MDPI

Article

A Hybrid Short-Term Load Forecasting Model Based on a Multi-Trait-Driven Methodology and Secondary Decomposition

Yixiang Ma [1], Lean Yu [1,2,3,*] and Guoxing Zhang [4]

1   School of Economics and Management, Beijing University of Chemical Technology, Beijing 100029, China
2   WQ-UCAS Graduate School of Business, Binzhou Institute of Technology, Binzhou 256600, China
3   WQ-UCAS Joint Lab, University of Chinese Academy of Sciences, Beijing 100190, China
4   School of Management, Lanzhou University, Lanzhou 730000, China
*   Correspondence: yulean@amss.ac.cn

Abstract: To improve the prediction accuracy of short-term load series, this paper proposes a hybrid model based on a multi-trait-driven methodology and secondary decomposition. In detail, four steps were performed sequentially, i.e., data decomposition, secondary decomposition, individual prediction, and ensemble output, all of which were designed based on a multi-trait-driven methodology. In particular, the multi-period identification method and the judgment basis of secondary decomposition were designed to assist the construction of the hybrid model. In the numerical experiment, the short-term load data with 15 min intervals was collected as the research object. By analyzing the results of multi-step-ahead forecasting and the Diebold–Mariano (DM) test, the proposed hybrid model was proven to outperform all benchmark models, which can be regarded as an effective solution for short-term load forecasting.

Keywords: multi-trait-driven methodology; secondary decomposition; multiple periodicity patterns; hybrid model; energy forecasting

1. Introduction

As a clean energy carrier, electricity plays an essential role in global energy conservation and emission reduction. Accordingly, its demand is also increasing rapidly. Based on research from the International Renewable Energy Agency, electricity will become the central energy carrier and account for nearly 50% of final consumption by 2050. Therefore, the rational planning of the deployment schedule to meet the ever-increasing electricity demand is a challenging task in the power system, which requires maintaining an optimal balance between the generated and consumed energy [1,2]. However, affected by economic activities, weather conditions, and calendar effects, the time series of short-term load exhibits periodicity, nonlinearity, and mutability traits [3,4]. This brings uncertainty to power dispatch. Correspondingly, accurate forecasting of short-term load is helpful to balance electricity supply and demand [5], as well as improve energy efficiency and reduce grid operating costs [6]. Therefore, an accurate forecast of the short-term load is vital for power dispatching. To this end, different models have been constructed to improve the prediction performance of the short-term load. Overall, these models can broadly fall into three groups: (1) classical statistical models, (2) machine learning techniques, and (3) hybrid learning methodologies.

As for classical statistical methods, autoregressive (AR) model [7], exponential smoothing (ES) model [8], seasonal autoregressive integrated moving average (SARIMA) model [9], and kernel ridge regression (KRR) model [10] have been utilized by the electricity demand [11]. For example, Cai et al. [12] combined the support vector regression (SVR) model and the SARIMA model to predict short-term load. By contrast, Elamin and Fukushige [13]

incorporated exogenous and interaction variables into the SARIMA model to model load accurately. However, limited by strict statistical assumptions, classical statistical models often exhibit unsatisfactory performance when dealing with nonlinear load time series [14].

In the second category, different machine learning techniques have been proven to be effective in exploring nonlinear mapping relationships between input and output datasets, which are constructed to predict the short-term load, such as support vector regression (SVR) model [15], random vector functional link network (RVFL) model [16], artificial neural network (ANN) model [17], extreme learning machine (ELM) model [18], long short-term memory (LSTM) model [19], and bi-directional long short-term memory (Bi-LSTM) model [20]. For instance, Zhu et al. [21] employed the extreme gradient boosting (XGBoost) model and achieved a higher prediction accuracy for holiday load forecasting. Nevertheless, machine learning techniques also suffer from their limitations, such as parameter sensitivity and local optimal [14,22].

As for the third category, guided by the idea of "divide and conquer" [23], various hybrid learning methodologies have been developed to reduce modeling complexity and combine the advantages of different prediction models, thereby achieving better performance in forecasting accuracy [24]. In detail, different decomposition methods, including ensemble empirical mode decomposition (EEMD) [25], wavelet transform (WT) [26], variational modal decomposition (VMD) [27,28], and complete ensemble empirical mode decomposition adaptive noise (CEEMDAN) [29] are first extracted as regular modes from complex load time series. Then, various statistical methods or machine learning techniques are selected to predict each mode, which are further integrated to output the prediction of short-term load. For example, Alipour et al. [30] proposed a hybrid learning methodology that combines the WT and deep neural networks (DNNs) for load forecasting. From the existing literature, the hybrid learning methodologies exhibit superior prediction performance compared with classical statistical methods and machine learning techniques [31].

To effectively explore the inner factors of the data, Yu et al. [32] proposed the principle of data-trait-driven modeling, which requires that the selected model needs to match the intrinsic traits of the data. Based on this principle, different research frameworks have been proposed for energy forecasting and demonstrated their effectiveness [23,33]. However, the existing literature only focuses on a specific aspect of the data, such as memory trait or complexity trait, and fails to analyze the overall traits of the data. Considering that the short-term load is affected by economic activities and weather conditions, it not only exhibits main multi-period trait (such as daily and weekly periodicities), but also has nonlinearity and mutability traits [3,34]. Therefore, it is necessary to analyze the multi-trait of the short-term load data. Then, the appropriate hybrid model can be constructed to improve the forecasting accuracy. In addition, combined data decomposition methods such as secondary decomposition (SD) have been employed to decrease the interference of the energy data, which improve the prediction accuracy compared with the separate decomposition method [35–37]. Nevertheless, the aforementioned studies directly selected the high-frequency decomposition mode for secondary decomposition, which lacks the judgment basis for the implementation of SD.

In summary, different research frameworks have been constructed to predict short-term load, but there exist some research gaps that need to be further improved. Firstly, affected by economic activities and weather conditions, although the short-term load time series exhibits apparent multiple periodicities, the existence of nonlinearity and mutability traits lead to the high complexity of the short-term load data. Meanwhile, affected by multiple periodicity patterns and calendar effects [38], the period length of the short-term load data is neither constant nor integral, which increases the difficulty of analyzing periodicity patterns. Therefore, the multiple traits of the short-term load data need to be recognized first, and correspondingly extract modes with different periodicity patterns or data traits. Secondly, for the construction of SD, reasonable criteria are required to determine which modes need to be further decomposed to reduce modeling complexity. Additionally, given that each decomposed mode has different data traits, the appropriate

prediction model should be selected for different modes, thereby improving the prediction accuracy of short-term load.

Given the aforementioned background, this paper aims to propose a hybrid model based on a multi-trait-driven methodology and secondary decomposition for short-term load forecasting. Specifically, four steps were performed sequentially, i.e., data decomposition, secondary decomposition, individual prediction, and ensemble output, all of which were designed based on a multi-trait-driven methodology. Firstly, after analyzing the multi-trait (e.g., stationarity trait, nonlinearity trait, and periodicity trait, etc.) and recognizing the multiple periodicities of the short-term load data, the improved Multiple Seasonal-Trend decomposition using Loess (IMSTL) model was performed to decompose the short-term load series. Secondly, through detecting the data traits of each decomposition mode, the mode with high-level complexity was further decomposed by the improved CEEMDAN (ICEEMDAN) model for simplifying. Thirdly, the K-fold walk forward validation (FV) method was employed to determine the forecasting model for the decomposed modes with different data traits. Finally, the prediction result of the short-term load can be obtained by integrating the forecasting results of each decomposed mode. Furthermore, various benchmark models and multi-step-ahead forecasting (i.e., one step, two steps, and three steps) are performed for comparison purposes.

Different from existing literature, the scientific contributions of the proposed methodology can be summarized below.

(1)　Based on the multi-trait analysis of the short-term load data, an appropriate decomposition method was constructed to extract modes with different periodicity patterns or data traits;

(2)　By embedding the method of multiple periodicities recognition, the IMSTL model can effectively identify and extract the multi-period patterns of the data;

(3)　A suitable judgment criterion is proposed to determine whether the modes need to be further decomposed, thereby improving the performance of the hybrid model.

The paper is organized as follows: The proposed multi-trait-driven methodology is introduced in Section 2. The results of the experiments and further discussion are shown in Section 3. Section 4 presents the conclusions and future research directions.

2. Material and Methodology

The proposed hybrid model based on multi-trait-driven methodology and secondary decomposition is introduced in this section. Specifically, the research framework is described in Section 2.1. The details of each step are elaborated in Sections 2.2–2.5.

2.1. The Framework of the Proposed Model

To explore the multi-trait (e.g., multi-period trait, nonlinearity trait, mutability trait, etc.) of short-term load data and improve the prediction performance, this paper proposes a hybrid model based on a multi-trait-driven methodology and secondary decomposition. In detail, the main variation patterns of the short-term load data are first identified by performing a multi-feature analysis. Meanwhile, based on the data traits analysis and multi-period recognition, the IMSTL model with a rolling mechanism is utilized to extract modes with different periodicity patterns or data traits. Then, different trait testing methods are selected to explore each decomposed mode, while the mode with high-level complexity is further decomposed. Meanwhile, considering the inherent serial correlation of the time series, the FV is employed to select the prediction model for each mode with different data traits. At last, the forecasting results of each mode are ensembled by an efficient method, simple addition (ADD), to output the final prediction of the short-term load.

As illustrated in Figure 1, the four steps are performed sequentially, i.e., data decomposition, secondary decomposition, individual prediction, and ensemble output, all of which are designed based on a multi-trait-driven methodology. The details of each step are elaborated as follows.
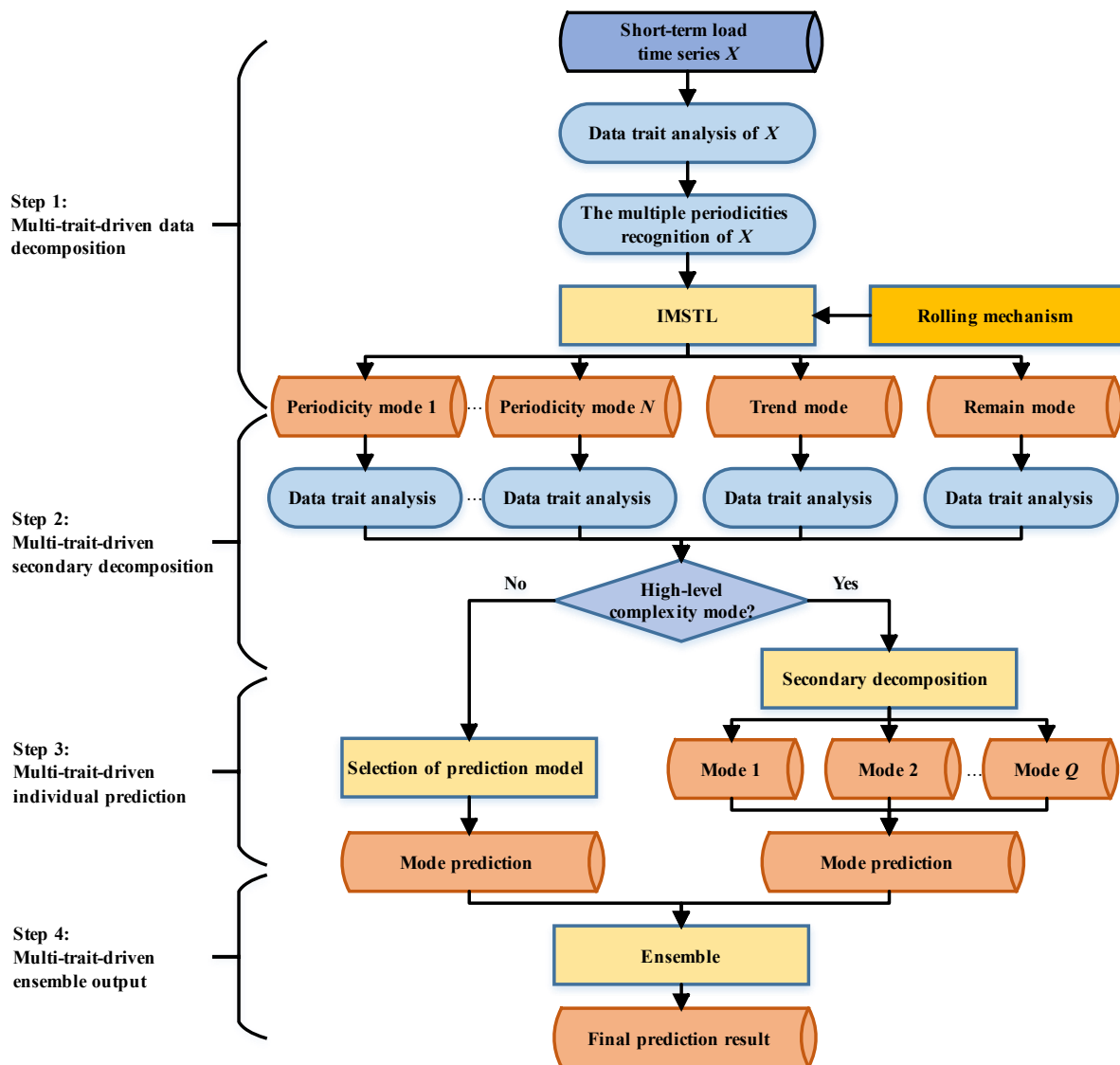
Figure 1. The framework of the proposed model.

Step 1: Multi-trait-driven data decomposition

In Step 1, the multi-trait and main variation pattern of the short-term load are analyzed by using different data trait metrics. Correspondingly, the nature traits such as nonstationarity trait, nonlinearity trait, and complexity trait, as well as pattern traits (i.e., periodicity trait, mutability trait, and randomicity trait) of the raw data can be explored. Then, given the potential multi-period trait of the data, an appropriate procedure is designed to recognize the multiple periodicities and determine the length of each periodicity. Finally, based on the data traits analysis and multi-period recognition, the raw data are decomposed by the IMSTL model into a series of modes with different periodicity patterns or data traits, as shown in Step 1 of Figure 1.

Step 2: Multi-trait-driven secondary decomposition

In multi-trait-driven secondary decomposition, the data traits of each decomposed mode are investigated to determine whether the modes need to be further decomposed. Specifically, the mode with the nonstationary trait, nonlinearity trait and complexity trait, as well as dominated by randomicity pattern can be judged to have high-level complexity, which needs to be further decomposed for simplifying. Considering that the ICEEMDAN model can effectively deal with the complexity data by solving the end effect issue. There-

fore, the ICEEMDAN model is selected to perform the SD of the mode with high-level complexity into different components, as exhibited in Step 2 of Figure 1.

Step 3: Multi-trait-driven individual prediction

In Step 3, various types of models are determined to predict each decomposed mode with different periodicity patterns or data traits. Accordingly, the classical statistical methods (i.e., ES, KRR and SARIMA), and machine learning techniques (i.e., SVR, ELM, RVFL and Bi-LSTM), are first employed as candidate models. At the same time, considering the inherent serial correlation of the time series data, the FV method is employed to assist in determining the prediction model for the decomposed modes with different data traits. Finally, based on the prediction performance of candidate models in K-data subsets (i.e., K = 5), the most appropriate model for each mode is selected for individual prediction.

Step 4: Multi-trait-driven ensemble output

In Step 4, the individual prediction of each decomposed mode is ensembled to output the final forecasting result of short-term load. Given the completeness of the decomposition method, that is, each decomposed mode can restore the initial data through linear addition. Therefore, to reduce computation time and avoid amplifying prediction errors, the simple addition (ADD) method is performed as the ensemble model. In detail, after performing Step 1 to Step 3, the prediction results of short-term load can be output by summing the forecasting results of each decomposed mode. Moreover, the rolling mechanism is performed to avoid misusing future information.

2.2. Multi-Trait-Driven Data Decomposition

In this section, the details of the multi-trait-driven data decomposition are given, along with the introduction of the corresponding methods. In detail, the data trait testing techniques are introduced in Section 2.2.1, and the decomposition method used in Step 1 is given in Section 2.2.2.

2.2.1. Data Trait Testing Technique

Affected by economic activities and the behaviors of household consumption, short-term load data show apparent multiple periodicities. For example, the power load data have a regular daily periodicity pattern, and the consumption of electricity on weekends and weekdays is quite different [39,40]. However, due to the superposition of calendar effects (i.e., moving festivals and holidays) and other factors such as weather conditions and geography, the short-term load data also exhibit high volatility and mutability, as well as a certain degree of randomness [34,38], which are typical complex data. Therefore, it is necessary to analyze multiple traits of short-term load data for subsequent model selection and prediction.

Based on the existing literature, the data traits mainly include two categories: nature traits and pattern traits. In detail, nature traits are used to analyze the regularity of the data generation process, such as nonstationarity trait, nonlinearity trait, and complexity trait. In contrast, pattern traits, namely periodicity trait, mutability trait, and randomicity trait, are used to analyze the impacts of different change patterns on the original data [41]. For example, if the data do not have a complexity trait and mainly exhibits periodic variation patterns, it means that the data have strong regularity. Correspondingly, the methods of the Augmented Dickey–Fuller (ADF) test [42], sample entropy [43], and the method of surrogate data [44] are used to recognize the nonstationarity trait, nonlinearity trait and complexity trait, respectively. As for the detection of the pattern traits, the dummy variables of average seasonal and structural change are used to extract the periodicity pattern and mutability pattern of the data, while the remainder of the data are used as a measure of randomicity pattern [32,41]. Meanwhile, the variance ratio is calculated to judge the importance of each pattern trait.

Furthermore, due to the existence of calendar effects (i.e., holidays occur on different days in different years), coupled with the influence of factors such as weather conditions and geography, the short-term load data may have multiple periodicity patterns and the

period lengths are not constant or integral [38]. Therefore, an appropriate procedure was designed in this paper to recognize the multiple periodicities of the data and determine the length of each periodicity. As shown in Figure 2, the periodogram method [45] was first used to obtain the possible period lengths of the short-term load data. Then, the 95% confidence level of these period lengths was set as a threshold to filter out the pseudo-periods caused by the random permutation process [46]. Subsequently, an efficient method, the Qs test [47], was used to statistically test the remaining periodicity patterns. Meanwhile, considering that spectral leakage may lead to multiple similar periodicity patterns [48], the K-nearest neighbors (KNNs) method was used to solve this problem by clustering the period lengths that pass the periodicity test. Finally, the center point of each cluster was selected as the multi-period length of short-term load data. More details of each technique can be found in the literature [32,38,41–48].
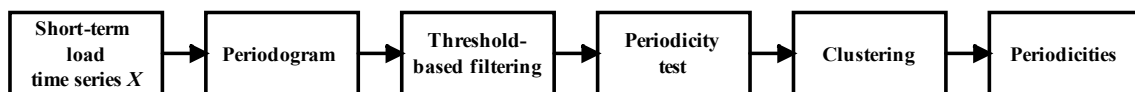

Figure 2. Flowchart for recognizing the multiple periodicities.

2.2.2. Decomposition Method

The short-term load has multiple data traits and change patterns, which is a typical complex data. Meanwhile, given the potential multi-period trait of the data, the candidate model needs to be able to efficiently extract modes with different periodicity patterns or data traits. Therefore, the MSTL model is chosen to decompose the short-term load data. As an extension of the Seasonal-Trend decomposition Procedure Based on Loess (STL) model [49], the MSTL model is capable of decomposing data with multiple periodicity patterns [50]. However, the multi-period length of the data must be determined in advance when using the MSTL model, and inappropriate extraction of multi-period patterns may ultimately degrade the prediction performance of the hybrid model. Therefore, the IMSTL model is proposed to efficiently recognize the multiple periodicities of the data and extract the multi-period modes accordingly, as shown in Figure 3.
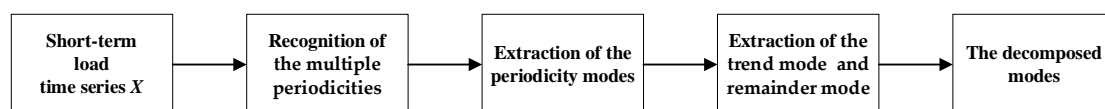

Figure 3. Flowchart of the proposed decomposition method.

In detail, the proposed IMSTL model can be performed by the following processes:

(1)　Based on the multi-period recognition procedure mentioned above, the length of each periodicity p_i, $(i = 1, \ldots, M)$, can be determined and sorted in ascending order.

(2)　Decompose the time series data $x(t)$ by STL to obtain their first periodicity mode with period length p_1:

$$r_1(t) = x(t) - Mode_1(t) \tag{1}$$

where $Mode_1(t)$ is the first periodicity mode, and $r_1(t)$ is the corresponding residual.

(3)　Calculate the *i*-th periodicity mode and the *i*-th residual until all periodicity modes are extracted:

$$r_{(i+1)}(t) = x(t) - \sum_{i=1}^{M} Mode_i(t) \tag{2}$$

(4)　Smooth $r_{(i+1)}(t)$ by loess to obtain the trend mode $T(t)$, and $x(t)$ can be expressed as:

$$x(t) = \sum_{i=1}^{M} Mode_i(t) + T(t) + R(t) \tag{3}$$

where $Mode_i(t)$ is the *i*-th periodicity mode, $T(t)$ is the trend mode, and $R(t)$ is the remainder mode.

2.3. Multi-Trait-Driven Secondary Decomposition

This section presents the judgment basis for implementing SD, which is utilized in Step 2 of Figure 1. Given existing literature, the implementation of SD can decrease the interference of the data, thereby improving the prediction accuracy as compared with the separate decomposition method [35–37]. However, there is no standard to decide which modes need to be further decomposed. To solve this problem, considering that the purpose of SD is to reduce the modeling complexity, the judgment basis can be formulated according to the data traits of each mode. Specifically, by performing data trait analysis on each decomposed mode obtained by the IMSTL model, the mode with the nonstationary trait, nonlinearity trait, and complexity trait, as well as dominated by randomicity pattern can be judged to have high-level complexity, which requires further decomposition for simplifying. Meanwhile, the ICEEMDAN model can effectively deal with the complexity data [51], which can be further improved by extending extreme value data to solve the end effect issue. Therefore, in the step of multi-trait-driven secondary decomposition, after analyzing the data traits of each mode, the ICEEMDAN model was used to decompose the mode with a high-level complexity.

2.4. Multi-Trait-Driven Individual Prediction

In Step 3 of Figure 1, the suitable model is selected to forecast each decomposed mode with different data traits. In detail, the classical statistical methods (e.g., ES, KRR and SARIMA), and machine learning techniques (e.g., SVR, ELM, RVFL and Bi-LSTM) are first employed as candidate models. Meanwhile, considering the inherent serial correlation of the time series data, the FV method [52] is employed to assist in determining the prediction model. Subsequently, by dividing the training dataset for each fold in chronological order, different candidate models are used to forecast each subset. Finally, according to the prediction performance of each model on the subset, the model that achieves the highest prediction accuracy is used to predict the corresponding mode. The detailed process of each prediction method can refer to the literature [8,10,15,16,18,20,53].

2.5. Intermittency-Trait-Driven Ensemble Output

In Step 4 of Figure 1, for the models that decompose the original short-term load data and the decomposed mode with high-level complexity, each mode can be linearly added to obtain the original data. Therefore, to reduce computational complexity and avoid amplifying prediction errors, the forecasting results of each decomposed mode are ensembled by the simple addition (ADD) approach. Accordingly, by linearly adding the predictions of each decomposed mode and quadratic decomposed component, the forecasting results of short-term load can be obtained. Furthermore, to avoid misusing the testing dataset information during decomposition, the rolling mechanism is performed [54].

3. Results

To demonstrate the effectiveness of the proposed model, the short-term load data with 15 min intervals was used in numerical experiments. Meanwhile, different benchmark models were chosen for the purpose of comparison. In particular, Section 3.1 explains the experimental design, the empirical results are analyzed in Section 3.2.

3.1. Experimental Design

In the numerical experiment, the short-term load data with 15 min intervals, collected from the Bonville Power Administration (BPA), were used as the research object. Given the data availability, the study period covers from 1 January to 7 January 2020, as presented in Figure 4. Meanwhile, according to the ratio of 7:3, the data were divided into a training dataset and a testing dataset. Among them, the training dataset was used for model selection and parameter estimation, and the testing dataset was applied to test the effectiveness of each model. The experiment of modeling and analysis were implemented on R 4.1.2 and the main libraries used include "forecast", "EEMDelm", and "seatests". Meanwhile,

as for the hyperparameters that need to be set by the user, the method of trial-and-error was selected to determine these parameters. Furthermore, the parameters of all benchmark models and the proposed model were set to be the same for a fair comparison.
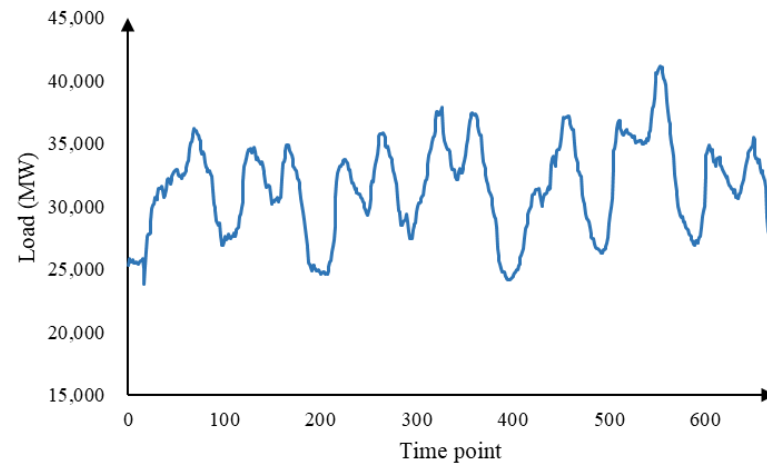
Figure 4. The short-term load data with 15 min intervals (MW).

Meanwhile, three popular metrics, namely root mean squared error (*RMSE*), mean absolute error (*MAE*), and mean absolute percentage error (*MAPE*), were calculated to measure the prediction accuracy of each model. The formulas of each metric are as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(x_t - \hat{x}_t)^2} \tag{4}$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|x_t - \hat{x}_t| \tag{5}$$

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{x_t - \hat{x}_t}{x_t}\right| \tag{6}$$

where x_t and $\hat{x}_t$ are the real value and the predicted value in t period ($t = 1, 2, \ldots, N$), respectively, N is the length of the data.

As for the selection of benchmark models, different single methods (e.g., ES, KRR, SARIMA, SVR, ELM, RVFL, and Bi-LSTM), coupled with different decomposition-ensemble models (e.g., IMSTL-SARIMA, IMSTL-ES, IMSTL-Bi-LSTM, and IMSTL-S) were selected to compare with the proposed hybrid model. Wherein, 'S' represents the multi-trait-driven individual prediction without SD. Furthermore, multi-step-ahead forecasting (i.e., one step, two steps and three steps) [55] and the Diebold–Mariano (DM) test [33] were utilized to demonstrate the effectiveness of the proposed model.

3.2. Experimental Results Analysis

As mentioned before, the experimental results are provided in this section. In detail, the multi-trait analysis of the short-term load data are presented in Section 3.2.1. Section 3.2.2 analyzes the data traits of each decomposed mode and performs the SD. The determination of the prediction model is reported in Section 3.2.3. The comparative analysis between the proposed hybrid model and the benchmark models is provided in Section 3.2.4.

3.2.1. Analysis of the Multi-Trait of the Short-Term Load Data

In this section, different nature traits and pattern traits metrics are calculated to analyze the multiple traits and main variation patterns of the training dataset. Accordingly, based on the data trait testing techniques, the corresponding results are listed in Table 1. Meanwhile, the symbol of the check mark indicates that the corresponding nature trait is possessed, otherwise it indicates that the corresponding nature trait is not possessed.

Table 1. Cyclicity test results of original time series data.

	Nature Trait			The Importance of Each Pattern Trait		
	Nonstationarity	Nonlinearity	Complexity	Periodicity	Mutability	Randomicity
The original data	×	√	√	0.575	0.163	0.262

According to Table 1, it can be found that the short-term load time series are typical complex data. In detail, for the testing results of nature trait, the short-term load data present nonstationarity trait, nonlinearity trait, and complexity trait, which indicates that the regularity of the data generation process is poor. Meanwhile, according to the importance of each pattern trait, periodicity is the main variation pattern of short-term load data, followed by randomicity and mutability patterns. Considering that the short-term load data mainly present the periodicity pattern, the multi-period recognition method was used to further identify the multiple periodicities and determine the length of each periodicity, as provided in Table 2.

Table 2. The results of multi-period recognition.

	Threshold-Based Filtering	Periodicity Test	Clustering
Possible periods	[19.200, 32.000, 43.636, 48.000, 53.333, 68.571, 80.000, 96.000, 120.000, 160.000, 240.000, 480.000]	[43.636, 48.000, 53.333, 96.000]	[48.323, 96.000]

Based on Table 2, four possible periods passed the periodicity test. However, there are some similar periodicity patterns (e.g., 43.636, 48.000 and 53.333), which may be caused by spectral leakage [48]. Furthermore, after clustering by the KNN method, the multi-period length of short-term load data was finally confirmed as 48.323 and 96.000, which are about half-day and one-day periods, respectively. Overall, affected by economic activities and weather conditions, although the short-term load series mainly presents the periodicity pattern, the irregular data generation process and variation pattern result in the high data complexity. Therefore, an appropriate decomposition model needs to be selected to extract the regular modes, thus improving the prediction accuracy of the short-term load series.

3.2.2. Data Decomposition Analysis

Considering that the short-term load data mainly presents the periodicity variation pattern and has multiple periodicities, the IMSTL model was utilized to extract modes with different periodicity patterns or data traits, as depicted in Figure 5.

According to Figure 5, it can be intuitively concluded that, compared with the original data, the regularity of the periodicity modes and the trend mode is better, while the change in the remainder mode is drastic. In order to determine which mode needs to be further decomposed, the data traits of each decomposed mode were analyzed, as can be seen in T.

From Table 3, it can be concluded that the periodicity mode, trend mode, and remainder mode have different nature traits and pattern traits. For example, although each periodicity mode is tested to have the nonstationary trait and nonlinearity trait, the dominant variation pattern of these modes is the periodicity pattern, with a pattern importance ratio of 0.857 and 0.932, respectively. Meanwhile, neither the periodicity mode nor the trend mode is detected to have the complexity trait, so these modes have strong regularity. Specific to the remainder mode, it not only has the nonstationarity trait, nonlinearity trait and complexity trait, but also its variation pattern is dominated by the randomicity pattern. Based on the data trait analysis results of each mode, the periodicity mode and the trend mode can be directly modeled to predict, while the remainder mode has high-level complexity and need to be further decomposed for simplifying. Therefore, the remainder mode was decomposed by the ICEEMDAN model to reduce the complexity.
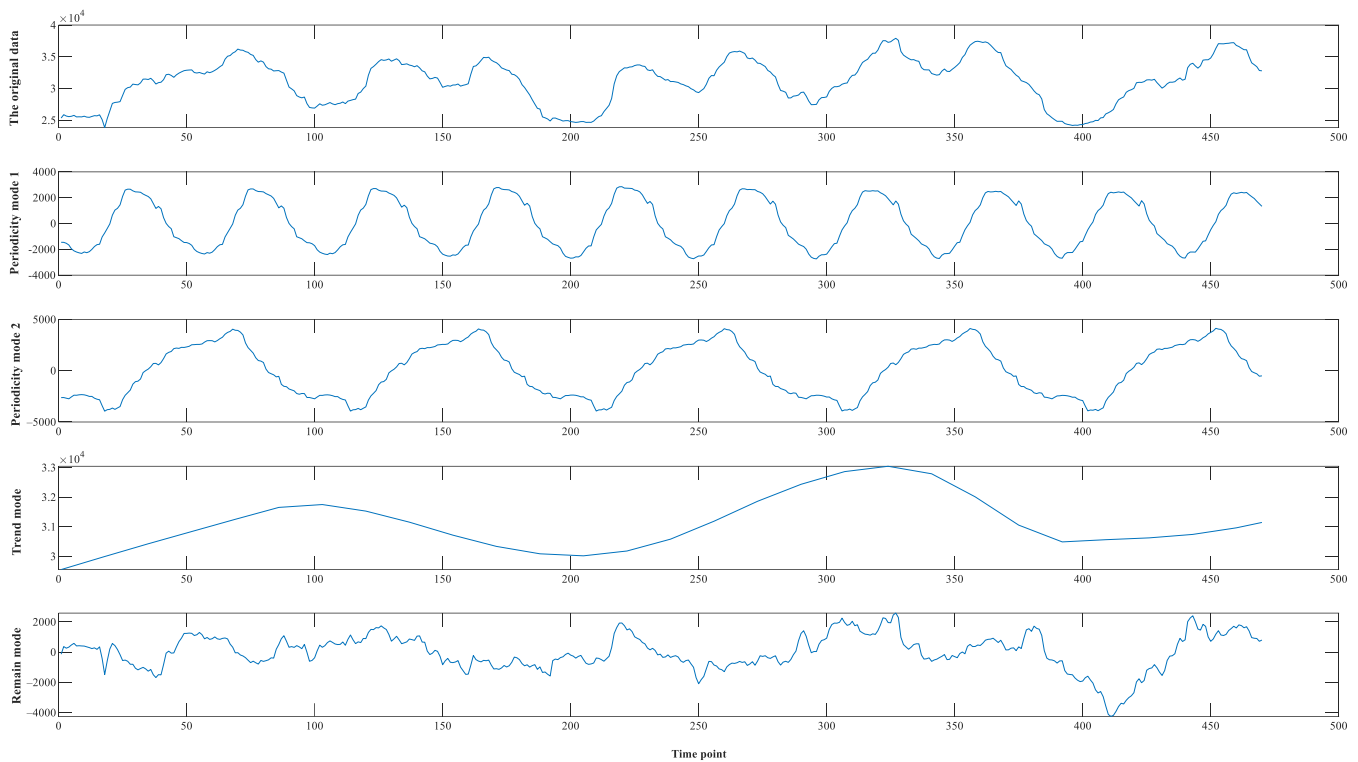
Figure 5. Decomposition modes obtained by IMSTL in the training dataset.

Table 3. Data trait testing of each mode.

	Nature Trait			The Importance of Each Pattern Trait		
	Nonstationarity	Nonlinearity	Complexity	Periodicity	Mutability	Randomicity
Periodicity mode 1	√	√	×	0.857	-	0.143
Periodicity mode 2	√	√	×	0.932	-	0.068
Trend mode	√	√	×	-	0.337	0.663
Remainder mode	√	√	√	-	0.096	0.904

3.2.3. Multi-Trait-Driven Prediction Model Selection

In this section, different candidate models and the FV method are used to determine the prediction model for each decomposed mode with different data traits. By calculating the evaluation metric values (e.g., *RMSE*, *MAE*, and *MAPE*) of the candidate models on each subset, the model with the highest prediction accuracy is selected to predict the corresponding mode. Specifically, the prediction performance of each candidate model is listed in Tables 4 and 5. Wherein, the prediction performance and average prediction performance of the candidate models on each subset i ($I = 1, 2, \ldots, 5$) are denoted by F_i and $\bar{F}$.

Table 4. Prediction performance of the candidate models on each subset.

Mode	Model	F1			F2			F3			F4			F5		
		MAPE	*RMSE*	*MAE*	*MAPE*	*RMSE*	*MAE*	*MAPE*	*RMSE*	*MAE*	*MAPE*	*RMSE*	*MAE*	*MAPE*	*RMSE*	*MAE*
Periodicity mode 1	Bi-LSTM	0.927	245.797	172.800	0.762	183.172	148.012	0.426	198.837	152.262	0.288	286.859	169.492	1.031	208.266	147.918
	ELM	0.381	131.336	96.408	0.837	199.056	143.797	0.494	164.562	130.427	0.275	165.683	125.588	1.355	158.329	121.387
	KRR	1.100	248.675	178.268	0.570	267.152	198.993	0.231	219.067	159.969	0.267	223.608	158.665	1.400	194.533	142.468
	SVR	1.206	273.146	212.811	0.643	288.488	226.164	0.336	240.510	192.332	0.295	233.485	174.498	1.497	199.946	155.590
	RVFL	0.435	119.259	90.766	0.765	200.696	144.254	0.416	152.906	122.954	0.285	161.419	128.649	1.437	156.563	124.844
	ES	0.157	206.243	149.063	0.898	331.332	250.421	1.198	304.998	241.116	0.490	280.956	221.201	2.099	278.239	225.278
	SARIMA	0.217	206.993	155.856	0.226	171.155	105.006	0.048	46.834	35.228	0.049	50.458	39.188	0.034	29.947	21.969
Periodicity mode 2	Bi-LSTM	0.407	328.710	213.988	0.348	303.087	222.715	0.104	284.487	188.135	0.328	307.913	245.182	0.358	278.914	227.661
	ELM	0.346	246.446	170.164	0.195	198.572	141.157	0.101	171.217	119.787	0.176	185.633	135.494	0.154	163.948	125.255
	KRR	0.477	269.428	184.659	0.285	282.065	210.790	0.136	220.858	151.651	0.237	238.050	175.000	0.210	224.272	163.845
	SVR	0.530	294.199	209.836	0.340	307.680	238.863	0.148	240.274	171.053	0.278	276.830	210.814	0.246	256.293	196.705
	RVFL	0.384	283.444	188.571	0.189	197.291	141.284	0.103	168.056	118.631	0.171	182.971	132.530	0.156	162.383	122.787
	ES	0.138	210.038	153.743	0.272	281.519	211.148	0.132	245.754	170.368	0.260	258.198	191.042	0.238	261.992	197.770
	SARIMA	0.173	216.284	153.938	0.070	157.681	73.852	0.004	7.010	5.479	0.006	5.966	4.875	0.006	6.499	4.946
Trend mode	Bi-LSTM	0.003	95.910	79.191	0.001	24.979	21.014	0.006	307.894	211.088	0.005	179.874	153.440	0.001	29.363	16.417
	ELM	0.003	2827.887	235.650	0.000	18.281	2.133	0.000	0.232	0.113	0.000	0.141	0.088	0.000	0.074	0.058
	KRR	0.001	66.677	66.139	0.001	74.555	73.408	0.001	51.457	50.941	0.001	69.497	69.406	0.001	66.456	66.385
	SVR	0.002	113.963	112.845	0.002	153.271	153.198	0.002	178.561	178.312	0.002	181.832	181.798	0.002	183.460	183.378
	RVFL	0.004	3524.889	291.528	0.000	44.607	5.317	0.000	0.244	0.116	0.000	0.142	0.079	0.000	0.076	0.058
	ES	0.001	19.060	17.462	0.000	2.661	0.590	0.001	32.204	31.108	0.001	38.091	34.116	0.000	9.852	8.427
	ARIMA	0.012	415.966	376.480	0.030	939.724	924.913	0.028	1043.801	873.919	0.034	1242.029	1056.441	0.014	464.305	427.642
Remainder mode	Bi-LSTM	0.616	193.057	154.081	0.280	179.005	144.683	0.354	199.253	136.666	0.636	245.584	188.516	0.174	325.771	252.912
	ELM	1.705	278.128	227.273	0.434	410.502	244.966	0.534	292.476	213.244	0.653	390.716	239.349	0.218	454.785	327.003
	KRR	1.893	251.111	200.934	0.356	249.167	181.776	0.508	253.489	199.368	0.578	277.362	186.589	0.202	347.059	266.996
	SVR	1.828	266.647	212.499	0.376	254.797	188.716	0.542	265.229	208.270	0.654	291.262	200.411	0.237	485.052	345.596
	RVFL	1.653	283.733	228.334	0.460	454.795	262.305	0.552	308.482	220.023	0.720	540.852	270.388	0.235	660.998	391.416
	ES	1.983	255.003	201.186	0.397	259.791	184.418	0.572	262.134	202.557	0.619	302.783	201.337	0.260	323.907	249.130
	ARIMA	2.215	246.489	190.728	0.480	247.095	182.575	0.815	240.575	189.495	0.709	272.573	182.989	0.320	292.273	221.723
	SD-Bi-LSTM	0.606	212.252	181.457	0.332	190.633	146.227	0.311	186.841	131.244	0.549	195.618	152.499	0.170	274.823	218.358

Table 5. Average prediction performance of the candidate models.

| Mode	Model	$\bar{F}$		
		MAPE	*RMSE*	*MAE*
Periodicity mode 1	Bi-LSTM	0.687	224.586	158.097
	ELM	0.668	163.793	123.521
	KRR	0.714	230.607	167.673
	SVR	0.795	247.115	192.279
	RVFL	0.668	158.169	122.293
	ES	0.969	280.354	217.416
	SARIMA	0.115	101.078	71.449
Periodicity mode 2	Bi-LSTM	0.309	300.622	219.536
	ELM	0.194	193.163	138.371
	KRR	0.269	246.935	177.189
	SVR	0.308	275.055	205.454
	RVFL	0.201	198.829	140.761
	ES	0.208	251.500	184.814
	SARIMA	0.050	77.495	47.643
Trend mode	Bi-LSTM	0.003	127.604	96.230
	ELM	0.001	569.323	47.608
	KRR	0.001	65.728	65.256
	SVR	0.002	162.217	161.906
	RVFL	0.001	713.992	59.420
	ES	0.001	19.845	17.693
	ARIMA	0.023	821.165	731.879
Remainder mode	Bi-LSTM	0.412	228.534	175.372
	ELM	0.709	365.321	250.367
	KRR	0.707	275.638	207.133
	SVR	0.727	312.597	231.098
	RVFL	0.724	449.772	274.493
	ES	0.766	280.723	207.726
	ARIMA	0.908	259.801	193.502
	SD-Bi-LSTM	0.394	212.033	165.957

From Tables 4 and 5, it can be demonstrated that the optimal prediction models are different for each decomposition mode with different data traits. Therefore, it is necessary to perform the multi-trait-driven prediction model selection. In detail, as for the periodicity modes, the performance of the SARIMA model is significantly better than other models. Therefore, considering that the SARIMA model can effectively predict the data with periodicity [9], the SARIMA model is selected to predict the periodicity modes. As for the trend mode, the averages of *MAPE*, *RMSE*, and *MAE* of the ES model in the FV method are 0.001, 19.845, and 17.693, respectively, which outperforms other alternative models. Given the low complexity of trend mode, the ES model is chosen for trend mode forecasting. When it comes to remainder mode forecasting, the Bi-LSTM model has the best prediction accuracy, where its *MAPE*, *RMSE*, and *MAE* are 0.412, 228.534, and 175.372, respectively. The possible reason is that the Bi-LSTM model can make full use the information of the complex time series data, thereby achieving higher prediction accuracy [20]. By contrast, by performing the SD to remainder mode, the SD-Bi-LSTM model can further improve the performance of the Bi-LSTM model. For example, compared with the Bi-LSTM model, the *MAPE*, *RMSE*, and *MAE* of the SD-Bi-LSTM model are reduced by 4.368%, 7.221%, and 5.369%, respectively. Therefore, the SD-Bi-LSTM model is utilized for the prediction of the remainder mode.

### 3.2.4. Comparison of Prediction Performance

In this section, the prediction performance of different models on the testing dataset are compared. Correspondingly, the comparison results are listed in Table 6 and plotted in Figures 6–8, from which four main conclusions can be drawn.

**Table 6.** Prediction performance of different models.

Model	Horizon 1			Horizon 2			Horizon 3		
	MAPE	*RMSE*	*MAE*	*MAPE*	*RMSE*	*MAE*	*MAPE*	*RMSE*	*MAE*
The proposed model	0.025	996.632	803.840	0.027	1075.094	863.915	0.029	1153.009	925.959
IMSTL-S	0.030	1097.379	945.976	0.032	1313.271	1053.374	0.039	1562.250	1273.193
IMSTL-SARIMA	0.032	1254.016	1014.147	0.033	1306.357	1057.062	0.034	1356.456	1098.243
IMSTL-ES	0.056	2026.460	1803.078	0.058	2141.056	1873.246	0.062	2273.182	1962.798
IMSTL-Bi-LSTM	0.043	1656.921	1340.371	0.052	2007.111	1661.422	0.057	2158.918	1775.344
Bi-LSTM	0.053	2591.693	1641.828	0.063	2688.045	2023.559	0.073	3191.726	2372.660
ELM	0.068	2981.484	2287.289	0.070	2901.366	2317.169	0.070	3077.101	2385.460
KRR	0.056	2261.984	1832.726	0.060	2427.685	1961.387	0.064	2589.605	2089.087
SVR	0.073	2868.298	2388.849	0.084	3291.700	2768.444	0.095	3719.329	3140.454
RVFL	0.079	2617.263	2569.141	0.077	2591.416	2484.866	0.078	2648.145	2501.780
ES	0.084	3218.942	2668.784	0.095	3394.825	2864.424	0.091	3251.437	2760.568
SARIMA	0.098	3888.001	3282.341	0.112	4119.156	3487.136	0.102	3897.657	3303.147
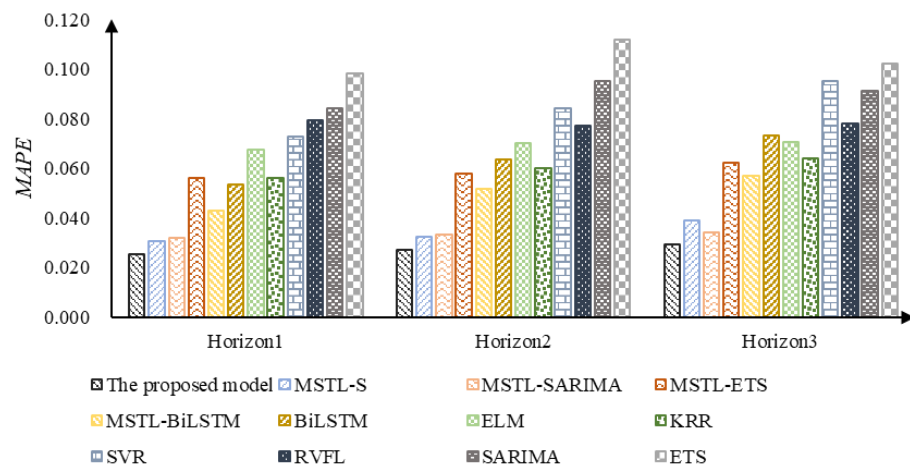
Figure 6. *MAPE* of each model in multi-step-ahead forecasting.
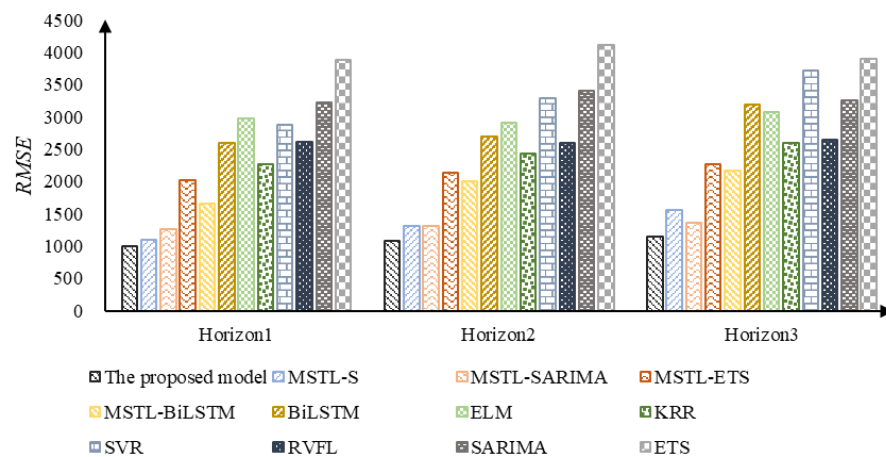
Figure 7. *RMSE* of each model in multi-step-ahead forecasting.
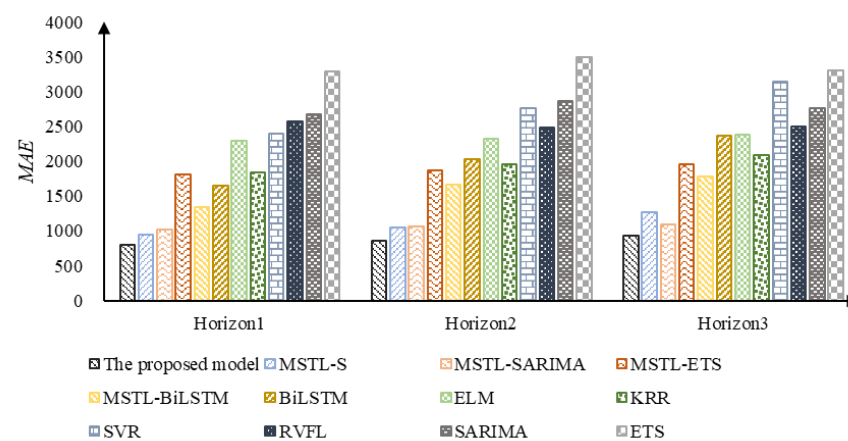
Figure 8. *MAE* of each model in multi-step-ahead forecasting.

Firstly, based on Figures 6–8 and Table 6, the proposed model outperforms all benchmark models in multi-step-ahead forecasting, as it has the lowest *MAPE*, *RMSE*, and *MAE*. In detail, compared with each single model and decomposition-ensemble model, the *MAPE*, *RMSE*, and *MAE* of the proposed model in multi-step-ahead forecasting are reduced by 55.866%, 54.636%, and 56.063% on average, respectively. Therefore, by performing multi-trait-driven modeling and SD, the proposed hybrid model can effectively improve the prediction accuracy of short-term load.

Secondly, the prediction performance of the decomposition-ensemble model outperforms the single model, as illustrated in Figures 6–8 and Table 6. For example, by comparing the performance of the single models, the *MAPE, RMSE*, and *MAE* of the decomposition-ensemble models are reduced by an average of 43.878%, 45.167%, and 44.543%, respectively. The main reason is that the divide-and-conquer framework can explore the inner factors of the data effectively [32], thereby achieving high accuracy of short-term load forecasting.

Thirdly, it can be demonstrated that the prediction model selection based on multi-trait-driven methodology can effectively improve the performance of the decomposition-ensemble model. For instance, compared with the accuracy of other decomposition-ensemble benchmark models, the *MAPE, RMSE*, and *MAE* of the IMSTL-S model are reduced by an average of 29.231%, 26.709%, and 27.981%, respectively. Therefore, the prediction accuracy of the decomposition-ensemble model can be improved by matching a suitable forecasting method for each decomposed mode with different data traits.

Fourthly, the SD of the decomposed mode with high-level complexity can further increase the accuracy of the prediction model, as can be seen in Figures 6–8 and Table 6. In detail, compared with the prediction accuracy of the IMSTL-S model, the proposed hybrid model achieves an average reduction in 19.311%, 17.837%, and 20.095% on *MAPE*, *RMSE*, and *MAE* metrics, respectively. Therefore, the implementation of SD can decrease the interference of the decomposed mode with high-level complexity [35], thereby outperforming the hybrid model with a separate decomposition method.

Furthermore, the DM test is used to statistically test the difference between the proposed model and the benchmark models, as displayed in Table 7. Accordingly, the above four main conclusions can be proved statistically. First, the proposed model outperforms all benchmark models with a 95% confidence interval, which statistically validates the effectiveness of the proposed model. Second, the prediction performance of each decomposition-ensemble model in the benchmark models is better than its corresponding single model under the 95% confidence interval. Thirdly, the decomposition-ensemble model combined with multi-trait-driven model selection achieves better prediction performance than its counterparts without model selection under the 99% confidence interval. Furthermore, focusing on the SD, the proposed model significantly outperforms the decomposition-ensemble benchmark models without SD under the 95% confidence interval.

Table 7. Results of the DM test.

Target Model	Benchmark Model										
	IMSTL-S	IMSTL-SARIMA	IMSTL-ES	IMSTL-Bi-LSTM	Bi-LSTM	ELM	KRR	SVR	RVFL	ES	SARIMA
The proposed model	−2.335 (0.011)	−7.354 (0.000)	−12.242 (0.000)	−7.053 (0.000)	−2.772 (0.003)	−7.574 (0.000)	−8.744 (0.000)	−10.875 (0.000)	−28.268 (0.000)	−11.407 (0.000)	−10.598 (0.000)
IMSTL-S		−4.563 (0.000)	−10.739 (0.000)	−5.871 (0.000)	−2.654 (0.004)	−7.229 (0.000)	−7.724 (0.000)	−10.008 (0.000)	−26.213 (0.000)	−11.079 (0.000)	−9.911 (0.000)
IMSTL-SARIMA			−8.909 (0.000)	−4.426 (0.000)	−2.475 (0.007)	−6.831 (0.000)	−6.871 (0.000)	−9.421 (0.000)	−21.491 (0.000)	−10.804 (0.000)	−9.527 (0.000)
IMSTL-ES				3.723 (0.000)	−1.251 (0.106)	−4.344 (0.000)	−1.293 (0.098)	−5.065 (0.000)	−8.575 (0.000)	−8.631 (0.000)	−7.706 (0.000)
IMSTL-Bi-LSTM					−1.957 (0.026)	−6.746 (0.000)	−8.067 (0.000)	−10.901 (0.000)	−13.675 (0.000)	−12.173 (0.000)	−7.963 (0.000)
Bi-LSTM						−0.955 (0.171)	0.964 (0.167)	−0.292 (0.385)	−0.063 (0.474)	−3.785 (0.000)	−1.715 (0.044)
ELM							5.571 (0.000)	2.391 (0.008)	2.022 (0.022)	−6.739 (0.000)	−1.072 (0.142)
KRR								−11.882 (0.000)	−4.983 (0.000)	−12.223 (0.000)	−5.981 (0.001)
SVR									0.815 (0.207)	−10.131 (0.000)	−3.084 (0.001)
RVFL										−6.728 (0.000)	−3.957 (0.000)
ES											3.028 (0.001)

4. Discussion

In this section, different multi-period decomposition models (e.g., Seasonal-Trend decomposition using Regression (STR) [56] and TBATS [57]), coupled with the STL model, are performed as the benchmark models (e.g., STR-P, TBATS and STL-P) for comparison purposes. Wherein, 'P' represents the model selection and SD based on multi-trait-driven modeling. Meanwhile, the related models in the existing literature [25,27,28,58] are also utilized to demonstrate the effectiveness of the proposed model, as provided in Tables 8 and 9 and Figures 9–11. Sequentially, the following three conclusions can be drawn.

Table 8. Prediction performance of different models in discussion.

| Model	Horizon 1			Horizon 2			Horizon 3		
	MAPE	*RMSE*	*MAE*	*MAPE*	*RMSE*	*MAE*	*MAPE*	*RMSE*	*MAE*
The proposed model	0.025	996.632	803.840	0.027	1075.094	863.915	0.029	1153.009	925.959
STR-P	0.030	1097.379	945.976	0.032	1313.271	1053.374	0.039	1562.250	1273.193
TBATS	0.032	1254.016	1014.147	0.033	1306.357	1057.062	0.034	1356.456	1098.243
STL-P	0.056	2026.460	1803.078	0.058	2141.056	1873.246	0.062	2273.182	1962.798
Pan et al. [25]	0.043	1656.921	1340.371	0.052	2007.111	1661.422	0.057	2158.918	1775.344
Jin et al. [58]	0.053	2591.693	1641.828	0.063	2688.045	2023.559	0.073	3191.726	2372.660
Tang et al. [27]	0.068	2981.484	2287.289	0.070	2901.366	2317.169	0.070	3077.101	2385.460
Lin et al. [28]	0.056	2261.984	1832.726	0.060	2427.685	1961.387	0.064	2589.605	2089.087
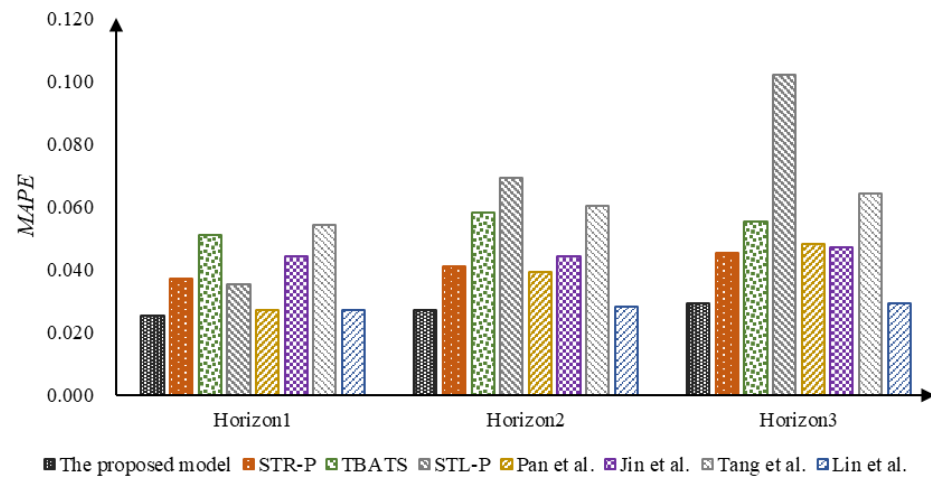
Figure 9. *MAPE* of each model in multi-step-ahead forecasting in discussion.
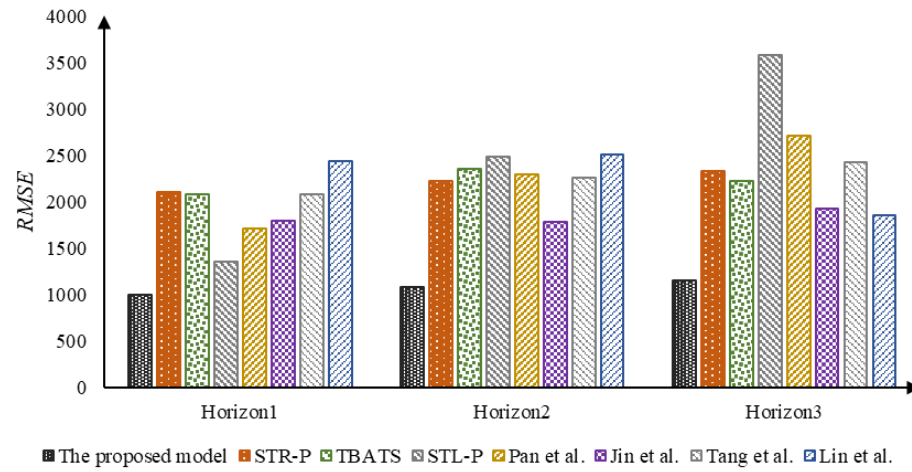

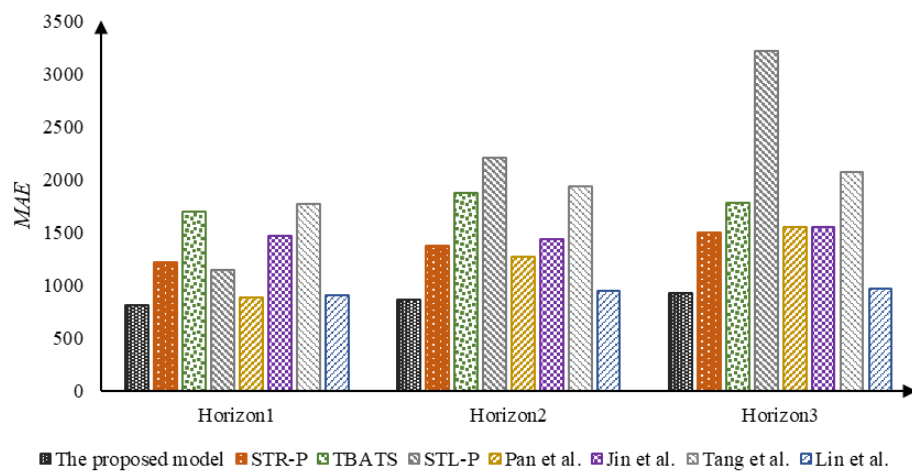**Figure 10.** *RMSE* of each model in multi-step-ahead forecasting in discussion.

Figure 11. *MAE* of each model in multi-step-ahead forecasting in discussion.

Table 9. Results of the DM test in discussion.

Target Model	Benchmark Model						
	STL-P	STR-P	TBATS	Pan et al. [25]	Jin et al. [58]	Tang et al. [27]	Lin et al. [28]
The proposed model	−4.419 (0.000)	−8.241 (0.000)	−6.293 (0.000)	−2.420 (0.008)	−6.761 (0.000)	−9.467 (0.000)	−1.668 (0.048)
STR-P		0.116 (0.453)	3.394 (0.000)	1.313 (0.095)	1.371 (0.086)	0.101 (0.459)	−0.487 (0.313)
TBATS			6.018 (0.000)	1.660 (0.049)	4.266 (0.000)	−0.062 (0.475)	−0.541 (0.294)
STL-P				−1.339 (0.091)	−4.009 (0.000)	−6.805 (0.000)	−1.374 (0.085)
Pan et al. [25]					−0.373 (0.354)	−1.725 (0.043)	−0.987 (0.162)
Jin et al. [58]						−5.657 (0.000)	−0.917 (0.179)
Tang et al. [27]							−0.542 (0.294)

First, as illustrated in Tables 8 and 9 and Figures 9–11, compared with the decomposition-ensemble model based on the single-period decomposition method, the proposed model obtains higher prediction accuracy. Specifically, compared with the STL-P model, the proposed model has an average reduction in 53.670%, 50.341%, and 53.839% on *MAPE*, *RMSE*, and *MAE* metrics, respectively. At the same time, as shown in Table 6, the proposed model outperforms the STL-P model under the 99% confidence interval. Therefore, the performance of short-term data can be improved by recognizing and extracting multi-period modes.

Secondly, according to the results demonstrated in Tables 8 and 9 and Figures 9–11, it can be found that the proposed model outperforms the decomposition-ensemble model based on other multi-period decomposition methods. In multi-step forward forecasting, the *MAPE*, *RMSE*, and *MAE* of the proposed model are reduced by an average of 43.546%, 51.458%, and 44.807% compared with its counterparts based on other multi-period decomposition methods. Meanwhile, the same conclusion can be confirmed by the results of the DM test at the 5% significance level. The possible reason is that the IMSTL model can extract the multi-period modes more effectively than other multi-period decomposition methods, thereby improving the prediction accuracy.

Thirdly, as can be seen in Tables 8 and 9 and Figures 9–11, compared with the existing prediction models [25,27,28,58], the proposed model has better prediction performance in multi-step forward forecasting. In detail, the *MAPE*, *RMSE*, and *MAE* of the proposed model in multi-step-ahead forecasting are reduced by an average of 36.450%, 49.937%, and 37.789%, respectively, compared with the existing methods. Furthermore, compared with existing methods, all p-values of the DM test for the proposed model are less than 0.05, which statistically validates the effectiveness of the proposed model.

5. Conclusions

To improve the prediction accuracy of short-term load series, this paper proposes a hybrid model based on a multi-trait-driven methodology and secondary decomposition. In detail, four steps are performed sequentially, i.e., data decomposition, secondary decomposition, individual prediction, and ensemble output, all of which are designed based on a multi-trait-driven methodology. In particular, the multi-period identification method and the judgment basis of secondary decomposition are designed to assist the construction of the hybrid model.

In the numerical experiment, the short-term load data with 15 min intervals are collected as the research object. Meanwhile, different single methods, decomposition-ensemble models, and the existing prediction models in the literature are utilized for comparison pur-

poses. By analyzing the results of multi-step-ahead forecasting and the Diebold–Mariano (DM) test, the proposed hybrid model is proven to outperform all benchmark models, which can be regarded as an effective solution for short-term load forecasting. Meanwhile, four insightful findings can be summarized. First, the proposed hybrid model based on multi-trait-driven methodology and secondary decomposition outperforms all benchmark models. Secondly, by combining with the multi-trait-driven model selection, the decomposition-ensemble models can obtain better prediction performance than their counterparts without this step. Thirdly, the IMSTL model can effectively recognize and extract the multi-period modes, thereby increasing the performance of short-term load forecasting. Finally, the proposed hybrid model is able to further increase the accuracy of short-term load forecasting by performing the SD for the mode with high-level complexity. However, there are some directions that are needed for further research. Considering the unavoidable prediction errors in forecasting, it is necessary to construct an appropriate method for uncertainty analysis to help the power system to formulate a dispatch plan. In addition, the feature selection of short-term load influencing factors would be an orientation to further improve the forecasting accuracy.

References

1. Walser, T.; Sauer, A. Typical load profile-supported convolutional neural network for short-term load forecasting in the industrial sector. *Energy AI* **2021**, *5*, 100104. [CrossRef]
2. Zhang, W.; Chen, Q.; Yan, J.; Zhang, S.; Xu, J. A novel asynchronous deep reinforcement learning model with adaptive early forecasting method and reward incentive mechanism for short-term load forecasting. *Energy* **2021**, *236*, 121492. [CrossRef]
3. Liu, H.; Shi, J. Applying ARMA–GARCH approaches to forecasting short-term electricity prices. *Energy Econ.* **2013**, *37*, 152–166. [CrossRef]
4. Chitalia, G.; Pipattanasomporn, M.; Garg, V.; Rahman, S. Robust short-term electrical load forecasting framework for commercial buildings using deep recurrent neural networks. *Appl. Energy* **2020**, *278*, 115410. [CrossRef]
5. Dai, Y.; Zhao, P. A hybrid load forecasting model based on support vector machine with intelligent methods for feature selection and parameter optimization. *Appl. Energy* **2020**, *279*, 115332. [CrossRef]
6. Wu, X.; Wang, Y.; Bai, Y.; Zhu, Z.; Xia, A. Online short-term load forecasting methods using hybrids of single multiplicative neuron model, particle swarm optimization variants and nonlinear filters. *Energy Rep.* **2021**, *7*, 683–692. [CrossRef]
7. Massana, J.; Pous, C.; Burgas, L.; Melendez, J.; Colomer, J. Identifying services for short-term load forecasting using data driven models in a Smart City platform. *Sustain. Cities Soc.* **2017**, *28*, 108–117. [CrossRef]
8. Rendon-Sanchez, F.; de Menezes, L.M. Structural combination of seasonal exponential smoothing forecasts applied to load forecasting. *Eur. J. Oper. Res.* **2019**, *275*, 916–924. [CrossRef]
9. Kazemzadeh, M.-R.; Amjadian, A.; Amraee, T. A hybrid data mining driven algorithm for long term electric peak load and energy demand forecasting. *Energy* **2020**, *204*, 117948. [CrossRef]
10. Dudek, G.; Pełka, P. Pattern similarity-based machine learning methods for mid-term load forecasting: A comparative study. *Appl. Soft Comput.* **2021**, *104*, 107223. [CrossRef]
11. Kychkin, A.V.; Chasparis, G.C. Feature and model selection for day-ahead electricity-load forecasting in residential buildings. *Energy Build.* **2021**, *249*, 111200. [CrossRef]
12. Cai, G.; Wang, W.; Lu, J. A novel hybrid short term load forecasting model considering the error of numerical weather prediction. *Energies* **2016**, *9*, 994. [CrossRef]
13. Elamin, N.; Fukushige, M. Modeling and forecasting hourly electricity demand by SARIMAX with interactions. *Energy* **2018**, *165*, 257–268. [CrossRef]

14. Li, C. A fuzzy theory-based machine learning method for workdays and weekends short-term load forecasting. *Energy Build.* **2021**, *245*, 111072. [CrossRef]

15. Wang, Z.; Zhou, X.; Tian, H.; Huang, T. Hierarchical parameter optimization based support vector regression for power load forecasting. *Sustain. Cities Soc.* **2021**, *71*, 102937. [CrossRef]

16. Hu, Y.; Qu, B.; Wang, J.; Liang, J.; Wang, Y.; Yu, K.; Li, Y.; Qiao, K. Short-term load forecasting using multimodal evolutionary algorithm and random vector functional link network based ensemble learning. *Appl. Energy* **2021**, *285*, 116415. [CrossRef]

17. Khwaja, A.S.; Anpalagan, A.; Naeem, M.; Venkatesh, B. Joint bagged-boosted artificial neural networks: Using ensemble machine learning to improve short-term electricity load forecasting. *Electric Power Syst. Res.* **2020**, *179*, 106080. [CrossRef]

18. Jiang, P.; Liu, F.; Song, Y. A hybrid forecasting model based on date-framework strategy and improved feature selection technology for short-term load forecasting. *Energy* **2017**, *119*, 694–709. [CrossRef]

19. Zang, H.; Xu, R.; Cheng, L.; Ding, T.; Liu, L.; Wei, Z.; Sun, G. Residential load forecasting based on LSTM fusing self-attention mechanism with pooling. *Energy* **2021**, *229*, 120682. [CrossRef]

20. Wang, S.; Wang, X.; Wang, S.; Wang, D. Bi-directional long short-term memory method based on attention mechanism and rolling update for short-term load forecasting. *Int. J. Electr. Power Energy Syst.* **2019**, *109*, 470–479. [CrossRef]

21. Zhu, K.; Geng, J.; Wang, K. A hybrid prediction model based on pattern sequence-based matching method and extreme gradient boosting for holiday load forecasting. *Electr. Power Syst. Res.* **2021**, *190*, 106841. [CrossRef]

22. Zhang, Z.; Hong, W.-C. Application of variational mode decomposition and chaotic grey wolf optimizer with support vector regression for forecasting electric loads. *Knowl.-Based Syst.* **2021**, *228*, 107297. [CrossRef]

23. Yang, D.; Guo, J.E.; Sun, S.; Han, J.; Wang, S. An interval decomposition-ensemble approach with data-characteristic-driven reconstruction for short-term load forecasting. *Appl. Energy* **2022**, *306*, 117992. [CrossRef]

24. Mohan, N.; Soman, K.P.; Sachin Kumar, S. A data-driven strategy for short-term electric load forecasting using dynamic mode decomposition model. *Appl. Energy* **2018**, *232*, 229–244. [CrossRef]

25. Pan, D.; Xu, B.; Jun, M.A.; Ding, Q.; Quan, D.I.; Jinjin, Z.H.; Qian, Z. Short-term load forecasting based on EEMD-approximate Entropy and ELM. In Proceedings of the 2019 IEEE Sustainable Power and Energy Conference (iSPEC) 2019, Beijing, China, 21–23 November 2019; pp. 1772–1775. [CrossRef]

26. Memarzadeh, G.; Keynia, F. Short-term electricity load and price forecasting by a new optimal LSTM-NN based prediction algorithm. *Electr. Power Syst. Res.* **2021**, *192*, 106995. [CrossRef]

27. Tang, J.; Zhao, J.; Zou, H.; Ma, G.; Wu, J.; Jiang, X.; Zhang, H. Bus load forecasting method of power system based on VMD and Bi-LSTM. *Sustainability* **2021**, *13*, 10526. [CrossRef]

28. Lin, Y.; Luo, H.; Wang, D.; Guo, H.; Zhu, K. An ensemble model based on machine learning methods and data preprocessing for short-term electric load forecasting. *Energies* **2017**, *10*, 1186. [CrossRef]

29. Zhang, Z.; Hong, W.-C. Electric load forecasting by complete ensemble empirical mode decomposition adaptive noise and support vector regression with quantum-based dragonfly algorithm. *Nonlinear Dyn.* **2019**, *98*, 1107–1136. [CrossRef]

30. Alipour, M.; Aghaei, H.; Norouzi, M.; Niknam, T.; Hashemi, S.; Lehtonen, M. A novel electrical net-load forecasting model based on deep neural networks and wavelet transform integration. *Energy* **2020**, *205*, 118106. [CrossRef]

31. Singh, S.N.; Mohapatra, A. Data driven day-ahead electrical load forecasting through repeated wavelet transform assisted SVM model. *Appl. Soft Comput.* **2021**, *111*, 107730. [CrossRef]

32. Yu, L.; Wang, Z.; Tang, L. A decomposition–ensemble model with data-characteristic-driven reconstruction for crude oil price forecasting. *Appl. Energy* **2015**, *156*, 251–267. [CrossRef]

33. Yu, L.; Ma, M. A memory-trait-driven decomposition–reconstruction–ensemble learning paradigm for oil price forecasting. *Appl. Soft Comput.* **2021**, *111*, 107699. [CrossRef]

34. Shah, I.; Iftikhar, H.; Ali, S.; Wang, D. Short-Term Electricity Demand Forecasting Using Components Estimation Technique. *Energies* **2019**, *12*, 2532. [CrossRef]

35. Wu, Z.; Xia, X.; Xiao, L.; Liu, Y. Combined model with secondary decomposition-model selection and sample selection for multi-step wind power forecasting. *Appl. Energy* **2020**, *261*, 114345. [CrossRef]

36. Xiang, L.; Li, J.; Hu, A.; Zhang, Y. Deterministic and probabilistic multi-step forecasting for short-term wind speed based on secondary decomposition and a deep learning method. *Energy Convers. Manag.* **2020**, *220*, 113098. [CrossRef]

37. Emeksiz, C.; Tan, M. Multi-step wind speed forecasting and Hurst analysis using novel hybrid secondary decomposition approach. *Energy* **2022**, *238*, 121764. [CrossRef]

38. Proietti, T.; Pedregal, D.J. Seasonality in high frequency time series. *Econ. Stat.* **2022**. [CrossRef]

39. Liu, X.; Iftikhar, N.; Huo, H.; Li, R.; Nielsen, P.S. Two approaches for synthesizing scalable residential energy consumption data. *Future Gener. Comput. Syst.* **2019**, *95*, 586–600. [CrossRef]

40. Bouktif, S.; Fiaz, A.; Ouni, A.; Serhani, M.A. Single and multi-sequence deep learning models for short and medium term electric load forecasting. *Energies* **2019**, *12*, 149. [CrossRef]

41. Tang, L.; Yu, L.; Liu, F.; Xu, W. An integrated data characteristic testing scheme for complex time series data exploration. *Int. J. Inf. Technol. Decis. Mak.* **2013**, *12*, 491–521. [CrossRef]

42. Lopez, J.H. The power of the ADF test. *Econ. Lett.* **1997**, *57*, 5–10. [CrossRef]

43. Tang, L.; Lv, H.; Yang, F.; Yu, L. Complexity testing techniques for time series data: A comprehensive literature review. *Chaos Solitons Fractals* **2015**, *81*, 117–135. [CrossRef]

44. Unsworth, C.P.; Cowper, M.R.; McLaughlin, S.; Mulgrew, B. A new method to detect nonlinearity in a time-series: Synthesizing surrogate data using a Kolmogorov–Smirnoff tested, hidden Markov model. *Phys. D Nonlinear Phenom.* **2001**, *155*, 51–68. [CrossRef]

45. Bartlett, M.S. Periodogram analysis and continuous spectra. *Biometrika* **1950**, *37*, 1–16. [CrossRef]

46. Li, T. Detection and estimation of hidden periodicity in asymmetric noise by using quantile periodogram. In Proceedings of the 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan, 25–30 March 2012; pp. 3969–3972. [CrossRef]

47. Daizadeh, I. Seasonal and secular periodicities identified in the dynamics of us FDA medical devices (1976–2020): Portends intrinsic industrial transformation and independence of certain crises. *Ther. Innov. Regul. Sci.* **2022**, *56*, 104–116. [CrossRef]

48. Puech, T.; Boussard, M.; D'Amato, A.; Millerand, G. A fully automated periodicity detection in time series [A]. In *International Workshop on Advanced Analysis and Learning on Temporal Data*; Springer: Cham, Switzerland, 2020; pp. 43–54. [CrossRef]

49. Cleveland, R.B.; Cleveland, W.S.; McRae, J.E. Terpenning I. STL: A seasonal-trend decomposition procedure based on loess. ournal of official statistics. *J. Off. Stat.* **1990**, *6*, 3–73.

50. Bandara, K.; Hyndman, R.J.; Bergmeir, C. MSTL: A seasonal-trend decomposition algorithm for time series with multiple seasonal patterns. *arXiv* **2021**, arXiv:2107.13462. [CrossRef]

51. Yu, L.; Ma, Y.; Ma, Y.; Zhang, G. A complexity-trait-driven rolling decomposition-reconstruction-ensemble model for short-term wind power forecasting. *Sustain. Energy Technol. Assess.* **2022**, *49*, 101794. [CrossRef]

52. Liu, X.; Zhang, Z.; Song, Z. A comparative study of the data-driven day-ahead hourly provincial load forecasting methods: From classical data mining to deep learning. *Renew. Sustain. Energy Rev.* **2020**, *119*, 109632. [CrossRef]

53. Chahkoutahi, F.; Khashei, M. A seasonal direct optimal hybrid model of computational intelligence and soft computing techniques for electricity load forecasting. *Energy* **2017**, *140*, 988–1004. [CrossRef]

54. Yu, L.; Ma, Y.; Ma, M. An effective rolling decomposition-ensemble model for gasoline consumption forecasting. *Energy* **2021**, *222*, 119869. [CrossRef]

55. Dong, Y.; Ma, X.; Fu, T. Electrical load forecasting: A deep learning approach based on K-nearest neighbors. *Appl. Soft Comput.* **2021**, *99*, 106900. [CrossRef]

56. Dokumentov, A.; Hyndman, R.J. STR: Seasonal-trend decomposition using regression. *INFORMS J. Data Sci.* **2021**, 1–13. [CrossRef]

57. De Livera, A.M.; Hyndman, R.J.; Snyder, R.D. Forecasting time series with complex seasonal patterns using exponential smoothing. *J. Am. Stat. Assoc.* **2011**, *106*, 1513–1527. [CrossRef]

58. Jin, Y.; Guo, H.; Wang, J.; Song, A. A hybrid system based on LSTM for short-term power load forecasting. *Energies* **2020**, *13*, 6241. [CrossRef]