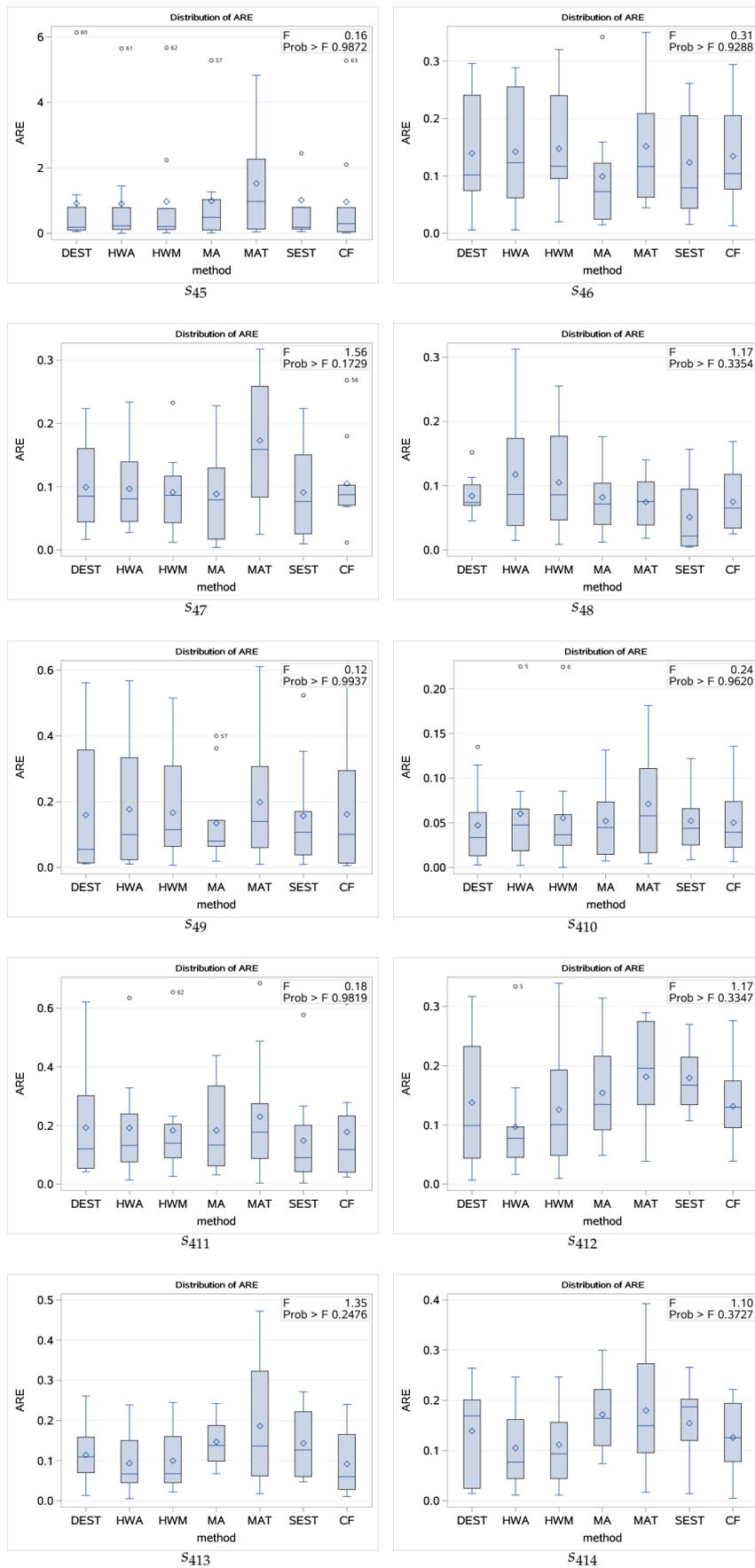


Figure 3. Cont.

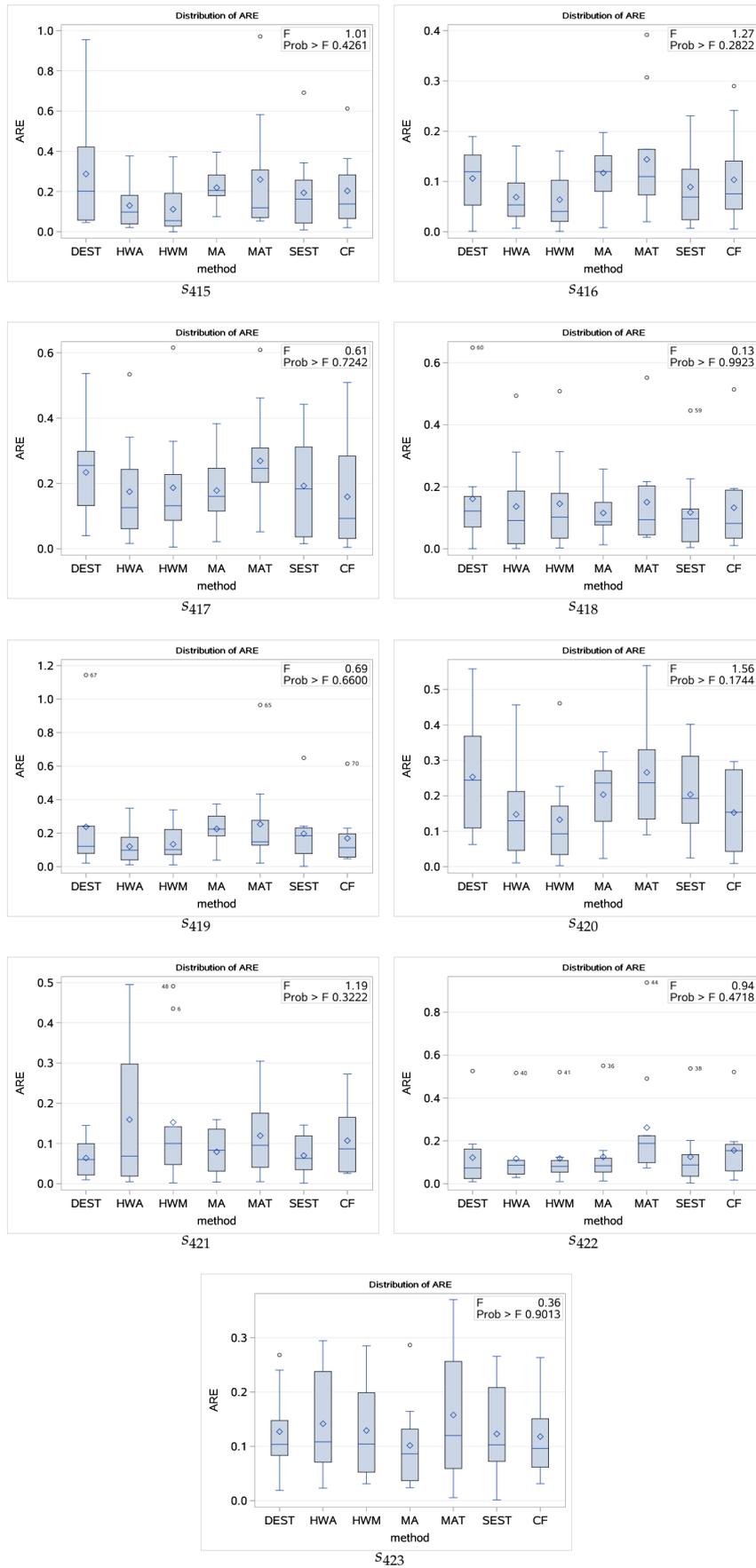


Figure 3. ARE distribution by ANOVA for every industry electricity demand.

Table 3. Method adaptability for each industry.

Data	The Best	The Second Best	The Worst	The Second Worst	CF Rank	Unsuitable for the Data
s <sub>00</sub>	SEST	DEST	HWA	HWM	2	
s <sub>11</sub>	SEST	DEST	HWA	HWM	4	
s <sub>12</sub>	HWA	HWM	WAT	SEST	3	
s <sub>21</sub>	HWA	HWM	MAT	DEST	3	
s <sub>22</sub>	SEST	DEST	HWA	HWM	5	
s <sub>23</sub>	HWA	HWM	MAT	DEST	3	
s <sub>24</sub>	HWA	HWM	MAT	DEST	4	
s <sub>25</sub>	HWA	HWM	MAT	DEST	4	
s <sub>31</sub>	MA	HWM	DEST	MAT	2	unsuitable
s <sub>32</sub>	SEST	DEST	HWA	HWM	3	
s <sub>33</sub>	DEST	SEST	MAT	HWA	4	
s <sub>34</sub>	HWA	HWM	MAT	HWA	4	
s <sub>35</sub>	HWA	HWM	MAT	SEST	3	
s <sub>36</sub>	HWM	HWA	MAT	DEST	5	
s <sub>37</sub>	SEST	MA	MAT	DEST	3	
s <sub>38</sub>	SEST	MA	MAT	DEST	2	
s <sub>41</sub>	DEST	SEST	HWA	HWM	1	
s <sub>42</sub>	DEST	MAT	HWA	HWM	1	unsuitable
s <sub>43</sub>	SEST	SEST	MAT	HWM	1	
s <sub>44</sub>	HWA	HWM	MAT	DEST	1	
s <sub>45</sub>	SEST	HWM	MAT	DEST	2	unsuitable
s <sub>46</sub>	MA	SEST	HWM	HWA	3	
s <sub>47</sub>	MA	SEST	MAT	DEST	6	
s <sub>48</sub>	SEST	MA	HWA	HWM	4	
s <sub>49</sub>	SEST	MA	MAT	DEST	3	
s <sub>410</sub>	HWM	HWA	MAT	SEST	4	
s <sub>411</sub>	SEST	MA	MAT	DEST	2	
s <sub>412</sub>	HWA	HWM	MAT	DEST	2	
s <sub>413</sub>	HWA	HWM	MAT	MA	1	
s <sub>414</sub>	HWA	HWM	MAT	MA	3	
s <sub>415</sub>	HWM	HWA	DEST	MAT	4	
s <sub>416</sub>	HWM	HWA	MAT	MA	6	
s <sub>417</sub>	MA	SEST	MAT	DEST	1	unsuitable
s <sub>418</sub>	SEST	MA	MAT	DEST	2	
s <sub>419</sub>	HWA	HWM	MAT	MA	3	unsuitable
s <sub>420</sub>	HWM	HWA	MAT	DEST	1	
s <sub>421</sub>	DEST	SEST	HWA	HWM	4	
s <sub>422</sub>	HWA	HWM	MAT	DEST	5	
s <sub>423</sub>	MA	DEST	MAT	HWA	1	

As can be seen in Figure 3, as the null hypothesis (same average estimates) was not able to be rejected at a 0.05 significance level, there were no major differences between the seven methods, which indicated that these classical time series forecasting methods could also be useful for forecasting electricity demand, which was consistent with the results in many previous studies. However, Figure 3 also indicates that MA also had good model fitness in some industries, and as the Q3 was lower than 10% and there were no outliers for  $s_{11}$ ,  $s_{22}$ ,  $s_{32}$ ,  $s_{34}$  and  $s_{410}$ , it also had a good forecasting performance. More importantly, an ARE lower than 10% is generally considered very satisfactory in engineering. Table 3 also shows that the MA had the best performance for  $s_{31}$ ,  $s_{46}$ ,  $s_{417}$  and  $s_{423}$  compared with the other individual methods. Unfortunately, MA also performed poorly; for  $s_{41}$ ,  $s_{413}$ ,  $s_{414}$  and  $s_{416}$  MA had the second worst performance compared with the other individual methods.

Karim and Alwi [31] concluded that the exponential smoothing technique had a better performance than the MA. Similarly, Figure 3 shows the effectiveness of SEST in some industries as the Q3 was lower than 0.1 and the ARE means were lower than 0.05 for  $s_{00}$ ,  $s_{11}$ ,  $s_{22}$ ,  $s_{32}$ ,  $s_{48}$  and  $s_{410}$ . As shown in the first column of Table 3, SEST was the best individual method 11 times (28.2%) and MA was the best individual method five times (12.8%).

Figure 3 also shows that DEST was also an effective method and was able to perform well in some industries; the ARE means were lower than 0.05 for  $s_{00}$ ,  $s_{11}$ ,  $s_{34}$ ,  $s_{49}$  and  $s_{410}$  and the Q3 was lower than 0.1 for  $s_{00}$ ,  $s_{11}$ ,  $s_{22}$ ,  $s_{32}$ ,  $s_{34}$ ,  $s_{410}$  and  $s_{421}$ . In addition, DEST had the best performance for  $s_{33}$ ,  $s_{41}$ ,  $s_{42}$  and  $s_{421}$  compared to the other individual methods.

Hussain et al. [5] proved that the Holt-Winters forecasting model could have robust results, and as shown in Table 3, this model outperformed some other methods in some industries. The MA, MAT, SEST, DEST, HWA and HWM models were respectively identified as the best individual methods five times (12.8%), zero times (0%), 11 times (28.2%), four times (10.3%), 14 times (35.9%) and five times (12.8%), with the HWA, and HWM together accounting for 49.2%. The MA, MAT, SEST, DEST, HWA, and HWM were respectively identified as the second best individual methods six times (15.4%), one time (2.6%), six times (15.4%), six times (15.4%), five times (12.8%) and 15 times (38.5%), the HWA and HWM together accounting for 51.3%. These results demonstrated that the Holt-Winters forecasting method outperformed the other individual methods in most cases.

However, it is somewhat difficult to explore the effectiveness and superiority of the Holt-Winters forecasting method as special attention needs to be paid to accuracy. In some cases, none of the methods was found to be suitable. As shown in Figure 3 and Table 3, however, for most of the rest, the Holt-Winters forecasting method produced accurate results, with the Q3 being lower than 0.1 and the median and means being lower than 0.05.

However, a further analysis of Figure 3 and Table 3 indicated that HWA and HWM were the worst individual methods eight times (20.5%) and one time (2.6%), and the second worst individual methods four times (10.25%) and nine times (23.1%). Therefore, as the Holt-Winters forecasting method was not available for all the time series data, it could be unhelpful for electricity demand forecasting in some industries.

From these analyses, it was concluded that as there was no one technique that was able to always obtain better forecasting results, different time series data sets require different individual methods. Therefore, there is no best forecasting model for the electricity data sets from different industries.

#### 4.2. Analysis of the Combination Forecasting Model

In the combination forecasting model proposed in this paper, inverse error weights were used so that the individual methods with less errors received an increased weight; that is, the advantages of each individual method were combined to improve the forecasting accuracy. Figure 3 shows that the CF had excellent forecasting performance as the Q3 was lower than 0.1 for  $s_{00}$ ,  $s_{11}$ ,  $s_{12}$ ,  $s_{22}$ ,  $s_{410}$  and the median was lower than 0.05 for  $s_{00}$ ,  $s_{34}$  and  $s_{410}$ .

Column 6 in Table 3 shows the CF Rank for the combination forecasting model compared to the other seven forecasting methods for each different industry. The CF model was respectively ranked from one to eight, 8, 7, 10, 9, 2, 3, and 0 times; that is, the combination forecasting method ranked in the top three 74.9% of the time, indicating its overall effectiveness. These results clearly indicated that forecasting combinations can yield unbiased forecasts even if the individual forecasts are biased. However, a more in-depth analysis of Figure 3 indicated that there were some industries in which none of the methods (including the combination model) were useful;  $s_{31}$ ,  $s_{45}$ ,  $s_{417}$ ,  $s_{419}$ ,  $s_{422}$ .

To determine the reason for this, the original electricity data from each industry were reanalyzed, from which it was found that there was a large gap between September 2009 and the other months in most industries. Therefore, it was concluded that a combination approach was not suitable when the data were highly volatile.

#### 4.3. Comparison Analysis of the Decomposition Experiments

To demonstrate the decomposition-based CF method, in this section, an ANOVA was used to analyze the electricity demand forecasting at the sub-subsector levels, the sub-sector levels, the sector levels, and the total aggregate level.

In the ANOVA, the 'B' following the parameter estimates indicated that the estimates were biased and did not represent a unique solution to the normal equations (Figure 4a). However, when using the orthogonal approach (ORTHOREG) to solve the least squares equations, more accurate estimates were produced compared to the other regression procedures. Figure 4b shows that the ORTHOREG fit achieved the same root mean square error (RMSE) as shown in the ANOVA table but avoided the

spurious singularities; therefore, the following analyses only considered the ANOVA. Furthermore, while there were no significant differences between DCF4 and the other methods when the other methods were superior to DCF4 in all situations, there were significant differences between DCF4 and other methods when DCF4 was superior to other methods in some situations. Figure 5a,b shows that the differences between DCF4 and the MAT groups (for  $s_{23}$ ) were marginal ( $p = 0.0959$ ), and that the differences between DCF4 and the MAT (for  $s_{35}$ ) were marginally significant ( $p = 0.0402$ ), which indicated that DCF4 may be more effective. A detailed analysis was then conducted, as outlined in the following.

Dependent Variable: F1

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	10	0.00315055	0.000315055	0.15	0.9986
Error	99	0.20257481	0.00204621		
Corrected Total	109	0.20572537			

R-Square	Coeff Var	Root MSE	F1 Mean
0.015314	88.63001	0.045235	0.051038

Parameter	Estimate		Standard Error	t Value	Pr >  t
Intercept	0.0496731790	B	0.01430458	3.47	0.0008
method CF	-.0054778167	B	0.02022973	-0.27	0.7871
method DCF1	0.0051766858	B	0.02022973	0.26	0.7986
method DCF2	0.0058244871	B	0.02022973	0.29	0.7740
method DCF3	0.0082023300	B	0.02022973	0.41	0.6860
method DEST	-.0002001710	B	0.02022973	-0.01	0.9921
method HWA	0.0013328606	B	0.02022973	0.07	0.9476
method HWM	-.0014874089	B	0.02022973	-0.07	0.9415
method MA	-.0006536303	B	0.02022973	-0.03	0.9743
method MAT	0.0103799429	B	0.02022973	0.51	0.6090
method SEST	-.0080833798	B	0.02022973	-0.40	0.6903
method DCF4	0.0000000000	B	.	.	.

(a) PROC GLM

Dependent Variable: F1

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	10	0.003150554	0.0003150554	0.15	0.9986
Error	99	0.2025748118	0.0020462102		
Corrected Total	109	0.2057253657			

Root MSE	R-Square
0.0452350552	0.015314368

Parameter	DF	Parameter Estimate	Standard Error	t Value	Pr >  t
Intercept	1	0.04967317902272	0.0143045805	3.47	0.0008
(method='CF')	1	-0.00547781674781	0.0202297317	-0.27	0.7871
(method='DCF1')	1	0.00517668575842	0.0202297317	0.26	0.7986
(method='DCF2')	1	0.0058244871403	0.0202297317	0.29	0.7740
(method='DCF3')	1	0.00820233001192	0.0202297317	0.41	0.6860
(method='DEST')	1	-0.00020017104435	0.0202297317	-0.01	0.9921
(method='HWA')	1	0.00133286056826	0.0202297317	0.07	0.9476
(method='HWM')	1	-0.0014874088854	0.0202297317	-0.07	0.9415
(method='MA')	1	-0.00065363025978	0.0202297317	-0.03	0.9743
(method='MAT')	1	0.01037994293493	0.0202297317	0.51	0.6090
(method='SEST')	1	-0.00808337977798	0.0202297317	-0.40	0.6903
(method='DCF4')	0	0	.	.	.

(b) PROC ORTHOREG

Figure 4. The GLM procedure and the ORTHOREG procedure for  $s_{00}$ .

Dependent Variable: F1

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	8	0.0488046698	0.0061005837	1.31	0.2529
Error	81	0.3786498901	0.00467469		
Corrected Total	89	0.4274545599			

Root MSE	0.0683717047
R-Square	0.1141751062

Parameter	DF	Parameter Estimate	Standard Error	t Value	Pr >  t
Intercept	1	0.08587496397816	0.0216210314	3.97	0.0002
(method='CF')	1	-0.0029831317626	0.0305767559	-0.10	0.9225
(method='DCF3')	1	0.00219893329617	0.0305767559	0.07	0.9428
(method='DEST')	1	0.04900803854892	0.0305767559	1.60	0.1129
(method='HWA')	1	-0.01628373269125	0.0305767559	-0.53	0.5958
(method='HWM')	1	-0.00145077170975	0.0305767559	-0.05	0.9623
(method='MA')	1	0.03429334644064	0.0305767559	1.12	0.2654
(method='MAT')	1	0.05151161281957	0.0305767559	1.68	0.0959
(method='SEST')	1	0.00640728664371	0.0305767559	0.21	0.8345
(method='DCF4')	0	0	.	.	.

(a)  $s_{23}$

Dependent Variable: F1

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	0.0753450371	0.0107635767	1.77	0.1068
Error	72	0.4380333079	0.0060837959		
Corrected Total	79	0.513378345			

Root MSE	0.0779986919
R-Square	0.1467631774

Parameter	DF	Parameter Estimate	Standard Error	t Value	Pr >  t
Intercept	1	0.10211039458111	0.0246653521	4.14	<.0001
(method='CF')	1	0.0117744205932	0.0348820755	0.34	0.7367
(method='DEST')	1	0.05590507774972	0.0348820755	1.60	0.1134
(method='HWA')	1	-0.01515159555613	0.0348820755	-0.43	0.6653
(method='HWM')	1	0.00211776701587	0.0348820755	0.06	0.9518
(method='MA')	1	0.04510189246904	0.0348820755	1.29	0.2002
(method='MAT')	1	0.07288908472264	0.0348820755	2.09	0.0402
(method='SEST')	1	0.05756160301778	0.0348820755	1.65	0.1033
(method='DCF4')	0	0	.	.	.

(b)  $s_{35}$

Figure 5. Parameter Estimate.

#### 4.3.1. Forecasting for the Sub-Subsector Electricity Demand

Based on Equation (2), the sub-subsector electricity demand was forecast using a decomposition-based CF at level 4.

$$G_{3j}^4 = \sum_v F_{4v}, \quad j = 1, \dots, 8,$$

where  $G_{3j}^4$  was the forecast for industry  $s_{3j}$  using the decomposition-based CF at level 4 and  $F_{4v}$  was the forecast for industry  $s_{4v}$  using the combination forecasting model. For example, the forecast using DCF4 for industry  $s_{31}$  was  $G_{31}^4 = F_{41} + F_{42} + F_{43} + F_{44} + F_{45}$ .

The ARE outputs using the eight methods at the sub-subsector level were compared using ANOVA, and the forecasting error distribution is shown in Figure 6. Figure 1 shows that there was no branch industry for  $s_{33}$ , indicating that no DCF4 existed for  $s_{33}$ -plus. The distribution of the forecast errors in Figure 6 shows the data discrepancy and biases as well as the outliers. For  $s_{32}$ -plus, DCF4 achieved the lowest ARE means compared to all other methods except SEST; however, SEST had two outliers. DCF4 also achieved the lowest median except for SEST and MA, but MA had an internal cap

that was far greater than for the DCF4. More importantly, 75% of the DCF4 AREs were lower than 10%, which was considered satisfactory.

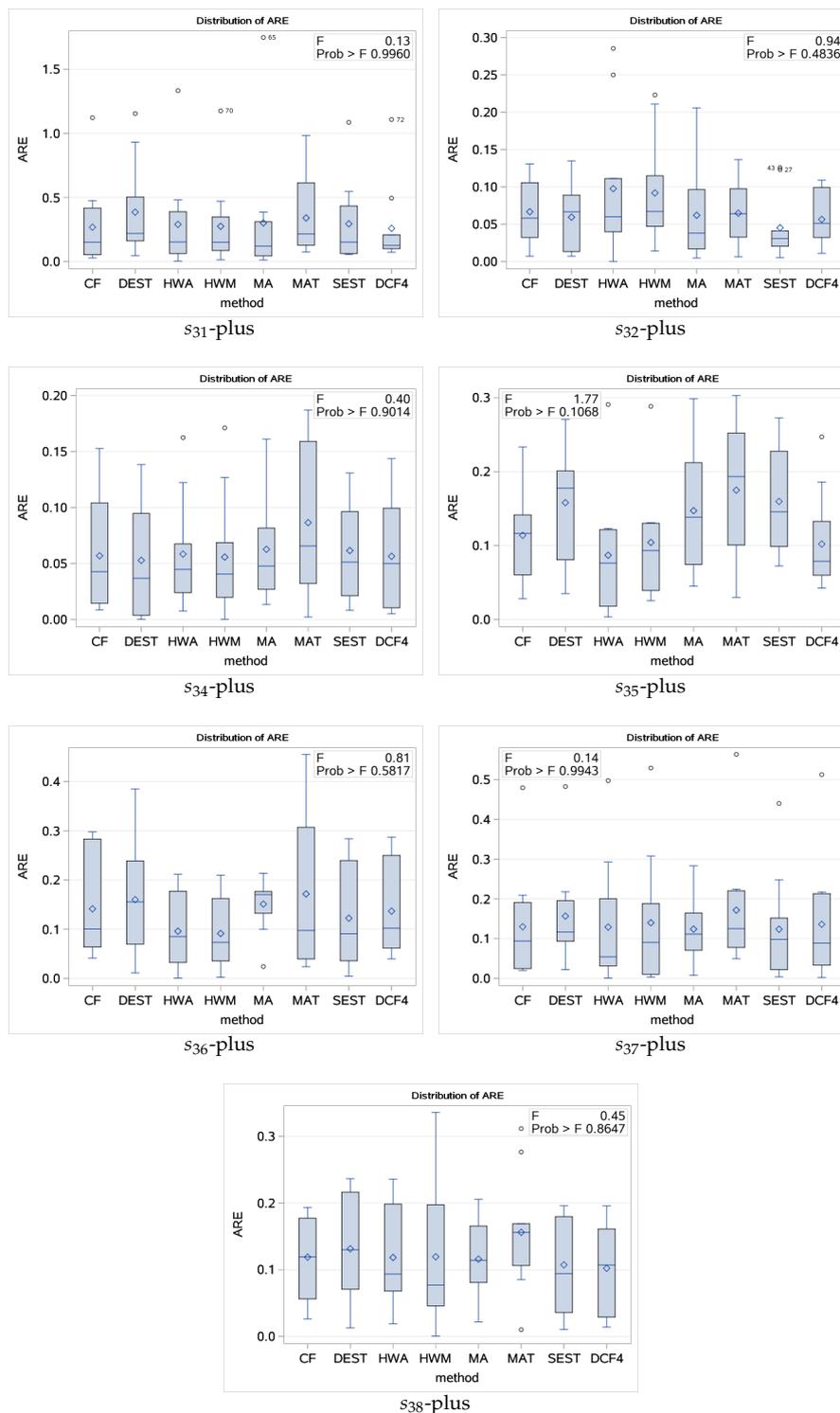


Figure 6. ARE distribution by ANOVA for sub-sector electricity demand.

For  $s_{38}$ -plus, DCF4 had the best forecasting performance as it achieved the lowest ARE means, Q1 and Q3, and had no outliers. For  $s_{34}$ -plus, the DCF4 had the third best performance after MA and DEST, as it achieved the lowest Q1 and means, and there were no outliers.

Even though all methods appeared to be unhelpful for industry  $s_{35}$ , DCF4 achieved the lowest means. However, no methods were suitable for the data associated with  $s_{31}$ -plus,  $s_{35}$ -plus,  $s_{36}$ -plus and  $s_{37}$ -plus. An analysis of the original electricity data indicated that there was a big gap between September 2009 and the other months in most industries.

From the above, it can be concluded that the forecasting effect tends to be better after adopting decomposition; however, a DCF with only one decomposition level is sometimes not suitable when the data are highly volatile

#### 4.3.2. Forecasting for the Sub-Sector Electricity Demand

Based on Equation (2), the sub-sector electricity demand was forecast using the decomposition-based CF at levels 4 and 3 using the following formula;

$$G_{2j}^4 = \sum_v F_{4v}, \quad j = 1, \dots, 5, \quad G_{2j}^3 = \sum_v F_{3v}, \quad j = 1, \dots, 5,$$

where  $G_{2j}^4$  and  $G_{2j}^3$  were respectively the forecasts for industry  $s_{2j}$  using the decomposition-based CF at levels 4 and 3, and  $F_{4v}$  and  $F_{3v}$  were the forecasts for industry  $s_{4v}$  and  $s_{3v}$  using the combination forecasting model.

Figure 7 compares the differences in the ARE output distributions for the nine sub-sector methods. From Figure 7, it can be seen that DCF4 achieved the lowest ARE means (approximately 0.05) compared to all other methods except for SEST, which had outliers in  $s_{22}$ -plus, 75% of the ARE outputs from DCF4 were lower than 10%, and DCF4 also had the lowest median, means and Q3 and no outliers compared to all other methods except for HWA and HWM, both of which have an outlier in  $s_{23}$ -plus. No methods, however, were found to be suitable for  $s_{21}$ . These results demonstrated that electricity demand forecasting uncertainties can be reduced using decomposition.

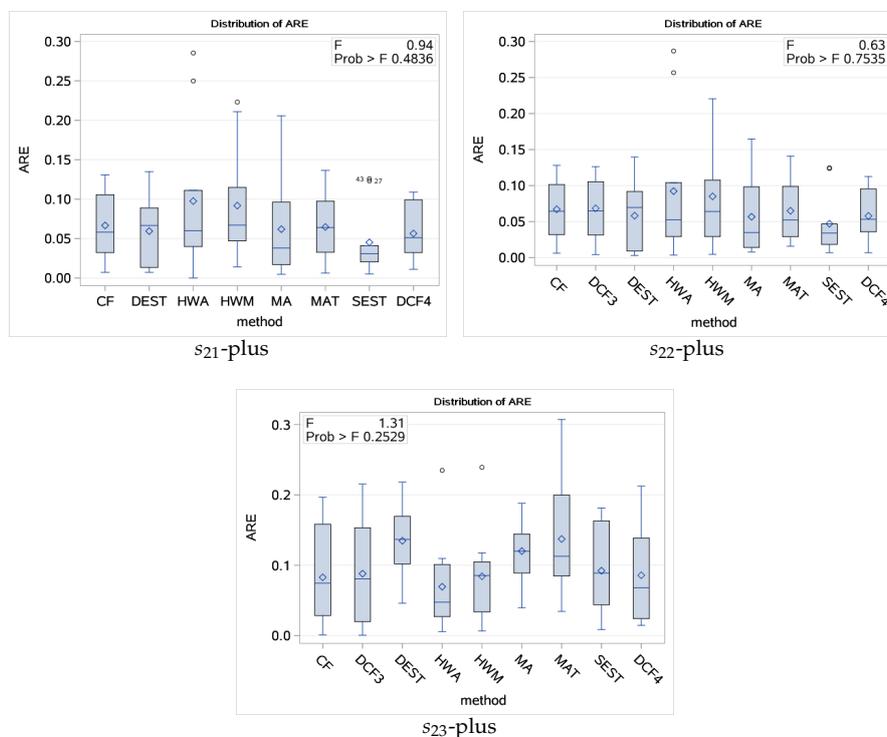


Figure 7. ARE distribution by ANOVA for sector.

### 4.3.3. Forecasting for Sector Demand

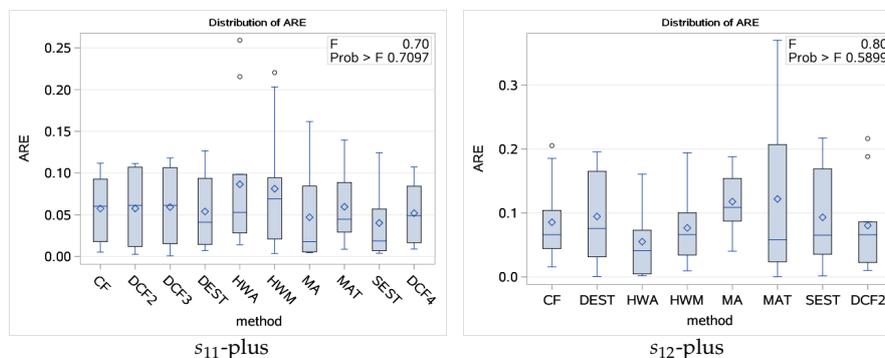
Based on Equation (2), the sector electricity demand was forecast using a decomposition-based CF at levels 4, 3 and 2:

$$G_{1j}^4 = \sum_v F_{4v}, \quad j = 1, 2, \quad G_{1j}^3 = \sum_v F_{3v}, \quad j = 1, 2,$$

$$G_{1j}^2 = \sum_v F_{2v}, \quad j = 1, 2,$$

where  $G_{2j}^4$ ,  $G_{2j}^3$  and  $G_{2j}^1$  were the respective forecasts for industry  $s_{1j}$  using DCF4, DCF3, and DCF2 respectively, and  $F_{4v}$ ,  $F_{3v}$  and  $F_{2v}$  were the forecasts for industry  $s_{4v}$ ,  $s_{3v}$  and  $s_{2v}$  using the combination forecasting model.

The ARE outputs for sector demand for the ten methods were compared using ANOVA, and the forecasting error distributions together with the schematic box-and-whisker plot representations are shown in Figure 8.  $s_{12}$  was residential electricity and was only decomposed from level 2 without DCF3 and DCF4, and  $s_{11}$  was decomposed from level 2 through to level 4.



**Figure 8.** ARE distribution by ANOVA for the sectors electricity demand.

A comparison in Figure 9 clearly shows that the decomposition-based CF model had the best forecasting performance for  $s_{11}$ , which was decomposed from three levels; however, the decomposition-based combination forecasting model had no advantages when forecasting  $s_{12}$  as there was only one decomposition level. DCF4 had the smallest maximum error when forecasting  $s_{11}$ , and DCF2 had a similar inner fence to the other methods. These results demonstrated that decomposition can adapt to the various characteristics of different data series.

Although Figure 8 indicates that HWA had the best performance for  $s_{12}$ -plus (which was consistent with the forecasting results in Hussain et al. [5]), for  $s_{11}$ -plus, both HWA and HWM had outliers, indicating that these methods were not suitable for the data set. In this case, for  $s_{11}$ -plus, DCF4 achieved the best performance because it had the lowest Q3 and means, the narrowest inner fence, and no high Q1 compared with all other methods except for SEST. Figure 9 also shows that DCF4 had smaller volatility than SEST.

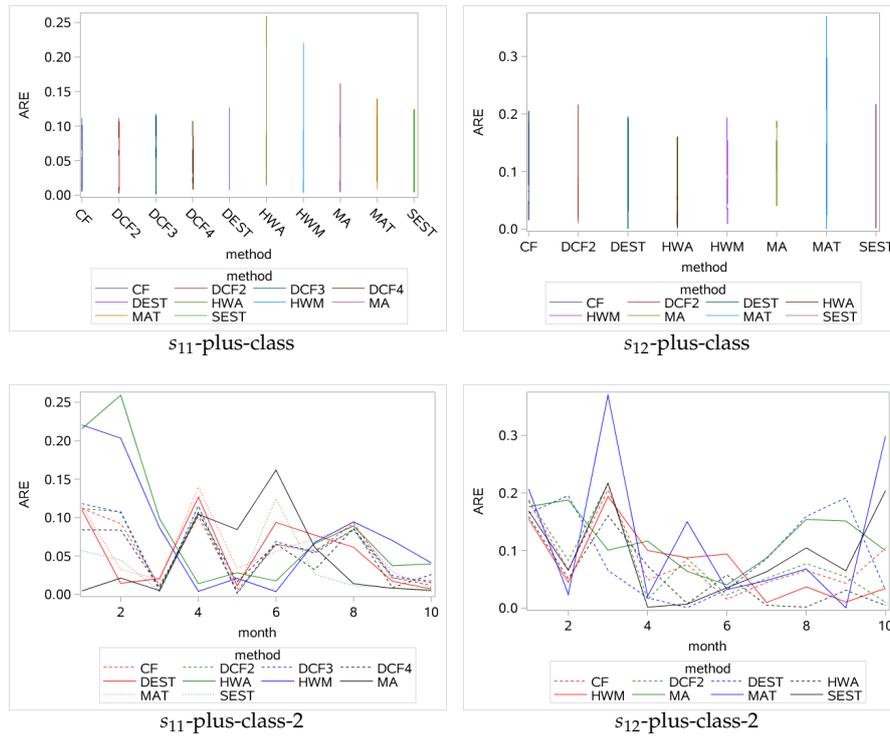


Figure 9. Sorted Data with Contrast Colors and Line Patterns Specified for  $s_{11}$  and  $s_{12}$ .

4.3.4. Forecasting for Total Electricity Demand

Governments are most interested in total electricity demand forecasting as this allows them to make future plans. Based on Equation (2), the electricity demand was forecast using the decomposition-based combination method described in Section 3.5.

Figure 10 shows that when only considering the means, the CF and SEST methods achieved better performance and the MAT and DCF3 methods had the worst performance. When considering the outliers, the CF, DCF4 and SEST methods achieved better performance, and the HWA, HWM, and MA methods contained outliers. For the degree of dispersion, the DCF4 achieved a lower  $Q1 - 1.5 * IQR$  and a lower  $Q3 + 1.5 * IQR$ , followed by SEST, CF, DCF1, DCF2. These results further supported the conclusion that decomposition-based combination forecasting at level 4 is better at forecasting total electricity demand.

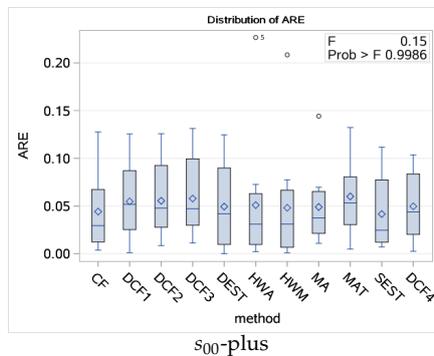


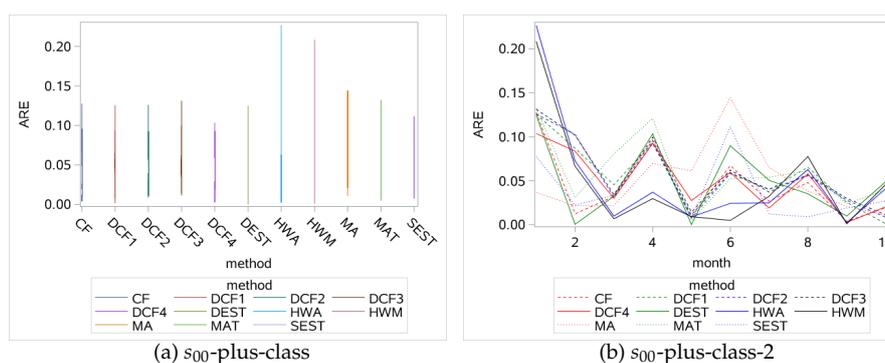
Figure 10. ARE distribution of by ANOVA for total electricity demand.

At any time, actual power generation must be able to accommodate at least a 15% increase in demand [32]. Figure 11a shows that even though the data are complex and uncertain, DCF4 performed excellently as all the AREs were lower than 10%. Figure 11b also shows that with an increase in the

learning period, the decomposition-based combination forecasting model errors tended to decrease, and the DCF4 remained more stable under volatility compared with the other methods. Furthermore, as can be seen from the ARE outputs for all methods for  $s_{00}$  in Table 4, there were more minimum AREs for the DCF4 model than for the other models. Even though they had fewer errors sometimes, the individual methods tended to display some instability, with errors ranging from very small to high; for example, the HWA had forecasting errors for  $s_{00}$  that were close to zero but as high as 20%. By default, compared to “accuracy”, stability is particularly important for engineering practice.

**Table 4.** The AREs of all the methods for  $s_{00}$ .

Month	MA	MAT	SEST	DEST	HWA	HWM	CF	DCF1	DCF2	DCF3	DCF4
25	0.0367	0.1323	0.0774	0.1246	0.2266	0.2084	0.1276	0.1255	0.1258	0.1312	0.1036
26	0.0214	0.0306	0.0223	0.0004	0.0727	0.0665	0.0124	0.0870	0.1027	0.1018	0.0837
27	0.0218	0.0805	0.0337	0.0322	0.0098	0.0067	0.0321	0.0454	0.0313	0.0343	0.0301
28	0.0697	0.1209	0.0969	0.1033	0.0370	0.0297	0.0957	0.0932	0.0926	0.0995	0.0927
29	0.0611	0.0048	0.0071	0.0001	0.0088	0.0091	0.0091	0.0075	0.0113	0.0143	0.0275
30	0.1442	0.0525	0.1115	0.0898	0.0243	0.0049	0.0672	0.0581	0.0590	0.0623	0.0595
31	0.0652	0.0540	0.0122	0.0505	0.0249	0.0323	0.0267	0.0398	0.0403	0.0371	0.0191
32	0.0383	0.0616	0.0090	0.0351	0.0629	0.0775	0.0480	0.0654	0.0556	0.0570	0.0578
33	0.0209	0.0239	0.0190	0.0098	0.0022	0.0010	0.0038	0.0253	0.0278	0.0299	0.0025
34	0.0109	0.0394	0.0269	0.0487	0.0409	0.0456	0.0193	0.0013	0.0086	0.0113	0.0203



**Figure 11.** Sorted Data with Contrast Colors and Line Patterns Specified for  $s_{00}$ .

#### 4.4. Comparison Analysis with Benchmark Methods

To further substantiate the forecasting results for DCF4, a comparison analysis was conducted with the benchmark methods ARIMA and ANN, both of which have been proven to be effective and highly precise [1,33]. For the comparison, the total electricity demand ( $s_{00}$ ) forecasting was executed by ARIMA and ANN.

The basic steps for ARIMA were as follows. First, the nonstationary series were transformed to stationary series using differencing, after which  $p$  (the order of the autoregressive part) and  $q$  (the order of the moving-average process) were determined using ACF and PACF and the optimal model determined from the check of the white noise residuals.

Three months' worth of electricity consumption data before the month to be predicted were taken as the input data for the ANN model, which had three input neurons. The number of hidden neurons were arranged to take values between 2 and 10, with the one output neuron being the consumption forecast for the subsequent month, and the leave-one-out cross-validation method used to evaluate the model. In addition, a rectified linear unit activation function was used as the hidden layer activation function, a linear function was used in the output layer, and an adaptive moment estimation was employed as the learning algorithm.

The analysis of variance results in Figure 12 indicated that there were no significant differences between the ANN, ARIMA, and DCF4 ( $p = 0.7544$ ). However, both Table 5 and Figure 13 show that the DCF4, ARIMA, and ANN forecasting results agreed. The ARE outputs for the ANN, ARIMA, and

DCF4 in Table 5 indicated that there were more minimum AREs for the DCF4 model than for either the ARIMA or ANN models, and that all the DCF4 errors were lower than 0.01 and tended to decrease with an increase in the learning period, as shown in Figure 13, all of which illustrated that the DCF4 was more stable, even when there was not enough data.

Table 5. The AREs of the ANN, ARIMA, and DCF4 for  $s_{00}$ .

Month	ANN	ARIMA	DCF4
25	0.029848707	0.084465171	0.103625946
26	0.015499663	0.036508881	0.083654997
27	0.124882225	0.031200544	0.030069406
28	0.055943052	0.101710947	0.092687115
29	0.010492949	0.004772692	0.027535572
30	0.086425693	0.108288557	0.059520543
31	0.052094060	0.028256914	0.019058591
32	0.105269199	0.00370881	0.057769253
33	0.043279300	0.01268148	0.002516652
34	0.047201343	0.034264153	0.020293710

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	0.00079130	0.00039565	0.28	0.7544
Error	27	0.03751559	0.00138947		
Corrected Total	29	0.03830689			

R-Square	Coeff Var	Root MSE	ARE Mean
0.020657	73.88484	0.037276	0.050451

Source	DF	Type III SS	Mean Square	F Value	Pr > F
method	2	0.00079130	0.00039565	0.28	0.7544

Parameter	Estimate	Standard Error	t Value	Pr >  t
Intercept	0.0496731790	0.01178756	4.21	0.0003
method ANN	0.0074204403	0.01667013	0.45	0.6598
method ARIMA	-.0050873642	0.01667013	-0.31	0.7626
method DCF4	0.0000000000	.	.	.

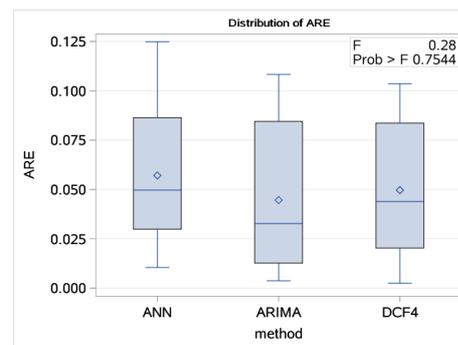


Figure 12. analysis of variance of ANN, ARIMA, and DCF4.

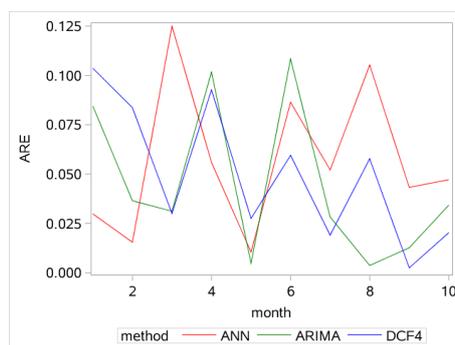


Figure 13. Sorted Data with Contrast Colors Specified for AREs by ANN, ARIMA, and DCF4.

### 5. Conclusions

With a focus on both prediction accuracy and stability, this paper developed a method for total electricity demand forecasting for nonlinear, nonstationary series. Using dynamic adaptive entropy-based weighting, a dynamic adaptive model was developed for total electricity demand forecasting that could be dynamically adapted to various forecast models through combination and to the various data characteristics through decomposition.

In this work, the total electricity demand in the target society was first decomposed based on industry categories, after which each component was forecast using a combination model that was developed using the entropy method, from which it was found that decomposition together with combination forecasting was able to improve forecasting accuracy.

The analyses of the individual methods in specific environments found that HWA was suitable for most data sets due to its varied applications; however, no significant differences were found between the forecast models based on the ANOVA. DCF4 was found to perform better than either the ARIMA or the ANN, especially when there was only small data and no need to handle outliers. From the schematic box-and-whisker plots and with the means, median, Q1, Q3, and outliers as the comparison indicators, it was found that the dynamic adaptive forecasting model performed the best, was highly accurate, had formidable forecasting ability, and would therefore be the best choice for complex, variable data. It was also concluded that if the raw data outliers were omitted, all methods analyzed here would perform better. It was also found that the large gap between September 2009 and the other months for most industries significantly affected the forecasting.

Although this paper was focused on the forecasting of monthly series data sets, the methods shown here could also be applied to quarterly data. As the proposed model performed well in forecasting one period, it could be used to develop electricity generation plans and associated energy policies. To further improve the methods, the forecast periods could be extended, and classical forecasting models and modern intelligent forecasting models combined to develop combination models for every industry; therefore, more work is necessary to fully refine the methodology.

**Author Contributions:** Z.H. designed the research, prepared the funding acquisition and reviewed the whole paper. J.M. investigated the area, prepared the original writing draft and edited latter changes. L.Y. analysed the data and carried out the experiments. X.L. revised the paper prior to the submission. M.P. provided the original data.

**Funding:** This research is funded by Basic Research Project of Philosophy and Social Sciences in Sichuan (Grant No. Xq14B04).

**Acknowledgments:** We would like to thank the two anonymous reviewers for their careful review and insightful recommendations. Especially, we appreciate the editors hard work for publishing this study.

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

## References

1. Laouafi, A.; Mordjaoui, M.; Haddad, S.; Boukelia, T.E.; Ganouche, A. Online electricity demand forecasting based on an effective forecast combination methodology. *Electr. Power Syst. Res.* **2017**, *148*, 35–47. [[CrossRef](#)]
2. Rallapalli, S.R.; Ghosh, S. Forecasting monthly peak demand of electricity in India-A critique. *Energy Policy* **2012**, *45*, 516–520. [[CrossRef](#)]
3. Hongye Guo, Q.C.; Qing Xia, C.K.; Zhang, X. A monthly electricity consumption forecasting method based on vector error correction model and self-adaptive screening method. *Int. J. Electr. Power Energy Syst.* **2018**, *95*, 427–439.
4. Li, W.; Yang, X.; Li, H.; Su, L. Hybrid Forecasting Approach Based on GRNN Neural Network and SVR Machine for Electricity Demand Forecasting. *Energies* **2017**, *10*, 44. [[CrossRef](#)]
5. Hussain, A.; Rahman, M.; Memon, J.A. Forecasting electricity consumption in Pakistan: The way forward. *Energy Policy* **2016**, *90*, 73–80. [[CrossRef](#)]
6. Li, W.; Bao, Z.; Huang, G.H.; Xie, Y.L. An Inexact Credibility Chance-Constrained Integer Programming for Greenhouse Gas Mitigation Management in Regional Electric Power System under Uncertainty. *J. Environ. Inform.* **2018**, *31*, 111–122. [[CrossRef](#)]
7. Brown, R.G. Exponential Smoothing. *Encycl. Oper. Res. Manag. Sci.* **2013**, *11*, 537–539. [[CrossRef](#)]
8. Box, G.E.P.; Tiao, G.C. Intervention Analysis with Applications to Economic and Environmental Problems. *Publ. Am. Stat. Assoc.* **1975**, *70*, 70–79. [[CrossRef](#)]

9. Akaike, H. Canonical Correlation Analysis of Time Series and the Use of an Information Criterion. *Math. Sci. Eng.* **1976**, *126*, 27–96. [[CrossRef](#)]
10. Kumar, U.; Jain, V.K. Time series models (Grey-Markov, Grey Model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in India. *Energy* **2010**, *35*, 1709–1716. [[CrossRef](#)]
11. Braun, M.R.; Altan, H.; Beck, S.B.M. Using regression analysis to predict the future energy consumption of a supermarket in the UK. *Appl. Energy* **2014**, *130*, 305–313. [[CrossRef](#)]
12. Khotanzad, A.; Hwang, R.C.; Abaye, A.; Maratukulam, D. An adaptive modular artificial neural network hourly load forecaster and its implementation at electric utilities. *IEEE Trans. Power Syst.* **1995**, *10*, 1716–1722. [[CrossRef](#)]
13. Son, H.; Kim, C. Short-term forecasting of electricity demand for the residential sector using weather and social variables. *Resour. Conserv. Recycl.* **2015**, *123*, 200–207. [[CrossRef](#)]
14. Van der Meer, D.; Shepero, M.; Svensson, A.; Widén, J.; Munkhammar, J. Probabilistic forecasting of electricity consumption, photovoltaic power generation and net demand of an individual building using Gaussian Processes. *Appl. Energy* **2018**, *213*, 195–207. [[CrossRef](#)]
15. Burger, E.M.; Moura, S.J. Gated ensemble learning method for demand-side electricity load forecasting. *Energy Build.* **2015**, *109*, 23–34. [[CrossRef](#)]
16. Clemen, R.T. Combining forecasts: A review and annotated bibliography. *Int. J. Forecast.* **1989**, *5*, 559–583. [[CrossRef](#)]
17. Xiao, L.; Shao, W.; Yu, M.; Ma, J.; Jin, C. Research and application of a combined model based on multi-objective optimization for electrical load forecasting. *Energy* **2017**, *119*, 1057–1074. [[CrossRef](#)]
18. Li, S.; Goel, L.; Wang, P. An ensemble approach for short-term load forecasting by extreme learning machine. *Appl. Energy* **2016**, *170*, 22–29. [[CrossRef](#)]
19. Wang, J.; Zhu, W.; Zhang, W.; Sun, D. A trend fixed on firstly and seasonal adjustment model combined with the -SVR for short-term forecasting of electricity demand. *Energy Policy* **2009**, *37*, 4901–4909. [[CrossRef](#)]
20. Shao, Z.; Chao, F.; Yang, S.L.; Zhou, K.L. A review of the decomposition methodology for extracting and identifying the fluctuation characteristics in electricity demand forecasting. *Renew. Sustain. Energy Rev.* **2017**, *75*, 123–136. [[CrossRef](#)]
21. Zhang, J.; Wei, Y.; Li, D.; Tan, Z.; Zhou, J. Short term electricity load forecasting using a hybrid model. *Energy* **2018**, *158*, 774–781. [[CrossRef](#)]
22. Li, C.; Tao, Y.; Ao, W.; Yang, S.; Bai, Y. Improving forecasting accuracy of daily enterprise electricity consumption using a random forest based on ensemble empirical mode decomposition. *Energy* **2018**, *165*, 1220–1227. [[CrossRef](#)]
23. Chen, J.S.; Chen, W.G.; Sun, P. A Generalized Model for Wind Turbine Faulty Condition Detection Using Combination Prediction Approach and Information Entropy. *J. Environ. Inform.* **2018**, *32*, 14–24. [[CrossRef](#)]
24. Che, J.X.; Wang, J.Z. Short-term load forecasting using a kernel-based support vector regression combination model. *Appl. Energy* **2014**, *132*, 602–609. [[CrossRef](#)]
25. Tang, P.; Chen, D.; Hou, Y. Entropy method combined with extreme learning machine method for the short-term photovoltaic power generation forecasting. *Chaos Solitons Fractals* **2016**, *89*, 243–248. [[CrossRef](#)]
26. Holt, C. *Forecasting Trends and Seasonal by Exponentially Weighted Averages*; Office of Naval Research Memorandum: Seattle, WA, USA, 1957.
27. Gooijer, J.G.D.; Hyndman, R.J. 25 years of time series forecasting. *Int. J. Forecast.* **2006**, *22*, 443–473. [[CrossRef](#)]
28. Gelper, S.; Fried, R.; Croux, C. Robust forecasting with exponential and Holt-Winters smoothing. *J. Forecast.* **2007**, *29*, 285–300. [[CrossRef](#)]
29. Anderson, D.R.; Sweeney, D.J.; Williams, T.A.; Camm, J.D.; Martin, R.K. *An Introduction to Management Science: Quantitative Approaches to Decision Making*; West Pub. Co.: New York, NY, USA, 1994; p. 128.
30. San-Juan, J.F.; San-Martín, M.; Pérez, I.; López, R. Hybrid perturbation methods based on statistical time series models. *Adv. Space Res.* **2015**, *57*, 1641–1651. [[CrossRef](#)]
31. Karim, S.A.A.; Alwi, S.A. Electricity load forecasting in UTP using moving averages and exponential smoothing techniques. *Appl. Math. Sci.* **2013**, *7*, 4003–4014. [[CrossRef](#)]

32. Williams, H.P. Model Building in Mathematical Programming. *News Med.* **2013**, *264*, 1132–1138.
33. Hamzaçebi, C. Forecasting of Turkey's net electricity energy consumption on sectoral bases. *Energy Policy* **2007**, *35*, 2009–2016. [[CrossRef](#)]



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).