

Utah State University

DigitalCommons@USU

All Graduate Plan B and other Reports

Graduate Studies

8-2020

Reinforcement Learning for Dynamic Futures Hedging

Evan Bullard

Utah State University

Follow this and additional works at: <https://digitalcommons.usu.edu/gradreports>



Part of the [Finance and Financial Management Commons](#)

Recommended Citation

Bullard, Evan, "Reinforcement Learning for Dynamic Futures Hedging" (2020). *All Graduate Plan B and other Reports*. 1479.

<https://digitalcommons.usu.edu/gradreports/1479>

This Report is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Plan B and other Reports by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.



Abstract

This paper focuses on oil hedging using near month crude oil futures. Hedging may allow a firm to reduce risks and focus on areas of comparative advantage. Hedging requires a firm to estimate ex-ante the correct hedge ratio. The portfolio optimization framework allows for OLS to be applied to the estimation of a hedge ratio. Reinforcement Learning is another method available to hedgers to estimate a hedge ratio. Three strategies using econometric tools and one using Reinforcement Learning are estimated and tested against 2019 oil price data.

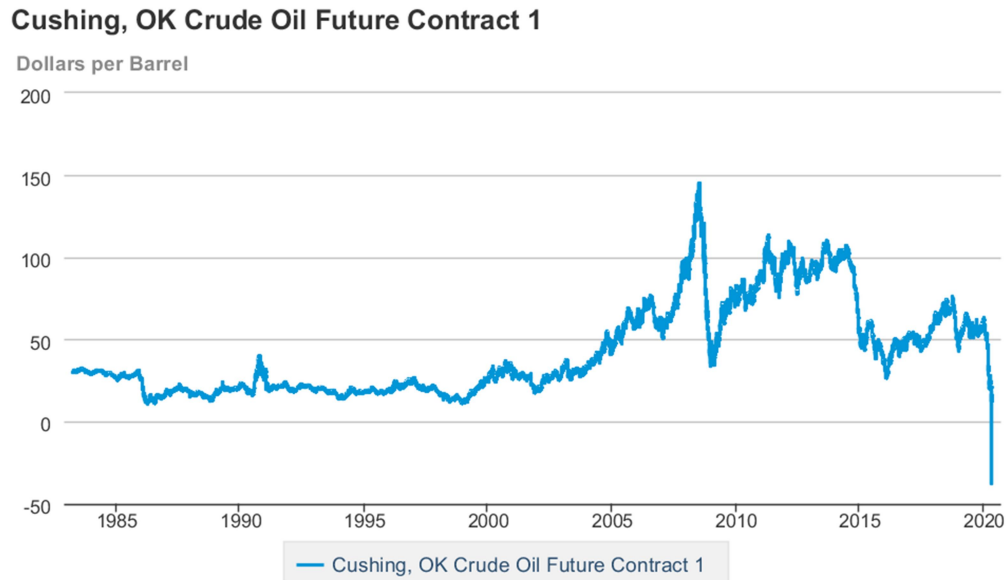
Introduction

This paper focuses on applying Deep Reinforcement Learning (RL) to dynamic oil hedging. Dynamic hedging is characterized by a change of strategy based upon the acquisition of additional information. An alternative to dynamic hedging is that of static hedging strategy. In static hedging, the strategy is fixed at the beginning of the hedge and does not change with new information.

Daily prices were obtained from the Energy Information Administration (EIA) and investing.com. The date range of the data is from January 1986 to December 2019. January 1986 to December 2018 is designated as the training data, while January 2019 to December 2019 is designated as the test data.¹

Recently, oil prices have experienced high volatility. On April 20, 2020, the price of oil fell to an all-time low of \$-37.63(eia.gov). Oil markets are typically in backwardation, which is characterized by futures prices below cash prices. Currently, oil markets are in contango, which is characterized by futures prices above cash prices. So oil prices for 2020 has not been included in the test data.

¹ Denise Bullard provided assistance in the filtering and data manipulation required of this project.



 Source: U.S. Energy Information Administration

Figure 1 (<https://www.eia.gov/dnav/pet/hist/RCLC1D.htm>)

Hedging

Hedging may help a firm to reduce the risks that it has exposure to. This risk reduction is accomplished by taking a position in the futures market. One such risk is that hedging allows a firm to reduce its price risk. By engaging in hedging a firm may more accurately predict future prices for a good. This removal of uncertainty through hedging helps allow a firm to plan and focus on areas where it possesses a competitive advantage. This removal of uncertainty can apply to firms that sell or purchase assets.

A firm that is involved in the sale of a physical commodity and wishes to hedge will select a short hedge. A short hedge involves selling a futures contract prior to selling the physical commodity. This transaction will effectively provide the firm with the current futures price as their cash price at the date of delivery. A firm that will purchase a commodity will engage in a long hedge. A long hedge involves the purchase of futures and then the subsequent purchase of

the commodity. This transaction provides a firm with the current futures price as the cost of the good purchased.

Regardless of the need for a short or long hedge, the question is the same. “Determination of the variance-minimizing hedge strategy for each t requires the solution of an extremely complex dynamic programming problem.”(Pirrong, 1997, p.547) The hedger must choose a method to solve this dynamic problem. The method of solving this complex problem will be addressed in two different ways in this paper. The first method is through an econometric framework focusing on previous observations and parameter estimation. The second method is through the creation of a RL environment.

Econometrics

Portfolio optimization is a powerful tool to be applied to the hedging problem. Implicit in this framework is a general quadratic utility function show in Equation 1.

$$U = E(r) - \frac{1}{2}A\sigma^2 \quad (1)$$

Where

U – is the utility
 $E(r)$ – is the expected return
 A – is the coefficient of risk aversion
 σ^2 – is the variance of the returns

(Bodie Kane Marcus, 2009, p.159)

The variance of the two assets in the portfolio can be expressed by the variance and covariance of the two assets. This relationship is demonstrated in Equation 2.

$$\sigma_P^2 = \sigma_S^2 + h^2\sigma_F^2 - 2h\rho\sigma_S\sigma_F \quad (2)$$

Where

σ_P^2 – is the variance of the portfolio
 σ_S^2 – is the variance of the spot price
 σ_F^2 – is the variance of the futures price
 h – is the hedge ratio
 ρ – is the coefficient of correlation

This expression of the variance can be differentiated. To find a hedge ratio that minimizes the variance of the portfolio the first derivative is solved for h . This process leads to a minimum portfolio variance of h^* as expressed in Equation 3.

$$h^* = \rho \frac{\sigma_S}{\sigma_F} \quad (3)$$

Where

ρ – is the correlation between the cash and futures
 σ_S – is the standard deviation of the change in the spot price
 σ_F – is the standard deviation of the change in the futures price

(Hull, 2012, p.57)

This result is the same relationship that is found in a parameter estimate produced by ordinary least squares. Hedging is a forward looking strategy, so there is no method of applying a hedging strategy ex-post. The hedging strategy must be estimated and implemented prior to the relationship between the assets being revealed. Without exact knowledge about the future, the hedger must estimate the future relationship. To perform this estimation, the only available information is from the past. Therefore, a linear regression on previous price changes is an appropriate estimation technique to apply to the development of a hedging strategy.

Reinforcement Learning

“Reinforcement learning is learning what to do ... so as to maximize a numerical reward signal.” (Sutton and Barto, 2018, p.1) The correct action is not provided to the learner, so the learner must determine the correct action through a method of trial and error. This perspective is particularly powerful for the formulation of a hedging strategy. Ex-ante, there does not exist an objectively correct hedge ratio or hedging strategy. It is only through time that the correct action can become clear.

The two main parts of a reinforcement learning system are the environment and the learner. The environment must provide the learner with information and allow the learner to take an action. Following the action, a result will be computed and provided as feedback to the learner. The learner will take an action, observe the results, and then adjust the behavior for the next action. The interaction between the agent and the environment is demonstrated in Figure 2.

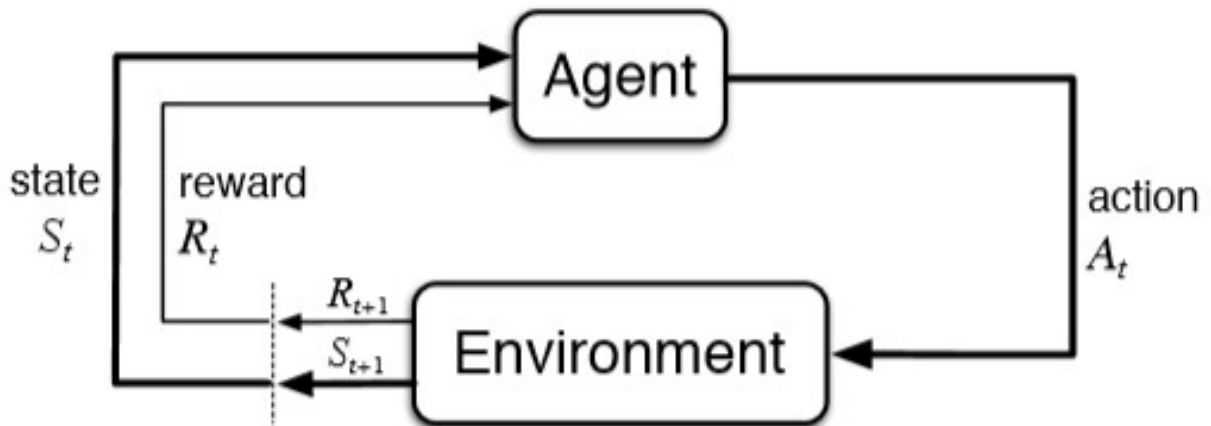


Figure 2 (Sutton and Barto, 2018, p.48)

Methodology

Static Hedge Strategy

The static minimum variance hedging strategy requires the calculation of a single hedge ratio. The method of estimation was accomplished by regressing price changes for the entire history of West Texas Intermediate (WTI). This provides an estimate of the minimum variance hedge with all available information as of the beginning of the test set. This single ratio is applied through the entire test set of 2019. The ratio that this method provides is 1.004161

Expanding Window Strategy

The creation of this strategy was accomplished by creating an expanding window that eventually took in price changes for the entire history of WTI futures and performed a regression on this expanding window of price changes. This provides an updating estimate of the minimum variance hedge ratio as more information becomes available. The change in the hedge ratio estimate over the test set is show in Figure 3.

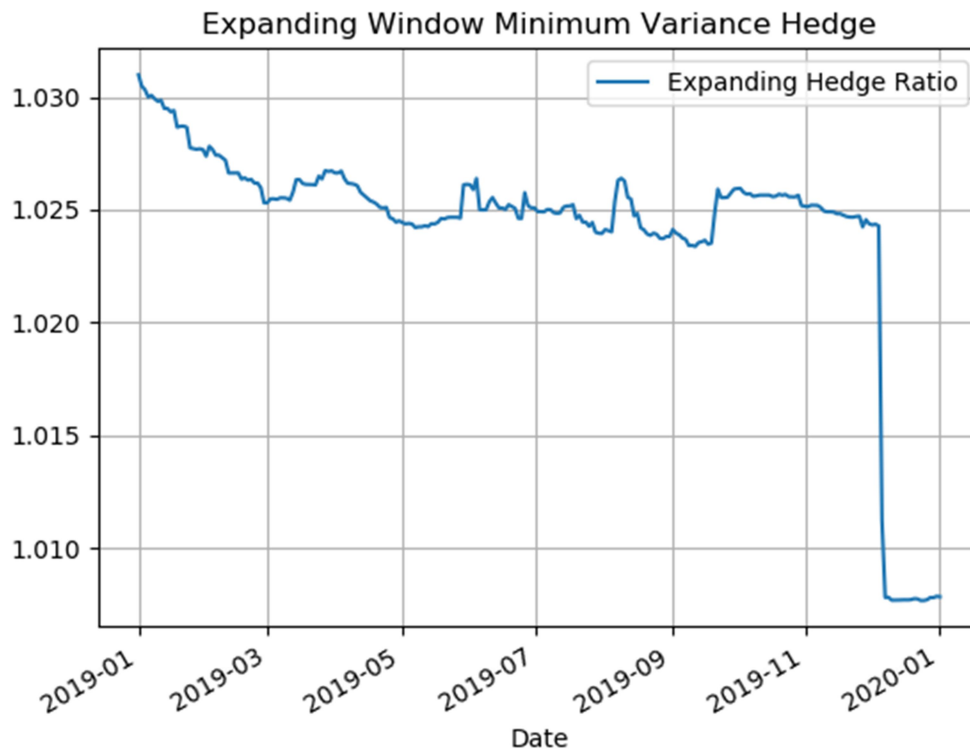


Figure 3

Rolling Window Strategy

This strategy was developed by creating a rolling window of 63 days of daily price changes through 2019. The size of this rolling window was selected to take account of the most recent quarter. The rolling window was applied through the entire data set, and the estimates to be applied to the test set are plotted in Figure 4.

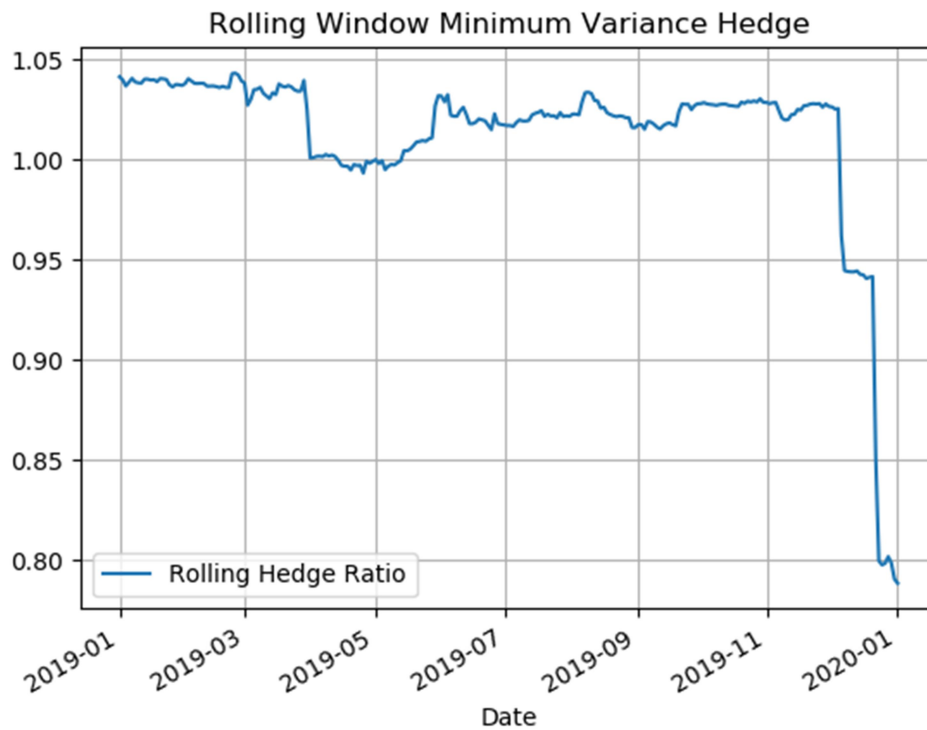


Figure 4

Reinforcement Learning Strategy

The RL environment that was created for this project consists of daily price data for crude oil and the near month crude oil futures contracts.² The action available to the learner is to change the hedge ratio. An economic utility function has been adapted for the reward calculation.³

The price of crude oil in this environment is defined as the price of WTI on the observation date. (eia.gov) The futures price within the environment is defined as the near month futures contracts for WTI. Near month futures are defined as the futures contract that has a

² This environment relies heavily on a stock market example created by Adam King. Mr. King's helpful article may be found here. <https://towardsdatascience.com/creating-a-custom-openai-gym-environment-for-stock-trading-be532be3910e>

³ Joshua Bullard provided technical assistance with the programming of the environment.

delivery date closest to the current date. The near month futures contract changes three days prior to the 25th of each month.(eia.gov)

The hedge ratio is the variable that the learner will control to maximize the reward received. The hedge ratio controls the number of futures contracts purchased or sold. The hedge ratio for this environment has been limited from 0 to 1.5. A hedge ratio of 0 represents an unhedged position. A hedge ratio of 1.5 is a speculative position in the futures market.

The reward function is based on a quadratic utility function. The specific reward function is given in Equation 4.

$$Reward = \Delta S - h * \Delta F - \frac{1}{2} A (\Delta S - h * \Delta F)^2 \quad (4)$$

Where

- ΔS – is the change in the spot price over the contract
- ΔF – is the change in the futures price over the contract
- h – is the hege ratio
- A – is the coefficient of risk aversion

A coefficient of risk aversion of 1 was chosen. The model was trained on historical price data for 70 epochs. The cumulative reward for this training is show in Figure 5. The trained model is then provided with the prices from 2019, and a daily hedge ratio is predicted. This series of hedge ratios are then applied to calculate a return if the RL strategy had been followed.

Cumulative_reward_VS_Epoch

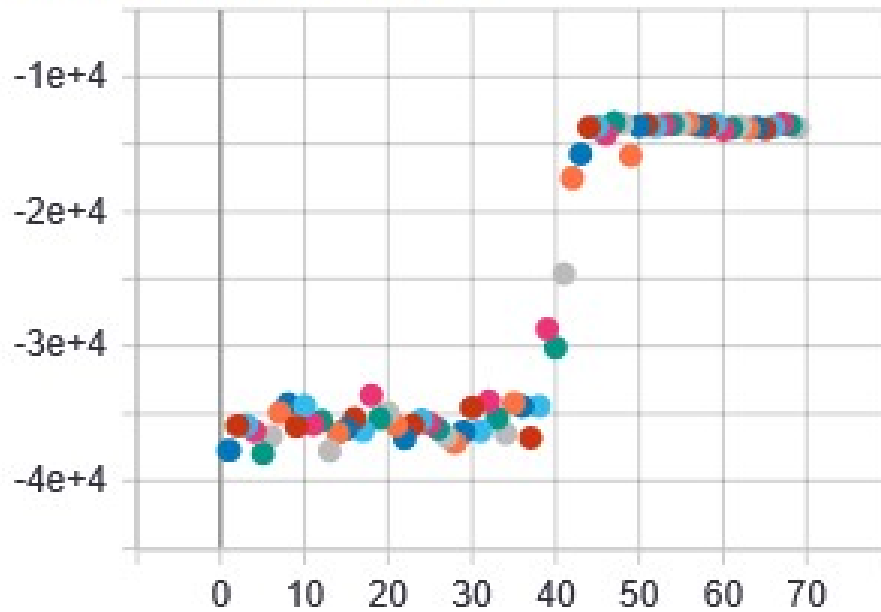


Figure 5

Results

To compare the different strategies, a common method was required. The strategies use a common utility function and therefore are comparable. The method chosen for evaluating which strategy produced the best results is the change in portfolio value. This calculation is given in Equation 5.

$$\Delta P = \sum((S_C - S_t) - h * (F_C - F_t)) \quad (5)$$

Where

S_C – is the spot price on the delivery date

S_t – is the current spot price

F_C – is the futures price on the delivery date

F_t – is the current futures price

h – is the hedge ratio selected for day t

The results of this evaluation are summarized in Table 1 below.

Abnormal Returns	
Strategy	Results
Static Minimum Variance Hedging	1.99
Dynamic Minimum Variance Hedge (Expanding Window)	12.55
Dynamic Minimum Variance Hedging (Rolling Window)	-9.79
RL Dynamic Hedging	-154.98

Table 1

Conclusion

Through this project, it has been demonstrated that it is possible to apply reinforcement learning to develop a dynamic hedging strategy given prices as the input. The development of a reinforcement learning environment allows more flexibility to the performance of forecasting regardless of whether it applies to hedging or other economic decisions.

In 2019, oil prices were relatively range bound. Spot prices increased from a low at the start of the year. Prices then reached a peak above \$65/bbl in the second quarter and then remained below this peak through the rest of the year. The price history for 2019 is shown in Figure 6. The strategy that provided the highest abnormal return within the test set was the expanding window. The RL strategy performed the worst of all of the strategies that were employed in this test.

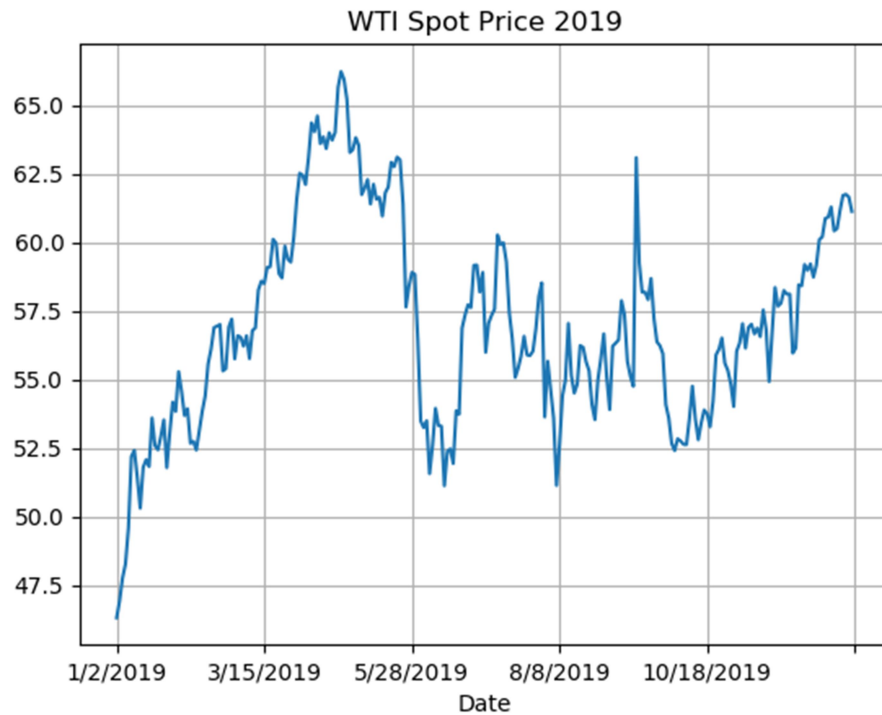


Figure 6

It may be possible to improve performance of the RL strategy by employing different reward functions and additional variables. This environment development is crucial to the final results of any artificial intelligence model.

With modern computational methods, it is possible to utilize powerful tools to answer new questions. With this expansion of ability, there remains a need for a firm understanding of economic theory and statistical practices. It is not enough to simply apply the latest methods and blindly accept the output. It is crucial to remember the questions that are being asked and the answers that are provided.

Further Research

With the advances in machine learning and artificial intelligence, it is now possible to investigate nonlinear relationships between assets. A logical extension of this paper is to examine cross hedging within energy markets. The same methods may be applied to long hedging and inventory management. Further refinements of the basic environment created here include the introduction of cross hedging, risk free borrowing rates, and differing contract terms.

References

- Alexander, Carol. (2008). Market Risk Analysis Vol. 3: Pricing, Hedging and Trading Financial Instruments. John Wiley & Sons Ltd.
- Arxiv Insights. (2018, October 1). Policy Gradient methods and Proximal Policy Optimization (PPO): Diving into Deep RL!. YouTube. <https://www.youtube.com/watch?v=5P7I-xPq8u8>
- Beeman, D. (2001, November 1). Multi-layer perceptrons (feed-forward nets), gradient descent, and back propagation. University of Colorado Boulder. <https://ecee.colorado.edu/~ecen4831/lectures/NNet3.html>
- Bodie, Zvi, Alex Kane, and Alan J. Marcus. (2009). Investments 8th edition. McGraw-Hill Irwin.
- Elliott, Graham and Allan Timmermann. (2008). Economic Forecasting. Journal of Economic Literature. Vol.46 No.1, pp. 3-56. <http://www.jstor.org/stable/27646946>
- Elliott, Graham and Allan Timmermann. (2016). Economic Forecasting. Princeton University Press.
- Hill, Ashley and Raffin, Antonin and Ernestus, Maximilian and Gleave, Adam and Kanervisto, Anssi and Traore, Rene and Dhariwal, Prafulla and Hesse, Christopher and Klimov, Oleg and Nichol, Alex and Plappert, Matthias and Radford, Alec and Schulman, John and Sidor, Szymon and Wu, Yuhuai. (2018). Stable Baselines(1). GitHub. <https://github.com/hill-a/stable-baselines>
- Holbrook, Working. (1953). Futures Trading and Hedging. The American Economic Review, Vol.43 No.3, pp. 314-343. <http://www.jstor.org/stable/1811346>
- Holbrook Working. (1962). New Concepts Concerning Futures Markets and Prices. The American Economic Review, Vol.52 No.3, pp.431-459. <http://www.jstor.org/stable/1810553>
- Hull, C. John. (2012). Options, Futures, and Other Derivatives 8th edition. Prentice Hall.
- Investing.com. (2020). Crude Oil WTI Futures Historical Data. <https://www.investing.com/commodities/crude-oil-historical-data>
- King, Adam. (2019, April 10). Create custom gym environments from scratch – A stock market example. towards data science. <https://towardsdatascience.com/creating-a-custom-openai-gym-environment-for-stock-trading-be532be3910e>
- Lence, H. Sergio and Dermot J. Hayes. (1994). The Empirical Minimum-Variance Hedge. American Journal of Agricultural Economics, Vol.76 No.1, pp. 94-104. <http://www.jstor.org/stable/1243924>
- Lence, H. Sergio and Dermot J. Hayes. (1994). Parameter-Based Decision Making Under Estimation Risk: An Application to Futures Trading. The Journal of Finance, Vol. 49 No.1, pp. 345-357. <http://www.jstor.org/stable/2329149>
- Lence, H. Sergio. (1995). The Economic Value of Minimum-Variance Hedges. American Journal of Agricultural Economics, Vol. 77 No. 2, pp. 353-364. <http://www.jstor.org/stable/1243545>

- McDonald, L. Robert. (2013). Derivatives Markets 3rd edition. Prentice Hall.
- Schulman, J., Wolski, F., Dhariwal, P., Radford, A., Klimov, O., (2017). Proximal Policy Optimization Algorithms, arXiv:1707.06347
- Sutton, S. Richard and Andrew G. Barto. (2018). Reinforcement Learning: An Introduction 2nd edition. The MIT Press.
- Pirrong, Stephen Craig. (1997). Metallgesellschaft: a prudent hedger ruined, or a wildcatter on NYMEX?. Journal of Futures Markets. Vol. 17 Issue 5, pp. 543-578.
- U.S. Energy Information Administration. (2020). Cushing, OK WTI Spot Price FOB.
<https://www.eia.gov/dnav/pet/hist/rwtcD.htm>