Kalman Filters in Carry and Inventory-Based Spread Trading

Raymond Luo

March 25th, 2021

1 Introduction

In considering oil carry as a proxy for inventories, it seemed possible that the perturbations in oil carrys' ability to represent inventory to be considered a signal for crude oil positions. Here, this paper incorporates the concept of using Kalman Filters to provide a dynamic hedge ratio between oil carry signals, namely $F_1 - F_2$ for WTI crude oil, and the properties of oil inventory, where we consider the capacity of oil storages used. By examining changes in the dynamic hedge ratio, this paper proposes a spread trading strategy analogous to the popular implementation of Kalman filters to examine hedge ratios between stocks. In this case, we examine the spread between the deviation (another spread) of the oil capacity to its 10 day moving average and the simple oil carry signal. We have the following notation:

$$
\begin{aligned}
&C_t(m,n) = F_{m,t} - F_{n,t}, \quad m < n \\
&I = \text{Current Crude Oil Inventory} \\
&C = \text{Total Available Capacity} \\
&I/C = \text{Percentage Inventory Used} \\
&I/C_{MA10} = \text{10 Day Moving Average of } I/C \\
&\Delta I/C = I/C_{dev} = I/C - I/C_{MA10} \\
&\pi(F_t, \Delta I/C) = \text{Trading Strategy based on } F_t, \Delta I/C \\
&A_t = \text{Kalman-filtered Expectation of State} \\
&e_t = y - C_t \cdot x = \text{Residuals from Kalman Filtering Predictions} \\
&Q_t = \text{Kalman Filtered Covariance} \\
&\epsilon = \text{Threshold control}
\end{aligned}
$$

1.1 Data Analysis

For current crude oil inventory, we use the weekly US stock of all commercial crude oil products along with Alaska transit products, this is the total US crude oil stock excluding the SPR crude oil stocks. For total available storage capacity, we used the monthly release of total working storage capacity and stocks in transit data from the Crude Oil Stocks and Storage Capacity (March 2011 - March 2020) report. Both data were taken from the U.S. Energy Information Administration (EIA) database. As the storage capacity data is restricted to March 2011 to March 2020, we reserve our analysis to this time frame. We specifically analyze our strategy on in the window of Jan 2015 to Jan 2020. We then backtest our results on the years preceding this time period from March 2011 to Dec 2015.

As the crude oil inventory is weekly and the capacity report is monthly, we interpolated our data to the daily basis in order to match the WTI crude oil futures that we use to trade and roll on a daily basis. It is important to note here, that we used linear interpolation which may prove to be very inaccurate for the weekly crude oil inventory data. While the monthly data is a lot longer to interpolate, as working storage capacity tends to fluctuate less as an infrastructural feature, linear interpolation still deemed suitable for our analysis. For our trading strategy, we will be taking positions in WTI Crude Oil futures from Cushing, Oklahoma and will be rolling the futures 5 days before their time of expiration.

We have the following two scatterplots colored by the date of data point for WTI Spreads ($C_{t,1,2}$) versus percent inventory used (I/C) and the spread of the inventory capacity to its 10 day moving-average ($\Delta I/C_{MA10}$):

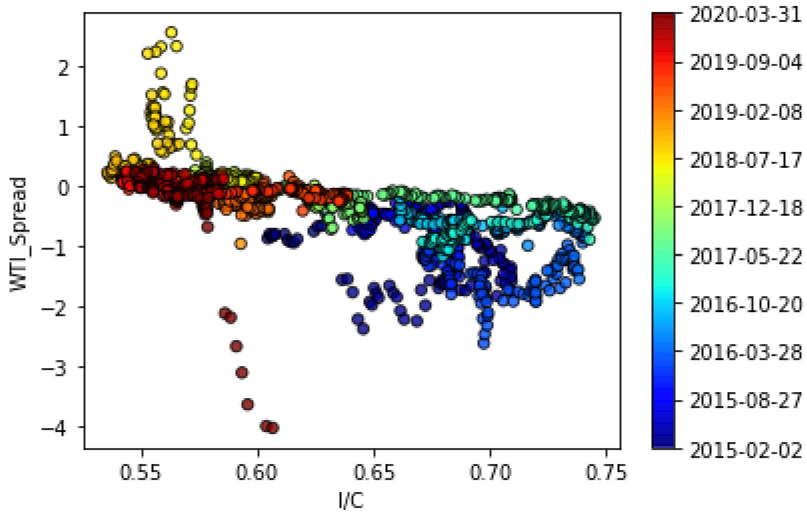

Figure 1: Scatterplot of WTI_{spread} to I/C

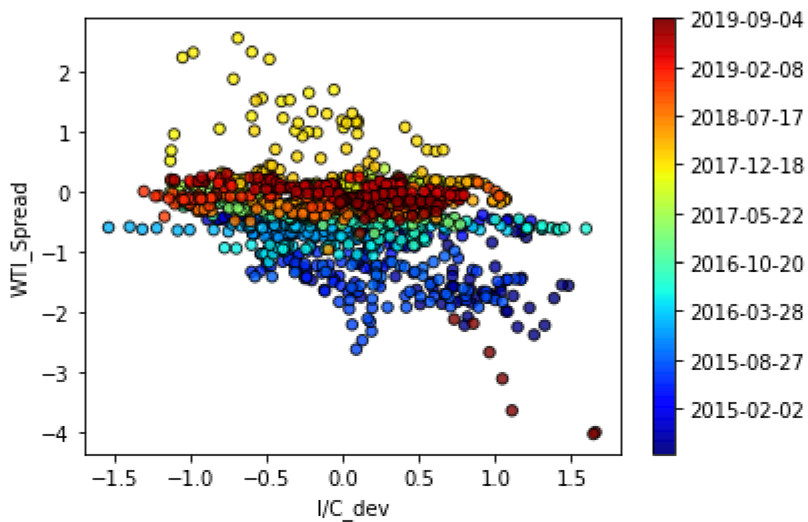

Figure 2: Scatterplot of WTI_{spread} to I/C_{dev}

In the exploration of applying kalman filters between different statistics, we first analyze the cointegration and stationarity of our data. Here, we collected a dataframe including total US petroleum product stock, total OECD US petroleum product stock, WTI 1-month futures contracts, WTI 2-month futures contracts, US crude oil product stock excluding SPR, and US total crude oil storage capacity data (and some derivations of these data).

We use the Johansen cointegration test to find a heat map of the cointegration levels across pairs of our dataframe. It becomes apparent that the majority of pairs have no significant cointegration level, proving it difficult to construct an effective pair-trading strategy. Of the pairs of our interest, the pairing of I/C spread to its moving average (denoted I/C_{dev}) and the WTI_spread has a range that is relatively close to significance. Looking at it more closely, we find that this pair has a 90% confidence interval of Johansen cointegration. This is why we decide to structure our pair trading strategy between the two.

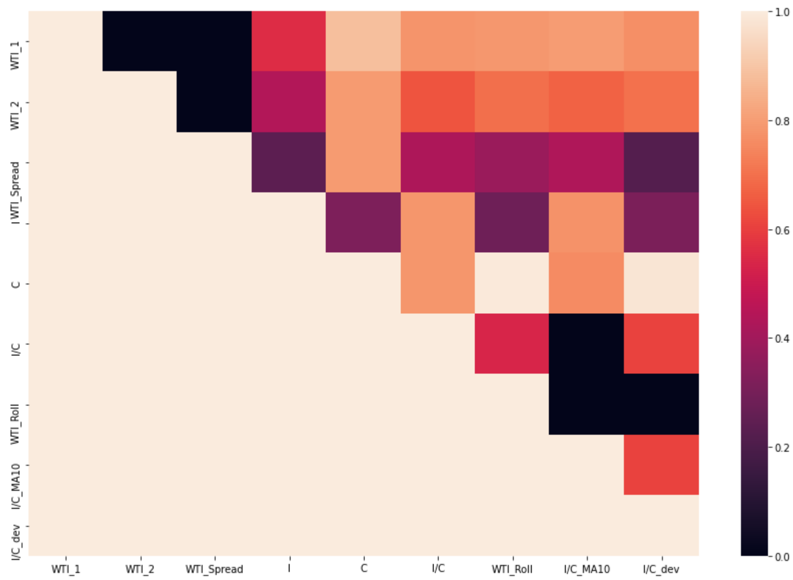

Figure 3: Cointegration Heat Map of Pairs

```
Trace Statistic: [57.518856    1.40094371]
Trace Critical Values: [[13.4294 15.4943 19.9349]
 [ 2.7055  3.8415  6.6349]]
Eigen Statistic: [56.1179123   1.40094371]
Eigen Critical Values: [[12.2971 14.2639 18.52  ]
 [ 2.7055  3.8415  6.6349]]
```

Figure 4: Cointegration Test on I/C_dev to WTI_spread

1.2 Kalman Filter Results

We implement Kalman Filtering with usage of the pykalman python package. After running Kalman filtering on our data, we have the following state means and state intercepts:

After using the Kalman Filter on this pairing, we can observe the residuals e_t which are the difference between the observed time-series WTI spreads and the predicted spreads according to the Kalman Filter. We run the Augmented Dickey–Fuller test to test for stationarity and we observe the Hurst Exponent, which represents a measure for mean-reversion and randomness. If the Hurst Exponent value is less than 0.5, then the residuals are mean-reverting (with strength indicated by how close it is to zero). We find in our results below that our residual statistic is both mean-reverting and stationary.

The diagram below indicates the covariance of state distribution at time t given observations from times [0, t] (denoted Q_t), for which we use for our trading strategy:

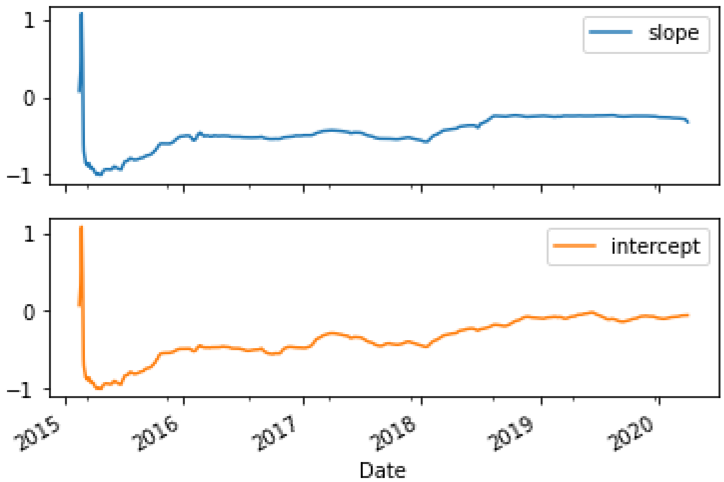

Figure 5: State means and State intercepts of Kalman Filter

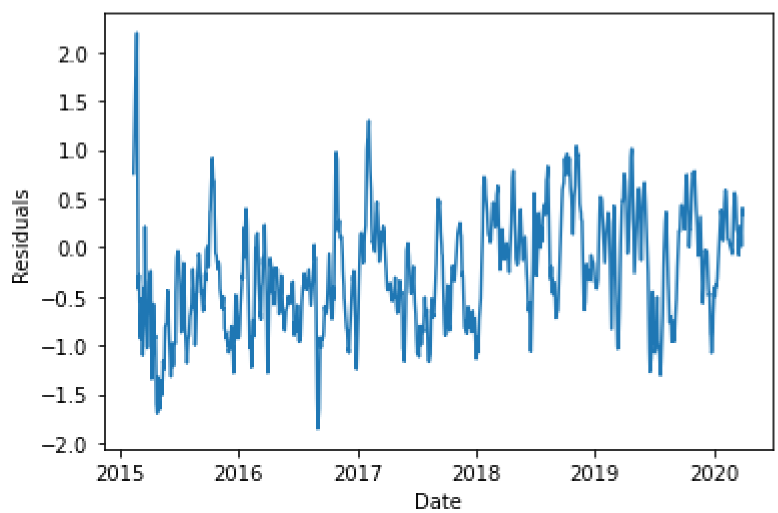

Figure 6: Residuals from Kalman Filter

1.3 Trading Strategy

From crude oil futures carry signal $C_t = F_{1,t} - F_{2,t}$ and inventory statistic $x = \Delta I/C$, for our choice of $\epsilon \in \{0.2x : x \in [0, 200]\}$ have the following Kalman filter trading strategy:

$$\pi(e_t, Q_t) = \begin{cases} 1, & \text{if } e_t < -\epsilon\sqrt{Q_t} \qquad \text{Equivalent to longing } C_t - x \\ -1, & \text{if } e_t > \epsilon\sqrt{Q_t} \qquad \text{Equivalent to shorting } C_t - x \\ 0, & \text{otherwise,} \end{cases}$$

Where $e_t = x_t - A_t \cdot x_{t-1}$, A_t is the mean of state distribution at time t given observations from times [0, t] and Q_t is the covariance of state distribution at time t given observations from times [0, t].

```
[ 0.07486439  0.34092156  0.63170688 ... -0.30194126 -0.3139269
 -0.32339096]
Results of Dickey-Fuller Test:
Test Statistic                   -5.432902
p-value                           0.000003
#Lags Used                       15.000000
Number of Observations Used    1263.000000
Critical Value (1%)              -3.435538
Critical Value (5%)              -2.863831
Critical Value (10%)             -2.567990
dtype: float64
Hurst Exponent 0.30561285430747165
```

Figure 7: Hurst Exponent and ADF Test

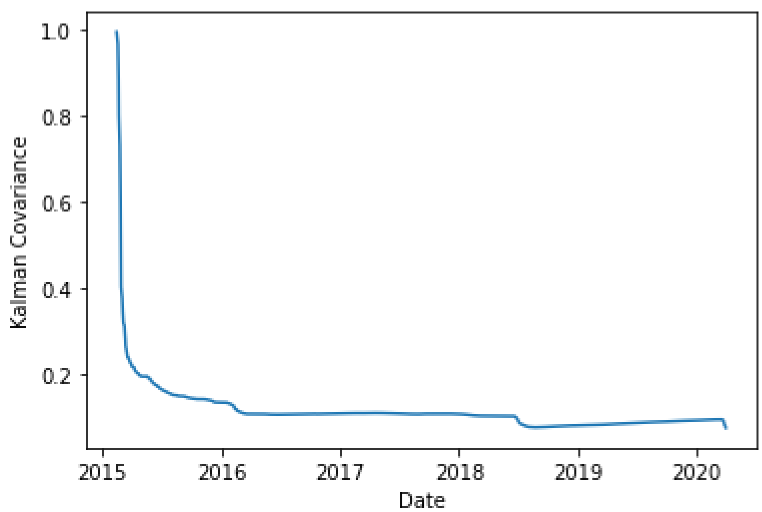

Figure 8: Kalman State Covariance

2 Empirical Results

We run our strategy for WTI futures data from Jan 2015 to March 2020 for varying values of $\epsilon = [0, 10]$ incrementing by 0.2 at a time. We record our terminal rolled PnL, Sharpe Ratio, and Drawdown for these varying values of ϵ to find the following diagrams.

We find that the optimal value of ϵ appears to be about 3 as the later spike in Sharpe ratio does not guarantee as high of a Return on Drawdown due to the drawdown difference. Using this $epsilon = 3$, we find that our strategy produces a result of:

Annualized PnL: 701760
Return on Drawdown: 0.3586
Sharpe Ratio: 0.4722
Max Drawdown: 1957000

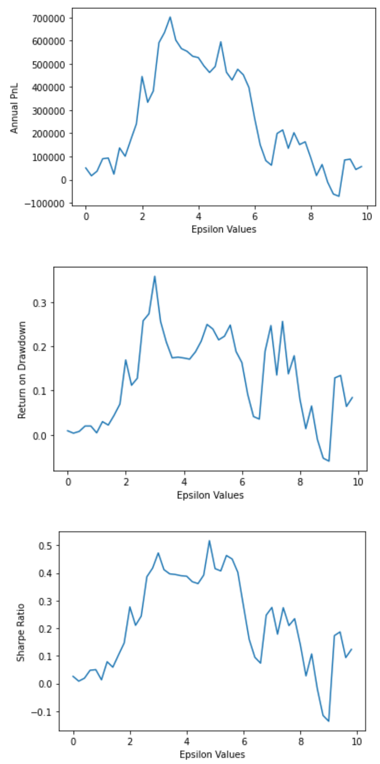

Figure 11: Drawdown for Different Epsilon

It is also interesting to note that while our Sharpe Ratio is a slight improvement to the baseline carry-signal trading strategy, our strategy involves a strategy that is not dominated by shorting and holding WTI. In fact, we observe that in our trading period, we have 589 positions in long, 238 positions in short, and 452 neutral positions. Our strategy also maintains a decent RoD, as it the drawdowns in our strategy do not seems too bad. However, due to the limitations in our capacity data, our strategy does not include the performance of the strategy over April 2020, when oil futures went negative.

We then take our strategy with $\epsilon = 3$ and run it on our test set for March 2011 to Dec 2015. The results are quite bad as we have the following statistics:

Annualized PnL: -674400.0
Return on Drawdown: -0.1711
Sharpe Ratio: -0.5733
Max Drawdown: 3941000.00

While this test period is quite brief, we can see that our strategy proves to be quite volatile depending on the regime that it is observing. We also note that our strategy provides a similar distribution of positions at 238 long positions, 235 short positions, and 465 neutral positions. There are likely various things to consider, such as the long term and short term rates of mean reversion for the inventory capacity and the possibility that the number of factors that were attempted to be brought in (as we considered the spread between WTI crude time spread and the inventory capacity momentum spread) added a lot of complexity that got in the way of the model. Further tests probably are also needed for the evaluation of the cointegration and mean reversion for statistics during the test set period.

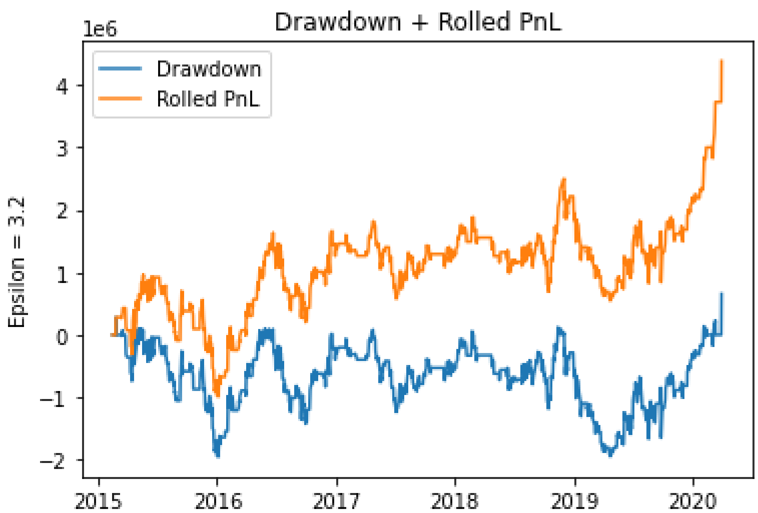

Figure 12: Strategy Results for $\epsilon = 0.3$