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

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[3]. 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[7].

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 [5]. Furthermore, ARIMA has been shown to generate forecasts in the 2000 CAISO DAM with an error of around 13% [8]. 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 [8]. 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, 4, 9, 10].
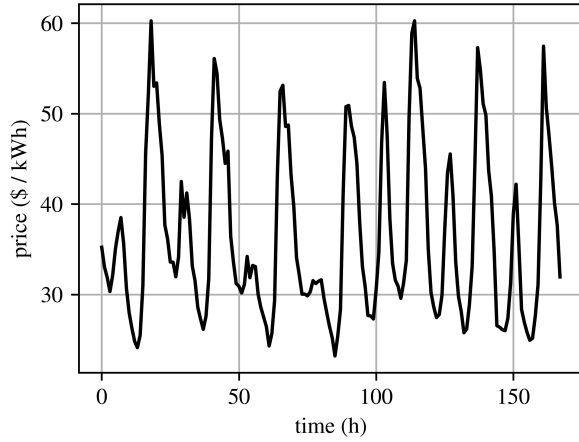
Figure 1: An example of time series behavior observed in the CAISO from January 1st, 2015 to January 7th, 2015.

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 along with a discussion of how RNN forecasting compares to existing methods and other insights from this project.
- Section 4: Major findings from the paper are summarized.

2 MODELING APPROACH

2.1 Problem Statement

We set out to use a recurrent neural network to forecast Day Ahead Market (DAM) prices of the California Independent System Operator (CAISO) in 2015. Specifically, we wanted to forecast 7 24 hour increments for a week in January for 50 different energy providers in the CAISO.

We also set out to get in depth information on one energy provider so we forecasted a full year (2015) of prices for one energy provider.

2.2 Neural Network Methodology

We used a recurrent neural network (RNN) for our time series prediction. We chose to use this architecture because RNNs are made to deal with sequential information. In other neural network architectures, it is assumed that all inputs are independent which would not be useful for a time series prediction where the order of the inputs is extremely important to accurately predict outputs. RNNs accomplish this by keeping a "memory" of inputs they have

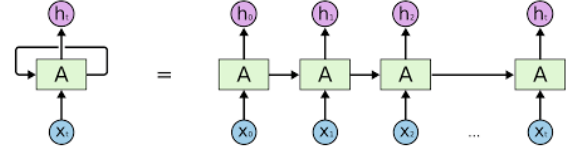seen at previous time steps to make decisions about future time steps [6].

Figure 2: An unrolled recurrent neural network cell [6].

Figure 2 depicts an unrolled RNN. For a sequence with t time steps, we will unroll our network into t corresponding layers. The blue nodes labelled x1, x2, etc. are our inputs at time step one, two, and so on. In the case of this paper, these would be energy prices at hour 1, hour 2, etc. The green nodes labelled A are our hidden states at each time step. This is how the network is able to have a memory. It is calculated with the previous hidden step and the current input. Because it takes this previous hidden step into account, it captures information about time steps that have already been processed. Finally, the purple nodes labelled with h's are the outputs at each time step that are calculated from the memory in their associated hidden node. For us, these outputs are the energy prices we are predicting for future time steps.

Unfortunately, basic RNNs are not perfect for this task because of the vanishing gradient problem. RNNs are trained with back propagation and gradient descent learning methods. When dealing with a huge number of time steps, as is the case in our dataset, calculated gradients can become so small that the neural network is unable to continue to train.

Our solution to this problem was using a Long-short Term Memory network (LSTM). These are a type of neural network that have input gates, output gates, and forget gates that give the network the ability to keep some information while forgetting other information. This cuts down the amount of data used and solves the vanishing gradient problem [6].

We wanted to be able to forecast a days worth of energy prices in advance or 24 output time steps. To accomplish this, we created a Multi-Step LSTM where our output was a vector of 24 prices corresponding to the 24 time steps in the future we predicted [2]. Our training data was the 168 time steps (one weeks worth) of data that came directly before the 24 hour testing period.

In order to get an accurate view of our network and match other current evaluation models for energy price forecasting, for each of the fifty nodes we analyzed, we performed a rolling horizon analysis to predict seven days worth of data. In other words, first we used t_0 through t_{167} to predict t_{168} through t_{191}, then we used t_{24} through t_{191} to predict t_{192} through t_{216}, and so on five more times until we had seven days of predictions or 168 hours.

To calculate error, we performed a standard average percent error calculation on our predicted prices vs the actual price. We chose percent error in order to be able to easily compare our results to other energy forecasting results on the same data from different prediction methods.

$$error = \frac{1}{7} \sum_{i=0}^{6} \frac{1}{24} \sum_{j=0}^{23} \frac{|p_{i,j} - a_{i,j}|}{a_{i,j}} \quad (1)$$

In the above equation $p_{i,j}$ is the predicted energy price on day i, hour j and $a_{i,j}$ is the actual energy price on day i, hour j.

As a supplemental inquiry, we also investigated forecasting a full year of energy prices for one energy provider. This was done almost identically, except for this study we used 336 input time steps as training to predict 24 hour time step outputs. We performed this calculation 356 times to see how our RNN performs on more long term data. Error was calculated using the same method above except for 365 days instead of 7:

$$error = \frac{1}{365} \sum_{i=0}^{364} \frac{1}{24} \sum_{j=0}^{23} \frac{|p_{i,j} - a_{i,j}|}{a_{i,j}} \quad (2)$$

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 January [5, 8, 9]. 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

This study was implemented using Keras with a TensorFlow backend. Specifically, we used the sequential library in the Keras packaged to define and train our LSTM model. We chose to use Keras because it is allowed us to access the deep learning resources of TensorFlow in an easy to use API. Because we were dealing with a huge amount of data, our study had a very high computation time. To make this more manageable, we submitted our code as jobs to the University of Notre Dame Center for Research Computing GPUs. This allowed faster and more convenient processing than if we were to run our model locally. Our entire code suite is available at https://github.com/celmore25/neural_net_predictions. This includes our class implementation of the model, our data, our evaluation tool, and every other piece of code we wrote in pursuit of completing this study. This research was supported in part by the Notre Dame Center for Research Computing through the use of available GPU systems.

3 RESULTS AND DISCUSSION

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.

3.1 Example Forecasts

It is important to notice that variations in market prices can significantly alter the ability of RNNs to accurately predict DAM prices. Figure 3 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 3 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 behave as well. Without the direct sinusoidal behavior, the RNN struggles to accurately predict the price at an error of 27%. Based on this analysis, it will be important to look at aggregated results from more than two forecasts in order to show how well an RNN actually performs in forecasting the CAISO DAM.

3.2 Multiple Node, Single Week Forecasts

With Figure 3 motivating the exploration into aggregated results from the dataset, Figure 4 shows a histogram of the forecasting results of 50 nodes in the CAISO for a single week in January. This histogram was created by training 50 separate RNNs on 168 hours of training data for each individual node. These RNNs then predict 24 hours into the future and an error measurement is taken based on the true data. The horizon is then rolled into the future for 7 days in a row, and the error measurements for a week are averaged to give the results in Figure 4. The histogram shows the error is reasonably dependent on the node being forecasted. This is most likely due to different spatial dependencies in price points across California. These influences can include weather, natural resource availability, etc. Because of these natural variations, it is expected that each node has a very different performance.

NOTE TO THE REVIEWER: We are currently testing how inclusion of more training data (336 hours) improves our forecasting. We will have another histogram here showing these results at a later date.

Table 1 provides some aggregated statistics for the single week performance of a RNN across 50 different nodes. An error rate of 14% is similar to literature values reported for certain single weeks in the CAISO from various literature findings that were mentioned in the introduction. However, this is not beating current standards of forecasting with traditional ARIMA models. In order to look further into how effective RNN forecasting is, a single node is investigated more in depth in the next section.

3.3 Single Node, Full Year Forecasts

As previously mentioned, it is desired to look at a single node's performance in RNN forecasting for the whole year in 2015. Figure 5 was constructed by using 336 hours of rolling training data for
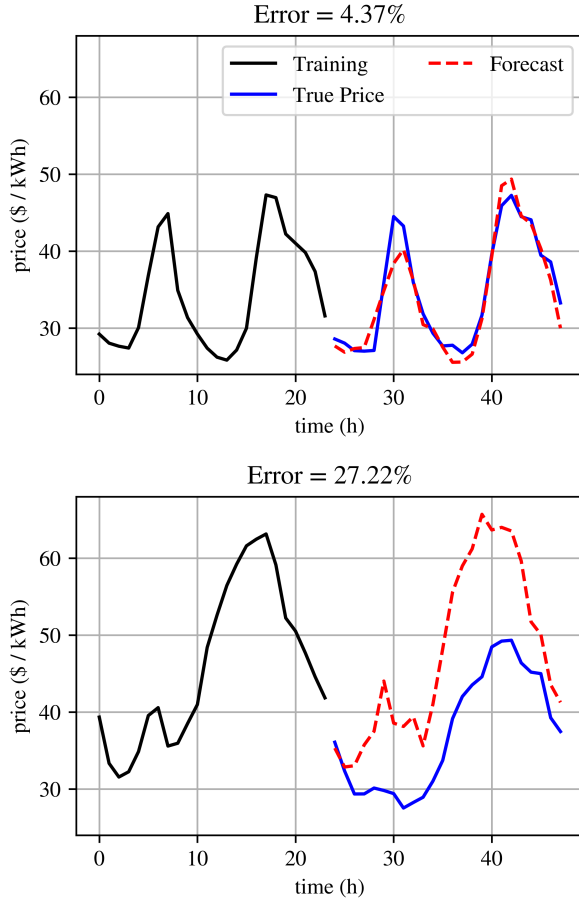
Error = 4.37%

Error = 27.22%

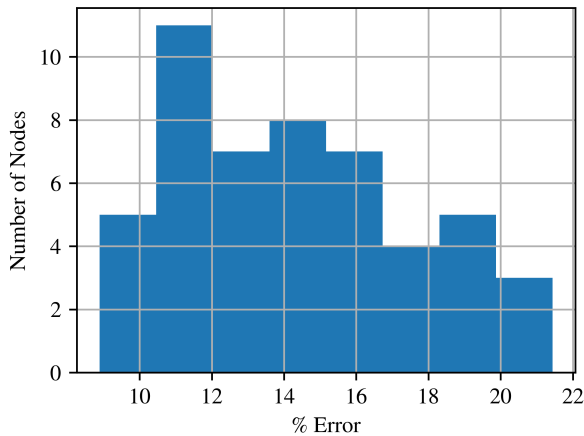**Figure 3: Example price forecasting using our RNN.**

Figure 4: Histogram of all nodes' results at forecasting a week of data in January 2015.

a single node in the CAISO. A different RNN was trained for each

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).

rolling horizon, and the subsequent 24 hour forecast was recorded. The histogram shows the error rate as a function of the number of times it occurred. A first observation on Figure 5 is that the dataset is very skewed to the right. From a market perspective, this means the RNN forecasts were generally quite accurate, but there were some large errors that did occur. It should also be noted that a large outlier at around 200% is not shown in this graph. However, the overall accuracy of the test was quite good. The median error for the full year was only 9% which is lower than previously reported for forecasting the CAISO via ARIMA and GARCH modeling [5, 8].
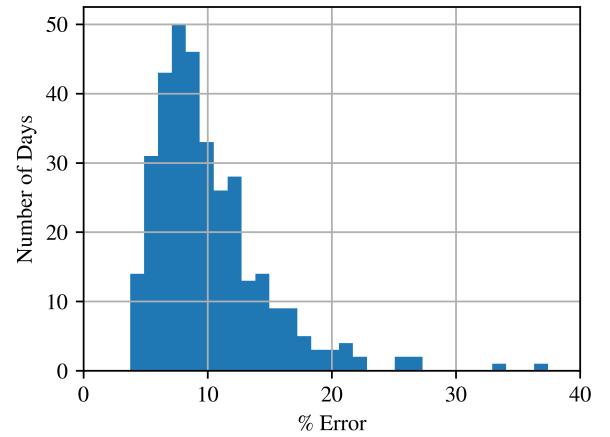

Figure 5: Histogram of a single node run through the whole year.

An important thing to look at in the CAISO market is the tendency for the market to change between the seasons. It is well documented that price ranges are quite different in the four quarters of the year. To connect this observation with how well our RNN performs, Figure 6 plots the error of each day in forecasting across the whole year of 2015. An observation made from this graph is that the error does not seem to be very dependent on the seasonality of the data. There are some abrupt error jumps, but those are most likely due to natural market disturbances, not seasonal variations.

Table 2 gives some aggregated results on the results shown in Figures 6 and 5. The qualitative observation that seasonality is not very important to RNN forecasting is held up by this table. A swing of only 1% error between seasons in both the median and the mean error shows that the RNN is very resistant to seasonal market changes. This bodes well for the usage of RNNs in further forecasting of DAM prices.
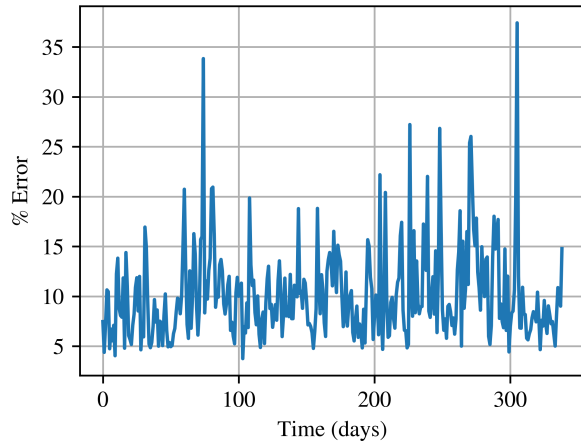
Figure 6: Time series plot of a single node run through the whole year.

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.

3.4 Key Insights

Now that all of these tests have been run and basic analysis has been performed, it is important to point out some key insights that can be determined. First, we have shown that RNN performance is extremely dependent on the input to the learning algorithm. If the algorithm is trained with well behaving market prices, then a poorly behaving market price is input into the network, the forecast can be expected to be poor. However, it is shown that using more market data can take decrease average error from 15% to less than 10%. Another important insight gained from this study is that RNN forecasting is quite resilient to seasonal changes in market dynamics. This is an interesting result as many literature finding suggest that seasonal differences are a large contributor to traditional forecasting method failures [5, 8]. With these things in mind, RNNs clearly show a lot of potential for forecasting DAM prices in the CAISO and potentially other markets.

4 CONCLUSIONS AND FUTURE WORK

In this paper we have shown that RNNs are well suited for energy market price predictions in the CAISO DAM in 2015. We report a median error of only 9% can be achieved using a rolling forecast for a single energy resource in California. This compares well to traditional price forecasting methods such as ARIMA and GARCH. Furthermore, we show that these forecasts can be extended to multiple resources across the CAISO for a single week analysis in January 2015. Lastly, we show that RNN forecasting is reasonably resistant

to market dynamic changes in seasonal changes in the CAISO.

Although the results from this work have been promising, we would like to pursue a few more studies in future work. First and foremost, it would be interesting to repeat the full year testing for more than one node in 2015. This would give a more wholistic view of RNN forecasting performance. Also, we would like to explore the possibility of training RNNs with data from multiple nodes at once in order to be able to forecast along multiple nodes instead of just one. We believe these future ventures would be helpful in making a more compelling argument for the usage of RNNs in energy market forecasting in the future.

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] Jason Brownlee. 2018. How to Develop LSTM Models for Time Series Forecasting. https://machinelearningmastery.com/how-to-develop-lstm-models-for-time-series-forecasting/
[3] California. [n. d.]. EIA - State Electricity Profiles. https://www.eia.gov/electricity/state/california/
[4] 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
[5] 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
[6] Tathagat Dasgupta. 2018. Time-Series Analysis Using Recurrent Neural Networks in Tensorflow. https://medium.com/themlblog/time-series-analysis-using-recurrent-neural-networks-in-tensorflow-2a0478b00be7
[7] 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
[8] 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
[9] 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
[10] 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