Predicting Crude Oil Volatility Utilizing GARCH and ANN-GARCH Models

Michael Napoli

May 2023

Abstract

This paper compares the performance of an Econometric GARCH model and a GARCH model hybridized with an Artificial Neural Network (ANN) in predicting future crude oil volatility using historical trends. Additionally, we discuss these models' theory, providing suggestions for further improving their performance. By examining the strengths and limitations of each model, this paper provides insight into the best practices for utilizing these approaches in the context of crude oil volatility prediction. These findings demonstrate the potential of combining traditional Econometric techniques with machine learning approaches for more accurate and robust volatility modeling.

Introduction

Imagine a world where you can consume marketplace goods and get paid for the privilege – needing to give nothing in return. One would think this utopian planet must have solved the fundamental economic problem; however, in April 2020, our strained, finite economy offered these fiction-like deals: specifically, negative crude oil prices (Johntson, 2022). As the beginning of the pandemic raged and communities began locking down, fewer people were commuting to work. This reduced vehicle traffic reduced the demand for oil, leaving oil producers scrambling to find facilities to store their surpluses. Ultimately, producers had nowhere to stockpile their oversupply since consumer demand had shifted too fast, forcing them to *pay* to have the product taken off their hands. The five supermajors lost $76 billion dollar combined because of the 2020 shocks. (*Big Oil incurred record loss in 2020*, 2021).

As illustrated, volatility can be very costly for various business and consumer interests. This paper discusses how we can utilize Econometric volatility modeling to help predict future uncertainty which can be leveraged to make more informed management decisions. These tools could help those ranging from public servants who want to protect constituents from unstable markets to producers themselves who want to mitigate losses from unsuspected volatility. The resulting research illustrates how pairing traditional Econometric time-series methods with machine learning methods helped drastically reduce the error in predicting future volatility.

The paper first discusses the background of the development of first Econometric ARCH, then GARCH volatility models, and how they were hybridized with neural networks to create ANN-GARCH models. The paper outlines how we prepared the data, our models' underlying mechanisms, and their specifications. Finally, the paper discusses how its findings could be improved by potentially using different types of GARCH models and by altering our neural network.

Literature Review

Published in *Econometrica*, Robert Engle (1982) discusses the shortcomings of assuming homoscedasticity in time series data, arguing that such models neglect key signals in a dataset: namely, changes in a set's variance over time. Engle demonstrates his newly proposed ARCH model (AutoRegressive Conditional Heteroscedasticity) on inflation in the United Kingdom spanning from January 1958 to December 1972. Engle's novel method outperformed traditional Econometric time series models, more accurately capturing dynamic changes. Within the next four years, researchers such as Domowitz and Hakkio (1985) were able to find similar predictive success in foreign exchange markets. Moreover, Weiss (1984) – combining ARMA (AutoRegressive Moving Averages) and Engle's ARCH method – was able to accurately model 13 different macroeconomic time series successfully.

Building off of Engle's work, Tim Bollerslev (1986) created the Generalized AutoRegressive Conditional Heteroscedasticity (GARCH) model. Bollerslev argues that GARCH models perform better on financial data as they can better capture the volatility clustering most financial data contains: much better than ARCH models which are overly "bursty." While an ARCH model only takes in the lagged squared residuals to predict future volatility, Bollerslev's GARCH also utilizes conditional variance squared laggers. These extra parameters make the model better predict consistent, clustered volatility.

Since then, researchers began hybridizing GARCH models with neural networks to make more accurate volatility predictions. Utilizing GARCH-ANN and RNN models, Kristjanpoller and Minutolo (2016) predicted oil prices, concluding that hybridized models do perform better than the traditional models – performing better out of sample as well. Kim and Won (2018) then integrated LSTM to predict the volatility of various currencies and the stock market, further increasing the accuracy of their predictions.

Data

West Texas Intermediate (ticker WTI) is a type of crude oil commonly used as a benchmark for oil prices. The resource is highly valued for being easy to process as it is light – referring to the oil being more "runny" – and sweet – referring to the oil's low sulfur content (Zeihan, 2016). WTI oil is produced in the United States, primarily in the Permian Basin of Texas. The resource is traded on the New York Mercantile Exchange and is used as the basis for pricing most of the crude oil produced in North America. Most importantly for this paper, West Texas Intermediate's price is often used as a reference point for the prices of other crude oils around the world; the supply and demand for WTI, as well as its price fluctuations, can have a significant impact on global oil markets and the overall economy (Chen, 2022). The historical prices used for this paper come from the Bloomberg terminal.

To prepare the data, we first find the historical daily volatility, which we can do by finding the raw returns from one day to the next (Hayes, 2023). Additionally, we split our data into train, validation, and test data: creating a 70/15/15 split. Splitting our data helps prevent overfitting in our neural network by preventing our model from relying on random noise to reduce errors rather than properly understanding underlying rules, the latter of which would be more generalizable to other datasets and predicting outside our sample. In Figure 1 and Figure 2, you can see WTI's historical price over the last 40 years as well as its daily volatility. An important note that becomes important later in this paper is the extreme outlier in 2020 resulting from the pandemic.

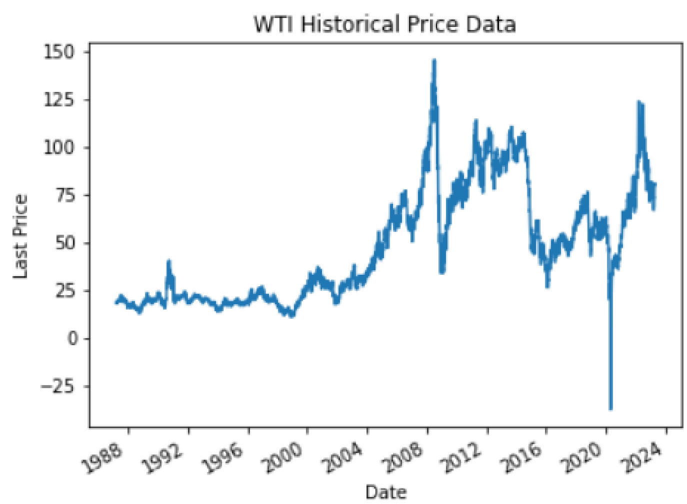

Figure 1

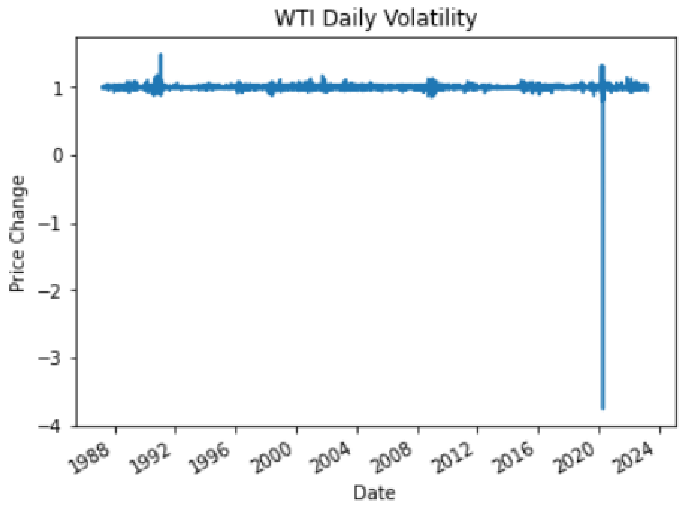

Figure 2

Empirical Specification and Theory

The Garch(p,q) model's general equation is as follows:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{1}$$

Where σ_t^2 is the conditional variance at time t: our dependent variable. Our independent variables include all the variables on the right-side of the equation. ω is the constant term, representing the baseline variance in the data. α_1 is the coefficient of the squared residual term and ϵ_{t-i}^2 is the squared residual at time period $t-i$. β_1 is the coefficient of the conditional variance, and σ_{t-j}^2 is the conditional variance itself: the latter taken at time $t-j$. In this paper we use a Garch(1,1) model, meaning that the upper limits for both summation functions are 1. These upper limits include one squared residual lagger and one conditional variance squared lagger into our model, respectively.

In our data set we found the Garch(1,1) specification that worked best on the WTI data looks as follows:

$$\sigma_t^2 = 1.13 \times 10^{-5} + .1\,\epsilon_{t-i}^2 + .88\,\sigma_{t-j}^2 \tag{2}$$

In our model, each of the parameters are considered statistically significant: well beyond the 5% level.

Now hybridizing the above Econometric model with supervised machine learning techniques, we will first discuss how the Long Short Term Memory (LSTM) neural network works under the hood. The first important aspect of understanding a LSTM is how feeding forward and back-propagation work on a more simple plain-vanilla neural network. In the simpler network, we utilize perceptrons. These types of neurons take independent variables, illustrated as "inputs" in Figure 3. We then multiply these neurons by their respective weights, feeding them and a bias through a summation function, and then finally transforming that output through an activation function. Activation functions come in many forms – RELU, leaky RELU, tanh, and sigmoid are pretty common – however, most activation functions essentially are continuous functions between two variables (usually being 0, 1, or -1), standardizing the output.

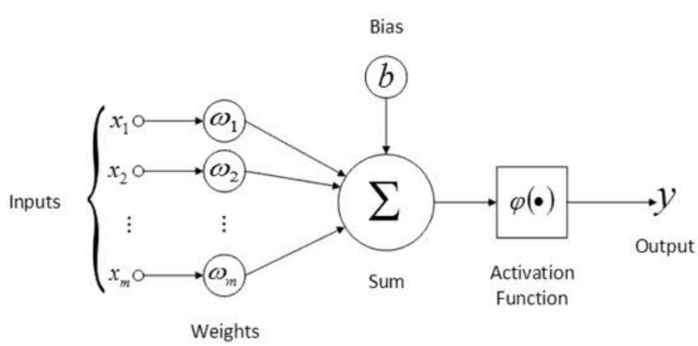

Figure 3: Neural Neuron (Sampaio et al., 2021)

These perceptrons are then chained one after another in layers, as illustrated in Figure 4. The first layer is called the input layer which takes in all the inputs, the middle layers are called hidden layers, and the last layer is called the output layer. We then "feed forward" the inputs through the neural network, allowing the perceptrons to perform their respective operations on the data. Once at the output layer, we compare our predicted output with the data's label – essentially our expected output – finding the residual which is also known as the gradient. We then take the gradient and use the chain rule from calculus to see how much each weight in a given layer "wants" to be nudged to make the predicted output slightly closer to the actual output, going sequentially from the last layer back to the first. This "gradient descent" process of forward propagation and backpropagation repeats until the error in the validation set is minimized.

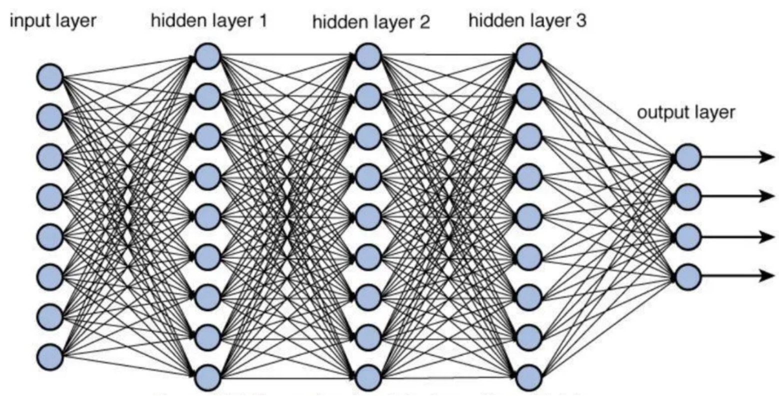

Figure 4: Artificial Neuron (Parmar, 2018)

A Recurrent Neural Network (RNN) is similar to the multilayer neural network illustrated above, except the former performs a lot better on sequential and time series data because their neurons are altered to pass a "hidden state" forward in the network, carrying additional information from previous layers to later ones. This process also works in the reverse during backpropagation, helping earlier neurons learn more than they would otherwise in the more simple perceptron. A LSTM is a type of RNN that adds extra gates which make the network forget unimportant information while allowing important signals to be retained. You can see these various gates in Figure 5 and how the neurons take two inputs and two outputs: a cell state (x_t, z_t and c_t) and a hidden state (h_{t-1} and h_t).

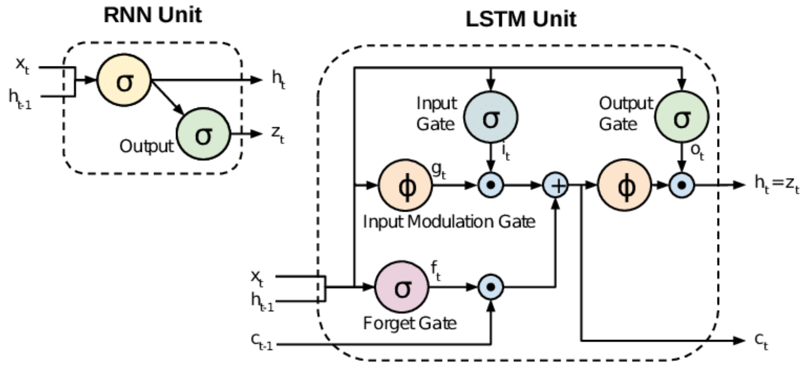

Figure 5: RNN and LSTM Neurons (Rassem et al., 2017)

An important note is that these LSTM neurons are still chained similarly and retain the same properties – including weights, biases, and activation functions –of the simpler perceptrons.

For the ANN in our case, I found four layers worked best: two being hidden layers, one being an input layer, and one being the output layer. The output layer only has one neuron because we only try to predict one value for each time period. Moreover, I used a linear activation function because my data was already standardized in percent form. I implemented a Huber loss function for the gradient decent process, which is less sensitive to outliers than other objective functions, preventing the model from learning weird behaviors trying to account for negative crude oil prices in 2020. Lastly, I fed six laggers into the neural network – three representing the previous three days of predicted volatility from our GARCH model and three representing the last three days' residual – and also fed in our three parameters from our specified GARCH model.

Results

Our hybridized GARCH model outperformed the traditional GARCH model when measured using mean squared error, which aligns with the existing literature. However, the neural network was prone to overfitting and various unexpected behaviors, such as setting all variables to 1 if we trained the network too much. This likely results from us using a symmetric GARCH model while most financial data is asymmetric – as the market often has a negativity bias. Our data's extreme outlier of negative oil prices in 2020 definitely compounded this effect. Using the Huber loss function reduced the erosiveness of our outliers. Still, testing and hybridizing the various asymmetric GARCH models – such as EGARCH and TGARCH – would likely demonstrate much better performance. Lastly, you can see in Figure 6 that when plotted, the ANN-GARCH model closer fit the true daily volatility than the GARCH model by itself.

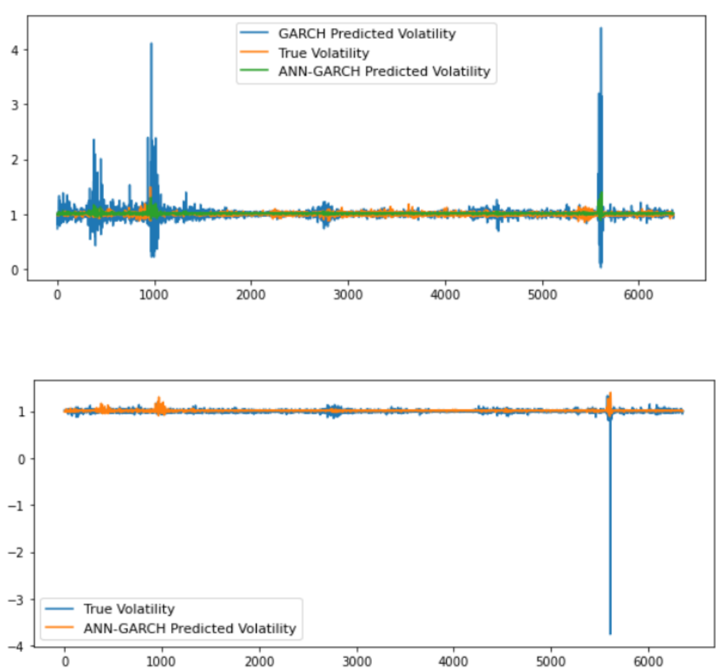

Figure 6: Resulting Daily Volatility

Conclusion

ANN-GARCH models prove to be a powerful tool to predict crude oil volatility. Pairing this tool with macroeconomic data such as GDP, inflation, and exchange rates could provide interesting results on how volatility affects the rest of the economy as well, which could provide more use cases to stakeholders. One example could be if researchers empirically demonstrated to what extent volatility in the crude oil market affects a country's overall GDP, these findings could be used to calculate the optimal amount of resources a government should provide to strategic reserve programs to mitigate these losses in production. Other stakeholders that could use this tool include insurance firms, oil producers, energy distributors, and portfolio managers.

References

Big Oil incurred record loss in 2020. (2021, February). Retrieved 2023-05-09, from https://www.ogj.com/general-interest/economics-markets/article/14197855/big-oil-incurred-record

Bollerslev, T. (1986, April). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327. Retrieved 2023-05-07, from https://linkinghub.elsevier.com/retrieve/pii/0304407686900631 doi: 10.1016/0304-4076(86)90063-1

Chen, J. (2022). *West Texas Intermediate (WTI): Definition and Use as a Benchmark.* Retrieved 2023-05-06, from https://www.investopedia.com/terms/w/wti.asp

Domowitz, I., & Hakkio, C. (1985). Conditional variance and the risk premium in the foreign exchange market. *Journal of International Economics*, *19*(1-2), 47–66. Retrieved 2023-05-09, from https://econpapers.repec.org/article/eeeinecon/v$_3a19_3ay_3a1985_3ai_3a1-2_3ap_3a47-66.htm$ (Publisher: Elsevier)

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, *50*(4), 987–1007. Retrieved 2023-05-07, from https://www.jstor.org/stable/1912773 (Publisher: [Wiley, Econometric Society]) doi: 10.2307/1912773

Johntson, M. (2022). *What Happened to Oil Prices in 2020.* Retrieved 2023-05-09, from https://www.investopedia.com/articles/investing/100615/will-oil-prices-go-2017.asp

Kim, H., & Won, C. (2018, August). *Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models - ScienceDirect.* Retrieved 2023-05-09, from https://www.sciencedirect.com/science/article/pii/S0957417418301416

Kristjanpoller, W., & Minutolo, M. C. (2016, December). Forecasting volatility of oil price using an artificial neural network-GARCH model. *Expert Systems with Applications*, *65*, 233–241. Retrieved 2023-05-09, from `https://www.sciencedirect.com/science/article/pii/S0957417416304420` doi: 10.1016/j.eswa.2016.08.045

Parmar, R. (2018, September). *Training Deep Neural Networks.* Retrieved 2023-05-08, from `https://towardsdatascience.com/training-deep-neural-networks-9fdb1964b964`

Rassem, A., El-Beltagy, M., & Saleh, M. (2017, June). Cross-Country Skiing Gears Classification using Deep Learning.

Sampaio, P. S., Almeida, A. S., & Brites, C. M. (2021, December). Use of Artificial Neural Network Model for Rice Quality Prediction Based on Grain Physical Parameters. *Foods*, *10*(12), 3016. Retrieved 2023-05-08, from `https://www.mdpi.com/2304-8158/10/12/3016` (Number: 12 Publisher: Multidisciplinary Digital Publishing Institute) doi: 10.3390/foods10123016

Weiss, A. A. (1984). Arma Models with Arch Errors. *Journal of Time Series Analysis*, *5*(2), 129–143. Retrieved 2023-05-09, from `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9892.1984.tb00382.x` (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9892.1984.tb00382.x) doi: 10.1111/j.1467-9892.1984.tb00382.x

Zeihan, P. (2016). *The Absolute Super Power – The Shale Revolution and a World without America.*