Postprint: Applied Research on Two-layer Decomposition Technique in Electricity Price Forecasting

Authors: Yin Hao, Zeng Yun, Huang Shengquan, Dong Zhen

Date: 2018-05-18T00:00:00+00:00

Abstract

To address the problem of large electricity price fluctuations and low prediction accuracy, a multi-step electricity price forecasting model composed of a two-layer decomposition technique and neural network is proposed. The model first employs ensemble empirical mode decomposition to decompose the original electricity price series into a series of components, where variational mode decomposition further decomposes the highest frequency component generated from the first-layer decomposition into a series of modal components. All components are predicted using a neural network model, with the parameters of the neural network optimized using a crisscross optimization algorithm. Finally, all sub-sequences are superimposed to obtain the forecasted electricity price values. Simulation results demonstrate that the proposed model exhibits superior forecasting performance compared to other hybrid models and possesses high practical value.

Full Text

Research on Two-Layer Decomposition Technique for Electricity Price Forecasting

Yin Hao, Zeng Yun, Huang Shengquan, Dong Zhen
(College of Automation, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: To address the problems of high volatility and low prediction accuracy in electricity prices, this paper proposes a multi-step electricity price forecasting model that combines a two-layer decomposition technique with neural networks. The model first employs Ensemble Empirical Mode Decomposition (EEMD) to decompose the original electricity price series into multiple components. Variational Mode Decomposition (VMD) then further decomposes the highest-frequency component generated in the first layer into a series of modal components. All components are predicted using a neural network model, with the network parameters optimized by the Crisscross Optimization (CSO) algorithm. Finally, all sub-sequence predictions are superimposed to obtain the forecasted electricity price. Simulation results demonstrate that the proposed model exhibits superior forecasting performance compared to other hybrid models and possesses high practical value.

Keywords: two-layer decomposition; crisscross optimization algorithm; multi-step prediction; neural network; electricity price forecasting

0 Introduction

Electricity is an essential green and pollution-free energy source in today's world, playing a vital role in daily life. In increasingly competitive electricity markets, electricity can be freely traded as a commodity, making electricity price a crucial reflection of supply and demand dynamics in power markets. However, electricity prices are influenced by numerous complex factors including load, weather, business transactions, daily activities, and generation-side bidding strategies, resulting in random, unstable, and nonlinear characteristics that pose significant challenges for accurate price forecasting. The magnitude of electricity price fluctuations is closely related to market risk, making accurate forecasting critically important for both power systems and electricity markets.

Electricity price forecasting methods fall into two categories: causal prediction methods that identify relationships between independent variables and forecast values, and time series methods that assume historical prices correlate with future values. Given the difficulty of establishing causal relationships in practice, many researchers have turned to time series approaches. The most commonly used time series methods include statistical models, hybrid models, and artificial intelligence models. Numerous AI techniques have been applied to various forecasting domains, such as artificial neural networks and Support Vector Machines (SVM). For instance, one study used Particle Swarm Optimization (PSO) to optimize neural network thresholds and weights, proposing a PSO-BP model for marginal electricity price forecasting. Comparisons between optimized and non-optimized neural network models show that optimized models produce results closer to actual data. However, optimizing neural network parameters represents a large-scale multimodal optimization problem where PSO suffers from slow convergence and susceptibility to local optima.

Since many data series exhibit non-stationary and nonlinear characteristics (e.g., electricity prices, wind speed, load), single prediction models cannot accurately forecast such complex data. Consequently, scholars have employed various decomposition techniques to address these complexities and improve forecasting accuracy. For example, one study proposed a hybrid model combining Wavelet Packet Transform (WPT) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) for day-ahead electricity price forecasting, demonstrating that WPT significantly improves prediction accuracy. Another study developed a novel hybrid model based on Wavelet Transform (WT) and Genetic Algorithm (GA) optimized SVM. While single decomposition techniques like Empirical Mode Decomposition (EMD) and Ensemble EMD have been widely applied, high-frequency Intrinsic Mode Functions (IMFs), particularly IMF1, are easily affected by random factors, increasing modeling difficulty. Moreover, many nonlinear and non-stationary original sequences generate numerous high-frequency IMFs, making single-layer decomposition insufficient for handling irregular data series.

To address these limitations, this paper combines the advantages of single decomposition techniques and proposes a novel hybrid forecasting model (EEMD-VMD-CSO-BP) based on Ensemble Empirical Mode Decomposition, Variational Mode Decomposition, and a CSO-optimized neural network for multi-step electricity price forecasting. EEMD decomposes the original price series into IMFs with decreasing frequencies and a residual component. To reduce non-stationarity, VMD further decomposes high-frequency IMFs into multiple modes, while CSO optimizes the neural network' s weights and thresholds to improve forecasting accuracy. Using electricity price data from New South Wales and Queensland, Australia as case studies, the model performs one-step, three-step, and five-step forecasting, with results demonstrating its superior performance.

1 EEMD-VMD-CSO-BP Hybrid Model

The simulation data originates from actual electricity price measurements in the Australian electricity markets of New South Wales (Market 1) and Queensland (Market 2). Regional differences in population, geography, industrial structure, and climate characteristics create significant variations in price volatility, enabling comprehensive evaluation of the proposed model' s effectiveness and practicality. With one data point per half-hour, each day contains 48 observations. Using price records from the first ten days of a month (720 observations), Figure 1 [Figure 1: see original paper] shows the price series for both markets. The first 600 data points serve as training samples for forecasting points 600–700.

The EEMD-VMD-CSO-BP hybrid model framework is illustrated in Figure 2 [Figure 2: see original paper]. All experiments use half-hourly data from Market 1 as examples. The detailed steps are:

a) In the first decomposition layer, EEMD decomposes the original price series into multiple components to reduce nonlinearity and non-stationarity. The number of EEMD decompositions is determined by evaluating the Mean Absolute Percentage Error (MAPE) of the EEMD-CSO-BP hybrid model. As shown in Figure 3 [Figure 3: see original paper], MAPE decreases as decomposition number increases, but rises slightly when exceeding 8. Therefore, this model uses 8 EEMD components: IMF1, IMF2, ⋯, IMF7, and one residual. Figure 4 [Figure 4: see original paper] shows IMFs arranged from high to low frequency, where each IMF reflects different oscillation patterns in the series. IMF1 represents the highest-frequency signal containing detailed price information, while the residual represents the lowest-frequency signal showing the basic trend.

b) In Figure 5 [Figure 5: see original paper], except for IMF1 with high-frequency oscillation characteristics containing significant noise, other IMFs and the residual exhibit relatively stable fluctuations that are easier to forecast accurately. Since IMF1 increases prediction difficulty, VMD further decomposes it in the second layer to improve performance. Based on the proposed model, VMD decomposition number is determined as shown in Figure 4, where MAPE stabilizes after 9 components. Thus, IMF1 is decomposed into 9 modes: mode1, mode2, ···, mode9, as illustrated in Figure 6 [Figure 6: see original paper].

c) The CSO-optimized neural network forecasts all sub-sequences including mode1–mode9, IMF2–IMF7, and the residual. The final forecast is obtained by superimposing all sub-sequence predictions.

2.1 Ensemble Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) is an adaptive signal time-frequency processing method that decomposes nonlinear, non-stationary complex signals into Intrinsic Mode Functions (IMFs) with decreasing frequencies and a residual. Ensemble Empirical Mode Decomposition (EEMD) improves upon EMD by adding Gaussian white noise to address mode mixing while preserving the original signal. In this study, EEMD’ s three key parameters are set as: white noise K = 0.4, ensemble number n = 9, and replication times M = 200. Detailed EEMD principles are referenced in [16]. The decomposition steps are:

a) Add Gaussian white noise to the original data sequence.

b) Decompose the noisy sequence into multiple IMFs and a residual using EMD.

c) Repeat steps (1) and (2) with different white noise realizations to obtain corresponding IMFs and residuals (repeat n times).

d) Use the mean of all IMFs and the mean of all residuals as the final decomposition results.

2.2 Variational Mode Decomposition

Variational Mode Decomposition (VMD), proposed by Dragomiretskiy in 2014 [17], is a novel non-recursive signal processing technique that adaptively decomposes real-valued signals into discrete modes with specific sparse properties.

VMD is essentially a variational problem where all modes are treated as finite bandwidths with different center frequencies. Using the Alternating Direction Method of Multipliers (ADMM), all modes are gradually modulated to their corresponding baseband frequencies, yielding each mode and its center frequency.

To evaluate each mode' s bandwidth, the following steps are taken:

a) For the original signal $f(t)$, apply Hilbert transform for time-frequency processing to obtain the analytic signal of each modal function $u_k(t)$ and its single-sided spectrum.

b) To modulate each mode' s spectrum to baseband, multiply $u_k(t)$ by an exponential term $e^{-j\omega_k t}$ to adjust to the estimated center frequency ω_k.

c) Calculate each mode' s bandwidth using the Gaussian smoothness H1 of the modulated signal. This yields a constrained variational problem:

$$\min_{\{u_k\},\{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \quad \text{s.t.} \quad \sum_k u_k = f$$

where $f(t)$ is the original data signal, $\{u_k\} = \{u_1, u_2, \ldots, u_K\}$ are the K modal components obtained after decomposition, and $\{\omega_k\} = \{\omega_1, \omega_2, \ldots, \omega_K\}$ are the corresponding center frequencies.

Considering the penalty term α and Lagrange multiplier λ_k, the constrained problem can be converted to an unconstrained one, forming the augmented Lagrangian expression:

$$\mathcal{L}(\{u_k\},\{\omega_k\},\lambda) = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle$$

where α is the data fidelity constraint balance parameter that maintains reconstruction accuracy even in the presence of noise, and λ is the Lagrange operator. The saddle point of this expression (the optimal solution to Eq. (1)) can be found using the ADMM method in each iteration [18]. The update expressions for u_k and ω_k are:

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega}$$

where ω is the signal frequency, n represents the iteration number, $\hat{f}(\omega)$ is the Fourier transform of $f(t)$, and $\hat{u}_k(\omega)$ is the Fourier transform of $u_k(t)$.

3 Crisscross Optimization Algorithm

The Crisscross Optimization (CSO) algorithm consists of two operators: horizontal crossover and vertical crossover. Each iteration performs both operations, with the resulting offspring called moderate solutions ($MShc$, $MSvc$). The algorithm retains superior solutions (horizontal crossover superior solution $DShc$, vertical crossover superior solution $DSvc$) by comparing offspring with their parents based on fitness.

3.1 Horizontal Crossover Operation

Horizontal crossover, similar to crossover in genetic algorithms (with probability typically set to 1), performs arithmetic crossover between two particles randomly selected from the same dimension. Compared to traditional crossover, this enhances edge search capabilities for locally optimal particles and expands the search range. The offspring generation formulas are:

$$MS_{hc}(d,i) = r_1 \cdot X(i,d) + (1-r_1) \cdot X(j,d) + c_1 \cdot (X(i,d) - X(j,d))$$

$$MS_{hc}(d,j) = r_2 \cdot X(j,d) + (1-r_2) \cdot X(i,d) + c_2 \cdot (X(j,d) - X(i,d))$$

where $i, j \in \{1, 2, \dots, M\}$ are random particle indices; $d \in \{1, 2, \dots, D\}$ is a random dimension; $r_1, r_2 \in [0, 1]$ and $c_1, c_2 \in [-1, 1]$ are random numbers; $X(i,d)$ and $X(j,d)$ represent the d-th dimension of parent particles i and j; $MS_{hc}(d,i)$ and $MS_{hc}(d,j)$ are the d-th dimensional offspring after horizontal crossover.

From a sociological perspective, the first term in equations (4)-(5) represents the particle' s current best value, while the second term indicates parent particle interaction. These two terms are well-balanced by the inertia weight factor. The third term searches the population edges with a small probability, expanding the search range. The resulting offspring compete with their parents, with superior fitness values retained in $DShc$ for the next iteration.

3.2 Vertical Crossover Operation

Vertical crossover helps particles escape local optima without interfering with other dimensions. Since different dimensions have different ranges, all dimensions in the population must be normalized before crossover. The offspring formula after each vertical crossover operation is:

$$MS_{vc}(d_1,i) = r \cdot X(i,d_1) + (1-r) \cdot X(i,d_2)$$

where $i \in \{1,2,\ldots,M\}$, $d_1, d_2 \in \{1,2,\ldots,D\}$ are random dimensions, and $r \in [0,1]$ is a random number. The vertical crossover probability p_v (typically 0.2-0.8) enables particles to break free from local optima. After crossover, a competition mechanism compares the generated offspring with their parent $X(i,d_1)$, retaining superior particles in $DSvc$.

4 CSO Optimization of Neural Networks

The BP neural network [19] is a feedforward network based on error backpropagation, consisting of three layers: input, hidden, and output. The hidden layer connects inputs and outputs. Due to its strong self-learning capability and noise tolerance, BP has been widely applied across forecasting domains. However, BP' s gradient descent approach for weight and threshold adjustment often leads to local optima, with computational complexity increasing and convergence slowing as interference factors and sample data grow.

The CSO algorithm' s excellent global search capability significantly improves neural network forecasting performance. As shown in Figure 7 [Figure 7: see original paper], comparing the optimal convergence curves for one-step forecasting between CSO-BP and PSO-BP models (both with population size 20 and 100 iterations using Market 1 data), CSO-BP demonstrates superior performance.

Assuming the neural network has n input nodes, m hidden nodes, and h output nodes, the variable dimension is $D = n \times m + m \times h + m + h$. Each particle' s fitness is calculated using the mean square error formula:

$$fitness = \frac{1}{N}\sum_{t=1}^{N}(p_t - \hat{p}_t)^2$$

where p_t is the actual neural network output, $\hat{p}_t$ is the target output, and N is the number of training samples.

The CSO optimization process for neural networks is summarized as:

a) Determine BP network layers and neuron numbers based on training samples, then perform real-number encoding.

b) Set key parameters: population size, maximum iterations, and vertical crossover probability p_v. Randomly generate the initial population X in the encoded solution space.

c) Calculate fitness values for all particles using equation (7).

d) Perform horizontal crossover using equations (4)-(5), store offspring in matrix $MShc$, calculate their fitness, and compare with parent generation X (except first generation) to retain superior particles in $DShc$.

e) Perform vertical crossover using equation (6), store solutions in matrix $MSvc$, calculate fitness, and compare with parent X to retain superior particles in $DSvc$.

f) Check termination criteria. If iteration count exceeds the maximum, terminate and set the best solution in $DSvc$ as the neural network' s weights and thresholds. Otherwise, return to step d).

5 Case Study Analysis

To validate the effectiveness of the proposed EEMD-VMD-CSO-BP model for one-step and multi-step electricity price forecasting, two experimental studies were conducted. All models use 100 iterations and population size $M = 20$, with identical neural network parameters.

Experiment 1 uses New South Wales price data as a benchmark, comparing the proposed model against its components: BP, CSO-BP, VMD-CSO-BP, and EEMD-CSO-BP to verify hybrid strategy performance. **Experiment 2** comprehensively evaluates the proposed model using both New South Wales and Queensland price data, comparing it against three existing models: WT-GA-SVM, EMD-GA-BP, and PSO-BP. Three evaluation criteria are adopted: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and MAPE, expressed in equations (8)-(10):

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|X_t - \hat{X}_t|$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(X_t - \hat{X}_t)^2}$$

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{X_t - \hat{X}_t}{X_t}\right| \times 100\%$$

where N is the number of test data points, X_t is the actual electricity price at time t, and $\hat{X}_t$ is the forecasted price.

Table 1 shows prediction errors for the proposed model and benchmark models in one-step (S1), three-step (S2), and five-step (S3) forecasting. The proposed model achieves the smallest MAE, RMSE, and MAPE values across all horizons, confirming its superior decomposition capability.

To further analyze the advantages of the two-layer decomposition technique and CSO optimization, three comparative analyses were performed on single-step predictions (Table 2):

Comparison 1: VMD-CSO-BP and EEMD-CSO-BP reduce average percentage error by 46% and 62% respectively compared to CSO-BP, demonstrating the advantage of single decomposition techniques.

Comparison 2: EEMD-VMD-CSO-BP reduces average percentage error by 51% and 30% compared to VMD-CSO-BP and EEMD-CSO-BP respectively, proving that the proposed two-layer decomposition outperforms single-layer approaches. Figures 8 [Figure 8: see original paper]–10 [Figure 10: see original paper] show that as the forecasting horizon increases, the impact of two-layer decomposition on accuracy decreases because error accumulation in other IMFs (besides IMF1) grows, making VMD' s role in the second layer increasingly important.

Comparison 3: CSO-BP reduces average percentage error by 11% compared to BP-NN, confirming CSO' s positive impact on neural networks by optimizing weights and thresholds for higher accuracy.

Experiment 2 results (Table 3) show the proposed model achieves the smallest prediction errors across one-step, three-step, and five-step forecasts in both markets, further proving its superior forecasting capability and applicability.

6 Conclusion

This paper first analyzed the significance of electricity price forecasting for risk assessment and decision management in electricity markets. After reviewing existing forecasting models, a novel EEMD-VMD-CSO-BP hybrid model was proposed for multi-step price forecasting. The model consists of data preprocessing and forecasting stages. Experimental results demonstrate that the two-layer decomposition technique outperforms single-layer models, and the CSO algorithm positively impacts the BP model by optimizing input-hidden and hidden-output layer weights and thresholds, thereby improving forecasting accuracy. Compared with other hybrid models, the EEMD-VMD-CSO-BP hybrid model exhibits the best performance, proving highly suitable for non-stationary electricity price forecasting.

References

[1] Xiong Shangfei, Zou Xiaoyan. Review of research on electricity market price risk value and volatility forecasting [J]. Power System Protection and Control, 2014 (2): 146-153.

[2] Weron R. Electricity price forecasting: A review of the state-of-the-art with a look into the future [J]. International Journal of Forecasting, 2014, 30 (4): 1030-1081.

[3] Zhang Xian, Wang Xifan. Time-of-use short-term electricity price forecasting [J]. Proceedings of the CSEE, 2005, 25 (15): 1-6.

[4] He K, Yu L, Tang L. Electricity price forecasting with a BED (Bivariate EMD Denoising) methodology [J]. Energy, 2015, 91: 601-609.

[5] Santos N, Rui R. Performance of state space and ARIMA models for consumer retail sales forecasting [M]. Pergamon Press, Inc. 2015.

[6] Lei M, Luan S, Jiang C. A review on the forecasting of wind speed and generated power [J]. Renewable & Sustainable Energy Reviews, 2009, 13 (4): 915-920.

[7] Panapakidis I P, Dagoumas A S. Day-ahead electricity price forecasting via the application of artificial neural network based models [J]. Applied Energy, 2016; 172: 132-151.

[8] Ren Hong' e, Huo Mandong. Research on PSO-optimized SVM prediction application [J]. Application Research of Computers, 2009, 26 (3): 867-869.

[9] Yang Bo, Zhao Zunlian, Chen Yunping, et al. A new marginal price forecasting method based on PSO-evolved neural network ensemble [J]. High Voltage Engineering, 2007, 33 (10): 162-166.

[10] Deng Jiajia, Huang Yuansheng, Song Gaofeng. Application of non-parametric GARCH time series model in day-ahead electricity price forecasting [J]. Power System Technology, 2012, 36 (4): 190-196.

[11] Wang JZ, Wang Y, Jiang P. The study and application of a novel hybrid forecasting model: a case study of wind speed forecasting in China [J]. Applied Energy, 2015; 143: 472–488.

[12] Liu Xingjie, Mi Zengqiang, Yang Qixun, et al. A short-term wind speed multi-step forecasting method based on EMD [J]. Transactions of China Electrotechnical Society, 2010, 25 (4): 165-170.

[13] Chen Wenchi, Liu Fei. A hybrid denoising method for process data based on EEMD [J]. Application Research of Computers, 2012, 29 (4): 1368-1370.

[14] Liu Z, Sun W, Zeng J. A new short-term load forecasting method of power system based on EEMD and SS-PSO [J]. Neural Computing & Applications, 2014, 24 (3-4): 973-983.

[15] Huang S Z, Chang J X, Huang Q, et al. Monthly stream flow prediction using modified EMD-based support vector machine [J]. Journal of Hydrology, 2014, 511 (7): 764-775.

[16] Zhaohua W U, Huang N E. Ensemble empirical mode decomposition: a noise-assisted data analysis method [J]. Advances in Adaptive Data Analysis, 2009, 1 (1): 1-41.

[17] Dragomiretskity K, Zosso D. Variational mode decomposition [J]. IEEE Transactions on Signal Processing, 2014, 62 (3): 531-544.

[18] Carrion M, Arroyo J M, Conejo A J. A bilevel stochastic programming approach for retailer futures market trading [J]. IEEE Transactions on Power

Systems, 2009, 24 (3): 1446-1456.

[19] Sheng Li, Liu Xiyu, Gao Ming. A new globally optimized BP network [J]. Application Research of Computers, 2006, 23 (2): 211-212.

Note: Figure translations are in progress. See original paper for figures.