INL/CON-24-77208-Revision-0

Short-term Electricity Price Forecasting with Constrained Regressors

April 2024

Changing the World's Energy Future

Mucun Sun, Li Zhang

INL/CON-24-77208-Revision-0

Short-term Electricity Price Forecasting with Constrained Regressors

Mucun Sun, Li Zhang

April 2024

Idaho National Laboratory
Idaho Falls, Idaho 83415

http://www.inl.gov

Prepared for the
U.S. Department of Energy
Under DOE Idaho Operations Office
Contract N/A

Short-term Electricity Price Forecasting with Constrained Regressors

Mucun Sun* Li Zhang†

Abstract

The volatility of electricity price presents a challenge to market participants as their decision-making process are highly depend on the accuracy of price forecasts. However, there is growing empirical evidence of increasing price volatility and price spikes in electricity markets as a result of variable renewable energy generation, extreme weather events, and other factors. The distribution shift caused by spikes in electricity price data differentiates the forecasting tasks from other renewable energy sources. Moreover, the observations may be compromised by cyberattacks and thus not available in the testing phase. To this end, we propose a **S**imilarity-**E**nhanced **E**lectricity **D**ecomposition **F**orecasting model (SEED-Forecaster) to address the missing response problem and spikes capturing in short-term electricity price forecasting. The effectiveness of the proposed framework is tested on real-world electricity price data from California Independent System Operator (CAISO). Numerical results of case studies show that the proposed SEED-Forecsater can enhance forecasting performance, particularly in capturing electricity spikes, even under conditions without regressors during testing stage.

1 Introduction

Accurate electricity prices forecasting (EPF) [23, 5, 16], is important for participants in the wholesale electricity markets. EPF can guide them in formulating bidding strategies, asset allocation, negotiating bilateral contracts, hedging against risks, and planning for facility investments[3, 11, 8]. Additionally, market operators leverage EPFs to gauge market power indexes, monitor the behaviour of market participants[20, 25, 1]. At the same time, various factors affect electricity price such as load, demand, natural gas prices and weather conditions[15, 4]. Moreover, the algorithms determining pricing by balancing supply and demand can lead to electricity price spikes[7, 2]. These spikes can result in huge losses for unwary business owners and are difficult to handle by traditional forecasting models[13].

Despite the inherent challenges of EPF, the process is further complicated by the electricity price being derived from the DC Optimal Power Flow (DCOPF) solution[14]. This solution is contingent upon factors such as grid topology, generation costs, and real-time demand. Consequently, the system is vulnerable to an expanded cyberattack surface[10]. Specifically, attackers can manipulate the topology information to alter the solution to DCOPF leading to alteration of electricity price, thus achieving financial gain. This vulnerability introduces a significant challenge for EPF, as testing electricity price may be compromised by cyberattacks, making it different from traditional EPF problems. To this end, our paper aims to explore a specific scenario for EPF, where electricity price has been targeted by attacks and is consequently unusable during testing stage.

In summary, the key research questions that we aim to address are:

1. How to predict the unusually high spike without compromising overall performance?
2. How can we enhance the performance of non-autoregressive models using the existing training data?

To answer these research questions, we propose a **S**imilarity-**E**nhanced **E**lectricity **D**ecomposition **F**orecasting model (SEED-Forecsater). To reduce the effect of not having autoregressive data, we propose a nearest neighbor shortcut module to directly inject the historically similar input sequence as a shortcut feature. To handle the extreme values, we use r-th root transformation a preprocessing unit to transform the extreme values. The paper is organized as follows. In Section 4, we present the overall proposed framework and detailed the proposed method. Section 5 applies the developed model to the electricity price dataset and compares the model performance across several benchmark deep neural network forecasting models. Concluding remarks and future work are discussed in Section 7.

2 Related Work

Extensive study have been focus on Long Sequence Time Series Forecasting. Typically, such models are based on Transformer model which originally are from natural language processing. A series of work have proposed different sparse attention [17, 27, 24, 19] to alleviate the memory cost problem or rely on a decomposition

*Idaho National LaboratoryID, USA Mucun.Sun@inl.gov
†The University of Texas Rio Grande Valley, li.zhang@utrgv.edu

module [28] to better capture information. Recently, Zeng et al. [26] proposed an MLP with a moving-average module to overcome the distribution shift. However, in electricity price forecasting, the forecasting length is usually small (within 24 sample points).

In the context of renewable energy forecasting problem, several studies have been conducted in the literature to address the aforementioned problem based on pooling, transfer learning, data generation, and other methods. For example, Bottieau *et al.* proposed a cross-learning forecasting approach to predict solar power generation without observation during both training and testing. By leveraging common patterns learned from neighboring monitored solar power production profiles, the proposed approach fits a single, generic forecasting function across the entire panel of monitored solar power time series based only on series-specific features. Moon *et al.* [18] propose a tree-based method to tackle limited regressor issue in short-term load forecasting. Ali *et al.* [12] propose a transfer learning-based framework enhanced by deep generative model for forecasting of residential electric vehicle charging behavior with limited regressor. The middle layers of the pre-trained deep neural networks are freezed and make a shortcut between the output layer and input layer during the backpropagation for newly committed electric vehicles with limited historical charging records. However, most of the aforementioned work are more focused on solving limited regressor issue on training data, which is referred to as 'cold-start' problem. There is very limited work on handling missing regressor problem in energy forecasting.

It is worthy to note that spike or extreme value forecasting has been addressed in electricity price forecasting [22, 21], electrical system load forecasting [9], and wind speed forecasting [6]. However, limited prior research has tackled the forecasting of electricity price with large spikes which are easy to be overfitted, especially under the specific scenario where electricity price data is unusable during testing stage.

3 Problem Formulation

We start with the background of the time series and then formulate problem setting.

We formulate the time series forecasting problem. Given n instances of time series $\mathbf{X} = [\mathbf{x}^1, \mathbf{x}^2, \cdots, \mathbf{x}^n] \in \mathcal{R}^{n \times d \times p}$, where d is the time series dimension, p is the sliding window length. Given $\mathbf{x}^i$, we want to predict a target time series $\mathbf{y}^i$ of length q. A time series forecasting problem aims to learn a model $F(., \theta)$ such that $F(\mathbf{x}^i)$ is close to $\mathbf{y}^i$.

Particularly, in this paper, we are dealing with a situation which $\mathbf{y}^i$ will be completely missing in the testing phase (i.e. the reading is completely not available at testing time, or the corresponding sensor was completely removed.) Thus, unlike auto-regressive models, we assume $\mathbf{y}^i$.is not a subset of Particularly, in this paper, we are dealing with a situation which $\mathbf{y}^i$ will be completely missing in the testing phase (i.e. the reading is completely not available at testing time, or the corresponding sensor was completely removed.) Thus, unlike auto-regressive models, we assume $\mathbf{y}^i$.is not a subset of $\mathbf{x}^i$. Thus, models under such a forecasting setting will be able to predict the target series while experiencing failure, delay, and corruption on the target reading without being affected.

4 Methodology

The overall flowchart of the proposed day-ahead electricity price forecasting framework is illustrated in Figure 1. Overall, the proposed forecasting model consists a extreme value aware normalization module and a Similarity Enhanced Shot-cut module to fuse the forecasting result.

4.1 Extreme Value Aware Normalization Unlike regular seasonal time series, it is well known that the target electricity price time series contains significant extreme values. Forecasting such extreme values is extremely important given the potential cost associated with it. However, leaving such extreme values in the time series also pose a unavoidable challenge – data normalizations step that widely used in time series forecasting to ensure the health of model gradient, is not reliable. For example, the z-score normalization and norm-based normalization will simply transfer all normal data into almost a flat line.

Therefore, in this proposed framework, we propose to use a Extreme Value Aware Normalization module to address this challenge. Specifically, we proposed to first perform r_{th} root of $\mathbf{y}_i$ points:

$$\tilde{\mathbf{y}}_i = \mathbf{y}_i^{\frac{1}{r}} \tag{4.1}$$

The extreme value and the normal value in $\tilde{\mathbf{y}}_i$ will have way smaller value-gap as the r_{th} root greatly reduce the extreme price value. Then the z-score normalization is performed to obtained the normalized value.

$$\bar{\mathbf{y}}_i = \frac{\tilde{\mathbf{y}}_i - \mu}{\sigma} \tag{4.2}$$

where μ and σ is the mean and standard deviation of series $\tilde{\mathbf{y}}$ respectively. Because $\mathbf{x}_i$ is typical features that associated with price y_i and do not have significant

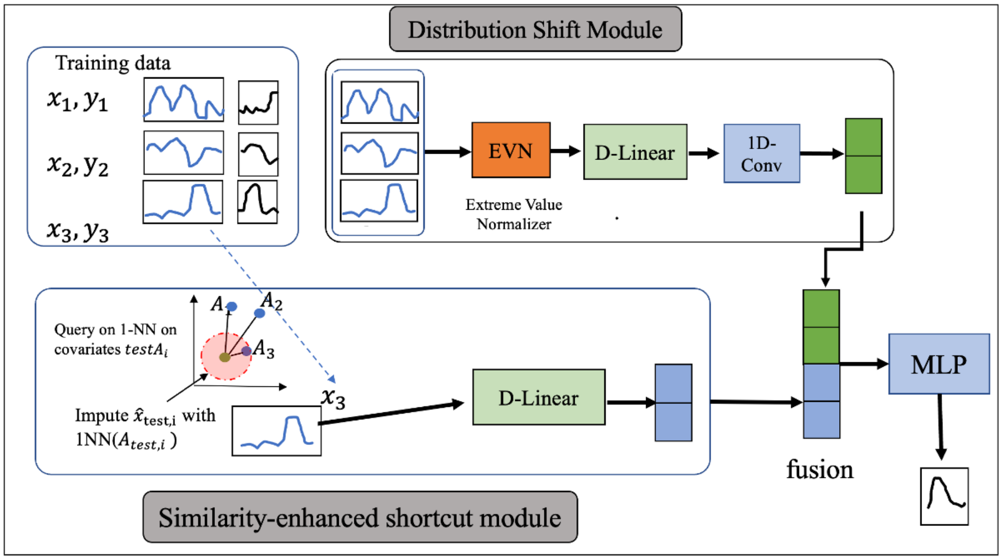

Figure 1: Overview flowchart of SEED-Forecsater

extreme value, we will use regular z-score normalization in the proposed framework:

$$\bar{\mathbf{x}}_i = \frac{\mathbf{x}_i - \mu}{\sigma} \tag{4.3}$$

4.2 Similarity Enhanced Shot-cut module Next, we discuss the proposed similarity-enhanced short-cut module. Intutively, this module aims at aggregating information gathered from $\mathbf{x}_i$ and its previously seen 1-NN $\mathbf{x}_i^{1nn}$ to perform forecasting task. To be more specific, given a input time series $\mathbf{x}_i$, the model first run through a query on a causal 1-nearest neighbor (Causal 1-NN) on $\mathbf{x}_k \quad k = 1 \dots i-1$. The closest match $\mathbf{x}_i^{1nn}$ along with $\mathbf{x}_i$ are passed into two seperate decomposed linear (DLinear)[26] forecasting module to obtain latent embedding:

$$\mathbf{h}_i = \text{DLinear}(\mathbf{x}_i) \tag{4.4}$$

$$\mathbf{h}_i^{1nn} = \text{DLinear}(\mathbf{x}_i^{1nn}) \tag{4.5}$$

$\mathbf{h}_i, \mathbf{h}_i^{1nn}$ are then concatenated to form $z = [\mathbf{h}_i, \mathbf{h}_i^{1nn}]$. Lastly, the forecasting is made through a MLP fused module:

$$\hat{\mathbf{y}}_i = MLP(z) \tag{4.6}$$

5 Experiments and Case Studies

5.1 Experiment Setting We conducted experiments on real-world electricity price data. We use approximately 700 days for training, 100 days for validation, and 293 days for testing. The forecasting performance was evaluated by RMSE and MAE metrics. We assume our model is executed every day at 8AM to predict the next day's hourly prices using the latest available data in the training phase. In the testing phase, we predict the next 24 hours' electricity price given the last 48 hours' other variables without using the electricity price in the testing data. We compare with our backbone model D-Linear [26] as well as MLP model.

5.2 Evaluation Metrics We use two most commonly used evaluation metrics: mean absolute error (MAE) and the root mean square error (RMSE). The mathematical expressions of the two metrics are expressed as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |\hat{y}_i - y_i| \tag{5.7}$$

$$RMSE = \frac{1}{n}\sqrt{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2} \tag{5.8}$$

where $\hat{y}_n$, and y_n, are the forecast wind power and wind power observation, respectively. For these metrics, a smaller value indicates better forecasting performance.

We report both overall and average hourly MAE and RMSE, where the hourly metrics are used to evaluate fine-grained performance.

Table 1: i^{th} Hour Head Performance Comparison

i^{th} hour ahead	Proposed		Dlinear		MLP	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
1	**1.06**	**1.71**	1.20	2.03	1.42	2.22
2	1.05	1.71	1.06	1.85	**0.93**	**1.21**
3	**1.02**	**1.71**	1.03	1.83	1.20	1.92
4	1.01	1.73	0.99	1.79	**0.58**	**0.97**
5	**1.02**	**1.76**	1.00	1.81	1.39	2.34
6	1.06	1.82	1.02	1.86	**1.00**	**1.48**
7	1.04	1.84	**0.96**	**1.78**	1.01	1.84
8	0.98	1.70	**0.88**	**1.63**	3.08	4.54
9	**0.85**	**1.44**	0.90	1.59	3.71	5.30
10	**0.76**	**1.27**	0.84	1.45	2.71	3.79
11	**0.94**	**1.49**	1.25	2.04	3.25	7.10
12	**0.82**	**1.34**	1.32	2.17	2.22	3.82
13	**0.69**	**1.19**	1.10	1.93	4.10	6.99
14	**0.61**	**1.11**	0.95	1.79	4.46	6.30
15	**0.58**	**1.05**	0.82	1.55	6.67	9.42
16	**0.58**	**1.09**	0.64	1.30	2.88	4.46
17	**0.57**	**1.10**	0.61	1.26	2.08	2.59
18	**0.50**	**1.03**	0.54	1.19	2.21	2.94
19	0.52	**1.09**	**0.50**	1.14	0.85	1.44
20	**0.49**	**1.04**	**0.49**	1.11	1.02	1.23
21	0.50	1.05	**0.48**	1.10	0.85	**1.01**
22	0.50	**1.03**	**0.49**	1.10	2.97	4.92
23	**0.64**	**1.22**	0.68	1.38	2.92	4.88
24	**1.02**	**1.72**	1.07	1.89	3.28	4.25
# of Win	**16**	**17**	6	2	3	3

5.3 Data Summary We use electricity price data obtained from real-world hourly usage. There are in total 12 features for approximately three years. Our goal is to predict electric prices in every day 8 am and the forecasting horizon is the next 24 hours. The electricity price data is collected at the pricing nodes of NP-15 zone from the CAISO market. The prices are affected by the dynamic interplay of supply and demand in both the CAISO real-time and day-ahead markets. These prices encapsulate the comprehensive cost of electricity generation, factoring in elements like fuel costs, renewable resource availability, and transmission constraints. It's worth noting that distinct pricing nodes may exhibit divergent prices due to factors such as local supply and demand conditions, transmission constraints, and the generation mix.

We first perform normalization on the entire time series data by subtracting the data mean and data standard deviation, then applied EVN module to reduce the effect of extreme values. The result is reported over all methods on the transformed scale.

Table 2: Average Performance Comparison

	Proposed	Dlinear	MLP
MAE	**0.78**	1.41	2.36
RMSE	**1.41**	2.5	4.24

6 Results and Discussion

Table 1 and Table 2 show our forecasting results evaluated by hourly and overall RMSE and MAE against DLinear [26] and MLP on CAISO data. MAE is known as a metric reflecting the performance of capturing extreme values, while RMSE is known as a metric bias toward the performance of average values. Our proposed method win 16 out 24 hours for MAE and 17 out of 24 hours for RMSE. Our largest MAE value is 1.06, which is lower than D-Linear and MLP by 13.2% and 33.6% respectively, while the largest error of MLP is 9.42, and DLinear is 1.32. Our overall MAE is 0.78, improved 46.7% and 67% of the result of DLinear and MLP. Our method also shows 45% and 66.9% of improvement on RMSE over DLinear and MLP. We can also see MLP fails in many cases and produced large error of more than 3, even reach to 9.42. That is because there is an outlier and DLinear and our method is robust to this. The results show that our proposed method significantly improves the ability to fit the extreme data as well as normal cases by infusing the similarity module.

7 Conclusion

This paper developed a novel model for short-term electricity price forecasting. This model is specifically designed to tackle a special problem–where testing regressors are unavailable–and to accurately predict spikes in electricity prices. Results of the case study showed that: (i) the proposed method improves both MAE and RMSE compared to baseline models even under scenarios with missing regressor during the testing stage; (ii) the proposed model excels in accurately capturing electricity spikes. Future work will explore: (i) generalizing the model for other renewable energy forecasting including wind, solar, and electric vehicle charging, and (ii) adapting the model to address missing regressor issues in clustering problems.

References

[1] M. Barmack, E. Kahn, S. Tierney, and C. Goldman. Econometric models of power prices: An approach to market monitoring in the western us. *Utilities Policy*, 16(4):307–320, 2008.

[2] R. Becker, S. Hurn, and V. Pavlov. Modelling spikes in electricity prices. *Economic Record*, 83(263):371–382, 2007.

[3] T. K. Boomsma, N. Juul, and S.-E. Fleten. Bidding in sequential electricity markets: The nordic case. *European Journal of Operational Research*, 238(3):797–809, 2014.

[4] C. Brandoni, M. Renzi, F. Caresana, and F. Polonara. Simulation of hybrid renewable microgeneration systems for variable electricity prices. *Applied thermal engineering*, 71(2):667–676, 2014.

[5] J. P. d. S. Catalão, S. J. P. S. Mariano, V. Mendes, and L. Ferreira. Short-term electricity prices forecasting in a competitive market: A neural network approach. *electric power systems research*, 77(10):1297–1304, 2007.

[6] P.-C. Chang, R.-Y. Yang, and C.-M. Lai. Potential of offshore wind energy and extreme wind speed forecasting on the west coast of taiwan. *Energies*, 8(3):1685–1700, 2015.

[7] T. M. Christensen, A. S. Hurn, and K. A. Lindsay. Forecasting spikes in electricity prices. *International Journal of Forecasting*, 28(2):400–411, 2012.

[8] A. K. David. Competitive bidding in electricity supply. In *IEE proceedings C (generation, Transmission and distribution)*, volume 140, pages 421–426. IET, 1993.

[9] X. Deng, A. Ye, J. Zhong, D. Xu, W. Yang, Z. Song, Z. Zhang, J. Guo, T. Wang, Y. Tian, et al. Bagging–xgboost algorithm based extreme weather identification and short-term load forecasting model. *Energy Reports*, 8:8661–8674, 2022.

[10] J. Duan, W. Zeng, and M.-Y. Chow. Economic impact of data integrity attacks on distributed dc optimal power flow algorithm. In *2015 North American Power Symposium (NAPS)*, pages 1–7. IEEE, 2015.

[11] G. Federico and D. Rahman. Bidding in an electricity pay-as-bid auction. *Journal of Regulatory Economics*, 24(2):175–211, 2003.

[12] A. Forootani, M. Rastegar, and H. Zareipour. Transfer learning-based framework enhanced by deep generative model for cold-start forecasting of residential ev charging behavior. *IEEE Transactions on Intelligent Vehicles*, 2023.

[13] H. Geman and A. Roncoroni. Understanding the fine structure of electricity prices. *The Journal of Business*, 79(3):1225–1261, 2006.

[14] A. Kargarian, J. Mohammadi, J. Guo, S. Chakrabarti, M. Barati, G. Hug, S. Kar, and R. Baldick. Toward distributed/decentralized dc optimal power flow implementation in future electric power systems. *IEEE Transactions on Smart Grid*, 9(4):2574–2594, 2016.

[15] D. Keles, M. Genoese, D. Möst, and W. Fichtner. Comparison of extended mean-reversion and time series models for electricity spot price simulation considering negative prices. *Energy economics*, 34(4):1012–1032, 2012.

[16] J. Lago, F. De Ridder, and B. De Schutter. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. *Applied Energy*, 221:386–405, 2018.

[17] S. Li, X. Jin, Y. Xuan, X. Zhou, W. Chen, Y.-X. Wang, and X. Yan. Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting. *Advances in neural information processing systems*, 32, 2019.

[18] J. Moon, J. Kim, P. Kang, and E. Hwang. Solving the cold-start problem in short-term load forecasting using tree-based methods. *Energies*, 13(4):886, 2020.

[19] Y. Nie, N. H. Nguyen, P. Sinthong, and J. Kalagnanam. A time series is worth 64 words: Long-term forecasting with transformers. *arXiv preprint arXiv:2211.14730*, 2022.

[20] A. Rahimi and A. Y. Sheffrin. Effective market monitoring in deregulated electricity markets. *IEEE Transactions on Power systems*, 18(2):486–493, 2003.

[21] I. Shah, S. Akbar, T. Saba, S. Ali, and A. Rehman. Short-term forecasting for the electricity spot prices with extreme values treatment. *IEEE Access*, 9:105451–105462, 2021.

[22] W. Shi, Y. Wang, Y. Chen, and J. Ma. An effective two-stage electricity price forecasting scheme. *Electric Power Systems Research*, 199:107416, 2021.

[23] R. Weron. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International journal of forecasting*, 30(4):1030–1081, 2014.

[24] H. Wu, J. Xu, J. Wang, and M. Long. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting. *Advances in Neural Information Processing Systems*, 34:22419–22430, 2021.

[25] H. Zareipour, K. Bhattacharya, and C. A. Canizares. Electricity market price volatility: The case of ontario. *Energy policy*, 35(9):4739–4748, 2007.

[26] A. Zeng, M. Chen, L. Zhang, and Q. Xu. Are transformers effective for time series forecasting? 2023.

[27] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, and W. Zhang. Informer: Beyond efficient transformer for long sequence time-series forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pages 11106–11115, 2021.

[28] T. Zhou, Z. Ma, Q. Wen, X. Wang, L. Sun, and R. Jin. Fedformer: Frequency enhanced decomposed transformer for long-term series forecasting. *arXiv preprint arXiv:2201.12740*, 2022.