12th IFAC International Symposium on
Advanced Control of Chemical Processes
July 14-17, 2024. Toronto, Canada

A New Modeling Framework for Real-Time Extreme Electricity Price Forecasting

You Peng *,† Zhenyu Wang ** Ivan Castillo *
LaGrande, Gunnell *** Shengli Jiang ****

* *Chemometrics, AI and Statistics, The Dow Chemical Company*
** *Energy and Climate Technology Center, The Dow Chemical Company*
*** *Department of Chemical Engineering, Brigham Young University*
**** *Department of Chemical and Biological Engineering, Princeton University*
† *Corresponding author: ypeng4@dow.com*

Abstract: In recent years, extreme electricity prices have occurred with greater frequency and magnitude. Accurately predicting extreme electricity prices is of great interests to market participants. This paper aims to forecast real-time electricity prices for the next 24 hours for the Houston load zone in Electric Reliability Council of Texas (ERCOT), targeting at providing accurate prediction of potential extreme prices. Historical energy prices from ERCOT and weather data from the National Oceanic and Atmospheric Administration (NOAA) were used. A new modeling framework that takes forecasted upper bound of exogenous variables to predict both (1) real-time prices with uncertainty using a temporal fusion transformer (TFT) model as well as (2) likelihood of having extreme prices in the forecasting horizon is proposed. Additionally, a concatenated model fusion strategy is applied as an additional step to further increase the framework's capability of accurately forecasting extreme prices. Our proposed method showed better forecasting capability with a RMSE of 35 compared to other state of the art methods such as auto-encoder long short-term memory (AE-LSTM) and PatchTST (RMSE of 49 and 42 respectively) on the same testing period. In addition, we quantified the uncertainty of the predictions leveraging the quantile output from the TFT model and found that 97% of the time, the 98th quantile of the forecasting horizon contains the actual real time price. Our proposed framework provided an accurate and robust approach for forecasting normal and extreme electricity prices that could have significant economic benefits to electricity market participants.

1. INTRODUCTION

The price of electricity is a key indicator of market activity. Electricity price forecasting has been a major focus of practitioners and researchers in the energy market because it is essential for market participants to set up bidding strategies and create appropriate products. Nevertheless, electricity price forecasting is volatile and it can be influenced by a variety of factors such as holidays, weather and policy. As electricity markets become increasingly integrated, it becomes more difficult to forecast the price. Moreover, extreme price (i.e., price peaks) happens from time to time mostly caused by rare events such as winter storms and extreme heat, can result in significant losses for market participants if it cannot be accurately forecasted. Most wholesale electricity marketplace includes two markets that work together in a multi-settlement system (Kamat and Oren, 2002). The day-ahead (DA) energy market allows market participants to commit to buy wholesale power the day before the operating day to help avoid price volatility. The real-time (RT) energy market allows market participants to buy wholesale power during the actual operating day. The RT energy market balances the difference between DA commitments and actual RT electricity demand and production by generating a separate financial statement. For demand that deviates from DA commitments, it establishes a RT Locational Marginal Price (LMP) that is charged to participants in the DA energy market. Therefore, it is necessary to forecast electricity prices in the RT energy markets accurately in order to plan electricity demand rationally and avoid economic losses. An example of such opportunity is shown in Figure 1 when DA price is much lower than RT price from 12pm to 6pm and being able to predict a coming peak of higher RT price to increase purchase of electricity at DA price will result in significant economic savings.

Extensive research efforts have contributed to the development and utilization of advanced technologies for electricity price forecasting. Statistical methods, such Auto-Regressive Integrated Moving Average (ARIMA) (Nguyen and Hansen, 2017) and Generalized Auto-regressive Conditional Heteroskedastic (GARCH) (Garcia et al., 2005), are widely used to forecast electricity prices. However, many statistical methods are limited in dealing with complex or nonlinear time series problems, especially extreme electricity prices. Compared to conventional statistical models, machine learning models due to their superior performance has boosted their popularity in the recent years.

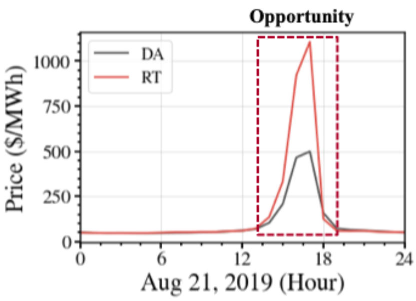


Fig. 1. Example of Economic Opportunity when Large Gap exists between DA and RT Electricity Price

Especially since the mid-2010s, least absolute shrinkage and selection operator (LASSO) (Uniejewski et al., 2016) methods have become popular due to its ability to automatically select and reduce the dimension of the regressors to a small subset of key relevant ones. Therefore, it has also been adopted in energy price forecasting, and one recent example is the development of LASSO-Estimated AutoRegressive (LEAR) model that combines statistical and machine learning methods (Lago et al., 2021).

In the recent decade, as computational resources becomes readily available that makes training of deep neural networks (DNN) feasible, DNNs especially RNNs (recurrent neural networks) have quickly become the most popular energy price forecasting methods among other disciplines. Li et al. applied the long short-term memory (LSTM) architecture combined with feature selection algorithms (Li and Becker, 2021). Chang et al. built a hybrid model based on wavelet transform and ADAM-optimized LSTM neural network (Chang et al., 2019). Yang et al. combined convolutional neural network (CNN) and gated recurrent neural network (GRU) to form a Tri-Branch CNN-GRU (Yang and Schell, 2022a). Haoling et al. developed a quadruple branch, CNN-based autoencoder (QCAE) framework to address the issue of noise and high dimensional feature space (Yang and Schell, 2022b).

In the past few years, there also have been a lot of advancement in the development of general time series forecasting models with the heavy adoption of a new type of attention based architecture – transformers (Vaswani et al., 2017) that are firstly introduced for language models. Since both text and time series belong to the type of sequential data, transformers also quickly get extended to time series modeling tasks especially for long-horizon forecasting as the self-attention mechanism can capture long-term dependencies. Different variants such as Autoformer (Wu et al., 2021) and Informer (Zhou et al., 2021) have been developed to address the issue of sparsity, computational complexity and limitations of encoder-decoder structure that is intrinsic to the conventional transformer architecture. Additionally, Temporal Fusion Transformer (TFT) (Lim et al., 2021) introduced in 2021 was developed to provide both interpretability and long-horizon forecasting that deals with various types of data sources. Most recently, Channel-Independent Patch Time Series Transformer (PatchTST) (Nie et al., 2022) is demonstrated to outperform most state of the art methods on benchmark datasets through retaining local semantic information in the embedding by leveraging the patching idea introduced in VisionTransformer.

Despite the great success of using DNNs for electricity price forecasting, most methods do not highlight forecasts of extreme prices, which is especially important as such situations often implies great economic opportunities.

To address the problem of extreme price prediction, we introduced a new modeling framework (Figure 2) based on the latest TFT model that takes in (1) static (i.e. time-invariant) covariates: S (2) known future covariates (use both past and future value): $Z_{t-L:t+H}$ (3) exogenous variables that are only observed in the past: $X_{t-L:t}$ (4) exogenous variables in (3) forecasted into the future: $X^f_{t:t+H}$. With (2) and (4), we also trained a separate classification model to predict the likelihood of having extreme price for the entire forecasting period. Combining the likelihood prediction and the quantile output from the TFT model, we were able to perform a fusion strategy that further improves the model performance. The proposed framework is robust as it consistently outperform other methods under different market conditions.

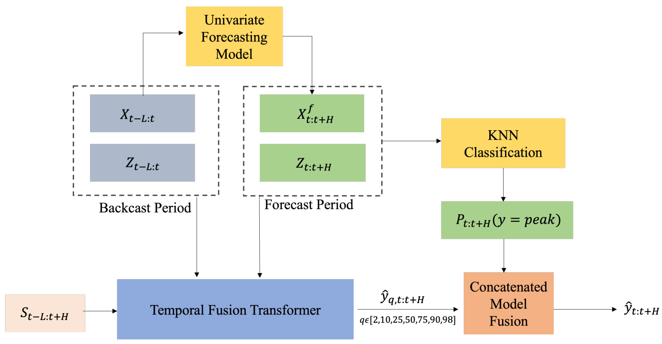


Fig. 2. Proposed Modeling Framework for Real-Time Electricity Price Prediction. y: real-time electricity price, P: probability of extreme price, q: quantile levels of the predicted y

2. DATA SETS

The hourly day-ahead and real-time electricity price for the Houston load zone is obtained from ERCOT. The historical weather data is acquired from the National Oceanic and Atmospheric Administration (NOAA). The station we use is the George Bush Intercontinental Airport (KIAH). The data collected ranges from Jan 1st, 2017 to December 31st, 2022.

2.1 Feature Selection

To accurately forecast real-time electricity prices, we incorporated features such as historical day-ahead prices ($/MWh), date information, weather information, grid load, energy index and natural gas prices. Specifically, date information includes 7 features such as hour of the day (HOD), day of the week (DOW), week of the year (WOY), month of the year (MOY), day of the year (DOY), seasonality and whether the specific day is a holiday. Weather information such as dew point (DP), temperature, precipitation (in), relative humidity (%), barometric pressure (inHg), wind speed (mph), visibility as well as dry bulb

(DBT) and wet bulb temperatures (WBT) were included (all temperatures are in unit of F). Figure 3 shows the price and some of the weather information from 2021. We can observe a significant price increase caused by the Texas winter storm in February 2021.

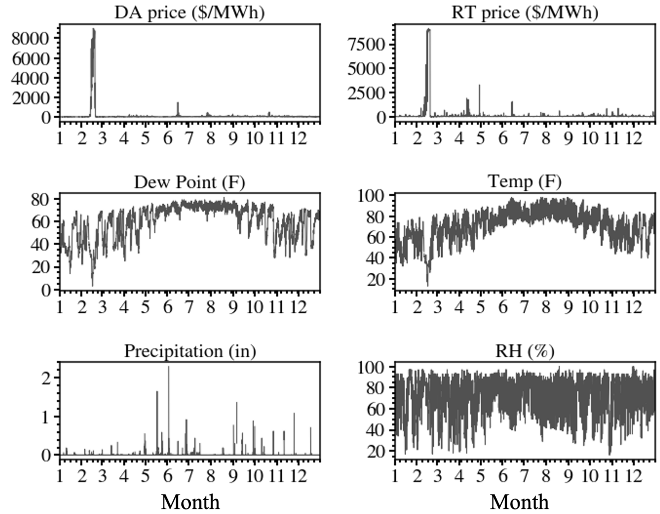


Fig. 3. Historical Electricity Price and Examples of Weather Data for the Houston Load Zone in 2021

For energy index (EI) and natural gas (NG) prices, they are only available on a daily and monthly basis, therefore a constant value throughout the day and the month respectively were used to convert these variables to the same hourly frequency as the rest of the dataset.

At each time step t, with a backcasting period (L) of 7 days and a forecasting horizon (H) of 24 hours (Figure 4), all the input features can be divided into three main categories:

- Static (time-invariant) Covariate (S): NG price
- Dynamic (time-varying) Covariate with future known ($Z_{t-L:t+H}$)
  - · Continuous: DA Price, EI
  - · Categorical: HOD, DOW, WOY, MOY, DOY, Seasonality, Holiday
- Dynamic (time-varying) Covariate with future unknown ($X_{t-L:t-1}$): DPT, Temperature, Grid Load, Wind Speed, Relative Humidity, Precipitation, Pressure, Visibility, DBT, WBT

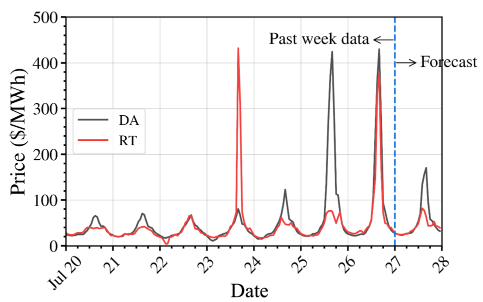


Fig. 4. Illustration of using Backcasting Period for Energy Price Forecasting

2.2 Data Split

The dataset used for this work consists of features mentioned above collected from 2017 to 2022 (6 years). We split the data into training and test set around 2022 September 1st which means roughly 4 month of data is used as the testing set to access model performance. Within the training set, a typical 5-fold cross validation is applied to select model hyper-parameters when necessary.

2.3 Prepossessing

To facilitate the training of machine learning methods, we perform the necessary pre-processing on all features. Since electricity prices do not follow a Gaussian distribution and contain extreme values (range from $<$ \$10 all the way to \$8000), an area (or inverse) hyperbolic sine variance stabilizing transformation (VST) (Uniejewski et al., 2017) is utilized here:

$$y_{vst} = \text{arcsinh}(y) = \log(y + \sqrt{y^2 + 1}) \tag{1}$$

Where y is the normalized original electricity price and y_{vst} is the price after VST. This transformation is very effective at reducing spike severity and stabilizing the variance.

Besides electricity price, the remaining modeling features are scaled to have a zero mean and unit standard deviation.

3. MODELING STRATEGY

As shown in Figure 2, the proposed framework consists of four modeling parts (1): Univariate Forecasting Model for Exogenous Variables; (2): The Temporal Fusion Transformer (TFT) model; (3): A K-Nearest Neighbour (KNN) based peak classification model and (4): a concatenated fusion between the quantile forecasting from (2) and the predicted peak likelihood from (3). The details of each part of the modeling framework is discussed below.

3.1 Exogenous Variable Forecasting

One of the key contribution of the proposed framework in this work is to include future forecasting of the dynamic covariates where the actual/measured future value is generally unknown ($X_{t-L:t-1}$). This is especially important as typically peaks or extreme prices happen on a very short time frame (i.e. 3-6 hours) compared to the forecasting horizon of 24 hours. Therefore, if only information from the past is available, it is very difficult to predict coming price peaks happening within the prediction horizon. Moreover, many of the exogenous variables especially the weather properties have significant seasonality thus having a dedicated forecasting model for these variables can help better capture their behavior in the next 24 hours.

The univariate Prophet forecasting model proposed by Taylor et.al (Taylor and Letham, 2018) from FAIR is used here. On a high level, the Prophet model uses a decomposable time series that that captures 3 key components in an additive fashion:

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t \tag{2}$$

where $g(t)$ captures the general non-periodic trend as a function of time, $s(t)$ accounts for the periodic changes at multiple resolution such as yearly, weekly and daily, $h(t)$ represents the effect of pre-programmed holidays with the flexibility to include custom known past and future events. Specifically for the trend model, a piece-wise Bayesian linear growth rate model is used with automatic detection of the change points using a sparse prior. More details of the model formulation can be found in the original paper (Taylor and Letham, 2018).

Example of the prophet model components for the Dew Point Temperature is shown in Figure 5. Clearly, the base growth rate is almost constant as quantified by the trend term with a strong seasonal/yearly periodicity behavior.

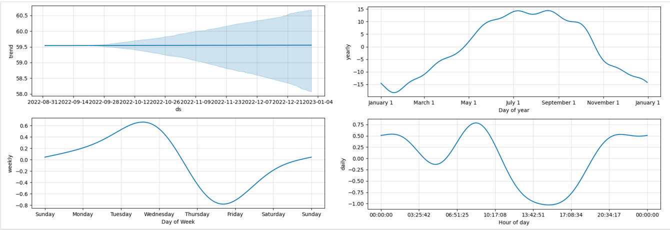
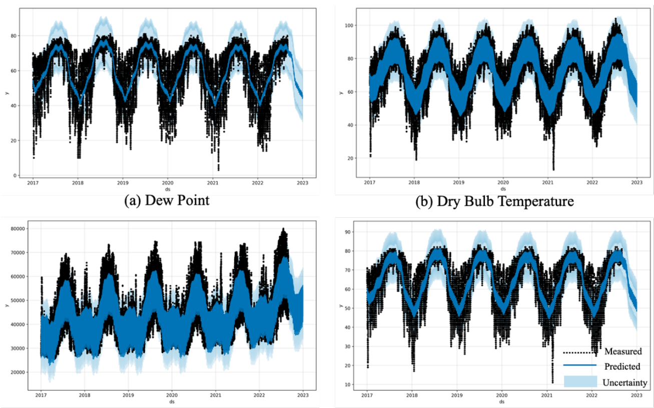


Fig. 5. Prophet Model Components of the Dew Point Temperature: Trend and Seasonality

Not all of the dynamic covariates with unknown future can be modeled well in this case using the prophet model. Pressure, precipitation and visibility were found to have large variation with little periodic trend. Therefore these variables were not forecasted. The remaining ones that show typical cyclic trends were well captured by the univariate forecasting models as shown in Figure 6. To make sure the forecasted variables are not underestimated, the forecasted upper bound instead of the mean is fed to the models discussed in the next section.



(a) Dew Point (b) Dry Bulb Temperature (c) Electric Grid Load (d) Precipitation

Fig. 6. Individual Univariate Model Fit and Forecast on Dynamic Covariates with Unknown Future

3.2 Temporal Fusion Transformer

Temporal Fusion Transformer (TFT) is first introduced by Bryan et al (Lim et al., 2021) in 2021. It is a multi-headed attention model that aims to provide interpretable insights based on the temporal dynamics of the multi-variate time series. The structure consists of many GRU based variable selection blocks before feeding into a LSTM encoder-decoder structure and makes use of a series of gating layers to skip over unused components to enable accurate forecasting performance over a wide range of situations. By default, the model uses quantile loss for training and produce quantile prediction at each time point of the forecasting horizon instead of a point estimate, which naturally provides uncertainty quantification to better capture the true price. TFT implementation in the pytorch-forecasting package is used for this study which by default outputs the 2nd, 10th, 25th, 50th, 75th, 90th and 98th quantile.

3.3 Extreme Price Classification

In practice, often times it is more important to know if an extreme price (or peak price) point is expected to show up in the next 24 hours rather than the exact numerical value of the extreme price. Therefore, on top of forecasting the price directly, a classification model is trained separately to predict the likelihood of expecting an extreme price for every hour of the next 24 hours in the prediction horizon. Here, extreme price is defined to be any electricity price above $150. Initially, a time series classification model is proposed to predict if at least one of the next 24 hour is expected to be an extreme price using the backcast 7 day period. However, with a single point prediction of the likelihood, it is not as straightforward to perform the fusion step with the 24 hour forecasted hourly quantile price. Therefore, a regular classification model that predicts if an extreme price is expected at every hour is trained using both the known and forecasted dynamic covariates ($Z_{t:t+H}$ and $X^f_{t:t+H}$) instead.

Alternatively, instead of extreme price, it might be more useful to know if the real-time electricity price will be above or below the given day-ahead price. In that case, the classification model can be modified to predict the likelihood of exceeding the provided day-ahead price. We have also tested that for our study and for this particular case, we do not see much difference in the final improved model performance compared to the extreme price classification model.

3.4 Concatenated Model Fusion

To further improve the performance of the forecasting model, a fusion strategy is carried out between the forecasted quantile prices and the predicted likelihood of observing an extreme price. Here, simple concatenated model fusion is adopted which means a regression model is fitted between the forecasted quantiles, day-ahead price and the predicted likelihood:

$$y_t = f(y_{q,t}, DA_t, P_t) \tag{3}$$

Various models such as Linear, LASSO, Ridge, Random Forest (RF) regressions were tried and most of them give very similar performance with RF models being slightly better. Therefore, a RF model is chosen for the concatenated model fusion.

4. RESULTS

Forecasting performance is discussed in this section. The impact of including forecasted dynamic covaraites as well as the fusion strategy is summarized by evaluating the model performance in terms of three commonly used metrics for time series forecasting: root mean squared error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) (Equation (4), (5), (6)).

$$RMSE = \sqrt{\frac{1}{H}\sum_{i=t}^{t+H}(y_i - \widehat{y}_i)^2} \tag{4}$$

$$MAE = \frac{1}{H}\sum_{i=t}^{t+H}|y_i - \widehat{y}_i| \tag{5}$$

$$MAPE(\%) = \frac{1}{H}\sum_{i=t}^{t+H}|\frac{y_i - \widehat{y}_i}{y_i}| \times 100 \tag{6}$$

All models' training and testing were performed on a Macbook with M1 Chip, 8G of memory and 8-core CPU.

4.1 Model Performance with Forecasted Future Covariates

TFT is selected for this study not only due to its ability to account for different types of data sources, but also due to its better baseline performance compared to other state of the art methods such as AE-LSTM and PatchTST. For these two models, an RMSE of 49 and 42 were obtained on the testing period, which is worse than the vanilla TFT model (RMSE of 38).

The performance of the TFT model with and without including the forecasted future covariates were summarized in table 1. Two examples of a 5-day period (a moving window of 24 hours is applied during forecasting) randomly selected from the testing period is shown in Figure 7. Overall, the model with the forecasted covariates performed better based on all three metrics. For certain regions such as region 1, the relative improvement of RMSE, MAE and MAPE are over 50%.

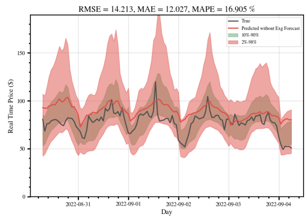
Region	RMSE		MAE		MAPE	
	Without	With	Without	With	Without	With
1	14.21	7.67	12.02	5.97	16.90%	8.04%
2	15.79	14.25	9.25	7.95	15.99%	13.85%
Overall	38.42	37.12	15.17	14.81	20.04%	18.04%

Table 1. Model Forecasting Performance with and without Forecasted Future Covariates

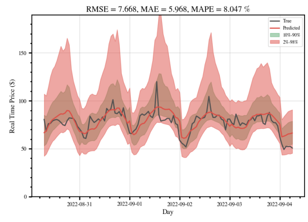
Additionally, with the forecasted quantile price, 84% of the time the 90th quantile contains the true real-time price, and this number increases to 97% if the 98th quantile is used. This uncertainty quantification provides valuable information for market participants in making more economically beneficial decisions of wholesale power the day before the operating date.

4.2 Extreme Price Classification Model Performance

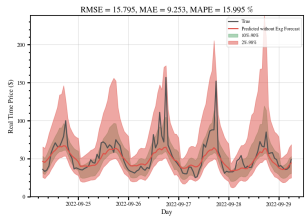
As discussed in section 3, a K-nearest neighbour (KNN) classification model is suggested with number of neighbours = 10 selected from a 5-fold cross-validation on the training set. KNN is selected among various classification model types due to its simplicity and overall performance. The performance of the model on the test set is shown in table 2. Overall, the accuracy for correctly predicting an extreme price (a.k.a price >$150) is only slightly over 50% as the dataset is extremely unbalanced with roughly 1:10 ratio between the number of extreme peaks and non-peaks. Other types of classification models such as random forest or ridge classifier with ROCKET (convolutional kernels) transformation (Dempster et al., 2020) of the features were also tried but no better performance is obtained.



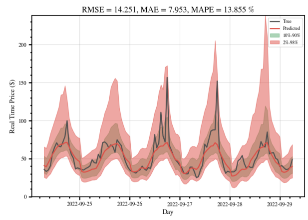
(a) Without Forecasted Covariates, Region 1



(b) With Forecasted Covariates, Region 1



(c) Without Forecasted Covariates, Region 2



(d) With Forecasted Covariates, Region 2

Fig. 7. Model Forecasting with and without Forecasted Future Covariates, 2 Examples of a 5-day Period from the Testing Set

True Label/Predicted Label	Peak	None-Peak
Peak	0.51	0.49
None-Peak	0.04	0.96

Table 2. Accuracy of KNN Model on Test Set

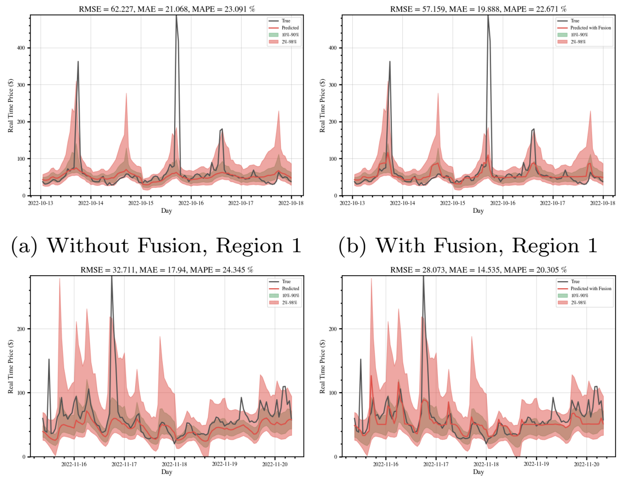
With the rather poor model performance, it might not be desired to include the classification model as part of the proposed framework. However, we have tested if the accuracy of the extreme price classification can be improved from the current 51% to 80-90%, it will drastically increase the final model performance after fusion. This is because as mentioned above, the upper quantile forecasting contained the true real time price for more than 97% of the time. Hence if a better knowledge regarding when to use the upper quantile for prediction is present, the proposed framework would be more effective.

4.3 Concatenated Model Fusion Performance

The performance of the proposed framework with and without fusion were summarized in table 3 (TFT model with future covariates were used for both cases). Two examples of a 5-day period selected from the testing set is shown in Figure 8. Overall, the model with the concatenated fusion performed better based on all three metrics. For certain regions such as region 2, the relative improvement of RMSE, MAE and MAPE are over 10%.

Region	RMSE		MAE		MAPE	
	Without	With	Without	With	Without	With
1	62.22	57.16	21.07	19.89	23.09%	22.67%
2	32.71	28.07	17.94	14.53	24.34%	20.3%
Overall	37.12	35.57	14.81	12.84	18.04%	17.21%

Table 3. Model Forecasting Performance Without and With Fusion



(a) Without Fusion, Region 1 (b) With Fusion, Region 1

(c) Without Fusion, Region 2 (d) With Fusion, Region 2

Fig. 8. Model Forecasting with and without fusion, 2 examples of 5-day period from the Test Set.

5. CONCLUSION

A new modeling framework that takes forecasted upper bound of exogenous variables to predict both real-time prices with uncertainty using a TFT model and likelihood of having extreme prices in the forecasting horizon is proposed. Additionally, a model fusion strategy is applied on top of the model predictions to further increase the framework's capability of accurately forecasting extreme prices. Our proposed method showed better performance on the same testing period in all metrics. For certain regions, more than 50% relative improvement is observed with the inclusion of forecasted covariates. With the proposed fusion strategy, an additional 10% relative improvement is achieved for certain testing period. Moreover, we were able to quantify the uncertainty of the predictions leveraging the quantile output from the TFT model and found that almost 97% of the time, the 98th quantile of the forecasting horizon contains the actual real time price. Our proposed framework provided an accurate and robust approach for forecasting extreme electricity prices that could have significant economic benefits to electricity market participants. Even though the study is conducted using ERCOT data from only the Houston load zone, the framework can be applied to other ERCOT regions.

REFERENCES

Chang, Z., Zhang, Y., and Chen, W. (2019). Electricity price prediction based on hybrid model of adam optimized lstm neural network and wavelet transform. *Energy*, 187, 115804.

Dempster, A., Petitjean, F., and Webb, G.I. (2020). Rocket: exceptionally fast and accurate time series classification using random convolutional kernels. *Data Mining and Knowledge Discovery*, 34(5), 1454–1495.

Garcia, R.C., Contreras, J., Van Akkeren, M., and Garcia, J.B.C. (2005). A garch forecasting model to predict day-ahead electricity prices. *IEEE transactions on power systems*, 20(2), 867–874.

Kamat, R. and Oren, S.S. (2002). Multi-settlement systems for electricity markets: Zonal aggregation under network uncertainty and market power. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences*, 739–748. IEEE.

Lago, J., Marcjasz, G., De Schutter, B., and Weron, R. (2021). Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark. *Applied Energy*, 293, 116983.

Li, W. and Becker, D.M. (2021). Day-ahead electricity price prediction applying hybrid models of lstm-based deep learning methods and feature selection algorithms under consideration of market coupling. *Energy*, 237, 121543.

Lim, B., Arık, S.Ö., Loeff, N., and Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4), 1748–1764.

Nguyen, H. and Hansen, C.K. (2017). Short-term electricity load forecasting with time series analysis. In *2017 IEEE International Conference on Prognostics and Health Management (ICPHM)*, 214–221. IEEE.

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

Taylor, S.J. and Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72(1), 37–45.

Uniejewski, B., Nowotarski, J., and Weron, R. (2016). Automated variable selection and shrinkage for day-ahead electricity price forecasting. *Energies*, 9(8), 621.

Uniejewski, B., Weron, R., and Ziel, F. (2017). Variance stabilizing transformations for electricity spot price forecasting. *IEEE Transactions on Power Systems*, 33(2), 2219–2229.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.

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

Yang, H. and Schell, K.R. (2022a). Ghtnet: Tri-branch deep learning network for real-time electricity price forecasting. *Energy*, 238, 122052.

Yang, H. and Schell, K.R. (2022b). Qcae: A quadruple branch cnn autoencoder for real-time electricity price forecasting. *International Journal of Electrical Power Energy Systems*, 141, 108092. doi: https://doi.org/10.1016/j.ijepes.2022.108092.

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