

Article

Short-Term Electricity Price Forecasting via a Reinforcement Learning-Based Dynamic Soft Ensemble Strategy

Yan Wang ¹, Yongxi Zhao ^{1,*}, Kun Liang ¹ and Hong Fan ²

¹ Faculty of Artificial Intelligence, Shanghai University of Electric Power, Shanghai 200090, China; a15830610206@mail.shiep.edu.cn (Y.W.)

² School of Electrical Engineering, Shanghai University of Electric Power, Shanghai 200090, China

* Correspondence: zhao_yongxi@shiep.edu.cn

Abstract

To address the high volatility of spot market prices and the feature extraction limitations of single models, a short-term electricity price forecasting method based on a reinforcement learning dynamic soft ensemble strategy is proposed. First, a complementary dual-branch architecture is constructed: the CNN-LSTM-Attention branch mines local temporal features, while the Transformer branch captures long-range global dependencies. Second, the Q-learning algorithm is introduced to model weight optimization as a Markov Decision Process. An intelligent agent perceives fluctuation states to adaptively allocate weights, overcoming the rigidity of traditional ensembles. Case studies on PJM market data demonstrate that the proposed model outperforms advanced benchmarks in MAE and RMSE metrics. Notably, prediction accuracy is significantly improved during price spikes and negative price periods. The results verify that the strategy effectively copes with market concept drift, supporting reliable bidding and risk mitigation.

Keywords: short-term electricity price forecasting; dynamic soft ensemble strategy; reinforcement learning; multi-architecture fusion; negative electricity prices; dynamic weight allocation

1. Introduction

With the deepening of market-oriented reforms in the electricity sector, a diversified competitive landscape centered on the spot market has initially taken shape [1,2]. In the new power system, influenced by the high integration of renewable energy and instantaneous changes in supply and demand, electricity price sequences exhibit characteristics of strong volatility [3], non-stationarity, and multiple time scales. Extreme phenomena, such as the coexistence of “negative electricity prices” and price spikes, occur frequently [4,5]. Negative electricity prices typically arise when high renewable energy generation (e.g., wind and solar) coincides with periods of low electricity demand. Under these conditions, inflexible baseload generators (such as nuclear or large coal units) often bid negatively to avoid the high economic and technical costs associated with shutting down and restarting, resulting in negative clearing prices. As the guiding signal for market resource allocation, price volatility directly impacts the operating efficiency of generation and sales companies and the energy costs of users. Therefore, constructing a short-term electricity price forecasting model with high precision and strong robustness is not only a guarantee for the efficient operation of the electricity market but also a key technical means for market entities to formulate scientific bidding strategies and mitigate spot market price risks.



Academic Editor: Seung-Hoon Yoo

Received: 9 January 2026

Revised: 26 January 2026

Accepted: 30 January 2026

Published: 1 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Currently, electricity price forecasting technologies are mainly categorized into statistical methods and artificial intelligence methods. Early statistical methods, such as Autoregressive Integrated Moving Average (ARIMA) [6] and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models, as well as shallow machine learning models like Support Vector Machines (SVMs), possess certain advantages when processing stationary sequences. However, when facing massive, high-dimensional, and nonlinear electricity price time-series data, the feature extraction capabilities and generalization performance of such models often fail to meet the requirements for high-precision forecasting.

In recent years, deep learning technology has become mainstream [7,8] due to its powerful nonlinear mapping capabilities. Convolutional Neural Networks (CNNs) [9], due to their local perception characteristics, can effectively capture short-term fluctuation patterns of electricity prices; Long Short-Term Memory (LSTM) networks solve the long-sequence dependency problem through gating mechanisms; and the Transformer architecture further achieves parallel capture of global dependencies in sequences via multi-head self-attention mechanisms [10,11]. However, single models often struggle to balance local abrupt changes and global trends: CNNs and LSTMs find it difficult to capture long-range dependencies, while Transformers have relatively insufficient sensitivity [12,13] when extracting local extreme features, leading to limited prediction accuracy at price spikes.

To break through the performance bottlenecks of single models, combination prediction strategies have emerged. Existing studies mostly employ linear weighting or static ensemble methods [14] based on the inverse of errors. However, the electricity market environment possesses significant “concept drift” characteristics [15]—meaning data distribution dynamically switches between stable periods, oscillating periods, and extreme negative price periods. Traditional static weighting strategies lack environmental perception capabilities and cannot adjust model emphasis in a timely manner when market states shift, leading to significantly amplified prediction deviations. Unlike existing static ensemble methods that rely on fixed weights or global error minimization, the proposed framework introduces a state-aware agent. This agent dynamically adjusts the fusion strategy at each time step, specifically targeting local volatility shifts that traditional methods fail to capture.

To address the aforementioned challenges, this paper proposes a short-term electricity price forecasting method based on a reinforcement learning dynamic soft ensemble strategy. Distinct from traditional static ensembles, this paper models the dynamic optimization problem of multi-model weights as a Markov Decision Process (MDP). First, a complementary dual-branch prediction architecture containing CNN-LSTM-Attention and the Transformer is constructed, which are responsible for the refined extraction of local temporal features and global long-range dependencies, respectively. Second, an intelligent ensemble agent based on Q-learning [16–18] is designed. By perceiving electricity price fluctuation states (stable, oscillating, or extreme) in real time, it adaptively learns the optimal weight allocation strategy through interaction with the environment. This method endows the prediction model with environmental perception and dynamic decision-making capabilities, effectively solving the problem of prediction inaccuracy during negative electricity prices and peak moments, providing a new engineering solution for electricity price forecasting in complex market environments. The remainder of this paper is organized as follows: Section 2 details the proposed dual-branch deep learning architecture, the reinforcement learning-based dynamic ensemble strategy, and the data description. Section 3 focuses on the experimental results, including performance analysis under extreme market conditions. Section 4 discusses the mechanism analysis and practical implications. Finally, Section 5 concludes the paper.

2. Methodology

2.1. Dual-Branch Deep Learning Architecture

The proposed soft ensemble strategy is built upon the complementary advantages of heterogeneous neural networks. This section details the mathematical characteristics of each fundamental component.

2.1.1. CNN-LSTM Module

The first component of the dual-branch architecture adopts a series structure of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, which are designed to extract local fluctuation features and short-term temporal dependencies. The model first utilizes CNN convolutional layers to perform local feature extraction on the input price sequence, filtering out high-frequency random noise via weight sharing and extracting underlying local trend features. The processed feature vectors are then fed into the LSTM network. Through gating mechanisms, LSTM effectively addresses the vanishing gradient problem in long-sequence training. The specific structure is illustrated in Figure 1. The state update formulas for the LSTM unit are as follows:

$$\begin{cases} f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ \tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \\ C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \\ o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t = o_t * \tanh(C_t), \end{cases} \quad (1)$$

where x_t represents the CNN feature input; h_{t-1} is the hidden state at the previous time step; C_t and $\tilde{C}_t$ denote the cell state and candidate cell state, respectively; f_t , i_t , and o_t are the states of the forget gate, input gate, and output gate, respectively; W and b represent the weight matrices and bias vectors corresponding to each state; σ denotes the Sigmoid activation function; and $*$ represents the Hadamard product.

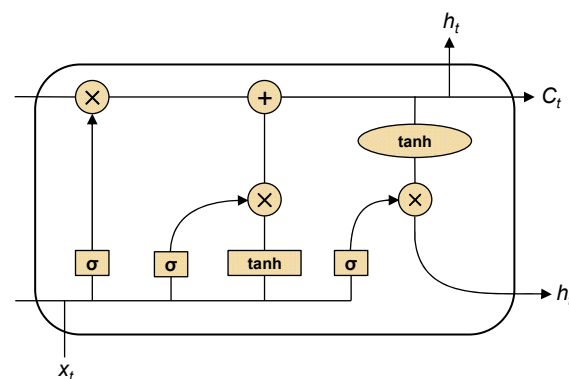


Figure 1. Schematic diagram of the LSTM unit structure.

2.1.2. Integration of Attention Mechanism

To complete the first branch and significantly enhance its ability to capture critical information such as price spikes, an Attention mechanism is introduced subsequent to the LSTM layer. This integration forms the CNN-LSTM-Attention branch [19]. The mechanism dynamically allocates weights to time steps by calculating the correlation between query

and key vectors, allowing the model to focus on historical moments with the greatest impact on future prices. The mathematical expression is calculated as follows:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \quad (2)$$

where Q , K , and V represent the query vector, key vector, and value vector, respectively; d_k denotes the dimension of the key vector [20]; and $\sqrt{d_k}$ serves as the scaling factor.

2.1.3. Transformer Branch

The second branch employs the Transformer architecture [21], which is designed to overcome the time-step limitations of recurrent networks and capture long-range global dependencies of the electricity price sequence in parallel. Unlike the recursive processing of LSTM, the Transformer is based entirely on the multi-head self-attention mechanism, enabling it to directly associate distant historical price features with the current prediction target across the temporal dimension. The structure of the multi-head attention mechanism is illustrated in Figure 2.

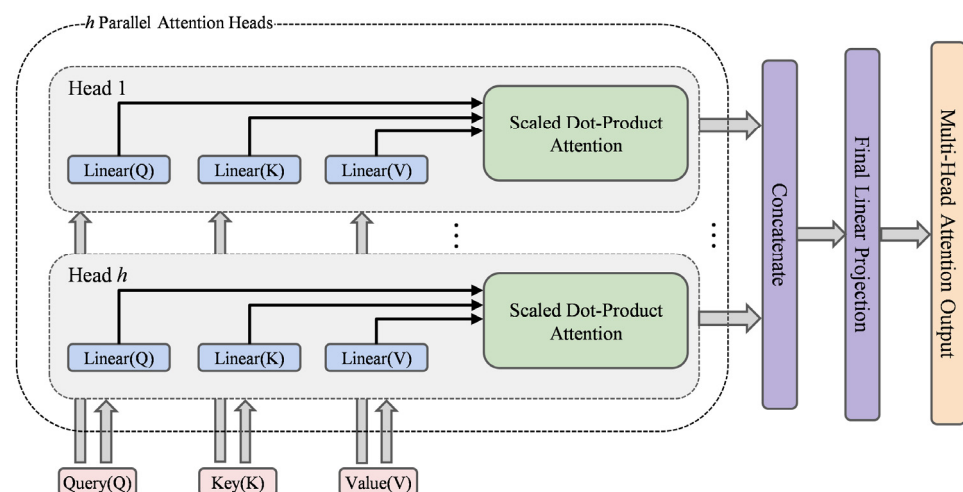


Figure 2. Structure of the multi-head attention mechanism.

The specific architectural hyperparameters for both branches, including the lookback window size, CNN kernel settings, and attention head configurations, are detailed in Section 2.3.

2.2. Reinforcement Learning-Based Soft Ensemble Strategy

2.2.1. Soft Ensemble Framework

To address the varying adaptability of single models across different market environments, this study proposes a soft ensemble framework that integrates dual-branch parallel prediction with dynamic weight optimization. The overall architecture and workflow are illustrated in Figure 3.

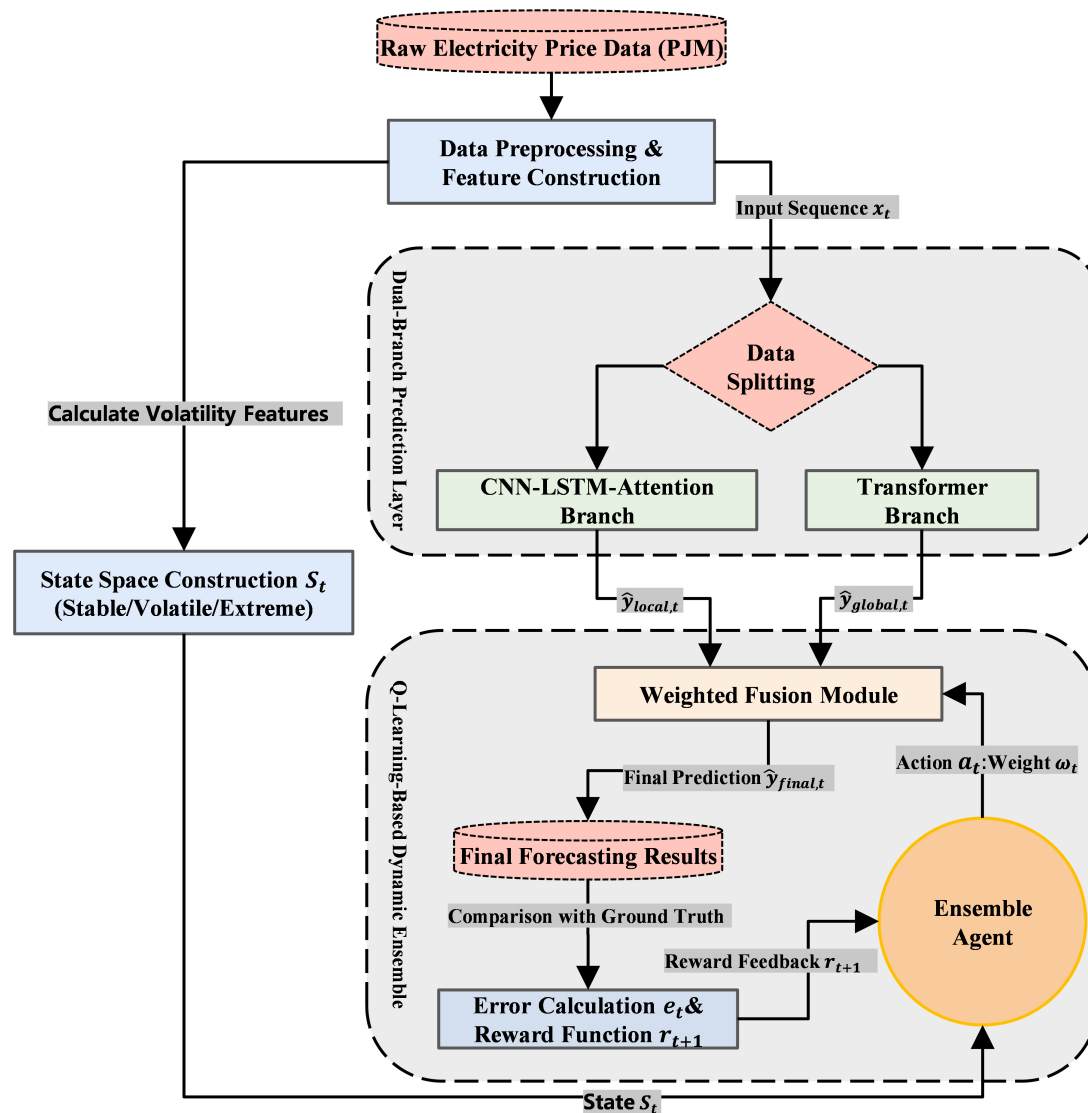


Figure 3. Dynamic soft ensemble forecasting framework based on reinforcement learning.

The process involves three specific stages:

1. **Dual-Branch Prediction:** Preprocessed sequences are fed into both the CNN-LSTM-Attention branch and the Transformer branch to generate basic predictions $\hat{y}_{local,t}$ and $\hat{y}_{global,t}$, respectively;
2. **Dynamic Weight Optimization:** An Ensemble Agent utilizes the Q-learning algorithm to dynamically determine the optimal weight coefficient ω_t based on the current fluctuation state;
3. **Weighted Fusion and Feedback:** The final prediction result is derived using the weighted fusion formula:

$$\hat{y}_{final,t} = \omega_t \cdot \hat{y}_{local,t} + (1 - \omega_t) \cdot \hat{y}_{global,t}, \quad (3)$$

where the prediction error is subsequently used to construct a reward signal to update the agent's strategy online.

2.2.2. Dynamic Weight Optimization via Q-Learning

The electricity spot market exhibits significant “concept drift” characteristics. To enable the model to actively adapt to environmental changes, the dynamic weight allocation

problem is modeled as a Markov Decision Process (MDP) [22] and solved using the Q-learning algorithm. The specific definitions are as follows:

- **State Space (S):** The state S_t characterizes the market fluctuation environment at time t . The standard deviation σ_t calculated over a 24-h sliding window of historical prices serves as the observation value. Based on thresholds θ_1 and θ_2 , the continuous volatility is discretized into three typical market states:

$$S_t = \begin{cases} 0 \text{ (Stable)}, & \text{if } \sigma_t < \theta_1 \\ 1 \text{ (Oscillating)}, & \text{if } \theta_1 \leq \sigma_t < \theta_2, \\ 2 \text{ (Extreme)}, & \text{if } \sigma_t \geq \theta_2 \end{cases} \quad (4)$$

- **Action Space (A):** The action a_t represents the weight coefficient ω_t of the CNN-LSTM-Attention branch. To balance computational efficiency and optimization precision, the continuous weight interval $[0, 1]$ is discretized into N candidate values:

$$A = \{0.0, 0.1, \dots, 1.0\}, \quad (5)$$

Upon selecting action a_t , the weight for the Transformer branch is automatically set to $1 - \omega_t$.

- **Reward Function (R):** A compound reward mechanism is designed to guide the model toward reducing overall error while penalizing prediction failures. As shown in Equation (6), a positive reward is granted if the prediction error decreases (i.e., $e_t < e_{t-1}$); otherwise, a penalty is applied. An additional heavy penalty λ is imposed via the indicator function $\mathbb{I}(\cdot)$ when the true price y_t exceeds the spike threshold δ or drops below zero:

$$r_{t+1} = \begin{cases} +1, & \text{if } e_t < e_{t-1} \\ -1 - \lambda \cdot \mathbb{I}(|y_t| > \delta \vee y_t < 0), & \text{otherwise} \end{cases} \quad (6)$$

- **Strategy Update:** The agent maintains a Q-table $Q(s, a)$ and updates it online using the Bellman equation to minimize the temporal difference error:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_{t+1} + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t)], \quad (7)$$

where α is the learning rate and γ is the discount factor. An ϵ -greedy strategy is employed to balance exploration and exploitation.

2.3. Experimental Setup and Evaluation Metrics

Experiments were conducted on a GPU-based computing platform using the TensorFlow 2.0 deep learning framework. Model training utilized the Adam optimizer [23] with an initial learning rate of 0.001 and a batch size of 64. Early stopping was implemented to prevent overfitting [24]. In the reinforcement learning module, the state space discretization thresholds were set to $\theta_1 = 5$ and $\theta_2 = 15$, determined by the 50th and 90th percentiles of the historical volatility distribution, respectively; the learning rate α was set to 0.1, and the discount factor γ was set to 0.9. Regarding the reward function, the penalty coefficient λ was set to 10, and the spike threshold δ was defined as the mean plus three standard deviations ($\mu + 3\sigma$) of the historical window.

Specifically, the model input was set to a lookback window of 168 h to effectively capture weekly periodicities, and the forecast horizon was set to 1 h (single-step prediction). The Mean Squared Error (MSE) was utilized as the loss function for model training. In terms

of network architecture, the CNN-LSTM-Attention branch employs 64 convolutional filters (kernel size of 3) followed by 128 LSTM hidden units. Parallely, the Transformer branch is configured with 2 encoder layers, 4 attention heads, and a feed-forward dimension of 256 with a dropout rate of 0.1.

To quantitatively assess forecasting performance, the Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE) were employed as core evaluation metrics:

$$\text{MAE} = \frac{1}{T} \sum_{t=1}^T |y_t - \hat{y}_t|, \quad (8)$$

$$\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2}, \quad (9)$$

$$\text{MAPE} = \frac{1}{T} \sum_{t=1}^T \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times 100\%, \quad (10)$$

where T denotes the total number of samples in the test set, y_t represents the actual electricity price at time t , and $\hat{y}_t$ is the predicted value.

It is important to note that the MAPE metric can be numerically unstable or undefined when actual prices (y_t) approach zero, and it lacks physical interpretability for negative prices. Therefore, while we report MAPE for overall performance comparison, this study prioritizes MAE and RMSE for the evaluation of negative price periods and spike scenarios. These metrics provide a robust measure of absolute error magnitude regardless of the price polarity.

2.4. Data Description and Preprocessing

This study utilizes public market data from the PJM Interconnection [25,26]. The dataset covers the period from 1 January 2023 to 31 December 2024 at an hourly resolution, totaling 17,544 sample points. To intuitively illustrate the market characteristics prior to modeling, a representative price sequence is visualized in Figure 4. As shown, the dataset exhibits significant non-stationarity and multi-scale volatility. Influenced by the instantaneous balance between supply and demand and the high penetration of renewable energy, the price sequence demonstrates frequent irregular fluctuations. Notably, the data is characterized by an extreme distribution where price spikes and negative electricity prices coexist.

To fully capture the evolution patterns of electricity prices, the model input features include multi-source variables such as historical price sequences, system load forecasts, and zonal load forecasts. In the preprocessing stage, the Min-Max normalization method is employed to map all data into the $[-1, 1]$ range. To strictly prevent future information leakage, the scaler is fitted solely on the training dataset. The validation and test datasets are subsequently transformed using the statistical parameters derived from the training set. Finally, the dataset is partitioned in chronological order, as detailed in Table 1.

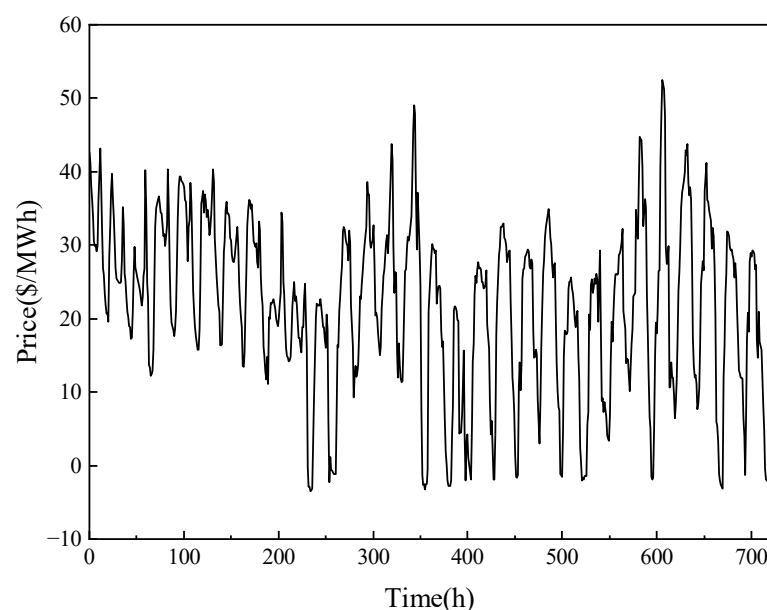


Figure 4. Visualization of typical electricity price fluctuations in the PJM market (representative 700-h window).

Table 1. Dataset partition and sample distribution of the PJM electricity market.

Dataset	Time Period	Sample Size
Training Set	1 January 2023–30 June 2024	13,128
Validation Set	1 July 2024–30 September 2024	2208
Test Set	1 October 2024–31 December 2024	2208

3. Results

3.1. Overall Forecasting Performance

To verify the performance of the proposed soft ensemble model, three deep learning models were selected as baselines: a Temporal Convolutional Network (TCN) [27], CNN-LSTM-Attention, and the Transformer [28]. The latter two serve as independent branches of the proposed framework, enabling an ablation study to validate the necessity of multi-architecture fusion. All models were trained and evaluated on the unified PJM dataset. The overall forecasting performance on the test set is presented in Table 2.

Table 2. Comparison of overall forecasting performance among different models on the test set.

Prediction Model	MAE (\$/MWh)	RMSE (\$/MWh)	MAPE (%)
TCN	2.342	3.586	9.152
CNN-LSTM-Attention	2.156	3.245	8.243
Transformer	2.048	3.120	7.821
Proposed Soft Ensemble	1.735	2.562	6.544

As shown in Table 2, the proposed soft ensemble model outperforms the comparison models across all metrics. Compared to the best-performing single baseline model (Transformer), the soft ensemble model reduces MAE and RMSE by 15.3% and 17.9%, respectively, demonstrating a significant improvement in MAPE.

Figure 5 illustrates the tracking curves of each model over continuous time periods. It reveals that while all models maintain good tracking performance during stable fluctuation intervals, single models exhibit performance bottlenecks during rapid jump and spike

periods (e.g., hours 80–100 and near hour 250). In contrast, the soft ensemble model achieves the best fit and response speed in non-stationary intervals.

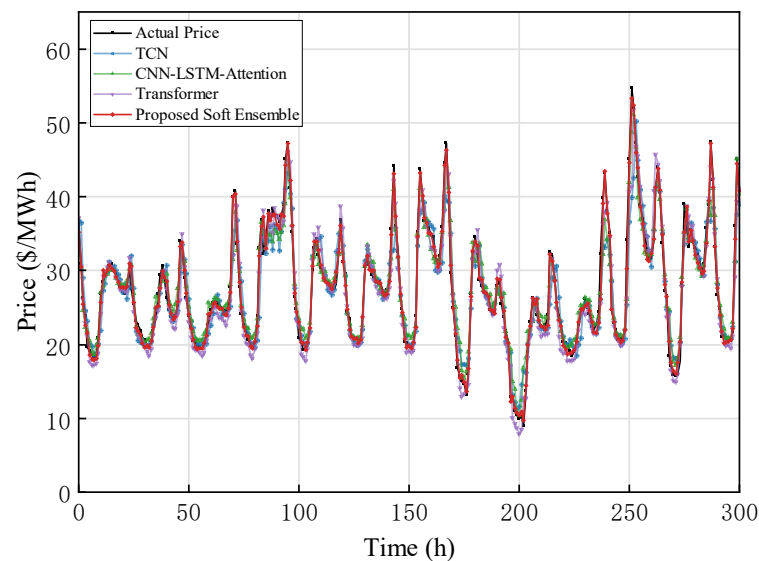


Figure 5. Comparison of forecasting results on the test set among different models.

3.2. Performance in Typical Scenarios

Extreme fluctuation scenarios in electricity price sequences represent critical dimensions for evaluating model robustness. This section provides an in-depth analysis of two typical extreme conditions: price spikes and negative electricity prices.

3.2.1. Price Spike Forecasting

Price jumps and spike moments typically entail high market risks [29]. In this study, extreme samples were strictly screened based on a rigorous statistical criterion. Specifically, 35 typical abrupt change points, where the absolute price deviation exceeded three times the rolling standard deviation of the preceding 24-h window, were selected from the test set for focused analysis. The tracking performance of each model in such extreme scenarios is shown in Figure 6.

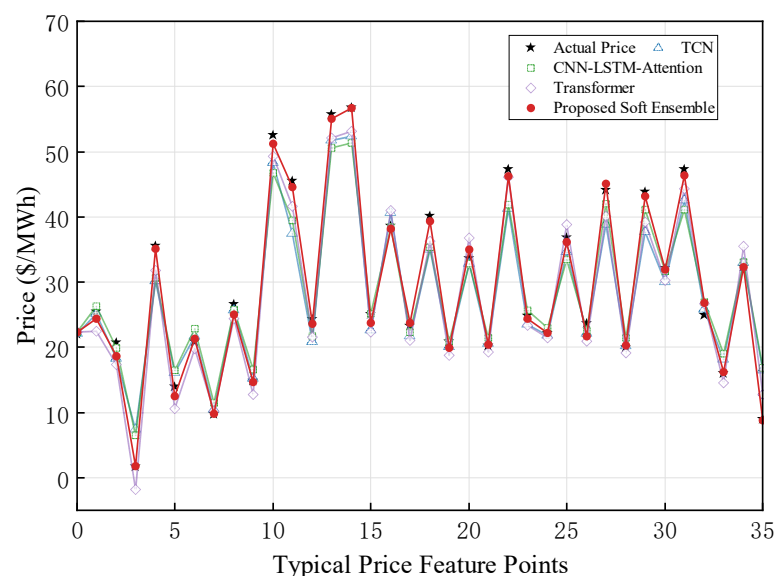


Figure 6. Comparison of forecasting performance under price spike scenarios.

As indicated in Table 3, the forecasting errors of single models escalate significantly under highly volatile price fluctuation scenarios. Conversely, the soft ensemble model demonstrates superior robustness. Its MAE and RMSE metrics are reduced by 39.3% and 37.1%, respectively, compared to the Transformer baseline model.

Table 3. Statistics of forecasting errors under price spike scenarios.

Prediction Model	MAE (\$/MWh)	RMSE (\$/MWh)
TCN	4.125	6.854
CNN-LSTM-Attention	3.962	6.420
Transformer	3.650	6.108
Proposed Soft Ensemble	2.215	3.842

3.2.2. Negative Electricity Price Forecasting

With the increasing penetration of renewable energy, negative electricity prices have evolved into a common phenomenon in the electricity market. This section evaluates the tracking capability of the models in these counter-intuitive extreme value ranges. The forecasting performance of each model during negative price periods is shown in Figure 7.

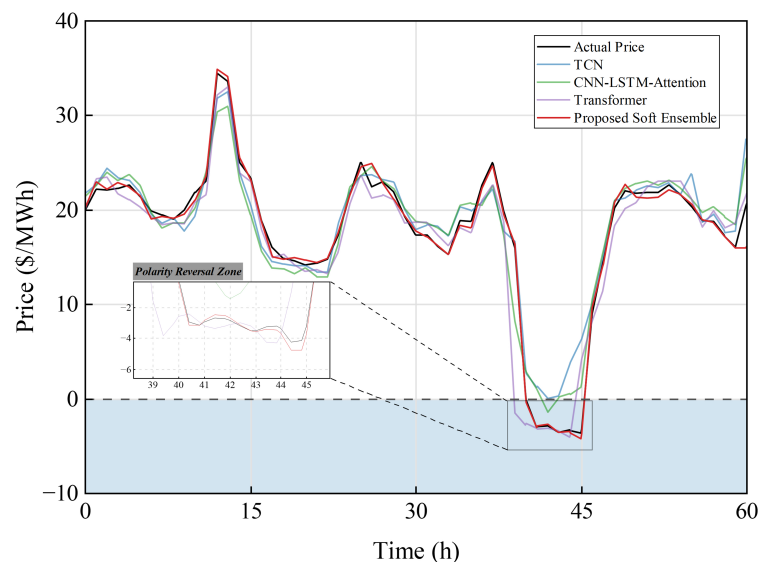


Figure 7. Comparison of forecasting performance under negative price extreme scenarios. The zoomed-in inset highlights the model's superior tracking capability during the polarity reversal phase (Time indices 39–45 h).

The micro-trajectory analysis in Figure 7 reveals significant heterogeneity in the failure modes of baseline models. The TCN model fails to respond effectively to polarity reversal signals, resulting in fundamental sign discrimination errors. Although the CNN-LSTM-Attention model captures the downward trend, it exhibits significant amplitude underfitting.

The Transformer model approximates the true value in amplitude but shows noticeable phase-leading discrepancies (i.e., predicting the drop earlier than it occurs). In contrast, the soft ensemble model balances magnitude depth and temporal accuracy, accurately fitting the negative price trajectory. Table 4 provides the error statistics for this scenario.

Table 4. Statistics of forecasting errors under negative price extreme scenarios.

Prediction Model	MAE (\$/MWh)	RMSE (\$/MWh)
TCN	3.854	5.420
CNN-LSTM-Attention	3.520	5.115
Transformer	3.186	4.850
Proposed Soft Ensemble	2.052	3.150

As shown in Table 4, the soft ensemble model exhibits excellent generalization capabilities. Its MAE and RMSE metrics are reduced by 35.6% and 35.1%, respectively, compared to the best-performing Transformer model.

4. Discussion

4.1. Mechanism Analysis of Deep Learning Branches

The core finding of this study is that the proposed soft ensemble strategy significantly outperforms single-structure models in handling extreme market volatility. This performance gap can be attributed to the inherent structural limitations of individual base models. While the CNN-LSTM-Attention branch excels at extracting high-frequency local oscillation features, its recursive nature limits its ability to model long-term dependencies spanning hundreds of hours. Conversely, the Transformer branch captures global context via self-attention but tends to smooth out local anomalies, leading to an “insensitivity” to instantaneous price spikes (as observed in the underfitting in Figure 6). Furthermore, the TCN baseline failed to predict negative prices accurately, primarily because its causal convolution operation lacks a specific mechanism to handle the abrupt “polarity reversal” of price signals, resulting in significant sign errors.

4.2. Efficacy of the RL-Based Dynamic Weighting Strategy

The superior performance of the soft ensemble model validates the necessity of dynamic weight allocation in addressing concept drift in electricity markets. Unlike traditional static ensemble methods (e.g., linear weighting or inverse-error weighting), the proposed Q-learning agent acts as an active decision-maker [30]. Through the designed reward function (Equation (6)), the agent receives heavy penalties when prediction failures occur during extreme events. It is crucial to highlight that a traditional static ensemble strategy (e.g., simple linear averaging) lacks this punishment-driven adaptability. In scenarios where one branch exhibits significant error (e.g., the Transformer branch failing to capture a sudden spike), a static ensemble would inevitably propagate this deviation into the final result. In contrast, the agent learns to identify “precursor states” of volatility. For instance, during price spikes, the agent adaptively increases the weight of the CNN-LSTM branch, which is more sensitive to local abrupt changes, while relying more on the Transformer branch during stable trends. This environmental perception capability allows the model to switch forecasting logic dynamically, ensuring robustness across conflicting market scenarios.

Furthermore, to verify the robustness of the proposed framework, a sensitivity analysis was conducted on the volatility state thresholds θ_1 and θ_2 . The experimental results indicate that the model’s forecasting performance remains stable when θ_1 varies within the range of 3 to 7 and θ_2 within 12 to 18, with MAE demonstrating negligible fluctuations. This confirms that the chosen parameters effectively capture market state transitions and that the agent is not overly sensitive to specific threshold settings.

Regarding the stability concern often associated with reinforcement learning [31], unlike methods that use DRL to directly generate price curves, our agent functions as a high-level weight scheduler. By leveraging the inherent stability of the supervised

base models (the CNN and Transformer), the proposed method effectively avoids the convergence difficulties and waveform oscillations common in generative RL approaches.

4.3. Practical Implications for Electricity Markets

From an engineering perspective, the proposed method offers substantial value for market participants. The accurate prediction of negative electricity prices is particularly critical for renewable energy generators and energy storage systems. By anticipating negative price intervals, operators can adjust bidding strategies [32]—such as reducing output or charging storage units—to convert potential economic losses into arbitrage opportunities. Similarly, the precise capture of price spikes assists utility companies in managing procurement risks and optimizing demand-side response measures. Moreover, this aligns with the emerging paradigm of “value-oriented forecasting” [33], where the prediction objective shifts from minimizing statistical errors to maximizing economic utility. By accurately identifying negative price intervals, the proposed model provides a reliable technical basis for market entities to identify potential arbitrage opportunities, facilitating a transition from passive acceptance to proactive risk management. Therefore, this study provides a reliable technical reference for algorithmic trading and risk management in modern power systems characterized by high renewable penetration.

5. Conclusions

To address the challenges of multi-scale volatility and distribution drift in electricity price forecasting for modern power systems, this paper proposes a soft ensemble forecasting method based on reinforcement learning. By constructing an adaptive mapping mechanism for state perception and weight reconstruction, this study overcomes the paradigm constraints of traditional static combination models. It achieves complementary advantages among heterogeneous sub-models under varying market conditions, effectively circumventing the structural limitations of single architectures.

Validation results indicate that while the proposed model delivers consistent improvements in general error metrics, its primary contribution is the superior robustness and transient response capability demonstrated specifically during high-frequency price spikes and negative price reversal scenarios, successfully resolving the engineering bottleneck of balancing prediction accuracy with generalization stability. This study verifies the effectiveness of the dynamic ensemble paradigm in processing non-stationary time-series data, providing valuable technical support for risk quantification and decision-making in electricity spot trading. However, certain limitations should be noted. First, the current framework relies on offline training, meaning its adaptability to unforeseen “black swan” events—extreme outliers beyond the historical distribution—remains to be verified. Second, while the reinforcement learning hyperparameters were validated on PJM data, application to other markets with different fluctuation characteristics may necessitate market-specific parameter tuning.

Future work will focus on incorporating more complex external variables (e.g., extreme weather events) into the state space and extending the current single-step model to multi-step-ahead forecasting to further enhance its applicability.

Author Contributions: Conceptualization, Y.W. and Y.Z.; methodology, Y.W.; software, Y.W.; validation, Y.W., K.L. and H.F.; formal analysis, Y.W.; investigation, Y.W.; resources, Y.Z. and K.L.; data curation, Y.W.; writing—original draft preparation, Y.W.; writing—review and editing, Y.Z. and H.F.; visualization, Y.W.; supervision, Y.Z. and H.F.; project administration, Y.Z.; funding acquisition, Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (Youth Program), grant number 62403303.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy restrictions.

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

Abbreviations

The following abbreviations are used in this manuscript:

ARIMA	Autoregressive Integrated Moving Average
CNN	Convolutional Neural Network
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MDP	Markov Decision Process
PJM	PJM Interconnection
RMSE	Root Mean Square Error
TCN	Temporal Convolutional Network

References

1. Ganczarek-Gamrot, A.; Gorczyca-Goraj, A.; Pilot, K.; Kania, K. Electricity price volatility and the performance of machine learning forecasting models in European energy markets. *Energies* **2025**, *18*, 6535. [\[CrossRef\]](#)
2. O'Connor, C.; Bahloul, M.; Prestwich, S.; Hebrard, E. A review of electricity price forecasting models in the day-ahead, intra-day, and balancing markets. *Energies* **2025**, *18*, 3097. [\[CrossRef\]](#)
3. Pavlík, M.; Kurimský, F.; Ševc, K. Renewable energy and price stability: An analysis of volatility and market shifts in the European electricity sector (2015–2025). *Appl. Sci.* **2025**, *15*, 6397. [\[CrossRef\]](#)
4. 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. [\[CrossRef\]](#)
5. Zamudio López, M.; Zareipour, H.; Quashie, M. Forecasting the occurrence of electricity price spikes: A statistical-economic investigation study. *Forecasting* **2024**, *6*, 115–137. [\[CrossRef\]](#)
6. Karabiber, O.A.; Xydis, G. Electricity price forecasting in the Danish day-ahead market using the TBATS, ANN and ARIMA methods. *Energies* **2019**, *12*, 928. [\[CrossRef\]](#)
7. Jasiński, T. A review of recent trends in electricity price forecasting using deep learning techniques. *Energies* **2025**, *18*, 6422. [\[CrossRef\]](#)
8. Pikus, M.; Waś, J.; Kozina, A. Using deep learning in forecasting the production of electricity from photovoltaic and wind farms. *Energies* **2025**, *18*, 3913. [\[CrossRef\]](#)
9. Alsmadi, L.; Lei, G.; Li, L. Forecasting day-ahead electricity demand in Australia using a CNN-LSTM model with an attention mechanism. *Appl. Sci.* **2025**, *15*, 3829. [\[CrossRef\]](#)
10. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Volume 30.
11. Li, S.; Jin, X.; Xuan, Y.; Zhou, X.; Chen, W.; Wang, Y.X.; Yan, X. Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 8–14 December 2019; Volume 32.
12. Lim, B.; Arik, S.O.; Loeff, N.; Pfister, T. Temporal fusion transformers for interpretable multi-horizon time series forecasting. *Int. J. Forecast.* **2021**, *37*, 1748–1764. [\[CrossRef\]](#)
13. Llorente, O.; Portela, J. A Transformer approach for electricity price forecasting. *arXiv* **2024**, arXiv:2403.16108. [\[CrossRef\]](#)
14. Wei, X.; Sun, X.; Sun, J. A comparative study of multi-model ensemble forecasting accuracy between equal- and variant-weight techniques. *Atmosphere* **2022**, *13*, 526. [\[CrossRef\]](#)
15. Jagait, R.K.; Fekri, M.N.; Grolinger, K.; Mir, S. Load forecasting under concept drift: Online ensemble learning with recurrent neural network and ARIMA. *IEEE Access* **2021**, *9*, 98992–99008. [\[CrossRef\]](#)
16. Feng, C.; Zhang, J. Reinforced deterministic and probabilistic load forecasting via Q-learning dynamic model selection. *IEEE Trans. Smart Grid* **2020**, *11*, 1377–1388. [\[CrossRef\]](#)

17. Pannakkong, W.; Vinh, V.T.; Tuyen, N.N.M.; Sriboonchitta, S. A reinforcement learning approach for ensemble machine learning models in peak electricity forecasting. *Energies* **2023**, *16*, 5099. [\[CrossRef\]](#)
18. Wang, J.; Liu, H.; Zheng, G. Short-term load forecasting based on outlier correction, decomposition, and ensemble reinforcement learning. *Energies* **2023**, *16*, 4401. [\[CrossRef\]](#)
19. Tusher, E.H.; Md Akbar, J.U.; Rabbi, R.I.; Islam, M.S. Energy forecasting utilizing CNN-LSTM attention mechanism: Empirical evidence from the Spanish electricity market. In *Machine Learning Technologies on Energy Economics and Finance: Energy and Sustainable Analytics*; Springer: Cham, Switzerland, 2025; Volume 2, pp. 51–72.
20. Laitsos, V.; Vontzos, G.; Bargiotas, D.; Daskalopulu, A. Data-driven techniques for short-term electricity price forecasting through novel deep learning approaches with attention mechanisms. *Energies* **2024**, *17*, 1625. [\[CrossRef\]](#)
21. Cantillo-Luna, S.; Moreno-Chuquen, R.; Lopez-Sotelo, J. An intra-day electricity price forecasting based on a probabilistic transformer neural network architecture. *Energies* **2023**, *16*, 6767. [\[CrossRef\]](#)
22. Sutton, R.S.; Barto, A.G. *Reinforcement Learning: An Introduction*, 2nd ed.; MIT Press: Cambridge, MA, USA, 2018.
23. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
24. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
25. PJM Interconnection. In *PJM Manual 11: Energy & Ancillary Services Market Operations*; PJM Interconnection: Norristown, PA, USA, 2024.
26. PJM Data Miner. Real-Time & Day-Ahead Energy Market Data. Available online: <https://dataminer2.pjm.com> (accessed on 10 January 2025).
27. Bai, S.; Kolter, J.Z.; Koltun, V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv* **2018**, arXiv:1803.01271. [\[CrossRef\]](#)
28. Tschora, L.; Pierre, E.; Plantevit, M.; Robardet, C. Electricity price forecasting on the day-ahead market using machine learning. *Appl. Energy* **2022**, *313*, 118752. [\[CrossRef\]](#)
29. Zamudio López, M.; Zareipour, H. Modeling the duration of electricity price spikes using survival analysis. *Energies* **2025**, *18*, 5255. [\[CrossRef\]](#)
30. Khan, S.; Aslam, S.; Mustafa, I. Short-term electricity price forecasting by employing ensemble empirical mode decomposition and extreme learning machine. *Forecasting* **2021**, *3*, 28. [\[CrossRef\]](#)
31. Chen, X.; Liu, Y.; Fan, N.; Lei, W. An operating profit-oriented medium-term planning method for renewable-integrated cascaded hydropower. *Energy* **2025**, *338*, 138686. [\[CrossRef\]](#)
32. Kılıç, D.K.; Nielsen, P.; Thibbotuwawa, A. Intraday electricity price forecasting via LSTM and trading strategy for the power market: A case study of the West Denmark DK1 grid region. *Energies* **2024**, *17*, 2909. [\[CrossRef\]](#)
33. Zhang, Y.; Jia, M.; Wen, H.; Bian, Y.; Shi, Y. Toward value-oriented renewable energy forecasting: An iterative learning approach. *IEEE Trans. Smart Grid* **2025**, *16*, 1962–1974. [\[CrossRef\]](#)

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