


Article

Advanced Optimal System for Electricity Price Forecasting Based on Hybrid Techniques

Hua Luo ¹ and Yuanyuan Shao ^{2,*} 

¹ School of Economics and Finance, Shanghai International Studies University, Shanghai 201620, China; loha_u@shisu.edu.cn

² School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China

* Correspondence: sy1079746994@163.com

Abstract: In the context of the electricity sector's liberalization and deregulation, the accurate forecasting of electricity prices has emerged as a crucial strategy for market participants and operators to minimize costs and maximize profits. However, their effectiveness is hampered by the variable temporal characteristics of real-time electricity prices and a wide array of influencing factors. These challenges hinder a single model's ability to discern the regularity, thereby compromising forecast precision. This study introduces a novel hybrid system to enhance forecast accuracy. Firstly, by employing an advanced decomposition technique, this methodology identifies different variation features within the electricity price series, thus bolstering feature extraction efficiency. Secondly, the incorporation of a novel multi-objective intelligent optimization algorithm, which utilizes two objective functions to constrain estimation errors, facilitates the optimal integration of multiple deep learning models. The case study uses electricity market data from Australia and Singapore to validate the effectiveness of the algorithm. The forecast results indicate that the hybrid short-term electricity price forecasting system proposed in this paper exhibits higher prediction accuracy compared to traditional single-model predictions, with MAE values of 7.3363 and 4.2784, respectively.

Keywords: electricity price forecasting; hybrid model; improved decomposition technique; novel multi-objective optimization algorithm; artificial intelligence



Citation: Luo, H.; Shao, Y. Advanced Optimal System for Electricity Price Forecasting Based on Hybrid Techniques. *Energies* **2024**, *17*, 4833. <https://doi.org/10.3390/en17194833>

Academic Editor: Seung-Hoon Yoo

Received: 9 August 2024

Revised: 8 September 2024

Accepted: 24 September 2024

Published: 26 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, renewable energy generation, including wind and solar photovoltaics, has seen rapid expansion and active integration into the electricity market. Due to the uncontrollability of their power generation, it breaks the real-time supply and demand balance of electric energy, which causes the fluctuation of electricity prices in the power market to become more severe [1,2]. Electricity prices are considered an important reference for decision making in the market. Accurate short-term electricity price forecasting can help market participants and power suppliers flexibly adjust pricing strategies according to changes in electricity prices, thereby increasing profits [3]. Moreover, electricity purchasers can use the difference between peak and off-peak electricity prices to carry out dynamic cost management to save expenses. At the same time, it also provides an important scientific basis for the real-time supervision of regulatory authorities, ensuring the normal operation of electricity [4].

Our study focuses on enhancing the precision and robustness of hybrid models for forecasting electricity prices, employing cutting-edge optimization techniques. The motivation for our research lies in the potential impact of reliable short-term electricity price forecasts on the energy market, fostering intelligent data-driven strategies beneficial to energy managers, suppliers, and consumers, especially in an era of rapid technological advancements and severe resource wastage. The essence of our research lies in developing

time-series forecasting models suitable for further academic investigation, thereby underscoring the goal of the rigorous evaluation and enhancement of current predictive methodologies.

Electricity prices exhibit non-stationary and variable characteristics within the energy market, coupled with notable anomalies. Such factors introduce increased uncertainty and complexity to the electricity market, complicating the task of short-term price forecasting [5,6]. Generally, the prevailing approaches to forecasting encompass statistical methods, along with models based on machine learning or deep learning, and those that integrate hybrid methodologies.

Models based on statistical analysis usually analyze historical data mathematically and statistically to build probabilistic models to predict future trends. Examples include the hidden Markov model (HMM) [7], autoregressive integrated generalized autoregressive conditional heteroskedasticity (Autoregressive-GARCH) model [8], and autoregressive integrated moving average (ARIMA) model [9]. These models are simple to implement, require fewer parameters for adjustment, and demand fewer computing resources. They can predict the trend in electricity price changes to a certain extent. However, their effectiveness is diminished by their reliance on assumptions, such as data smoothness and autocorrelation, which render them less capable of accommodating a nonlinear series.

To address the limitations inherent in statistical analysis models, scholars have increasingly shifted their focus toward machine learning and deep learning methodologies [10]. These techniques have proven to be highly effective in identifying and processing complex, nonlinear characteristics, thus offering superior performances in applications. Machine learning models mainly include artificial neural networks (ANN) [11], support vector machine (SVM) [12], and random forest (RF) [13]. Machine learning can adjust weights to approximate multivariate functions and capture complex, dynamic, nonlinear features of electricity prices. However, these methods require a significant amount of historical data and effective features. Additionally, the algorithm's robustness needs improvement. More researchers opt for deep learning models when it comes to forecasting tasks [14]. Typical studies include the following: Rezaei, Rajabi, & Estebesar (2022) utilized a Stacked Auto Encoder (SAE) to extract features from denoised electricity price data efficiently. To further refine the accuracy of electricity price forecasts and minimize the potential for overfitting, this model incorporated a Gated Recurrent Unit (GRU) [15]. And B. Wang, Wei, & Su (2022) employed Long Short-Term Memory (LSTM) networks, which led to the selection of more accurate forecast samples, resulting in a reduction in both relative and absolute forecasting errors [16]. Furthermore, three different artificial intelligence algorithms were applied to forecast day-ahead electricity prices in the Turkish electricity market [17]. The experimental findings demonstrated that, in the context of error metrics, the model based on Extreme Gradient Boosting Decision Trees (XGBoost) exhibited superior performance compared to its counterparts. Deep learning approaches typically outperform traditional machine learning algorithms and encounter notable difficulties [18]. However, the process of directly incorporating exogenous variables through partial models risks overwhelming the system and obscuring recognizable patterns. Furthermore, the conventional method of initializing weights randomly in deep learning frameworks might underemphasize vital features and moments, thereby reducing the effectiveness of feature extraction. This necessitates a refined strategy to ensure balanced weight distribution and optimal pattern discernment, enhancing the overall efficacy of models.

Researchers have employed a variety of hybrid strategies to address the limitations inherent in singular models by leveraging the unique advantages of different methodologies to "learn from the strong to offset one's weaknesses effect" [19,20]. These efforts can be categorized into three distinct groups.

The first group focuses on optimizing from a singular model perspective by combining singular models with optimization models or by integrating singular models with variable selection models [21]. For instance, Qu et al. (2024) introduced an improved Wavelet Neural Network (WNN) model, building upon the foundation laid by ELM to construct

classification prediction models for various electricity pricing schemes, effectively overcoming the traditional WNN issues of slow or even non-converging rates [22]. Similarly, J. Wang, Liu, Song, & Zhao (2016) updated the dynamic choice artificial neural network (DCANN) using the Cuckoo Search algorithm (CS) to address the initial parameter shortcomings of DCANN [23]. This approach also involved the identification of poor training sample sets through the Index of Bad Samples Matrix (IBSM) and the selection of appropriate inputs for each desired output. While such models have achieved commendable predictive performance, they struggle to capture the complex characteristics of electricity price series.

To address this issue, a second group of models was introduced, incorporating signal decomposition strategies to forecast different trend variables independently and subsequently integrate the final predictions. Typical studies include the following: H. Zhang et al. (2021) introduced an innovative hybrid model that combined Temporal Convolutional Networks (TCNs) and ARIMA with empirical mode decomposition (EMD) [24]. This method targets high- and low-frequency data parts and significantly enhances accuracy. However, EMDs may experience modal aliasing at different frequencies when processing signals, especially when the signals contain similar frequency components. To address this challenge, P. Jiang, Nie, Wang, & Huang (2023) implemented the Improved Complete Ensemble EMD with Adaptive Noise (ICEEMDAN) technique in their research [25]. This approach ameliorates the mode mixing problem by introducing additional noise signals. However, the selection of noise levels within the algorithm significantly impacts the outcome, and the lack of a unified standard for choosing the optimal noise level complicates the utilization of this method. Conversely, variational mode decomposition (VMD) operates through predetermined bandwidth constraints to decompose signals, effectively reducing mode mixing between different modes [26]. The decomposition results are less susceptible to signal noise, enhancing the precision and interpretability of the decomposition [27]. Compared to EMD and its variants, VMD requires fewer parameters, making the algorithm easier to adjust and implement [28]. The predictive accuracy of second-group hybrid models demonstrates an improvement, yet they struggle to surmount the intrinsic limitations of singular models and lack the capacity for adaptive parameter adjustments in response to variations in input samples.

Consequently, contemporary research has expanded to encompass a third group of models that advocate for the integration of multiple predictive models, employing optimization algorithms to facilitate a weighted amalgamation of diverse models. Z. Yang, Ce, & Lian (2017) introduced the hybrid methodology aimed at enhancing predictive accuracy by leveraging the distinct characteristics of each model [29]. Specifically, our approach integrated the wavelet transform with the kernel extreme learning machine (KELM), which was further optimized through self-adapting particle swarm optimization techniques and the autoregressive moving average (ARMA) model. X. Zhang, Wang, & Gao (2019) assessed a groundbreaking hybrid forecasting framework in the context of the New South Wales electricity market [30]. This framework innovatively melds a variety of predictive methodologies, encompassing seasonal ARIMA, Backpropagation (BP) neural networks, and SVM, aiming to bolster the precision of forecasts. Y. Xu, Li, Wang, & Du (2024) adeptly utilized an advanced version of the Multi-Objective Tuna Swarm Optimization (MOTSO) algorithm [31]. This innovative approach was pivotal in determining the combined weights assigned to each model. The third group of models had significantly improved predictive accuracy; however, the selection of the model and its parameter configuration greatly influenced the operational efficiency and predictive performance. Thus, it is imperative to engage in additional simulation experiments to identify more optimal methods for short-term electricity price forecasts. Moreover, conducting multi-step forecasting experiments is crucial for maintaining the integrity of the research framework. The representative articles, models, and findings discussed in the preceding literature review are systematically outlined and summarized in Table 1.

Table 1. A comprehensive overview based on the representative literature.

Categorization	Related Studies	Main Technology	Discovery
Statistical analysis-based categorization	Liu, Wei, Yang, & Guan, 2013 [7]	HMM	Using the firing matrix of the HMM to correct the residuals of the combined prediction.
	Girish, 2016 [8]	Autoregressive-GARCH	A study on modeling and forecasting prices in a market with demand exceeding supply.
	Zhao, Wang, Nokleby, & Miller, 2017 [9]	ARIMA	Incorporating an exogenous time series into the model enhances precision.
Machine learning models	Singh & Sahay, 2018 [11]	ANN	Accuracy can be maintained even in the presence of high volatility in the raw data.
	Wu, He, Zhang, & Du, 2021 [12]	DE-SVM	The differential evolutionary (DE) enhanced the precision of the SVM model while concurrently diminishing the computational time required.
	P. Wang et al., 2022 [13]	RF	An adaptive online prediction methodology that employed RF algorithms was designed to dynamically adjust the sizes of its training datasets.
Deep learning models	Rezaei, Rajabi, & Estebarsari, 2022 [15]	GRU	This approach resulted in stable and robust forecasting performance.
	B. Wang, Wei, & Su, 2022 [16]	LSTM	LSTM networks led to the selection of more accurate forecast samples.
	Yorat, Ozbek, Zor, & Saribulut, 2023 [17].	XGBoost	The test results indicated that the XGBoost model outperformed the others in terms of error metrics.
Hybrid models	Qu et al., 2024 [22]	ELM-WNN (IWNN)	A two-stage forecasting algorithm that integrates pattern recognition and classification-based forecasting was proposed.
	P. Jiang, Nie, Wang, & Huang, 2023 [25]	ICEEMDAN-MSSA	Anchoring electricity price forecasts to a loading series marked by minimal variability.
	Y. Xu, Li, Wang, & Du, 2024 [31]	IMOTSO	This approach yields probability density estimation curves that closely align with the actual values, demonstrating minimal deviation.

In reviewing the literature, it is evident that hybrid forecasting methods are gaining acknowledgment for their significant advantages in predicting electricity prices. However, despite their rising importance, these models have limitations. The following gaps exist in current research: (1) The electricity price dataset is extensive and rife with extraneous information. Conventional time-series decomposition methods frequently generate mixed-frequency components, compromising the precision of predictive analyses. (2) In the realm of model combination, the absence of a scientifically rigorous weighting scheme can yield less-than-ideal forecasting outcomes. Moreover, the utilization of antiquated weighting strategies may impede optimal model parameter tuning and prolong the computational runtime.

Therefore, this study employed an innovative approach to address the existing gaps in research by integrating trend decomposition techniques with multi-objective intelligent optimization algorithms, effectively optimizing prediction accuracy. The principal contributions of the study are as follows:

(1) Introduce an innovative trend decomposition technique that enhances the efficiency of feature extraction for predictive analysis. By employing an improved variational mode decomposition, we can dissect electricity price sequences into trend components, cyclical components, and irregular fluctuating components. This allows for the separate prediction of each component, facilitating a more effective capture of the distinct characteristics inherent in electricity price data.

(2) Apply a novel multi-objective intelligent optimization technique to further refine the predictive outcomes of deep learning models through scientific weighting, thereby increasing the forecasting accuracy for non-stationary sequences. By constraining the estimation error through an objective function and seeking Pareto optimal solutions, we promote the integration of multiple deep learning models, leveraging the unique predictive strengths of each model.

(3) The proposed hybrid short-term electricity price forecasting system demonstrates adaptability and scalability, making it applicable to various types of power markets.

Through cases and discussions conducted across 14 comparative models in markets experiencing different levels of volatility, we validated the effectiveness of the proposed model's forecasting capabilities.

This paper is structured as follows: Section 1 is the introduction presenting the research motivation, related literature, current research gaps, and innovations; Section 2 introduces the basic concepts of model preparation; Section 3 demonstrates the main framework of this study; Section 4 gives the details of the experimental setup, data presentation, numerical results, and case study; Section 5 discusses the results of the present study; and Section 6 presents the conclusion and future research plans.

2. Basic Concepts of Model Preparation

2.1. Trend Decomposition Technique

Electricity price series exhibit complexities such as instability, nonlinearity, and unpredictability, posing a significant challenge to accurate forecasting [32]. Decomposition algorithms become essential in addressing prediction inaccuracies due to noise, elucidating crucial electricity price attributes, and augmenting the efficacy of predictive models in various domains [33]. The essence of selecting apt decomposition algorithms lies in their ability to discern and utilize the principal characteristics of electricity prices efficiently. However, traditional signal decomposition methods, such as Fourier transform and wavelet transform, may have limitations in some cases. For example, they usually assume that the signal is linear and stationary, which is less effective when dealing with nonlinear and non-stationary signals. In order to overcome these limitations, K. Dragomiretskiy & K. Zosso (2014) proposed the variational mode decomposition (VMD) algorithm [34].

VMD stands out as a premier technique for the iterative separation of multi-component signals into the quasi-orthogonal intrinsic mode function. VMD is notably superior to the adaptive decomposition method EMD as it circumvents the issues of recursive computation errors and premature termination. Therefore, for the analysis of electricity price data, VMD emerges as the methodology of choice [35].

The price sequence $\gamma(t)$ can be decomposed into P sub-signals (modes) $\{E_p(t)\} = \{E_1(t), \dots, E_p(t)\}$, where each mode is a set of band-limited data with a center frequency. $\{\omega_p\} = \{\omega_1, \dots, \omega_p\}$ and $E_p(t)$ are able to be transformed into their analytical form by $E_{A,p}(t) = E_p(t)j\mathbf{H}E_p(t) = ((t) + j/\pi t) * E_p(t)$, where \mathbf{H} is the Hilbert transform operator, $\delta(\cdot)$ represents the use of the Dirac function, and $*$ embodies the convolution.

To derive the $E_p(t)$ in the frequency domain, we can recenter the spectrum of $E_p(t)$ and apply a low-pass filter. The formulation of the recenter process is $E_{A,p}(t)e^{-j\omega_p t} = [(\delta(t) + j/\pi t) * E_p(t)]e^{-j\omega_p t}$.

Following Fourier transform, $e^{-j\omega_p t}$ should be converted to $\delta(\omega + \omega_p)$ in the frequency domain. The bandwidth of $E_p(t)$ can be determined by optimizing the demodulated spectrum. Thus, the objective and constraint are formed in (1).

$$\begin{cases} \min \left\{ \sum \left\| \partial_t [(\delta(t) + j/\pi t) * E_p(t)] e^{-j\omega_p t} \right\|_2^2 \right\} \\ \text{s.t. } \sum E_p(t) = \gamma(t) \end{cases} \quad (1)$$

The utilization of the augmented Lagrangian method (\mathcal{L}) assists in transforming the aforementioned constrained optimization problem to an unconstrained form:

$$\mathcal{L} = \alpha \sum \left\| \partial_t [(\delta(t) + j/\pi t) * E_p(t)] e^{-j\omega_p t} \right\|_2^2 + \left\| \gamma(t) - \sum E_p(t) \right\|_2^2 + \lambda [\gamma(t) - \sum E_p(t)] \quad (2)$$

where α is the balancing parameter of the data-fidelity constraint, and λ is the Lagrangian multiplier.

To simplify (2), a more feasible approach would be to transpose \mathcal{L} into the frequency domain using the Fourier transform as follows (3):

$$\min_{\alpha} \sum j\omega \left\| \left[(1 + \text{sgn}(\omega + \omega_p)) E_p(\omega + \omega_p) \right] \right\|_2^2 + \left\| \gamma(\omega) - \sum E_p(\omega) + \lambda(\omega)/2 \right\|_2^2 \quad (3)$$

where the multiplication of the exponential term becomes the frequency spectrum shift operation, the convolution becomes multiplication, the derivative operation becomes multiplying the objective function with $j\omega$, and sgn means the sign function.

Employing the Alternating Direction Method of Multipliers (ADMMs) [36], we computed the spectral properties of mode $E_p(\omega)$ by executing a series of iterative refinement with:

$$E_p^{r+1}(\omega) = \underset{E_p}{\text{argmin}} \left\{ \alpha \left\| j\omega (1 + \text{sgn}(\omega + \omega_p^r)) E_p^r(\omega + \omega_p^r) \right\|_2^2 + \left\| \gamma(\omega) - \sum_{l=1}^{p-1} E_l^{r+1}(\omega) - \sum E_l^r(\omega) + \lambda^r(\omega)/2 \right\|_2^2 \right\} \quad (4)$$

The iteration number is denoted by r , where l is mode index, P is the total number of modes, and Hermitian symmetry is applied to simplify the following:

$$E_p^{r+1}(\omega) = \underset{E_p}{\text{argmin}} \left\{ \int_0^\infty 4\alpha (\omega - \omega_p^r)^2 \left| E_p^r(\omega) \right|^2 + 2 \left| \gamma(\omega) - \sum_{l=1}^{p-1} E_l^{r+1}(\omega) - \sum_{l=p}^P E_l^r(\omega) + \lambda^r(\omega)/2 \right|^2 d\omega \right\} \quad (5)$$

The optimal solution is letting the derivative equal 0. Thus,

$$E_p^{r+1}(\omega) = \frac{\gamma(\omega) - \sum_{l=1}^{p-1} E_l^{r+1}(\omega) - \sum_{l=p}^P E_l^r(\omega) + \lambda^r(\omega)/2}{1 + 2\alpha (\omega - \omega_p^r)^2} \quad (6)$$

The center frequency is $\omega_p^{r+1} = \underset{\omega_p}{\text{argmin}} \left\{ \int_0^\infty (\omega - \omega_p^r)^2 \left| E_p^{r+1}(\omega) \right|^2 d\omega \right\}$. The solution is found by the following equations: $\omega_p^{r+1} = (\int_0^\infty \omega \left| E_p^{r+1}(\omega) \right|^2 d\omega) / (\int_0^\infty \left| E_p^{r+1}(\omega) \right|^2 d\omega)$ and $\lambda^{r+1}(\omega) = \lambda^r(\omega) + \tau(\gamma(\omega) - \sum E_p^{r+1}(\omega))$, where τ is the penalty parameter. And they run until the convergence criterion $\sum_v \left\| E_p^{r+1}(\omega) - E_p^r(\omega) \right\|_2^2 / E_p^r(\omega)_2^2 \leq \varepsilon$ is met.

Following a thorough examination of the data, this study employed a K-means clustering methodology to categorize dataset $E_p^r(\omega)$ into n distinct groups. Each chronological day within the dataset was assigned a specific pattern label to facilitate analysis. The Euclidean distance metric served as the basis for assessing the similarity between patterns. In the process of de-duplication and revision, the d -dimensional vector space of data points $E_{i,D} = \{E_i \mid i = 1, 2, \dots, t\}$ was partitioned by minimizing the objective function: $F = \sum (\text{argmin} \left\| E_i - c_j^\forall \right\|^2)$. Here, c_j^\forall represents the k -cluster centroids in the central set $C^\forall = \{c_j^\forall \mid j = 1, 2, \dots, k\}$. This method helps to identify the optimal number of clusters for the dataset, ensuring the efficient and accurate partitioning of the data points [37]. The final result obtained contained several trend components; hence, the method is called tVMD.

2.2. Multi-Objective Wild Horse Optimization Algorithm

The living behavior of stallions, several mares, and foals in the wild horse society inspired the wild horse optimization algorithm (WHO) [38]. The decency behavior of horses, such as grazing, chasing, dominance, leadership, and mating, has been used to construct a new method. In comparison to other intelligence algorithms, the optimization process of horses departing from the group and mating with horses from other groups efficiently avoids the algorithm slipping into local optimization and balances the algorithm's exploration and development. On this basis, this paper introduces the concepts of the Pareto optimal solution and external archiving mechanism into WHO and creatively proposes

a multi-objective wild horse swarm optimization algorithm (MOWHO) to solve complex forecasting problems. Figure 1 illustrates the key link and the flowchart of the algorithm.

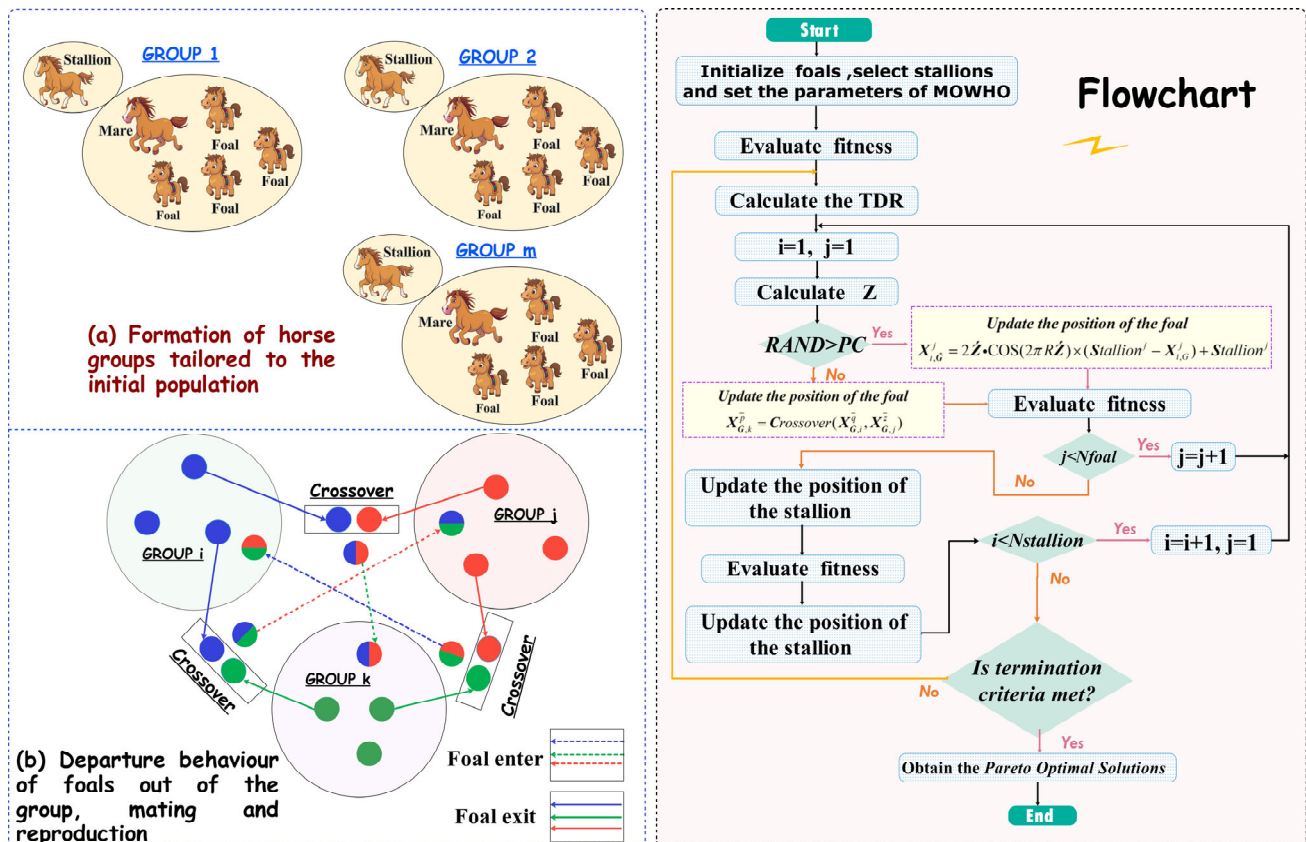


Figure 1. Key link and the flowchart of MOWHO.

(1) *Creating an initial population.*

An initialized random population is constructed $X = \langle x_1, x_2, \dots, x_n \rangle$ and determines the objective value $\vec{O}_i = \langle \vec{O}_{i1}, \vec{O}_{i2}, \dots, \vec{O}_{in} \rangle$. Then, the groups are divided. Let the population size be \bar{N} ; then, the number of groups is $\vec{G} = \lceil \bar{N} * PS \rceil$, where PS indicates the proportion of stallions. The remaining populations $\bar{N} - \vec{G}$ are divided equally into these groups according to the number of stallions, *Stallion* (Leader). Figure 1a shows an example of this population division.

(2) *Grazing behavior.*

The algorithm positions the stallion \vec{G} as the central point of the grazing area, with group members moving and searching around the *Stallion's* perimeter by $X_{i,\vec{G}}^j = 2\vec{Z} * \text{COS}(2\pi R\vec{Z}) \times (\text{Stallion}^j - X_{i,\vec{G}}^j) + \text{Stallion}^j$, where $X_{i,\vec{G}}^j$ is the position of the mare and foal and \vec{Z} is an adaptive mechanism: $\vec{Z} = R_2 \Theta \vec{I} + R_2 \Theta (\vec{I})$. $R_i \in [-2, 2]$ is a uniform random number. The COS function enables the population to move in a 360-degree direction. The \vec{I} index satisfies the condition ($P == 0$) and this results in $P = R_1 < \text{TDR}$, $\vec{I} = (P == 0)$ and $\vec{Z} = R_2 \Theta \vec{I} + R_2 \Theta (\vec{I})$, where TDR is the parameter regulator by $\text{TDR} = 1 - \text{Iter} \times (1/\text{IMax})$, which gradually decreases to 0 with the number of iterations.

(3) *Improved nonlinear update strategy.*

To overcome the shortage of the only local search, a new nonlinear $\overline{\text{TDR}}_{non}$ was proposed by $\overline{\text{TDR}}_{non} = 2 \times (1 - (\text{Iter}/\text{IMax})^{1/3})^{1/3} \times (\text{Rand} - 1)$, where $\text{Rand} \in [0, 1]$ is a random number. Through the nonlinear updating mode, the algorithm performs global and local searches in the whole iterative process. Therefore, $\text{TDR} = \overline{\text{TDR}}_{non}$ in the paper.

(4) *Horse mating behavior.*

Horses have unique forms of decent behavior, i.e., the foal will leave the original population when it is about to mature, the male horse will join the single horse group, and the female horse will join another family group. Such decent behavior avoids inbreeding among horses. WHO uses the mean crossover operator to simulate this behavior in horses: $X_{\bar{G}_k}^p = Crossover(X_{\bar{G}_i}^q, X_{\bar{G}_j}^z)$, where $X_{\bar{G}_k}^p$ is the distance from group k ; $X_{\bar{G}_i}^q$ and $X_{\bar{G}_j}^z$ are two maturing foals from i group and j group, respectively. Figure 1b shows this mating and departure behavior.

(5) *Group leadership.*

Each group has a leader (*Stallion* \bar{G}_i) who leads the group to occupy the water hole \mathbf{H} and occupy a dominant position. (\mathbf{H} is the current optimal solution).

$$Stallion_{\bar{G}_i} = \begin{cases} X_{\bar{G}_i} & \text{if } \text{COS}[t(X_{\bar{G}_i})] > \text{COS}[t(Stallion_{\bar{G}_i})] \\ Stallion_{\bar{G}_i} & \text{if } \text{COS}[t(X_{\bar{G}_i})] < \text{COS}[t(Stallion_{\bar{G}_i})] \end{cases} \quad (7)$$

$$\overline{\overline{Stallion}}_{\bar{G}_i} = \begin{cases} 2\bar{Z} * \text{COS}(2\pi R\bar{Z}) \times (\mathbf{H} - Stallion_{\bar{G}_i}) + \mathbf{H} & \text{if } R_3 > 0.5 \\ 2\bar{Z} * \text{COS}(2\pi R\bar{Z}) \times (\mathbf{H} - Stallion_{\bar{G}_i}) - \mathbf{H} & \text{if } R_3 \leq 0.5 \end{cases} \quad (8)$$

where $\overline{\overline{Stallion}}_{\bar{G}_i}$ and $Stallion_{\bar{G}_i}$ are the next and current positions of the leader of the i group. In addition, for the selection of leaders, $\overline{\overline{Stallion}}_{\bar{G}_i}$ in each group, we randomly select leaders $\overline{\overline{Stallion}}_{\bar{G}_1}$ in the initial stage to maintain the diversity of algorithms. In the later stage, the position of the leader $\overline{\overline{Stallion}}_{\bar{G}_i}, i = 2, 3, \dots$ is exchanged in the group according to the Pareto dominance relationship.

(6) *Adaptive T-distribution variation strategy.*

The variation operator in the algorithm uses the T-distribution with the number of iterations (*Iter*) serving as the degree of freedom. This aims to balance exploration and development by $\overline{\overline{Stallion}}_{\bar{G}_i} = \overline{\overline{Stallion}}_{\bar{G}_i} + T(Iter) * \overline{\overline{Stallion}}_{\bar{G}_i}$. The probability density function is as follows:

$$t_n(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, -\infty < x < \infty \quad (9)$$

where $\Gamma(\cdot)$ is the Gamma function.

(7) *Multi-objective version introduction.*

This paper innovatively introduces the Pareto optimal solution \mathcal{A}^* , and external Archiving mechanism \mathbf{Arc} into WHO, and constructs a multi-objective wild horse optimization algorithm (MOWHO) suitable for time-series forecasts. In the current iteration, \mathbf{Arc} serves primarily as a repository for the non-dominated solution \mathcal{A}'' , thus establishing an upper limit $\bar{N} \leq N$. Upon acquiring a new non-dominated solution \mathcal{A}^{nd} via the iterative method, it undergoes comparison with the archived member $\{\mathcal{A}''_1, \mathcal{A}''_2, \dots, \mathcal{A}''_N\}$. If $F(\mathcal{A}^{nd}) \succ F(\mathcal{A}''_i)$, $i \in [1, N]$, it shall incorporate \mathcal{A}^{nd} in \mathbf{Arc} . To refine \mathbf{Arc} , the optimization algorithm will retain the superior solution within \mathbf{Arc} while discarding the least effective one. The selection of the most advantageous starting point for the subsequent iteration is conducted through the application of a roulette-based mechanism $\mathbf{p}_i^f = \zeta / \bar{A}_i^f, \zeta > 1$, where ζ is considered a constant, while \bar{A}_i^f represents the number of solutions that are closely related to the i solution. \mathbf{p}_i^f assigns higher weights to solutions located in regions with fewer samples at the Pareto optimality boundary. This strategy aims to expand the coverage to provide a more diverse and comprehensive solution. It is worth noting that the two optimization objectives $[\overline{\overline{Obf}}_1^M, \overline{\overline{Obf}}_2^M]$ and constraints of this study are as follows: $\overline{\overline{Obf}}_1^M = \text{MinMAE}$ and $\overline{\overline{Obf}}_2^M = \text{Minstd}(EP'' - EP)$.

3. The Main Structure

The principal architecture is depicted in Figure 2, with the specific workflow delineated as follows. Initially, trend decomposition techniques are employed to extract trends, seasonal, and irregular components from raw electricity price data. Subsequently, the three most accurate forecasting models are selected for multiple simulations and parameter adjustments to ensure the stability of results. Ultimately, the multi-objective optimization algorithm is utilized to execute a weighted combination of the three models. And the comparative results are obtained by analyzing the error evaluation indexes.

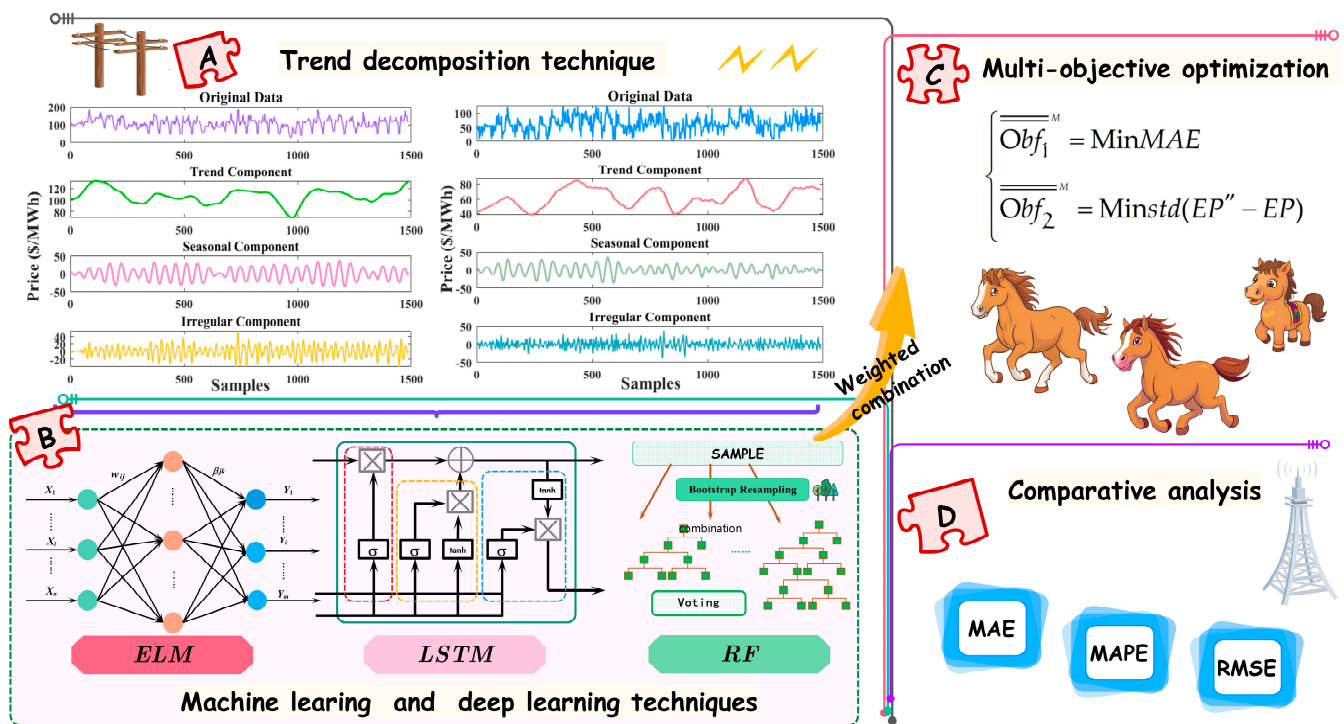


Figure 2. The main structure of this study.

4. Case Study and Evaluations

4.1. Datasets and Preprocessing

In this comparative analysis, we leveraged datasets derived from authentic electricity markets. The Singapore electricity market (SGP-EP), inaugurated in 2003, represents Asia's pioneering deregulated market. Recent years have seen the Singaporean government adapt to the global movement towards energy transition, enacting long-term strategies for the decarbonization of its electricity sector. This initiative has culminated in the development of a sophisticated framework for electricity generation and retailing. Conversely, the Australian National Electricity Market (AUS-EP), initiated in 1998, extends over 5000 km, making it the world's most extensive alternating current electricity system. Characterized by its leading position in global price volatility, it ranks among the most unstable electricity markets internationally.

In the context of electricity markets, the data on electricity prices exhibit pronounced volatility and are frequently characterized by a surplus of peak values. Utilizing unadjusted electricity price series as inputs for predictive models inherently compromises the precision of future price forecasts. To mitigate this challenge, our approach encompasses the identification and subsequent adjustment of aberrant outlier data points. For example, within the Singaporean electricity market, outliers are ascertained utilizing the 3σ criterion. Conversely, the Australian market's outlier data, distinguished by their considerable deviation resulting in atypical standard deviations, necessitates outlier identification through box plot analysis. Following outlier adjustment, data imputation is conducted employing the piecewise cubic Hermite interpolation polynomial (PCHIP) method.

It aims to approximate the original data by a series of cubic polynomials, each of which is responsible for interpolating between two points in the dataset. The polynomial is expressed as follows: $H_3(x) = [(1 + 2\chi)y_0 + (x - x_0)y'_0]\psi^2 + [(1 + 2\psi)y_1 + (x - x_1)y'_1]\chi$, where x_0, x_1 are the positions of the two neighboring points of the point to be interpolated, $\chi = (x - x_0)/(x_1 - x_0)$, $\psi = (x - x_1)/(x_0 - x_1)$, y_0, y_1 corresponds to the dependent variables of the independent variable x_0, x_1 , and y'_0, y'_1 are the corresponding derivatives. This methodology ensures the interpolation curve's monotonicity and smoothness, effectively eliminating the introduction of supplementary local extrema. Such a feature renders it exceptionally applicable to contexts necessitating the mitigation of undue oscillations and the retention of inherent data trends and attributes.

Our dataset encompasses 31 days from 1 January to 31 January 2024 (sample interval: 15 min). We adopted a stratified split for each dataset into training, validation, and test sets in a 6:2:2 ratio, ensuring rigorous methodology and comprehensive analysis. The specific features of the datasets are shown in Table 2.

Table 2. Statistical features of the datasets.

Dataset	Samples	Numbers	Statistical Indicator (USD/MWh)				
			Max	Min	Median	MEAN	Std.
AUS-EP	All samples	1488	127.83	4.15	59.98	61.54	25.03
	Training set	912	127.83	4.15	58.00	60.10	26.53
	Validation set	288	125.72	10.00	61.02	65.93	21.83
	Testing set	288	127.81	9.54	60.77	61.69	22.55
SGP-EP	All samples	1488	183.62	47.48	112.17	109.23	25.75
	Training set	912	183.54	29.08	107.68	101.38	33.22
	Validation set	288	177.17	49.43	116.27	107.55	26.19
	Testing set	288	183.62	29.08	113.12	107.38	27.58

4.2. Evaluation Metrics

In assessing the predictive model, three essential metrics were utilized: the Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE) [39,40]. The equation and indicator strengths are shown in Table 3.

Table 3. Three indicators.

Equation	Strengths
$MAE = \frac{1}{n} \sum_{t=1}^n EP_t - EP'_t $	The average difference between predicted and actual values is directly provided, relatively unaffected by extreme values.
$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{ EP_t - EP'_t }{EP_t} * 100\%$	Expression of the error as a percentage allows for fair comparisons between datasets of different scales.
$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (EP_t - EP'_t)^2}$	As the error is squared and then averaged, a larger penalty is given to large prediction errors.

Note: EP_t and EP'_t represent actual and predicted prices, t represents the sample index in chronological order, and $t = 1, 2, \dots, n$, and n denotes the total number of samples.

4.3. Experimental Setup

In this segment, we test the proposed hybrid short-term electricity price forecast (HSEPF) model through experimental studies. We begin by outlining the model's complex operational settings, followed by a thorough examination of the results obtained from these processes. The first case demonstrates the model's significant ability to integrate various forecasting elements, showing a clear advantage over traditional methods. The second case highlights an improvement in performance compared to models based on different decomposition techniques. The third case emphasizes the model's exceptional accuracy in predictions thanks to its novel approach of intelligently weighting multi-objectives.

4.3.1. Operational Settings

In the field of deep learning, model parameters play a pivotal role in determining both the computational efficiency and the accuracy of experimental simulations. This study has rigorously tested these parameters across a series of simulation experiments, with the definitive configurations enumerated in Table 4.

Table 4. Parameter settings.

Model	Parameter	Symbol	Value	Reason
ELM	Input-layer nodes	I_{Elm}	5	Trial and error
	Hidden-layer nodes	H_{Elm}	15	Trial and error
	Output-layer nodes	O_{Elm}	1	Target outcome
	Transfer function	Tf_{Elm}	<i>sigmoid</i>	Common preset
LSTM	Input-layer nodes	I_{Lstm}	5	Trial and error
	Hidden-layer nodes	H_{Lstm}	150	Trial and error
	Output-layer nodes	O_{Lstm}	1	Target outcome
	Gradient threshold	Γ_{tau}^{Lstm}	2	Common preset
	Maximum iteration	$Iter_{max}^{Lstm}$	500	Trial and error
RF	Max_depth	$Depth_{max}^{RF}$	50	Trial and error
	Min_samples_leaf	$Leaf_{max}^{RF}$	5	Trial and error
tVMD	Moderate bandwidth constraint	α_{tVMD}^{\vee}	2000	Trial and error
	Modes K	K_{tVMD}^{\vee}	5	Trial and error
	categories	C_{tVMD}^{\vee}	3	outcome
MOWHO	Population number	NP_{MOWHO}	100	Trial and error
	Maximum iteration	$Iter_{max}^{MOWHO}$	100	Trial and error
MODA	Population number	NP_{MODA}	100	Parameter consistency
	Maximum iteration	$Iter_{max}^{MODA}$	100	Parameter consistency

4.3.2. Case I: Comparison with Predictive Classical Models

Case I involved conducting a comparative experiment with a single model. In this scenario, ARIMA, BP, ELM, LSTM, and RF single models were first used for direct prediction. Subsequently, considering that combining multiple models would increase the complexity of the prediction framework, this paper intended to use the three best-performing single models for decomposition-ensemble [41]. The prediction results are shown in Table 5.

(a) For AUS-EP forecasting, the developed HSEFP system was evaluated alongside five benchmark models. This comparison provides a fundamental insight into the advanced capabilities of the proposed system. Specifically, when making a 1-step prediction—which involves forecasting the upcoming half-hour—the HSEFP system demonstrates superior performance with parameter settings of $\alpha_{tVMD}^{\vee} = 2000$, $K_{tVMD}^{\vee} = 5$, $C_{tVMD}^{\vee} = 3$, $NP_{MOWHO} = 100$, and $Iter_{max}^{MOWHO} = 100$. It achieved $MAE_{one-step}^{HSEFP} = 7.3362$, $MAPE_{one-step}^{HSEFP} = 14.9100\%$, and $RMSE_{one-step}^{HSEFP} = 9.7486$. In contrast, the benchmarks reported MAE values above 10, MAPE values ranging from 22% to 27%, and RMSE values exceeding 15. The prediction accuracies are significantly better relative to all the base single models. Figure 3 shows the forecast trends and error indicators of each model.

(b) For SGP-EP forecasting, the overall forecast error was smaller because this market is more stable than the AUS-EP. When 1-step forecasting was performed, the best forecasting performance in the benchmark model was ELM, with parameters $I_{Elm} = 5$, $H_{Elm} = 15$, $O_{Elm} = 1$, and $Tf_{Elm} = sigmoid$ resulting in $MAE_{one-step}^{ELM} = 5.9693$, $MAPE_{one-step}^{ELM} = 6.0769\%$, and $RMSE_{one-step}^{ELM} = 9.3607$. HSEFP has a better forecasting performance (with parameters $NP_{MOWHO} = 100$, $Iter_{max}^{MOWHO} = 100$, $H_{Elm} = 15$, $Tf_{Elm} = sigmoid$, $H_{Lstm} = 150$, $Iter_{max}^{Lstm} = 500$, $Depth_{max}^{RF} = 50$, and $Leaf_{max}^{RF} = 5$), where the prediction is expressed as $MAE_{one-step}^{HSEFP} = 4.2784$, $MAPE_{one-step}^{HSEFP} = 4.1918\%$, and $RMSE_{one-step}^{HSEFP} = 5.9268$. There is still a gap in the forecasting effectiveness of the single model.

Table 5. Results of the compared models for different markets.

Electricity Markets	Model	1-Step			2-Step			3-Step		
		MAE (USD/MWh)	MAPE (%)	RMSE (USD/MWh)	MAE (USD/MWh)	MAPE (%)	RMSE (USD/MWh)	MAE (USD/MWh)	MAPE (%)	RMSE (USD/MWh)
AUS-EP	ARIMA	11.3460	21.8877	16.0503	15.1869	30.1169	20.3130	17.0614	34.3129	22.6320
	BP	11.6282	26.0983	15.8673	14.8692	32.6167	19.6171	15.6677	37.4229	20.5253
	ELM	10.8203	22.6797	15.4311	13.8873	30.9370	18.9472	14.8129	34.4319	19.7708
	LSTM	10.7900	22.5088	15.4000	13.9089	30.9113	18.9242	14.6578	34.2829	19.5256
	RF	11.1553	24.9916	15.4319	14.4085	32.7613	18.7385	15.2152	36.8230	19.5678
	tVMD-ELM	7.3713	15.3919	9.9353	9.9147	20.8753	12.9356	10.4908	22.4343	13.5512
	tVMD-LSTM	7.5255	15.6030	10.2262	9.8997	20.7652	13.0770	10.2958	22.0460	13.2776
	tVMD-RF	8.8530	18.8079	11.8251	10.1543	21.7829	13.3063	10.7188	23.1796	13.5515
	Proposed HSEPF	7.3362	14.9100	9.7486	9.8568	20.9750	12.9024	10.2954	21.9254	13.2731
SGP-EP	ARIMA	6.5799	6.6092	10.3703	10.5916	10.6513	15.5312	13.9822	14.2917	19.7375
	BP	6.2245	6.3944	9.5537	9.3644	9.5499	13.0753	11.4359	11.9542	15.4864
	ELM	5.9693	6.0769	9.3607	8.9450	9.0558	12.7482	11.1910	11.5506	15.3115
	LSTM	6.0180	6.0780	9.4193	9.1182	9.2607	12.9609	11.1937	11.6502	15.4007
	RF	6.3716	6.2949	9.4881	9.3393	9.4269	12.6135	10.9317	11.4619	14.7077
	tVMD-ELM	4.3112	4.2667	5.9489	5.2712	5.1873	6.9389	6.8767	6.8401	9.0130
	tVMD-LSTM	4.5092	4.4401	6.2313	5.1233	5.0226	6.8330	6.7955	6.7179	8.8900
	tVMD-RF	5.4662	5.5410	7.3617	6.3240	6.4909	8.2456	7.8141	8.1212	9.9183
	Proposed HSEPF	4.2784	4.1918	5.9268	5.0703	4.8966	6.7367	6.7639	6.5111	9.1110

Note: This table shows the prediction results for five single models, three models based on the tVMD trend decomposition afterward and the proposed HSEPF, where $MAE = \frac{1}{n} \sum_{t=1}^n |EP_t - EP'_t|$, $MAPE = \frac{1}{n} \sum_{t=1}^n (|EP_t - EP'_t|) / (EP_t) * 100\%$, and $RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (EP_t - EP'_t)^2}$. Bold numbers indicate that the model performed better in all comparison results.

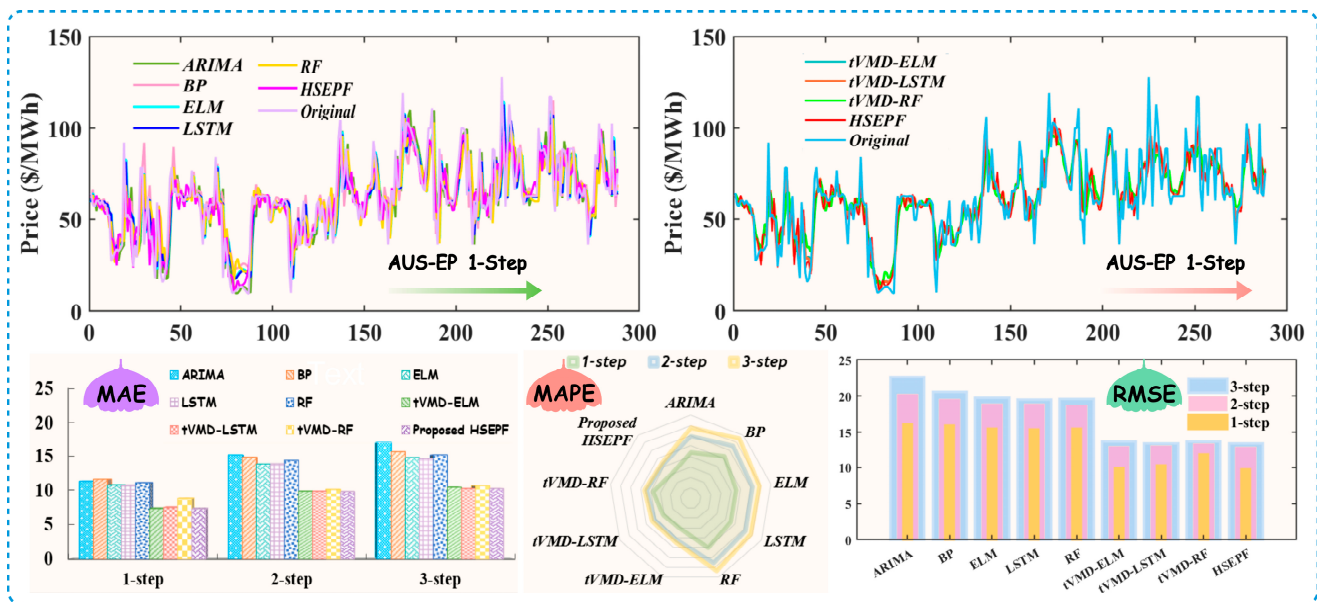


Figure 3. The forecasting performances in Case I for the AUS market.

(c) To compare the capability of the data decomposition–integration strategy, the three models ELM, LSTM, and RF, which performed better in the benchmark model, were trend decomposed, and each trend variable was predicted individually. As shown in Table 5, the prediction accuracy improved after integration. From the forecasting results for the AUS-EP, to ensure objectivity in the model comparisons, the ELM was unified with parameters $I_{Elm} = 5$, $H_{Elm} = 15$, $O_{Elm} = 1$, and $Tf_{Elm} = sigmoid$; the changes in the indicators at 1-step prediction were $MAE_{\{ELM,tVMD-ELM\}}^{one-step} = \{10.8203 \rightarrow 7.3713\}$, $MAPE_{\{ELM,tVMD-ELM\}}^{one-step}$

= {22.6797%→15.3919%}, and $RMSE_{\{ELM,tVMD-ELM\}}^{one-step} = \{15.4311 \rightarrow 9.9353\}$. The LSTM uniform parameters were $H_{Lstm} = 150$, $\Gamma_{\tau}^{Lstm} = 2$, and $Iter_{max}^{Lstm} = 500$; the changes in the indicators at one-step prediction were as follows: $MAE_{\{LSTM,tVMD-LSTM\}}^{one-step} = \{10.7900 \rightarrow 7.5255\}$, $MAPE_{\{LSTM,tVMD-LSTM\}}^{one-step} = \{22.5088\% \rightarrow 15.6030\% \}$, and $RMSE_{\{LSTM,tVMD-LSTM\}}^{one-step} = \{15.4000 \rightarrow 9.8997\}$. The RF model unification parameters were $Depth_{max}^{RF} = 50$ and $Leaf_{max}^{RF} = 5$; the changes in the indicators were $MAE_{\{RF,tVMD-RF\}}^{one-step} = \{11.1553 \rightarrow 8.8530\}$, $MAPE_{\{RF,tVMD-RF\}}^{one-step} = \{24.9916\% \rightarrow 18.8079\% \}$, and $RMSE_{\{RF,tVMD-RF\}}^{one-step} = \{15.4319 \rightarrow 11.8251\}$. From the performance of the metrics, the tVMD decomposition technique effectively improved the accuracy prediction. In the multi-step simulation results, it seemed that 88.89% of the indicators could reflect the superiority of the HSEPF system.

(d) The fitting effect of the SGP-EP models was satisfactory. Additionally, decomposing the model using tVMD significantly enhanced the performance. Under the parameters $I_{Elm} = 5$, $H_{Elm} = 15$, $O_{Elm} = 1$, and $Tf_{Elm} = sigmoid$, there were $MAE_{\{ELM,tVMD-ELM\}}^{one-step} = \{5.9693 \rightarrow 4.3112\}$, $MAPE_{\{ELM,tVMD-ELM\}}^{one-step} = \{6.0769\% \rightarrow 4.2667\% \}$, and $RMSE_{\{ELM,tVMD-ELM\}}^{one-step} = \{9.3607 \rightarrow 5.9489\}$. The LSTM uniform parameters were $I_{Lstm} = 5$, $H_{Lstm} = 150$, $O_{Lstm} = 1$, $\Gamma_{\tau}^{Lstm} = 2$, and $Iter_{max}^{Lstm} = 500$; the results were $MAE_{\{LSTM,tVMD-LSTM\}}^{one-step} = \{6.0180 \rightarrow 4.5092\}$, $MAPE_{\{LSTM,tVMD-LSTM\}}^{one-step} = \{6.0780\% \rightarrow 4.4401\% \}$, and $RMSE_{\{LSTM,tVMD-LSTM\}}^{one-step} = \{9.4193 \rightarrow 6.2313\}$. The errors of the RF model were reduced (parameters were $Depth_{max}^{RF} = 50$ and $Leaf_{min}^{RF} = 5$) and the results were $MAE_{\{RF,tVMD-RF\}}^{one-step} = \{6.3716 \rightarrow 5.4662\}$, $MAPE_{\{RF,tVMD-RF\}}^{one-step} = \{6.2949\% \rightarrow 5.5410\% \}$, and $RMSE_{\{RF,tVMD-RF\}}^{one-step} = \{9.4881 \rightarrow 7.3617\}$. The prediction errors all decreased significantly, indicating that the introduction of the tVMD decomposition technique significantly enhanced the hybrid model.

Remark 1. *The effectiveness of the HSEPF system in predicting future electricity prices was validated through experimental results. This system surpassed other single models in forecasting accuracy, as evidenced by the results that support the validity and effectiveness of the proposed modeling theory and methodology. Additionally, the use of the tVMD decomposition technique significantly improved forecasting accuracy. Crucially, this method supports automatic adjustments based on incoming data, enhancing their applicability across various forecasting domains and its practical implementation.*

4.3.3. Case II: Comparison with Other Data Decomposition Methods

This experiment validated the ability of the tVMD method proposed in this paper to decompose trends in electricity price data, comparing it with the following three models: one without decomposition and two with other varieties of decomposition methods (EMD and EEMD models). The research results for the AUS-EP and SGP-EP markets are summarized in Table 6 and Figure 4, with specific comparative analyses as follows.

(a) Due to various uncertainties in the electricity market, electricity price series are typical examples of nonlinear and non-stationary data. Decomposition techniques can significantly improve predictive performance. For the AUS-EP, which is characterized by significant price fluctuations, the multi-step forecasting results of a combined deep learning and machine learning approach without data decomposition are $MAE_{ELsR}^{step} = \{10.7941, 13.8730, 14.6527\}$, $MAPE_{ELsR}^{step} = \{22.4224\%, 31.1126\%, 34.2981\% \}$, and $RMSE_{ELsR}^{step} = \{15.3988, 18.7547, 19.3794\}$. And for the relatively stable SGP-EP, the combined model ELsR without data decomposition processing can obtain $MAE_{ELsR}^{step} = \{5.8931, 8.7458, 10.9286\}$, $MAPE_{ELsR}^{step} = \{5.8611\%, 8.7855\%, 11.0888\% \}$, and $RMSE_{ELsR}^{step} = \{9.0542, 12.3424, 14.7562\}$ in the calculation of error indicators for multi-step prediction. However, the use of decomposition techniques in the forecasting model reduces the MAPE, MAE, and RMSE values.

Table 6. Results of the compared models for data decomposition methods.

Electricity Markets	Model	1-Step			2-Step			3-Step		
		MAE (USD/MWh)	MAPE (%)	RMSE (USD/MWh)	MAE (USD/MWh)	MAPE (%)	RMSE (USD/MWh)	MAE (USD/MWh)	MAPE (%)	RMSE (USD/MWh)
AUS-EP	ELsR	10.7941	22.4224	15.3988	13.8730	31.1126	18.7547	14.6527	34.2981	19.3794
	EMD-ELsR	8.1432	17.3928	10.2257	10.8504	23.2847	14.1286	13.9258	30.5456	18.3599
	EEMD-ELsR	9.6894	20.2068	13.2696	11.2209	23.7322	15.3640	11.6149	24.7575	15.7821
	Proposed HSEPF	7.3362	14.9100	9.7486	9.8568	20.9750	12.9024	10.2954	21.9254	13.2731
SGP-EP	ELsR	5.8931	5.8611	9.0542	8.7458	8.7855	12.3424	10.9286	11.0888	14.7562
	EMD-ELsR	4.9054	4.9167	7.3040	5.5805	5.6481	7.5667	8.4750	8.7852	11.2588
	EEMD-ELsR	4.3753	4.5751	6.4219	5.8097	5.8772	8.6673	8.1902	8.1718	11.7028
	Proposed HSEPF	4.2784	4.1918	5.9268	5.0703	4.8966	6.7367	6.7639	6.5111	9.1110

Note: ELsR represents the combination model of ELM, LSTM, and RF, which is modeled without any decomposition processing. This table shows the prediction results for different decompositions afterward, where $\mathbf{MAE} = \frac{1}{n} \sum_{t=1}^n |EP_t - EP'_t|$, $\mathbf{MAPE} = \frac{1}{n} \sum_{t=1}^n (|EP_t - EP'_t|) / (EP_t) * 100\%$, and $\mathbf{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (EP_t - EP'_t)^2}$. Bold numbers indicate that the model performed better in all comparison results.

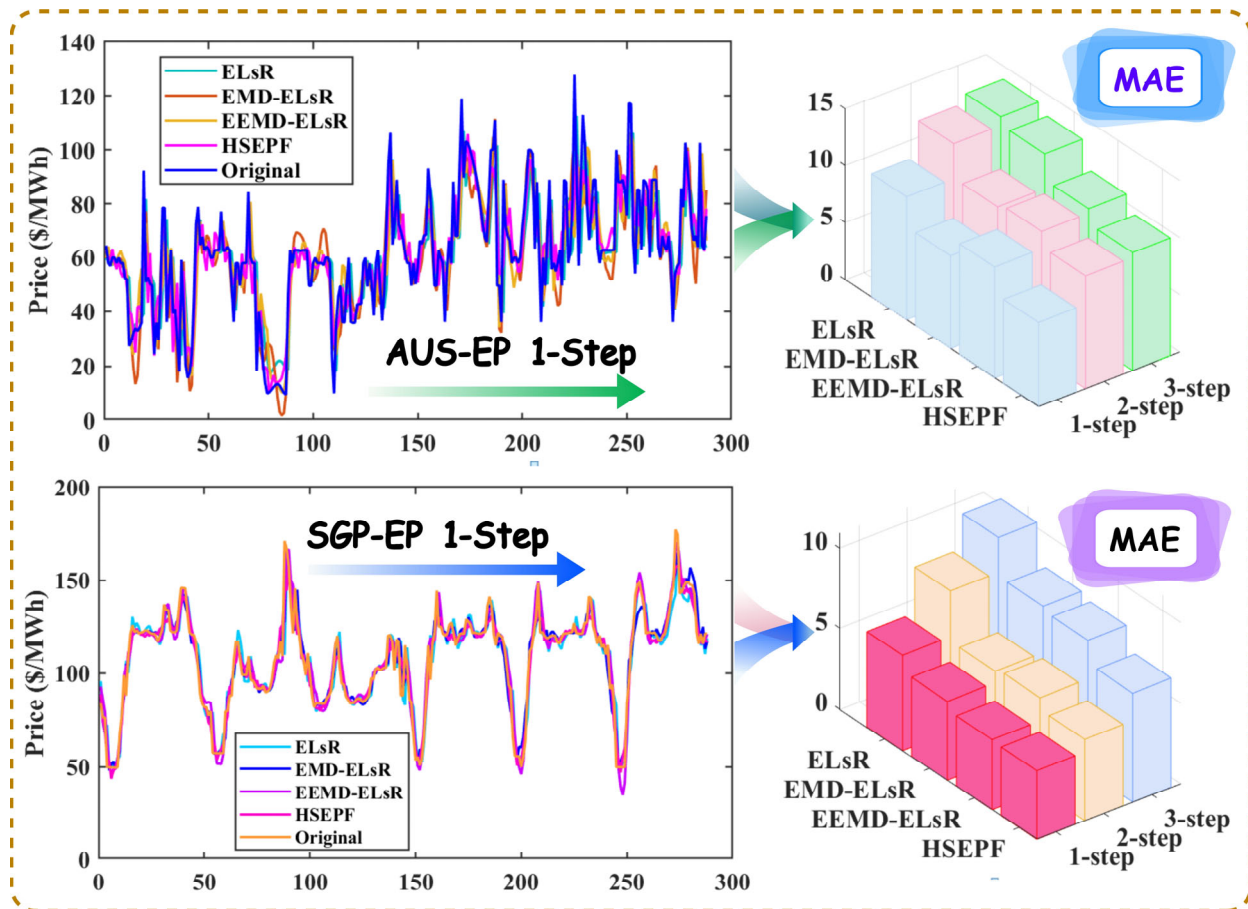


Figure 4. The forecasting performances in Case II.

(b) For the 1-step forecasting analysis using two market datasets, the accuracy improvements were most pronounced following decomposition through the tVMD method. Notably, there was a significant reduction in all error metrics for the AUS-EP: $\nabla MAPE_{AUS-EP}^{HSEPF} = 7.5124\%$, $\nabla RMAE_{AUS-EP}^{HSEPF} = 5.6502$ USD/MWh, and $\nabla MAE_{AUS-EP}^{HSEPF} = 3.4579$ USD/MWh. In the context of SGP-EP, forecasts derived from tVMD decomposition were most satisfactory, showing improved prediction precision with $\nabla RMSE_{SGP-EP}^{HSEPF} = 3.12174$ USD/MWh, $\nabla MAPE_{SGP-EP}^{HSEPF} = 1.6693\%$, and $\nabla MAE_{SGP-EP}^{HSEPF} = 1.6147$ USD/MWh.

(c) In multi-step forecasting performances across two major market datasets, models decomposed via the tVMD method consistently maintained excellent predictive performance. There are significant improvements in the 2-step and 3-step forecasting accuracy in AUS-EP: $\nabla MAPE_{AUS-EP\{2,3\}}^{HSEPF} = \{10.1376\%, 12.3727\%\}$. The improvement effects of EEMD and EMD decomposed models are similar in the 2-step, but EEMD performs better at the 3-step, though it still lags behind the proposed HSEPF model by about 3% in terms of the MAPE index. For SGP-EP, the forecasting errors decreased by $\nabla MAPE_{SGP-EP\{2,3\}}^{HSEPF} = \{3.8889\%, 4.5777\%\}$ for 2-step and 3-step forecasting, where the improvement effects of EEMD and EMD were closer.

Remark 2. Before integration into predictive models, it is crucial to stabilize electricity price data and to adaptively isolate several components that exhibit periodicity. Comparative analyses robustly validate that forecasting precision enhancements afforded by tVMD outperform those achieved with alternative decomposition methodologies. Furthermore, the benefits of this decomposition approach intensify progressively with the increase in forecast steps. The potential of tVMD as a superior tool for improving the reliability and accuracy of forecasting models is underscored by this finding.

4.3.4. Case III: Comparison with Other Combinatorial Weighting Methods

In this case, we compared the HSEPF model proposed in this paper with a combinatorial weighting strategy based on three different principles. Among them, simple average (SA) weightings were the most common method of model mixing, which reflects the average performance of multiple model combinations. The root mean square (RMS) weighting strategy amplifies large deviations by squaring them and emphasizes the volatility or difference in these values. The multi-objective Dragonfly optimization algorithm (MODA) is widely used in the fields of energy and financial forecasting. Compared with the model prediction results in Table 7, it was found that the HSEPF method achieves higher prediction accuracy and a smaller prediction error. Figure 5 shows the forecast trends and error indicators of each method for the SGP-EP market.

(a) In 1-step forecasting, the HSEPF technique outperformed the other weighting strategies in the forecasting performance in Singapore and Australian markets. tVMD-ELsR-SA, tVMD-ELsR-RMS, and tVMD-ELsR-MODA all performed well when applied to AUS-EP prediction, with tVMD-ELsR-SA being the strategy with the smaller forecasting errors, $MAE_{AUS-EP}^{SA} = 7.7246$, $MAPE_{AUS-EP}^{SA} = 16.2421\%$, and $RMSE_{AUS-EP}^{SA} = 10.4957$. However, there was a gap relative to HSEPF, $MAE_{AUS-EP}^{HSEPF} = 7.3362$, $MAPE_{AUS-EP}^{HSEPF} = 14.9100\%$, and $RMSE_{AUS-EP}^{HSEPF} = 9.7486$. On the other hand, among the comparison models in Singapore, tVMD-ELsR-RMS was the strategy with less prediction error with $MAE_{SGP-RP}^{RMS} = 4.4409$, $MAPE_{SGP-RP}^{RMS} = 4.3885\%$, and $RMSE_{SGP-RP}^{RMS} = 6.1485$. Similarly, the prediction performance lost out to the proposed HSEPF system with $MAE_{SGP-RP}^{HSEPF} = 4.2784$, $MAPE_{SGP-RP}^{HSEPF} = 4.1918\%$, and $RMSE_{SGP-RP}^{HSEPF} = 5.9268$.

(b) In the 2-step prediction of the AUS-EP, the prediction accuracy of the tVMD-ELsR-SA in MAPE index was better: $MAPE_{AUS-EP}^{SA} = 20.8713\%$. At this time, with $MAPE_{AUS-EP}^{HSEPF} = 20.9750\%$, $MAE_{AUS-EP}^{HSEPF} = 9.8568$, and $RMSE_{AUS-EP}^{HSEPF} = 12.9024$, the HSEPF model was slightly better than the tVMD-ELsR-SA model in the performance of MAE and MAPE prediction errors. In SGP-EP, the HSEPF model still maintained a good forecasting ability, where $MAE_{SGP-RP}^{HSEPF} = 5.0703$, $MAPE_{SGP-RP}^{HSEPF} = 4.8966\%$, $RMSE_{SGP-RP}^{HSEPF} = 6.7367$. The overall performance was better than that of the best-performing comparative models: tVMD-ELsR-MODA with $MAE_{SGP-RP}^{MODA} = 5.1215$, $MAPE_{SGP-RP}^{MODA} = 4.9873\%$, and $RMSE_{SGP-RP}^{MODA} = 6.8570$. Therefore, in general, the overall performance of the HSEPF model was better and can be widely used in different electricity price markets.

(c) For a 3-step prediction with a longer prediction period, the HSEPF model had excellent performance in model comparison. In the AUS-EP market, the forecast errors were $MAE_{AUS-EP}^{HSEPF} = 10.2954$, $MAPE_{AUS-EP}^{HSEPF} = 21.9254\%$, and $RMSE_{AUS-EP}^{HSEPF} = 13.2731$. The tVMD-ELsR-SA, tVMD-ELsR-RMS, and tVMD-ELsR-MODA models had closer prediction accuracies, and the MAPE errors were all above 22%. The tVMD-ELsR-RMS model had a better RMSE in the SGP-EP market, which was lower than the HSEPF model with $\nabla RMSE_{SGP-EP}^{RMS} = 0.1586$ USD/MWh. However, the HSEPF model of the other two indicators still maintained low prediction errors as $MAE_{SGP-EP}^{HSEPF} = 6.7639$, $MAPE_{SGP-EP}^{HSEPF} = 6.5111\%$.

Remark 3. *In this comprehensive study, we substantiated the robust theoretical foundation of the HSEPF model and its superiority in enhancing the predictive accuracy of complex systems. Comparative analyses revealed that other models failed to guarantee stable electricity price forecasts across varied market conditions. In contrast, the model proposed herein consistently demonstrates heightened precision in numerous forecasting scenarios, proficiently responding to the challenges presented by the significant fluctuations in electricity prices.*

Table 7. Results of the compared models for other combinatorial weighting methods.

Electricity Markets	Model	1-Step			2-Step			3-Step		
		MAE (USD/MWh)	MAPE (%)	RMSE (USD/MWh)	MAE (USD/MWh)	MAPE (%)	RMSE (USD/MWh)	MAE (USD/MWh)	MAPE (%)	RMSE (USD/MWh)
AUS-EP	tVMD-ELsR-SA	7.7246	16.2421	10.4957	9.8595	20.8713	12.9840	10.4218	22.4290	13.3598
	tVMD-ELsR-RMS	7.7316	16.2596	10.5043	9.8595	20.8732	12.9843	10.4232	22.4332	13.3602
	tVMD-ELsR-MODA	8.2481	17.7936	10.7628	10.3643	23.2164	13.4367	10.4906	22.8998	13.3782
	Proposed HSEPF	7.3362	14.9100	9.7486	9.8568	20.9750	12.9024	10.2954	21.9254	13.2731
SGP-EP	tVMD-ELsR-SA	4.6101	4.5774	6.3866	5.4094	5.3809	7.1779	6.9354	7.0019	9.0398
	tVMD-ELsR-RMS	4.4409	4.3885	6.1485	5.2488	5.1829	7.0080	6.8546	6.8528	8.9524
	tVMD-ELsR-MODA	4.5325	4.4305	6.3215	5.1215	4.9873	6.8570	6.8631	6.6923	8.9942
	Proposed HSEPF	4.2784	4.1918	5.9268	5.0703	4.8966	6.7367	6.7639	6.5111	9.1110

Note: ELsR represents the combination model of ELM, LSTM, and RF, which is modeled without any decomposition processing; SA is the simple average method; and RMS is the root mean square. $\mathbf{MAE} = \frac{1}{n} \sum_{t=1}^n |EP_t - EP_t''|$, $\mathbf{MAPE} = \frac{1}{n} \sum_{t=1}^n (|EP_t - EP_t''|) / (EP_t) * 100\%$, and $\mathbf{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (EP_t - EP_t'')^2}$. Bold numbers indicate that the model performed better in all comparison results.

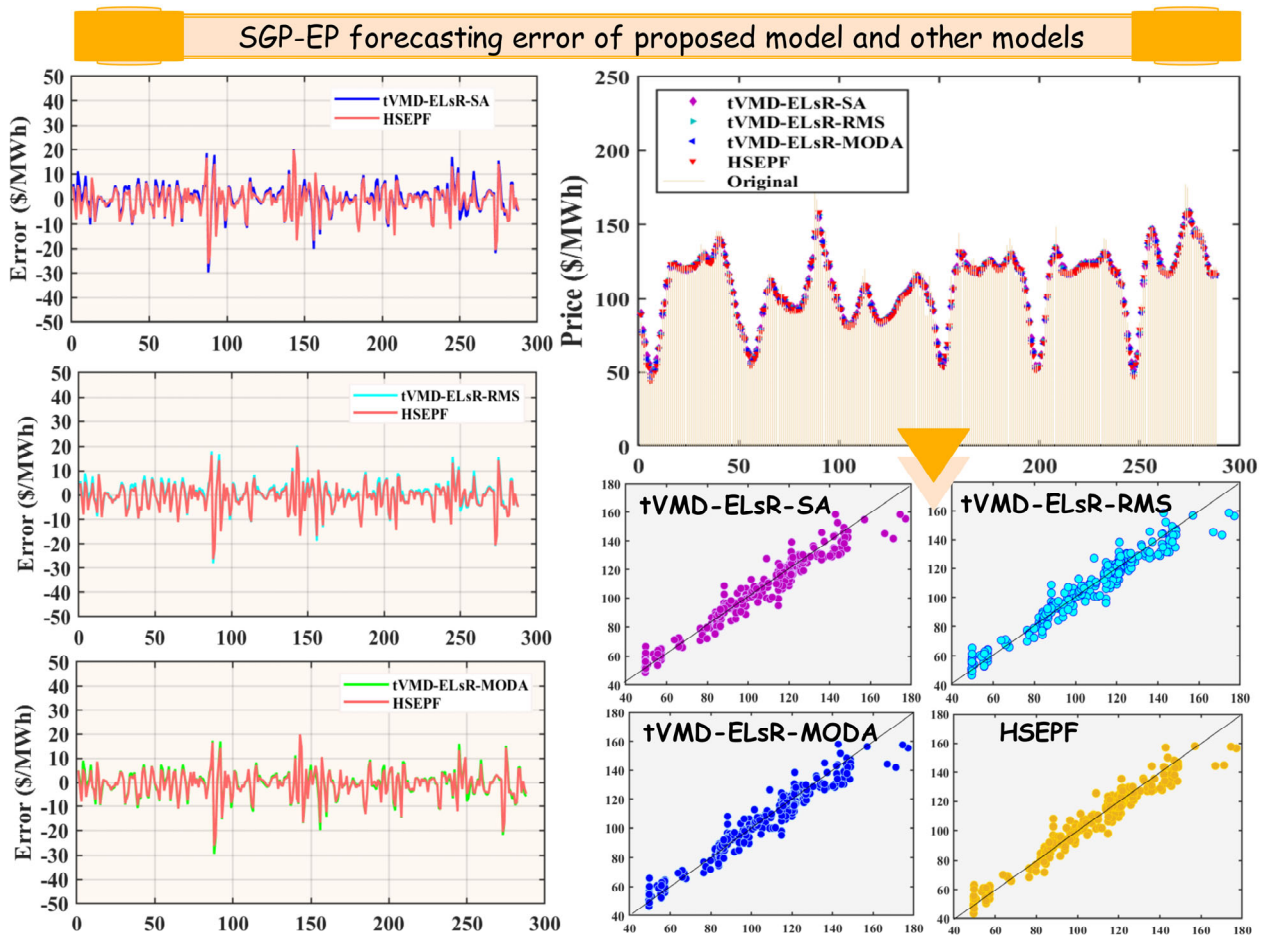


Figure 5. The forecasting performances in Case III for the SGP market.

5. Discussion

To further demonstrate the accuracy of the proposed HSEPF, we perform two additional sets of model tests.

5.1. DM Test

From the comparison of error indicators, we can understand the model prediction performance but cannot discern whether the model is statistically significant or not. The Diebold–Mariano (DM) test can solve this problem, where DM applies to the comparison of predictive ability between two or more models [42]. We can assume that the true value is $\{EP_t; t = 1, \dots, T\}$, the two predicted values in the different models are $\{EP''_{1,t}; t = 1, \dots, T\}$ and $\{EP''_{2,t}; t = 1, \dots, T\}$, respectively, and the prediction errors of the two models are $\{\mathfrak{S}_{i,t}; i = 1, 2\}$, where $\mathfrak{S}_{i,t} = EP_t - EP''_{i,t}$, $i = 1, 2$. We can assume that the loss associated with prediction i is a function of the prediction error, $\mathfrak{S}_{i,t}$, and denotes the loss function by $\{\Psi(\mathfrak{S}_{i,t}), i = 1, 2\}$. $\delta_t^\Delta = \Psi(\mathfrak{S}_{1,t}) - \Psi(\mathfrak{S}_{2,t})$, where δ_t^Δ is a first-order difference function of the loss function, if and only if $G(\delta_t^\Delta) = 0$, $EP''_{1,t}$ and $EP''_{2,t}$ have the same prediction accuracy, so we construct the null hypothesis $H_0 : G(\delta_t^\Delta) = 0 \forall t$ and the alternative hypothesis $H_1 : G(\delta_t^\Delta) \neq 0 \forall t$. Here, H_0 means that the two prediction models have the same prediction accuracy, and H_1 states that the two do not have related prediction accuracies. The DM statistic can be constructed as $\sqrt{T}(\bar{\delta}^\Delta - \mu) \xrightarrow{\delta^\Delta} N(0, 2\pi f_\delta(0))$, where $\bar{\delta}^\Delta$ is the mean of δ_t^Δ , μ is the overall mean of the loss function, and $f_\delta(0)$ is the spectral line density function when the loss difference is 0. The expression of $f_\delta(0)$ is as follows: $f_\delta(0) = \sum_{-\infty}^{+\infty} \gamma_\delta(\phi)$, where $\gamma_\delta(\phi)$ is the order self-covariance of δ_t^Δ , and the specific func-

tional equation is $\gamma_\delta(\phi) = E[(\delta_t^\Delta - \mu)(\delta_{t-\phi}^\Delta - \mu)]$, where $DM = \bar{\delta}^\Delta / \sqrt{2\pi f_\delta^\Delta(0)/T}; N(0, 1)$, $f_\delta^\Delta(0)$ and $f_\delta(0)$ are the consistent estimators when the null hypothesis is valid; the DM statistic obeys $N(0,1)$, HSEPF is used as the benchmark model for the DM test, and the other 14 control models are selected for the DM test. Table 8 presents the outcomes of the tests.

Table 8. DM test results.

Electricity Market	AUS-EP			SGP-EP		
	Model	1-Step	2-Step	3-Step	1-Step	2-Step
ARIMA	6.8043 ***	7.4900 ***	8.0509 ***	4.8452 ***	8.4343 ***	8.8529 ***
BP	7.1934 ***	9.0850 ***	8.8016 ***	3.9281 ***	7.5013 ***	8.2048 ***
ELM	7.1672 ***	8.3486 ***	8.6239 ***	3.8637 ***	7.2651 ***	8.1112 ***
LSTM	7.1313 ***	8.3954 ***	8.2850 ***	3.9831 ***	7.3610 ***	8.2098 ***
RF	7.1672 ***	8.7916 ***	8.3099 ***	4.2884 ***	6.2514 ***	6.4664 ***
tVMD-ELM	1.8208 *	0.5748	2.3776 **	0.4241	2.1572 **	0.7557
tVMD-LSTM	3.5418 ***	1.8477 *	0.1518	3.8626 ***	1.3480 *	1.8091 *
tVMD-RF	5.6180 ***	2.3662 **	1.8415 *	6.1373 ***	5.3650 ***	3.1619 ***
ELsR	7.1362 ***	8.5006 ***	8.1303 ***	3.5945 ***	6.4505 ***	6.8781 ***
EMD-ELsR	0.9941	2.0821 **	7.2362 ***	3.0622 ***	2.5157 **	6.1797 ***
EEMD-ELsR	5.8997 ***	3.9738 ***	3.5646 ***	1.1119	2.4545 **	3.2436 ***
tVMD-ELsR-SA	3.7340 ***	1.5210 *	1.5346 *	4.1060 ***	3.0546 ***	0.5165
tVMD-ELsR-RMS	3.7501 ***	1.5189 *	1.5313 *	3.0160 ***	2.4181 **	1.2954 *
tVMD-ELsR-MODA	3.3949 ***	2.3408 **	1.1199	4.5819 ***	1.6640 *	0.9010

Note: ***, **, and * represent the significance levels under 99%, 95%, and 90% confidence level, respectively. $DM = \bar{\delta}^\Delta / \sqrt{2\pi f_\delta^\Delta(0)/T}; N(0, 1)$. Here, $z_{0.01/2} = 2.576$, $z_{0.05/2} = 1.96$, and $z_{0.1/2} = 1.645$.

(a) In the comparison of single-mode models, whether for 1-step, 2-step, or 3-step predictions, the results uniformly achieved statistical significance at the 1% level. This outcome indicates a disparity in predictive capabilities between the proposed hybrid system and single models. Further analysis revealed that the predictive strength of models diminishes as the forward prediction steps increase. Moreover, the DM test values exhibited an increasing trend in 85% of the cases. This trend signifies that in 85% of instances, the decline in the predictive ability of single models is notably more pronounced than that of the HSEPF model, underscoring the superior multi-step forecasting stability of the hybrid model.

(b) In the comparison of machine learning models decomposed by tVMD, undecomposed deep learning integration models, and hybrid deep learning models based on other data decomposition techniques, it was found that there were cases where individual models performed better in prediction, but the prediction performance was not stable in multi-step prediction. Among them, the hybrid model ELsR without data decomposition had the highest DM values: $DM_{ELsR}^{AUS-EP} = \{7.1362, 8.5006, 8.1303\}$ and $DM_{ELsR}^{SGP-EP} = \{3.5945, 6.4505, 6.8781\}$, which are both greater than $z_{0.01/2} = 2.576$, suggesting that the prediction without data decomposition is the weakest. Further analysis revealed that the decomposition technique helped improve the prediction accuracy, whether it was a single model or a hybrid model. For example, in the 2-step prediction in AUS-EP, tVMD-ELM has $DM_{tVMD-ELM}^{AUS-EP} = 0.5748$, and in the 3-step prediction model, tVMD-LSTM has a statistically non-significant difference with the model proposed in this paper, $DM_{tVMD-LSTM}^{AUS-EP} = 0.1518$. It suggests that the forecasting performance of deep learning models, after undergoing trend decomposition through tVMD, closely matches that of the optimal model in specific instances.

(c) In the optimization comparison of different weighting strategies, the DM results in the 1-step prediction for the Australian and Singaporean markets, which both exceed $z_{0.01/2} = 2.576$, indicating a significant difference between the HSEPF model and others at a 99% confidence level. This suggests the effectiveness of the HSEPF model. In the multi-step prediction performance, the majority of models have DM values exceeding $z_{0.1/2} = 1.645$, signifying a significant difference between models at a 90% confidence level. The HSEPF model demonstrates a leading advantage in electricity price prediction. These findings

underscore the robustness and superiority of the HSEFP model in forecasting electricity prices, positioning them as a valuable tool in the energy market.

Remark 4. According to the results in Table 8, 90.48% of the models passed the test in the AUS-EP, while 88% of the models passed the test in the SGP-EP. This demonstrates a significant difference between the comparative models, confirming the predictive effectiveness of the HSEFP model.

5.2. Performance Improvement Analysis

In predictive research, the improvement percentage of predictive indicators serves as a crucial metric for assessing the predictive capabilities of models. This is quantified as $\overline{p_{improvement}} = \left| (Metric_{compared} - Metric_{HSEFP}) / Metric_{compared} \right|$, where $Metric_{compared}$ denotes the evaluation metric of the reference model, and $Metric_{HSEFP}$ signifies the predictive indicator from the composite system developed in this study. $\overline{p_{improvement}}$ quantifies the enhancement in predictive performance, with higher values reflecting the superior predictive accuracy of HSEFP. The comparative outcomes of models in AUS-EP and SGP-EP are summarized in Table 9.

(a) Upon analyzing the results of multi-step forecasting, it was evident that the HSEFP prediction system had significantly improved the accuracy and stability of electricity price forecasting in Australia. Particularly, for the single prediction model, the minimum values of the model average improvement rates were $\overline{p_{LSTM}^{Avg.MAE}} = 30.30\%$, $\overline{p_{ARIMA}^{Avg.MAPE}} = 32.78\%$, and $\overline{p_{RF}^{Avg.RMSE}} = 33.38\%$. However, when predicting using the single model after tVMD model decomposition, the model improvement rates were all below 10%, indicating the significant impact of tVMD technology on enhancing model prediction accuracy. The average improvement rates of models decomposed by EMD and EEMD are the minimal values of $\overline{p_{EMD-ELsR}^{Avg.RMSE}} = 13.68\%$ and $\overline{p_{EEMD-ELsR}^{Avg.MAE}} = 15.93\%$, respectively, suggesting that these two decomposition techniques have weaker capabilities in decomposing electricity price sequences compared to the tVMD technology proposed in this study. Among the three different combination weighted models, only $\overline{p_{tVMD-ELsR-MODA}^{Avg.MAPE}}$ exceeded 10%, indicating that the hybrid weighting strategy effectively enhanced model prediction capability, yet the HSEFP proposed in this study still outperformed the other comparative strategies.

(b) The electricity price sequence in Singapore exhibited relative stability, with small prediction errors. However, the improvement in forecast accuracy with the HSEFP model was also significant when considering the enhanced percentage. Compared to ARIMA, the HSEFP model demonstrates a superiority of $\overline{p_{ARIMA}^{Avg.RMSE}} = 51.10\%$. In the models decomposed by tVMD, the improvement rates were all above 10%, with $\overline{p_{tVMD-RF}^{Avg.MAE}} = 18.33\%$, $\overline{p_{tVMD-RF}^{Avg.MAPE}} = 22.91\%$, and $\overline{p_{tVMD-RF}^{Avg.RMSE}} = 15.31\%$, showcasing the excellent predictive performance of the MOWHO-weighted strategy. This finding highlights the capability of tVMD to handle electricity price data, significantly enhancing the predictive accuracy of deep learning algorithms and positioning it as a crucial tool for advancing predictive analytics in the field.

Remark 5. The improvement in percentage through model enhancement further confirms the predictive ability of the HSEFP system. It validates the significant effect of the tVMD technique in enhancing model prediction accuracy, and compared with other weighting strategies, the MOWHO algorithm has clear advantages, which can be further extended to the analysis of time-series prediction.

Table 9. Improvement ratio results (%).

Electricity Markets	Model	1-Step			2-Step			3-Step			Avg.		
		MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE
AUS-EP	ARIMA	35.34	31.88	39.26	35.10	30.35	36.48	39.66	36.10	41.35	36.70	32.78	39.03
	BP	36.91	42.87	38.56	33.71	35.69	34.23	34.29	41.41	35.33	34.97	39.99	36.04
	ELM	32.20	34.26	36.82	29.02	32.20	31.90	30.50	36.32	32.87	30.57	34.26	33.86
	LSTM	32.01	33.76	36.70	29.13	32.14	31.82	29.76	36.05	32.02	30.30	33.98	33.51
	RF	34.24	40.34	36.83	31.59	35.98	31.14	32.33	40.46	32.17	32.72	38.92	33.38
	tVMD-ELM	0.48	3.13	1.88	0.58	0.48	0.26	1.86	2.27	2.05	0.97	1.96	1.40
	tVMD-LSTM	2.52	4.44	4.67	0.43	1.01	1.34	0.00	0.55	0.03	0.98	2.00	2.01
	tVMD-RF	17.13	20.72	17.56	2.93	3.71	3.04	3.95	5.41	2.05	8.00	9.95	7.55
	ELsR	32.04	33.50	36.69	28.95	32.58	31.20	29.74	36.07	31.51	30.24	34.05	33.14
	EMD-ELsR	9.91	14.27	4.67	9.16	9.92	8.68	26.07	28.22	27.71	15.05	17.47	13.68
	EEMD- ELsR	24.29	26.21	26.53	12.16	11.62	16.02	11.36	11.44	15.90	15.93	16.42	19.48
	tVMD-ELsR-SA	5.03	8.20	7.12	0.03	0.50	0.63	1.21	2.25	0.65	2.09	3.65	2.80
	tVMD-ELsR-RMS	5.11	8.30	7.19	0.03	0.49	0.63	1.23	2.26	0.65	2.12	3.68	2.83
tVMD-ELsR-MODA	11.06	16.21	9.42	4.90	9.65	3.98	1.86	4.26	0.79	5.94	10.04	4.73	
SGP-EP	ARIMA	34.98	36.58	42.85	52.13	54.03	56.62	51.62	54.44	53.84	46.24	48.35	51.10
	BP	31.27	34.45	37.96	45.86	48.73	48.48	40.85	45.53	41.17	39.32	42.90	42.54
	ELM	28.33	31.02	36.68	43.32	45.93	47.16	39.56	43.63	40.50	37.07	40.19	41.45
	LSTM	28.91	31.03	37.08	44.39	47.12	48.02	39.57	44.11	40.84	37.62	40.76	41.98
	RF	32.85	33.41	37.53	45.71	48.06	46.59	38.13	43.19	38.05	38.90	41.55	40.73
	tVMD-ELM	0.76	1.76	0.37	3.81	5.60	2.91	1.64	4.81	1.09	2.07	4.06	1.46
	tVMD-LSTM	5.12	5.59	4.89	1.03	2.51	1.41	0.47	3.08	2.49	2.21	3.73	2.93
	tVMD-RF	21.73	24.35	19.49	19.82	24.56	18.30	13.44	19.83	8.14	18.33	22.91	15.31
	ELsR	27.40	28.48	34.54	42.03	44.26	45.42	38.11	41.28	38.26	35.84	38.01	39.41
	EMD-ELsR	12.78	14.74	18.86	9.14	13.31	10.97	20.19	25.89	19.08	14.04	17.98	16.30
	EEMD- ELsR	2.21	8.38	7.71	12.73	16.68	22.27	17.41	20.32	22.15	10.79	15.13	17.38
	tVMD-ELsR-SA	7.20	8.42	7.20	6.27	9.00	6.15	2.47	7.01	0.79	5.31	8.14	4.71
	tVMD-ELsR-RMS	3.66	4.48	3.61	3.40	5.52	3.87	1.32	4.99	1.77	2.79	5.00	3.08
tVMD-ELsR-MODA	5.61	5.39	6.24	1.00	1.82	1.75	1.45	2.71	1.30	2.68	3.30	3.10	

6. Conclusions and Future Work

The HSEPF system, introduced in this study, is a novel approach for forecasting short-term electricity prices. Initially, acknowledging the high volatility and irregular movements characteristic of electricity price sequences, the HSEPF system applied a stabilizing transformation to the raw data before forecasting. This involves the adaptive extraction of several feature components with distinct central frequencies. Specifically, this study employed the tVMD technique to isolate the trend, cyclical, and irregular components of the dataset. This approach significantly enhanced the efficiency of feature extraction in the forecasting model. Comparative analysis and discussion provided compelling evidence that the application of tVMD exerts a scalable impact on forecasting accuracy, warranting further exploration and validation in a broader data context.

Subsequently, the model inputs the decomposed sequence components into both linear and nonlinear machine learning models for separate predictions before integrating the results from each component. The final stage involved selecting the top three performing single models and employing the MOWHO model to obtain a weighted amalgamation of their outcomes. The experimental results demonstrated that the proposed HSEPF model was capable of predicting electricity prices across different market types. The forecast results in the Singapore market were as follows: $MAE_{\text{one-step}}^{\text{HSEPF}} = 4.2784$, $MAPE_{\text{one-step}}^{\text{HSEPF}} = 4.1918\%$, and $RMSE_{\text{one-step}}^{\text{HSEPF}} = 5.9268$. Australian market data fluctuated dramatically but still obtained smaller errors as follows: $MAE_{\text{one-step}}^{\text{HSEPF}} = 7.3362$, $MAPE_{\text{one-step}}^{\text{HSEPF}} = 14.9100\%$, and $RMSE_{\text{one-step}}^{\text{HSEPF}} = 9.7486$. This system not only enhances prediction accuracy through innovative decomposition and optimization techniques but also demonstrates its universality and effectiveness under different market conditions. This study highlights the potential of combining advanced analytical methods with deep learning models to improve predictive performance in the non-stationary, complex field of electricity price forecasting, providing valuable insights for academia and industry stakeholders.

In future research, two aspects can be explored. Firstly, a generalized and precise electricity price prediction model can be developed by considering various factors, such as user behavior, holidays, and weather. Secondly, future work can focus on peak electricity price prediction based on the current research to achieve higher prediction accuracy.

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

Funding: This research was funded by the HSSF of the Chinese Ministry of Education (No. 23YJA790053) and the NSFC (No. 12071302).

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Acknowledgments: The authors extend their gratitude to the editors and reviewers for their invaluable comments and suggestions.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Jiang, H.; Dong, Y.; Wang, J. Electricity Price Forecasting Using Quantile Regression Averaging with Nonconvex Regularization. *J. Forecast.* **2024**, *43*, 1859–1879. [[CrossRef](#)]
2. Indira, G.; Bhavani, M.; Brinda, R.; Zahira, R. Electricity Load Demand Prediction for Microgrid Energy Management System Using Hybrid Adaptive Barnacle-Mating Optimizer with Artificial Neural Network Algorithm. *Energy Technol.* **2024**, *12*, 2301091. [[CrossRef](#)]
3. Iftikhar, H.; Turpo-Chaparro, J.E.; Canas Rodrigues, P.; López-Gonzales, J.L. Forecasting Day-Ahead Electricity Prices for the Italian Electricity Market Using a New Decomposition—Combination Technique. *Energies* **2023**, *16*, 6669. [[CrossRef](#)]
4. Tan, Y.Q.; Shen, Y.X.; Yu, X.Y.; Lu, X. Day-Ahead Electricity Price Forecasting Employing a Novel Hybrid Frame of Deep Learning Methods: A Case Study in NSW, Australia. *Electr. Power Syst. Res.* **2023**, *220*, 109300. [[CrossRef](#)]

5. Sai, W.; Pan, Z.; Liu, S.; Jiao, Z.; Zhong, Z.; Miao, B.; Chan, S.H. Event-Driven Forecasting of Wholesale Electricity Price and Frequency Regulation Price Using Machine Learning Algorithms. *Appl. Energy* **2023**, *352*, 121989. [[CrossRef](#)]
6. Zhang, T.; Tang, Z.; Wu, J.; Du, X.; Chen, K. Short Term Electricity Price Forecasting Using a New Hybrid Model Based on Two-Layer Decomposition Technique and Ensemble Learning. *Electr. Power Syst. Res.* **2022**, *205*, 107762. [[CrossRef](#)]
7. Liu, D.; Wei, Y.; Yang, S.; Guan, Z. Electricity Price Forecast Using Combined Models with Adaptive Weights Selected and Errors Calibrated by Hidden Markov Model. *Math. Probl. Eng.* **2013**, *2013*, 648101. [[CrossRef](#)]
8. Girish, G.P. Spot Electricity Price Forecasting in Indian Electricity Market Using Autoregressive-GARCH Models. *Energy Strategy Rev.* **2016**, *11–12*, 52–57. [[CrossRef](#)]
9. Zhao, Z.; Wang, C.; Nokleby, M.; Miller, C.J. Improving Short-Term Electricity Price Forecasting Using Day-Ahead LMP with ARIMA Models. In Proceedings of the 2017 IEEE Power & Energy Society General Meeting, Chicago, IL, USA, 16–20 July 2017; pp. 1–5.
10. Imani, M.H.; Bompard, E.; Colella, P.; Huang, T. Forecasting Electricity Price in Different Time Horizons: An Application to the Italian Electricity Market. *IEEE Trans. Ind. Appl.* **2021**, *57*, 5726–5736. [[CrossRef](#)]
11. Singh, A.; Sahay, K.B. Short-Term Demand Forecasting by Using ANN Algorithms. In Proceedings of the 2018 International Electrical Engineering Congress, Krabi, Thailand, 7–9 March 2018; pp. 1–4. [[CrossRef](#)]
12. Wu, S.; He, L.; Zhang, Z.; Du, Y. Forecast of Short-Term Electricity Price Based on Data Analysis. *Math. Probl. Eng.* **2021**, *2021*, 6637183. [[CrossRef](#)]
13. Wang, P.; Xu, K.; Ding, Z.; Du, Y.; Liu, W.; Sun, B.; Zhu, Z.; Tang, H. An Online Electricity Market Price Forecasting Method Via Random Forest. *IEEE Trans. Ind. Appl.* **2022**, *58*, 7013–7021. [[CrossRef](#)]
14. Huang, S.; Shi, J.; Wang, B.; Lyu, J.; Nie, N.; Dong, X.; Li, H.; Zhang, S.; Ren, X. A Meta-Learning Based Method for Day Ahead Electricity Price Forecasting in Markets with Renewable Energy Resources. In Proceedings of the 2023 IEEE 12th Global Conference on Consumer Electronics, Nara, Japan, 10–13 October 2023; pp. 288–290. [[CrossRef](#)]
15. Rezaei, N.; Rajabi, R.; Estebarsari, A. Electricity Price Forecasting Model Based on Gated Recurrent Units. In Proceedings of the 2022 IEEE International Conference on Environment and Electrical Engineering and 2022 IEEE Industrial and Commercial Power Systems Europe, Prague, Czech Republic, 28 June–1 July 2022; pp. 1–5. [[CrossRef](#)]
16. Wang, B.; Wei, W.; Su, W. Short-Term Electricity Price Forecasting Based on Data Mining. In Proceedings of the 2022 2nd International Conference on Algorithms, High Performance Computing and Artificial Intelligence, Guangzhou, China, 21–23 October 2022; pp. 394–397. [[CrossRef](#)]
17. Yorat, E.; Ozbek, N.S.; Zor, K.; Saribulut, L. Day-Ahead Electricity Price Forecasting Using Artificial Intelligence-Based Algorithms. In Proceedings of the 2023 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies, Sakheer, Bahrain, 20–21 November 2023; pp. 121–126. [[CrossRef](#)]
18. Xu, H.; Hu, F.; Liang, X.; Gunmi, M.A. Attention Mechanism Multi-Size Depthwise Convolutional Long Short-Term Memory Neural Network for Forecasting Real-Time Electricity Prices. *IEEE Trans. Power Syst.* **2024**, *39*, 6277–6289. [[CrossRef](#)]
19. Yang, W.; Sun, S.; Hao, Y.; Wang, S. A Novel Machine Learning-Based Electricity Price Forecasting Model Based on Optimal Model Selection Strategy. *Energy* **2022**, *238*, 121989. [[CrossRef](#)]
20. Bozlak, Ç.B.; Yaşar, C.F. An Optimized Deep Learning Approach for Forecasting Day-Ahead Electricity Prices. *Electr. Power Syst. Res.* **2024**, *229*, 110129. [[CrossRef](#)]
21. Najafi, A.; Homaee, O.; Jasinski, M.; Golshan, M.; Leonowicz, Z. Application of Extreme Learning Machine-Autoencoder to Medium Term Electricity Price Forecasting. In Proceedings of the 2022 IEEE International Conference on Environment and Electrical Engineering and 2022 IEEE Industrial and Commercial Power Systems Europe, Prague, Czech Republic, 28 June–1 July 2022; pp. 1–4. [[CrossRef](#)]
22. Qu, Z.; He, L.; Ge, X.; Wang, F.; Xu, F.; Lu, J. A Two-Stage Forecasting Approach for Day-Ahead Electricity Price Based on Improved Wavelet Neural Network with ELM Initialization. *IEEE Trans. Ind. Appl.* **2024**, *60*, 5061–5073. [[CrossRef](#)]
23. Wang, J.; Liu, F.; Song, Y.; Zhao, J. A Novel Model: Dynamic Choice Artificial Neural Network (DCANN) for an Electricity Price Forecasting System. *Appl. Soft Comput. J.* **2016**, *48*, 281–297. [[CrossRef](#)]
24. Zhang, H.; Hu, W.; Member, S.; Cao, D.; Member, S.; Huang, Q.; Chen, Z.; Blaabjerg, F. A Temporal Convolutional Network Based Hybrid Model of Short-Term Electricity Price Forecasting. *CSEE J. Power Energy Syst.* **2021**, *10*, 1119–1130. [[CrossRef](#)]
25. Jiang, P.; Nie, Y.; Wang, J.; Huang, X. Multivariable Short-Term Electricity Price Forecasting Using Artificial Intelligence and Multi-Input Multi-Output Scheme. *Energy Econ.* **2023**, *117*, 106471. [[CrossRef](#)]
26. Wang, J.; Yang, W.; Du, P.; Niu, T. Outlier-Robust Hybrid Electricity Price Forecasting Model for Electricity Market Management. *J. Clean. Prod.* **2020**, *249*, 119318. [[CrossRef](#)]
27. Zhang, J.; Tan, Z.; Wei, Y. An Adaptive Hybrid Model for Short Term Electricity Price Forecasting. *Appl. Energy* **2020**, *258*, 114087. [[CrossRef](#)]
28. Huang, C.J.; Shen, Y.; Chen, Y.H.; Chen, H.C. A Novel Hybrid Deep Neural Network Model for Short-Term Electricity Price Forecasting. *Int. J. Energy Res.* **2021**, *45*, 2511–2532. [[CrossRef](#)]
29. Yang, Z.; Ce, L.; Lian, L. Electricity Price Forecasting by a Hybrid Model, Combining Wavelet Transform, ARMA and Kernel-Based Extreme Learning Machine Methods. *Appl. Energy* **2017**, *190*, 291–305. [[CrossRef](#)]
30. Zhang, X.; Wang, J.; Gao, Y. A Hybrid Short-Term Electricity Price Forecasting Framework: Cuckoo Search-Based Feature Selection with Singular Spectrum Analysis and SVM. *Energy Econ.* **2019**, *81*, 899–913. [[CrossRef](#)]

31. Xu, Y.; Li, J.; Wang, H.; Du, P. A Novel Probabilistic Forecasting System Based on Quantile Combination in Electricity Price. *Comput. Ind. Eng.* **2024**, *187*, 109834. [[CrossRef](#)]
32. Yang, W.; Wang, J.; Niu, T.; Du, P. A Novel System for Multi-Step Electricity Price Forecasting for Electricity Market Management. *Appl. Soft Comput. J.* **2020**, *88*, 106029. [[CrossRef](#)]
33. Kavakci, G.; Cicekdag, B.; Ertekin, S. Time Series Prediction of Solar Power Generation Using Trend Decomposition. *Energy Technol.* **2024**, *12*, 2300914. [[CrossRef](#)]
34. Dragomiretskiy, K.; Zosso, D. Variational Mode Decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544. [[CrossRef](#)]
35. Zhang, C.; Fu, Y.; Gong, L. Short-Term Electricity Price Forecast Using Frequency Analysis and Price Spikes Oversampling. *IEEE Trans. Power Syst.* **2023**, *38*, 4739–4751. [[CrossRef](#)]
36. Rockafellar, R.T. A Dual Approach to Solving Nonlinear Programming Problems by Unconstrained Optimization. *Math. Program.* **1973**, *5*, 354–373. [[CrossRef](#)]
37. Shi, C.; Wei, B.; Wei, S.; Wang, W.; Liu, H.; Liu, J. A Quantitative Discriminant Method of Elbow Point for the Optimal Number of Clusters in Clustering Algorithm. *EURASIP J. Wirel. Commun. Netw.* **2021**, *2021*, 31. [[CrossRef](#)]
38. Naruei, I.; Keynia, F. *Wild Horse Optimizer: A New Meta-Heuristic Algorithm for Solving Engineering Optimization Problems*; Springer: London, UK, 2022; Volume 38, ISBN 0123456789.
39. Wang, X.; An, Y.; Wang, J.; Shao, Y. A Novel Prediction-Integration Forecasting System for Short Wind Speed Based on Combined Data Preprocessing Technique and Weight Determination Strategy. *Energy Technol.* **2024**, *12*, 2300889. [[CrossRef](#)]
40. Zhang, W.; Kou, M.; Lv, M.; Shao, Y. Improved Combined System and Application to Precipitation Forecasting Model. *Alex. Eng. J.* **2022**, *61*, 12739–12757. [[CrossRef](#)]
41. Wang, J.; Guo, H.; Song, A. Photovoltaic Power Combination Prediction System Based on Improved Multi-Objective Optimization Algorithm and Nonlinear Weighting Strategy. *Expert Syst.* **2023**, *40*, e13209. [[CrossRef](#)]
42. Diebold, F.X.; Mariano, R.S. Comparing Predictive Accuracy. *J. Bus. Econ. Stat.* **1995**, *13*, 253–263. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.