A GARCH Forecasting Model to Predict Day-Ahead Electricity Prices

Jose A. Maldonado, Quentin Cardone

*Abstract*-With the information age upon us, the demand for electricity all over the world is increasing. Determining the price of electricity is necessary for both long-term contracts and companies trying to maximize their gain. Due to electricity markets being unpredictable, new methods need to be discovered that can deal with the rapid variation of price in such a short amount of time. The purpose of this paper is to introduce Generalized Autoregressive Conditionally Heteroskedastic(GARCH) forecasting, which is derived from Autoregressive Conditionally Heteroskedastic(ARCH), to predict next-day electricity prices. In this paper, ARCH is introduced to better understand GARCH and its methodology. Along with the introduction to GARCH, we will also focus on estimating our parameters and validating our model to obtain the most optimal forecast. To fully understand how powerful GARCH can be, we will use two different data sets, one from Compañia Operada del Mercado Español, which consists of hourly electricity prices from the months of June and October of 2000. The other dataset is the data provided by California ISO (CAISO) for May and June of 2018. Using GARCH on this data will let us see how the forecast model handles data that is specially volatile during seasonal periods of the year.

I. INTRODUCTION

An electricity market contains pool systems that allow companies to bid or trade for power and energy. Participants bid by price and quantity until the total demand is met, which causes volatility. From this, companies involved that are able to forecast these electricity prices can change their own production and prices to their favor. Another way electricity is traded is by the bilateral contract system. A buyer and seller determine a fixed price and amount to be traded at a specific time. Electricity markets are very complex and knowledge of future prices is extremely useful. Therefore, Generalized Autoregressive Conditionally Heteroskedastic(GARCH) processes will be used to forecast prices.

To understand GARCH models, certain basic terms and concepts must be clarified. First and foremost, GARCH is a process that is a part of time series analysis. In time series analysis, models use previous years to forecast a value. The whole point of the forecast is to predict the exact value meaning the error term of the model has a zero mean. However, these electricity markets are volatile, liable to change rapidly and be unpredictable, meaning the data has a high variance. This paper focuses on creating 24-hour ahead forecasts based on historical data from both Spain and California’s electricity marketsusing the GARCH methodology approach (Compañia Operadora del Mercado Español de Electricidad and CAISO from Energy Online).

II. GARCH METHODOLOGY

GARCH and Autoregressive Conditionally Heteroskedastic(ARCH) create models on the variance of the time series. ARCH literally uses an AR(p) model on the variance. An AR(p) model;

yt = c + ϕ1yt-1+ ϕ2yt-2 + … + ϕpyt-p + εt;

is a regression of the variable of interest against itself. In order to perform these processes on the data, the time series must be conditionally heteroskedastic(CH). Conditional heteroskedastic is when an increase in variance is correlated to a further increase of variance. Conditional heteroskedastic is not stationary because the variance is not constant over time and that is why an AR model is used to model the variance for ARCH. However, to determine if ARCH is needed, a model must first be fitted to the time series making the residuals appear to be discrete white noise. Once the model is fitted, the residuals must be squared and the ACF must be observed to determine if an ARCH model is needed (“Time Series Analysis for Financial Data VI”, 2017).

Going more in depth, a GARCH model is an ARMA (p, q) model applied to the variance of the time series. In other words, a GARCH (0, q) model is the exact same as an ARCH (q) model. An ARMA model adds moving average to model the variance of the actual process. An ARMA model is as follows;

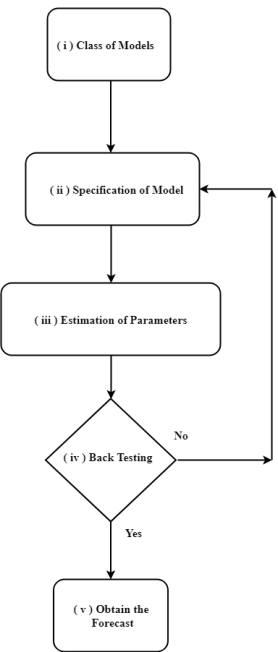
and the assumptions are that the error terms contain a zero mean and a constant variance. Since the data is heteroskedastic, the data does not necessarily need to have a constant variance. Moving forward, the GARCH model uses values of the past squared observations and past variances to model the variance at time *t* (Reinaldo et al., 2005)*.*A prime example would be;

y2t-1 +

which is a GARCH(1,1) model that is commonly used in financial time series analysis. As we can see from the above equation, sigma *t* squared is the predicted value that is derived from a constant () plus some alpha times the squared return of the previous period along with some beta times the squared volatility from the period before. To identify if an ARMA model is needed, the ACF and PACF of the squared time series will have decaying lags meaning a GARCH model is needed. Using the flowchart from Reinaldo et al. (2005) as a base line we created a new methodology model that worked best for our GARCH models since different software was used.

FIGURE 1

Flowchart of GARCH methodology using RStudio.



(i.) In the first step, careful inspection was done on each time series of our data set. Ultimately, GARCH was identified as a better method than ARIMA to model our forecast because of the characteristics of the datasets such as the nonzero mean and variance. (ii.) In the second step model specifications were created. By studying the PACF and ACF plots for all data sets different combinations of ARMA were tested alongside different GARCH models such as GARCH (1,1), GARCH (1,2), and GARCH(1,3). GARCH (1,3) was chosen along with ARMA(1,1) for three of the four data sets. For the California Data in May, GARCH(1,3) was used with ARMA (1,3) so it can better handle the abnormal spikes in the data. (iii.) Once we obtain our model specifications our parameters were obtained using a package called rugarch in RStudio, Table I contains our parameters. (iv.)To validate our chosen model, a number of tests were conducted such as Ljung- Box test, Lagrage Multiplier Test as well as a train and test set were created using the rugarch package. The train and test set known as GARCH Roll can be seen in Figures 2-5 where the last week of the months were used to test for accuracy. (v.) Once all the steps were followed and tested several times with different configurations, a 24 - hour forecast was obtained using the ugarchboot function, which is also available in the rugarch package.

III. RESULTS

Applying the GARCH processes to the Spanish electricity market for the months of October and June in 2000 and the Californian market for May and June in 2018. ARIMA was also used to for forecasting but it did not perform well when there were price spikes in the data. A model was created for each month of each market. Table 1 explains that for all of our models, a GARCH (1, 3) model was chosen with different values on the coefficients.

**Table I**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Spanish Market | | | Californian Market | | |
| June | | October | May | | June |
| *C* | 24.39909 | 33.04992 | *C* | 20.41034 | 22.16640 |
| **1 | 0.594340 | 0.00000 | **1 | 0.19423 | 0.031151 |
| **2 | 0.000000 | 0.00000 | **2 | 0.24855 | 0.000000 |
| **3 | 0.313229 | 0.71184 | **3 | 0.54944 | 0.000000 |
| ** | 0.069848 | 0.10949 | ** | 0.00000 | 0.528987 |
| GARCH(1,3) | | | GARCH(1,3) | | |

GARCH Model Coefficients

Although GARCH (1, 3) models are the best for all of them, the model, GARCH(1, 1) for June of the Californian market also produced great results. Since all of the other models needed GARCH (1, 3) to produce forecasts to our liking, a GARCH (1, 3) model was used for June as well for the Californian market even though GARCH (1, 1) still produce good forecasts and is simpler.

Figures 2-5 are graphs of the real prices versus forecasted prices to show the effectiveness and power of GARCH. The figures have an x-interval as the hour with the corresponding y-interval being the price, which is either €/MWh or $/MWh. By analyzing the Spanish market, the forecasts for both weeks of the months appear to be good with the forecasts coming close the true value. From all of the figures, figure 3 is very hard to distinguish the forecast from the real near the end of the week and this is due largely because of the high variation between hours.

FIGURE 2 FIGURE 3

Spanish Market from June 24-30, 2000 Californian Market from May 19-28, 2018

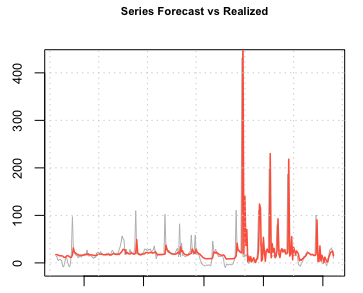
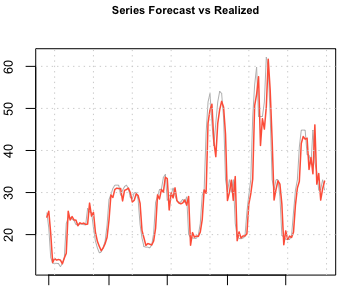
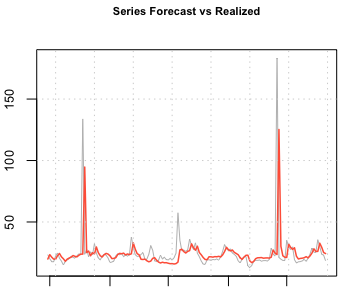
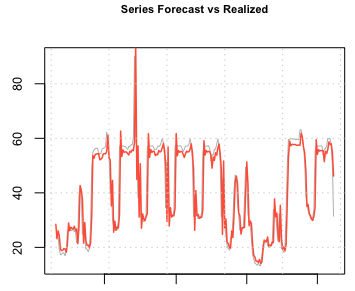


FIGURE 4 FIGURE 5

Spanish Market from October 21-28, 2000 Californian Market from June 24-30, 2018



Now, others performing a GARCH on this same timeline may receive better results because they may possibly use a larger dataset or a more representable dataset. Comparing these results to the same results obtained by Reinaldo et al. (2005), our results have far more error and this is because our dataset is smaller and all from the same month making our predictions weaker. In their study, they used the same data but for two whole years. Specifically, during the weeks mentioned in the figures, Reinaldo et al. used data points from as long as a year prior where the model could identify certain trends and seasonality’s. Our error is not terrible but their results show it could be better. The size of data is important when performing GARCH models because it can change the results significantly.

As for the figures regarding the Californian market, they perform well but because of the drastic price spikes all happening in rapid succession, the GARCH model was able to perform as well. Still, the results found were more than sufficient.

To exemplify the power of the GARCH models once again, tables 2-5 represent the real price of the data and the forecasted price from the model along with the amount of error. The average error for the Spanish market on June 29this 12.07% and for October 25th is 4.21%. Comparing these results to Reinaldo et al. (2005), the difference between the errors is about 3% meaning the forecasts obtained are not the best, but still usable considering the size of the data.

**Table II Table III**

Errors During Peak Hours of June 29 in Spanish Market Errors During Peak Hours of May 21 in the Californian Market

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Real Price | Forecasted price | Error (%) |  | Real Price | Forecasted price | Error (%) |
| Hour 9 | 29.51 | 27.25 | 7.65 | **Hour 9** | 16.28 | 16.72 | 2.63 |
| Hour 10 | 33.06 | 29.70 | 10.16 | **Hour 10** | 21.90 | 16.70 | 23.74 |
| Hour 11 | 48.08 | 33.19 | 30.96 | **Hour 11** | 26.93 | 18.19 | 32.45 |
| Hour 12 | 55.61 | 50.67 | 8.88 | **Hour 12** | 22.79 | 19.73 | 13.42 |
| Hour 13 | 59.80 | 53.31 | 9.33 | **Hour 13** | 20.33 | 18.87 | 7.18 |
| Hour 14 | 48.08 | 57.54 | 16.44 | **Hour 14** | 19.72 | 18.11 | 8.16 |
| Hour 15 | 48.08 | 41.19 | 14.33 | **Hour 15** | 19.09 | 17.87 | 1.15 |
| Hour 16 | 48.08 | 47.56 | 1.08 | **Hour 16** | 21.12 | 17.65 | 16.78 |
| Hour 17 | 51.54 | 45.08 | 12.53 | **Hour 17** | 18.99 | 18.14 | 4.47 |
| Hour 18 | 62.20 | 50.38 | 19.00 | **Hour 18** | 28.53 | 17.65 | 38.13 |
| Hour 19 | 59.80 | 61.69 | 3.06 | **Hour 19** | 33.95 | 20.10 | 40.79 |
| Hour 20 | 48.08 | 54.28 | 11.42 | **Hour 20** | 43.66 | 21.91 | 49.81 |

**Table IV Table V**

Errors During Peak Hours of October 25 in Spanish Market Errors During Peak Hours of June 29 in the Californian Market

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Real Price | Forecasted price | Error (%) |  | Real Price | Forecasted price | Error (%) |
| Hour 9 | 57.30 | 61.69 | 7.11 | **Hour 9** | 19.30 | 29.03 | 33.51 |
| Hour 10 | 57.19 | 52.98 | 7.36 | **Hour 10** | 16.65 | 22.26 | 25.20 |
| Hour 11 | 57.14 | 55.93 | 2.11 | **Hour 11** | 17.54 | 19.90 | 11.85 |
| Hour 12 | 57.18 | 54.83 | 4.10 | **Hour 12** | 17.78 | 20.27 | 12.28 |
| Hour 13 | 57.25 | 55.27 | 3.45 | **Hour 13** | 18.34 | 20.46 | 10.36 |
| Hour 14 | 55.94 | 55.20 | 1.32 | **Hour 14** | 19.62 | 20.84 | 5.85 |
| Hour 15 | 55.94 | 53.61 | 4.16 | **Hour 15** | 17.98 | 21.72 | 17.21 |
| Hour 16 | 55.94 | 54.17 | 3.16 | **Hour 16** | 20.54 | 20.72 | 0.86 |
| Hour 17 | 57.14 | 53.97 | 5.54 | **Hour 17** | 23.85 | 22.31 | 6.45 |
| Hour 18 | 57.14 | 55.52 | 2.83 | **Hour 18** | 28.81 | 24.63 | 14.50 |
| Hour 19 | 57.25 | 54.97 | 3.98 | **Hour 19** | 25.29 | 28.10 | 10.00 |
| Hour 20 | 58.46 | 55.30 | 5.40 | **Hour 20** | 25.23 | 26.10 | 3.33 |

As for the forecasts based on the Californian market, the average error is extremely high, about 19.9% for May 21 and 12.6% for June 29. This is mainly due to the price spikes happening in rapid succession in such a small amount of time. GARCH is meant to deal with this volatility, but there is just so much unpredictability happening so fast, most models will have difficulty forecasting.

IV. CONCLUSION

In this paper, GARCH processes have been established in order to forecast hourly electricity prices in both the Californian and Spanish markets. The GARCH model is clearly a powerful tool when attempting to forecast volatile data. A specific GARCH model was obtained for each month from the two different markets.

The errors from the Spanish market are acceptable although they can be improved. Additionally, how GARCH deals with extreme volatility was tested by the model created for the time of May 19-May 28. The forecasted prices for all of the models are usable considering the variation of the data used as well as the size of data. The data used is only about 700 points for each market. Results would be better if the amount of data was in the thousands or at least enough to identify trends and seasonality’s. Future studies could take this into account when creating GARCH models for electricity markets.

REFERENCES

CompañíaOperadora del Mercado Español de Electricidad (Spanish  Market Operator) [Online]. Available: <http://www.omel.es>.

Energy Online [Online]. CAISO (California ISO). Availabe: <http://www.energyonline.com/Data/GenericData.aspx?DataId=20&CAISO___Average_Price>.

Garcia, R. C., J., S., I., Akkeren, M. V., & Garcia, J. C. (2005, May). A GARCH Forecasting Model to Predict Day-Ahead Electricity Prices. *IEEE TRANSACTIONS ON POWER SYSTEMS,* 20(2).

Time Series Analysis for Financial Data VI – GARCH model and predicting SPX returns. (2017, December 13). Retrieved from <https://medium.com/auquan/time-series-analysis-for-finance-arch-garch-models-822f87f1d755>.