Stochastic Energy Market Price Forecasting with Recurrent Neural Networks

Elmore, Clay
celmore1@nd.edu
University of Notre Dame

Kopp, Grace
gkopp@nd.edu
University of Notre Dame

ABSTRACT

Energy market forecasting is a necessary for the functional operation of electricity markets across the globe. Because of large price swings in energy prices, sometimes as much as 4 times, there is a large economic opportunity for accurate forecasting to incur profit for energy users. However, these prices are notoriously hard to predict accurately because of the large stochastic nature inherent to the problem. In this paper we use Recurrent Neural Networks to forecast Day Ahead Market (DAM) prices in the California Independent System Operator (CAISO) in 2015. The study consists of forecasting 24 hour increments for a week in January for 50 different energy providers in the CAISO. Furthermore, a full year of prices is forecasted for a single Californian energy resource for all of 2015. We show that a full year's worth of forecasting in California has an average error of around 10%. This is on par with current standards in traditional forecasting methods. It is also shown that Recurrent Neural Networks are more resistant to changes in market dynamics involving seasonal changes than traditional methods such as ARIMA or GARCH.

KEYWORDS

machine learning, stochastic forecasting, recurrent neural networks, energy markets

ACM Reference Format:
Elmore, Clay and Kopp, Grace. 2019. Stochastic Energy Market Price Forecasting with Recurrent Neural Networks. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Energy markets are present all around the globe. Prices in these markets are extremely volatile due to a variety of different factors including but limited to market demand, erratic weather, natural resource availability, and so much more. Because of the large swings in these prices, participating in these energy markets presents an economic opportunity to both energy buyers and sellers alike. For an energy provider, knowing when the price of energy will be high is critical to knowing when to use valuable resources to produce energy. For an energy user, knowing when to buy energy on the market at a low price can be the difference between an industrial user being profitable or going under. With these considerations in mind, it is clear that energy price forecasting is a worthy venture for academics to study because of its direct applicability to the commercial world.

In this paper, we will explore a specific energy market: the California Independent System Operator (CAISO). The CAISO is an enormous, complex energy market system in which energy buyers and sellers schedule a multitude of energy related transactions. The CAISO had some 40 billion dollars in sales, and more than 260,000 gigawatts sold in 2015[2]. This complex system makes for a data-rich problem worthy for study by academics and industry professionals alike. A particularly interesting problem in the CAISO is the forecasting of Day-Ahead Market (DAM) prices which are notoriously stochastic. DAM prices are used to forecast energy prices a day in advance which energy buyers and sellers then use to schedule when they will buy and sell energy on the CAISO market. It has been shown that accurate forecasting is absolutely critical to profiting off DAM market prices[5].

Although prices in the CAISO DAM are quite volatile, they have very clear periodic tendencies. Figure 1 shows an example week of energy prices for an energy resource in the CAISO DAM for the first week of 2015. Note how the price is clearly periodic, but also has stochastic tendencies as well. Furthermore, Figure 1 shows the economic opportunity available to sell at high prices and buy at low prices. Some price swings are almost double the price. An accurate forecast would allow for large amounts of profit to be extracted from these systems.

Many standard time series forecasting methods have been used to predict DAM prices in the CAISO and other energy markets across the globe. For example, standard autoregressive integrated moving average (ARIMA) forecasting methods have been used to differing degrees of success in the CAISO market and the Spanish DAM. Conejo et al. have used ARIMA models to predict prices in the Spanish market with a 10% error for the 2002 market [4]. Furthermore, ARIMA has been shown to generate forecasts in the 2000 CAISO DAM with an error of around 13% [6]. Another standard method for time series forecasting that has been used in the Generalized autoregressive conditional heteroskedasticity (GARCH) method. Similar to ARIMA models, an error rate of around 10% for the Spanish DAM has been reported when using GARCH [6]. More advanced hybrid methods have also been used for DAM forecasting. These methods include the usage of wavelet decompositions in order remove noise from data as well as the incorporation of ARIMA models in conjunction with GARCH methods [1, 3, 7, 8].

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

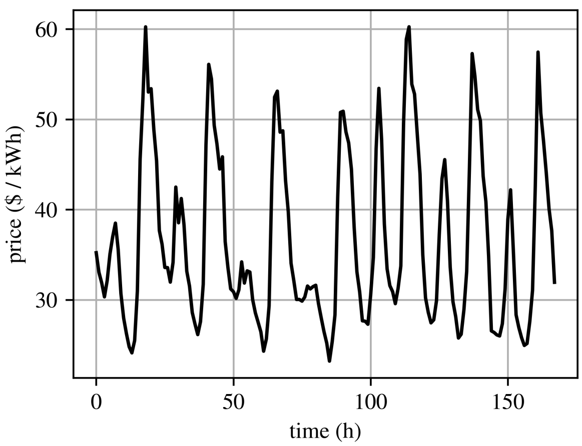

Figure 1: An example of time series behavior observed in the CAISO from January 1st, 2015 to January 7th, 2015.

These various hybrid methods have error rates reported that range from as low as 1% to as high as 25%.

Recent advances in Recurrent Neural Networks (RNN) in the past 5 to 10 years have motivated this project to investigate their potential ability to predict CAISO DAM prices. The remainder of the paper will be organized in the following manner.

- Section 2: A problem statement along with the design of the RNN, an explanation of the dataset, and software implementation details.
- Section 3: Results of the forecasting performance of our RNN forecaster is presented.
- Section 4: A discussion of how RNN forecasting compares to existing methods and other insights from this project.
- Section 5: Major findings from the paper are summarized.

2 MODELING APPROACH

2.1 Problem Statement

2.2 Neural Network Methodology

Outline:

- Basic explanation of an RNN.
- How does an RNN relate to energy prices.
- Figure: How does our RNN look in relation to energy prices.
- What are the specifics of the RNN we use.
- What was the methodology we used to determine the this design.
- INCLUDE HOW ERROR IS CALCULATED!
- Probably should cite the towards data science link we used.

2.3 Data Sources

The dataset for this project was kindly provided by the Dowling Lab for Uncertainty Quantification and Mathematical Optimization at the University of Notre Dame in the Department of Chemical and Biomolecular Engineering. The dataset is comprised of over 6500 energy vendors, called nodes, participating in the CAISO. There is a full year's worth of data for each node in 2015, and price measurements are recorded at 1 hour intervals. Latitude and longitude coordinates are known for around 2000 of these nodes for visualization purposes. Because of the limited timeline of this project, only 50 of these nodes will be used, and only a single week of data in January 2015 will be used for testing purposes. This is a reasonable thing to do as many papers report error values for single weeks in Janurary [4, 6, 7]. However, a single node will be used to predict prices 24 hours into the future for a full year in order to give a more accurate representation of error quantification across a larger time span. Furthermore, these papers only report results for single nodes in the DAM. Because our results include 50 nodes, it can be concluded that the results obtained for this dataset are statistically significant.

2.4 Software Implementation

Outline:

- Tensorflow
- CRC GPUs
- Code available at: GitHub
- Find the CRC thing we have to mention

3 RESULTS

Now that we have shown how we are going to forecast market prices with a RNN, we will now enter the results section where we analyze the effectiveness of RNNs.

It is important to notice that variations in market prices can significantly alter the ability of RNNs to accurately predict DAM prices. Figure 2 was constructed by training a RNN with 336 hours of training data, or two weeks. This 336 hour training set was then split into 265 training instances. These instances consisted of one 24 hour input, and the corresponding 24 hour output that would be the next day's market data. Two different forecasts are shown in Figure 2 to illustrate how different inputs to the RNN can drastically alter its ability to forecast market prices. In the first plot (top), there is well behaving periodic wave input to the RNN. The output is a forecast that is extremely successful with a small error of only 4.4%. However, the second plot (bottom) shows that when there is an input that does not as

Training Data Used	1 week	2 weeks
Median	14.49	N/A
Mean	14.34	N/A
Standard Deviation	3.13	N/A

Table 1: Aggregated statistical performance for a single week in January when 1 week of training data was used vs. 2 weeks of training data (2 week data is running currently).

4 DISCUSSION

Figures and Tables we will need:

- Figure: Graphical comparison of how backcasting does not do what the RNN can do.
- Figure/Table: Comparison to techniques that have been mentioned earlier in the paper and backcasting.

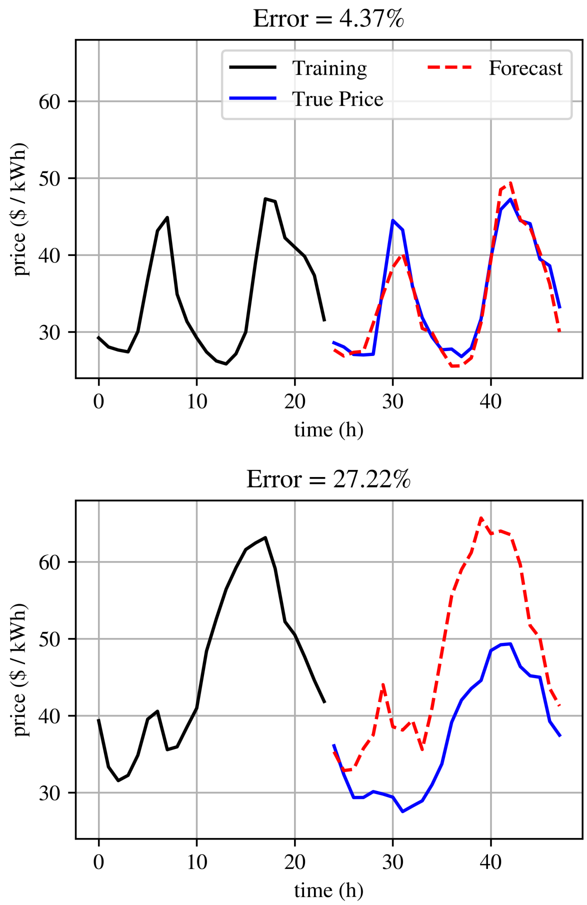

Figure 2: Example price forecasting using our RNN.

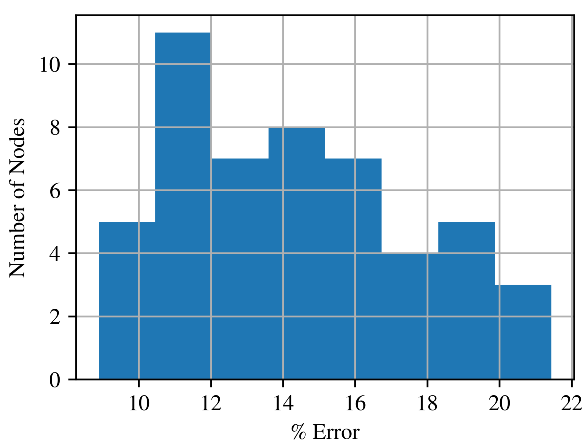

Figure 3: Histogram of all nodes.

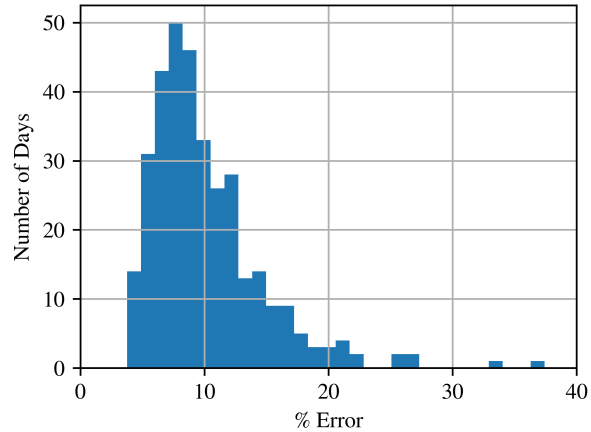

Figure 4: Histogram of a single node run through the whole year.

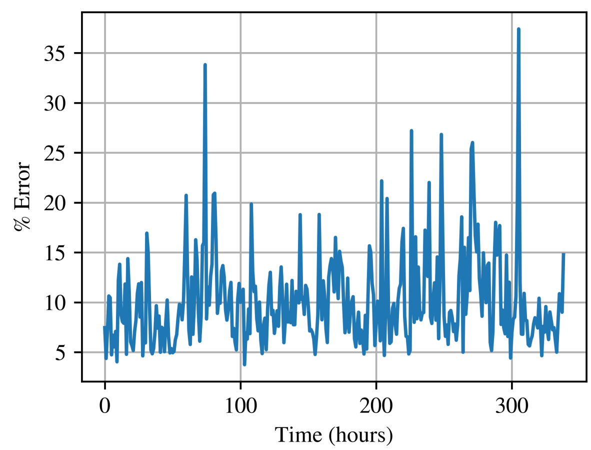

Figure 5: Time series plot of a single node run through the whole year.

Quarter	All	1	2	3	4
Median	8.89	8.39	9.34	9.29	8.80
Mean	10.14	9.60	9.67	10.53	10.74
Std.	4.59	4.78	3.03	4.67	5.41

Table 2: Aggregated statistical performance for a single node in 2015.

5 CONCLUSIONS AND FUTURE WORK

REFERENCES

[1] Nima Amjady and Farshid Keynia. 2008. Day ahead price forecasting of electricity markets by a mixed data model and hybrid forecast method. *International Journal of Electrical Power & Energy Systems* 30, 9 (nov 2008), 533–546. https://doi.org/10.1016/J.IJEPES.2008.06.001

[2] California. [n. d.]. EIA - State Electricity Profiles. https://www.eia.gov/electricity/state/california/

[3] A.J. Conejo, M.A. Plazas, R. Espinola, and A.B. Molina. 2005. Day-Ahead Electricity Price Forecasting Using the Wavelet Transform and ARIMA Models. *IEEE Transactions on Power Systems* 20, 2 (may 2005), 1035–1042. https://doi.org/10.1109/TPWRS.2005.846054

[4] Antonio J. Conejo, Javier Contreras, Rosa Espínola, and Miguel A. Plazas. 2005. Forecasting electricity prices for a day-ahead pool-based electric energy market. *International Journal of Forecasting* 21, 3 (jul 2005), 435–462. https://doi.org/10.1016/J.IJFORECAST.2004.12.005

[5] Alexander W. Dowling, Ranjeet Kumar, and Victor M. Zavala. 2017. A multi-scale optimization framework for electricity market participation. *Applied Energy* 190 (mar 2017), 147–164. https://doi.org/10.1016/J.APENERGY.2016.12.081

[6] R.C. Garcia, J. Contreras, M. VanAkkeren, and J.B.C. Garcia. 2005. A GARCH Forecasting Model to Predict Day-Ahead Electricity Prices. *IEEE Transactions on Power Systems* 20, 2 (may 2005), 867–874. https://doi.org/10.1109/TPWRS.2005.846044

[7] Zhongfu Tan, Jinliang Zhang, Jianhui Wang, and Jun Xu. 2010. Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models. *Applied Energy* 87, 11 (nov 2010), 3606–3610. https://doi.org/10.1016/J.APENERGY.2010.05.012

[8] Yudong Wang and Chongfeng Wu. 2012. Forecasting energy market volatility using GARCH models: Can multivariate models beat univariate models? *Energy Economics* 34, 6 (nov 2012), 2167–2181. https://doi.org/10.1016/J.ENECO.2012.03.010