Spot Market Electricity Price Forecast via the Combination of Transformer and Ornstein-Uhlenbeck Process

Zhouhe Zhang[1], Haochen Hua[1, *], Xingying Chen[1], Jia Shao[2], Jin Zheng[3], Bo Wang[1], Lei Gan[1], Kun Yu[1]
[1] School of Electrical and Power Engineering, Hohai University, Nanjing, 211100, China
[2] School of Mathematics, University of Birmingham, Edgbaston Birmingham, B15 2TT, UK
[3] School of Engineering Mathematics and Technology, University of Bristol, Bristol, BS8 1TW, UK
* Correspondence: huahc16@tsinghua.org.cn

Abstract—**Spot market electricity price fluctuations can expose market participants to substantial financial risks if not accurately forecasted. Traditional statistical models (e.g., ARIMA) can capture linear trends but struggle with complex nonlinear relationships, while pure neural network models (e.g., Transformer) are insufficiently sensitive to random fluctuations. To address both persistent price trends and transient volatility, this paper combines Transformer with stochastic differential equations driven by Ornstein-Uhlenbeck process. Compared to ARIMA, the proposed model achieves an average MAE reduction of approximately 45%, along with comparable improvements in RMSE and MAPE and a significant boost in R^2. Against a standalone Transformer, it also exhibits substantial performance gains across all key metrics.**

Keywords—Transformer, stochastic differential equations, electricity price forecast, spot market

I. INTRODUCTION

In recent years, European countries have been deeply troubled by severe spot market electricity price fluctuations [1]. On December 11, 2024, Germany's spot market electricity price surged to a historic €936/MWh, marking the highest level recorded since the onset of the energy crisis [2]. However, just one month later, a significant surge in renewable energy generation caused Germany to experience four consecutive hours of negative electricity prices [2]. Such extreme phenomena are not unique to Germany, countries like the United Kingdom, France, and Spain also recorded unprecedented durations of negative prices in 2024. The drivers of this intense volatility lie in the dual pressures of renewable energy variability and supply-demand imbalances [3]. Extreme weather often leads to sharp reductions in wind and solar power generation, leading to supply shortages during periods of high demand. Conversely, favourable weather conditions can produce an oversupply of renewable energy which exceeding the grid's capacity to absorb it, triggering frequent occurrences of negative prices. These supply-demand imbalances, compounded by power plants' operational inflexibility, which limits their ability to quickly adjust output, destabilize energy markets. This instability causes electricity prices to fluctuate dramatically between spikes and plunges, posing significant challenges to the stability of energy systems and the operation of socioeconomic frameworks [4].

In such a highly volatile market, electricity price forecasting becomes critically important as a key tool to address price instability, balance interests, and maintain market stability. For power plants, accurate price forecasts enable the optimization of generation plans, avoiding resource wastage and improving profitability. For instance, power plants can increase output during periods of anticipated high prices to maximize revenues or scale back production during expected low-price periods to minimize losses [5]. For electricity consumers, price forecasting helps reduce costs by facilitating proactive energy storage or production adjustments, allowing them to avoid unnecessary expenses during peak price periods [6]. Furthermore, for electricity market operators, accurate forecasting aids in the early identification of market risks, optimizing dispatch strategies, and adjusting market rules to mitigate the impact of extreme price events [7]. Accurate electricity price forecasting not only enhances the economic efficiency of individual market participants but also supports the stability and resilience of the broader energy market [8],[9].

To help market participants better understand the dynamic changes in electricity prices, researchers worldwide have developed mathematical models based on historical electricity price data and various influencing factors to describe the temporal variation of electricity prices. These models can be broadly classified into data-driven electricity price models [10]-[19] and stochastic differential equation (SDE)-based models [20]-[23]. Data-driven models can be further divided into statistical models [10],[11], machine learning models [12]-[14], and deep learning models [15]-[19].

Statistical models typically rely on autoregressive assumptions for electricity price forecasting. Common approaches include Autoregressive (AR), Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA), Vector Autoregression (VAR), and Multiple Linear Regression (MLR). Statistical models offer advantages such as simplicity in computation and low parameter requirements. However, they demand high-quality data with specific distributional properties [10]. Furthermore, these models primarily consider the recursive relationships in price time series, often neglecting external factors such as supply and demand, which limits their predictive accuracy. This limitation becomes especially pronounced when handling high-frequency or nonlinear data, as traditional statistical models struggle to capture complex dynamics [11].

Electricity price data inherently exhibit nonlinear characteristics [12]. To address the limitations of traditional statistical methods, researchers have developed machine learning models, such as Artificial Neural Networks (ANNs), Gradient Boosting (GBoost), and Random Forests (RF). These models leverage mappings from input features to output predictions to effectively capture nonlinear relationships in electricity price data. For instance, reference [13] utilizes XGBoost to perform real-time automated price forecasting for wholesale electricity prices. Despite their strong generalization capabilities and robustness, these models are relatively simple in structure and face challenges in extracting deep, latent features of electricity price data [14].

With the advancements in computational technologies, deep learning models have become a research focus [15], [16]. These models include Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNNs), Deep Neural Networks (DNNs), and Gated Recurrent Units (GRUs). Compared to traditional machine learning approaches, deep learning models exhibit superior data learning and pattern extraction capabilities through the stacking of multiple nonlinear layers [17], [18]. Recently, models based on the Transformer architecture have gained attention due to their self-attention mechanism, which can directly compute the correlation between any two points in a time series. This feature makes them particularly suitable for handling long-term dependencies in data, such as seasonal variations and market patterns. For example, [19] proposed a new Transformer-based model to assist the short-term trading strategies of market players. While these models have demonstrated superior performance in electricity price forecasting, they still face challenges in modeling short-term stochasticity.

Complementing machine learning and deep learning approaches, stochastic differential equation (SDE)-based models provide a mathematically rigorous framework for characterizing the intrinsic stochastic volatility of electricity prices [20]. By establishing dynamic equations for stochastic processes, these models excel in capturing short-term price randomness [21]. However, they face limitations in modeling long-term price trends, requiring frequent parameter updates to adapt to dynamic changes over time [22], [23]. Striking a balance between stochastic description and long-term trend modeling remains a critical research challenge.

Given the strengths and weaknesses of the aforementioned models, existing single-model approaches struggle to balance long-term pattern extraction and short-term fluctuation modeling in electricity price forecasting. To address this issue, this paper innovatively proposes a hybrid electricity price model that integrates Transformer and SDE. Specifically, the Transformer component captures the overall trend of electricity prices, while the SDE component models short-term stochastic volatility. Compared to using Transformer alone, the addition of SDE enables the dynamic simulation of rapid, random fluctuations. Conversely, integrating Transformer with SDE overcomes the inefficiency of modeling countless short intervals solely with SDE. The proposed hybrid model is validated using real data from the Nord Pool electricity market. Simulation results demonstrate that the model performs well in both trend prediction and stochastic volatility modeling.

The paper is structured as follows. Section II describes the construction process of the electricity price forecasting model. Section III presents the simulation results validating the performance of the proposed model. Finally, Section IV provides the conclusions.

II. METHOD

A. *Transformer*

Transformer, introduced by Vaswani et al. in the 2017 paper “Attention Is All You Need” [24], is a deep learning model that replaces traditional recurrent structures (e.g., RNNs and LSTMs) with self-attention mechanisms and parallel computation. It has demonstrated exceptional performance in natural language processing and time series analysis due to its ability to capture global dependencies and perform efficient parallel computations. The model consists of several key components: input embedding with positional encoding for temporal representation, self-attention mechanisms for capturing long-range dependencies, multi-head attention for extracting diverse features across subspaces, and feedforward neural networks (FFNs) for processing complex patterns. Residual connections and layer normalization stabilize training, while the encoder-decoder structure allows the model to extract global features and generate sequential predictions.

In electricity price forecasting, the Transformer leverages its advanced architecture to improve prediction accuracy. Time series data is preprocessed through segmentation, normalization, and the addition of positional encodings. Multi-dimensional features are mapped into vector spaces, and the encoder captures dependencies and patterns, while the decoder refines multi-step predictions. The model’s optimization uses mean squared error as the loss function, with the Adam optimizer and learning rate scheduling enhancing efficiency. This approach enables the Transformer to effectively model both short-term fluctuations and long-term trends, providing a robust framework for electricity price forecasting tasks.

B. *Electricity price prediction based on Transformer*

In this paper, temperature and humidity are identified as key factors influencing spot market electricity prices and are incorporated into a two-dimensional input feature set for the Transformer-based electricity price forecasting model. The model workflow involves several steps: the dataset is first split into training (70%), validation (20%), and testing (10%) sets using sequential partitioning, with input features normalized to [0,1] range. Initial model parameters are specifically configured with 4 attention heads, 3 encoder/decoder blocks, 32-dimensional embeddings, 64 hidden units in feed-forward layers, 0.01 dropout rate, and 10 historical time steps for single-step prediction (T=1). The model is trained with Adam optimizer using 0.001 learning rate, 32 batch size, and 50 maximum epochs. Training the Transformer model using these specified parameters involves iterative optimization with early validation monitoring. Training continues iteratively until either the accuracy is sufficient or the epoch limit is reached. Finally, the trained model is used to predict electricity prices based on the test dataset, effectively integrating temperature and humidity as critical inputs to improve forecasting precision through multivariate temporal modeling.

The hyperparameters of the Transformer model were selected through a combination of empirical validation and computational efficiency considerations. A grid search (2–8 heads) revealed that 4 heads provided the best trade-off between model complexity and performance, balancing multi-scale feature extraction without overfitting. Three blocks were chosen after testing deeper architectures (up to 6 blocks). Deeper models showed marginal performance gains but significantly increased training time. Embedding`s dimensions were optimized via cross-validation on the validation set. Smaller embeddings (e.g., 16D) underfit, while larger ones (e.g., 64D) led to overfitting.

C. Electricity price forecasting based on Transformer and Ornstein-Uhlenberg processes

To account for the short-term stochastic fluctuations of electricity prices while simultaneously modeling both their long-term trends and short-term random variability, the specific model is defined as follows:

$$x(t)=x_T(t)+x_S(t). \tag{1}$$

where $x(t)$ represents the simulated electricity price at time t, $x_T(t)$ is the overall trend component of the electricity price, predicted using the Transformer-based model introduced in Section *B*, $x_S(t)$ represents the random fluctuation component, capturing the deviation of the actual electricity price from the predicted value at time t, modeled using the Ornstein–Uhlenbeck(OU) process.

In the field of stochastic analysis, there are various forms of SDEs, such as those driven by Wiener processes or Lévy processes. In [25], the OU process was used to model the output power of PVs. Similarly, this paper adopts the OU process for the stochastic component $x_S(t)$. The expression for $x_S(t)$ is as follows [25]:

$$dx_S(t)=k_S(\mu_S-x_S(t))dt+\sigma_S dW(t). \tag{2}$$

where k_S is the mean-reverting level of electricity prices, μ_S is the mean-reversion rate, σ_S is the volatility of electricity prices, and $W(t)$ is the standard Brownian motion. The OU process models a mean-reverting behavior. When $x_S(t)$ deviates above μ_S, the next state value $x_S(t+\Delta t)$ tends to decrease. Conversely, when $x_S(t)$ is below μ_S, $x_S(t+\Delta t)$ tends to increase. Thus, any deviation from the mean is gradually pulled back toward.

The parameters μ_S, k_S, σ_S in (2) can be estimated from historical electricity price data using the maximum likelihood estimation (MLE) method [21], [25]. Assume there is a sequence of N random fluctuation components $x_S=\left\{x_S^1,x_S^2,\ldots,x_S^{N-1},x_S^N\right\}$, with a time step Δt between two consecutive components. The discrete form of (2) is given by:

$$\Delta x_S=x_S^t-x_S^{t-1}=\mu_S k_S\Delta t+(1-k_S\Delta t)x_S^t+\sigma_S x_S^t(W(t+\Delta t)-W(t)). \tag{3}$$

where $W(t+\Delta t)-W(t)$ follows a standard normal distribution [26]. Based on this, let $P(x_S^t|x_S^{t-1},\mu_S,k_S,\sigma_S)$ denote the transition probability density function from $t-1$, conditioned on the parameters μ_S, k_S, σ_S. This transition probability follows a normal distribution with mean $x_S^{t-1}+k_S(\mu_S-x_S^{t-1})\Delta t$. Thus, the transition probability density function is expressed as:

$$P\left(x_S^t|x_S^{t-1},\mu_S,k_S,\sigma_S\right)=\frac{1}{\sqrt{2\pi\sigma_S^2\Delta t}}e^{\frac{-\left(\Delta x_S-k_S\left(\mu_S-x_S^{t-1}\right)\Delta t\right)^2}{2\sigma_S^2\Delta t}}. \tag{4}$$

From (4), the log-likelihood function is obtained as:

$$ln\left(L\left(\mu_S,k_S,\sigma_S\right)\right)=ln\left(\prod_{t=1}^N P\left(x_S^t|x_S^{t-1},\mu_S,k_S,\sigma_S\right)\right)=-\frac{N}{2}ln\,(2\pi\Delta t)-N\,ln\,\sigma_E-\frac{1}{2\sigma_S^2 h}\sum_{t=1}^N\left(x_S^t-k_S\mu_S\Delta t+(k_S\Delta t-1)x_S^{t-1}\right)^2. \tag{5}$$

To maximize the likelihood function, the parameters μ_S, k_S,σ_S are determined by solving for the values that maximize the log-likelihood function in (5). By taking partial derivatives of (5) with respect to μ_S, k_S,σ_S, the system of (6) is obtained, and solving it yields the implicit expressions for the parameters in (7):

$$\begin{cases}\frac{\partial\, ln\left(L\left(\mu_S,\sigma_S,k_S\right)\right)}{\partial\mu_S}=0\\ \frac{\partial\, ln\left(L\left(\mu_S,\sigma_S,k_S\right)\right)}{\partial\sigma_S}=0\\ \frac{\partial\, ln\left(L\left(\mu_S,\sigma_S,k_S\right)\right)}{\partial k_S}=0\end{cases}. \tag{6}$$

$$\begin{cases}\sigma_S^2=\frac{1}{N\Delta t}\sum_{t=1}^N\left(x_S^t-k_S\mu_S\Delta t+(k_S\Delta t-1)x_S^{t-1}\right)^2\\ k_S=-\frac{1}{\Delta t}\frac{\sum_{t=1}^N\left(x_S^t-x_S^{t-1}\right)\left(x_S^{t-1}-\mu_S\right)}{\sum_{t=1}^N\left(x_S^{t-1}-\mu_S\right)^2}\\ \mu_S=\frac{1}{k_S N\Delta t}\sum_{t=1}^N\left(x_S^t+(k_S\Delta t-1)x_S^{t-1}\right)\end{cases}. \tag{7}$$

D. Model Simulation Error Evaluation Metrics

To evaluate the accuracy of model predictions against actual data, several commonly used metrics are employed, including the mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and the coefficient of determination (R^2).

MAE quantifies the average absolute difference between predicted and observed values, representing the model's overall deviation. It is calculated using the following equation:

$$E_{MAE}=\frac{1}{n}\sum_{i=1}^{n}|\hat{y}(i)\text{-}y(i)|. \quad (8)$$

where E_{MAE} is the mean absolute error, n is the number of samples in the test set, $\hat{y}(i)$ represents the i-th simulated data, $y(i)$ represents the i-th actual data.

RMSE evaluates the standard deviation of prediction errors by computing the square root of the mean squared differences between predictions and observations. Its equation is:

$$E_{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}(i)\text{-}y(i))^2}. \quad (9)$$

where E_{RMSE} is the root mean square error.

MAPE assesses the relative error between predictions and actual values as a percentage, providing a normalized measure of accuracy. The equation for MAPE is as follows:

$$E_{MAPE}=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y(i)\text{-}\hat{y}(i)}{y(i)}\right| \times 100\%. \quad (10)$$

where E_{MAPE} is the mean absolute percentage error.

R^2 quantifies the proportion of variance in observed data that is explained by the model. Unlike the error-based metrics, R^2 values range from 0 to 1, with higher values indicating better model fit. The equation is:

$$E_{R^2}=1\text{-}\frac{\sum_{i=1}^{n}(y(i)\text{-}\hat{y}(i))^2}{\sum_{i=1}^{n}(y(i)\text{-}\bar{y})^2}. \quad (11)$$

where E_{R^2} is the coefficient of determination, $\bar{y}$ represents the mean value of the actual data.

III. Result

The year 2024 exhibited unprecedented price spikes and prolonged negative prices in European markets, providing a robust testbed for evaluating model performance under extreme conditions. For this paper, Växjö, in the Swedish province of Kronoberg, within the jurisdiction of Nord Pool, is selected as the research focus. The time span of this study covers a full year, specifically from January 1, 2024, to December 31, 2024. The spot market electricity price daily data used in this paper is sourced from Nord Pool [27], a leading power market operator in Europe that facilitates electricity trading across multiple regions. Fig. 1 illustrates the spot electricity price data for Sweden during 2024, obtained directly from Nord Pool.

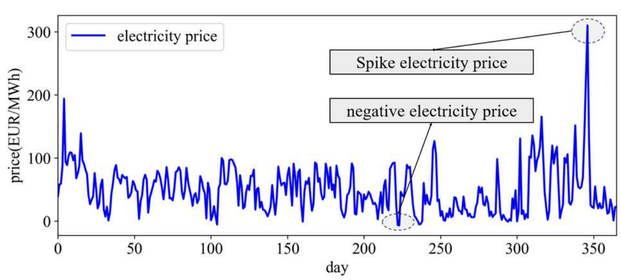

Fig. 1. Spot electricity price for one year in Växjö.

The temperature and humidity data used for the simulation in this paper were obtained from the World Weather Online [28]. Fig. 2 illustrates the temperature variation in Växjö over this period, while Fig. 3 depicts the corresponding humidity trends. These data serve as crucial input.

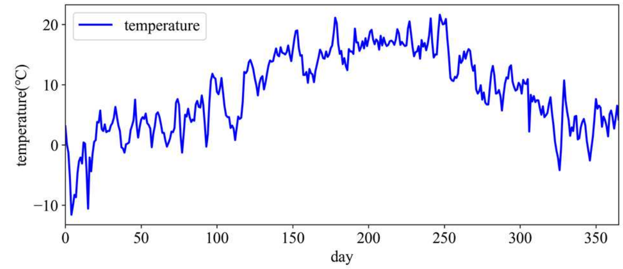

Fig. 2. Temperatures from January 1, 2024 to December 31, 2024 in Växjö.

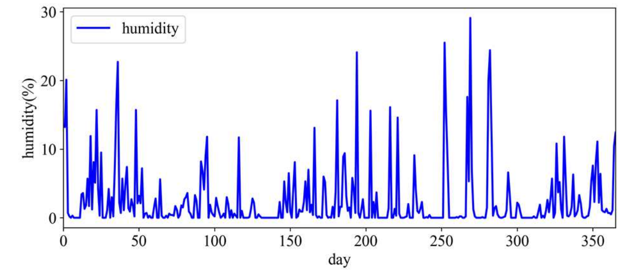

Fig. 3. Humidity from January 1, 2024, to December 31, 2024 in Växjö.

Fig. 4 illustrates the electricity price forecasting results obtained using Transformer-based model, while Fig. 5 presents the forecasting results derived from the proposed Hybrid-Driven model, which integrates Transformer with the Ornstein-Uhlenbeck process. According to the spot electricity price data in Växjö, the parameters obtained via MLE (i.e., (4)-(7)) for dynamic electricity price model (i.e., (2)) are k_S=0.021 , μ_S=-2.15, σ_S=34.72. In Fig. 4 and Fig. 5, the actual electricity price data are depicted as a solid black line, the most probable price curves calculated by the proposed methods are shown as a solid red line, and the gray shaded regions represent the most probable price ranges obtained through 200 simulations of proposed methods. It can be observed that the majority of the actual electricity price curves fall within the gray shaded regions, thereby validating the effectiveness of the electricity price forecasting models proposed in this paper.

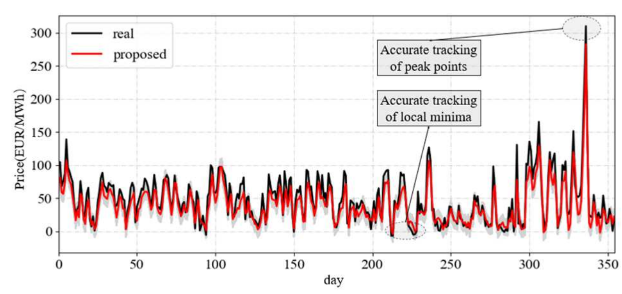

Fig. 4. Electricity price prediction model based on Transformer

As shown in Fig. 4, the Transformer-based model can generally capture the overall fluctuation patterns of electricity prices but exhibits relatively low accuracy in tracking random fluctuations, particularly in predicting the extreme maximum and minimum values of electricity prices. In contrast, Fig. 5 demonstrates that the electricity price forecasting model proposed in this paper not only reflects the overall fluctuation trends but also accurately captures random fluctuations, achieving higher accuracy in tracking sharp fluctuations in electricity prices.

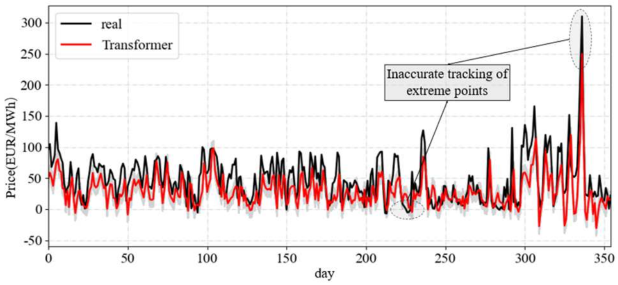

Fig. 5. Electricity price prediction model based on Hybrid-Driven

To further validate the feasibility of the electricity price forecasting model proposed in this paper, another simulation method, ARIMA, is employed to predict electricity prices, and the results are compared with those of the proposed method. The stationarity of the data was validated through the Augmented Dickey-Fuller unit root test. The test statistic value of |-9.496| significantly exceeds the critical value of |-2.571| at the 10% significance level. Based on this result, we reject the null hypothesis of the presence of a unit root in the series with 90% confidence, concluding that the first-order differenced time series exhibits stationarity and satisfies the fundamental requirements for ARIMA modeling. Further model identification was conducted using autocorrelation function and partial autocorrelation function plots. The ACF plot displays a distinct truncation characteristic after lag 2, while the PACF plot shows clear truncation after lag 1. By applying the Bayesian Information Criterion minimization principle, the optimal model structure was determined as ARIMA(1,1,2), with an autoregressive order of 1, a differencing order of 1, and a moving average order of 2.

Fig. 6 illustrates the comparison between the electricity prices calculated by the proposed model, the actual prices, and the prices predicted by the ARIMA model. Compared with the ARIMA model, the proposed method significantly improves prediction accuracy.

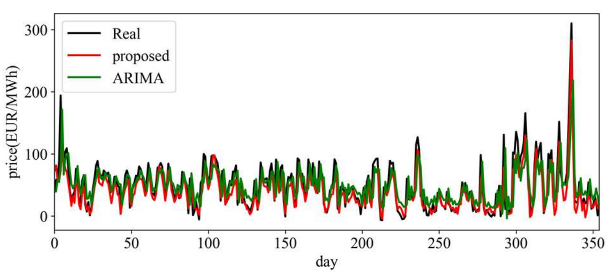

Fig. 6. Comparison of Forecasting Performance Between Hybrid-Driven Electricity Price Model and the ARIMA Model

Specifically, compared to the forecasting results of the ARIMA model, the proposed model achieved a reduction in MAE by 41.60% to 48.85%, RMSE by 47.64% to 48.97%, and MAPE by 50.60% to 66.27%, while improving R^2 by 83.31% to 111.49%. Furthermore, when compared to the standalone Transformer model, the proposed model demonstrated reductions in MAE by 21.52% to 34.39%, RMSE by 12.57% to 18.42%, and MAPE by 44.15% to 50.91%, along with an increase in R^2 by 44.70% to 58.56%.

TABLE. I. Comparison of Evaluation Metrics for Simulation Errors Across Different Models

Method	Case	MAE	RMSE	MAPE	R^2
ARIMA	a week	20.8391	27.7323	2.7099	0.4349
	a month	21.3654	27.5312	2.4392	0.4501
	half a year	20.8818	29.1133	2.2943	0.4061
	one year	20.5803	28.5278	2.1279	0.4234
Transformer	a week	15.5053	17.3725	1.8535	0.5801
	a month	16.6576	16.4872	1.6697	0.5949
	half a year	16.2202	16.4872	1.9396	0.6037
	one year	15.7562	17.1752	1.7373	0.5639
Proposed (Transformer+SDE)	a week	12.1683	14.1733	0.9139	0.9198
	a month	10.9280	14.4149	0.9325	0.8251
	half a year	10.8670	14.8560	1.1737	0.8736
	one year	11.5034	14.8065	1.0501	0.8409

Although incorporating the stochastic differential equation to capture random fluctuations adds some computational overhead, the parallel processing capabilities of the Transformer component ensure that overall training and prediction times remain comparable to those of a standalone Transformer model. Based on our current experimental results, the hybrid model exhibits both high forecasting accuracy and computational efficiency, theoretically making it suitable for real-time electricity market prediction. However, further model optimization and deployment testing are needed to fully realize its potential in a real-time setting.

IV. Conclusion

This paper proposes a hybrid electricity price forecasting model by integrating Transformer with the Ornstein-Uhlenbeck process. The proposed model leverages the Transformer to predict the overall trend of electricity prices, while employing a stochastic differential equation driven by the Ornstein-Uhlenbeck process to capture short-term price fluctuations. A comparative analysis was conducted between the proposed model, a standalone Transformer-based electricity price forecasting model, and an ARIMA-based forecasting model. Compared to ARIMA, the proposed model achieves average reductions of 45% in MAE, 48% in RMSE, and 58% in MAPE, and an increase of 97% in R^2. Against a standalone Transformer, it reduces MAE by 28%, RMSE by 15%, and MAPE by 48%, while boosting R^2 by 52%, thereby validating its robust forecasting capabilities over weekly, monthly, and yearly horizons.

Future research will focus on three key areas to enhance the model's practicality and universality. First, we aim to explore Lévy-driven stochastic differential equations to better capture heavy-tailed price distributions and extreme price spikes, thereby improving the modeling of rare but impactful events. Second, the model will be validated across diverse electricity markets, such as those in the Asia-Pacific (APAC) and North America, to assess its generalizability under varying market mechanisms and volatility patterns. Third, integrating numerical weather prediction (NWP) data as exogenous inputs will refine the model's responsiveness to abrupt renewable energy fluctuations by dynamically adjusting key meteorological variables like temperature and humidity.

REFERENCES

[1] F. Moraglio and C. S. Ragusa, “Day-Ahead Electricity Price Forecasting in the Contemporary Italian Market,” in IEEE Access, vol. 12, pp. 72062-72078, 2024.

[2] “Energy Markets for Europe.” https://www.europex.org/.

[3] S. Xu, J. Shen, H. Hua, F. Li, K. Yu, Z. Li, X. Gao, and X. Dong, “Trading portfolio strategy optimization via mean-variance model considering multiple energy derivatives,” Processes, vol. 11, art. no. 532, 2023.

[4] C. Wu, X. Chen, H. Hua, K. Yu, L. Gan, and B. Wang, “Optimal energy sharing for renewable portfolio standard and carbon cap-and-trade scheme: A two-step energy trading approach,” International Journal of Electrical Power & Energy Systems, vol. 162, 110334, Nov. 2024.

[5] M. Zhang, Y. Xu and H. Sun, “Optimal Coordinated Operation for a Distribution Network With Virtual Power Plants Considering Load Shaping,” in IEEE Transactions on Sustainable Energy, vol. 14, no. 1, pp. 550-562, Jan. 2023.

[6] D. Liu, Z. Qin, H. Hua, Y. Ding, and J. Cao, “Incremental incentive mechanism design for diversified consumers in demand response,” Applied Energy, vol. 329, Jan. 2023, 120240.

[7] H. Hua, C. Du, X. Chen, H. Kong, K. Li, Z. Liu, P. Naidoo, M. Lv, N. Hu, M. Fu, and B. Li, “Optimal dispatch of multiple interconnected-integrated energy systems considering multi-energy interaction and aggregated demand response for multiple stakeholders,” Applied Energy, vol. 376, Part A, 124256, Dec. 2024.

[8] C. Zhang, Y. Fu and L. Gong, "Short-Term Electricity Price Forecast Using Frequency Analysis and Price Spikes Oversampling," in IEEE Transactions on Power Systems, vol. 38, no. 5, pp. 4739-4751, Sept. 2023

[9] H. Hua, X. Wu, X. Chen, H. Kong, Y. Sun, Q. Yang, M. Tavares, and P. Naidoo, “Carbon reduction oriented regional integrated energy system optimization via cloud-edge cooperative framework,” CSEE Journal of Power and Energy Systems, in press.

[10] J. Wang, Y. Yu, B. Zeng and H. Lu, “Hybrid ultra-short-term PV power forecasting system for deterministic forecasting and uncertainty analysis,” Energy, vol. 288, pp. 129898, Feb. 2024.

[11] J. Lago, F. Ridder, B. Schutter, “Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms,” Applied Energy, vol. 221, pp. 386-405, 2018.

[12] Z. Qin, H. Hua, H. Liang, R. Herzallah, Y. Zhou and J. Cao, “Optimal Electricity Trading Strategy for a Household Microgrid,” 2020 IEEE 16th International Conference on Control & Automation (ICCA), Singapore, pp. 1308-1313, Oct. 2020.

[13] W. Sai, Z. Pan, S. Liu, Z. Jiao, Z. Zhong, B. Miao and S. Chan, “Event-driven forecasting of wholesale electricity price and frequency regulation price using machine learning algorithms,” Applied Energy, vol. 352, pp. 121989, 2023.

[14] H. Wu, Y. Liang, X. Gao, J. Heng, “Bionic-inspired oil price prediction: Auditory multi-feature collaboration network,” Expert Systems with Applications, vol. 244, pp. 122971, 2024.

[15] H. Hua, Z. Wei, Y. Qin, T. Wang, L. Li, and J. Cao, “Review of distributed control and optimization in energy Internet: From traditional methods to artificial intelligence-based methods,” IET Cyber-Physical Systems: Theory & Applications, vol. 6, no. 2, pp. 63-79, 2021.

[16] J. Lago, F. Ridder, B. Schutter, “Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms,” Applied Energy, vol. 221, pp. 386-405, 2018.

[17] J. Bedi and D. Toshniwal, “Deep learning framework to forecast electricity demand,” Applied energy, vol. 238, pp. 1312-1326, 2019.

[18] H. Hua, Y. Qin, C. Hao, and J. Cao, “Optimal energy management strategies for energy Internet via deep reinforcement learning approach,” Applied Energy, vol. 239, pp. 598-609, Apr. 2019.

[19] J. Bottieau, Y. Wang, Z. De Grève, F. Vallée and J. -F. Toubeau, “Interpretable Transformer Model for Capturing Regime Switching Effects of Real-Time Electricity Prices,” in IEEE Transactions on Power Systems, vol. 38, no. 3, pp. 2162-2176, May 2023

[20] H. Hua, J. Cao, G. Yang, and G. Ren, “Voltage control for uncertain stochastic nonlinear system with application to energy Internet: non-fragile robust H_∞ approach,” Journal of Mathematical Analysis and Applications, vol. 463, no. 1, pp. 93-110, Jul. 2018.

[21] Y. Qin, H. Hua, and J. Cao, “Stochastic optimal control scheme for battery lifetime extension in islanded microgrid via a novel modeling approach,” IEEE Transactions on Smart Grid, vol. 10, no. 4, pp. 4467-4475, Jul. 2019.

[22] B. Gashi and H. Hua, “Optimal regulators for a class of nonlinear stochastic systems,” International Journal of Control, vol. 96, no. 1, pp. 136-146, 2023.

[23] H. Hua, B. Gashi, and Moyu Zhang, “Robust risk-sensitive control,” International Journal of Robust and Nonlinear Control, vol. 33, no. 10, pp. 1-26, Jul. 2023.

[24] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. Gomez, Ł. Kaiser, I. Polosukhin, “Attention is all you need,” Advances in Neural Information Processing Systems, 2017.

[25] H. Hua, Y. Qin, C. Hao, and J. Cao, “Stochastic optimal control for energy Internet: A bottom-up energy management approach,” IEEE Transactions on Industrial Informatics, vol. 15, no. 3, pp. 1788-1797, Mar. 2019.

[26] Yu. K, Li. F, Chen. X, Hua. H, Yin. M, Yang. Q, Y. Jiang, J. Shao and Naidoo. P, “Mean-variance Trading Portfolio Selection for A Class of Energy Retailers,” CSEE Journal of Power and Energy Systems , in press.

[27] “Nord Pool.” https://www.nordpoolgroup.com/en/.

[28] “World Weather Information Service.” https://worldweather.wmo.int/zh/home.html.