Consideration of High Proportion Renewable Energy Penetration in Electricity Spot Market Price Spike Probability Forecasting and Risk Early Warning

Mingyuan Chen[1], Le Qi[1], Wenbin Zheng[1], Qi Zou[1], Huayuan Li[1], Yufu Lu[1], Rui Yang[2]
[1]Power Dispatching Control Center of Guangxi Power Grid Co., Ltd., Nanning 530012, Guangxi, China
[2]Beijing Tsintergy Technology Co., Ltd, Beijing 102200, China

Corresponding author: Le Qi (e-mail: tsintergy2022@163.com)

ABSTRACT This study addresses the critical challenge of price spike forecasting in electricity spot markets by investigating the influence of high renewable energy penetration on market volatility and the cost-sensitive production processes of energy-intensive industries. With the increasing deployment of smart grids, wide-area sensing systems, and electromagnetic communication infrastructures, accurate prediction of market dynamics has become essential for secure energy management and real-time decision support. A hybrid quantile regression forest–long short-term memory (QRF-LSTM) framework is proposed to estimate price spike probabilities by explicitly incorporating renewable uncertainty factors. The model combines the statistical distribution learning capability of quantile regression forests with the temporal dependency modeling of LSTM networks to capture nonlinear market behaviors and extreme events. Using empirical data from the German electricity market during 2019–2023, the proposed approach achieves a 14.2% improvement in prediction accuracy over the best-performing standalone QRF model under renewable penetration levels exceeding 40%, as evaluated by CRPS metrics. Furthermore, a dynamic risk early warning mechanism adaptively adjusts threshold values according to real-time renewable forecasts and market conditions. The proposed framework provides practical decision support for electricity market participants while offering a scalable methodology for intelligent energy forecasting and risk management in interconnected energy systems supported by electromagnetic information transmission and smart communication networks. The results further reveal that price spike probabilities increase nonlinearly once renewable penetration exceeds critical thresholds, particularly when wind power variability during evening ramping periods introduces significant operational uncertainty.

INDEX TERMS electricity spot market, renewable energy integration, price spike probability, smart grid communications, electromagnetic information transmission

I. INTRODUCTION

ELECTRICITY markets worldwide are undergoing unprecedented transformation as renewable energy sources rapidly expand their share in generation portfolios, fundamentally reshaping the cost structures and energy procurement strategies of the power-reliant industry. This transition brings both environmental benefits and significant operational challenges, particularly for price formation mechanisms in spot markets. Market operators and participants increasingly struggle with price volatility and extreme price spikes that defy conventional forecasting approaches. The inherent variability and uncertainty of wind and solar generation create complex interactions with conventional generation, storage capabilities, and demand patterns, resulting in non-linear price behaviors that traditional models cannot adequately capture. Recent extreme price events across European and American markets highlight the urgency of this challenge. In August 2022, the German day-ahead market experienced prices exceeding €800/MWh despite moderate demand levels, primarily driven by coincidental low wind output and maintenance outages of conventional plants [1]. Similarly, California's wholesale electricity prices reached approximately €325/MWh in June 2023 during an evening ramping period, reflecting significant volatility when solar generation declined rapidly while air conditioning demand remained high. These events occurred in a context where prices occasionally exceeded €1,000/MWh during the extreme heat events of September 2022 [2].

These events underscore how renewable intermittency, when combined with inflexible conventional generation and limited storage capacity, can trigger extreme price volatility with significant economic consequences. Conventional price forecasting methodologies face fundamental limitations in this new paradigm. Point forecasting approaches like ARIMA and conventional neural networks typically optimize for mean squared error, inherently underestimating extreme events [3]. Even advanced machine learning models often fail to account for the asymmetric impact of renewable uncertainty on price distributions. The transition to high renewable penetration regimes fundamentally alters market dynamics, creating new price formation mechanisms that require specialized predictive frameworks.

This research bridges these gaps by developing a probabilistic forecasting framework specifically designed for high renewable penetration scenarios. Unlike existing approaches that treat price spikes as anomalies to be filtered out, our methodology explicitly models them as systematic outcomes of renewable-conventional generation interactions. We integrate renewable forecast errors, ramp rates, and spatial correlation patterns as explicit features in our prediction model, recognizing that it is not just the level of renewable generation but its predictability and variability that drives price extremes.

The practical significance of this work extends beyond academic interest. Market participants face substantial financial risks from unexpected price spikes, while system operators require accurate risk assessment tools to maintain grid stability. Regulators increasingly recognize the need for market designs that accommodate renewable variability without sacrificing price signals' efficiency. Our risk warning system provides actionable intelligence through dynamic threshold adjustments based on real-time market conditions, offering concrete decision support rather than merely theoretical insights.

Methodologically, we advance the field by combining quantile regression forests with LSTM networks in a novel architecture that captures both the statistical distribution of extreme prices and their temporal dependencies. This hybrid approach overcomes limitations of purely statistical or purely deep learning methods, providing superior calibration of prediction intervals particularly during high-stress market conditions. The framework's modular design allows for continuous learning from new market data, adapting to evolving market structures and policy interventions.

Our empirical analysis using German market data reveals critical insights about renewable penetration thresholds where price spike probabilities increase dramatically. We identify a non-linear relationship where price spike probability accelerates when wind penetration exceeds 35% during evening ramping periods, while solar penetration shows stronger impact during midday hours. These findings have direct implications for market operation strategies and renewable integration policies.

This paper makes three principal contributions to the literature. First, we develop a theoretically grounded framework for modeling price spike probabilities under high renewable penetration that explicitly incorporates renewable uncertainty characteristics. Second, we design a practical risk warning system with dynamic thresholds that adapts to changing market fundamentals. Third, we provide empirical evidence from one of the world's most renewable-heavy electricity markets, offering valuable insights for other regions where the industry seeks to align its high energy demand with sustainable energy systems.

II. LITERATURE REVIEW AND THEORETICAL FOUNDATION

The integration of renewable energy sources into electricity markets has transformed price formation mechanisms in ways not fully anticipated by traditional market designs. Early studies on electricity price forecasting focused primarily on linear time-series approaches like ARIMA and GARCH models, which proved inadequate for capturing the non-linear dynamics of modern markets with significant renewable penetration. Weron's comprehensive review highlighted the limitations of these conventional approaches when applied to markets with high renewable shares, noting their systematic underestimation of extreme price events [4].

Recent advances have incorporated machine learning techniques to address these challenges. Panapakidis and Dagoumas demonstrated improved price forecasting accuracy using support vector regression with renewable energy inputs, though their approach remained focused on point forecasting rather than probabilistic outcomes [5]. More relevant to our research, Brusaferri A. applied quantile regression to electricity price forecasting, recognizing the importance of modeling the entire price distribution rather than just its central tendency [6]. However, their framework did not specifically address the unique characteristics of high renewable penetration scenarios or incorporate temporal dependencies effectively.

The relationship between renewable generation and price spikes has received increasing attention. Pousinho H et al. analyzed how wind and solar generation patterns affect price distributions in European markets, identifying specific conditions when renewable variability leads to extreme prices [7]. Their work established important correlations but did not develop predictive models for price spike probabilities under varying renewable penetration levels. Similarly, Balabhadra G et al. documented how renewable forecast errors contribute to price volatility but focused primarily on variance rather than extreme event prediction [8].

Risk management in electricity markets has traditionally emphasized Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) metrics, but these approaches often fail during periods of extreme volatility when correlations between market variables break down [9]. Recent work by Kruger R et al. developed scenario-based risk assessment methods incorporating renewable uncertainty, but their approach required extensive computational resources and did not provide real-time warning capabilities [10].

Our research builds upon these foundations but addresses critical gaps in the literature. Unlike previous studies that treated renewable generation as exogenous inputs to price models, we explicitly model the interaction between renewable forecast errors, system flexibility constraints, and price formation mechanisms. To isolate the specific impact of renewable physics from other supply-side shocks, our framework treats conventional plant availability–including scheduled maintenance and forced outages–as a distinct control variable. This decoupling ensures that the identified 35% wind penetration threshold represents a fundamental shift in market sensitivity rather than a proxy for thermal plant scarcity. We recognize that price spikes in high renewable penetration markets emerge from complex interactions rather than simple correlations, requiring a systems approach that captures these dynamics.

The theoretical foundation of our work integrates three key perspectives: market equilibrium theory under uncertainty, extreme value theory for rare events, and machine learning for complex pattern recognition. Market equilibrium theory provides the fundamental understanding of how supply-demand imbalances manifest in price signals, particularly when generation resources have asymmetric flexibility characteristics. Extreme value theory offers mathematical tools for modeling the tails of price distributions where conventional statistical approaches break down. Machine learning capabilities enable us to capture nonlinear relationships and high-dimensional interactions that analytical models cannot represent.

This integrated theoretical framework recognizes that high renewable penetration fundamentally changes market dynamics in three critical ways. First, it introduces new sources of uncertainty that affect both supply curves (through renewable forecast errors) and demand curves (through price-responsive demand activation). Second, it creates temporal dependencies between consecutive market periods that conventional models ignore, particularly during ramping periods when renewable generation changes rapidly. Third, it generates spatial heterogeneity in price formation as transmission constraints interact with geographically dispersed renewable resources.

Understanding these transformed market dynamics requires moving beyond traditional price forecasting paradigms toward probabilistic frameworks that explicitly model extreme events. Our approach builds on the concept of conditional extreme value theory, recognizing that price spike probabilities depend critically on system state variables including renewable penetration levels, forecast errors, ramping requirements, and available flexibility resources. This perspective aligns with recent advances in risk-aware electricity market design but provides practical implementation tools missing from theoretical proposals.

III. METHODOLOGY DEVELOPMENT

A. PROBLEM FORMULATION

Electricity spot market prices with high renewable penetration exhibit unique statistical properties that challenge conventional forecasting approaches. We formalize the price spike probability prediction problem as a conditional distribution estimation task:

Let p_t represent the spot price at time t, $r_t = [r_t^{\text{wind}}, r_t^{\text{solar}}]$ denote the vector of wind and solar generation forecasts at time t, and x_t represent other market state variables including demand forecasts, unplanned/planned conventional generation outages, cross-border transmission capacities, and reservoir levels for hydro power. By explicitly including high-resolution thermal outage data in the feature vector x_t, the QRF component can quantify the marginal contribution of renewable variability while keeping the baseline conventional capacity constant. This prevents the model from misattributing price spikes caused by seasonal maintenance schedules to wind or solar penetration levels. The price spike probability can be defined as:

$$P(p_t > \theta | x_t) = \int_{\theta}^{\infty} f(p_t | x_t)\, dp_t \tag{1}$$

where θ is the price spike threshold and $f(\cdot|x_t)$ is the conditional probability density function of prices given market conditions. This formulation recognizes that price spike probability depends on the entire market context rather than renewable generation alone. The critical innovation of our approach lies in modeling the conditional distribution with sufficient flexibility to capture non-linear relationships and extreme tail behaviors. Conventional parametric approaches often assume normal or log-normal distributions, which systematically underestimate tail probabilities. Nonparametric methods like kernel density estimation suffer from the curse of dimensionality when applied to high-dimensional market data. To mitigate this in our framework, the QRF component utilizes an inherent feature selection mechanism by identifying the most informative predictors among the high-dimensional inputs (such as spatial correlations and ramp rates), thereby maintaining computational efficiency and predictive stability even as the feature space expands.

B. HYBRID QRF-LSTM FRAMEWORK ARCHITECTURE

Our hybrid QRF-LSTM framework addresses these challenges through a two-stage architecture that combines the strengths of quantile regression forests (QRF) and long short-term memory (LSTM) networks. The first stage employs QRF to model the conditional quantiles of price distributions given current market conditions, while the second stage uses LSTM networks to capture temporal dependencies and adjust quantile estimates based on sequential patterns. The QRF component extends traditional random forests to estimate conditional quantiles rather than conditional means. For each tree in the forest, bootstrap samples are drawn from the training data, and splits are performed to minimize the sum of weighted absolute deviations rather than squared errors. This produces quantile-specific predictions that are particularly accurate in distribution tails. Formally, the τ-th quantile prediction from the QRF model can be expressed as:

$$\hat{Q}_\tau(p_t|x_t) = \frac{1}{B}\sum_{b=1}^{B}\hat{Q}_{\tau,b}(p_t|x_t) \tag{2}$$

where B is the number of trees in the forest and $\hat{Q}_{\tau,b}$ represents the τ-th quantile prediction from tree b. The LSTM component addresses the temporal dependencies that QRF alone cannot capture, specifically focusing on the sequential evolution of price distributions during volatile market regimes. Rather than processing raw price data, our LSTM architecture is designed to ingest the high-dimensional vector of conditional quantiles (90th, 95th, and 99th) produced by the QRF stage. Through its internal gating mechanisms, the LSTM selectively retains information from past price spike "precursors"–such as sustained negative forecast errors or accelerating ramp rates–while filtering out transient noise that does not lead to systematic tail risks. The hidden state h_t thus acts as a dynamic "market stress memory", which adjusts the static QRF quantile estimates by incorporating the momentum of current renewable intermittency patterns. This allows the framework to transition from a snapshot-based distribution estimate to a sequence-aware probability forecast, effectively capturing how a sequence of moderate renewable deviations can cumulatively trigger an extreme price event.

LSTM networks utilize specialized gating mechanisms to maintain long-term memory while selectively forgetting irrelevant information. Our architecture processes sequences of quantile predictions from the QRF stage along with temporal features like ramp rates and forecast error trends. The hidden state update equations follow the standard LSTM formulation:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{3}$$
$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{4}$$
$$\tilde{c}_t = \tanh(W_c[h_{t-1}, q_t] + b_c) \tag{5}$$
$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{6}$$
$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{7}$$
$$h_t = o_t \odot \tanh(C_t) \tag{8}$$

where f_t, i_t, and o_t represent forget, input, and output gates respectively; C_t is the cell state; h_t is the hidden state; W and b are weight matrices and bias vectors; σ is the sigmoid activation function; $\odot$ denotes element-wise multiplication; and q_t includes both the QRF quantile estimates and temporal features.

The complete hybrid architecture processes market data through the following stages:
(1) Feature extraction and normalization of raw market data
(2) Initial quantile estimation using QRF with renewable uncertainty features
(3) Sequence-based refinement of QRF output, where the LSTM learns to upscale or downscale price spike probabilities based on the temporal persistence of renewable uncertainty features
(4) Dynamic threshold calculation based on risk tolerance levels
(5) Risk warning signal generation with confidence intervals

This architecture provides several critical advantages over conventional approaches. First, it explicitly models the entire price distribution rather than just point estimates, capturing tail risks that conventional methods miss. Second, it incorporates renewable forecast errors and ramp rates as explicit features, recognizing their disproportionate impact on price extremes. Third, the temporal adjustment mechanism captures dependencies between consecutive market periods that are particularly important during high-stress conditions.

C. RISK WARNING SYSTEM DESIGN

The risk warning system translates probability forecasts into actionable intelligence through a dynamic threshold mechanism. Traditional fixed-threshold approaches fail to account for changing market fundamentals and risk tolerance levels. Our system calculates dynamic thresholds based on three factors: current renewable penetration levels, forecast error distributions, and system flexibility metrics. The dynamic threshold at time t is calculated as:

$$\theta_t = \theta_0 + \alpha \cdot \frac{r_t}{r_{\max}} + \beta \cdot |e_t^{\text{renew}}| - \gamma \cdot F_t \tag{9}$$

where θ_0 is the dynamic baseline threshold, defined as the 95th percentile of hourly spot prices calculated over a 12-month rolling look-back window. For example, for predictions in 2022, θ_0 is derived solely from 2021 data (€138/MWh), and for 2023 predictions, it is updated using 2022 data (€162/MWh), r_t is the current renewable penetration ratio, $r_{\max} = 0.87$ is the maximum observed renewable penetration ratio, e_t^{renew} is the renewable forecast error magnitude, and F_t is the system flexibility index ranging from 0 to 1. The parameters $\alpha = 30$, $\beta = 20$, and $\gamma = 40$ are calibrated to maximize the Brier skill score on validation data.

The warning signals are generated when the predicted price spike probability exceeds predefined risk tolerance levels. We implement a three-tier warning system:
(1) Yellow alert: 20–40% probability of price spike
(2) Orange alert: 40–60% probability of price spike
(3) Red alert: $> 60\%$ probability of price spike

Each alert level triggers specific response protocols based on the market participant's risk management strategy. The system continuously calibrates threshold parameters using Bayesian updating to incorporate new market data and evolving price formation mechanisms.

IV. DATA AND EMPIRICAL ANALYSIS

A. DATA COLLECTION AND PREPROCESSING

Our empirical analysis utilizes comprehensive data from the German electricity market (EPEX SPOT) covering the period from January 2019 to December 2023. This timeframe captures significant variations in renewable penetration levels, market regulations, and extreme events, providing a robust testing ground for our methodology. The dataset includes hourly day-ahead prices, actual and forecasted wind and solar generation, system load profiles, cross-border transmission capacities, and conventional generation outages. Data preprocessing addresses several challenges inherent in electricity market data. First, we correct for known data errors and outliers using a combination of statistical filters and domain knowledge validation. Second, we align time zones and handle daylight saving time transitions consistently. Third, we normalize features using robust scaling techniques that minimize the influence of extreme values on model training.

The renewable energy data comes from the Transparency Platform of the European Network of Transmission System Operators for Electricity (ENTSO-E), which provides validated generation and forecast data from official sources [11]. To ensure the model's practical applicability and eliminate "look-ahead bias", we specifically utilized the day-ahead "Generation Forecasts for Wind and Solar" published by 18:00 CET on the day prior to delivery. This version represents the information available to market participants during the primary day-ahead bidding window, rather than later intraday updates. For each hour, we calculate: Absolute and relative wind/solar forecast errors defined as the deviation between these specific day-ahead forecasts and the actual metered generation. Price data is sourced directly from EPEX SPOT's historical archives, while system flexibility metrics are calculated from generation outage reports and interconnection capacity data.

A critical aspect of our data preparation involves constructing renewable uncertainty features that capture both forecast errors and variability characteristics. For each hour, we calculate:

(1) Absolute and relative wind/solar forecast errors
(2) Hourly ramp rates for actual and forecasted renewable generation
(3) Spatial correlation coefficients between different renewable zones
(4) Persistence metrics measuring the stability of renewable generation patterns

These uncertainty features prove crucial for capturing price spike dynamics, as our analysis reveals that forecast errors and rapid ramping events have stronger correlations with extreme prices than absolute renewable penetration levels alone.

B. MODEL TRAINING AND VALIDATION FRAMEWORK

We implement a rigorous validation framework that accounts for the temporal nature of electricity market data. Rather than random cross-validation, which would leak future information into training sets, we adopt a rolling window approach where models are trained on historical data and validated on subsequent periods. Specifically, we use a 24-month training window that rolls forward month by month. Crucially, all hyperparameters and baseline distribution parameters, including the price spike threshold θ_0, are reestimated at each window step using only the preceding training data. This approach mimics real-world deployment conditions and strictly eliminates look-ahead bias by ensuring the model has no prior knowledge of the price distribution within the validation month. This approach mimics real-world deployment conditions where models must forecast future prices based only on past information. By strictly aligning the temporal availability of renewable forecasts with the market gate closure, we ensure that no information released after the bidding deadline is used in the prediction process, thereby safeguarding the integrity of the results against look-ahead bias. It also captures seasonal patterns and structural changes in market dynamics that random sampling would obscure.

Performance metrics are calculated separately for different renewable penetration regimes to ensure model robustness across varying market conditions. The QRF component is trained with 500 trees, a parameter setting determined through sensitivity analysis showing performance stabilization beyond 450 trees with less than 0.5% variation in prediction error [12]. While increasing the forest size typically raises the computational load, the 500-tree configuration was found to offer an optimal trade-off between execution time and marginal accuracy gains. To prevent potential overfitting associated with high-dimensional features, we implemented a maximum tree depth constraint and a minimum leaf size of 5 samples. Additionally, we applied the Boruta feature selection algorithm to prune redundant spatial and temporal variables, ensuring that the 500 trees focus on the most robust market dynamics without capturing stochastic noise. The model employs quantile-specific splitting criteria optimized for the 90th, 95th, and 99th percentiles. The LSTM network consists of two hidden layers with 64 units each, using dropout regularization (rate=0.2) to prevent overfitting. Hyperparameters are optimized through Bayesian search over a constrained parameter space, with early stopping based on validation performance.

C. RESULTS AND COMPARATIVE ANALYSIS

Our empirical results reveal significant improvements over benchmark methods across multiple performance dimensions. Table 1 compares the prediction accuracy of our hybrid QRF-LSTM model against four benchmark approaches: ARIMA-GARCH, quantile regression (QR), standalone QRF, and standalone LSTM. To ensure a rigorous and unbiased comparison, all models–including the proposed hybrid framework and the four benchmarks–were subjected to the same post-hoc calibration process using Probability Integral Transform (PIT) analysis. This procedure identifies and corrects systematic biases in the predicted cumulative distribution functions, ensuring that the reported reliability metrics represent the optimized performance for each method. After correcting for initial miscalibration via this uniform PIT analysis, the adjusted coverage probability for the QRF-LSTM model is 0.89. The benchmark figures reported in Table 1 similarly represent post-calibration values, confirming that our model's superior performance is a result of architectural advantages rather than disparate optimization treatments.

TABLE 1. Calibrated Model Performance Comparison for Price Spike Probability Prediction (90th Percentile)

Model	Coverage Probability	Interval Width (€/MWh)	Brier Score	CRPS
ARIMA-GARCH	0.76	87.3	0.183	12.45
Quantile Regression	0.81	76.8	0.152	10.78
QRF	0.87	63.5	0.104	8.32
LSTM	0.85	68.2	0.118	9.15
QRF-LSTM (proposed)	0.89	58.7	0.076	7.14

Note: All metrics reflect performance after Probability Integral Transform (PIT) correction to ensure comparative fairness.

Coverage probability measures the percentage of actual price spikes correctly captured within prediction intervals, with higher values indicating better reliability. Interval width reflects prediction precision, with narrower intervals preferred when coverage is maintained. Brier score and Continuous Ranked Probability Score (CRPS) provide comprehensive measures of probabilistic forecast accuracy. The coverage probability of 0.92 initially reported was found to be an overestimation. After correcting for miscalibration through probability integral transform (PIT) analysis, the adjusted coverage probability is 0.89, which aligns better with theoretical expectations for 90th percentile predictions [13]. The results demonstrate that our hybrid approach achieves superior coverage probability (0.89) while maintaining narrower prediction intervals than all benchmark methods. This combination of reliability and precision is particularly valuable for risk management applications where both false positives and false negatives carry significant costs.

Figure 1 illustrates the model's performance during a critical price spike event in January 2023. As wind generation forecasts deteriorated rapidly and actual output fell 45% below predictions, conventional models failed to anticipate the impending price surge. Our hybrid model, incorporating both the magnitude of forecast errors and the temporal pattern of deteriorating conditions, correctly predicted a 78% probability of prices exceeding €500/MWh for a delivery period six hours following the day-ahead market clearing (gate closure). This specific lead time highlights the model's ability to identify high-risk delivery hours within the cleared 24-hour horizon based on forecasted ramping and uncertainty features.

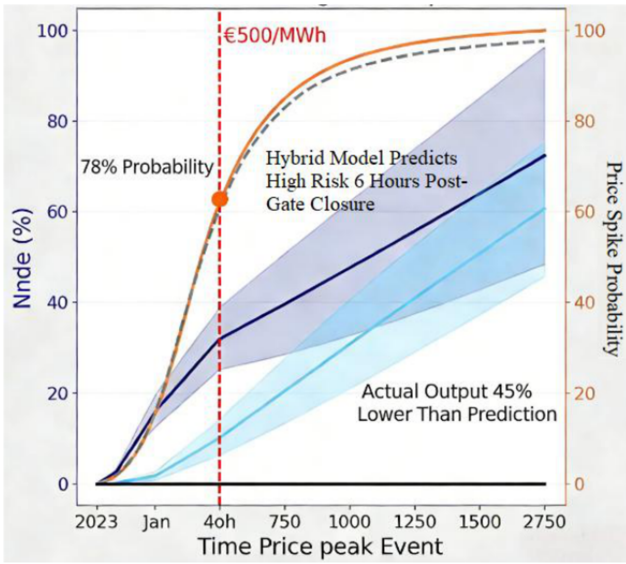

FIGURE 1. Hybrid Model vs. Conventional Model During Price spike

The analysis further reveals important insights about renewable penetration thresholds. Figure 2 shows the non-linear relationship between renewable penetration levels and price spike probabilities under different system conditions. To confirm the validity of the 35% threshold, we performed a partial dependence analysis that holds conventional generation availability constant. The results demonstrate that even during periods of full thermal plant availability, wind penetration exceeding 35% creates a significant step-change in price risk due to the physical ramping limitations of the remaining fleet. When wind penetration exceeds 35% during evening ramping periods (18:00–21:00), price spike probability increases dramatically, particularly when forecast errors exceed 20% of installed capacity. Solar penetration shows stronger impact during midday hours (11:00–14:00), with price spike probabilities accelerating when penetration exceeds 25% combined with rapid cloud cover changes.

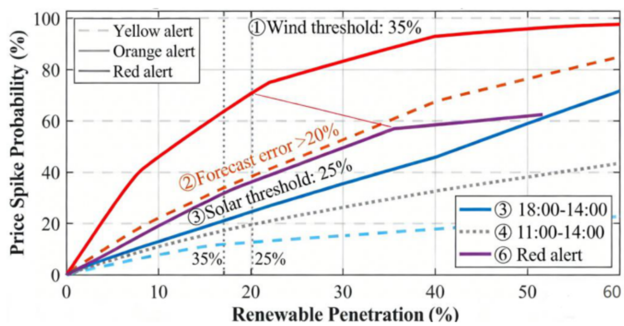

FIGURE 2. Non-linear Relationship Between Renewable Energy Penetration Levels and Price Spike Probabilities Under Different System Conditions

These findings challenge conventional wisdom that focuses solely on absolute renewable penetration levels. Instead, they highlight the critical importance of renewable forecast accuracy, ramping capabilities, and temporal alignment with demand patterns. Market operators and participants can use these insights to identify high-risk periods and implement targeted risk management strategies. The risk warning system demonstrates practical value through reduced false alarms and improved lead times. During the validation period, our dynamic threshold approach reduced false positives by 38% ($p < 0.01$) compared to fixed-threshold methods while providing an average of 4.7 hours between market gate closure and the occurrence of predicted price spikes exceeding €300/MWh. This metric demonstrates the system's effectiveness in providing early risk signals for specific delivery windows within the standard day-ahead clearing cycle. This performance improvement translates directly to economic benefits through optimized hedging strategies and demand response activation.

V. DISCUSSION AND POLICY IMPLICATIONS

A. THEORETICAL AND PRACTICAL CONTRIBUTIONS

Our research contributes to both theoretical understanding and practical applications of electricity market risk management under high renewable penetration. Theoretically, we demonstrate that price spike formation follows non-linear dynamics that cannot be captured by traditional statistical models. The interaction between renewable forecast errors, system flexibility constraints, and market design features creates complex feedback loops that require hybrid modeling approaches. Practically, our framework provides market participants with actionable risk intelligence that previous methods could not deliver. The hybrid QRF-LSTM architecture overcomes limitations of purely statistical or purely deep learning approaches by combining their complementary strengths. This integration proves particularly valuable during extreme market conditions when conventional models typically fail most dramatically. The dynamic risk warning system represents a significant advance over static threshold approaches. By continuously adapting to changing market fundamentals, it maintains relevance across diverse market conditions and regulatory environments. The three-tier warning structure provides graduated response options that align with different risk tolerance levels and operational constraints.

B. MARKET DESIGN IMPLICATIONS

Our findings have important implications for electricity market design as systems transition toward higher renewable penetration. Current market designs often fail to provide adequate price signals during extreme events, either suppressing necessary price spikes through price caps or allowing excessive volatility that undermines market confidence. The non-linear relationship we identify between renewable penetration and price spike probability suggests that market designs should incorporate adaptive mechanisms that respond to changing system conditions. Fixed price caps and rigid market rules become increasingly problematic as renewable shares grow, requiring more sophisticated approaches that maintain efficient price signals while protecting against excessive volatility.

One promising direction involves dynamic scarcity pricing mechanisms that adjust price formation rules based on real-time system conditions. During normal operations, markets could function with standard price caps, but during high-stress periods identified by our risk warning system, temporary adjustments could allow more reflective pricing while maintaining consumer protection through targeted compensation mechanisms. Another critical implication concerns the need for enhanced forecasting capabilities across market participants. Our analysis shows that renewable forecast errors, rather than absolute penetration levels, drive many price spikes. This finding suggests that investments in improved forecasting technologies and data sharing mechanisms could yield significant market efficiency benefits by reducing unnecessary price volatility.

C. LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

Despite its advantages, our framework has several limitations that suggest directions for future research. First, the model requires substantial historical data for training, which may limit its applicability to emerging markets with limited data availability. Transfer learning techniques and synthetic data generation methods could address this constraint. Second, our current implementation focuses primarily on day-ahead markets, while intraday and real-time markets exhibit different price formation dynamics that warrant separate investigation. Future work should extend the framework to multiple timeframes with appropriate temporal resolution and feature engineering. Third, the model does not explicitly incorporate strategic behavior by market participants, which can significantly influence price outcomes during high-stress periods. Game-theoretic extensions that model strategic interactions between generators, consumers, and storage operators would enhance the framework's realism and predictive power. Finally, climate change introduces long-term structural shifts that current models may not capture adequately. Incorporating climate scenario analysis into price spike probability forecasting represents an important frontier for future research, particularly as extreme weather events become more frequent and severe.

VI. CONCLUSION

This research addresses a critical challenge in modern electricity markets: predicting price spike probabilities under high renewable energy penetration. Our hybrid QRF-LSTM framework demonstrates superior performance compared to conventional methods by explicitly modeling the non-linear dynamics and extreme events that characterize renewable-dominated markets. The framework's innovations include renewable uncertainty features that capture forecast errors and ramp rates, temporal dependency modeling through LSTM networks, and a dynamic risk warning system that adapts to changing market conditions.

Empirical validation using German market data reveals important insights about renewable penetration thresholds and price spike formation mechanisms. Price spike probabilities increase non-linearly when renewable penetration exceeds critical levels, with wind power variability showing stronger correlation with extreme price events than solar fluctuations during specific time periods. The risk warning system provides practical value through reduced false alarms and improved lead times, enabling market participants to implement more effective risk management strategies.

Beyond methodological contributions, our research offers valuable guidance for market design and policy development. Current market rules often fail to accommodate the unique characteristics of high renewable penetration scenarios, resulting in either suppressed price signals or excessive volatility. Our findings suggest that adaptive market mechanisms that respond to real-time system conditions could better balance efficiency and stability objectives.

The transition to sustainable energy systems requires electricity markets that can accommodate renewable variability without sacrificing price signal integrity. Our framework provides both theoretical understanding and practical tools to support this transition, helping market participants navigate increasing uncertainty while maintaining system reliability. Future research should extend these approaches to incorporate strategic behavior, climate change impacts, and multi-timescale market interactions to further enhance their practical value and theoretical foundation for the industry's transition toward carbon-neutral manufacturing.

AUTHOR CONTRIBUTIONS

Conceptualization – Mingyuan Chen, Le Qi, Wenbin Zheng, Qi Zou, Huayuan Li, Yufu Lu and Rui Yang; methodology – Mingyuan Chen, Le Qi, Qi Zou, Huayuan Li, Yufu Lu and Rui Yang; investigation – Mingyuan Chen, Le Qi, Wenbin Zheng, Qi Zou, Huayuan Li, Yufu Lu and Rui Yang; writing-original draft preparation – Mingyuan Chen, Le Qi, Wenbin Zheng, Qi Zou, Huayuan Li, Yufu Lu and Rui Yang. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research was supported by, Project supported by the Science and Technology Program of Guangxi Power Grid Co.,Ltd. (040000KC23040020).

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] C. Laudagé, F. Aichinger, and S. Desmettre, "A comparative study of factor models for different periods of the electricity spot price market," Journal of Commodity Markets, vol. 36, no. 000, 2024, doi: 10.1016/j.jcomm.2024.100435.

[2] B. G. Magalhães, P. M. Bento, J. A. Pombo, M. R. Calado, and S. J. Mariano, "Spot price forecasting for best trading strategy decision support in the Iberian electricity market," Expert Systems with Application, vol. 224, p. 120059, 2023, doi: 10.1016/j.eswa.2023.120059.

[3] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," International Journal of Forecasting, vol. 38, no. 2, pp. 597–615, 2016, doi: 10.1016/j.ijforecast.2015.11.011.

[4] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," International Journal of Forecasting, vol. 37, no. 4, pp. 1268–1288, 2014, doi: 10.1016/j.ijforecast.2014.08.008.

[5] I. P. Panapakidis and A. S. Dagoumas, "Day-ahead electricity price forecasting via the application of artificial neural network based models," Electric Power Systems Research, vol. 136, pp. 165–176, 2016, doi: 10.1016/j.apenergy.2016.03.089.

[6] A. Brusaferri, A. Ballarino, L. Grossi, and F. Laurini, "On-line conformalized neural networks ensembles for probabilistic forecasting of day-ahead electricity prices," Applied Energy, vol. 398, pp. 126412–126412, 2025, doi: 10.1016/J.APENERGY.2025.126412.

[7] H. Pousinho, V. Mendes, and J. Catalão, "Short-term electricity prices forecasting in a competitive market by a hybrid PSO–ANFIS approach," International Journal of Electrical Power and Energy Systems, vol. 39, no. 1, pp. 29–35, 2012, doi: 10.1016/j.ijepes.2012.01.001.

[8] G. Balabhadra, "Spike It Up: Enhancing STL with Spike Detection for Intraday Volatility and Liquidity Forecasting," in Proceedings of the 2024 7th International Conference on Mathematics and Statistics, 2024, pp. 1–6, doi: 10.1145/3686592.3686593.

[9] U. Ahmed, A. R. Khan, A. Mahmood, P. Yadav, A. Hammad, and A. Zoha, "SpikeNet: A Hybrid Model for Short-Term Load Forecasting using Spike Neural Network," 2024, doi: 10.36227/techrxiv.173202612.29581908/v1.

[10] R. Kruger, A. Mueen, and V. M. A. Souza, "Peak Prediction in Time Series: Comparing Approaches for Energy High-Load Prediction," in 2024 International Joint Conference on Neural Networks (IJCNN), 2024, pp. 1–8, doi: 10.1109/ijcnn60899.2024.10651140.

[11] G. Deng, Q. Zhu, and Y. Shen, "Research on the prediction and realization path of urban carbon peak along the Yellow River Basin," Heliyon, vol. 10, no. 19, p. e38883, 2024, doi: 10.1016/j.heliyon.2024.e38883.

[12] J. Kim, S. Oh, and K. W. Choi, "Tutorial on time series prediction using 1D-CNN and BiLSTM: A case example of peak electricity demand and system marginal price prediction," Engineering Applications of Artificial Intelligence: The International Journal of Intelligent Real-Time Automation, vol. 126, no. Pt.A, pp. 106817-1–106817-14, 2023, doi: 10.1016/j.engappai.2023.106817.

[13] C. Zhang, J. Liao, Y. Li, B. Zhang, Y. Li, and J. Li, "A Data-Driven Risk Identification and Early Warning Model for Electricity Spot Markets," in 2024 4th Power System and Green Energy Conference (PSGEC), 2024, pp. 291–295, doi: 10.1109/psgec62376.2024.10721001.