Negative Price Forecasting in Australian Energy Markets using gradient-boosted Machines: Predictive and Probabilistic Analysis

Devinder Kaur*, Gagangeet Singh Aujla†, and Md. Apel Mahmud‡
*School of Engineering, Deakin University, Australia
†Department of Computer Science, Durham University, Durham, United Kingdom
‡Department of Mathematics, Physics and Electrical Engineering, Northumbria University, UK
E-mail: devinderkaur@deakin.edu.au, gagangeet.s.aujla@durham.ac.uk, md.a.mahmud@northumbria.ac.uk

***Abstract*—With the integration of distributed energy resources such as roof-top solar panels and wind turbines into the grid, power generation can surpass demand-generation and thus, gives rise to the negative pricing events, especially during summer months. In this regard, a scientific case study is conducted in this paper to analyse and predict the increasing instances of negative energy prices against demand-generation in Australian energy markets (AEMs) using real-time energy data from the Hornsdale power reserve, South Australia. A robust machine learning method, Light gradient boosting machine (LightGBM) is utilised to detect and predict negative prices at different quantiles to quantify the outliers in the pricing data. The implementation results demonstrate that predicting the prices at different quantiles can tackle outliers (negative prices) effectively with the help of extracted upper and lower bounds using quantile regression-based approach. The case study is further extended to learn the complex statistical relationships between different data features using Naive-Bayes Tree Augmented (NB-TAN) algorithm considering 'price' as the dependent feature against the independent features such as demand-generation, battery charging/discharging, and frequency control ancillary services.**

Keywords - Australian energy markets, battery storage systems, light gradient-boosted machines, negative pricing, quantile regression, renewable energy generation.

I. Introduction

With more renewable energy sources (RESs) getting integrated into the electricity grid, increased instances of wholesale electricity prices going negative have been observed in the Australian energy markets (AEMs) in the past 5 years[1]. Negative pricing usually occurs when the demand is low but the energy generation is high mainly due to renewable energy sources such as solar and wind power [1], [2]. Lately, in the United States of America (USA), a notable trend has also been reported in the negative pricing events particularly in the states of Texas and California due to the increased supply of renewable energy from the wind and solar power plants [3], [4]. Statistically, South Australia has maximum instances of negative prices observed i.e., 2017 instances in the year 2019, which is more than any other states in Australia[2]. Fig. 1 reflects the analyses of electricity pricing values against battery discharging (MW) and energy demand for the month of Dec., 2021 in South Australia. According to the stats from this month, 24.9% values of the total pricing data are reported negative with a minimum of -\$999/MWh against 864.78 MW demand-generation with an interval of 5 minutes.

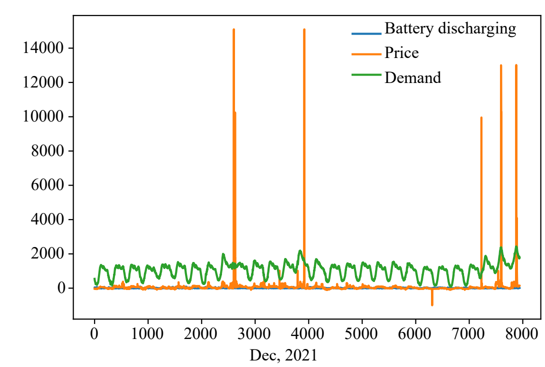


Fig. 1. Battery discharging, price, and demand trends for Dec, 2021

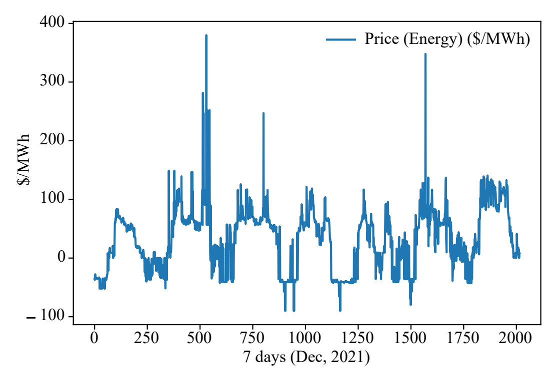


Fig. 2. Observed prices for 7 days

Furthermore, Fig. 2 shows a clearer image of price fluctuations for the first week of Dec, 2021. It can be observed that there is a frequent trend of negative pricing, especially

[1]https://flowpower. com.au/positives-of-negative-prices/

[2]https://www.energycouncil.com.au/analysis/ increases-in-negative-prices-is-it-a-positive/

during the middle of the day. As during mid-day, especially on warm and windy days, RESs produce energy with high intensity leading to demand and supply deregulation in the power grid.

To tackle this demand and supply variation, either generators producing energy should be turned off or consumers need to be encouraged to consume more energy. However, turning off the energy generators is a slow and not so cost effective process. Instead, energy is being sold at lower prices and thus, causing a rise in negative pricing instances, i.e. consumers get paid to consume energy guaranteeing dispatch. As RES generators get fixed turnovers, so they are happy to sell it at lower prices causing the rise in negative pricing instances. However, it can impact grid operations and planning services as stakeholders seek to manage and mitigate the risks associated with the negative pricing events in advance.

An effective way to tackle the negative pricing is by utilising battery storage systems. As, it is more cost effective to pay the battery operators than to shut the RES generators. When the battery is getting charged, the energy gets stored and it can later be discharged to be used at the consumer-end according to the demand. However, to execute this solution, an optimal battery charging technique needs to be implemented which further relies on planing and predicting negative pricing instances effectively against demand-generation and other key features in the energy data such as battery charging/discharging, frequency control ancillary services (FCAS) prices etc.

In the past few years, the Australian national electricity market (NEM) has started to use automated methods involving data-driven intelligent techniques [5] such as machine learning algorithms for energy forecasting applications [6] to aid bidding operations as opposed to conventional trading tools. This has helped to increase the NEM revenue by 10% according to a recent report from the Australian clean energy council[3].

In this direction, electricity price forecasting is one of the key applications to support decision making process in the energy markets to support energy trading operations. In [7], the authors have conducted a detailed study about existing price forecasting methods, their strengths and weaknesses. The authors highlighted that data-driven intelligent methods are the future for energy forecasting domain [8]. Currently, the most commonly used methods for price forecasting are statistical models and non-linear machine learning models such as linear regression (LR) and exponential smoothing [9]. Furthermore, non-linear artificial intelligence-based neural networks are gaining popularity in their ability to learn hidden patterns effectively in the data [10]. Although these methods provide a reasonable computational performance and accuracy, these are not robust enough when dealing with the outliers (negative instances) [11]. Therefore, the authors emphasised the need to propose a robust and holistic technique for effective price forecasting, especially when the energy data distribution is skewed and has outliers. In this case, conditional quantiles need to be quantified in the form of prediction intervals rather than just the conditional mean.

In this regard, quantile regression (QR)-based methods such as quantile regression averaging (QRA) and factor QR averaging (FQRA) have been proven successful to compute the prediction intervals and generate probabilistic forecasts in the Global Energy Forecasting Competition 2014 [12]. In specific, QRA which is a forest combination approach achieved 3rd rank for solar, wind, and price forecasting. While QRA and FQRA were proven to be effective to capture the conditional quantiles for the target feature, they lack interpretability and flexibility. In this regard, machine-learning based Light gradient boosted machines (LightGBMs) are found to be more scalable and can be of vital importance with their superior feature selection and handling capabilities [13].

A. Motivation

Considering the aforementioned research challenges and lack of effective research proposals, it can be inferred that there is a dire need of a robust negative price forecasting method quantifying outliers in the pricing data. In addition, statistical and probabilistic analysis among various features for the target 'price' variable needs to be entailed with the forecasting technique.

B. Key Contributions

To be specific, we make the following contributions:

- A QR-based Light gradient boosted machine (LightGBM) is adopted to predict negative prices in Australian energy markets. A case study is presented using real-time energy dataset taken from Australian energy market operator (AEMO) and Hornsdale Power reserve, South Australia. The target variable 'price' in the data is predicted against energy demand-generation feature at 10%, 50%, and 90% percentiles.
- In addition, statistical and probabilistic relationships between various features in the given data are modeled and analysed using Naive-Bayes Tree Augmented (NB-TAN) algorithm. The analysis verifies higher dependencies between demand-generation and 'price' (target) feature.
- Furthermore, to support the efficiency of LightGBM, a comparative analysis is performed for two different months (Dec., 2021 and Sep., 2022) of data against LR, QRA, and FQRA using root-mean square and Pinball errors. The numerical and graphical results demonstrate that predicting energy prices at different percentiles can cover more instances in advance using the extracted upper and lower bounds.

C. Organisation

The rest of the paper is organised as follows. Section II elaborates the proposed methodology including preliminaries of QR using LightGBM and feature analysis using NB-TAN algorithm. Section III presents implementation results and discussions. Finally, Section IV concludes the paper and outlines the future work.

[3]https://reneweconomy.com.au/australias-wind-and-solar-farms-seek-to-dodge-negative-prices-and-grid-costs/

II. Proposed methodology

A. Preliminaries of QR

QR is a type of regression analysis used when conditions of LR are not met and the conditional quantiles are of interest. In case of high variability it becomes challenging for LR to accurately generate predictions in the presence of outliers. In this regard, QR algorithms can be of prime importance as they can quantify outliers with the help of upper and lower bounds extracted outside the mean of the data. In a regression problem, with y as dependent (target) variable and x as independent variable, the regression equation is given as:

$$y = f(x) \tag{1}$$

The linear cost function for above regression problem is computed by the squared loss (l), as:

$$l = (y - \hat{y})^2 \tag{2}$$

$$l = (y - \hat{f(x)})^2 \tag{3}$$

where $\hat{y}$ symbolizes the predicted value and y denotes the actual value for the dependent variable. Furthermore, l is used to compute the root-mean square error ($RMSE$) to access the performance of deterministic or point-based regression methods with total n number of samples as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(l)} \tag{4}$$

On the other hand, to assess the efficacy of QR, quantile loss (Q_l) is utilised, as:

$$Q_l = \begin{cases} \tau(y - \hat{y}) & if(y - \hat{y}) \geq 0 \\ (\tau - 1)(y - \hat{y}) & if(y - \hat{y}) < 0 \end{cases} \tag{5}$$

where τ can take any value between 0 to 100 percent or $(0, 1)$. First case in (5) signifies when actual values are greater than or equal to predicted values. When going for lower bound percentiles, higher values are penalised. Second case signifies when predicted values are higher than the actual values, i.e., when dealing with higher percentiles.

Therefore, the cost function formulates to minimise the error between predicted and actual values, as given below:

$$\min_{\phi} Q_l \tag{6}$$

where, ϕ represents model parameters for the training process such as learning rate, no. of estimators or trees, maximum depth, etc.

B. Gradient Boosting using LightGBM

Gradient boosting is an ensemble machine learning technique that combines multiple weak prediction learners. In this paper, we utilise light gradient boosted machine (LightGBM) framework to implement QR for negative price forecasting [14]. LightGBM supports faster training with the help of parallel and distributed learning module in it. Furthermore, it uses smart featuring and advanced sampling methods to obtain optimised results using various techniques such as binning [15].

LightGBM utilises ensemble approach for predictive modeling working with M number of weak learners (decision trees), say m_0, m_1 , $\cdots$, m_n and m_0 being the baseline model/learner. To begin with, m_0 is used to model the actual values and errors are computed as the end product. Consecutively, the other models are trained on these errors one by one and after combining the results it entails a more efficient and accurate final model, as presented by the following equation:

$$F(x) = \sum_{m=1}^{M} f_m(x) \tag{7}$$

where $F(x)$ is the final model and $f_m(x)$ is the output value of the m^{th} weak regression tree.

It is important to note that the tree in LightGBM grows leaf-wise as shown in Fig. 3, which helps in effective loss reduction and improves the accuracy. While other boosting algorithms grow trees level-wise and horizontally, LightGBM grow decision trees vertically.

Algorithm 1 QR using LightGBM for negative price forecast

No. of models to fit: M
No. of observations: N
Parameters: ϕ
Performance metrics: Q_l, $RMSE$

1: Initialise the model with actual/constant values;
2: Fit the baseline model (m_0) with N;
3: **for** m=1 to M **do**:
4: Compute the pseudo-residuals;
5: Fit the learners on ϕ using (7);
6: Compute Q_l and $RMSE$ using (4) and (5), res.;
7: Update m;
8: **end for**

Algorithm 1 represents pseudo-code for tree-based LightGBM method. As given, the algorithm involves M number of models/learners/decision trees to be iterated and N number of actual data points over parameters ϕ. The performance metrics used to access the performance of algorithm are pinball error (Q_l) and $RMSE$. The model is initialised with constant values (line 1), m being the index of each decision tree (line 3), each $m + 1$ attempts to improve the errors of its predecessor m. GBM learns the errors or residuals in previous rounds and tries to improve the errors in every round. The new weak model is then fitted to the negative gradient, aiming to reduce the overall error.

C. Feature analysis using NB-TAN method

This section presents analysis conducted on real-time pricing data to model dependencies between features. Various statistical and probabilistic properties of different features in the pricing data are learned which can further support the detection and predictive analysis of negative price forecasting in Australian energy markets. We use probability distribution

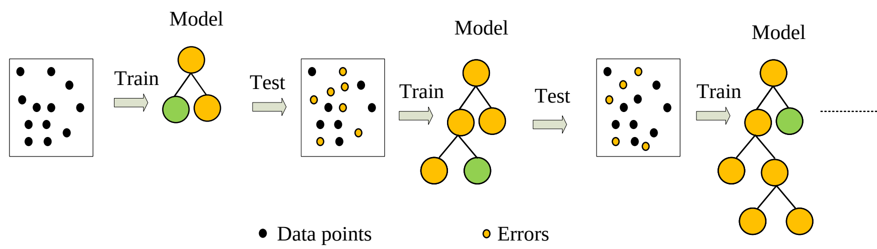


Fig. 3. Leaf-wise tree growth in LightGBM

visualiser also known as ‘Genie’ software to conduct this task. This module is basically used to create and refine network models and includes various parameter learning algorithms such as Naive Bayes classifier. Fig. 4 depicts the representation of relationships between different features in pricing data and how they are statistically impacted by each other.

The main implementation steps are described as follows:

- *Data description:* The main features considered for this analysis are time, price (energy) (\$/MWh), demand, price (FCAS regulation raise) (\$/MWh), price (FCAS regulation raise) (\$/MWh), battery discharging (MW) and battery charging (MW).

- *Data discretisation:* This step involves discretising various features into different bins using hierarchical method based on data distribution of a particular feature. To be precise, for demand and price features, bin size of 4 is taken. While for FCAS and battery related variables, bin size of 2 is considered as reflected in Fig. 4. It is important to note that to discretise the target variable 'price', uniform counts method is used instead of hierarchical to give equal importance to the pricing values below 0.
- *Feature selection:* In this step, the dependent (target) and independent variables are selected. As reflected in the figure, price is the target variable and rest of the five variables are selected as the independent variables/features.
- *Parameter learning:* For parameter learning, NB-TAN algorithm is utilised.
- *Validation:* Furthermore, k-fold validation is utilised to validate the parameter learning process.
- *Extract the statistics and probabilities:* It can be inferred from the nodes and arrows of the directed cyclic graph in Fig. 4 that 'price' is the parent node which is closely related to all other variables (nodes). Here, a special case is demonstrated, where negative (s1_below_0) 'price' values are considered and its impact on other variables is analysed. It can be observed that negative prices arise when 72% of total demand-generation comes from below 863 kW; and 99% of battery discharges from below 48 MW.

  Furthermore, in Fig. 5, it can be seen that when demand is highest, which is above 2066 kW, the energy prices are high as well i.e., more than 72 \$/MWh, which is 98%. Therefore, it is inferred that price and demand-generation are two closely related features and for forecasting purposes, demand-generation has been considered as independent variable to perform the predictive analysis.

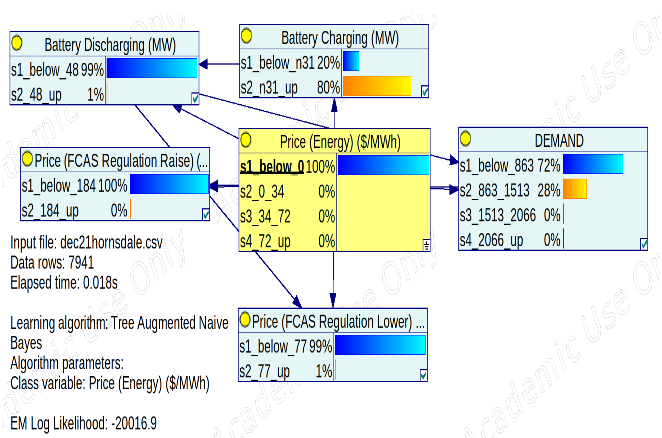


Fig. 4. Complex relationships and probabilistic visualisations between features (with energy prices less than zero (negative))

III. Case Study

This section presents an elaborated and reproducible case study for negative price forecasting carried on real-time energy data using state-of-the-art LR and QR-based algorithms.

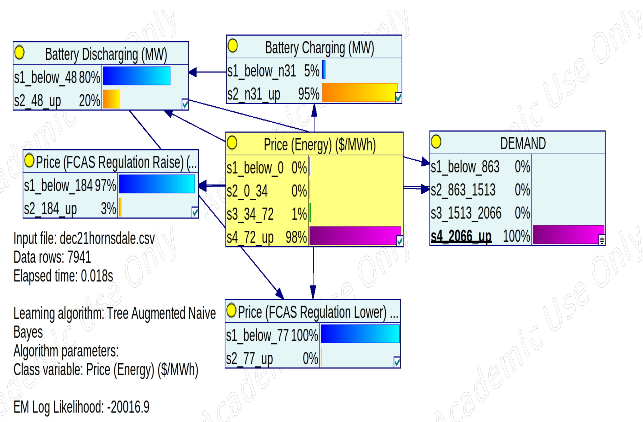


Fig. 5. Complex relationships and probabilistic visualisations between features (with high demand i.e., greater than 2066 kW)

A. Dataset generation and description

The proposed method is implemented on real-time pricing data against demand-generation with a training to testing ratio of 80:20 generated from AEMO[4]. In addition, battery (150 MW) dataset is also considered from Hornsdale power reserve, South Australia[5] to verify the statistics. The data gets generated at the interval of 5 minute and includes battery charging/discharging, demand-generation, price (energy), price (FCAS regulation lower) and price (FCAS regulation raise) values. As of now, for the predictive analysis, 'demand-generation' is considered as the independent variable and 'price' is considered as the target variable.

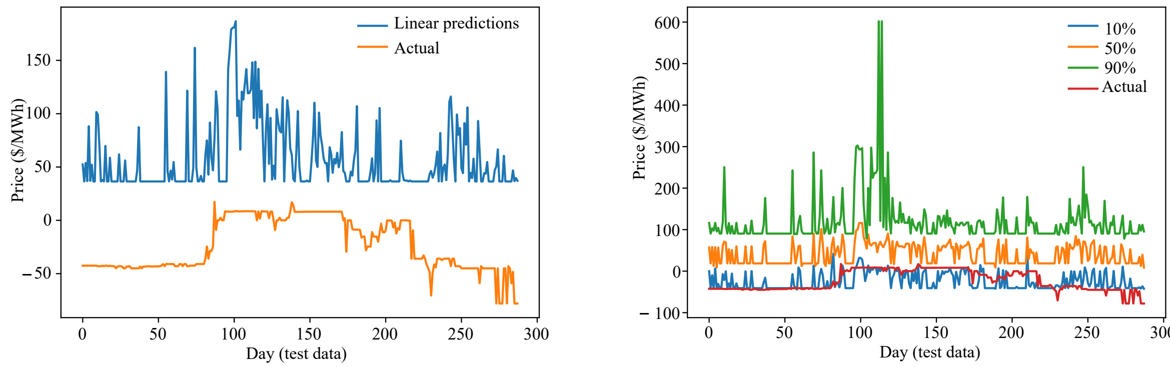
For implementation purposes, the months of Dec., 2021 and Sep., 2022 are considered for brevity. Furthermore, to test the efficacy, a comparative analysis is provided in which LightGBM has been tested against state-of-the-art LR, QRA and FQRA algorithms. The implementation is executed using Python machine learning libraries such as scikit-learn.

B. Results and discussions

This section presents implementation results and discussions for the predictive analysis conducted on pricing data using state-of-the-art LightGBM method. Figs. 6(a) and 6(b) show comparative plots for actual and obtained predictions for energy prices for Dec., 2021 by LR and QR using LightGBM, respectively. As reflected in Fig. 6(a), LR provides only mean prediction values, i.e., at (50%), which are not accurate enough. On the other hand, plot in Fig. 6(b) provides predictions at different percentiles, and thus, covers prices at larger variance from negative to larger prices at percentile values 10, 50, and 90. Also, it is important to note that at 10^{th} percentile more accurate predictions are obtained with lesser error. Similarly, Figs. 7(a) and 7(b) outline the graphical representations of linear vs LightGBM methods to predict the negative prices for the month of Sep., 2022 for South Australia. It can be observed from these figures that

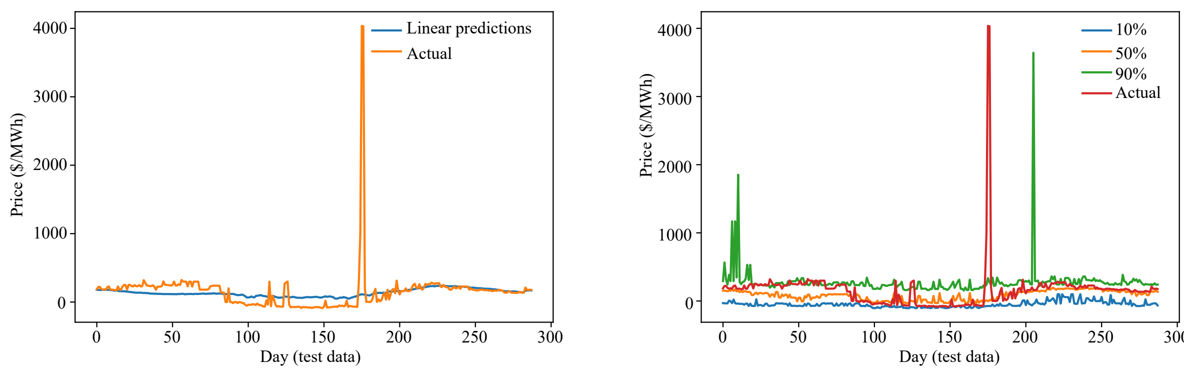
[4]https://www. aemo.com.au/about

[5]https: //hornsdalepowerreserve.com.au/



(a) Actual vs linear predictions (b) 10%, 50% and 90% percentiles (LightGBM) vs Actual prices

Fig. 6. Dec., 2021



(a) Actual vs linear predictions (b) 10%, 50% and 90% percentiles (LightGBM) vs Actual prices

Fig. 7. Sep. 2022

the outlier at $4000 cannot be quantified by LR. However, using LightGBM, at three different percentiles, prices can be predicted accurately within a range.

A comparative analysis of the price prediction errors using LR, QRA, FQRA, and LightGBM is provided in Table I. It is important to note that for LR the RMSE value is slightly lower than QR-based methods as the average of three percentiles, i.e. at 10, 50, and 90 is taken for the latter. To evaluate the performance of QR, Pinball score (avg.) is considered using Eq. (5). Also, please note that for LR, pinball loss can not computed as just the conditional mean is considered and thus, the space is left blank. From the numerical values under pinball scores, it can be observed that LightGBM outperforms other state-of-the-art methods namely, QRA, and FQRA.

Furthermore, Fig. 8 shows the scatter-plot for a comparison between actual and predicted prices with respect to battery discharging values at the horizontal axis. It is plotted against entire test data mainly to show that the apart from few outliers, rest of the data points are covered effectively by LightGBM with large variance in the price values, ranging from -$999 to $15100. Most values are covered around the mean pricing, which is $73.20.

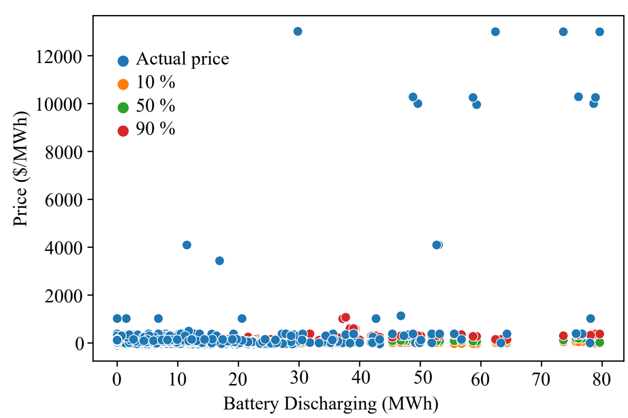


Fig. 8. Actual vs predictions with respect to battery discharging

IV. Conclusion

In this work, a predictive and probabilistic case study is presented to analyse and predict the increasing rate of negative prices in Australian energy markets, for the months of Dec., 2021 and Sep., 2022. The state-of-the-art LightGBM method is

TABLE I
LINEAR VS QR FOR PRICE FORECASTING

Sr. no.	Method	RMSE (quantile avg.)	Pinball (avg)	RMSE (quantile avg.)	Pinball (avg)
		Dec., 2021		Sep., 2022	
1	LR	946.64	-	150.52	-
2	QRA	1024.23	142.54	269.04	124.91
3	FQRA	1104.12	181.24	280.34	134.89
4	LightGBM	955.93	88.74	247.56	73.90

employed at 10, 50 and 90 percentiles to highlight the importance of QR methods contrary to LR in order to detect outliers (negative prices) effectively in the energy data. In addition, feature analysis is performed using NB-TAN algorithm. The future work will focus on how battery storage systems can be benefited from effective negative price forecasting.

REFERENCES

[1] R. P. McKone and P. J. Wolfs, "Analysis of historical spot prices for the australian national energy market, growing concerns about negative prices & postulations on energy storage," in *2019 29th Australasian Universities Power Engineering Conference (AUPEC)*. IEEE, 2019, pp. 1–6.

[2] W. Chen, B. Gath, E. Kinross-Smith, H. Toole, E. Yusofi, and K. N. Hasan, "Analysis of negative electricity price to identify demand management opportunity for consumers in renewable-rich power systems," in *2021 IEEE PES Innovative Smart Grid Technologies-Asia (ISGT Asia)*. IEEE, 2021, pp. 1–5.

[3] J. Seel, D. Millstein, A. Mills, M. Bolinger, and R. Wiser, "Plentiful electricity turns wholesale prices negative," *Advances in Applied Energy*, vol. 4, p. 100073, 2021.

[4] M. Bajwa and J. Cavicchi, "Growing evidence of increased frequency of negative electricity prices in us wholesale electricity markets," in *IAEE Energy Forum*, vol. 2017, 2017.

[5] D. Kaur, S. N. Islam, M. A. Mahmud, and M. E. Haque, "Interval-based probabilistic load forecasting for individual households: Clustering approach," in *2022 IEEE PES 14th Asia-Pacific Power and Energy Engineering Conference (APPEEC)*. IEEE, 2022, pp. 1–6.

[6] D. Kaur, S. N. Islam, and M. A. Mahmud, "A bayesian deep learning technique for multi-step ahead solar generation forecasting," *arXiv preprint arXiv:2203.11379*, 2022.

[7] Y. Li, N. Yu, and W. Wang, "Machine learning-driven virtual bidding with electricity market efficiency analysis," *IEEE Transactions on Power Systems*, vol. 37, no. 1, pp. 354–364, 2021.

[8] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," *International journal of forecasting*, vol. 30, no. 4, pp. 1030–1081, 2014.

[9] D. Kaur, S. N. Islam, M. A. Mahmud, M. E. Haque, and Z. Y. Dong, "Energy forecasting in smart grid systems: recent advancements in probabilistic deep learning," *IET Generation, Transmission & Distribution*, vol. 16, no. 22, pp. 4461–4479, 2022.

[10] R. Juban, H. Ohlsson, M. Maasoumy, L. Poirier, and J. Z. Kolter, "A multiple quantile regression approach to the wind, solar, and price tracks of gefcom2014," *International Journal of Forecasting*, vol. 32, no. 3, pp. 1094–1102, 2016.

[11] H. Deng, F. Yan, H. Wang, L. Fang, Z. Zhou, F. Zhang, C. Xu, and H. Jiang, "Electricity price prediction based on lstm and lightgbm," in *2021 IEEE 4th International Conference on Electronics and Communication Engineering (ICECE)*. IEEE, 2021, pp. 286–290.

[12] B. Uniejewski and R. Weron, "Regularized quantile regression averaging for probabilistic electricity price forecasting," *Energy Economics*, vol. 95, p. 105121, 2021.

[13] Y. Wang, J. Chen, X. Chen, X. Zeng, Y. Kong, S. Sun, Y. Guo, and Y. Liu, "Short-term load forecasting for industrial customers based on tcn-lightgbm," *IEEE Transactions on Power Systems*, vol. 36, no. 3, pp. 1984–1997, 2020.

[14] M. Mohamed, F. E. Mahmood, M. A. Abd, A. Chandra, and B. Singh, "Dynamic forecasting of solar energy microgrid systems using feature engineering," *IEEE Transactions on Industry Applications*, 2022.

[15] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "Lightgbm: A highly efficient gradient boosting decision tree," *Advances in neural information processing systems*, vol. 30, 2017.