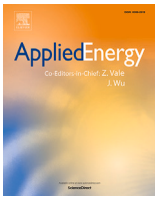
Applied Energy 395 (2025) 126197

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apen



Novel deep hybrid model for electricity price prediction based on dual decomposition

Sujan Ghimire [a], Thong Nguyen-Huy [b,c], Ravinesh C. Deo [a,b], David Casillas-Pérez [d,*], A. A. Masrur Ahmed [e], Sancho Salcedo-Sanz [f,a]

[a] Artificial Intelligence Applications Laboratory, School of Mathematics, Physics, and Computing, University of Southern Queensland, Springfield, QLD, 4300, Australia
[b] Centre for Applied Climate Sciences, University of Southern Queensland, Toowoomba, 4350, QLD, Australia
[c] Faculty of Information Technology, Thanh Do University, Kim Chung, Hoai Duc, Ha Noi, 100000, Viet Nam
[d] Department of Signal Processing and Communications, Universidad Rey Juan Carlos, Fuenlabrada, 28942, Madrid, Spain
[e] Science, Economics, and Insights Division, Department of Climate Change, Energy, the Environment and Water, NSW Government, Sydney, 2141, New South Wales, Australia
[f] Department of Signal Processing and Communications, Universidad de Alcalá, Alcalá de Henares, 28805, Madrid, Spain

HIGHLIGHTS

- Developed a data preprocessing approach termed the dual decomposition strategy.
- Designed an innovative prediction system employing Deep Residual Networks for forecasting electricity prices.
- Enhanced both the accuracy and stability of the forecasting system concurrently.
- Verified the applicability and effectiveness in the context of a real electricity market in Australia.

ARTICLE INFO

MSC:
0000
1111

PACS:
0000
1111

Keywords:
Deep learning
Convolutional neural network
Variational mode decomposition
Empirical wavelet transform
Residual connection
Bayesian optimization

ABSTRACT

Electricity price (EP) forecasting is vital for effective market operation, strategic planning, and risk management in deregulated energy systems. However, the inherent volatility and complexity of electricity prices, shaped by demand supply dynamics, weather variability, and regulatory interventions, pose substantial challenges to accurate prediction. This study introduces a novel hybrid framework designed to improve forecasting accuracy by leveraging both signal decomposition and deep learning techniques. Specifically, the method integrates Variational Mode Decomposition (VMD) and Empirical Wavelet Transform (EWT) for noise reduction and feature extraction, followed by a Multi Resolution Convolution (MRC) layer and a Bidirectional Long Short Term Memory (BiLSTM) network to capture multiscale temporal patterns in electricity price data. The model is applied to half hourly electricity price data from South Australia spanning January 2018 to December 2022. Its performance is benchmarked against a suite of traditional and hybrid models using a comprehensive set of twelve evaluation metrics. The results reveal that the proposed hybrid model consistently outperforms all baselines across seasons and forecast horizons. Notably, during the spring period, it achieved a Normalized Root Mean Square Error of ≈ 4.87 %, a Mean Absolute Percentage Error of ≈ 12.09 %, and a Global Performance Index of ≈ 3.22. These improvements demonstrate the model's ability to effectively handle the non-linear and nonstationary nature of EP. Overall, the findings underscore the potential of combining advanced decomposition methods with deep learning architectures to deliver more accurate and robust EP forecasts, thereby offering valuable support for decision making in complex and evolving energy markets.

* Corresponding author.
Email addresses: sujan.ghimire@unisq.edu.au (S. Ghimire), thong.nguyen-huy@unisq.edu.au, nhthong@thanhdouni.edu.vn (T. Nguyen-Huy), ravinesh.deo@unisq.edu.au (R.C. Deo), david.casillas@urjc.es (D. Casillas-Pérez), masrur.ahmed@environment.nsw.gov.au (A.A. Masrur Ahmed), sancho.salcedo@uah.es (S. Salcedo-Sanz).

https://doi.org/10.1016/j.apenergy.2025.126197
Received 28 December 2024; Received in revised form 22 April 2025; Accepted 23 May 2025
Available online 4 June 2025

List of acronyms

ACBFS Adaptive copula-based feature selection
Adam Adaptive moment estimation optimization
ADMM Alternating direction method of multipliers
AEMO Australian energy market operator
AI Artificial intelligence
ANEM Australian national electricity market
ANFIS Adaptive neuro-fuzzy inference system
ANN Artificial neural network
AR Autoregressive
ARD Automatic relevance determination
ARIFMA Autoregressive fractionally integrated moving average
ARIMA Autoregressive integrated moving average
ARMA Autoregressive moving average
ATT Attention mechanism
AUD Australian dollar
BDL Bayesian deep learning
BiLSTM Bi-Directional long short-term memory
BNN Bayesian neural network
BO Bayesian optimization
BOHB Bayesian optimization and hyperband
BP Backpropagation neural network
CEEMD Complete ensemble empirical mode decomposition
CEEMDAN Complementary ensemble empirical mode decomposition with adaptive noises
CNN Convolutional neural network
D-GCNN Deep graph convolutional neural network
DBN Deep belief network
DE Differential evolution
DL Deep learning
DNN Deep neural networks
EEMD Ensemble empirical mode decomposition
ELM Extreme learning machine
EMD Empirical mode decomposition
ENN or ELMAN Elman neural network
EP Electricity prices (AUD/MWh)
ERCRF Error correction random forest
ETR Extra tree regression
EWT Empirical wavelet transform
FA Firefly algorithm
FEEMD Fast ensemble empirical mode decomposition
FFNN Feed forward neural networks
GARCH Generalized autoregressive conditional heteroskedasticity
GBM Gradient boosting machine
GRU Gated recurrent unit
GVM Grey verhulst model
GWO Grey wolf optimizer
ICEEMDAN Improved complementary ensemble empirical mode decomposition with adaptive noises
IMD Improved mahalanobis distance
LASSO Least absolute shrinkage and selection operator
LEAR LASSO estimated autoregressive
LR Linear regression
LSSVM Least squares support vector machine
LSTM Long short-term memory
LSTR Logistic smooth transition regression
MAE Mean absolute error
MAPE Mean absolute percentage error
ML Machine learning
MLP Multi-layer perceptron
MOBBSA Multi-objective binary backtracking search algorithm
MoDWT Maximal overlap discrete wavelet transform
MOGWO Multi-objective grey wolf optimizer
MSVR Modified support vector regression
NARX Auto regressive model with exogenous
NN Neural network
PACF Partial autocorrelation function
PSO Particle swarm optimization
RBFNN Radial basis function neural network
Res. Residual term
RF Random forest regression
RMSE Root mean square error
RNN Recurrent neural network
RVM Relevance vector machine
SAE Stacked autoencoder
SARMAX Seasonal autoregressive moving average with exogenous variables
sMAPE Symmetric mean absolute percentage error
SSDAE Stacked sparse denoising autoencoder
TabNet Tabular neural network
VMD Variational mode decomposition
WNN Wavelet neural network
WT Wavelet transform
XGB eXtreme gradient boost

1. Introduction

The global electric power system is pivotal to national economies, with electricity prices (EP) shaped by factors like supply and demand dynamics, fuel prices, weather patterns, and regulatory frameworks [1]. These fluctuations influence investment decisions made by electricity producers, consumers, and associated industries worldwide. Additionally, the attainment of Sustainable Development Goal 7 (SDG 7) [2], which aims for universal access to affordable, reliable, sustainable, and modern energy, heavily relies on effective electricity pricing. High prices can restrict access for low-income households, impacting their ability to afford essential services and hindering economic progress. Conversely, well-planned pricing strategies can stimulate investments in robust infrastructure and promote the adoption of renewable energy sources, thereby advancing sustainability [3]. Competitive and stable electricity pricing also encourages energy efficiency improvements and the adoption of modern technologies [4]. Hence, the precise prediction of EP trends is essential for facilitating informed decisions on investments and consumption in a dynamic and competitive market environment.

Traditional approaches to EP forecasting include Multi-Agent Models, Fundamental Models, and Reduced-Form Models. Multi-Agent Models such as the Nash Cournot Framework, Supply Function Equilibrium, Strategic Production Cost Models, and Agent-Based Simulations aim to capture the strategic behavior of market participants [5,6]. These models can offer insights into market dynamics and policy effects, but their reliance on assumptions like perfect information and rational behavior can reduce their effectiveness in rapidly changing environments [7]. Fundamental Models, which are based on physical and economic relationships, are useful for long-term forecasting and scenario analysis [8–10]. However, they may not perform well in short-term prediction tasks and often require intensive computation [11]. Reduced-Form Models such as Jump Diffusion and Markov Regime Switching are designed to represent volatility and sudden price changes [12,13]. While effective in capturing regime shifts and stochastic behavior,

these models are sensitive to parameter estimation and require careful calibration [8].

Empirically, a wide range of forecasting methods has been applied, including classical time series models, regression techniques, and data-driven methods such as artificial neural networks and other machine learning algorithms [14]. As noted in the literature, no single method consistently outperforms others across all situations [15]. The selection of a suitable model depends on the characteristics of electricity price data, including seasonality, stationarity, linearity, and trends. Additionally, the forecasting task has become more complex due to external factors such as economic fluctuations, the COVID-19 pandemic, and natural disasters linked to climate change [16,17]. These challenges affect both electricity demand and price behavior. As a result, forecasting approaches in recent studies are generally grouped into three main categories: time series econometric models, machine learning and deep learning models, and hybrid models.

1. **Time Series Econometrics Models:** These models focus on statistical methods rooted in time-series analysis and econometrics principles. They often account for traditional factors such as seasonal patterns, trends, and autocorrelation [18].
2. **Machine and Deep Learning Models:** These modern approaches leverage advanced computational techniques to capture complex patterns and dependencies within data. They are capable of handling non-linearities and are increasingly applied due to their flexibility and ability to learn from large datasets [19].
3. **Hybrid Models:** Typically, these models integrate multiple machine learning (ML) or deep learning (DL) techniques, incorporating strategies such as data decomposition, feature selection, clustering, ensemble, and heuristic optimization to effectively estimate model hyperparameters [20]. This approach leverages the strengths of different methodologies to enhance predictive accuracy and robustness across diverse datasets and scenarios.

Each category offers distinct advantages and challenges, catering to different aspects of prediction accuracy and computational efficiency in varying contexts. Until recently, econometric models have dominated empirical applications, primarily focusing on linear time series models such as Auto Regressive Moving Average (ARMA), Logistic Smooth Transition Regression (LSTR), Auto Regressive Integrated Moving Average (ARIMA), Seasonal Auto Regressive Integrated Moving Average (SARIMA), Seasonal Autoregressive Moving Average with Exogenous Variables (SARMAX) and Vector Auto Regressive (VAR) [21]. In Ref. [22], an LSTR framework was employed to reveal that electricity spot price dynamics are influenced by non-linear intra-day relationships with key fundamental drivers, which differ across various time segments. Specifically, carbon prices were significant during off-peak hours, reserve margins played a crucial role during morning peak periods, and market concentration became more influential during the evening peak. These findings emphasize the need to account for regime-specific structural features in electricity price modeling. Building on this, a subsequent study [23] evaluated two hybrid models for day-ahead EP forecasting: one integrating fundamental supply stack modeling with econometric methods, and an enhanced version incorporating LSTR to capture structural regime shifts. The hybrid-LSTR model outperformed conventional SARMA, SARMAX, and stand-alone LSTR models in out-of-sample forecasts, although its prediction intervals were too narrow, indicating an under-representation of uncertainty. These studies collectively highlight the benefits of combining traditional and econometric techniques to improve forecasting accuracy in complex and evolving market environments. However, despite their computational efficiency and straightforward estimation, such models often fail to fully capture the non-linear dynamics and sharp fluctuations typical of electricity markets, limiting their robustness and reliability in practice [24].

In contrast, Machine Learning (ML) and Deep Learning (DL) models have recently emerged as powerful alternatives to traditional econometrics [25]. These AI-based approaches are particularly adept at handling complex, multivariate datasets with significant non-linearities, thereby offering an enhanced capability to capture and model intricate patterns. ML models have gained widespread application across various prediction domains, demonstrating their effectiveness in addressing the limitations of traditional statistical models in EP prediction [26]. Common ML models used for EP prediction include Random Forest (RFR), Support Vector Regression (SVR), and Artificial Neural Networks (ANN). For instance, a modified SVR (MSVR) has been shown to outperform Feed-Forward Neural Networks (FFNN) and Grey Verhulst Model (GVM) in long-term EP prediction [27], offering superior generalization ability to efficiently reach global solutions. However, MSVR's scalability for larger datasets remains limited. ANN, with its powerful multivariable mapping capability, addresses scalability issues but often gets trapped in local optima, losing sight of the internal influences of the data, thereby restricting further improvements in prediction accuracy.

Conversely, DL models address the limitations of traditional ML models by effectively handling large datasets and capturing complex underlying patterns within the data. DL is particularly proficient at feature extraction and modeling temporal dependencies. Feature extraction reduces the number of parameters, making the model more concise and efficient. Popular DL models for feature extraction in EP prediction include Stacked Autoencoder (SAE), Deep Belief Network (DBN), and Convolutional Neural Network (CNN). CNNs have demonstrated remarkable results in time series feature extraction due to their convolutional kernels, which autonomously mine relevant information from the data. For instance, in Ref. [28], the SEPNet model, which integrates CNN, Variational Mode Decomposition (VMD), and Gated Recurrent Unit (GRU), was proposed for one-hour-ahead EP prediction. The results demonstrated that SEPNet outperformed other models, with Mean Absolute Percentage Error ($MAPE$) and Root Mean Square Error ($RMSE$) reduced by $\approx$ 84 % and $\approx$ 81 %, respectively, across different seasons. Another study in Ref. [29] introduced Long Short-Term Memory (LSTM) optimized with Sequence Model-Based Optimization (SMBO) for EP prediction, achieving the lowest $RMSE$ of $\approx$ 1.72. Similarly, an LSTM model in Ref. [30] achieved a minimal prediction error of less than $\approx$ 1.5 % for one-day-ahead EP prediction. Moreover, the proposed LSTM network demonstrated superior performance compared to linear regression, Decision Trees, Deep Neural Networks (DNN), and Deep Feed-Forward Neural Networks (DFFNN).

In many instances, traditional single prediction models struggle to effectively analyse intricate relationships within non-stationary and non-linear datasets [31]. Consequently, researchers often employ hybrid models to enhance prediction accuracy either through model combination or by leveraging decomposition techniques [32]. For instance, in [33], a hybrid approach integrating Extreme Learning Machine (ELM) and Maximum Likelihood Estimation (MLE) was proposed for day-ahead EP prediction in the Australian electricity market. Similarly, in [34] authors introduced a hybrid evolutionary adaptive methodology combining Wavelet Transform (WT), Mutual Information (MI), Adaptive Neuro-Fuzzy Inference System (ANFIS), and Particle Swarm Optimization (PSO) for short-term EP prediction. In [35], a novel hybrid model is proposed employing Fast Ensemble Empirical Mode Decomposition (FEEMD), VMD, and Back-Propagation (BP) neural network optimized by the Firefly Algorithm (FA) for multi-step-ahead EP prediction. In [36] the authors uses Empirical Mode Decomposition with Support Vector Regression for electricity price prediction. In [37], Whale Optimization Algorithm (WOA)-based multivariate exponential smoothing Grey-Holt (GMHES) model is proposed for EP forecasting.

Recently in [38], ensembles of ML models have been tested for electricity price forecasting.

Moreover, the integration of CNN and LSTM models has emerged as a prominent approach in EP prediction. For example, in [39,40], CNN-LSTM was utilised to predict EP one hour ahead using the Pennsylvania–New Jersey-Maryland Interconnection (PJM) dataset, while Ref. [41] focused on short-term prediction up to one day ahead in the German electricity market, demonstrating superior average prediction performance. Additionally, Ref. [42] introduced a hybrid model that combines Wavelet Transform (WT), Sparse Autoencoder (SAE), and LSTM for EP prediction. In another study, Ref. [43] proposed a novel error compensation framework using VMD and RFR for half-hourly EP prediction in the Australian market. The research showed significant improvements with increases in Legates and McCabe Index by 15.97 %, 16.31 %, 20.23 %, 10.24 %, and 14.03 % for winter, autumn, spring, summer, and yearly predictions, respectively, compared to standalone ML and DL models. Furthermore, Ref. [44] presented a hybrid model based on WT, ARMA, and Kernel-based ELM for day-ahead EP prediction, demonstrating both linear and nonlinear prediction capabilities validated across three real electricity markets. In another study [45], an adaptive hybrid prediction model that combines VMD with self-adaptive particle swarm optimization (SAPSO) for predicting EP was proposed. This innovative approach aims to enhance prediction accuracy by effectively decomposing the original EP sequence using VMD and then DBN model was optimized by SAPSO. The model's performance was validated using quarterly data from diverse electricity markets including Australia, PJM, and Spain, demonstrating robust stability and high accuracy in prediction. Furthermore, a separate study [46] proposed a combined prediction model that integrates VMD with General Regression Neural Network (GRNN), focusing on modeling both EP and load series dynamics. This hybrid approach was specifically validated in PJM and Spanish electricity markets. By leveraging the capabilities of VMD to extract meaningful features and the predictive power of GRNN, the model aims to capture complex relationships within the EP and demand data, thereby improving the accuracy of short-term EP prediction.

Recent advancements in hybrid forecasting models have led to notable improvements across various domains such as EP forecasting, carbon pricing, water quality, building energy consumption, wind energy, and hydrological flow prediction. For instance, in carbon price forecasting, a framework combining the Maximal Information Coefficient (MIC), Improved Variational Mode Decomposition (IVMD), Autoformer, and Extreme Learning Machine (ELM) achieved outstanding accuracy with Coefficient of Determination values exceeding 0.99 for 15–25-day forecasts [47]. In water quality prediction, a hybrid model using MIC, decomposition–reconstruction strategies, Temporal Convolutional Networks (TCNs), and Gated Recurrent Units (GRUs) attained Nash–Sutcliffe Efficiency values ranging from 0.88 to 0.99 [48]. In the context of building energy consumption, a hybrid approach using an Improved Dung Beetle Optimization (IDBO) algorithm, VMD, and Bidirectional LSTM networks enhanced forecasting [49]. For wind energy prediction, combining Singular Spectral Analysis and VMD in a multi-stage hybrid framework reduced $MAPE$ by over 50 % for 72-h forecasts [50]. VMD-based models such as VMD-BiGRU and VMD-LSTM also outperformed conventional methods for short- to medium-term energy consumption forecasting across multiple countries [51]. In hydropower reservoir flow forecasting, coupling VMD with Echo State Networks (ESNs), Deep Echo State Networks (DeepESNs), and Next Generation Reservoir Computing (NGRC) resulted in significant error reductions for 7–21-day horizons [52]. A novel ensemble model combining VMD, AR, Elman Neural Network (ELMAN), Improved Bidirectional LSTM, and attention-based Convolutional Neural Networks further improved energy prices forecasting accuracy [53]. Furthermore, empirical wavelet transform (EWT) paired with an attention-based LSTM has demonstrated superior performance for monthly energy consumption prediction [54]. Recent research has also explored the integration of multi-head self-attention mechanisms, decomposition algorithm and CNN-based architectures, offering additional enhancements in EP forecasting capabilities [55–57].

Despite the advancements, these models typically rely on a single decomposition approach for data preprocessing, potentially limiting their ability to fully capture the main features of EP series. Moreover, challenges associated with predicting high-frequency signals remain largely unaddressed in previous studies, highlighting the need for improved data preprocessing techniques to enhance prediction performance. Therefore, enhancing data preprocessing techniques to boost prediction performance necessitates further investigation and refinement in future studies.

In response to the challenges outlined above and with the goal of advancing beyond existing methodologies, this study introduces a novel and robust model for predicting electricity prices (EP) tailored specifically for the electricity market. The framework encompasses four core modules: data preprocessing, optimization, prediction, and evaluation. Notably, the data preprocessing module integrates a novel dual decomposition approach utilising Variational Mode Decomposition (VMD) and Empirical Wavelet Transform (EWT). This innovative strategy addresses the shortcomings of prior decomposition techniques, yielding substantial improvements in preprocessing effectiveness. Simultaneously, to achieve dual objectives of enhancing prediction accuracy and stability, a sophisticated hybrid model is proposed. This model combines Multi-scale Residual Convolutional Neural Network (MRC-CNN) with a stacked Bi-directional Long Short-Term Memory (BiLSTM), optimized using Bayesian Optimization (BO) within the dedicated optimization module. This hybrid architecture significantly boosts the precision and stability of predictions for each decomposed component. Ultimately, the combined predictions from these components are aggregated to derive the final predicted values.

This study contributes significantly by addressing gaps in existing literature on "decomposition-prediction" methodologies and rectifying deficiencies observed in prior research. Unlike conventional hybrid models, the VMD-EWT-MRC-BiLSTM model developed in this study offers several distinctive advantages:

- Firstly, a two-layer decomposition approach incorporating VMD and EWT is applied to EP series. The residual after VMD decomposition often includes noise and other stochastic components of the EP series. By isolating and modeling these separately, the predictive model can reduce the impact of noise, leading to more stable and reliable predictions. EWT is further employed to decompose the residual terms post-VMD decomposition, thereby enhancing the accuracy of residual term prediction and overall hybrid model performance.
- Secondly, a novel hybrid model, Multi-scale Residual Convolutional Neural Network with Bi-directional Long Short-Term Memory (MRC-BiLSTM) is introduced. This innovative model architecture involves stacking BiLSTM layers on top of the MRC-CNN. The integration of these two advanced neural network structures allows the model to leverage the strengths of both: MRC-CNN's powerful feature extraction capabilities and BiLSTM's proficiency in handling sequential data. MRC-CNN effectively captures intricate patterns and features from the data through its multi-scale residual learning approach, ensuring robust and nuanced feature extraction. On the other hand, BiLSTM excels in sequence modeling by capturing long-term dependencies and bidirectional relationships within the data sequence. By combining these capabilities, the MRC-BiLSTM model enhances overall prediction accuracy, making it highly effective for complex and dynamic EP prediction. This hybrid approach not only improves the model's ability to understand and predict complex patterns but also ensures that both spatial and temporal dependencies within the data are effectively addressed.

- Lastly, the Bayesian Optimization Algorithm is employed to fine-tune the parameters of the MRC-BiLSTM model during the training process. This optimization technique systematically searches the parameter space and efficiently identifies the optimal set of hyper-parameters that enhance the model's performance. By leveraging Bayesian Optimization, the model can achieve a higher level of stability and reliability in its predictions. This method not only ensures that the MRC-BiLSTM model is accurately calibrated but also enhances its robustness against overfitting and underfitting. Consequently, the use of Bayesian Optimization contributes significantly to improving the model's predictive accuracy and generalization capability, making it a more effective tool for predicting EP.

The developed EP prediction model is rigorously validated across four seasonal as well as twelve months prediction scenarios, demonstrating resilience to seasonal effects and robust prediction performance. Experimental results underscore the effectiveness and reliability of the proposed system, positioning it as a valuable tool for risk analysis and decision-making within real-world electricity markets. Thus, this robust prediction system holds significant potential for widespread application, offering enhanced insights for strategic decision-making in electricity markets.

The structure of the paper unfolds as follows: Section 2 introduces methodologies. In Section 3, the dataset and the developed dual decomposition-based hybrid model for electricity price prediction are detailed. Section 4 provides an analysis of the results obtained and offers the authors' insights on the experimental findings. Finally, Section 5 provides the concluding remarks for the paper.

2. Methodologies

This section provides an overview of the study's primary focus on decomposition and deep hybrid models, specifically the Variational Modal Decomposition (VMD), Empirical Wavelet Transform (EWT), Multi-scale Residual Convolutional block (MRC) and Bidirectional Long Short-Term Memory (BiLSTM) models. Additionally, it is worth noting that detailed explanations of Deep learning models, such as Long Short-term Memory (LSTM), Convolutional Neural Networks (CNN) are discussed in more detail in [40,43,58–61], as they are widely recognised and well-established models for time-series prediction.

2.1. Variational modal decomposition

The Variational Modal Decomposition (VMD) technique is a relatively recent quasi-orthogonal method designed for the decomposition of input signals into intrinsic mode functions (IMFs) [62]. These IMFs serve to segregate distinct frequency components within the processed signal. The VMD method achieves this through the computation of a one-way frequency spectrum utilising the Hilbert Transform (HT) and subsequently shifting individual modes to the baseband. The bandwidth of each mode is estimated by evaluating the Dirichlet energy of the demodulated signal [63]. VMD departs from Empirical Mode Decomposition's (EMD) cubic spline curve-fitting approach for decomposition and instead adopts a novel method involving the formulation and resolution of a variational model [64]. This transformation eliminates the need to address the fitting challenge associated with local extreme points. Given this fundamental shift in the decomposition approach, VMD exhibits clear advantages over conventional signal processing techniques, particularly in terms of mitigating modal mixing and noise robustness [65]. It also presents specific computational benefits.

To obtain the bandwidths of the multiple modes after the VMD method, the original signal, denoted as $f(t)$, can be decomposed into K intrinsic mode functions (IMFs), represented by $u_i(t), i = 1, 2, \ldots, K$:

$$\min_{\{u_i\},\{\omega_i\}} \left\{ \sum_{i=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_i(t) \right] e^{-j\omega_i t} \right\|_2^2 \right\} \quad \text{subject to} \quad \sum_{i=1}^{K} u_i(t) = f(t) \tag{1}$$

where K represents the number of IMFs, $u_k(t)$ represents the set of the i^{th} IMF, ω_i represents the set of center frequencies corresponding to the i^{th} mode, and $\delta(t)$ is the Dirac function,$*$ for convolutional operations. To solve this problem, the optimization problem mentioned earlier, which includes equality constraints Eq. (1), can be converted into an unconstrained optimization problem through the augmentation of the Lagrangian function Eq. (2).

$$L\left(\{u_i\},\{\omega_i\},\lambda\right) = \alpha \sum_{i=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_i(t) \right] e^{-j\omega_i t} \right\|_2^2 + \left\| f(t) - \sum_{i=1}^{K} u_i(t) \right\|_2^2 + \lambda(t) \left(f(t) - \sum_{i=1}^{K} u_i(t) \right) \tag{2}$$

where λ is the Lagrange multiplier and α is the quadratic penalty term. Utilising the Alternating Direction Multiplication Method (ADMM) to solve the aforementioned problem results in the determination of both $\{u_i\}$ and $\{\omega_i\}$. The efficacy of VMD is contingent upon the choice of parameter $[K, \alpha]$. A high value for K may introduce noise redundancy, whereas an excessively low value may yield inadequate decomposition and mode aliasing. Likewise, a small α value might cause mode aliasing, while a large α value may result in substantial signal loss.

2.2. Empirical wavelet transform

Gilles introduced the Empirical Wavelet Transform (EWT) algorithm in 2013, offering a novel method for the adaptive analysis of time series signals that exhibit non-linearity and non-stationarity [66]. The fundamental idea at the heart of this algorithm involves the creation of a customised wavelet filter bank that dynamically adapts to the unique features of the input signal, leading to its segmentation into distinct modes. The EWT algorithm typically encompasses two primary phases:

- **Phase 1: Spectrum Computation and Boundary Identification:** The first step involves calculating the spectrum of the original signal through the Fast Fourier Transform (FFT) algorithm. Subsequently, local maxima within the spectrum are identified, serving as critical reference points for the partitioning process.
- **Phase 2: Empirical Wavelet Construction:** The second step entails the creation of empirical wavelets based on the previously determined boundaries. These empirical wavelets are instrumental in extracting the individual signal modes, effectively characterising the signal's various components.

During the segmentation process using EWT, specialised orthogonal wavelet filters are designed to enhance the amplitude and frequency modulation (AM-FM) characteristics in the Fourier spectrum. Subsequently, the Hilbert transform is applied to AM-FM to yield instantaneous frequency amplitude [67]. From a Fourier transform perspective, this transformation is analogous to creating a series of bandpass filters. The theoretical foundation of the EWT algorithm is elucidated as follows: it postulates that the Fourier spectrum can be partitioned into continuous N segments, with each segment denoted as $\Lambda_n = [\omega_{n-1}, \omega_n]$ where ω_n $(\omega_0 = 0, \omega_N = \pi)$ represents the boundaries of these segments. In this context, the empirical wavelet functions are considered as the wavelet filters corresponding to each Λ_n. The scaling function Eq. (3) and the empirical wavelet function Eq. (4) are subsequently derived using the Meyer wavelet, and their specific mathematical expressions are detailed below.

$$\widehat{\phi}_n = \begin{cases} 1, & \text{if} \quad |\omega| \le (1-\gamma)\omega_n, \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{|\omega|-(1-\gamma)\omega_n}{2\gamma\omega_n}\right)\right], & \text{if} \quad (1-\gamma)\omega_n \le |\omega| \le (1+\gamma)\omega_n, \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

$$\widehat{\psi}_n = \begin{cases} 1, & \text{if} \quad (1+\gamma)\omega_n \le |\omega| \le (1-\gamma)\omega_{n+1}, \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{|\omega|-(1-\gamma)\omega_{n+1}}{2\gamma\omega_{n+1}}\right)\right], & \text{if} \quad (1-\gamma)\omega_{n+1} \le |\omega| \le (1+\gamma)\omega_{n+1}, \\ \sin\left[\frac{\pi}{2}\beta\left(\frac{|\omega|-(1-\gamma)\omega_n}{2\gamma\omega_n}\right)\right], & \text{if} \quad (1-\gamma)\omega_n \le |\omega| \le (1+\gamma)\omega_n, \\ 0, & \text{otherwise,} \end{cases} \tag{4}$$

where $\beta(x)$ is defined as:

$$\beta(x) = \begin{cases} 0, & \text{if} \quad x \le 0, \\ 1, & \text{if} \quad x \ge 1, \\ x^4\left(35 - 84x + 70x^2 - 20x^3\right), & \text{if} \quad x \in [0,1]. \end{cases} \tag{5}$$

and

$$\beta(x) + \beta(1-x) = 1, \quad \forall x \in [0,1] \tag{6}$$

The approximation coefficients ($W_f^{\varepsilon}(0,t)$) are given by the inner product with the scaling function Eq. (7):

$$W_f^{\varepsilon}(0,t) = \langle f, \varphi_1 \rangle = \int f(\tau)\,\overline{\varphi_{1\,n}(\tau - t)}\,d\tau. \tag{7}$$

The detailed coefficients ($W_f^{\varepsilon}(n,t)$) are given by the inner products with the empirical wavelets Eq. (8):

$$W_f^{\varepsilon}(n,t) = \langle f, \varphi_n \rangle = \int f(\tau)\,\overline{\varphi_n(\tau - t)}\,d\tau \tag{8}$$

The original signal can be reconstructed as follows:

$$\begin{aligned} f(t) &= W_f^{\varepsilon}(0,t) * \varphi_1(t) + \sum_{n=1}^{N} W_f^{\varepsilon}(n,t) * \psi_n(t) \\ &= \mathcal{F}^{-1}\left(\widehat{W}_f^{\varepsilon}(0,\omega)\cdot\widehat{\varphi}_1(\omega) + \sum_{n=1}^{N}\widehat{W}_f^{\varepsilon}(n,\omega)\cdot\widehat{\psi}_n(\omega)\right) \end{aligned} \tag{9}$$

where $\widehat{W}_f^{\varepsilon}(n,\omega)$ represents the Fourier transform of $W_f^{\varepsilon}(n,t)$, $\widehat{\varphi}_1(\omega)$ is the Fourier transform of the scaling function, $\widehat{\psi}_n(\omega)$ is the Fourier transform of the wavelet functions, and $\mathcal{F}^{-1}$ refers to the inverse Fourier transform.$*$ represents convolution. According to the above formula, the empirical modes are obtained by Eqs. (10) and (11):

$$f_0(t) = W_f^{\varepsilon}(0,t) * \varphi_1(t) \tag{10}$$

$$f_k(t) = W_f^{\varepsilon}(k,t) * \psi_k(t) \tag{11}$$

Hence, the Empirical Wavelet Transform (EWT) is a dynamic decomposition of the Fourier spectrum, and it may not always align precisely with the positions of distinct modes. Employing the local max-min technique, the algorithm can effectively identify the period boundaries within the Fourier supports based on the information extracted from the processed spectrum.

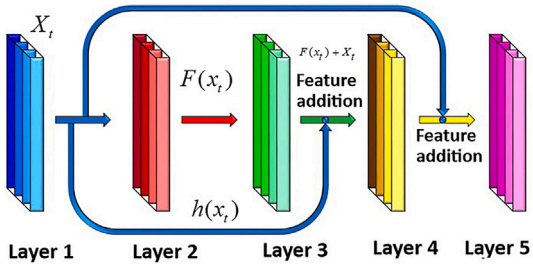


Fig. 1. Residual learning framework with skip connections.

2.3. Deep residual networks

In Deep Learning models, adding more layers enhances feature separation and identification, resulting in improved network performance. However, increasing the network's depth introduces challenges like vanishing or exploding gradients, which can hinder convergence of the optimization algorithm [68]. These issues can be mitigated through techniques like batch normalization and appropriate initialisation strategies, enabling the training of deep networks with numerous layers. While these methods facilitate the training of deeper networks, an issue called degradation arises as the network's depth increases. Degradation refers to the situation where the accuracy of the model on the training set decreases with more layers, leading to deep networks under-performing compared to shallow ones [69]. To address this degradation problem, deep residual networks were developed. The DNN-ResNet, proposed by He et al. [70], comprises multiple residual units, as illustrated in Fig. 1. Each residual block can be conceptualized as a shallow Neural Network. A standard residual block comprises weighted layers, activation functions, and an activation function, and the shortcut connection. Illustrated in Fig. 1, the shortcut connection allows the input to bypass one or more layers, facilitating the transmission of data to deeper layers in the neural network without any loss of information. The study conducted in [71] effectively demonstrates the pivotal role played by these shortcuts in the context of residual networks. The following Eqs. (12) and (13) encapsulate the representation of each residual unit:

$$y_t = h(x_t) + F(x_t,\ w_t) \tag{12}$$

$$x_{t+1} = f(y_i) \tag{13}$$

Where F is a residual function, f is an activation function ($ReLU$), w_t is the weight matrix, and x_t and y_t are the inputs and outputs of the t^{th} layer. The function h is an identity mapping given by (14):

$$h(x_t) = x_t \tag{14}$$

Skip connections, akin to residual connections, are shortcuts within the model's architecture. Through skip connections, each residual block can directly transmit its output to the final node of the model. This output is then combined with outputs from other levels and the single final output that has traversed all block transformations [72]. Employing skip connections in time series data integration enables the incorporation of features from earlier layers, both before and after model nonlinear transformations, fostering a comprehensive understanding of data trends and seasonality. Consequently, this contributes to improved network convergence [73]. Inspired by the concept of Residual units, our proposed model incorporates skip connections.

2.4. Deep hybrid MRC-BiLSTM model architecture

Drawing inspiration from ResNet [74] and driven by the necessity to address time series challenges, this study has innovatively devised the Multi-scale Residual Convolutional block (MRC). This building

block is specifically tailored for one-dimensional Convolutional Neural Network (CNN) applications. The MRC is seamlessly integrated with a Bi-directional LSTM network, giving rise to a novel deep hybrid network architecture referred to as MRC-BiLSTM, engineered for the task of predicting half-hourly electricity price time series. In the subsequent sections, we will commence by providing an exposition on the design principles behind the multi-scale residual block. Subsequently, we will elaborate on the structure of the MRC-BiLSTM model, elucidating how this new building block collaborates with BiLSTM to yield electricity price predictions.

2.4.1. The multi-scale residual CNN block

Convolutional Neural Network (CNN) is a widely recognised deep feature extraction framework comprising convolutional, pooling, fully connected, and regression layers. The convolutional layers execute convolutions, employing various filters to generate feature maps. Following convolution, a pooling operation is carried out to reduce feature map resolution. Subsequently, fully connected layers come into play, wherein each neuron is interconnected with neurons in the subsequent layer. The primary concept underpinning the use of fully connected layers is to encapsulate the learned global features distribution within a unified space for higher-level reasoning, thereby aiding in the prediction of the ultimate output through the regression layer. The mathematical representations of the Convolution layer and the pooling operation are denoted in Eqs. (15) and (16), respectively.

$$C^{l}_{m,n,o} = F\left(\left(W^{l}_{o}\right)^{t} X^{l}_{m,n} + b^{l}_{k}\right) \tag{15}$$

$$P^{l}_{m,n,o} = Pool\left(X^{l}_{m+n,o}\right) \tag{16}$$

where b^{l}_{o} represents the bias of the convolutional filter in the l^{th} layer, $X^{l}_{m,n}$ represents the locations of the input region in the l^{th} layer, and F is the activation function.

In practice, the inclusion of multiple CNN layers is a common approach for addressing complex problems and enhancing accuracy performance. The rationale behind incorporating additional layers is to extract intricate features from the input data. However, increasing the number of layers in the network can lead to higher training and testing errors, heightened complexity, and diminished performance. To resolve the challenges associated with training deep-layer networks, the concept of residual networks was introduced. These networks utilise a residual connection, which links the output of an earlier convolutional layer to the input of a subsequent convolutional layer located several layers further along the network. This connection effectively skips a number of intermediate convolutional steps, as depicted in Fig. 2. In terms of how the extracted features from the mainstream model input and the input from the preceding convolutional layer are combined, this study has adopted a structure similar to DenseNet [75]. Instead of summing the features, they are concatenated. This concatenation of features offers the advantage of allowing each additional layer in the network to incorporate information from some subset of prior residual block outputs. This results in increased information available for the model's utilisation, ideally leading to a more precise model fit [76].

2.4.2. MRC-BiLSTM model

The deep hybrid architecture, comprised of Multi-scale Residual CNN (MRC) and a stacked Bi-directional LSTM (BiLSTM), effectively captures the intricate and irregular trends while extracting complex spatiotemporal features from the electricity price (EP) data (Fig. 3). In this process, the residual CNN layers play a pivotal role in feature extraction [77,78], forwarding the extracted features to the stacked BiLSTM to uncover the sequential patterns connecting them. BiLSTM, an enhanced variant of the LSTM model [79], is employed due to its suitability for EP prediction. This choice arises from the limitations of traditional Recurrent Neural Networks (RNN), such as vanishing gradient problems. LSTM

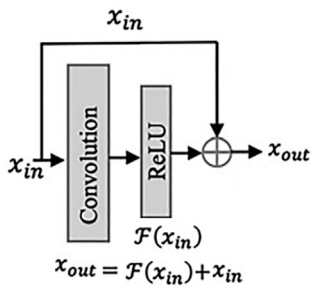


Fig. 2. An example of a skip connection, connecting the input to the output of one convolution block.

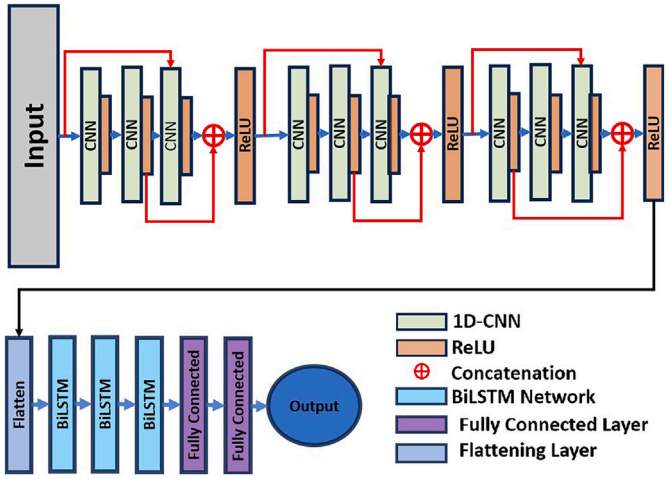


Fig. 3. The overall framework of Multi-scale residual CNN and BiLSTM to predict half-hourly electricity price.

addresses this issue by facilitating the learning of both short-term and long-term dependencies. Information is propagated through the forget gate(f_t), input gate(i_t), and output gate(o_t), thereby controlling the extent of memory retention and information forgetting regarding prior and current data. The forget gate determines the exclusion of irrelevant information. Specifically, it processes the output from the previous time step and the input from the current time step to compute the forget gate's output. The standard expression of the LSTM is provided in the Eqs. (17)–(22):

$$i_t = \sigma_g\left(W_i s_{t-1} + U_i x_t + b_i\right) \tag{17}$$

$$o_t = \sigma_g\left(W_o s_{t-1} + U_o x_t + b_o\right) \tag{18}$$

$$f_t = \sigma_g\left(W_f s_{t-1} + U_f x_t + b_f\right) \tag{19}$$

$$\tilde{s} = \phi\left(W\left(o_t \odot s_{t-1}\right) + U x_t + b\right) \tag{20}$$

$$s_t = f_t \odot s_{t-1} + i_t \odot \tilde{s}_t \tag{21}$$

$$y_t = o_t \odot \sigma_y\left(s_t\right) \tag{22}$$

where s_t and s_{t-1} are the current and previous states; W, U are the weights in the networks; b is the bias; σ_g, ϕ, σ_y are activation functions; and $\odot$ denotes the element-wise multiplication. For BiLSTM, the architecture consists of both forward and backward LSTM layers,

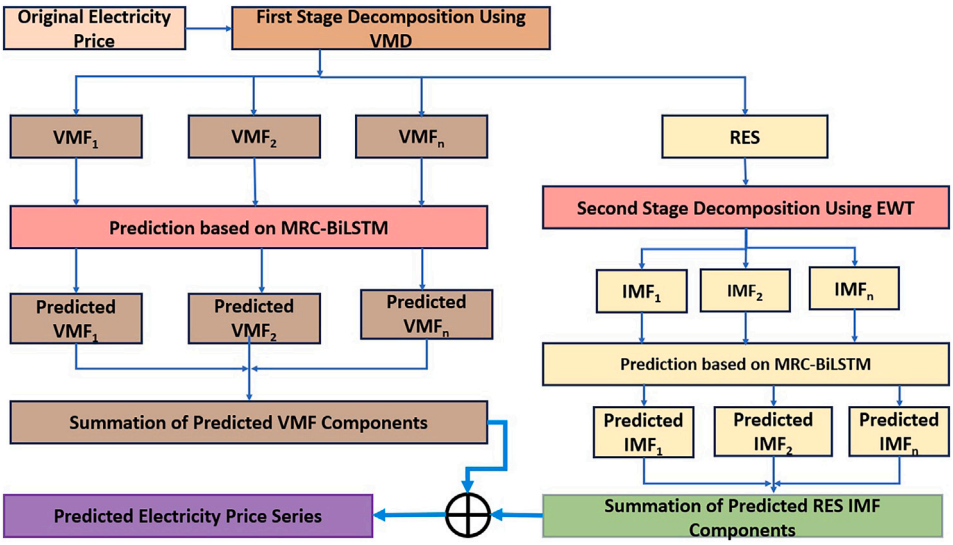


Fig. 4. Flow chart of the proposed VMD-EWT-MRC-BiLSTM deep hybrid model.

where the inputs from the forward and backward layers are handled simultaneously by the output layer as follows:

$$\overrightarrow{h_t} = H\left(W_1 x_t + W_2 \overrightarrow{h_{t-1}} + \vec{b}\right) \tag{23}$$

$$\overleftarrow{h_t} = H\left(W_3 x_t + W_5 \overleftarrow{h_{t-1}} + \overleftarrow{b}\right) \tag{24}$$

$$y_t = W_4 \overrightarrow{h_t} + W_6 \overleftarrow{h_t} + b_y \tag{25}$$

where $\overleftarrow{h_t}$, $\overrightarrow{h_t}$ and y_t are the vectors for backward propagation, forward propagation and an output layer, respectively; W_1, W_2, W_3, W_4, W_5 and W_6 are the corresponding weight coefficients; $\overleftrightarrow{b}$, $\vec{b}$ and b_y are the corresponding bias vectors.

2.5. Proposed VMD-EWT-MRC-BiLSTM model

As previously mentioned, the electricity price sequence exhibits a multitude of complex characteristics, and relying on a single prediction method often results in limited accuracy. The VMD technique is particularly valuable in decomposing intricate signals into several regular modal components, significantly enhancing prediction accuracy. However, in contrast to typical residuals obtained through other decomposition methods, the residual term derived from VMD tends to be more intricate. In previous research employing the VMD technique, insufficient attention has been given to the wealth of information contained within the residual term, leading to diminished overall prediction accuracy of the model. To address this issue, this study introduces a two-layer decomposition technique to perform comprehensive decomposition and prediction on the residual term produced by VMD. This approach aims to bolster the prediction performance of the residual term. The deep hybrid model presented in this paper (Fig. 4), VMD-EWT-MRC-BiLSTM, consists of four essential modules:

- **Pre-processing Module:** The VMD method is applied to decompose the initial *EP* sequence into individual modal *VMF* components. The cumulative sum of these *VMF* components is then subtracted from the original time series data to derive the residual term (*RES*) subsequent to the VMD decomposition.
- **Normalization and optimization:** Each *VMF* component is normalized, suitable training and testing datasets are chosen, and the training of each *VMF* component is carried out using the MRC-BiLSTM model. Within the MRC-BiLSTM training process, the model's parameters are fine-tuned using the Bayesian Optimization (BO) algorithm. This parameter optimization and training procedure lead to the generation of prediction results for each VMF component sub-sequence.
- **Second Decomposition:** The *RES* undergoes the EWT decomposition, followed by utilising the MRC-BiLSTM model for the prediction of each subsequence resulting from the EWT decomposition.
- **Integration Module:** The predictions for each *VMF* component and residual component are combined to yield the final *EP* prediction.

3. Electricity price data and predictive model development

Electricity price sequences obtained from the Australian Energy Market Operator (https://www.aemo.com.au/) are employed in this study to assess the predictive performance of the VMD-EWT-MRC-BiLSTM model. The focus is primarily on the half-hour *EP; AUD/MWh* of South Australia (SA). The dataset encompasses observations from January 1, 2018, to December 31, 2022, totaling 77,446 data points. Within this dataset, data from January 1, 2018, to December 31, 2021 (63,540 data points) serve as the training dataset, while data from January 1, 2021, to December 31, 2022 (13,906 data points) are designated as the test dataset. Descriptive statistical outcomes for *EP* series are detailed in Table 1. The distributional characteristics of the *EP* series generally exhibit non-normality. As shown in Table 1, the *EP* series display positive skewness (*Skew*), and their kurtosis (*Kurt*) surpasses three, indicating a leptokurtic distribution. To assess the normality of the daily

Table 1
The statistical description of the electricity price(*EP*) dataset of South Australia. Note *Mean, Max, Min and Std. dev.* refers to the mean, maximum, minimum and standard deviation value of *EP* in *AUD/MWh* respectively.

Mean	Max.	Min.	Std. dev.	Skew	Kurt	JB_{stat}	JB_{pval}	ADF_{stat}	ADF_{pval}
63.32	981.65	1.00	56.40	3.41	25.40	400,064.14	0.00	−8.06	0.00

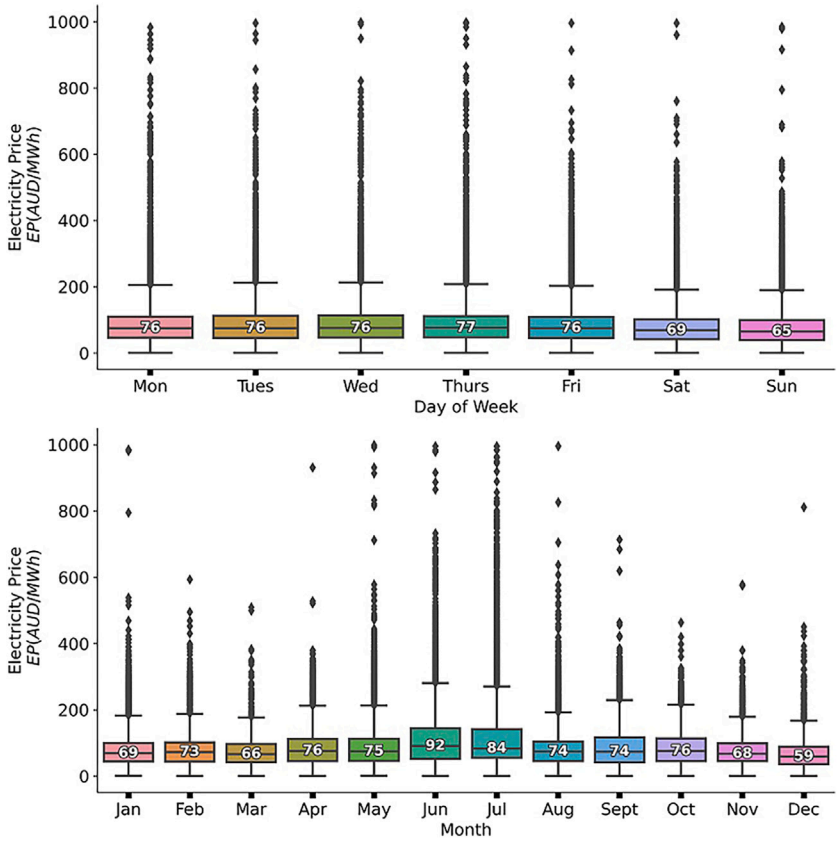


Fig. 5. Box-plot of daily and monthly cycle attributes of the *EP*.

distribution of *EP*, the Jarque–Bera statistic test (JB_{stat}) [80] was employed, as presented in Table 1. All the $p-values$ are below the 0.05 level of significance, leading to the rejection of the null hypothesis. This suggests that the half-hourly *EP* series are not adequately approximated by a normal distribution. Additionally, each price series undergoes a unit root test using the augmented Dickey-Fuller (ADF) test. With $p-values$ below 0.05, it can be concluded that the *EP* series exhibit stationarity.

Moreover, Figs. 5 and 6 employ box plots to provide a visual representation of noteworthy attributes related to *EP* concerning factors such as the day of the week, month, hourly variations, and annual trends. In the top and bottom panels of Fig. 5, the box plots illustrate the weekly and monthly seasonality in hourly prices. The top panel clearly illustrates that, on average, weekdays exhibit higher prices compared to Saturdays and Sundays. When considering annual seasonality, it aligns with expectations, showing that winter months (June, July, and August) tend to experience higher prices in contrast to summer months (December, January, and February). Notably, the price trajectory initiates an increase at approximately 6:00 a.m. as people wake up and the workday commences. The mean prices steadily rise throughout the day, reaching a peak at around 6:00 p.m., and then decrease rapidly after dark (Fig. 6 top panel). This observed price pattern is most likely a result of increased demand for air conditioning during these hours. Furthermore, Fig. 6 lower panel displays the annual *EP* distribution spanning from 2018 to 2022. It's essential to observe that the *EP* for the year 2020 stands out as notably lower compared to other years due to the impact of the COVID-19 pandemic. On March 11, 2020, the World Health Organisation [81] declared the COVID-19 pandemic, leading to stringent lockdown measures implemented by Australian government (23-March-2020–15-May-2020). These measures included restrictions on movement, event cancellations, and the suspension of numerous commercial activities. These changes in societal behaviour also had repercussions on the electricity sector, resulting in alterations in consumption patterns. Notably, in Australia, total electric demand during the pandemic decreased by 5.6 % [82], with a significant 40 % [83] increase in residential demand.

3.1. Decomposition of electricity price series

In this study, we initially employed VMD to decompose the original *EP* series, characterised by complex variations, into subsequences known as $VMFs$, which exhibit simpler variation characteristics. Despite this decomposition, the variation characteristics of the residual sequence (VMD_RES) that remained after VMD were still somewhat complex. To further reduce the complexity, we applied EWT to decompose the VMD_RES into $IMFs$. This additional decomposition resulted in a reduction of the non-stationary and non-linear characteristics in the decomposed subsequences, making them more amenable to prediction. However, the proper determination of parameters during VMD decomposition is crucial for achieving effective results. In our study, for VMD, we selected a penalty factor (α) of 4000, and the mode number (K) was determined using the central frequency iteration rule [84]. Fig. 7 illustrates the central frequency of each sub-component obtained by VMD for different values of mode number K. Notably, Fig. 7 reveals a consistent central frequency pattern when the decomposition mode number K falls within the range of 13 to 18, indicating an over-decomposition phenomenon. Therefore, for this study, we set the decomposition mode number K of VMD to 12.

Fig. 8 presents the original *EP* series and the $VMFs$ (IMFs of VMD) after the decomposition of the *EP* series using VMD. $VMF1$-$VMF12$

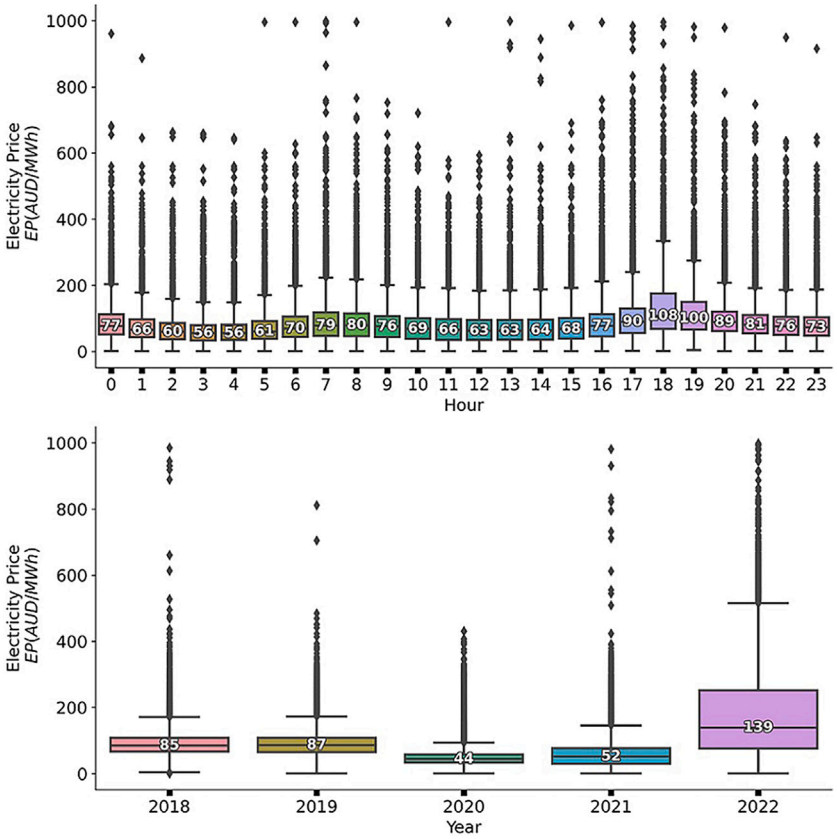


Fig. 6. Box-plot of hourly and yearly cycle attributes of the *EP*.

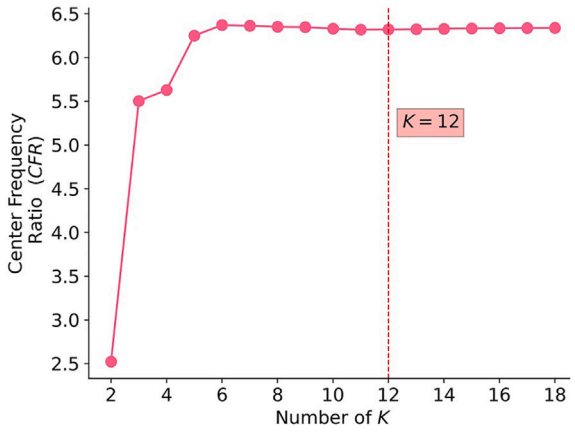


Fig. 7. Central frequency of each sub-signal (VMF) obtained by VMD with different decomposition mode number (K).

represent the components of the decomposed *EP* series, ranging from low to high frequency. Among these components, $VMF1$ and $VMF2$, representing the lowest frequency signals, reflect the trend changes in the original sequence, while $VMF3$–$VMF12$, which represent the highest frequency signals, reflect local volatility trends. Importantly, each mode retained the original *EP* series' characteristics and mitigated mode aliasing. Additionally, we observed that the residual component of VMD (VMD_RES), displayed in Fig. 8, exhibited fluctuations within the range of 0–200. This disordered and irregular residue, remaining after VMD decomposition, could potentially impact the accuracy and stability of prediction models.

To address this challenge, we applied the EWT algorithm to further decompose the VMD_RES. EWT reduced the randomness and fluctuation in the initial VMD_RES, enhancing the series' modeling capabilities and ultimately improving prediction accuracy. Before implementing EWT, we needed to pre-determine the mode number (N). To select the mode number, we calculated the Root Mean Square Error ($RMSE$) of EWT reconstruction and chose the mode that exhibited the lowest $RMSE$ error. Based on the findings presented in Fig. 9, we established the mode number for EWT decomposition as 8, with the aim of further reducing the complexity of the residual sequence and lessening the computational burden. The decomposition results of the residual sequence are displayed in Fig. 10, demonstrating a reduction in complexity following EWT decomposition. It is worth noting that the both training data and test data are separately decomposed to prevent the leakage of future information. This practice is essential for practical applications, particularly when the test dataset is unknown.

3.2. Data normalization

To ensure the prediction effectiveness of the VMD-EWT-MRC-BiLSTM model, we need to normalize the data for each decomposed series of every modal component(VMF and IMF) after decomposition. This study employs the min-max deviation normalization method for data transformation, as expressed below:

$$IMF' = \frac{IMF - IMF_{min}}{IMF_{max} - IMF_{min}} \tag{26}$$

where IMF' is normalized subsequence data, IMF is its original value, and IMF_{max} and IMF_{min} are its maximum and minimum values, respectively.

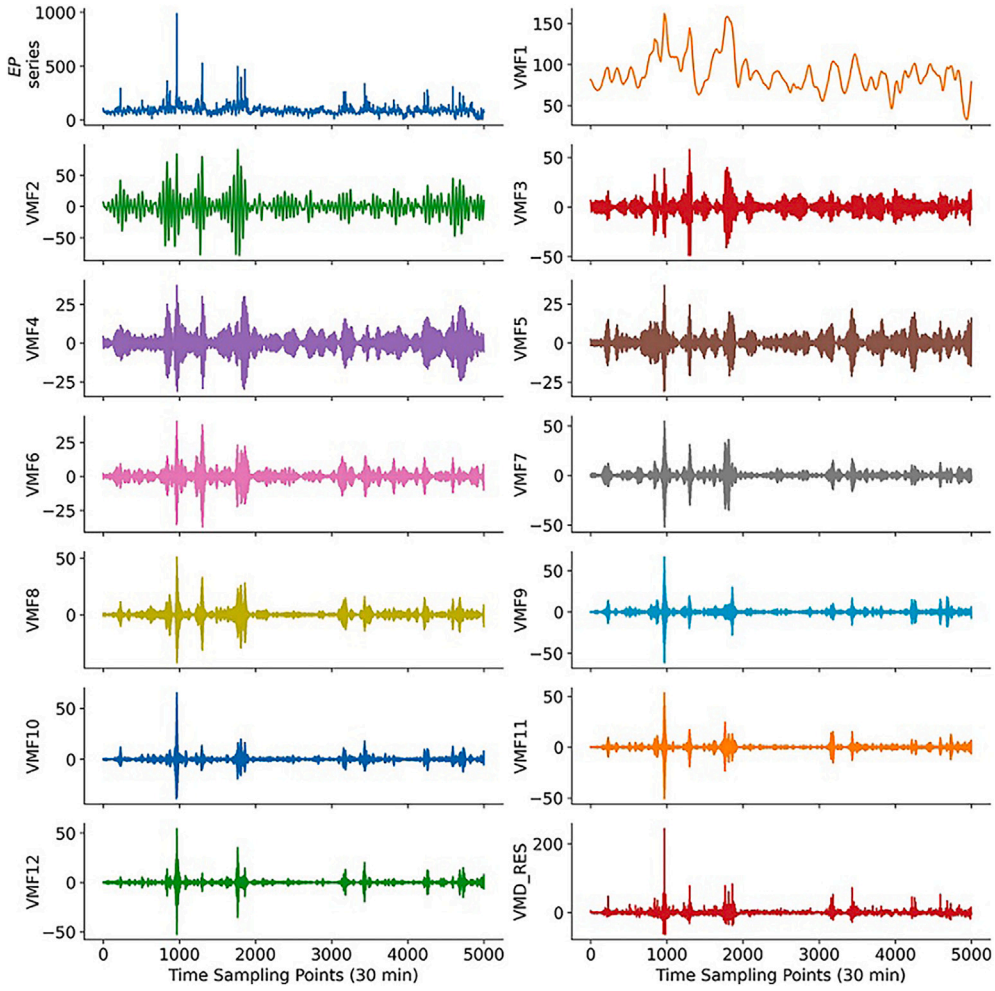


Fig. 8. Decomposition of electricity price (*EP*) series using the VMD algorithm.

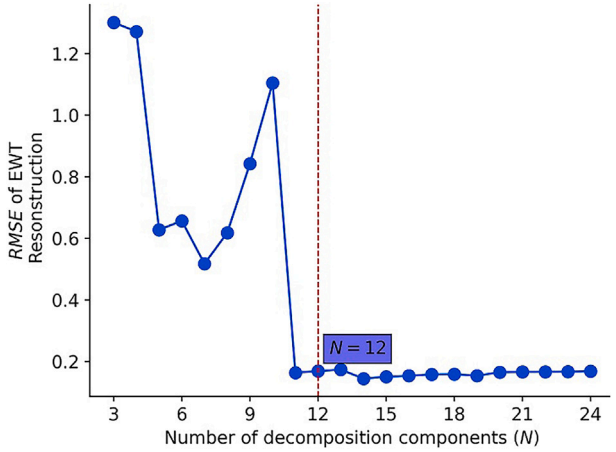


Fig. 9. Analysis of EWT decomposition modes and reconstruction $RMSE$ error.

3.3. Extraction of significant time-lagged inputs for predictive model development

Fig. 11 illustrates the results of Partial Autocorrelation Function (PACF) analysis for the normalized decomposed VMF and IMF series. The selection of input variables was based on an analysis of the PACF diagrams, as seen in the plots of PACFs with varying lag lengths (Fig. 11). In the iterative process of half-hourly EP prediction, the input variables consisted of 24 sub series (twelve VMF and twelve IMF), and the number of inputs for the prediction models was determined, as outlined in Table 2.

3.4. Predictive model development

Once the best predictors have been identified, the decomposition based prediction model (MRC-BiLSTM) is constructed for each sub-series (VMF and IMF). To achieve the optimal prediction results for different sub-series, the optimal hyper-parameters are selected using Bayesian Optimization [85]. In the proposed MRC-BiLSTM, parameters can be fine-tuned, and various types of layers can be adjusted to optimize the model. The model comprises three types of layers: residual convolutional blocks, BiLSTM layers, and dense layers. Parameters such as kernel size, the number of units in LSTM, and the number of filters in CNN can be adjusted in each layer, potentially affecting learning speed and performance. Table 3 provides details on the layer types, parameters, and kernel sizes for each layer. The lagged series (Table 2) of each decomposed component is then passed to the residual CNN. The residual CNN architecture is employed to extract spatial information from processed historical data. A Rectified Linear Unit (ReLU) activation function is applied in each convolutional layer and after adding a layer to the architecture. The output feature from the residual CNN is flattened and then forwarded to stacked BiLSTM layers to capture temporal information. Finally, fully connected layers are integrated to produce the predicted output values. A block diagram of the MRC-BiLSTM is

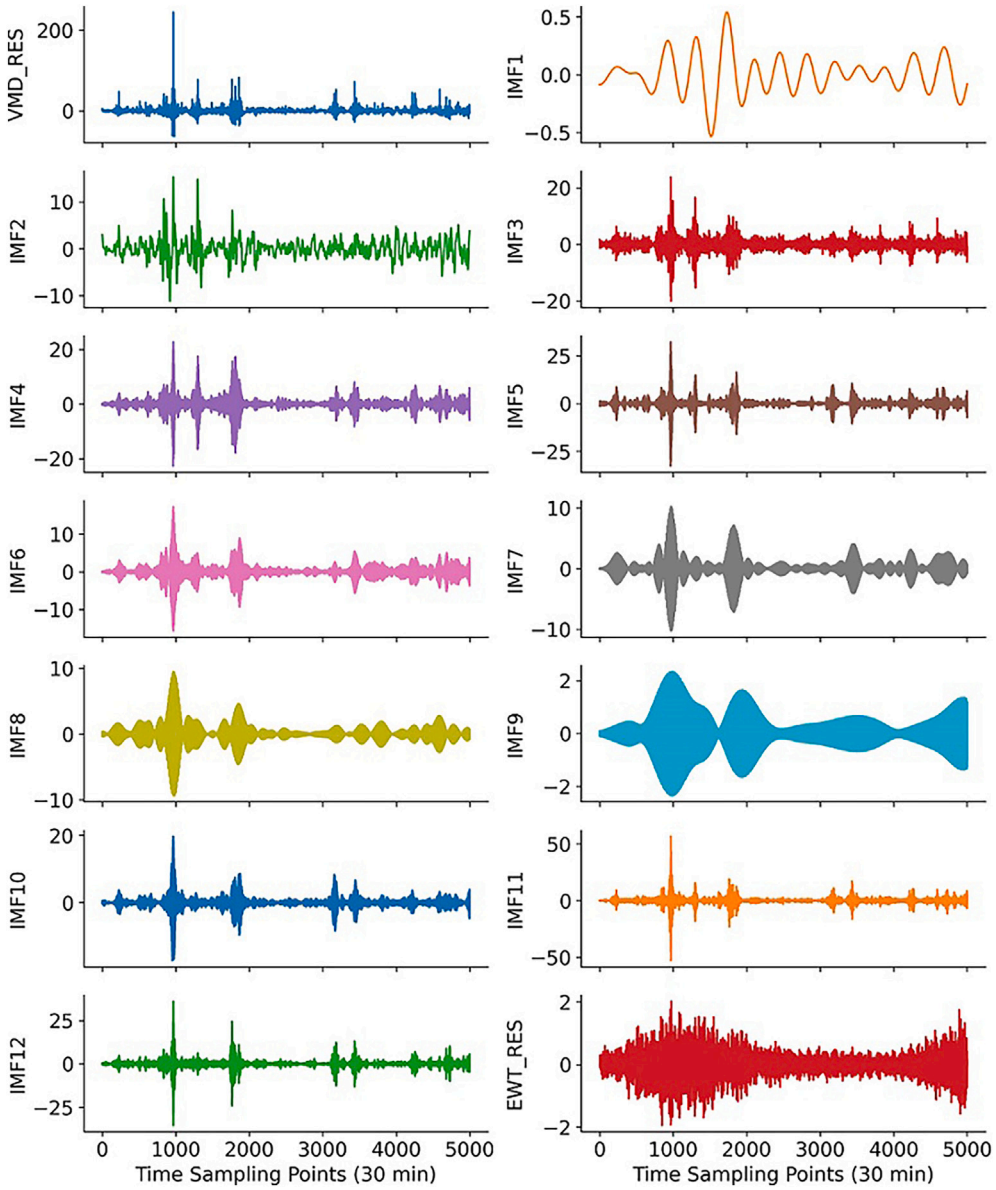


Fig. 10. Decomposition of VMD residual (VMD_{RES}) sequences into IMF components using the EWT algorithm.

presented in Fig. 12, illustrating the training and testing procedure of the model.

3.5. Hyperparameter tuning

As previously mentioned, Bayesian Optimization [85] is utilised to fine-tune the hyperparameters of the proposed model. Bayesian optimization (BO) operates based on Bayes' theorem, as described by Eq. (27). BO involves two crucial components when selecting the next hyperparameter configuration: a probabilistic surrogate model for modeling the objective function and an acquisition function that explores new areas in the sample space while exploiting known areas for optimal results [86]. This characteristic makes BO more efficient compared to grid search and random search.

$$p(w|D) = \frac{p(D|w)\,p(w)}{p(D)} \tag{27}$$

where w indicates an unseen value, $p(w)$ refers to the preceding distribution, $p(D|w)$ indicates the probability and $p(w|D)$ denotes the posterior distribution. The fundamental steps of BO are as follows [87]:

1. Build a probabilistic surrogate model of the objective function.
2. Determine the optimal hyperparameter values on the surrogate model.
3. Apply the chosen hyperparameter values to the real objective function for evaluation.
4. Update the surrogate model with the new results.
5. Repeat steps (2–4) until the maximum number of iterations is reached.

Following each evaluation of the objective function, BO updates the surrogate model. Common surrogate models for BO include Gaussian process (GP) [88], Random Forest (RF) [89], and the tree-structured

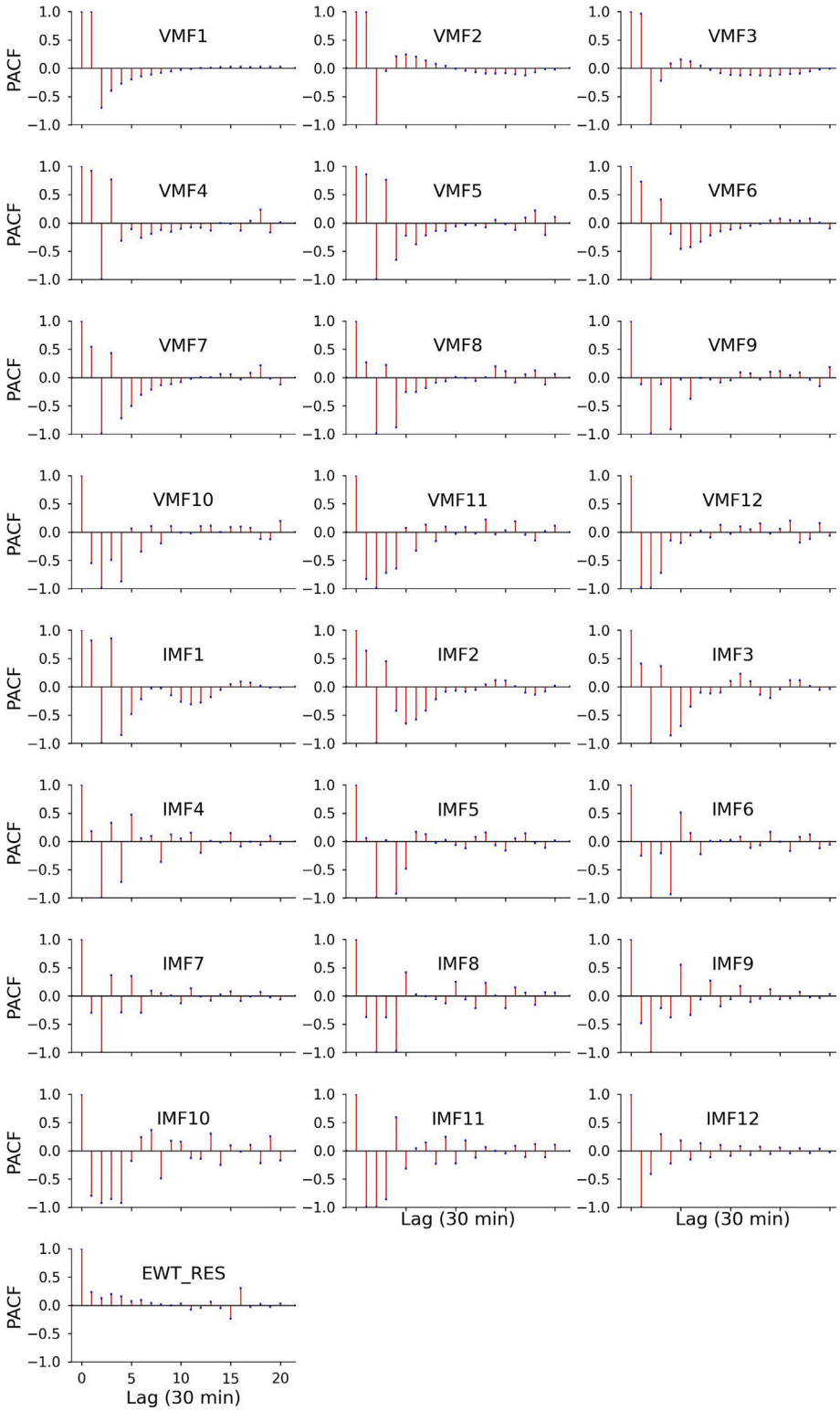


Fig. 11. Partial autocorrelation function (PACF) plot of the decomposed *EP* series using VMD (VMF) and EWT (IMF).

Parzen estimator (TPE) [90]. In this study, the TPE surrogate function is employed to model the objective function due to its lower time complexity [91]. In this study, we harnessed the Hyperopt [92] framework to apply the TPE algorithm for the automatic optimization of hyperparameters for all models. Python [93] was the chosen programming language, and TensorFlow 2.0 [94] served as the deep learning framework.

Table 2
The input features and output targets for the VMD and EWT decomposition sequences for predictive model development.

Decomposition algorithm	Feature term	Model input	Number of inputs	Output
VMD decomposition	VMF1	X(t-i), i = 0, 1, 2, 3, 4, 5, 6, 7, 8	9	X(t+1)
	VMF2	X(t-i), i = 0, 1, 2, 4, 5, 6, 7, 8	8	X(t+1)
	VMF3	X(t-i), i = 0, 1, 2, 3, 4, 5, 6, 7, 8	9	X(t+1)
	VMF4	X(t-i), i = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10	11	X(t+1)
	VMF5	X(t-i), i = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9	10	X(t+1)
	VMF6	X(t-i), i = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9	10	X(t+1)
	VMF7	X(t-i), i = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9	10	X(t+1)
	VMF8	X(t-i), i = 0, 1, 2, 3, 4, 5, 6, 7, 8	9	X(t+1)
	VMF9	X(t-i), i = 0, 1, 2, 3, 4, 5, 6	7	X(t+1)
	VMF10	X(t-i), i = 0, 1, 2, 3, 4, 6, 8	7	X(t+1)
	VMF11	X(t-i), i = 0, 1, 2, 3, 4, 6, 8	7	X(t+1)
	VMF12	X(t-i), i = 0, 1, 2, 3, 4, 5	6	X(t+1)
EWT decomposition	IMF1	X(t-i), i = 0, 1, 2, 3, 4, 5, 6	7	X(t+1)
	IMF2	X(t-i), i = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9	10	X(t+1)
	IMF3	X(t-i), i = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9	10	X(t+1)
	IMF4	X(t-i), i = 0, 1, 2, 3, 4, 5, 8	7	X(t+1)
	IMF5	X(t-i), i = 0, 1, 2, 4, 5, 6, 7	7	X(t+1)
	IMF6	X(t-i), i = 0, 1, 2, 3, 4, 5, 6, 7	8	X(t+1)
	IMF7	X(t-i), i = 0, 1, 2, 3, 4, 5, 6	7	X(t+1)
	IMF8	X(t-i), i = 0, 1, 2, 3, 4, 5	6	X(t+1)
	IMF9	X(t-i), i = 0, 1, 2, 3, 4, 5, 6, 8	8	X(t+1)
	IMF10	X(t-i), i = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10	11	X(t+1)
	IMF11	X(t-i), i = 0, 1, 2, 3, 4, 5, 7, 8, 9	9	X(t+1)
	IMF12	X(t-i), i = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9	10	X(t+1)
	EWT_RES	X(t-i), i = 0, 1, 2, 3, 4	5	X(t+1)

Table 3
Internal architecture of residual CNN model with stacked BiLSTM.

Layer Name	Parameters	Hyperparameter search space	Fixed parameter
Conv 1, 2, 3	Number of filters	[32, 64,128,256]	
	Kernel size		4, 3, 2
Add layer [Conv 2 and Conv 3]	Number of filters	[32, 64,128,256]	
Conv 4, 5, 6	Number of filters	[32, 64,128,256]	
	Kernel size		4, 3, 2
Add Layer [Conv 5 and Conv 6]	Number of filters	[32, 64,128,256]	
Conv 7, 8, 9	Number of filters	[32, 64,128,256]	
	Kernel size		4, 3, 2
Add layer [Conv 8 and Conv 9]	Number of filters	[32, 64,128,256]	
Flatten layer			
BiLSTM layer	Number of units	[20,015,010,050]	
BiLSTM layer	Number of units	[1,501,005,020]	
BiLSTM layer	Number of units	[100,755,020]	
Dense layer		[1,005,040]	
Dense layer		[1,005,040]	
Activation layer			ReLu
	Batch size	[300, 500, 1000, 1500, 2000]	
	Optimizer		Adam

3.6. Benchmark model development

To assess the effectiveness of the proposed VMD-EWT-MRC-BiLSTM model, it is compared with several benchmark and competitive models.

1. VMD-LSTM: The EP series are first decomposed using the VMD algorithm, followed by predictions made with the LSTM model. A three-layer stacked LSTM model is utilised, with each hidden layer having a Bayesian optimization search range of units between 50 and 100, and a batch size search range between 200 and 1000. The ReLU activation function is applied.
2. VMD-CLSTM: The EP series is initially decomposed using the VMD algorithm, and subsequently, predictions are generated using the deep hybrid CNN-LSTM model. This model consists of a two-layer CNN and a three-layer stacked LSTM. For the LSTM hidden layers, a Bayesian optimization search range for units is set between 50 and 100, with a batch size search range between 200 and 1000. The ReLU activation function is applied. Likewise, for the CNN, the search range for the number of filters is between 50 and 200.
3. VMD-ALSTM: The EP series undergo initial decomposition using the VMD algorithm, and subsequently, predictions are made using the Attention-based LSTM model. This model includes a three-layer stacked LSTM, with each hidden layer having a Bayesian optimization search range for units between 50 and 100, and a batch size search range between 200 and 1000. The ReLU activation function is applied.
4. VMD-GRU: The initial step involves decomposing the EP series using the VMD algorithm, after which predictions are generated using the GRU model. The GRU model employed in this study is a four-layer stacked configuration, with each hidden layer having its units determined via Bayesian optimization, within a range of 20–200 units. The batch size is also optimized within a range of 100–1000. Furthermore, the ReLU activation function is employed in this model.

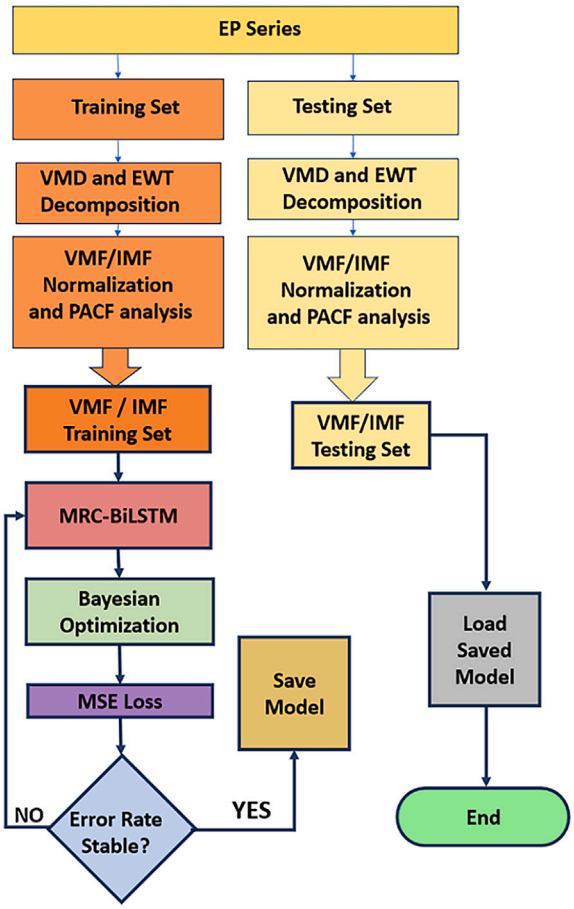


Fig. 12. Model Development phase of the deep hybrid VMD-EWT-MRC-BiLSTM model for half-hourly EP prediction.

5. Standalone models: In this study, the CLSTM, ALSTM, LSTM, and GRU models are also used for comparative analysis. It is important to emphasize that these standalone models utilise lagged EP values as input features, and Bayesian Optimization (BO) is employed to fine-tune their respective hyperparameters.

3.7. Model performance criteria

An extensive set of scoring metrics was applied to conduct a thorough evaluation of the newly developed deep hybrid VMD-EWT-MRC-BiLSTM model in comparison to benchmark decomposition-based models and standalone models. This study employs a diverse array of visual and statistical metrics during the independent testing phase. The effectiveness of a model hinges primarily on the relationship between projected and observed values. However, these criteria are often determined without due consideration of the model's goals and objectives. The statistical metrics are categorized into two classes: Class A, where the ideal value is 1, and Class B, where the ideal value is 0 [95–97]. The evaluation encompasses five Class A indicators Coefficient of Determination (R^2, Eq. 28), Willmott's Index (WI, Eq. 29), Nash–Sutcliffe Efficiency (NS, Eq. 30), Legates–McCabe's Index (LM, Eq. 31), and Kling–Gupta Efficiency (KGE, Eq. 32) and seven Class B indicators (Root Mean Square Error ($RMSE$, Eq. 33), Mean Absolute Error(MAE, Eq. 34), Normalized $RMSE$($nRMSE$, Eq. 35), Relative MAE ($RMAE$, Eq. 36), Symmetric Mean Absolute Percentage Error ($sMAPE$, Eq. 41), Theil's Inequality Coefficient (TIC, Eq. 38), and Absolute Percentage Bias (APB, Eq. 39)). The Class A indicators measured the discrepancies between the predicted and actual EP values, whereas the Class B error indicators evaluated the model's inherent bias. Since both bias and variance contribute to correctable errors, the models were assessed in terms of their capacity to simultaneously reduce bias and variance.

Coefficient of Determination (R^2):

$$R^2 = \left(\frac{\sum_{i=1}^{N} (EP^a - \langle EP^a \rangle)(EP^p - \langle EP^p \rangle)}{\sqrt{\sum_{i=1}^{n} (EP^a - \langle EP^a \rangle)^2} \sqrt{\sum_{i=1}^{n} (EP^p - \langle EP^p \rangle)^2}} \right)^2, \tag{28}$$

Wilmott's Index (I_{WI}):

$$WI = 1 - \frac{\sum_{i=1}^{N} (EP^a - EP^p)^2}{\sum_{i=1}^{N} (|EP^p - \langle EP^a \rangle| + |(EP^o - \langle EP^p \rangle|)^2}, \tag{29}$$

Nash–Sutcliffe Index (I_{NS}):

$$NS = 1 - \frac{\sum_{i=1}^{N} (EP^a - EP^p)^2}{\sum_{i=1}^{N} (EP^a - \langle EP^a \rangle)^2}, \tag{30}$$

Legates and McCabe Index (I_{LM}):

$$LM = 1 - \frac{\sum_{i=1}^{N} |EP^p - EP^a|}{\sum_{i=1}^{N} |EP^a - \langle EP^a \rangle|}, \tag{31}$$

Kling–Gupta Efficiency:

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\langle EP^p \rangle}{\langle EP^a \rangle} - 1 \right)^2 + \left(\frac{CV^p}{CV^a} \right)^2} \tag{32}$$

Root Mean Square Error ($RMSE$):

$$RMSE(AUD/MWh) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (EP^p - EP^a)^2}, \tag{33}$$

Mean Absolute Error (MAE):

$$MAE(AUD/MWh) = \frac{1}{N} \sum_{i=1}^{N} |EP^p - EP^a|, \tag{34}$$

Normalized Root Mean Square Error($RRMSE$):

$$nRMSE(\%) = \frac{RMSE}{EP^a} \times 100\ \% \tag{35}$$

Relative Mean Absolute($RMAE$):

$$RMAE(\%) = \frac{MAE}{EP^a} \times 100\ \% \tag{36}$$

Symmetric Mean Absolute Percentage Error ($sMAPE$):

$$sMAPE = \frac{1}{N} \sum_{i=1}^{N} \frac{|EP^a - EP^p|}{(|EP^a| + |EP^p|)/2}, \tag{37}$$

Theil's Inequality Coefficient:

$$TIC = \frac{\sqrt{\frac{1}{n} \times \sum_{i=1}^{n} (EP^p - EP^a)^2}}{\left(\sqrt{\frac{1}{n} \times \sum_{i=1}^{n} (EP^a)^2} + \sqrt{\frac{1}{n} \times \sum_{i=1}^{n} (EP^p)^2} \right)} \tag{38}$$

Absolute Percentage Bias (APB):

$$APB(\%) = \left| \frac{\sum_{i=1}^{n} (EP^a - EP^p)}{\sum_{i=1}^{n} EP^a} \right| \cdot 100, \quad (39)$$

Notably, EP^a and EP^p represent actual and predicted half-hourly EP while $\langle EP^a \rangle$ and $\langle EP^p \rangle$ represent actual and predicted mean EP, N = number of tested data points and CV is the Coefficient of Variation.

3.7.1. Global performance indicator

This study has incorporated the Global Performance Indicator (GPI) as a technique for ranking the performance of machine learning models [98]. The GPI integrates various metrics to produce a score that assesses the superior model. The GPI formula encompasses a constant α, which is set to -1 for Class A metrics and $+1$ for Class B metrics. The normalized value of each statistical indicator is denoted as n_{ik} and m_k representing the median of the normalized statistical indicator j across all models (where k ranges from 1 to 12). Higher GPI values signify superior model performance

$$GPI = \sum_{k=1}^{12} \alpha_k \left(m_k - n_{ik} \right), \quad (-\infty < GPI < +\infty) \quad (40)$$

3.7.2. Skill score

In addition to comparing the models, this study used Skill Score metrics to assess them against a persistence-based approach. The Skill Score metric ($RMSE_{ss}$) is defined in Eq. (41), where 'M' represents the error ($RMSE$) resulting from the predictions made by the models being studied, and 'P' corresponds to the $RMSE$ of the persistence model. The persistence model assumes that the EP prediction for the next time step depends on the EP at the current time step.

$$RMSE_{SS} = 1 - \frac{RMSE_M}{RMSE_P} \quad (41)$$

3.7.3. Directional statistic

The directional statistic (D_{stat}) serves to quantify the likelihood that the predicted series and the actual series exhibit changes in the same direction. This index is extensively employed in the evaluation of prediction models for price time series, owing to its ability to effectively assess the models' capacity to predict both upward and downward movements. In general, a higher D_{stat} value indicates superior prediction performance. The equation representing this metric is provided below:

$$D_{stat} = \frac{1}{n} \sum_{t=2}^{n} d_t \times 100\ \% \quad (42)$$

where:

$$d_t = \begin{cases} 1 & \text{if } (EP_t^a - EP_{t-1}^a)(EP_t^p - EP_{t-1}^a) > 0 \\ 0 & \text{otherwise} \end{cases} \quad (43)$$

3.7.4. The giacomini–white test

The Giacomini-White (GW) test [99] can be viewed as an extension of the Diebold-Mariano test [100], assessing the Conditional Predictive Ability (CPA) instead of Unconditional predictive ability [13]. Similar to the DM test, the GW test gauges the statistical significance of disparities between predictions produced by two different models. This study utilises the GW tests to assess the significance of differences in prediction accuracy. The mathematical formulation of the GW tests can be found in Ref. [101].

3.8. Model implementation environment

The computational setup used for this research is outlined in Table 4. It is important to note that training times may differ based on the specific computational environment employed. Table 5 provides information regarding the time spent for model construction, including training and validation, as well as the time required for processing new inputs during testing. According to the data presented in Table 4, the average time for building the VMD-EWT-MRC-BiLSTM model was 265 minutes, which is relatively lengthy. However, once the model is established, it can be utilised over an extended period, considering the testing dataset spans 13,902 data points. Moreover, the processing time for a testing is under 30 s, ensuring the practical feasibility of applying our model.

Table 4
Computational environment.

Hardware/software	Specifications
CPU	Intel i7-9700k 3.80 GHz
Memory	64 GB
Compile	Python 3.12
DL framework	Tensorflow 2.0
Hyperopt	0.2.7
Keras	2.14.0

Table 5
Average of computation time.

Model	Construction time (Training and validation)	Prediction time (Test data)
VMD-CLSTM	185 min	25 s
VMD-LSTM	178 min	29 s
VMD-GRU	162 min	28 s
VMD-ALSTM	157 min	28 s
CLSTM	80 min	26 s
ALSTM	76 min	32 s
LSTM	58 min	31 s
GRU	63 min	32 s
Proposed model (VMD-EWT-MRC-BiLSTM)	265 min	29 s

4. Results and discussion

The evaluation of the proposed VMD-EWT-MRC-BiLSTM model and benchmark with other decomposition-based and standalone models were carried out through a comprehensive set of performance criteria. Generally, the extensive experimental results demonstrate that the proposed VMD-EWT-MRC-BiLSTM model outperforms standalone models, including GRU, ALSTM, CLSTM, and LSTM, across various predictive metrics. Even their hybrid models, which also use VMD but without EWT and MRC components, including VMD-LSTM, VMD-CLSTM, VMD-ALSTM and VMD-GRU, also performed well but did not reach the same level of accuracy as the VMD-EWT-MRC-BiLSTM model.

Table 6 shows the evaluation results of the proposed model against other comparative models during the testing phase, highlighting the significant improvements in predictive performance across all Class A metrics. The proposed VMD-EWT-MRC-BiLSTM model achieved an R^2

Table 6
Evaluation of proposed (i.e. VMD-EWT-MRC-BiLSTM) vs. all other comparative models in the testing phase, using the *Class A* Metrics (Coefficient of Determination (R^2, Eq. 28), Willmott's Index (I_{WI}, Eq. 29), Nash–Sutcliffe Efficiency (I_{NS}, Eq. 30), Legates-McCabe's Index (I_{LM}, Eq. 31), and Kling–Gupta Efficiency (KGE, Eq. 32).

Predictive models	R^2	WI	NS	LM	KGE
VMD-EWT-MRC-BiLSTM	**0.999**	**0.995**	**0.994**	**0.938**	**0.988**
VMD-LSTM	0.972	0.953	0.972	0.912	0.958
VMD-CLSTM	0.972	0.953	0.971	0.909	0.953
VMD-ALSTM	0.969	0.946	0.948	0.878	0.886
VMD-GRU	0.953	0.903	0.891	0.758	0.873
GRU	0.923	0.828	0.790	0.654	0.862
ALSTM	0.922	0.827	0.788	0.654	0.877
CLSTM	0.922	0.828	0.786	0.656	0.894
LSTM	0.948	0.848	0.806	0.673	0.896

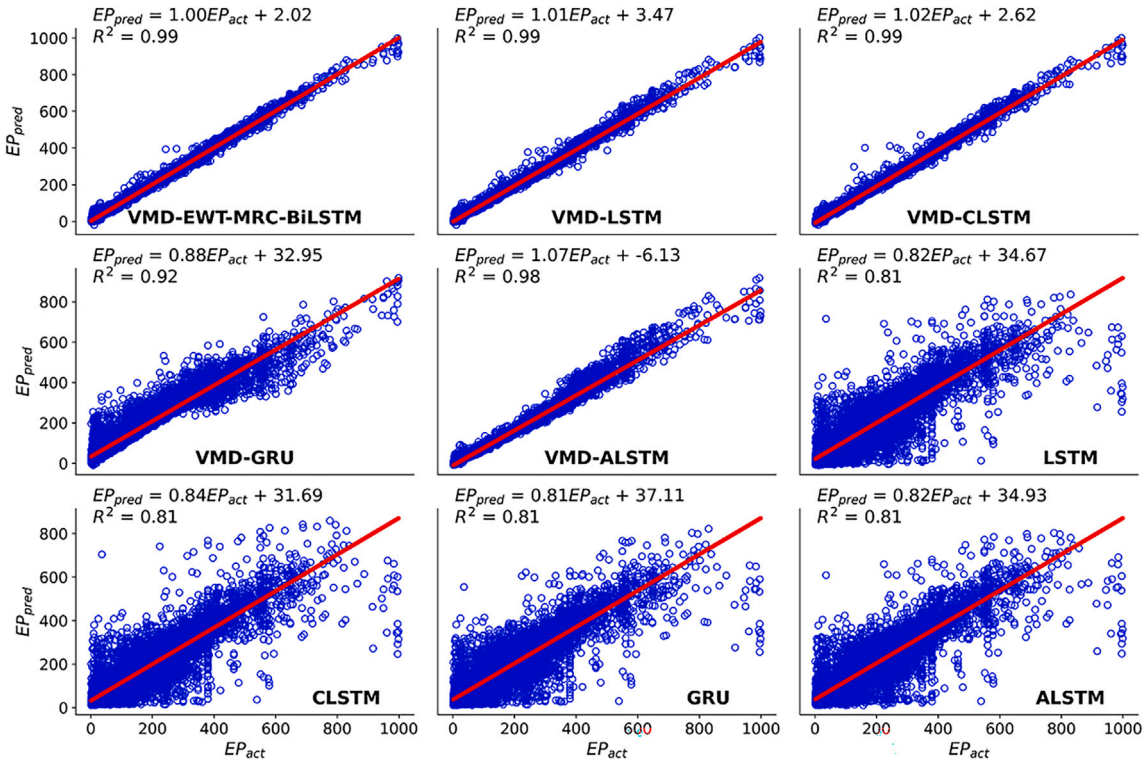


Fig. 13. Scatter plot of predicted and actual EP in testing phase using the proposed VMD-EWT-MRC-BiLSTM vs. other models for half-hourly EP prediction. Least square regression line and coefficient of determination (R^2) with a linear fit is shown in each sub-panel.

value of ≈ 0.999, suggesting that it can explain nearly all variability in the target EP variable, outperforming all other models in capturing the underlying patterns in the data. This is a substantial improvement compared to the next best models, VMD-LSTM and VMD-CLSTM, both of which recorded R^2 values of ≈ 0.972. The comparative analysis can be also visualised through Figs. 13 and 14 illustrating how well these models fit the predicted and actual EP. In terms of Willmott's Index, the VMD-EWT-MRC-BiLSTM model attained the highest score of ≈ 0.995. The comparative models, VMD-LSTM and VMD-CLSTM, both scored ≈ 0.953, showing that while they are effective, they are less precise than the proposed model. Similarly, the proposed model recorded the highest values of I_{NS} (≈ 0.994), I_{LM} (≈ 0.938), and KGE (≈ 0.988), significantly higher than the next best models, VMD-LSTM (I_{NS} ≈ 0.972, I_{LM} ≈ 0.912, and KGE ≈ 0.958 and VMD-CLSTM (I_{NS} ≈ 0.972, I_{LM} ≈ 0.909, and KGE ≈ 0.958). Standalone models, such as GRU, ALSTM, CLSTM and LSTM, have R^2 values ranging from ≈ 0.922 to ≈ 0.948 and other metric scores correspondingly lower.

The proposed model (VMD-EWT-MRC-BiLSTM) also demonstrated superior performance across all Class B error metrics, including $RMSE$, MAE, $nRMSE$, and $RMAE$, as shown in Table 7. These metrics are essential to understand the accuracy and reliability of the models in predicting EP. With an $RMSE$ ≈ 10.555 AUD/MWh and an MAE ≈ 6.707 AUD/MWh, the VMD-EWT-MRC-BiLSTM model showed lower error rates than the other models evaluated, confirming its precision in EP prediction. The decomposition-based VMD-LSTM model is the second-best model with an $RMSE$ ≈ 12.680 AUD/MWh and an MAE ≈ 7.396 AUD/MWh. Additionally, the lower $nRMSE$ and $RMAE$ of the proposed model indicate that its errors are smaller relative to the overall variation in the data and its predictions are closer to the actual prices, respectively. The results in Table 8 also provide an evaluation of the proposed VMD-EWT-MRC-BiLSTM model against comparative models, using Class B metrics. The lower $sMAPE$, TIC, and APB values indicate the capacity of the proposed model in capturing complex patterns in the data effectively, reducing both error and bias. These findings further highlight the consistently superior predictive performance of the proposed model, making it more reliable for practical applications.

The performance of the model for the prediction of electricity prices a half hour in advance across different time intervals was also notable. Table 9 presents the results of the month-wise prediction for the EP prediction in the half hour ahead. Overall, the proposed VMD-EWT-MRC-BiLSTM model consistently achieved higher $GPIs$ across all months compared to other models. Particularly, the GPI ≈ 2.38, 3.12, 2.55 and 2.24 in March, April, Nov and December, respectively, are markedly higher than those of the two second-best models, VMD-ALSTM (GPI ≈ 0.97, 1.72, 0.87 and 0.96) and VMD-CLSTM (GPI ≈ 1.00, 1.71, 0.86 and 0.96), reflecting exceptional predictive accuracy during these periods. The negative $GPIs$ of the VMD-GRU and other standalone models indicate their inadequacy in accurately predicting EP. Similarly, across all weekdays as described in Table 10, the VMD-EWT-MRC-BiLSTM model consistently outperforms the comparative models, achieving the lowest $RMSE$, $nRMSE$, $RMAE$ and $MAPE$ values, as well as the highest GPI scores. For example, on Sunday analysis, the proposed model achieved the lowest $RMSE$ ≈ 8.89, $nRMSE$ ≈ 5.62 %, $RMAE$ ≈ 10.09 %, $MAPE$ ≈ 11.01 % and the highest GPI ≈ 2.41. In contrast, the LSTM and ALSTM models perform poorly, with significantly higher values of $RMSE$, $nRMSE$, $RMAE$, $MAPE$ and lower values of GPI. Specifically, LSTM attained $RMSE$ ≈ 53.98, $nRMSE$ ≈ 33.46 %, $RMAE$ ≈ 29.71 %, $MAPE$ ≈ 51.50 % and GPI ≈ −6.67 while ALSTM scored $RMSE$ ≈ 54.58, $nRMSE$ ≈ 34 %, $RMAE$ ≈ 32 %, $MAPE$ ≈ 55 % and GPI ≈ −7. Fig. 14 also depicts the predicted residual errors for half-hourly EP prediction at 1-day period among the tested models, indicating superior performance of the proposed VMD-EWT-MRC-BiLSTM model.

Moreover, the seasonal analysis shown in Table 11 reveals that the VMD-EWT-MRC-BiLSTM model stands out in all seasons. For example, in spring, the proposed model achieved $nRMSE$ ≈ 4.87 %, $RMAE$

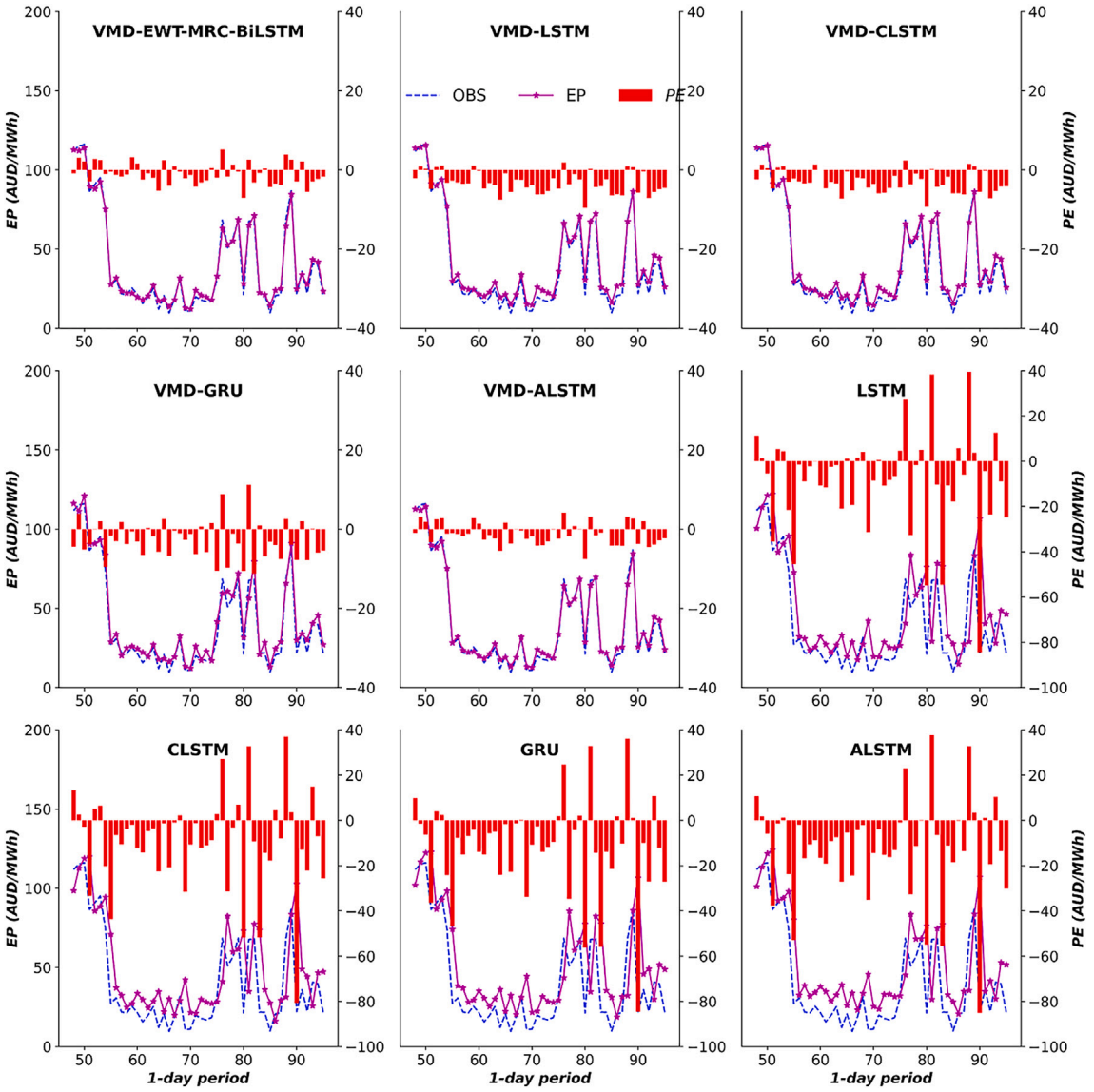


Fig. 14. The time series plot of predicted and actual EP for the proposed VMD-EWT-MRC-BiLSTM model and other comparative models for half-hourly EP prediction. Note: Errors (residuals) of each prediction is plotted on secondary axis.

Table 7
Evaluation of proposed (i.e. VMD-EWT-MRC-BiLSTM) vs. all other comparative models in the testing phase, using the *Class B* metrics (Root Mean Square Error (*RMSE*, Eq. 33;AUD/MWh), Mean Absolute Error(*MAE*, Eq. 34; AUD/MWh), Normalized *RMSE*(*nRMSE*, Eq. 35;%), and Relative *MAE* (*RMAE*, Eq. 36;%).

Predictive models	*RMSE*	*MAE*	*nRMSE*	*RMAE*
VMD-EWT-MRC-BiLSTM	**10.555**	**6.707**	**5.971**	**11.115**
VMD-LSTM	12.680	7.396	7.173	11.986
VMD-CLSTM	13.328	7.707	7.540	11.681
VMD-ALSTM	24.790	11.095	14.024	12.449
VMD-GRU	41.706	24.153	23.594	43.386
GRU	60.774	35.462	34.381	59.031
ALSTM	61.089	35.396	34.559	58.007
CLSTM	61.408	35.227	34.739	57.947
LSTM	61.424	35.571	34.748	56.357

Table 8
Evaluation of proposed (i.e. VMD-EWT-MRC-BiLSTM) vs. all other comparative models in the testing phase, using the *Class B* metrics (Symmetric Mean Absolute Percentage Error (*sMAPE*, Eq. 41;%), Theil's Inequality Coefficient (*TIC*, Eq. 38), and Absolute Percentage Bias (*APB*, Eq. 39;%).

Predictive models	*APB*	*SMAPE*	*TIC*
VMD-EWT-MRC-BiLSTM	**3.794**	**0.080**	**0.023**
VMD-LSTM	4.584	0.089	0.028
VMD-CLSTM	4.760	0.088	0.029
VMD-ALSTM	6.677	0.094	0.054
VMD-GRU	14.064	0.202	0.092
GRU	20.461	0.293	0.137
ALST	20.424	0.297	0.137
CLSTM	20.328	0.289	0.137
LSTM	20.123	0.285	0.138

$\approx$ 6.74 %, $MAPE \approx$ 12.09 % and $GPI \approx$ 3.22, outperforming other models by a significant margin. In general, the models VMD-LSTM and VMD-CLSTM frequently appear as the second-best performers. However, if we consider the overall presence across all seasons, VMD-LSTM appears most consistently in second place for multiple metrics. In contrast, ALSTM model consistently has the worst performance, followed by

Table 9
Month-wise prediction results for half-an-hour ahead electricity price forecasting (AUD/MWh). Models highlighted in bold demonstrate superior performance as assessed by the Global Performance Indicator (*GPI*) metrics.

	Models								
Month	**VMD-EWT-MRC-BiLSTM**	VMD-LSTM	VMD-CLSTM	VMD-GRU	VMD-ALSTM	LSTM	CLSTM	GRU	ALSTM
Jan	**0.93**	0.85	0.86	−0.83	0.87	−8.20	−8.23	−8.45	−8.73
Feb	**0.96**	0.92	0.93	−0.89	0.92	−7.92	−8.11	−8.22	−8.39
Mar	**2.38**	0.94	1.00	−0.93	0.97	−8.18	−8.34	−8.35	−8.58
Apr	**3.12**	1.68	1.71	−1.68	1.72	−7.78	−7.76	−7.73	−7.68
May	**2.78**	2.72	2.71	−2.55	2.53	−6.58	−6.56	−6.46	−6.55
Jun	**3.91**	3.64	3.67	−2.74	2.46	−5.20	−4.80	−5.29	−4.89
Jul	**5.40**	4.80	4.62	−3.60	1.56	−2.45	−1.88	−2.34	−1.83
Aug	**2.26**	2.29	2.30	−2.19	2.06	−6.96	−6.94	−7.02	−6.99
Sep	**2.05**	1.97	1.98	−1.96	2.04	−7.24	−7.22	−7.28	−7.46
Oct	**2.24**	1.49	1.38	−1.35	1.58	−7.77	−7.65	−7.66	−7.74
Nov	**2.55**	0.85	0.86	−0.83	0.87	−8.20	−8.23	−8.45	−8.73
Dec	**2.24**	0.92	0.96	−0.89	0.96	−7.92	−8.11	−8.22	−8.39

Table 10
Prediction results for half-an-hour ahead electricity price forecasting (AUD/MWh) bassed on weekdays.Models highlighted in bold demonstrate superior performance as assessed by the Root Mean Square Error (*RMSE*), Normalized *RMSE*(*nRMSE*), Relative Mean Absolute Error (*RMAE*),and Global Performance Indicator (*GPI*) metrics.

Weekday	Metrics	VMD-EWT-MRC-BiLSTM	VMD-LSTM	VMD-CLSTM	VMD-GRU	VMD-ALSTM	LSTM	CLSTM	GRU	ALSTM
Sunday	*RMSE*	**8.89**	9.79	11.96	33.08	11.14	53.98	53.96	53.75	54.58
	nRMSE	**5.62 %**	6.16 %	7.52 %	19.16 %	6.99 %	33.46 %	33.40 %	33.18 %	34 %
	RMAE	**10.09 %**	11.48 %	11.54 %	16.08 %	11.57 %	29.71 %	28.58 %	28.69 %	32 %
	MAPE	**11.01 %**	12.11 %	11.97 %	46.06 %	11.54 %	51.50 %	52.61 %	54.30 %	55 %
	GPI	**2.41**	2.26	2.11	−2.07	2.22	−6.67	−6.53	−6.79	−7.00
Monday	*RMSE*	**10.81**	12.13	12.88	41.89	20.63	67.47	67.47	67.02	66.74
	nRMSE	**5.91 %**	6.59 %	7.01 %	21.27 %	11.06 %	36.89 %	36.60 %	36.62 %	36.35 %
	RMAE	**7.05 %**	7.28 %	7.33 %	16.41 %	7.75 %	34.63 %	33.26 %	33.14 %	34.70 %
	MAPE	**11.03 %**	11.91 %	11.54 %	47.51 %	11.72 %	54.92 %	56.25 %	56.97 %	53.99 %
	GPI	**2.87**	2.75	2.71	−1.92	1.92	−6.61	−6.30	−6.69	−6.36
Tuesday	*RMSE*	**11.44**	13.69	14.71	45.32	29.30	64.80	64.38	63.57	63.73
	nRMSE	**6.30 %**	7.49 %	8.04 %	23.47 %	15.69 %	35.65 %	35.23 %	34.95 %	34.91 %
	RMAE	**8.31 %**	8.58 %	8.32 %	19.91 %	9.12 %	30.89 %	30.20 %	30.16 %	34.03 %
	MAPE	**13.77 %**	14.35 %	14.04 %	58.87 %	15.03 %	72.48 %	74.23 %	75.29 %	74.49 %
	GPI	**3.74**	3.48	3.39	−1.83	1.83	−5.54	−5.24	−5.52	−5.43
Wednesday	*RMSE*	**11.47**	14.61	14.44	45.79	27.27	62.52	62.55	61.87	62.36
	nRMSE	**6.11 %**	7.72 %	7.63 %	23.42 %	14.06 %	33.45 %	33.28 %	33.10 %	33.28 %
	RMAE	**7.67 %**	8.90 %	7.90 %	16.45 %	9.15 %	31.67 %	30.89 %	30.63 %	32.99 %
	MAPE	**11.09 %**	12.07 %	11.73 %	40.38 %	13.05 %	51.91 %	54.08 %	54.75 %	53.95 %
	GPI	**3.71**	3.34	3.34	−1.88	1.85	−5.51	−5.25	−5.58	−5.47
Thursday	*RMSE*	**11.19**	15.69	15.30	50.37	35.73	66.33	66.24	65.59	65.66
	nRMSE	**5.66 %**	7.87 %	7.67 %	24.99 %	17.31 %	33.91 %	33.59 %	33.51 %	33.41 %
	RMAE	**6.41 %**	6.81 %	6.85 %	16.16 %	7.87 %	29.45 %	28.19 %	28.16 %	30.09 %
	MAPE	**9.58 %**	10.46 %	10.12 %	36.42 %	11.67 %	54.35 %	56.01 %	57.17 %	56.19 %
	GPI	**4.38**	3.91	3.89	−1.85	1.61	−4.96	−4.62	−5.02	−4.89
Friday	*RMSE*	**9.62**	10.13	11.32	35.02	22.34	55.16	54.91	54.53	55.07
	nRMSE	**5.32 %**	5.58 %	6.22 %	18.57 %	12.00 %	30.40 %	30.17 %	30.02 %	30.26 %
	RMAE	**5.99 %**	6.22 %	6.29 %	13.38 %	7.16 %	28.98 %	28.01 %	27.99 %	29.57 %
	MAPE	**9.36 %**	10.07 %	9.72 %	31.41 %	11.19 %	46.96 %	48.17 %	49.49 %	48.91 %
	GPI	**3.14**	3.05	2.87	−1.33	1.33	−6.15	−5.89	−6.25	−6.28
Saturday	*RMSE*	**9.98**	11.12	11.90	36.34	17.64	57.16	57.85	56.62	57.28
	nRMSE	**6.24 %**	6.92 %	7.40 %	21.17 %	10.81 %	35.21 %	35.56 %	34.79 %	35.05 %
	RMAE	**8.02 %**	8.60 %	8.58 %	15.73 %	8.56 %	31.63 %	30.79 %	30.65 %	33.30 %
	MAPE	**11.99 %**	12.99 %	12.71 %	43.25 %	12.87 %	62.20 %	64.03 %	65.09 %	63.89 %
	GPI	**2.76**	2.65	2.56	−1.78	1.78	−6.45	−6.42	−6.57	−6.61

VMD-GRU. Additionally, Fig. 15 shows an evaluation of the proposed VMD-EWT-MRC-BiLSTM model versus other comparative models for half-hourly EP prediction, using GPI. The plot effectively demonstrates the superior predictive capability of the proposed model in predicting half-hourly EP.

Fig. 16 provides box plots illustrating the Prediction Error ($|PE|$) generated by the VMD-EWT-MRC-BiLSTM and benchmarked models in the testing phase. Overall, the proposed model presents a smaller spread, lower median, and fewer outliers compared to the other models. These demonstrates more accurate predictions with minimal prediction errors and less variation, and better handling of extreme values of the VMD-EWT-MRC-BiLSTM model. The VMD-LSTM and VMD-CLSTM models also perform well but have a slightly larger spread and more outliers compared to VMD-EWT-MRC-BiLSTM. The VMD-GRU and VMD-ALSTM models show a larger spread in the ($|PE|$) values and a higher number of outliers, indicating less reliable performance. Finally, traditional models such as LSTM, GRU, and their variants (ALSTM, CLSTM) exhibit the largest spread and the highest number of outliers, suggesting they

Table 11
Prediction results for half-an-hour ahead electricity price forecasting (AUD/MWh) based on season. Models highlighted in bold demonstrate superior performance as assessed by the Root Mean Square Error (*RMSE*), Normalized *RMSE*(*nRMSE*), Relative Mean Absolute Error (*RMAE*), and Global Performance Indicator (*GPI*) metrics.

Season	Metrics	VMD-EWT-MRC-BiLSTM	VMD-LSTM	VMD-CLSTM	VMD-GRU	VMD-ALSTM	LSTM	CLSTM	GRU	ALSTM
Summer	*nRMSE*	**5.91 %**	6.86 %	6.94 %	18.86 %	6.82 %	47.24 %	47.43 %	45.71 %	46.02 %
	RMAE	**7.97 %**	8.88 %	8.61 %	17.49 %	8.68 %	33.43 %	32.27 %	32.45 %	33.50 %
	MAPE	**9.02 %**	11.86 %	11.18 %	30.43 %	11.31 %	55.18 %	58.27 %	61.90 %	64.11 %
	GPI	**2.66**	1.17	1.19	−1.17	1.20	−7.81	−7.89	−8.03	−8.31
Autumn	*nRMSE*	**4.22 %**	4.82 %	4.94 %	16.36 %	5.40 %	31.57 %	31.95 %	31.11 %	31.58 %
	RMAE	**5.53 %**	6.71 %	6.18 %	13.52 %	6.59 %	29.09 %	28.05 %	27.70 %	28.94 %
	MAPE	**7.71 %**	10.23 %	9.83 %	35.90 %	9.94 %	48.04 %	49.83 %	50.65 %	49.76 %
	GPI	**3.41**	1.97	2.01	−1.86	1.92	−7.35	−7.39	−7.37	−7.51
Winter	*nRMSE*	**5.43 %**	6.99 %	7.40 %	22.55 %	14.72 %	28.92 %	28.27 %	28.81 %	28.40 %
	RMAE	**6.88 %**	8.82 %	7.66 %	18.40 %	9.89 %	28.61 %	27.89 %	28.03 %	31.32 %
	MAPE	**10.02 %**	12.07 %	12.32 %	58.08 %	15.45 %	55.15 %	55.41 %	55.20 %	52.51 %
	GPI	**6.48**	4.44	4.35	−2.68	2.15	−4.15	−3.68	−4.13	−3.74
Spring	*nRMSE*	**4.87 %**	6.03 %	6.14 %	17.71 %	5.65 %	37.11 %	37.44 %	36.55 %	37.53 %
	RMAE	**6.74 %**	8.02 %	9.69 %	15.94 %	7.97 %	33.81 %	32.67 %	32.39 %	36.77 %
	MAPE	**12.09 %**	14.22 %	13.74 %	48.51 %	13.46 %	69.50 %	70.85 %	71.13 %	68.52 %
	GPI	**3.22**	1.67	1.63	−1.61	1.74	−7.54	−7.53	−7.53	−7.70

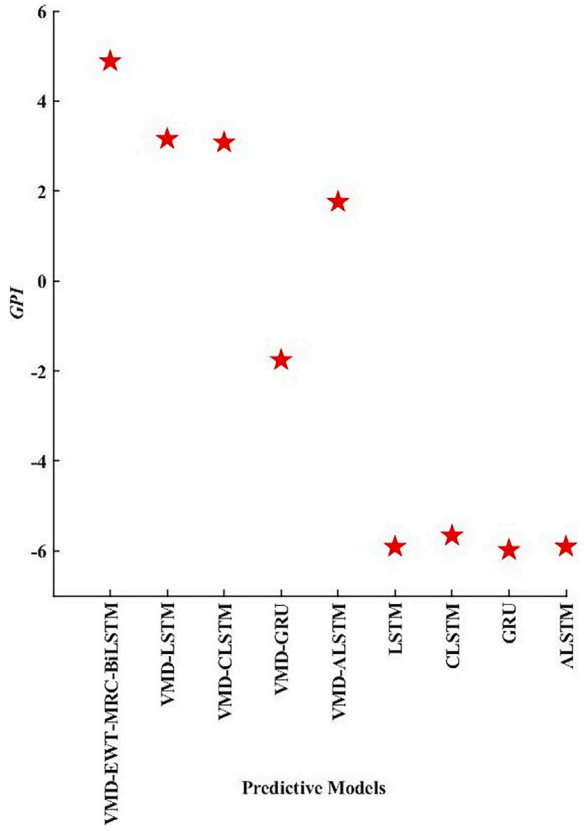


Fig. 15. Efficacy of proposed VMD-EWT-MRC-BiLSTM model vs. other comparative models for half-hourly EP prediction in terms of the Global Performance Indicator (GPI).

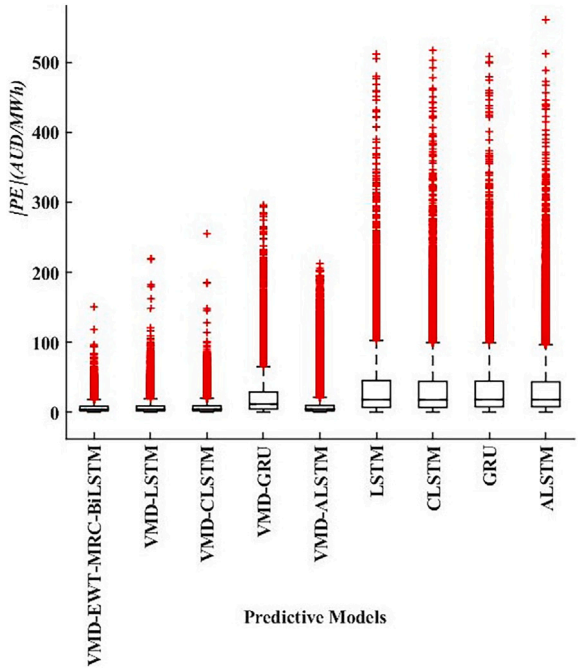


Fig. 16. Box plots of the Prediction Error ($|PE|$) generated by VMD-EWT-MRC-BiLSTM model vs other benchmarked models in the testing phase.

are the least accurate and most variable in their predictions. Fig. 17 provides further assessment of the ($|PE|$) distributions among models. It can be seen that the majority of errors ($\approx$ 48 %) generated by the VMD-EWT-MRC-BiLSTM lies in the first bin (0-5) compared to the other models.

Fig. 18 compares the performance of all models using two metrics: Skill Score ($RMSE_{SS}$) and Directional Symmetry (DS). The VMD-EWT-MRC-BiLSTM model achieved the highest $RMSE_{SS} \approx 97$ %, indicating it significantly reduces the $RMSE$ compared to the baseline model. VMD-LSTM, VMD-CLSTM, and VDM-ALSTM also performed well with $RMSE_{SS} \approx 92$ %, ≈ 91 %, and ≈ 85 %, respectively, showing substantial error reductions. VMD-GRU recorded $RMSE_{SS} \approx 59$ %, indicating moderate performance. LSTM, CLSTM, GRU, and ALSTM have much lower $RMSE_{SS}$, ranging from ≈ 11 % to ≈ 13 %, indicating they are less effective at reducing $RMSE$ compared to the hybrid and decomposed models. Similarly, the VMD-EWT-MRC-BiLSTM model also stand out in DS metric, achieving ≈ 96 %, which implies that it correctly predicts the direction of price changes ≈ 96 % of the time. VMD-LSTM follows with $DS \approx 87$ %, and VMD-CLSTM with $DS \approx 77$ %, suggesting these models also have good directional prediction capability. On the other hand, VMD-GRU and VMD-ALSTM show $DS \approx 73$ % and ≈ 70 %, respectively, showing reasonable performance. However, LSTM ($DS \approx 62$ %), CLSTM ($DS \approx 54$ %), GRU ($DS \approx 53$ %), and ALSTM ($DS \approx 53$ %) perform

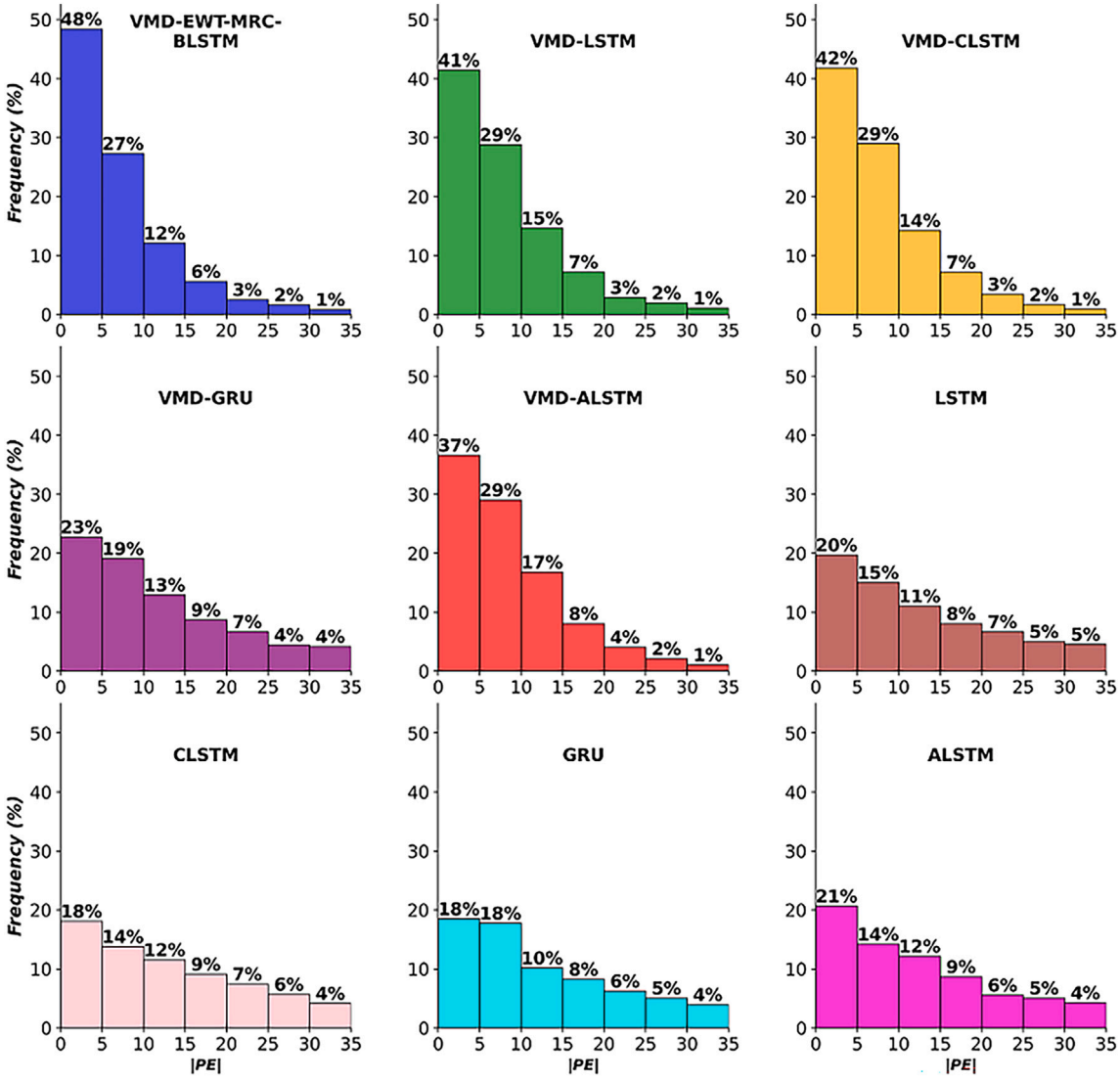


Fig. 17. Histogram illustrating the frequency (in percentages) of absolute Prediction Errors ($|PE|$) of the best performing VMD-EWT-MRC-BiLSTM model during test period compared with benchmarked models for the prediction of half-hourly EP.

significantly worse, indicating they are less reliable in predicting the direction of price changes.

The predictive ability among all models was also evaluated using the unconditional Giacomini–White test, illustrated in Fig. 19. The $p-values$ close to 0 (green colour) against all benchmark models suggest that the VMD-EWT-MRC-BiLSTM model is significantly superior in predictive accuracy. The absence of significant red areas in the first row indicates that no benchmark model significantly outperforms the VMD-EWT-MRC-BiLSTM model. VMD-LSTM displays a small red area against VMD-CLSTM, indicating a case where VMD-CLSTM is slightly more accurate. Furthermore, the hybrid models including VMD-LSTM, VMD-GRU, VMD-CLSTM and VMD-ALSTM perform well against traditional models (LSTM, CLSTM, GRU, ALSTM) with green areas.

According to numerous published scholars [102–104], hybrid models that integrate decomposition methods with advanced neural networks tend to outperform standalone models due to their ability to capture complex patterns and temporal dependencies in data. In this study, the enhanced performance reflects the contribution of the effective decomposition of input signals by VMD and EWT, which separate various underlying mode functions and relevant features. Specifically, VMD is effective in decomposing non-stationary signals into intrinsic mode functions, aiding in more accurate feature extraction for prediction models [62]. At the same time, EWT is beneficial for analysing localised variations in frequency components, making it suitable for electricity price data which is often volatile [66]. These decompositions allow the MRC and BiLSTM components to better capture temporal patterns and dependencies in the electricity price data. These findings highlight the importance of the hybrid approach for accurate modelling of electricity prices, which are inherently volatile and influenced by numerous factors such as demand–supply dynamics, weather conditions, and market regulations.

Despite promising results, the model performance may vary with different datasets and market conditions. Additionally, due to the complex nature of electricity prices, involving rapid price spikes, high volatility, and seasonality, the overall performance of various types of predictive models still continues to be competitive and cannot consistently outperform one another [105]. Compared to previous studies that employed single decomposition methods or standard neural network architectures, our results demonstrate that hybrid or multi-stage models are superior, which is in agreement with published scholars [28,45,106,107]. This improvement is particularly evident when benchmarked against price prediction models reported in recent literature,

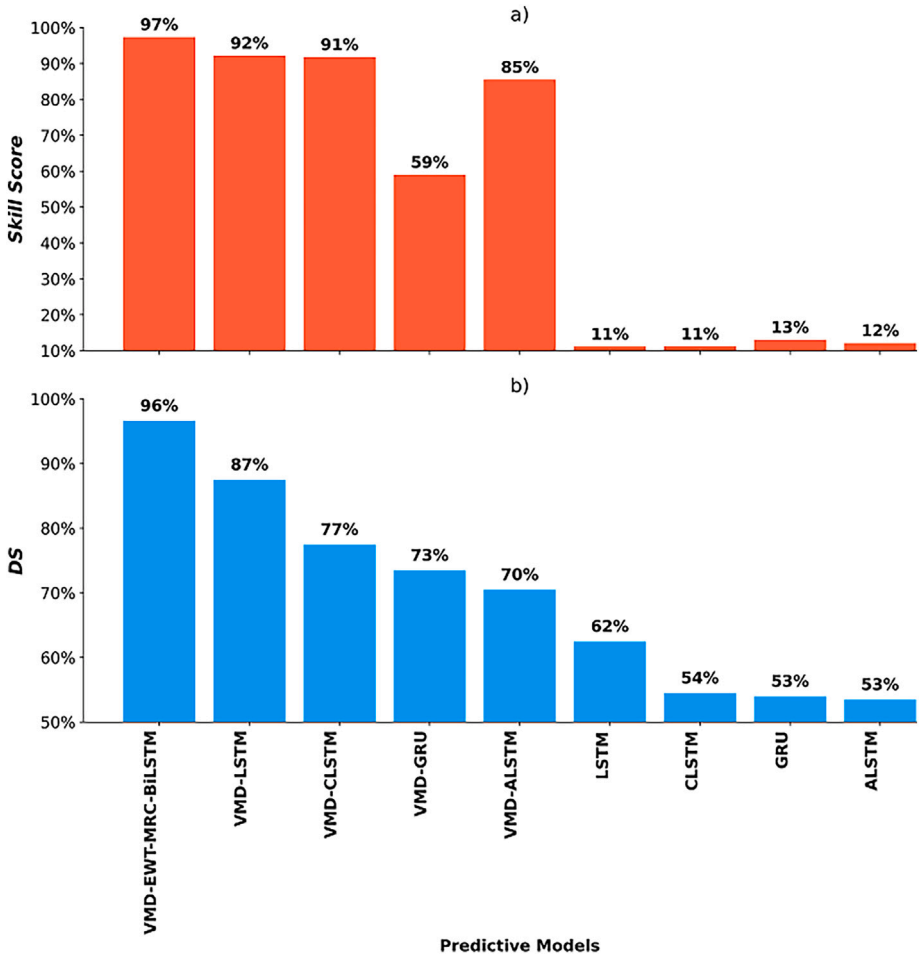


Fig. 18. The performance of the proposed VMD-EWT-MRC-BiLSTM model in comparison to other models using a) Skill score ($RMSE_{SS}$) and b) Directional symmetry (D_{stat}) criteria.

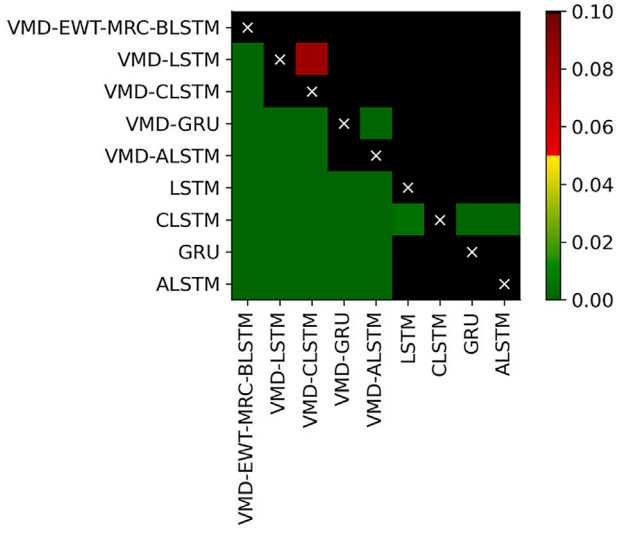


Fig. 19. The $p-values$ of the unconditional Giacomini–White test for superior predictive ability between the VMD-EWT-MRC-BiLSTM model and each of the benchmark models. $p-values$ close to 0 represent cases where the predicted output from the model on the x-axis is significantly more accurate than the forecast on the y-axis.

such as standard GRU and LSTM [105,108] or single-decomposition VMD-LSTM and VMD-GRU [109–112]. However, other studies stated that hybrid models do not necessarily outperform their standalone counterparts [113,114]. Therefore, additional exploration of this argument is highly recommended.

5. Conclusion and recommendations

This study presents a novel deep hybrid model for electricity price prediction, integrating Variational Modal Decomposition (VMD), Empirical Wavelet Transform (EWT), Multi-scale Residual Convolutional block (MRC), and Bidirectional Long Short-Term Memory (BiLSTM). This hybrid approach aims to leverage the strengths of each standalone algorithm in addressing the complex and non-linear nature of electricity price data, offering a comprehensive framework that improves prediction accuracy. The proposed VMD-EWT-MRC-BiLSTM model was evaluated and bench-marked against other decomposition-based and standalone models using a comprehensive set of performance criteria.

The proposed VMD-EWT-MRC-BiLSTM model represents a significant advancement in the field of electricity price prediction. By effectively blending decomposition techniques with advanced deep learning architectures, the model addresses the complex nature of electricity price data, achieving superior accurate and reliable predictions. The comprehensive evaluation of the proposed hybrid model against several decomposition-based and standalone alternatives has demonstrated its

superior performance across a variety of metrics and time scales. This research contributes to ongoing efforts to enhance predictive models in the energy sector, offering substantial benefits for various stakeholders, and providing the tool for more efficient and informed decision-making processes. For instance, energy traders can have better decision-making, allowing them to optimise their trading strategies and maximise profits. Furthermore, utility companies can benefit from precise forecasts to manage supply and demand more effectively, ensuring stability and reducing the risk of blackouts. Regulators and policymakers can use reliable price forecasts to design better market regulations and policies that promote fair competition and consumer protection. Finally, researchers can build upon this hybrid approach to further enhance predictive models in other domains of time series prediction.

There are multiple promising directions for future research that can extend the contributions of this study. First, optimizing the model's parameters and refining its architectural design could lead to further improvements in forecasting performance. In particular, the current study utilised empirical heuristics to determine decomposition parameters, such as the number of modes (K) in Variational Mode Decomposition (VMD) based on center frequency stabilization, and the number of bands (N) in Empirical Wavelet Transform (EWT) determined using reconstruction error measured by RMSE. While these approaches are practical and commonly used, they lack a systematic examination of parameter sensitivity. Future work should include a comprehensive sensitivity analysis to better understand the influence of these parameters on the model's output and to enhance the robustness, interpretability, and reproducibility of the approach. In addition, future studies could focus on minimizing the residual error between predicted and actual electricity prices while also incorporating methods to estimate uncertainty at various confidence levels. Incorporating error correction mechanisms and probabilistic forecasting techniques could improve the accuracy and reliability of the proposed model [115].

Another important direction involves the inclusion of additional exogenous variables such as weather data, electricity demand, generation capacity, fuel prices, economic indicators, and social or policy-related factors. Prior studies [13,113,116–118] have demonstrated that such variables can improve model performance, and their integration could further strengthen the forecasting capabilities of the proposed approach. Developing real-time forecasting systems [119] based on the proposed model would also offer substantial practical benefits, providing timely, data-driven insights to electricity market stakeholders and grid operators, thus enhancing their ability to make informed operational and strategic decisions. Furthermore, a more detailed analysis of how different models perform under specific scenarios, such as sharp short-term fluctuations or long-term trend turnarounds, is an important area for future research. Exploring how models behave in response to sudden price spikes, market crashes, or significant shifts in trends will provide a deeper understanding of their strengths and weaknesses, ultimately contributing to the optimization and practical application of these models. Finally, it is important to address the geographical limitation of the current study. The model was trained and tested using data from South Australia, and therefore its generalization to other regions with different market structures or environmental conditions remains uncertain. Future research should validate the proposed model using datasets from various geographic areas, both within and beyond Australia, to assess its scalability and applicability across diverse energy markets.

CRediT authorship contribution statement

Sujan Ghimire Writing – original draft, Software, Investigation, Data curation, Conceptualization. **Thong Nguyen-Huy:** Writing – review & editing, Validation, Supervision, Investigation. **Ravinesh C. Deo:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization. **David Casillas-Pérez** Writing – review & editing, Visualization, Validation, Investigation, Conceptualization. **A. A. Masrur Ahmed** Writing – review & editing, Validation, Supervision, Investigation. **Sancho Salcedo-Sanz:** Writing – review & editing, Supervision, Investigation, Funding acquisition, Conceptualization.

Declaration of competing 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.

Acknowledgement

The authors thank the data providers, all reviewers and Editor for their thoughtful comments, suggestions and review process. Partial support for this work was through a project PID2020-115454GB-C21 of the Spanish Ministry of Science and Innovation (MICINN). This work has been partially supported through the LATENTIA project PID2022-140786NB-C31 of the Spanish Ministry of Science, Innovation and Universities (MICINNU).

Appendix A. Literature review

Table A.12 summarizes the recent state-of-the-art models for EP prediction.

Table A.12
Summary of recent models for EP prediction.

Ref.	Model	Inputs	Prediction horizon	Datasets	Decomposition algorithm
[56]	STL-VMD-MLP RFR-TabNet	half hourly lagged EP values	Short-term	Australia	STL and VMD
[120]	VMD-GWO-ATT-LSTM	Lagged price data	Half-hourly	Australia market	VMD
[55]	Multi-head self-attention and Convolutional Neural networks	168 hourly lagged EP values	Day ahead	Ontario electricity prices	None
[121]	Deep Graph Convolutional Neural Network (D-GCNN)	Lagged price data	Half-hourly	Market for India, Singapore and Australia	EMD and WT
[122]	LASSO Estimated AutoRegressive (LEAR)	Preceding 8 h of EP prices	Day ahead	Market for Ireland and Northern Ireland	None
[123]	(CNN-GRU) with an attention mechanism	Cyclical time variables and other climatic variables	Hourly	German electricity market	None
[43]	VMD-CLSTM-VMD-ERCRF	Record of EP of the past 24 h	Half-hourly	Queensland, Australia	VMD

(*continued on next page*)

Table A.12 *(continued)*

Ref.	Model	Inputs	Prediction horizon	Datasets	Decomposition algorithm
[124]	CNN-GRU	Lagged EP, electricity load, generation, import and export, and weather data	Hourly	Ontario electricity prices	None
[125]	Quantile Regression and Kalman Filter	Lagged EP data and other climatic variables	Daily	Iranian wholesale market	None
[126]	Combined probability forecasting system (CPFS)	Lagged price data	Half-hourly	New South Wales, Australia and Singapore half-hourly electricity price dataset	None
[127]	ICEEMDAN-CNN-SSDAE	Lagged price data	Day ahead	Australia market	ICEEMDAN
[3]	RF-IMD-ICEEMD-VMD-Bi-LSTM	Historical data and other Renewable energy variables	Daily	Denmark market	ICEEMD and VMD
[39]	CNN-LSTM	168 hourly lagged EP values	Day ahead	Iranian electricity market	None
[128]	ACBFS-VMD-BOHB-LSTM	Lagged price data	Day ahead	PJM	VMD
[14]	VMD-Res.-EEMD-DE-ELM-DE-ELM	Lagged price data	Half-hourly	Australia market	VMD and EEMD
[42]	CNN-LSTM	24, 168, and 720 hourly lagged EP values	Day, week, and month ahead	German electricity spot price	None
[30]	LSTM	Record of EP for the past 96 h	Four hours ahead	Spanish electricity dataset	None
[129]	ARD-ETR	168 hourly lagged EP values	Short-term	Nord pool market	None
[28]	VMD-CNN-GRU	Record of EP of the past 24 h	Hour ahead	New York City in the United States	VMD
[130]	WT-SAE-LSTM	168 hourly lagged EP values	Short-term (Hour ahead)	Shandong province	WT
[131]	ICEEMDAN-VMD-MOGWO-ENN	Lagged price data	Half-hourly	Australia market	ICEEMDAN and VMD
[132]	WT-Adam-LSTM	168 hourly lagged EP values	Hour ahead, Day ahead	New South Wales of Australia and French	WT
[133]	MOBBSA-ANFIS	168 hourly lagged EP values	Short-term	Ontario market	None
[29]	LSTM	24, 168, hourly lagged EP values	Hour and day ahead	PJM	None
[41]	CNN-LSTM	Record of EP of the past 24 h	Short-term (Hour ahead)	PJM Regulation Zone Preliminary Billing Data	None
[35]	FEEMD-VMD-FA-BP	Lagged price data	Half-hourly	Australia market	FEEMD and VMD
[134]	RVMs-LR	24, 168, hourly lagged EP values	Day ahead	New England Electricity market	None

Data availability

Data were acquired from AEMO (https://www.aemo.com.au/).

References

[1] Yang H, Schell KR. Real-time electricity price forecasting of wind farms with deep neural network transfer learning and hybrid datasets. Appl Energy 2021;299:117242.
[2] Villavicencio Calzadilla P, Mauger R. The UN's new sustainable development agenda and renewable energy: the challenge to reach sdg7 while achieving energy justice. J Energy Nat Resour Law 2018;36(2):233–54.
[3] Wang K, Yu M, Niu D, Liang Y, Peng S, Xu X. Short-term electricity price forecasting based on similarity day screening, two-layer decomposition technique and BI-LSTM neural network. Soft Comput 2023;136:110018.
[4] Windler T, Busse J, Rieck J. One month-ahead electricity price forecasting in the context of production planning. J Clean Prod 2019;238:117910.
[5] Ventosa M, Baıllo A, Ramos A, Rivier M. Electricity market modeling trends. Energy Policy 2005;33(7):897–913.
[6] Day CJ, Hobbs BF, Pang J-S. Oligopolistic competition in power networks: a conjectured supply function approach. IEEE Trans Power Syst 2002;17(3):597–607.
[7] Chatzidimitriou KC, Chrysopoulos AC, Symeonidis AL, Mitkas PA. Enhancing agent intelligence through evolving reservoir networks for predictions in power stock markets. In: Agents and data mining interaction: 7th international workshop on agents and data mining interation, ADMI 2011, Taipei, Taiwan, May 2-6, 2011, revised selected papers 7; Springer; 2012. p. 228–47.
[8] Liebl D. Modeling and forecasting electricity spot prices: a functional data perspective. Ann Appl Stat 2013:1562–92.
[9] Kristiansen T. Forecasting nord pool day-ahead prices with an autoregressive model. Energy Policy 2012;49:328–32.
[10] Karakatsani NV, Bunn DW. Forecasting electricity prices: the impact of fundamentals and time-varying coefficients. Int J Forecast 2008;24(4):764–85.
[11] Carmona R, Coulon M. A survey of commodity markets and structural models for electricity prices. In: Quantitative energy finance: modeling, pricing, and hedging in energy and commodity markets. Springer; 2013. pp. 41–83.
[12] Hamilton JD. Regime switching models. In: Macroeconometrics and time series analysis. Springer; 2010. pp. 202–09.
[13] Weron R. Electricity price forecasting: a review of the state-of-the-art with a look into the future. Int J Forecast 2014;30(4):1030–81.
[14] 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.
[15] Kapoor G, Wichitaksorn N. Electricity price forecasting in New Zealand: a comparative analysis of statistical and machine learning models with feature selection. Appl Energy 2023;347:121446.
[16] Chai S, Li Q, Abedin MZ, Lucey BM. Forecasting electricity prices from the state-of-the-art modeling technology and the price determinant perspectives. Res Int Bus Finance 2023:102132.
[17] Prieto-Herráez D, Martínez-Lastras S, Frías-Paredes L, Asensio MI, González-Aguilera D. Eolo, a wind energy forecaster based on public information and automatic learning for the Spanish electricity markets. Measurement 2024;231:114557.
[18] Marcjasz G, Narajewski M, Weron R, Ziel F. Distributional neural networks for electricity price forecasting. Energy Econ 2023;125:106843.
[19] 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.
[20] Manfre Jaimes D, Zamudio López M, Zareipour H, Quashie M. A hybrid model for multi-day-ahead electricity price forecasting considering price spikes. Forecasting 2023;5(3):499–521.
[21] Grothe O, Kächele F, Krüger F. From point forecasts to multivariate probabilistic forecasts: the Schaake shuffle for day-ahead electricity price forecasting. Energy Econ 2023;120:106602.
[22] Gonzalez V, Contreras J, Bunn DW. Forecasting power prices using a hybrid fundamental-econometric model. IEEE Trans Power Syst 2011;27(1):363–72.
[23] Chen X, Dong ZY, Meng K, Xu Y, Wong KP, Ngan HW. Electricity price forecasting with extreme learning machine and bootstrapping. IEEE Trans Power Syst 2012;27(4):2055–62.
[24] Sai W, Pan Z, Liu S, Jiao Z, Zhong Z, Miao B, et al. Event-driven forecasting of wholesale electricity price and frequency regulation price using machine learning algorithms. Appl Energy 2023;352:121989.

[25] Sarwar S, Aziz G, Tiwari AK. Implication of machine learning techniques to forecast the electricity price and carbon emission: evidence from a hot region. Geosci Front 2024;15(3):101647.
[26] Hanif M, Shahzad MK, Mehmood V, Saleem I. Epfg: electricity price forecasting with enhanced gans neural network. IETE J Res 2023;69(9):6473–82.
[27] Abroun M, Jahangiri A, Shamim AG, Heidari H. Predicting long-term electricity prices using modified support vector regression method. Electr Eng 2024:1–12.
[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(2):2511–32.
[29] Zhou S, Zhou L, Mao M, Tai H-M, Wan Y. An optimized heterogeneous structure LSTM network for electricity price forecasting. IEEE Access 2019;7:108161–73.
[30] Torres JF, Martínez-Álvarez F, Troncoso A. A deep LSTM network for the Spanish electricity consumption forecasting. Neural Comput Appl 2022;34(13):10533–45.
[31] Reddy SS, Jung C-M. Short-term load forecasting using artificial neural networks and wavelet transform. Int J Appl Eng Res 2016;11(19):9831–36.
[32] Zhang X, Wang J, Zhang K. Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by cuckoo search algorithm. Electr Power Syst Res 2017;146:270–85.
[33] Wan C, Xu Z, Wang Y, Dong ZY, Wong KP. A hybrid approach for probabilistic forecasting of electricity price. IEEE Trans Smart Grid 2013;5(1):463–70.
[34] Osório GJ, Matias JCO, Catalão JP. Electricity prices forecasting by a hybrid evolutionary-adaptive methodology. Energy Convers Manag 2014;80:363–73.
[35] Wang D, Luo H, Grunder O, Lin Y, Guo H. Multi-step ahead electricity price forecasting using a hybrid model based on two-layer decomposition technique and bp neural network optimized by firefly algorithm. Appl Energy 2017;190:390–407.
[36] Yaslan Y, Bican B. Empirical mode decomposition based denoising method with support vector regression for time series prediction: a case study for electricity load forecasting. Measurement 2017;103:52–61.
[37] Sapnken FE, Tazehkandgheshlagh AK, Diboma BS, Hamaidi M, Noumo PG, Wang Y, et al. A whale optimization algorithm-based multivariate exponential smoothing grey-holt model for electricity price forecasting. Expert Syst Appl 2024;255:124663.
[38] Kitsatoglou A, Georgopoulos G, Papadopoulos P, Antonopoulos H. An ensemble approach for enhanced day-ahead price forecasting in electricity markets. Expert Syst Appl 2024;256:124971.
[39] Heidarpanah M, Hooshyaripor F, Fazeli M. Daily electricity price forecasting using artificial intelligence models in the Iranian electricity market. Energy 2023;263:126011.
[40] Deo RC, Grant RH, Webb A, Ghimire S, Igoe DP, Downs NJ, et al. Forecasting solar photosynthetic photon flux density under cloud cover effects: novel predictive model using convolutional neural network integrated with long short-term memory network. Stoch Environ Res Risk Assess 2022;36(10):3183–220.
[41] Kuo P-H, Huang C-J. An electricity price forecasting model by hybrid structured deep neural networks. Sustainability 2018;10(4):1280.
[42] Lehna M, Scheller F, Herwartz H. Forecasting day-ahead electricity prices: a comparison of time series and neural network models taking external regressors into account. Energy Econ 2022;106:105742.
[43] Ghimire S, Deo RC, Casillas-Pérez D, Salcedo-Sanz S. Two-step deep learning framework with error compensation technique for short-term, half-hourly electricity price forecasting. Appl Energy 2024;353:122059.
[44] 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.
[45] Zhang J, Tan Z, Wei Y. An adaptive hybrid model for short term electricity price forecasting. Appl Energy 2020;258:114087.
[46] Heydari A, Nezhad MM, Pirshayan E, Garcia DA, Keynia F, De Santoli L. Short-term electricity price and load forecasting in isolated power grids based on composite neural network and gravitational search optimization algorithm. Appl Energy 2020;277:115503.
[47] Wang B, Wang Z, Yao Z. Enhancing carbon price point-interval multi-step-ahead prediction using a hybrid framework of autoformer and extreme learning machine with multi-factors. Expert Syst Appl 2025:126467.
[48] Yao Z, Wang Z, Huang J, Xu N, Cui X, Wu T. Interpretable prediction, classification and regulation of water quality: a case study of Poyang lake, China. Sci Total Environ 2024;951:175407.
[49] Liu J, Lv Z, Zhao L. A dual-optimization building energy prediction framework based on improved dung beetle algorithm, variational mode decomposition and deep learning. Energy Build 2025;328:115143.
[50] Moreno SR, Seman LO, Stefenon SF, dos Santos Coelho L, Mariani VC. Enhancing wind speed forecasting through synergy of machine learning, singular spectral analysis, and variational mode decomposition. Energy 2024;292:130493.
[51] Ahmed AAM, Bailek N, Abualigah L, Bouchouicha K, Kuriqi A, Sharifi A, et al. Global control of electrical supply: a variational mode decomposition-aided deep learning model for energy consumption prediction. Energy Rep 2023;10:2152–65.
[52] Larcher JHK, Stefenon SF, dos Santos Coelho L, Mariani VC. Enhanced multi-step streamflow series forecasting using hybrid signal decomposition and optimized reservoir computing models. Expert Syst Appl 2024;255:124856.
[53] Lin Y, Lu Q, Tan B, Yu Y. Forecasting energy prices using a novel hybrid model with variational mode decomposition. Energy 2022;246:123366.
[54] Peng L, Wang L, Xia D, Gao Q. Effective energy consumption forecasting using empirical wavelet transform and long short-term memory. Energy 2022;238:121756.
[55] Pourdaryaei A, Mohammadi M, Mubarak H, Abdellatif A, Karimi M, Gryazina E, et al. A new framework for electricity price forecasting via multi-head self-attention and CNN-based techniques in the competitive electricity market. Expert Syst Appl 2024;235:121207.
[56] Ghimire S, Deo RC, Hopf K, Liu H, Casillas-Pérez D, Helwig A, et al. Half-hourly electricity price prediction model with explainable-decomposition hybrid deep learning approach. Energy AI 2025:100492.
[57] Ghimire S, Musaylh MSA, Nguyen-Huy T, Deo RC, Acharya R, Casillas-Pérez D, et al. Explainable deeply-fused nets electricity demand prediction model: factoring climate predictors for accuracy and deeper insights with probabilistic confidence interval and point-based forecasts. Appl Energy 2025;378:124763.
[58] Jayasinghe WJMLP, Deo RC, Ghahramani A, Ghimire S, Raj N. Deep multi-stage reference evapotranspiration forecasting model: multivariate empirical mode decomposition integrated with the boruta-random forest algorithm. IEEE Access 2021;9:166695–708.
[59] Jayasinghe WJMLP, Deo RC, Ghahramani A, Ghimire S, Raj N. Development and evaluation of hybrid deep learning long short-term memory network model for pan evaporation estimation trained with satellite and ground-based data. J Hydrol: Reg Stud 2022;607:127534.
[60] Ghimire S, Nguyen-Huy T, Prasad R, Deo RC, Casillas-Perez D, Salcedo-Sanz S, et al. Hybrid convolutional neural network-multilayer perceptron model for solar radiation prediction. Cognit Comput 2023;15(2):645–71.
[61] Ghimire S, Nguyen-Huy T, Musaylh MSA, Deo RC, Casillas-Pérez D, Salcedo-Sanz S. A novel approach based on integration of convolutional neural networks and echo state network for daily electricity demand prediction. Energy 2023;275:127430.
[62] Dragomiretskiy K, Zosso D. Variational mode decomposition. IEEE Trans Signal Process 2013;62(3):531–44.
[63] Isham MF, Leong MS, Lim MH, Ahmad ZA. Variational mode decomposition: mode determination method for rotating machinery diagnosis. J Vibroengineering 2018;20(7):2604–21.
[64] Sujadevi VG, Soman KP, Kumar SS, Mohan N, Arunjith AS. Denoising of phonocardiogram signals using variational mode decomposition. In: 2017 International conference on advances in computing, communications and informatics (ICACCI); IEEE; 2017. p. 1443–46.
[65] Liu Y, Huang S, Tian X, Zhang F, Zhao F, Zhang C. A stock series prediction model based on variational mode decomposition and dual-channel attention network. Expert Syst Appl 2023:121708.
[66] Gilles J. Empirical wavelet transform. IEEE Trans Signal Process 2013;61(16):3999–4010.
[67] Guoqiang S, Zhi L, Nayan Y, et al. Short-term wind power probability density prediction based on ewt and quantile regression forest. Electr Power Autom Equip 2018;292(8):165–72.
[68] Ghimire S, Deo RC, Casillas-Perez D, Salcedo-Sanz S. Improved complete ensemble empirical mode decomposition with adaptive noise deep residual model for short-term multi-step solar radiation prediction. Renew Energy 2022;190:408–24.
[69] Xu H, Wu L, Xiong S, Li W, Garg A, Gao L. An improved CNN-LSTM model-based state-of-health estimation approach for lithium-ion batteries. Energy 2023;276:127585.
[70] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 770–78.
[71] Liu T, Chen M, Zhou M, Du SS, Zhou E, Zhao T. Towards understanding the importance of shortcut connections in residual networks. Adv Neural Inf Process Syst 2019;32.
[72] Sun Y, He J, Ma H, Yang X, Xiong Z, Zhu X, et al. Online chatter detection considering beat effect based on inception and LSTM neural networks. Mech Syst Signal Process 2023;184:109723.
[73] Kim GI, Jang B. Petroleum price prediction with CNN-LSTM and CNN-GRU using skip-connection. Mathematics 2023;11(3):547.
[74] Targ S, Almeida D, Lyman K. Resnet in resnet: Generalizing residual architectures, arXiv preprint 2016 arXiv:1603.08029.
[75] Iandola F, Moskewicz M, Karayev S, Girshick R, Darrell T, Keutzer K. Densenet: implementing efficient convnet descriptor pyramids, arXiv preprint 2014 arXiv:1404.1869.
[76] Zhu Y, Newsam S. Densenet for dense flow. In: 2017 IEEE international conference on image processing (ICIP); IEEE; 2017. p. 790–94.
[77] Ghimire S, Deo RC, Casillas-Pérez D, Sharma E, Salcedo-Sanz S, Barua PD, et al. Half-hourly electricity price prediction with a hybrid convolution neural network-random vector functional link deep learning approach. Appl Energy 2024;374:123920.
[78] Ghimire S, Deo RC, Jiang N, Ahmed AAM, Prasad SS, Casillas-Pérez D, et al. Explainable deep learning hybrid modeling framework for total suspended particles concentrations prediction. Atmos Environ 2025:121079.
[79] Ghimire S, Deo RC, Casillas-Pérez D, Salcedo-Sanz S. Electricity demand error corrections with attention bi-directional neural networks. Energy 2024;291:129938.
[80] Jarque CM, Bera AK. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. Econ Lett 1980;6(3):255–59.
[81] World Health Organization and others. Who director-general's opening remarks at the media briefing on covid-19, (No Title) (2020.
[82] Farrow H. Commercial down v residential up: Covid-19's electricity impact. Energy Netw Aust 2020:1–23.
[83] Çelik D, Meral ME, Waseem M. The progress, impact analysis, challenges and new perceptions for electric power and energy sectors in the light of the covid-19 pandemic. Sustain Energy Grids Netw 2022;31:100728.
[84] Liu C, Wu YJ, Zhen CG. Rolling bearing fault diagnosis based on variational mode decomposition and fuzzy c means clustering. Proc CSEE 2015;35(13):3358–65.
[85] Močkus J. On bayesian methods for seeking the extremum. In: Optimization techniques IFIP technical conference: Novosibirsk, July 1–7, 1974; Springer; 1975. p. 400–04.

[86] Injadat M, Salo F, Nassif AB, Essex A, Shami A. Bayesian optimization with machine learning algorithms towards anomaly detection. In: 2018 IEEE global communications conference (GLOBECOM); IEEE; 2018. p. 1–6.
[87] Snoek J, Larochelle H, Adams RP. Practical bayesian optimization of machine learning algorithms. Adv Neural Inf Process Syst 2012;25.
[88] Seeger M. Gaussian processes for machine learning. Int J Neural Syst 2004;14(2):69–106.
[89] Hutter F, Hoos HH, Leyton-Brown K. Sequential model-based optimization for general algorithm configuration. In: Learning and intelligent optimization: 5th international conference, Lion 5, Rome, Italy, January 17-21, 2011. Selected Papers 5; Springer; 2011. p. 507–23.
[90] Bergstra J, Bardenet R, Bengio Y, Kégl B. Algorithms for hyper-parameter optimization. Adv Neural Inf Process Syst 2011;24.
[91] Elshawi R, Maher M, Sakr S. Automated machine learning: state-of-the-art and open challenges, arXiv preprint 2019 arXiv:1906.02287.
[92] Komer B, Bergstra J, Eliasmith C. Hyperopt-sklearn. Autom Mach Learn: Methods Syst Chall 2019:97–111.
[93] Van Rossum G, Drake Jr FL. Python tutorial. vol. 620. The Netherlands: Centrum voor Wiskunde en Informatica Amsterdam; 1995.
[94] Singh P, Manure A, Singh P, Manure A. Introduction to tensorflow 2.0. Learn TensorFlow 2.0: Implement Mach Learn Deep Learn Models Python 2020:1–24.
[95] Ghimire S, Deo RC, Casillas-Pérez D, Salcedo-Sanz S. Efficient daily electricity demand prediction with hybrid deep-learning multi-algorithm approach. Energy Convers Manag 2023;297:117707.
[96] Ghimire S, Nguyen-Huy T, Musaylh MSA, Deo RC, Casillas-Pérez D, Salcedo-Sanz S. Integrated multi-head self-attention transformer model for electricity demand prediction incorporating local climate variables. Energy AI 2023:100302.
[97] Ghimire S, Deo RC, Casillas-Perez D, Salcedo-Sanz S. Boosting solar radiation predictions with global climate models, observational predictors and hybrid deep-machine learning algorithms. Appl Energy 2022;316:119063.
[98] Behar O, Khellaf A, Mohammedi K. Comparison of solar radiation models and their validation under Algerian climate–the case of direct irradiance. Energy Convers Manag 2015;98:236–51.
[99] Giacomini R, White H. Tests of conditional predictive ability. Econometrica 2006;74(6):1545–78.
[100] Diebold FX, Mariano RS. Comparing predictive accuracy. J Bus Econ Stat 2002;20(1):134–44.
[101] Lago J, Marcjasz G, De Schutter B, Weron R. Forecasting day-ahead electricity prices: a review of state-of-the-art algorithms, best practices and an open-access benchmark. Appl Energy 2021;293:116983.
[102] 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.
[103] Li W, Becker DM. Day-ahead electricity price prediction applying hybrid models of LSTM-based deep learning methods and feature selection algorithms under consideration of market coupling. Energy 2021;237:121543.
[104] Huang Y, Dai X, Wang Q, Zhou D. A hybrid model for carbon price forecasting using garch and long short-term memory network. Appl Energy 2021;285:116485.
[105] Ugurlu U, Oksuz I, Tas O. Electricity price forecasting using recurrent neural networks. Energies 2018;11(5):1255.
[106] Shao Z, Zheng Q, Liu C, Gao S, Wang G, Chu Y. A feature extraction-and ranking-based framework for electricity spot price forecasting using a hybrid deep neural network. Electr Power Syst Res 2021;200:107453.
[107] 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.
[108] Ubrani A, Motwani S. LSTM-and GRU-based time series models for market clearing price forecasting of Indian deregulated electricity markets. In: Soft computing and signal processing: proceedings of ICSCSP 2018; vol. 2. Springer; 2019. p. 693–700.
[109] Jianwei E, Ye J, He L, Jin H. Energy price prediction based on independent component analysis and gated recurrent unit neural network. Energy 2019;189:116278.
[110] Huang Y, Deng Y. A new crude oil price forecasting model based on variational mode decomposition. Knowl-Based Syst 2021;213:106669.
[111] Sun W, Huang C. A novel carbon price prediction model combines the secondary decomposition algorithm and the long short-term memory network. Energy 2020;207:118294.
[112] Liu Y, Yang C, Huang K, Gui W. Non-ferrous metals price forecasting based on variational mode decomposition and LSTM network. Knowl-Based Syst 2020;188:105006.
[113] Lago J, De Ridder F, De Schutter B. Forecasting spot electricity prices: deep learning approaches and empirical comparison of traditional algorithms. Appl Energy 2018;221:386–405.
[114] Khosravi K, Golkarian A, Melesse AM, Deo RC. Suspended sediment load modeling using advanced hybrid rotation forest based elastic network approach. J Hydrol: Reg Stud 2022;610:127963.
[115] Zhang R, Li G, Ma Z. A deep learning based hybrid framework for day-ahead electricity price forecasting. IEEE Access 2020;8:143423–36.
[116] Zhang C, Li R, Shi H, Li F. Deep learning for day-ahead electricity price forecasting. IET Smart Grid 2020;3(4):462–69.
[117] Matsumoto T, Endo M. Electricity price forecast based on weekly weather forecast and its application to arbitrage in the forward market. In: 2021 11th International conference on power, energy and electrical engineering (CPEEE); IEEE; 2021. p. 104–11.
[118] Sgarlato R, Ziel F. The role of weather predictions in electricity price forecasting beyond the day-ahead horizon. IEEE Trans Power Syst 2022;38(3):2500–11.
[119] Mohsenian-Rad A-H, Leon-Garcia A. Optimal residential load control with price prediction in real-time electricity pricing environments. IEEE Trans Smart Grid 2010;1(2):120–33.
[120] Xu Y, Huang X, Zheng X, Zeng Z, Jin T. VMD-ATT-LSTM electricity price prediction based on grey wolf optimization algorithm in electricity markets considering renewable energy. Renew Energy 2024;236:121408.
[121] Rawal K, Ahmad A. Mining latent patterns with multi-scale decomposition for electricity demand and price forecasting using modified deep graph convolutional neural networks. Sustain Energy Grids Netw 2024;39:101436.
[122] O'Connor C, Collins J, Prestwich S, Visentin A. Electricity price forecasting in the Irish balancing market. Energy Strategy Rev 2024;54:101436.
[123] Laitsos V, Vontzos G, Bargiotas D, Daskalopulu A, Tsoukalas LH. Data-driven techniques for short-term electricity price forecasting through novel deep learning approaches with attention mechanisms. Energies 2024;17(7):1625.
[124] Ehsani B, Pineau P-O, Charlin L. Price forecasting in the Ontario electricity market via triconvgru hybrid model: univariate vs. multivariate frameworks. Appl Energy 2024;359:122649.
[125] Monjazeb MR, Amiri H, Movahedi A. Wholesale electricity price forecasting by quantile regression and kalman filter method. Energy 2024;290:129925.
[126] 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.
[127] Tan YQ, Shen YX, Yu XY, 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.
[128] Xiong X, Qing G. A hybrid day-ahead electricity price forecasting framework based on time series. Energy 2023;264:126099.
[129] Alkawaz AN, Abdellatif A, Kanesan J, Khairuddin ASM, Gheni HM. Day-ahead electricity price forecasting based on hybrid regression model. IEEE Access 2022;10:108021–33.
[130] Qiao W, Yang Z. Forecast the electricity price of US using a wavelet transform-based hybrid model. Energy 2020;193:116704.
[131] Yang W, Wang J, Niu T, Du P. A hybrid forecasting system based on a dual decomposition strategy and multi-objective optimization for electricity price forecasting. Appl Energy 2019;235:1205–25.
[132] Chang Z, Zhang Y, Chen W. Electricity price prediction based on hybrid model of adam optimized LSTM neural network and wavelet transform. Energy 2019;187:115804.
[133] Pourdaryaei A, Mokhlis H, Illias HA, Kaboli SHA, Ahmad S. Short-term electricity price forecasting via hybrid backtracking search algorithm and anfis approach. IEEE Access 2019;7:77674–91.
[134] Alamaniotis M, Bargiotas D, Bourbakis NG, Tsoukalas LH. Genetic optimal regression of relevance vector machines for electricity pricing signal forecasting in smart grids. IEEE Trans Smart Grid 2015;6(6):2997–3005.