ISSN: 3049-7159

Research Article

Journal of Business and Econometrics Studies

A Novel Approach to Electricity Price Forecasting in Market Economics: Three-stage Machine Learning Ensemble Utilizing Feed Forward Neural Network, XGBoost, and LightGBM

Lake Zhou

Phillips Exeter Academy, 2 Dolloff Farm Drive, Exeter, 03833, NH, United States

Corresponding author
Lake Zhou, Phillips Exeter Academy, 2 Dolloff Farm Drive, Exeter, 03833, NH, United States.

Received: November 21, 2025; **Accepted:** November 25, 2025; **Published:** December 02, 2025

ABSTRACT
Electricity Price Forecasting (EPF) is a critical and complex task, necessary for many electrical market participants. This paper proposes a novel approach to EPF utilizing a three-stage machine learning model, FXL3, including base learners, a stacking model, and a final residual correcting model. Using the EPF Toolbox Python library, we benchmark our approach against ten state-of-the-art EPF models, showing a significant improvement in accuracy. Then, to evaluate the efficiency of the proposed model, we conduct a series of variation experiments. Furthermore, in conjunction with the pro- posed model, this paper also introduces a novel and comprehensive dataset that includes five exogenous factors: hourly temperature, wind speed, natural gas price, shift factor, and location. Finally, using experimental results, this paper demonstrates the effectiveness of the three-stage approach and the high correlation between Shift Factor and electrical price, which has been commonly neglected in EPF research.

Keywords: Electricity Price Forecasting, EPF, Shift Factor, Machine Learning

Introduction

Electricity prices can have ripple effects across the broader economy, and having a reliable and transparent electricity market is vital to our modern economic infrastructure. To ensure market efficiency, market participants must be able to accurately forecast electricity prices, enabling them to make better-informed decisions regarding load scheduling, power purchasing, dispatch strategies, and demand response programs [1].

However, accurate Electricity Price Forecasting (EPF) is difficult due to a multitude of factors. Firstly, there is a wide scope of interrelated factors that can impact electrical prices, including weather, natural gas prices, transmission topology, and load [2]. Additionally, many of these factors themselves are inherently hard to predict, such as weather, resulting in noisy and inaccurate data. Furthermore, in many cases, prices exhibit non-linear relationships [3].

To address these challenges, this paper proposes a novel three-stage approach towards modeling electrical prices, including two base models, a stacking stage, and a final residual boosting stage. First, utilizing a feedforward neural network model and a tree-based gradient-boosted XGBoost model to make base predictions, then combining them with a tree-based machine learning model, LightGBM. Finally, another XGBoost model is trained to predict the residuals of this ensemble, producing the final predictions. This robust model can capture nonlinear dynamics well and effectively handle noisy data.

Novelty

This paper's novelty comes from two factors:

- A robust machine learning ensemble that proves to be more accurate than other state-of-the-art EPF models
- A novel dataset that factors in network topology via Shift Factor

Through a series of experiments, this paper will benchmark the proposed model against other state-of-the-art EPF models to evaluate its comparative precision. Additionally, this paper will

Citation: Lake Zhou. A Novel Approach to Electricity Price Forecasting in Market Economics: Three-stage Machine Learning Ensemble Utilizing Feed Forward Neural Network, XGBoost, and LightGBM. J Bus Econo Stud. 2025. 2(6): 1-7. DOI: doi.org/10.61440/JBES.2025.v2.99

show how Shift Factor, a way to quantify electrical topography, is a key input.

In previous works, other papers have recognized the importance of electrical topography but have rarely been able to accurately represent it; this paper uses an innovative approach to measure it, utilizing shift factor. Using DAZER's shift factor calculator, we can derive the shift factor based on present electrical constraints [4-6].

Related Work

In the past, there have been many attempts at EPF, each different and with its unique value. In this section, the paper provides an overview of previous work aimed at achieving accurate EPF.

When looking at previous model architectures, two main approaches present themselves: Machine learning and statistical modeling. In a study, Zhang et al. presented the use of an Integrated Long-term Recurrent Convolutional Network (ILRCN) to capture temporal and spatial dependencies in price data [7]. Whereas, Lago et al. utilized deep learning methods (LSTM, RNN) and traditional machine learning algorithms like XGBoost and Light- GBM, concluding that ensemble methods were superior in EPF [8]. On the other side of things, statistical models, such as ARIMA and VAR, have historically been employed, though they often fail to capture complex non- linear relationships due to their simplicity [9]. Furthermore, there have also been hybrid models that combine statistical baselines with advanced machine learning techniques that have emerged [10].

Prior studies have concluded that a variety of factors can influence electricity prices. Commonly used inputs include historical price data, load demand forecasts, weather variables, fuel prices, and economic indicators [10,11]. In addition, some studies also include market-related factors such as bidding strategies and auction results in forecasting models [12].

However, while the advancement of new models in the field of EPF has been plentiful, proper benchmarking and comparison of these models have not been as prominently established [13]. So, this paper will not only evaluate the proposed model, but also benchmark it against ten other state-of-the-art EPF models.

Methodology

This section will discuss the proposed methodology, which includes a novel dataset and model architecture.

Data Processing

Data Description

The novel dataset includes five inputs: Temperature, Wind Speed, Load, Shift Factor, and Location. With these five inputs, this dataset can comprehensively account for the factors influencing electrical prices. Temperature, Wind Speed, and Load encompass demand-side drivers while Shift Factor and Location cover physical supply-side constraints. To obtain this data, a variety of sources were used.

Table 1: Data Inputs and Their Sources

Data Input	Data Source
Electrical Price History	EROCT Public Information
Historical and Forecast Shift Factor	DAZER Calculations ICE
Historical and Forecast Natural Gas	Public Data WSI Trader
Historical and Forecast Temperature	WSI Trader
Historical and Forecast Wind Speeds	EROCT Public Information
Load History	EROCT Forecast
Load Forecast	

The dataset consists of data for all major hubs and load zones in the EROCT electrical market from 1/1/2024 to 6/25/2025. It is also important to note that this is not a time-series dataset and does not include time as an input. This dataset encourages models to predict prices purely from exogenous variables that fundamentally influence prices.

Data Preprocessing

To achieve better results with model training, a series of data preprocessing steps are applied. Firstly, data entries with prices higher than $200 MWh are eliminated from the datasets, as prices higher than $200 MWh are extremely rare. Looking at Figure 1 below, we can see how greater than $200 is uncommon, with most prices falling around the median of $22. This trimming does present a range limitation of this model; however, without this filter, machine learning models may have a hard time recognizing trends in the already challenging data.

Following this, more outlier detection is performed using a two-pronged approach, using z-score and filtering with an interquartile range technique. The dataset is filtered to include only points with a Z-score, defined as $z = \frac{X-\mu}{\sigma}$ of 1.2. On top of the Z-score, IQR, defined below, is applied. These outlier handling techniques are used to better focus on the noisy data and help the model concentrate on the real signals.

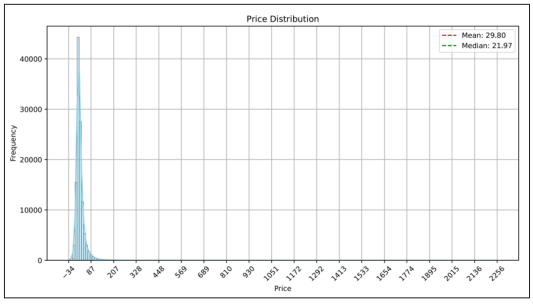

Figure 1: Prices Frequency in Dataset.

$$IQR = Q_3 - Q_1 \quad (1)$$

Lower Bound = $Q_1 - 1.5 \times IQR$ Upper Bound = $Q_3 + 1.5 \times IQR$

Outliers : x < Lower Bound or x > Upper Bound

When handling categorical features like location, they are converted into vectors. Each unique location is converted into a multidimensional vector, in this case, a five-dimensional vector.

Adjusting the number of dimensions can help to emphasize or de-emphasize the uniqueness between differing lo- cations.

Finally, the resulting dataset is divided into a training and a testing sub- set, following an 80-20 split between training and testing. The partitioning is random, ensuring an unbiased model.

Model Architecture

The proposed model architecture consists of 3 stages: base models, stack- ing, and residual boosting (See Figure 2 below).

Base Models

The FNN is one of the core base models alongside a tree-based XGBoost model. Constructed using PyTorch, it takes two streams of input: a set of numerical features as well as a dense vector that represents categorical features such as location.

Post embedding, inputs are then passed through a sequence of 24 deep layers, each containing 512 fully-connected neurons, with non-linear activation functions applied at every layer. In order to preserve negative values and avoid the "Dying ReLU" problem, which is common in deep neural networks such as this one, this model employs LeakyReLU as its activation function.

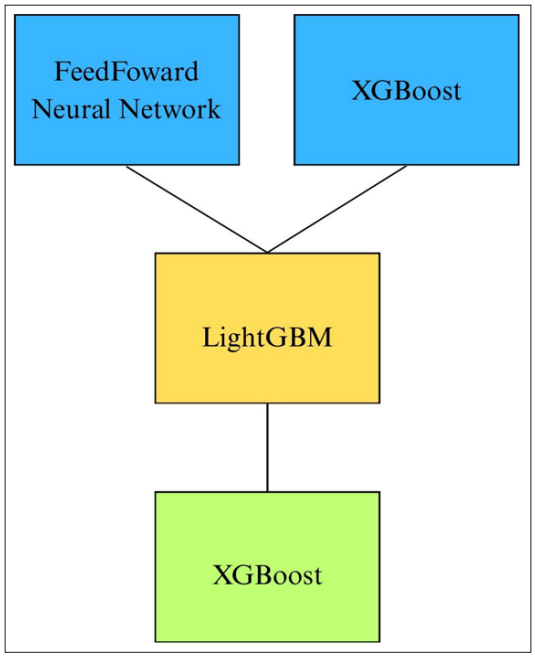

Figure 2: Proposed Model Structure.

$$f(x) = \begin{cases} x, & \text{if } x \geq 0 \\ \alpha x, & \text{if } x < 0 \end{cases} \quad (2)$$

The loss function for the FNN is Huber loss, which is a combination of MSE and MAE. It switches between quadratic and linear using a threshold (which can be adjusted), allowing it to act like MSE on smaller values and MAE on larger values. Huber Loss offers a powerful approach towards noisy, average-focused data, as is the case in EPF. Using Huber Loss with the FNN allows the base model to capture a wider range of prices by being more lenient on outliers, helping address the average-focused issue. Huber loss is defined mathematically in equation 3.

$$L_\delta(y, f(x)) = \begin{cases} \frac{1}{2}y - f(x)^2, & \text{if } |y - f(x)| \leq \delta, \\ \delta\,|y - f(x)| - 1\,\delta^2, & \text{otherwise} \end{cases} \quad (3)$$

Running in parallel and trained on the same data as the FNN, the XG- Boost model serves as the second base model. XGBoost constructs an ensemble of trees, with each tree attempting to correct the residuals of the prior one. The model uses Taylor's leverages, both first (Gradient) and second (Hessian) derivatives, of the loss function to assess optimal split locations, prioritizing high-gain split locations.

Additionally, to minimize overfitting in noisy, high-dimensional datasets, XGBoost utilizes L1 and L2 regularization. Where L1 is defined mathemati cally as: $\alpha\sum_{j=1}^{T}|w_j|$; and L2 is defined mathematically as: $\frac{\lambda\sum_{j=1}^{T}}{2}w_j^2$ These help to penalize complex trees that may be overfitting the dataset.

Compared to the FNN, XGBoost offers another robust model that can effectively handle outliers and complex, nonlinear trends. However, instead of embedding location, location is encoded as a corresponding integer value, where each location is matched to a unique integer.

Stacking

In this second stage, results from the two base models, FNN and XG- Boost, are combined with a LightGBM model. The role of this model is to learn an optimal mapping from these base model predictions to the true target price, correcting for deviations and biases.

LightGBM uses a leaf-wise tree growth strategy, which splits from in- dividual leaves that provide the most gain. This asymmetric growth means that the model can deepen in areas most impactful, which is especially useful when aggregating, as this can be a lot of data.

On top of this, LightGBM also has a variety of methods to accelerate training and reduce memory usage, including: histogram-based binning, Gradient-based One-Side Sampling (GOSS), and Exclusive Feature Bundling (EFB).

When stacking, the LightGBM model is able to dynamically adjust the weightings and interaction of the two base predictions based on their relative reliability across different data regions. In general, LightGBM performed well as a meta-learner that effectively optimized the results of the two base models, FNN and XGBoost.

Residual Boosting

As a third step, a final XGBoost model is trained on the residuals of the LightGBM stacked model, further enhancing the predictions. In this stage, residuals are calculated as the difference between the true target values and the predictions output by the LightGBM stacking model. The residual XGBoost model then also uses the original dataset, used by the base FNN and XGBoost models, as input, ensuring the model does not solely rely on the previous model's outputs.

In terms of architecture, this XGBoost model also follows a similar structure to the base XGBoost model. By constructing

an ensemble of decision trees that are optimized to minimize residual errors of the preceding ensemble, this model works to capture subtle patterns in the residuals. In the end, the final predictions are computed by adding the predicted residuals from the XGBoost residuals model and the LightGBM stacking model.

Experiments

In this section, this paper will describe the various experiments conducted on the proposed method in order to benchmark against other state-of-the-art models, evaluate the importance of shift factor, test the novel dataset and model architecture in combination, and test the efficiency of the proposed model with a variation study.

Baseline Against EPFBench

The first experiment evaluates the proposed model with the EPF Toolbox Python library. This library includes five standard datasets from Nord Pool, PJM, BE, FR, and DE.

In order to be compatible with EPF Toolbox, some adjustments had to be made to the model. Firstly, the location embedding was changed to embed a null column as EPF Toolbox does not consider location as an input. Secondly, all the data processing steps are skipped to ensure a fair comparison. Data is pulled straight from EPF Toolbox using their data extraction module and is split into train and test datasets based on time.

Benchmarking with EPF Toolbox has largely been split between presenting metrics in base units (Dollars/MWh) and normalized units. To be thorough, this study compares both in base units and normalized units.

The models compared with in base units are NBEATSx, Deep Neural Network, LEAR, and a Transformer model [13-15]. The comparison is made by taking the reported results from each of the respective studies and comparing them against the 3-stage ensemble model.

The models compared with in normalized units are TSMixer, I Transformer, MTSCI+SemGuide, AlignTime, and Periormer [16-20]. The comparison is made by taking the reported results from each of the respective studies and comparing them against the 3-stage ensemble model.

Testing On Custom Dataset

In this experiment, we combine the proposed model architecture with the novel dataset and evaluate its performance. These results are the most conclusive as they combine the proposed model and dataset with hyperparameter optimization.

Shift Factor Importance

To evaluate the importance of the shift factor in EPF relative to other inputs, this experiment isolates one factor and tests the impact of that input by removing it from the dataset and comparing the difference in accuracy across various metrics.

Model Variations

Finally, in order to test the efficiency of the proposed model, a series of variations was constructed and tested. For each stage of this three-stage approach, an FNN, XGBoost, and LightGBM were evaluated to see which was the best performing. Each model had the optimal hyperparameters that resulted in the smallest loss. The final proposed model combines the two best base models, the best stacking model, and the best residual boosting model.

Results

Metrics

In evaluating each experiment, this paper uses four metrics to benchmark against: MAE, RMSE, MAPE, and sMAPE.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \tag{4}$$

$$RMSE = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{5}$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n} \frac{y_i - \hat{y}_i}{y_i} \tag{6}$$

$$MAE = \frac{100}{n}\sum_{i=1}^{n} \frac{| y_i - \hat{y}_i |}{\frac{|y_i|+|\hat{y}_i|}{2}} \tag{7}$$

While MAPE is known to exaggerate loss, especially when near zero, it is included as a reference. Mainly, results should be interpreted in terms of MAE, RMSE, and sMAPE.

Baseline Against EPF Toolbox

When compared to other state-of-the-art EPF models, the proposed model's performance is on par and even better than many of the other state-of-the-art EPF models across a diverse set of data.

Table 2 presents the MAE, RMSE, MAPE, and sMAPE values in base units for each model across various datasets. In each metric per dataset, the best results are in bold. Blank values mean that they were unreported.

Table 2: Performance of 5 State-of-the-Art EPF Models and FXL3 Across 5 Datasets in Base Units.

Dataset	Metric	FXL3	DNN Ensemble	LEAR Ensemble	Transformer	NBEATSx-G	NBEATSx-I
NP	MAE	1.7846	1.683	1.738	2.33	**1.58**	1.62
	RMSE	3.5662	3.319	3.362	4.08	**3.16**	3.27
	MAPE (%)	**4.45**	5.384	5.533	–	–	–
	sMAPE (%)	**7.08**	4.88	5.009	7	4.62	4.70
PJM	MAE	**1.6815**	2.862	3.013	3.67	2.91	2.90
	RMSE	**3.3447**	5.040	5.127	5.85	5.02	4.84
	MAPE (%)	**17.19**	27.478	30.134	–	–	–
	sMAPE (%)	**7.08**	11.331	11.980	14	11.54	11.61

EPEX-BE	MAE	**4.9511**	5.870	6.140	6.54	5.95	6.11
	RMSE	17.4137	15.966	15.974	**16.68**	15.76	15.80
	MAPE (%)	45.84	24.892	**20.720**	–	–	–
	sMAPE (%)	**10.6**	13.446	14.546	15	13.86	14.02
EPEX-FR	MAE	2.9732	3.866	3.980	4.91	3.81	3.79
	RMSE	15.4036	11.867	10.676	12.67	11.50	11.25
	MAPE (%)	**7.89**	13.601	14.680	–	–	–
	sMAPE (%)	**7.78**	10.812	11.566	14	10.59	10.69
EPEX-DE	MAE	**2.6456**	3.413	3.955	4.03	3.31	3.29
	RMSE	**5.3174**	5.927	7.079	6.99	5.72	5.65
	MAPE (%)	127.49	**94.334**	122.412	–	–	–
	sMAPE (%)	**10.91**	14.078	15.747	17	13.99	13.99

Table 3 presents the MAE and MSE in normalized units compared to other state-of-the-art EPF models. While some other proposed models are competitive, the FXL3 still shows to be effective.

Table 3: Performance Metrics (MSE and MAE) Across Datasets for Different Models in Normalized Units.

Model	NP MSE	NP MAE	PJM MSE	PJM MAE	BE MSE	BE MAE	FR MSE	FR MAE	DE MSE	DE MAE
TSMixer (2023)	0.187	0.253	0.078	0.174	0.316	0.241	0.396	0.214	0.250	0.322
I Transformer (2024)	0.203	0.250	0.077	0.171	0.339	0.246	0.378	0.205	0.258	0.327
MTSCI+SemGuide (2025)	0.120	0.173	0.048	0.125	0.385	0.283	0.254	0.164	0.240	0.287
Align Time (2025)	0.279	0.302	0.106	0.214	0.431	0.284	0.389	0.210	0.490	0.445
Periormer (2024)	0.625	0.528	0.314	0.422	0.575	0.369	0.702	0.439	0.767	0.581
FXL3	0.123	0.226	0.022	0.081	0.531	0.192	0.463	0.122	0.103	0.167

Custom Dataset

When testing on the proposed dataset, the model achieved the results described in Table 4.

Table 4: Final Model Evaluation Metrics

MAE	RMSE	MAPE (%)	sMAPE (%)
1.02	1.4	9.16	6.71

To better visualize what these numbers mean, the figure below shows a graph of the predicted prices versus actual prices, along with the residual plot.

From these results, we can see how, when used in combinations, the proposed model and dataset can achieve state-of-the-art accuracy. However, at the same time, we also recognize that the range restriction of 200 MWh limits the model's performance in times when the price is greater than 200 MWh.

Input Factor Importance

The results of the experiment are displayed in Table 5 below. The best performing is bold, while the worst performing is underlined. The Input column refers to the variable that was excluded. The control includes all of the variables.

These results show a clear indication that shift factor and load are the two most impactful inputs. This makes sense as Shift Factor is arguably the most important supply-side variable, while load encompasses demand.

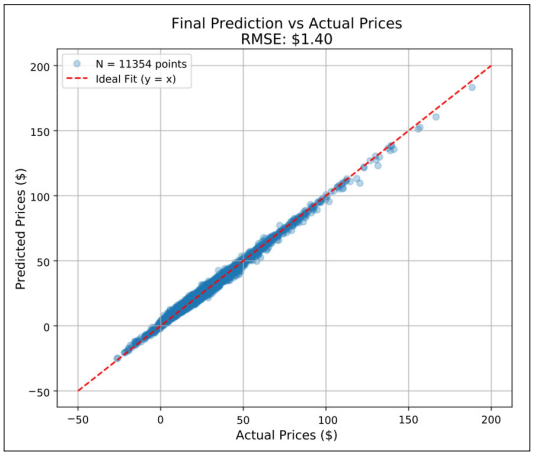

Figure 3: Plot of Predicted Prices vs Actual Prices

Table 5: Model Performance by Input Feature

Input	RMSE	MAE	MAPE (%)	sMAPE (%)
Control	**1.80**	**1.29**	15.91	**8.24**
Temp	2.66	1.91	22.56	11.43
Wind	2.64	1.92	21.63	12.72
Gas	2.57	1.84	29.20	11.32
Shift	3.09	2.20	19.61	13.12
Location	2.01	1.43	**14.69**	9.09
Load	2.90	2.06	29.88	13.19

From these results, it is more reasonable to categorize them into tiers as the Table 6 shows. However, this does not mean that lower-tier factors are not important or relevant. As the results show, the control, which includes all the factors, still outperforms all the variation cases. Additionally, to reaffirm these conclusions, analysis of the LightGBM's

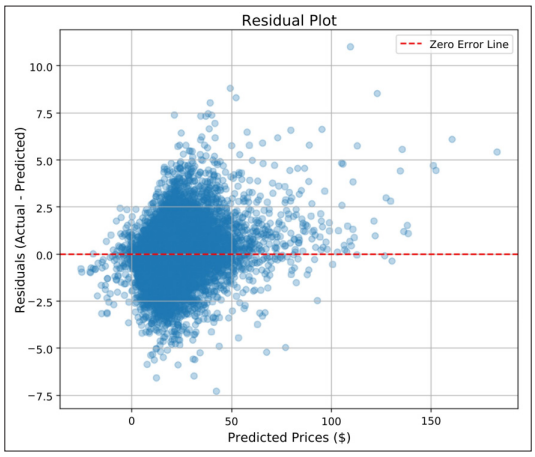

Figure 4: Predicted Prices vs Actual Prices Residual Plot

Feature Importance can reveal which factors contributed the most to increas- ing accuracy. Specifically, it sums together the total amount of gain achieved across all boosting rounds for a certain feature. These results are displayed in the figure below.

The Figure shows that LightGBM found Shift Factor was extremely im- pactful, resulting in almost 4 times more gain than the Gas. Furthermore, on a more general scale, these findings also agree with the tier list described above.

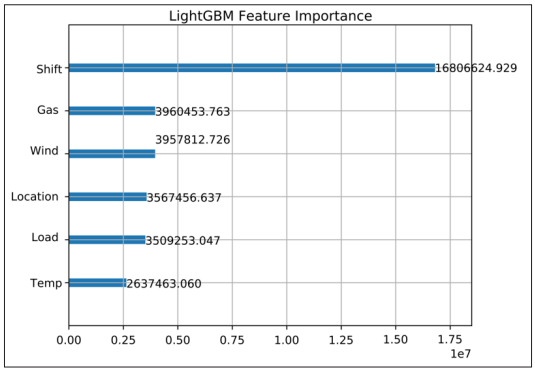

Figure 5: LightGBM Feature Importance

Table 6: Feature Groups by Importance

Tier	Features
1	Shift factor, Load
2	Wind, Gas
3	Location, Temp

However, LightGBM Feature Importance should be taken with a grain of salt, as a high gain can also indicate complexity rather than correlation. Shift factor's impact on electrical prices is not as clear-cut as other factors, for example, Load. When there is a high load, it is obvious that there will be a high price. Futhermore, this is only representative up to stage two of the proposed model meaning stage three changes are not represented.

Model Variations

In the first stage, base predictions, the results of each model are below. Table 7 shows how the FNN and XGBoost can achieve the smallest loss.

Table 7: Stage 1 Model Performance

Variation	MAE	RMSE	MAPE (%)	sMAPE (%)
FNN	5.5934	9.9795	52.18	26.92
LightGBM	5.7107	9.3876	82.26	27.69
XGBoost	5.2444	8.7931	66.13	25.69

So, when testing in the second stage, stacking, the FNN and XGBoost models are used as base models. Then, an FNN, XGBoost, and LightGBM were used to stack the results. When comparing the resulting loss, it is clear that LightGBM is the most effective at stacking. A possible explanation for this is LightGBM's leaf-wise growth strategy, which allows it to assign different weights to different inputs, making it better at handling the scope of inputs for the stacking stage.

Table 8: Stage 2 Model Performance

Variation	MAE	RMSE	MAPE (%)	sMAPE (%)
FNN	5.1150	8.7086	1382893.84	25.05
LightGBM	4.7352	7.5972	3312885.80	24.00
XGBoost	5.1345	8.6381	3194274.73	25.23

Finally, in testing the third stage, residual boosting, the FNN and XG- Boost models are used as base models while LightGBM is used as a stacker. Again, an FNN, XGBoost, and LightGBM are used as residual boosters. The results are in Table 8 below. From these results, we can conclude that XGBoost is the best model for residual boosting.

Table 9: Stage 3 Model Performance

Variation	MAE	RMSE	MAPE (%)	sMAPE (%)
FNN	4.56649	7.4698	3439057.80	23.28
LightGBM	3.7818	5.6593	1700484.90	20.49
XGBoost	2.1360	2.9616	125851617.00	12.90

Therefore, based on these variations, we can conclude that the proposed model is efficient.

Conclusion

This paper proposes a three-stage machine learning model to forecast wholesale electrical prices effectively. By first using a Feedforward Neural Network (FNN) and a gradient boosted XGBoost model to make base predictions that are then stacked utilizing a LightGBM model and finally boosted with another XGBoost model that predicts the residuals, the resulting model offers a unique approach that combines three different machine learning models' strengths to deliver robust electrical price forecasts. Along- side the proposed model is the proposed dataset, including factors such as Load, Wind Speed, Temperature, Shift Factor, and Location, which are able to effectively encompass both supply and demand factors influencing electrical prices.

When benchmarked against other state-of-the-art EPF models, the pro- posed model was able to achieve better accuracy across various metrics and datasets, proving to be an effective and robust approach. Furthermore, when combined with the comprehensive dataset proposed in this paper, the model is able to achieve high accuracy and accurately make new predictions for the future. This paper also showed the importance of Shift Factor and suggests that future studies should also consider it. Future work could also explore more models that could be fitted into this 3-stage approach: base models, stacking results, and residual boosting.

As the electrical market grows, the need for accurate electrical price forecasting grows even larger. Accurate price forecasting is essential to all market participants when making decisions on bidding strategies, congestion handling, cost-effective integration of renewable energy sources, and even where large-scale data centers should be built. Now, with the introduction of FXL3, EPF becomes one step better.

References

1. Weron R. Electricity price forecasting: A review of the state-of-the-art with a look into the future. International Journal of Forecasting. 2014. 30: 1030-1081.
2. Hajigholam Saryazdi S. A novel hybrid deep learning model for electricity price forecasting. SSRN Electronic Journal. 2024.
3. Laitsos I, Andreadis P, Vokas G, Karfopoulos E. Data-driven tech- niques for short-term electricity price forecasting through novel deep learning approaches with attention mechanisms. Energies. 2024. 17: 1625.
4. Conejo AJ, Contreras J, Espınola R, Plazas MA. Forecasting elec- tricity prices for a day-ahead pool-based electric energy market. Inter- national Journal of Forecasting. 2005. 21: 435-462.
5. Borges CL, Falc˜ao DM. Optimal distributed generation allocation for reliability, losses, and voltage improvement. International Journal of Electrical Power & Energy Systems. 2006. 28: 413-420.
6. Zimmerman RD, Murillo-Sa´nchez CE, Thomas RJ. Matpower: Steady-state operations, planning, and analysis tools for power systems research and education. IEEE Transactions on Power Systems. 2011. 26: 12-19.
7. Zhang J, Wang J, Wang W. Review on probabilistic forecasting of wind power generation. Renewable and Sustainable Energy Reviews. 2017. 32: 255-270.
8. Lago J, De Ridder F, De Schutter B. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algo- rithms. Applied Energy. 2018. 221: 386-405.
9. Taieb SB, Hyndman RJ. A gradient boosting approach to the kaggle load forecasting competition. International Journal of Forecasting. 2014. 30: 382-394.
10. Khosravi A, Nahavandi S, Creighton D, Atiya AF. Comprehensive review of neural network-based prediction intervals and new advances. IEEE Transactions on Neural Networks and Learning Systems. 2013. 22: 1341-1356.
11. Uniejewski B, Ziel F. Probabilistic forecasts of load, solar and wind for electricity price forecasting. arXiv .2025.
12. Ziel J, Steinert R, Husmann S. Forecasting day-ahead electricity prices using high-dimensional bid data: A lasso-based approach. Electric Power Systems Research. 2019. 170: 25-32.
13. 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. Applied Energy. 2021. 293: 116983.
14. Olivares KG, Challu C, Marcjasz G, Weron R, Dubrawski A. Neural basis expansion analysis with exogenous variables: Forecasting elec- tricity prices with nbeatsx. International Journal of Forecasting. 2023. 39: 84-900.
15. Chan JW, Yeo CK. A transformer-based approach to electric- ity load forecasting. The Electricity Journal. 2024. 37: 107370.
16. Chen SA, Li CL, Yoder N, Arik SO, Pfister T. Tsmixer: An all- mlp architecture for time series forecasting. arXiv. 2023.
17. Liu Y, Hu T, Zhang H, Wu H, Wang S, et al. Long, itransformer: Inverted transformers are effective for time series forecasting. arXiv. 2024.
18. Ding R, Meng H, Zhang Z, Yang J. Semantically-guided inference for conditional diffusion models: Enhancing covariate consistency in time series forecasting. arXiv. 2025.
19. Ren X, Zhao K, Takova K, Riddle P, Li L. Periormer: Periodic trans- former for seasonal and irregularly sampled time series, in: Proceedings of the 33rd ACM International Conference on Information and Knowl- edge Management, CIKM ’24, Association for Computing Machinery, New York, NY, USA. 2024.
20. Wang M, Wang H, Zhang F. Aligntime: Interperiodic phase alignment sampling for time-series forecasting. Information Processing & Management. 2026. 63: 104296.