**CUNY DATA 698: Master’s Research Project**

**Measuring a hurricane season-related risk premium for energy commodities [oil, gasoline, heating oil and refinery margins] with a focus on the Gulf**

**Table of Contents**

[**Introduction** 4](#_Toc57842013)

[**Literature Review** 5](#_Toc57842014)

[**Theory and Hypothesis** 7](#_Toc57842015)

[**Gulf Coast Petroleum Infrastructure** 7](#_Toc57842016)

[**Commodity Markets** 9](#_Toc57842017)

[**Hurricanes in Gulf Coast** 10](#_Toc57842018)

[**Hypothesis** 11](#_Toc57842019)

[**Data and Variables** 12](#_Toc57842020)

[**Commodities Futures Data** 12](#_Toc57842021)

[**Commodities Seasonality** 13](#_Toc57842022)

[**Hurricane Data** 15](#_Toc57842023)

[**Statistical Methods** 19](#_Toc57842024)

[**Findings** 20](#_Toc57842025)

[**Commodity Time Spread and Cracks Analysis** 20](#_Toc57842026)

[**Commodity Time Spread and Cracks Data Returns Analysis** 23](#_Toc57842027)

[**Hurricane Data Initial Analysis** 26](#_Toc57842028)

[**Discussions** 29](#_Toc57842029)

[**Strategy Design** 32](#_Toc57842030)

[**Conclusion** 34](#_Toc57842031)

[**Future Work** 35](#_Toc57842032)

[**References** 36](#_Toc57842033)

[Figure 1: Gulf Coast Energy Infrastructure EIA [25] 7](#_Toc57841978)

[Figure 2 Colonial Pipeline System [27] 8](#_Toc57841979)

[Figure 3 Colonial Pipeline Damage due to Harvey [29] 9](#_Toc57841980)

[Figure 4 Backwardation vs Contango 10](#_Toc57841981)

[Figure 5 Saffir-Simpsons Hurricane Scale [33] 11](#_Toc57841982)

[Figure 6: RBOB (Gasoline) Continuous M+2 Time-series 12](#_Toc57841983)

[Figure 7: RB (Gasoline) Month+2 Continuous Price Series chart 13](file:///C:\Users\Neil\Documents\MSDS-Data%20Science\DATA%20698\698Project\Neil_Shah_698_Final.docx#_Toc57841984)

[Figure 8: Imputed 3-2-1 refinery crack price chart 13](#_Toc57841985)

[Figure 9 Snippet of Custom Seasonality 14](#_Toc57841986)

[Figure 10: RB M+2 Seasonality 15](#_Toc57841987)

[Figure 11 Example HURDAT2 for Irene 16](#_Toc57841988)

[Figure 12 Wind Quadrants 16](#_Toc57841989)

[Figure 13 Cleaned up HURDATA in dataframe 17](#_Toc57841990)

[Figure 14 HURDATA with Imputed Category 17](#_Toc57841991)

[Figure 15 HURDAT2 Hurricane Map 18](#_Toc57841992)

[Figure 16 Summary Weekly Plots 21](#_Toc57841993)

[Figure 17 ADF Statistics for Prices 21](#_Toc57841994)

[Figure 18 Pearson Correlations 22](#_Toc57841995)

[Figure 19 Price ACFs 22](#_Toc57841996)

[Figure 20 Weekly Returns 23](#_Toc57841997)

[Figure 21 Returns Histogram 24](#_Toc57841998)

[Figure 22 Returns Summary Stats 24](#_Toc57841999)

[Figure 23 Returns ACF 25](#_Toc57842000)

[Figure 24 Pearson Returns 25](#_Toc57842001)

[Figure 25 Number of Hurricanes Per Year 26](#_Toc57842002)

[Figure 26 Monthly Hurricane Frequency and ACF 26](#_Toc57842003)

[Figure 27 Frequency of Hurricanes by Month 27](#_Toc57842004)

[Figure 28 Category Strength and Plots 28](#_Toc57842005)

[Figure 29 Bootstrap Code Snippet 29](#_Toc57842006)

[Figure 30 Bootstrap Returns for Seasons 30](#_Toc57842007)

[Figure 31 RB Category Returns 31](#_Toc57842008)

[Figure 32 HO Category Returns 31](#_Toc57842009)

[Figure 33 Crude returns in Category 32](#_Toc57842010)

[Figure 34 Crack returns in Category 32](#_Toc57842011)

[Figure 35 Hurricane Strength Indicator 33](#_Toc57842012)

# **Introduction**

# 

Natural disasters—hurricanes, floods, drought, and other weather-driven events—accounted for $232 billion worldwide damages in 2019 [1]. While these damages range from physical damage to infrastructures, there are externalities in financial risk due to disruption of business, and the driving up prices driven by fears of calamity. For example, hurricanes—which are prone to disruption of densely populated areas situated on the coast—there is a phenomenon of panic buying in essential items such as toilet paper, food, water and other supplies for emergency preparedness [2].

Such fluctuations in supply and demand are an essential part of the free market, and more importantly are prevalent in the commodities—natural or derived resources such as crude oil, natural gas, corn or soybeans—which frequently have swings in prices due to expectations of drought on crops, or more recently geopolitical tensions in the oil-rich Middle East on energy supplies [3]. Given that commodities are integral part of daily life, be it in direct form of agricultural commodities for consumption or indirectly as energy supplies for electricity use, it’s important to understand their precise relationship with natural disasters. This extends to risk mitigation for the entire commodity supply chain to forecast changes in commodity prices for insurance or planning purposes. In addition, as price volatility and liquidity tend to correlate, this relationship is important for the essential function of market speculators, who should understand these dynamics to provide liquidity in times of weather uncertainty.

The goal of this research is to specifically quantity and understand the role that hurricanes have on energy commodities [oil, gasoline, heating oil and refinery margins] with a focus on the Gulf of Mexico region. While physical energy suppliers generally have a rule of thumb to be vigilant during hurricane season, the author seeks to quantify that precise relationship and form insight into the weather signals that translate to price action, and explore whether this “weather risk premia” can form a reliable trading strategy. Seasonal time-series analysis, correlations and predictive modeling will be used between commodity data sets and past Gulf of Mexico storm data, to investigate whether the relationship is casual or there is a statistical significance behind weather risk.

# **Literature Review**

Early research into commodity prices centered around Keynes’s “Theory of backwardation” [4] [5] that explained backwardation as a frequent occurrence in commodity markets, due to commercial hedgers desire to eliminate risk. Working [6] introduced the “Theory of Storage” which incorporated said risk into cost-of-carry arbitrage within commodity markets, and Cooter [7] related the idea as a “risk premium” to compensate speculators for taking an opposing position of commercial hedgers. This risk premium was an excess return related to the transfer of risk. Bodie and Rosinsky [8] researched the historical performance of a wide basket of commodity futures and concluded a wide distribution and uncorrelated profile of returns that were also uncorrelated with tradition equity markets, and thus could be a source of diversification in investments. Hirsh Leifer [9] would later extend this work, and show that commodity risk premium would correlate with higher volatility or shocks to said markets.

In the early 2000s, investors clamored for higher returns, a hedge against inflation [10] and diversification from traditional equity markets, through allocations to different markets, including commodities. This gave rise to the modern commodities index, such as the Goldman Sachs Commodities Index (GSCI) and long-only commodity asset managers such as Gresham Investment Management. Such strategies centered around forming a long-only position in a long commodity and extracting value from the risk-premium through the tenor of said commodity’s term structure.

However Till [11] noted while such strategies could provide excess returns for investors, they needed to be actively managed with skill, due to the excess volatility and distribution of returns. Furthermore Till [12] [13] would later argue “value-add” through enhanced commodity strategies, that would deviate solely from a long only position, and instead target specific risk within an underlying commodity through short positions. One specific risk noted was “weather fear risk premia” which related fears of weather disruptions—for example excessive drought impacting crops or cold snaps in natural gas—and Till showed a statistical edge in returns for skilled managers.

Further research into the variance of risk premia [14] [15] [16] fluctuating long/short signals related to commodity risk-premia for enhanced returns, but as a function of statistical signals and not based on the fundamental—inventories, yield or other parameters—of the commodity. While trend-following or statistical oriented commodity asset managers (collectively called CTAs) are not new to the investment space, the rapid rise of computing power and quant revolution post 2008 has led to increasing allocations to systematic or quantitative strategies [17], which seek to trade commodities strictly off said statistical signals. Yet returns for said funds have stagnated or trailed the market [18], due to high volatility, rapidly changing commodity regimes and the disconnect of signals from underlying market fundamentals. The culmination of said inexperience was within a catastrophic blow up [19] of a Chicago based CTA due to unusual weather volatility in 2018.

Till revisited “weather risk premia” [20] [21] and noted superior risk returns and portfolios of commodities could be constructed, but needed to include fundamental information such as crop-balances or weather events, and once again be in the hands of a nimble and skilled manager. Furthermore, short term signals existed but instead of adopting a long-term strategy, tactical strategies would provide enhanced returns. Li [22] reinforced Till’s work by looking at passive strategies in agricultural markets, but found once again active management was key.

Thus, past literature has reinforced the existence and enhanced returns in weather risk premia for commodities, but focus has been on agricultural (corn, soy or wheat) markets through longer term drought modeling or weather forecast for natural gas, and not petroleum product commodities such as oil, gasoline or heating oil, or even intra-commodity spreads like crack margins. Finally, while weather risk has been previously modeled in the context of excess rains/drought or cold/hot weather, there has not been any analysis on tropical storm or hurricane activity, and its relationship to weather risk premium. While contemporary analysis [23] [24] of the frequency and strength of Gulf Coast Hurricanes concluded no discernible trend, the focus was not underpinned back to commodity markets.

The author seeks to extend the body of research by focusing on petroleum products and their pricing impacts due to hurricanes, due to natural link between infrastructure vulnerability and supply shocks to underlying market.

# **Theory and Hypothesis**

Speculation in the commodities market relies on the transfer of risk premia, and that the commodity markets are prone to shocks which spill-over to higher volatility, and thus higher risk premia. Furthermore, as weather events form a special group of risk premia, ‘weather risk premia”, and that previous literature has concluded the existence of excess return strategies that rely on tactfully trading said events by utilizing some sort of fundamental data. Thus, the author is led to believe that identifying or predicting a weather event could lead to profitable strategy with a statistical edge. More importantly from the literature review and the author’s own professional experience trading commodities, Atlantic basin hurricane activity could serve as consistent signals for a strategy based on petroleum products, due to the large percentage physical infrastructure situated in gulf coast which could be impacted.

## **Gulf Coast Petroleum Infrastructure**

The Gulf of Mexico is home to 17% of total United States crude oil production (both onshore and offshore), 5% of dry production, 45% of total refined products capacity and 51% of natural gas processing capacity [25]. In addition to the 175 + oil & gas rigs that are situated off-shore, the vast majority of refineries and processing plants are located within 5 miles from the coastline, clustered together within the cities of Houston, Port Arthur, Lake Charles and Baton Rouge.

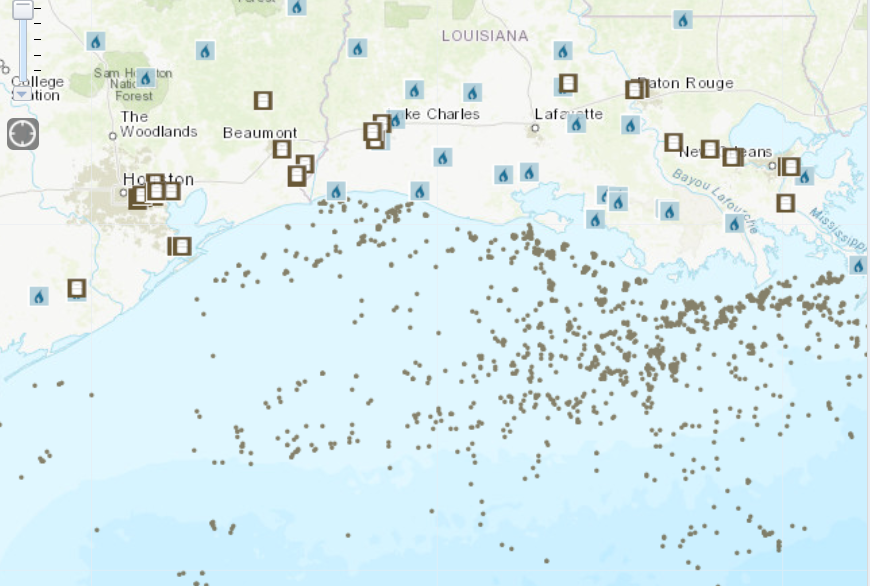


Figure 1: Gulf Coast Energy Infrastructure EIA [25]

Beyond refining and production assets, the Gulf coast is home to a network of pipelines that supply refined products from refineries, to destinations across the United States. One such critical pipeline is the Colonial Pipeline which is the largest pipeline in the United States and is responsible for over 100 million gallons [26] of daily fuel supply from the Gulf of Mexico to most of PADD1, and New York Harbor. Colonial Pipeline infrastructure begins with a series of pumping stations and connections to gulf coast refineries, and then carries products on two distinct lines to a hub in Greensboro, and then ultimately terminates in New York Harbor, supplying the vast majority of PADD3 and PADD1 with petroleum products [27].

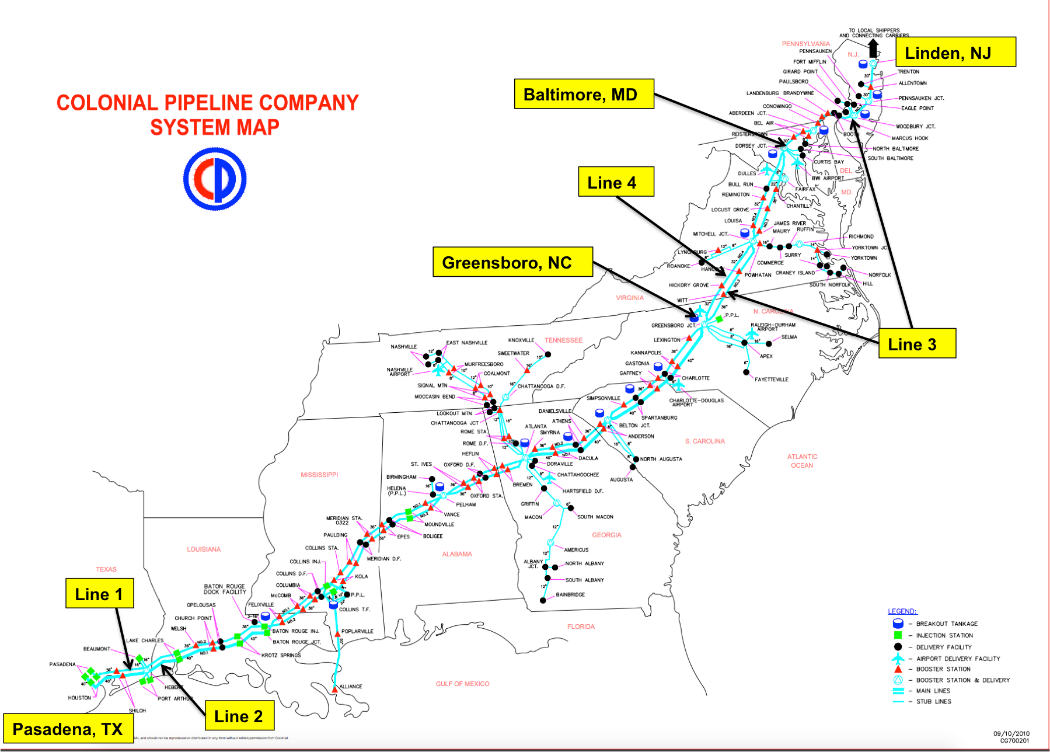


Figure 2 Colonial Pipeline System [27]

Given the closer proximity and density of both production, refining and transportation infrastructure in the Gulf of Mexico, there is a risk of supply shocks to the market by curtailed production through evacuations during a hurricane, unplanned power outages, and overall damage due to flooding, high-winds and storm surges. In fact surges in prices [28] have occurred during extreme hurricane events—in 2017 Hurricane Harvey flooded refineries and Colonial Pipeline operations, and even recently during 2020’s Hurricane Delta.

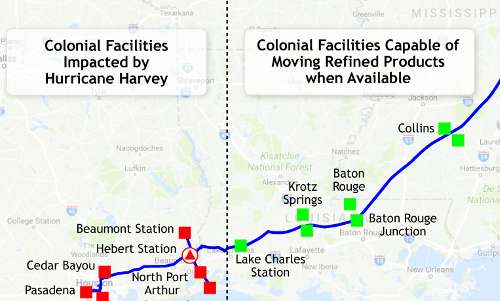


Figure 3 Colonial Pipeline Damage due to Harvey [29]

## **Commodity Markets**

Commodity markets cover a wide range of instruments that range from physical transactions, financial over-the counter (OTC) swaps and exchange cleared instruments, all of which are part of price discovery in the near $4.65 trillion dollar energy industry [30] . Futures are standardized contracts to buy or sell a pre-determined amount of a commodity—for example crude oil, gasoline or heating oil—and are the instrument of choice for speculators and commercial counterparties, due to liquidity, month specific expiry and ease of trading. Futures trade on the New York Mercantile Exchange (NYMEX) and have different month tenors for expiry, and thus are a reasonable approximate as the market’s outlook of price information and should reflect any price shocks due to hurricanes.

Price can be translated in different ways within the markets—a change in the simple price of a contract (referred to as flat price in the industry) is common and reflects a strictly directional change in value, and can’t be negative. A time spread involves the difference between the same commodities but at two different maturities and reflects the market’s view of the supply and demand of said commodity. When a time spread is positive and thus prompt or closer maturity future is pricing higher than a deferred counterpart, the curve is called **backwardation**, and thus reflects a shortage of supplies. The opposite, a negative spread between prompt and deferred, reflects **contango** and excess supplies. A near flat or close to 0 spread suggest short term balance or equilibrium.

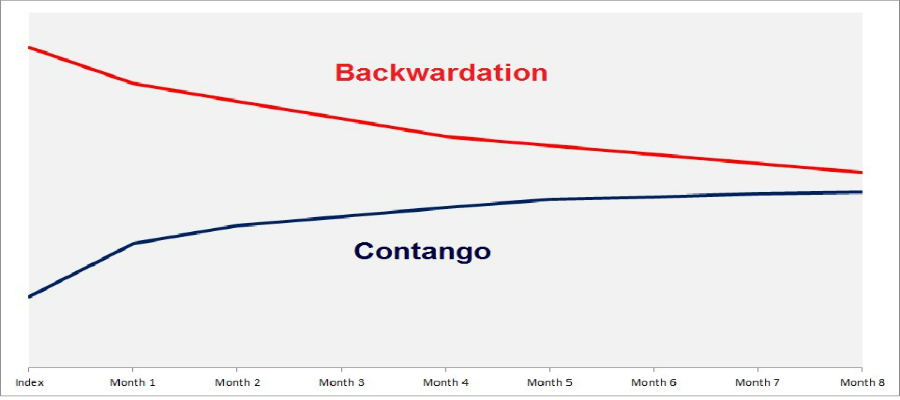


Figure 4 Backwardation vs Contango

Finally, ratios or spreads between different commodities of the same tenor are referred to as intra-commodity spreads and approximate the profitability of buying one commodity and selling another. More specifically as refineries process (or buy crude oil) and then sell refined products (heating oil and gasoline), the spread between them—commonly called the crack spread—is a proxy for how profitable a refinery is.

As mentioned previously, NYMEX contains futures that cover crude oil, gasoline, heating oil, and by imputation, crack spreads, and thus changes in the flat price or time-spread should reflect added risk, through primarily a surge in price (as a hurricane should raise the risk of infrastructure damage) and volatility.

## **Hurricanes in Gulf Coast**

Hurricane season typically runs from June to November each calendar year where weather conditions are conducive to storm creation in the Atlantic basin containing the Gulf of Mexico [31]. Peak storm intensity occurs from August to October. Storms taxonomy begins with a Tropical Wave which is a pressure aberration moving East to West in the Atlantic basic but can subsequently upgrade to a tropical depression, tropical Storm and ultimately a hurricane based on increasingly higher wind speeds [32]. Once a storm is categorized as a hurricane (more than 74 mph winds), the Saffir-Simpsons scale is used to delineate the strength and devastation of a hurricane (in terms of wind speed/storm surge potential) from 1 being weakest to 5 being strongest, and anything 3 or higher deemed a major hurricane [33].

Thus, the destructive power and therefore weather risk premium due to a hurricane should be a function of the strength of the hurricane (parameterized by the category scale) and the projected area of impact. While a higher category hurricane might be strong in nature, if it hits an area without significant infrastructure, for example southern Texas, there might be minimum impact if any to refining and other capability. However a hurricane that makes a direct hit on refining capability, for example the greater Houston area, can have a profound disruption in supply, even if the storm weakens to a non-major category—as was the case of Harvey which started out as a category 4 but stalled out over Houston [29].

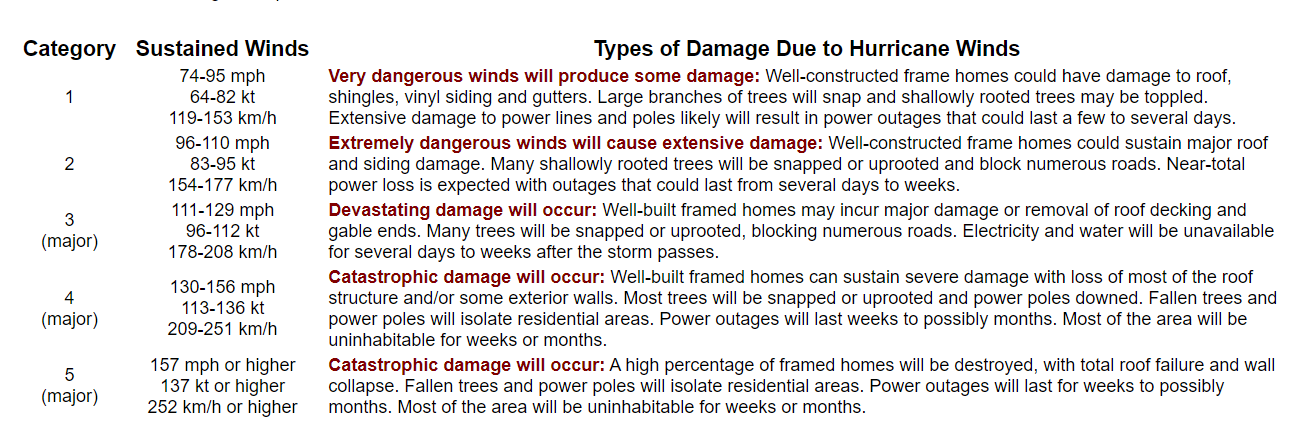


Figure 5 Saffir-Simpsons Hurricane Scale [33]

## **Hypothesis**

The previous background theory establishes the vulnerability of energy infrastructure to supply shock due to disruption by a hurricane. Naturally the hypothesis is that a higher weather risk premium should be reflected in the commodity’s future markets, and the possibility of a strategy and or predictor variables to exploit the higher weather risk premia in the form of excess returns.

Ultimately the author seeks to answer **whether hurricane season—be it the entire or sub season-- provides a statistically significant risk premia than the rest of the year or comparability to other years, whether such risk premia translates to higher commodity prices, and whether it is a viable signal or profitable strategy.**

# **Data and Variables**

This study will make use of two broad datasets—commodities time-series pricing data and National Hurricane Center (NHC) historical hurricane (HURDATA). The primary variables will be statistics of the commodity prices—for example returns or volatility—and their relationship with the hurricane data.

## **Commodities Futures Data**

Commodity data will be derived from Quandl [34] NYMEX end of day settles for the crude oil (CL), gasoline (RB), and heating oil (HO) futures contract, spanning the last 10 years (2010-2009) of active trading days. Quandl provides a Python friendly API to load time-series data for end of day futures, however as mentioned previously, futures contracts are defined for a specific month. This can be problematic as commodity future typically trade many months (or years) prior to their expiry, and thus for a single year there can be at least 12 different time-series per instrument, and thus unruly from a data-analysis perspective.

However since trading volume and liquidity [ability to transact] are typically highest on near-month contract, and risk events i.e. hurricanes typically impact the immediate future, the author decided to convert the aggregate time-series to a single time-series per instrument that reflects the near-month (month +1) price action and maximum trading volume/liquidity. This transformation is referred to as a continuous futures time-series and methodology was derived from Quantopian [35]. The resultant—as seen below—are a single time-series for each instrument.

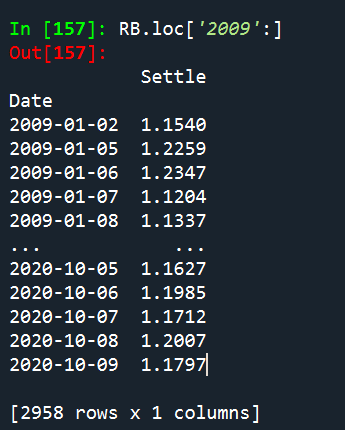


Figure 6: RBOB (Gasoline) Continuous M+2 Time-series

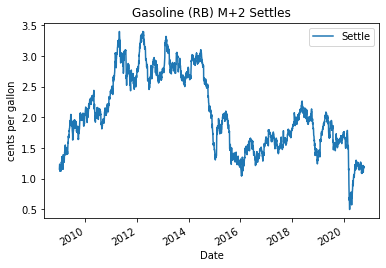
Time-spreads can be imputed by subtracting an instruments time-series versus the next month continuous time-series, and commodity spread (cracks) can be calculated by subtracting different instrument time-series. For simplicity, a refinery crack will be defined via the EIA 3-2-1 model, using a ratio of 3 crude oil futures contracts, to 2 gasoline (RB) and 1 heating oil (HO) contract.

Figure 7: RB (Gasoline) Month+2 Continuous Price Series chart

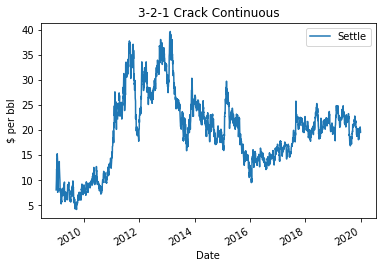


Figure 8: Imputed 3-2-1 refinery crack price chart

## **Commodities Seasonality**

To compare the seasonality, or differences between commodity prices for the same period for different years or frequencies, the author wrote a custom function to decompose any single time series by it is week, month, and year components. Pivot tables and plots, as seen below, can thus be easily constructed for comparisons.

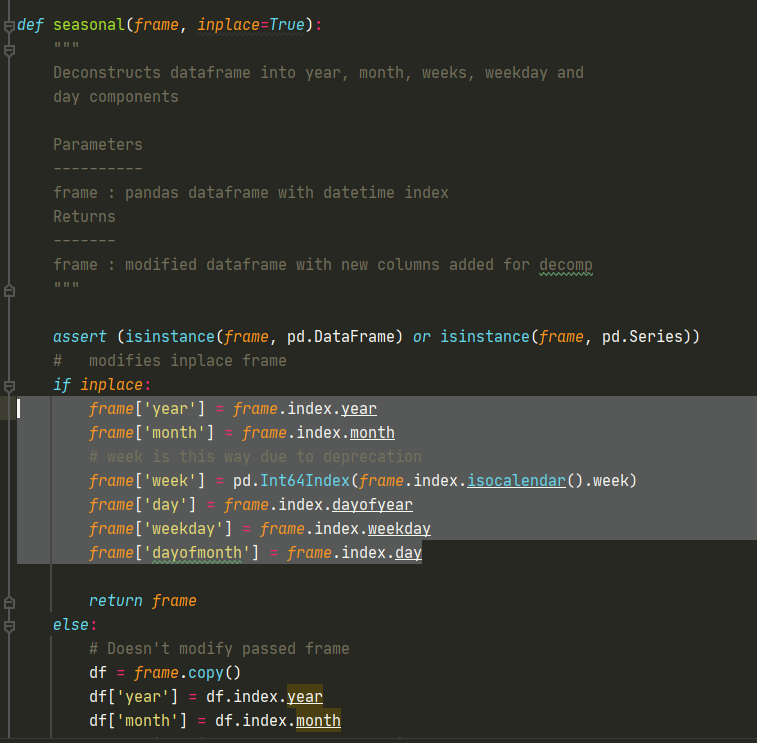


Figure 9 Snippet of Custom Seasonality

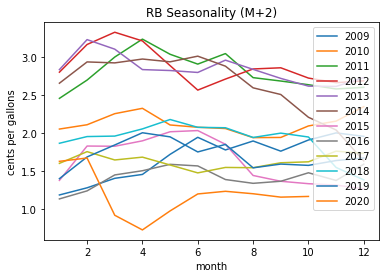


Figure 10: RB M+2 Seasonality

## **Hurricane Data**

The National Hurricane Center (NHC) provides detailed data on all tropical storms, depressions and hurricanes in a historical dataset text file called HURDAT2 [36]. HURDAT covers all activity from 1851 to the prior year (2019 in this case) and provides the following fields:

1. **Basin Location** (Atlantic, Northcentral Pacific of Northeast Pacific)
2. **Year and Number of Cyclone**
3. **Name (if applicable)**
4. **Time stamps** with tracking information for
   1. Landfall
   2. Classification Changes from a disturbance (DB) to Hurricane (HU)
   3. Latitude and Longitude positional data
   4. Maximum sustained winds
   5. Maximum pressure
   6. Wind speeds from quadrants of the storm

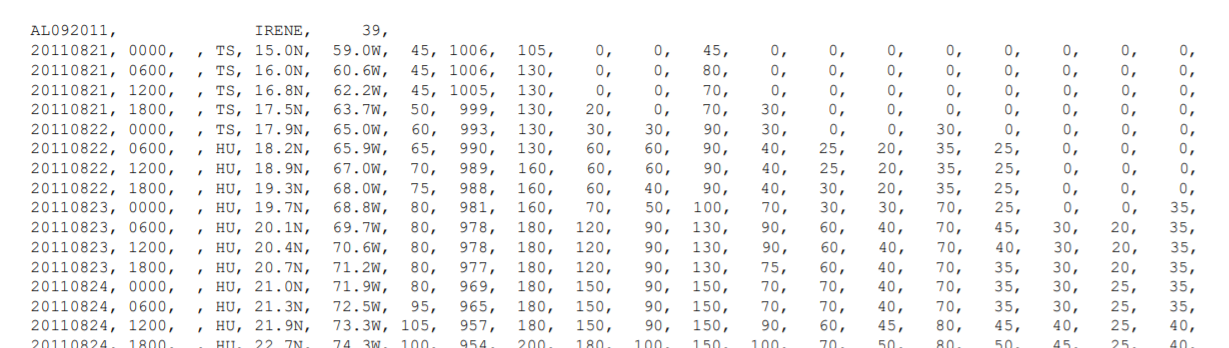


Figure 11 Example HURDAT2 for Irene

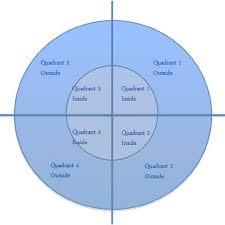


Figure 12 Wind Quadrants

The parameters provided by HURDAT2 give an effective timeline that characterizes the trajectory, strength, and evolution of a hurricane, and can thus be aligned with pricing data to see any impact.

Data was ingested via the HURDAT2 text file, and then cleaned into a “tidy” format through author derived helper functions.

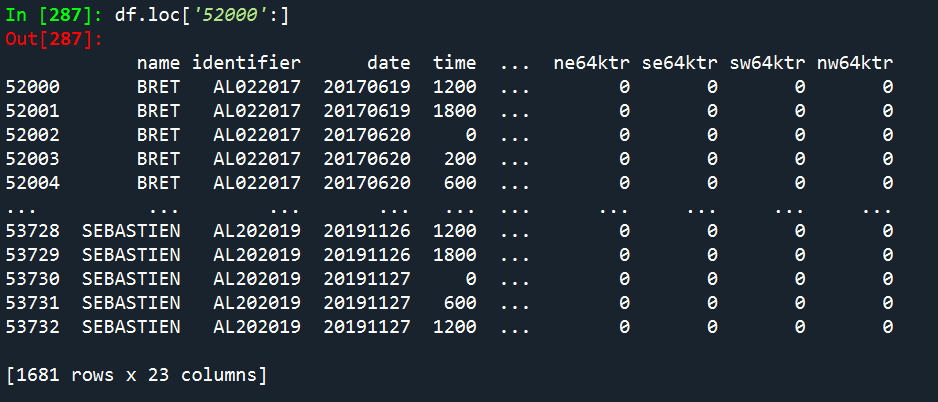


Figure 13 Cleaned up HURDATA in dataframe

The resultant was a time-series of each storm with position and meteorological parameters. While the Sariff-Simpson category is not explicitly contained in the data, the author wrote a function to classify a hurricane (marked HU) from 1 to 5 based on the current wind speed.

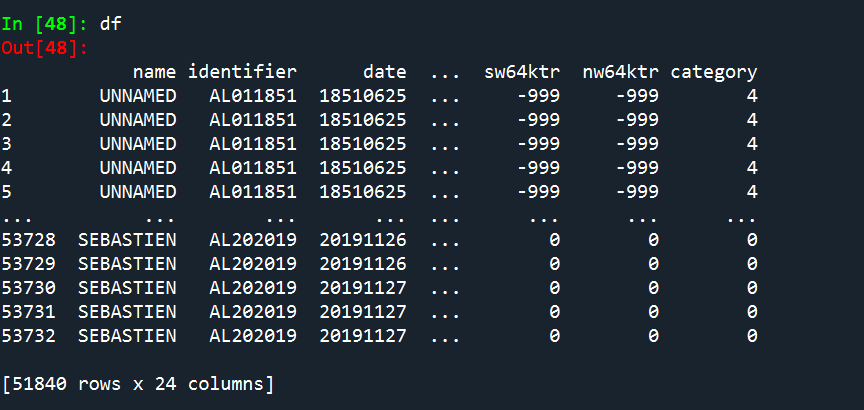


Figure 14 HURDATA with Imputed Category

The following is a snippet of the partial code used to impute and clean up the data-set.



Figure 15 HURDAT2 Hurricane Map

The cleaned and imputed data was saved to a .csv for quick loading and processing.

# **Statistical Methods**

The primary research goal is to identify relationships between energy commodity prices signals—time spreads , crack margin and hurricane activity in the Gulf of Mexico. As markets can be influenced by a multitude of factors, including prices between commodities, it will be necessary to analyze commodity prices and hurricane activity separately, and then establish (if any) there is a link between them.

For commodity prices—specifically time spreads for crude oil, heating oil, gasoline and refinery cracks the following time-series and statistical techniques will be used for the 2009-2019 Hurricane season. The primary variables will be price and returns. To simplify the scales, both RB and HO will be converted from cents per gallon to cents per barrel (42 gallons = 1 barrel).

* Boxplots, summary statistics and Distribution of prices and returns.
* ACF plots to see if there is trend, seasonality, or correlation within commodity time periods.
* Correlation between commodities.
* Normality tests (if appropriate) to compare mean returns and prices

For the hurricane analysis, the goal is to identify if there are any trends in overall hurricane frequency (be it formation, strength, or land fall) within each year, hurricane season or intra-season and will utilize the following techniques:

* Boxplots, summary statistics and Distribution of hurricane frequency, location genesis, landfall and category
* ACF plots test to see if there is trend, seasonality, or correlation within hurricane formation and strength

Finally, to establish (if any) there is a link between these two phenomena the following methods will be used for comparison for overall hurricane season and intra-season. As the thesis is that price should be impacted, a lagged series of commodity prices will be compared to hurricane metrics

* Comparison of price and returns, and correlation within identified trends within hurricane season
* Identify benchmark trading strategies based on said trends

# **Findings**

## **Commodity Time Spread and Cracks Analysis**

Dataframes containing time spreads of crude oil, heating oil, gasoline and flat price refinery cracks were created from the pricing dataset and then resampled to a weekly timeframe, as daily analysis tends to have significant noise. Missing NaN values represented less than .05% of the total datasets and were therefore dropped.

Figures 15-23 contain time-series plots, summary statistics, box plots, histogram, correlations and ACF for the price series

Initially the following were observed

* Refinery crack margins was the only series that took positive values and had the widest range in values and dispersion (standard deviations). This makes sense given that the crack margin is a spread between multiple commodities and refineries would not operate at negative margins.
* Time-spreads themselves varies from negative to positive values, which as an intra-commodity represents supply-demand shifts from backwardation to contango. Crude oil and heating oil mean over the timeframe were negative indication primarily an oversupplied market, while gasoline was positive indicating tendency for backwardation.
  + Refinery margin are not stationary as confirmed visually and by ADF test
  + All other time-spreads were stationary (ADF-test) which also makes sense as spreads tend to be more stable and in balance than just flat price.
* From the boxplots and histogram, gasoline time-spreads had much higher dispersion compared to crude oil and heating oil.
* Visual inspection of the price distributions reveals non-normality and further Shapiro-Wilkins test [Table 1] rejected the null hypothesis.
* All prices had a weak positive correlation with each other which makes sense due to the economic link of all them as oil products.
* ACF plots and time-series plot show that there is strong serial correlation/trend within refinery crack margins—which isn’t surprising as it’s once again a function of three independent price series and financial time-series tend to have autocorrelation/trend—serial correlation/trend in crude, weak seasonality in HO an RB.

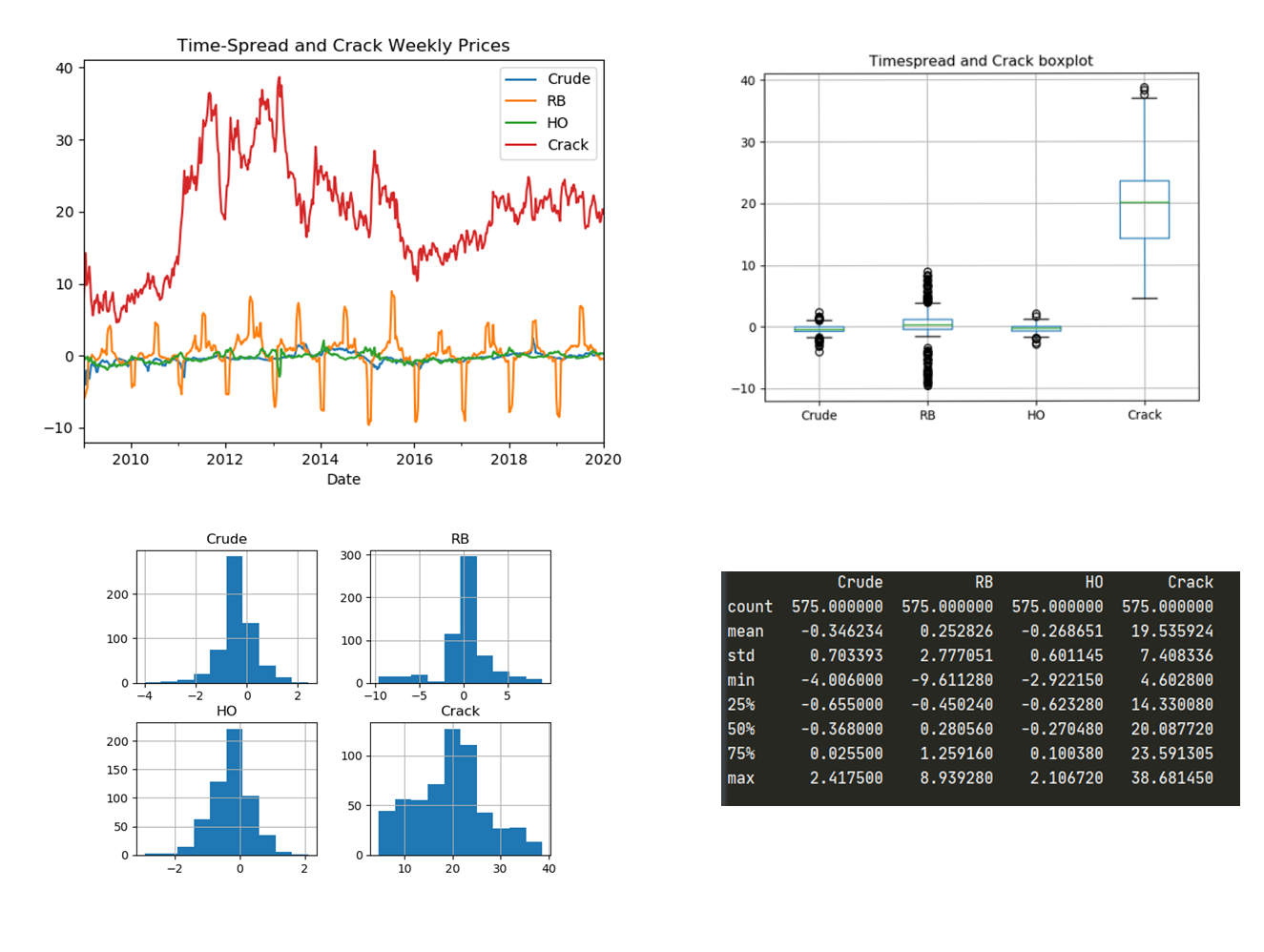


Figure 16 Summary Weekly Plots

|  |  |  |
| --- | --- | --- |
|  | **ADF** | **p** |
| **Crude** | -4.8 | 0.00005 |
| **HO** | -3.13 | 0.02300 |
| **RB** | -6.06 | 0.00000 |
| **Crack** | -2.4 | 0.13000 |

Figure 17 ADF Statistics for Prices

|  |  |  |
| --- | --- | --- |
|  | **Shapiro** | **p** |
| **Crude** | 0.937 | 7.58E-15 |
| **HO** | 0.941 | 3.00E-14 |
| **RB** | 0.924 | 2.10E-14 |
| **Crack** | 0.98 | 1.45E-06 |

Table 1 Shapiro Normality test for Prices

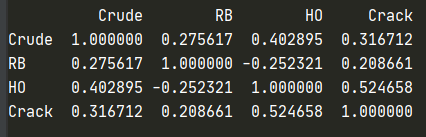


Figure 18 Pearson Correlations

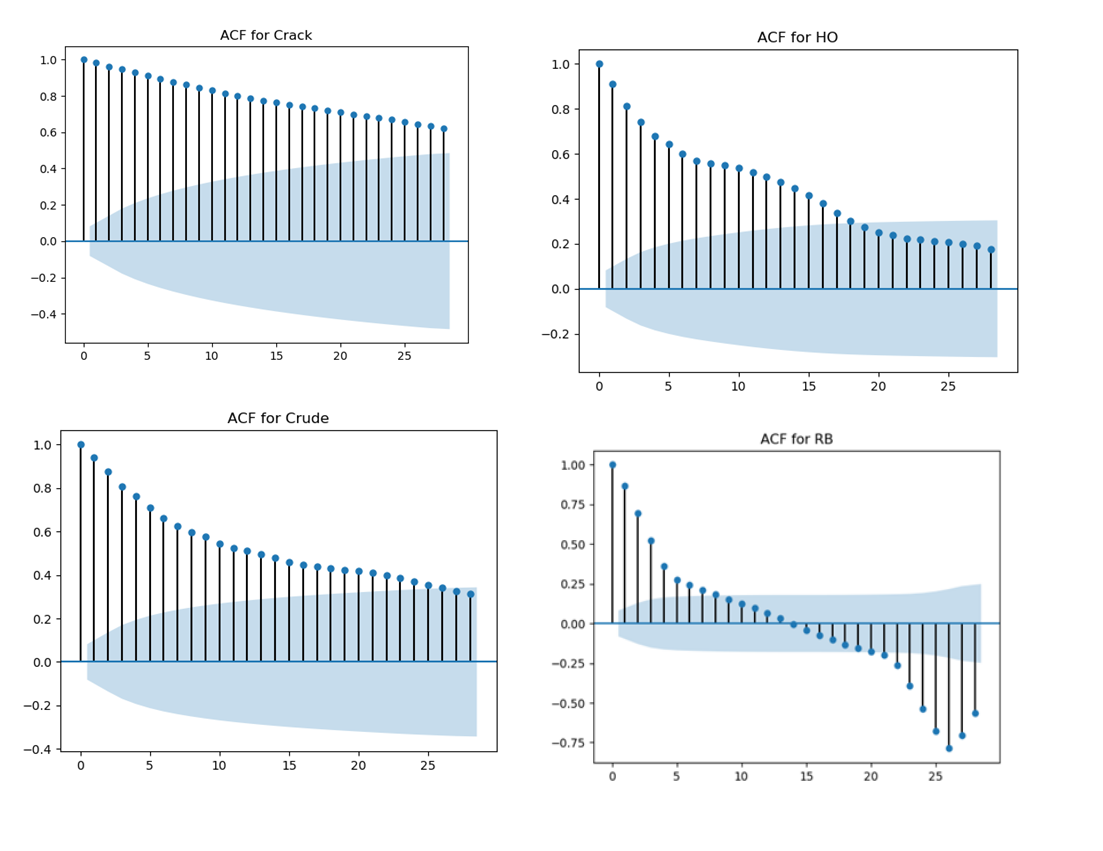


Figure 19 Price ACFs

## **Commodity Time Spread and Cracks Data Returns Analysis**

Trading strategies seek returns, not just price, and thus the dataset was converted to returns to demonstrate investment potential. Log-normal returns were used for cracks while time-spreads can be negative, simple arithmetic returns were used.

Initially the following were observed

* Weekly returns for all commodity groups appear to be non-normally distributed (Shapiro Test and with heavy tails (consistent with most assets). ACF plots show some degree of lack of serial correlation, and curiously enough RB and crude appear as discrete white noise.
* Returns were all stationary[ADF test]—this is curious given the time-series showed outliers.
* Mean weekly returns for Crude and HO were negative (losing money) while cracks and RB were positive. Most were centered or near 0, which is further corroborated by the ACF resembling discrete white noise, however they don’t appear normally distributed via a Shapiro test.
* Pearson correlations showed little cross-correlation return wise between the commodities.

Chart, histogram

Description automatically generated

Figure 20 Weekly Returns

Chart

Description automatically generated

Figure 21 Returns Histogram

|  |  |  |
| --- | --- | --- |
|  | **Shapiro** | **p** |
| **Crude** | 0.16 | 1.82E-44 |
| **HO** | 0.45 | 3.05E-44 |
| **RB** | 0.19 | 6.00E-44 |
| **Crack** | 0.95 | 1.29E-11 |

Table 2 Shapiro Normality Test for Returns

|  |  |  |
| --- | --- | --- |
|  | **ADF** | **p** |
| **Crude** | -23 | 0.00000 |
| **HO** | -21 | 0.00000 |
| **RB** | -24 | 0.00000 |
| **Crack** | -6 | 0.00000 |

Table 3 ADF results for Returns

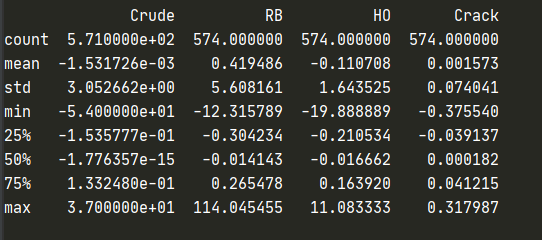


Figure 22 Returns Summary Stats

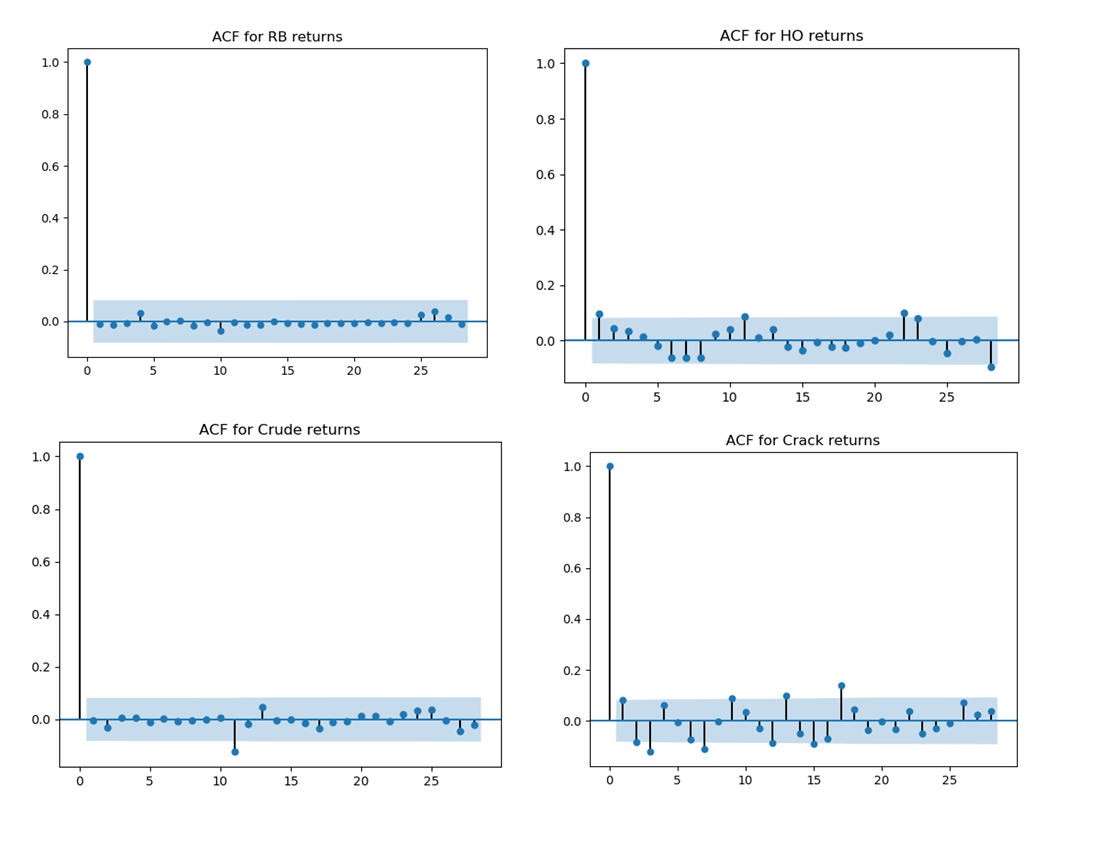


Figure 23 Returns ACF

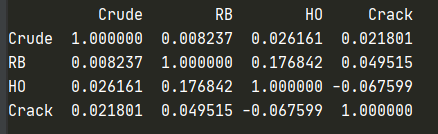


Figure 24 Pearson Returns

## **Hurricane Data Initial Analysis**

Turning to the Hurricane Data set, data was cleaned, formatted, category strength imputed (as mentioned in the previous sections) and cleared of NaNs. The first area of interest was the frequency of hurricanes and whether there was any truth to “more frequent hurricanes”. To analyze , plots for named hurricane genesis dates were created by year and month, as well as related ACF plots.

Chart, line chart

Description automatically generatedChart, box and whisker chart

Description automatically generated

Figure 25 Number of Hurricanes Per Year

On the yearly frequency or ACF there does not seem to any significant serial correlation though the ACF lags point to some cyclical nature. Moving to a more granular monthly analysis:

Chart, histogram

Description automatically generatedChart, histogram

Description automatically generated

Figure 26 Monthly Hurricane Frequency and ACF

Here it is much more apparent to see the cyclical nature per year of hurricane season [cyclical spikes and negative correlations of ACF spikes], and a degree of serial correlation over the years. While this dataset is only 10 years of data, it appears that in terms of frequency of more hurricanes occurring in the Gulf of Mexico, there are just as many that form now than 10 years ago. However there seems to be a cycle of more hurricanes forming, falling, and then repeating itself—which upon further research is reinforced by the La Nina/El Nino cyclical patterns which are the genesis of hurricanes. At the time of writing this the 2020 Hurricane season is the most active, which fits the ramp up from 2018 onward.

Moving on to a seasonal month study—a histogram of named hurricane counts per month was constructed, as follows, and clearly reinforces that NOAA’s definition of hurricane season from May/June to November of each year, with a peak formation in August-October.

**Chart, bar chart, histogram

Description automatically generated**

Figure 27 Frequency of Hurricanes by Month

As mentioned in the introduction, a Hurricane’s strength is characterized by the Saffir-Simpsons scale, and would give insights into whether “stronger” hurricanes are forming. Histograms and ACFs were analyzed based on the mean category (which were imputed previously) across all hurricane duration.

Chart

Description automatically generatedHistogram

Description automatically generated

Chart, histogram, box and whisker chart

Description automatically generated

Figure 28 Category Strength and Plots

The results here are much more interesting—not only are the mean category strength higher over the last 5 years [compared to the overall mean across the time period] but the ACF plot shows cyclical strength in categories—that is there is the same pronounced cycle of hurricane strength rising and then falling.

Based on this analysis there is evidence that while the **number** of hurricanes formed has not significantly changed and is indeed cyclical in nature, the **strength** recently has.

# **Discussions**

The previous findings showed that while frequency of hurricane seems to be cyclical, recent hurricanes have been stronger and there more pose more destructive power. To relate this phenomenon back to commodity returns, statistics across the overall year, hurricane season, and intra-season (as reinforced by the previous section) were compiled.

Non-Hurricane season was defined as week’s 1-22 (encompassing June to May), Hurricane season as (22-48) and Intra-Hurricane (32-44).

Given that returns and prices are not-normal, to compare significance in means, boot-strap aggregation (n=1000) were calculated—the author programmed his own bootstrapper function as seen below.



Figure 29 Bootstrap Code Snippet

The following were concluded for Returns:

* Crack mean returns outside of Hurricane season [.018] were positive, and significantly higher than returns in season [-.012] and intrasession [-0.013], which were statistically similar.
* Crude mean returns within all seasons were statistically similar, and all negative [Full- -0.28, Hurriane-.007 and Intra -.05]
* HO mean returns outside hurricane season [-.228] were statistically similar to within season [-.22] and intra-season [-.13]
* RB mean returns within hurricane season [.12] were statistically different than the full season [-.21] and [-.16].

Chart, histogram

Description automatically generated

Figure 30 Bootstrap Returns for Seasons

Next to see the impact of category strength the dataset was now decomposed into two sets 2009-2014, a period of weaker hurricanes (by category) and 2015-2019, a period of stronger hurricanes (by category).

For RB

* Mean returns within hurricane season were highest in weaker category [0.19 vs .043], but were positive and statistically different than non-hurricane season.

Chart, histogram

Description automatically generated

Figure 31 RB Category Returns

For HO

* Mean returns within a weak category were all statistically similar , and negative, while mean returns in a strong hurricane category were similar for non-hurricane/hurricane season, and both higher than intra-hurricane season.

Chart, histogram

Description automatically generated

Figure 32 HO Category Returns

For Crude

* Mean returns in both a strong and weak hurricane category were all statistically significant, and near 0.

Chart, histogram

Description automatically generated

Figure 33 Crude returns in Category

For Cracks

* Mean returns were highest, and positive for non-hurricane season [.016 and .02] and statistically significant/higher than the hurricane seasons.

Chart, histogram

Description automatically generated

Figure 34 Crack returns in Category

## **Strategy Design**

Based on the previous discussion the following lessons were gleamed

* RB are the only commodity to show positive returns within hurricane season, with slightly higher returns in a weaker storm.
* Cracks show highest positive returns outside hurricane season
* Heating oil and crude are both neutral, or negative within the time-frame.

A simple strategy was devised based on RB solely, due to its statistically significant positive returns within hurricane season, and simplicity is being modeled as a long position [hence positive returns]. A “category strength” momentum indicator was imputed as the difference between category strengths, to further fine-tune entry for weakness and strong category.

Chart, box and whisker chart

Description automatically generated

Figure 35 Hurricane Strength Indicator

To test the strategy the hurricane data was split-into test train splits randomly, but by year, meaning that each year’s data was consecutive but the training set could consist of out of order years; this was necessary due to the serial correlation common in financial time series.

The author intentionally left out tweaks to the strategy but both in-sample and out of sample returns were statistically significantly positive, and higher than the previous boot-strap returns

# **Conclusion**

Based on the studies and analysis of this report the following conclusions were drawn.

* While hurricane activity follows a cyclical pattern spanning 10 years there is evidence to suggest that recent hurricanes have indeed been stronger.
* Hurricane strength and activity does not have a significant impact on Heating Oil or Crude time-spreads, indicating that there isn’t any weather risk premia.
* However both hurricane activity, season and category do translate to weather-risk premia in gasoline (RB) spreads, indicating that said markets are more vulnerable to disruptions, and demand-shock is present.
* Cracks (refinery margins) have decreased weather risk premium within hurricane season.
* Strategies can be implemented to exploit said hurricane related phenomena, leading to significant returns.

# **Future Work**

The research conducted can be further enhanced in the following ways:

* While the focus was on time-spreads as they represent the entire curve balance, another analysis strictly on prices would highlight one-direction demand shocks
* The author focused on “long” returns in prices increasing due to demand shocks, but futures allow shorting (selling) and thus negative returns previous attributed to heating oil or cracks, can form a profitable strategy by shorting.
* Weekly samples were used for simplicity but higher frequency data could be analyzed for more detailed analytics.

# **References**

|  |  |  |
| --- | --- | --- |
| [1] | | STATISTA, "Global economic losses from natural disasters 2000-2019," [Online]. Available: https://www.statista.com/statistics/510894/natural-disasters-globally-and-economic-losses/#:~:text=As%20of%202019%2C%20the%20global,232%20billion%20U.S.%20dollars%20annually.&text=There%20are%20a%20lot%20of,hurricanes%2C%20earthquakes%2C%20and%20tsunamis.. |
| [2] | | E. Zwiebach, "HURRICANE BRINGS PANIC BUYING," Supermarket News, [Online]. Available: https://www.supermarketnews.com/archive/hurricane-brings-panic-buying. |
| [3] | | D. Hodari and R. Dezember, "Middle East Clashes Send Oil Sharply Higher," The Wall Street Journal, 20 June 2019. [Online]. Available: https://www.wsj.com/articles/oil-prices-rise-sharply-as-middle-east-tensions-worsen-11561022592. |
| [4] | | J. Keynes, "Some Aspects of Commodity Markets," *Manchester Guardian Commercial: European Reconstruction Series,* 1923. |
| [5] | | J. Keynes, "Theory of Normal Backwardation," in *Treatsie on Money*, 1930. |
| [6] | | H. Working, "Price Relations between July and September Wheat Futures at Chicago Since 1885," *Wheat Studies of the Food Research Institute,* 1933. |
| [7] | | P. Cootner, "Returns to Speculators: Telser versus Keynes," *Jouranl of Political Economy,* 1960. |
| [8] | | B. Zvi and V. Rosanksy, "Risk and Return in Commodity Futures," *Financial Analysts Journal,* pp. 27-31, 1980. |
| [9] | | D. Hirshleifer, "Determinants of Hedging and Risk Premia in Commodity Futures Markets," *The Journal of Financial and Quantitative Analysis,* pp. 313-331, 1989. |
| [10] | | R. Greer, "Commodities – Commodity Indices for Real Return and Diversification," in *The Handbook of Inflation Hedging Investments* , McGraw Hill, 2005. |
| [11] | | H. Till and J. Eagleeye, "The Risk of Commodity Investing," August 2006. [Online]. |
| [12] | | H. Till and J. Eageleye, "Commodities-Active Strategies to Enhance Return," *Journal of Wealth Management,* pp. 42-61, 2005. |
| [13] | | H. Till, "Case Studies and Risk Management in Commodity Derivatives Trading," in *Risk Management in Commodity Markets: From Shipping to Agriculturals and Energy,*, Chichester, John Wiley & Sons, 2008, pp. 255-291. |
| [14] | | M. Syzmankowsa and F. De Roon, "An Anatomy of Commodity Futures Risk Premia," *The Journal of Finance,* 2013. |
| [15] | | A. Trollers and E. Schwartz, "Variance Risk Premia in Energy Commodities," *The Journal of Derivatives ,* pp. 15-32, 2010. |
| [16] | | M. Prokopczuk and C. Wese Simen, "Variance Risk Premia in Commodity," 14 January 2013. [Online]. Available: https://efmaefm.org/0efmameetings/efma%20annual%20meetings/2013-Reading/papers/EFMA2013\_0247\_fullpaper.pdf. |
| [17] | | RCM-Alternatives, "RISE OF THE ROBOTS: THE QUANT REVOLUTION," 2017. [Online]. Available: https://www.rcmalternatives.com/2017/07/rise-of-the-robots-the-quant-revolution/. |
| [18] | | R. Wigglesworth and L. Fletcher, "Trend-following hedge funds struggle in topsy turvy year," Financial Times , [Online]. Available: https://www.ft.com/content/5ea09868-ecc0-47d1-aa5c-57d33af543f4. |
| [19] | | G. Banerji, "Energy Losses Ruin Options Firm," Wallstreet Journdal , November 2018. [Online]. |
| [20] | H. TIll, "Global Commodities Applied Research Digest," 2019. [Online]. | |
| [21] | | H. Till, "A Brief Primer on Commodity Risk Management," 2016. [Online]. Available: http://www.jpmcc-gcard.com/wp-content/uploads/2016/12/GCARD-CEC-Brief-Primer-Fall-2016.pdf. |
| [22] | | L. Ziran, D. Hayes and K. Jacobs, "The weather premium in the corn market," *Wiley Futures Market,* 2017. |
| [23] | | M. Bove, D. Zierden and J. O'Brien, "Are Gulf Landfalling Hurricanes," *Bulletin of the American Meteorological Society ,* 1998. |
| [24] | | P. Koltzbah, "Continental U.S. Hurricane Landfall Frequency and Associated Damage: Observations and Future Risks," *Bull. Amer. Meteor. Soc. ,* 2018. |
| [25] | | EIA, "Gulf of Mexico Fact Sheet," 22 June 2020. [Online]. Available: https://www.eia.gov/special/gulf\_of\_mexico/data.php. |
| [26] | | C. Pipeline, "About Us," Colonial Pipeline, [Online]. Available: https://www.colpipe.com/about-us. |
| [27] | | RBN-Energy, "Colonial Pipeline," [Online]. Available: https://rbnenergy.com/space-oddity-congestion-on-the-colonial-refined-products-pipeline. |
| [28] | | A. Sider, "Gasoline Prices Jump in Harvey’s Wake," The Wall Street Journal, [Online]. Available: https://www.wsj.com/articles/gasoline-prices-surge-as-harveys-impact-is-felt-1504202779. |
| [29] | | N. A. Pipelines, "Colonial Pipeline to Restore Full Operations After Hurricane Harvey," [Online]. Available: https://napipelines.com/colonial-restore-operations-harvey/. |
| [30] | | C. Casella, "When The Fossil Fuel Industry Pops, It Will Be Way Bigger Than The 2008 Financial Crisis," 5 June 2018. [Online]. Available: https://www.sciencealert.com/fossil-fuel-industry-pops-way-bigger-than-2008-financial-crisis-renewable-energy. |
| [31] | | Wikipedia, "Atlantic Hurricane Season," [Online]. Available: https://en.wikipedia.org/wiki/Atlantic\_hurricane\_season. |
| [32] | | N. W. Service, "TROPICAL DEFINITIONS," [Online]. Available: https://www.weather.gov/mob/tropical\_definitions#:~:text=A%20tropical%20storm%20is%20a,(34%20to%2063%20knots).&text=A%20hurricane%20is%20a%20tropical,(64%20knots%20or%20greater).. |
| [33] | | N. H. Center, "Saffir-Simpsons Hurricane Scale," [Online]. Available: https://www.nhc.noaa.gov/aboutsshws.php. |
| [34] | | Quandl, "Quandl," [Online]. Available: https://www.quandl.com/. |
| [35] | | Quantopian, "Continuous Futures," [Online]. Available: https://www.quantopian.com/tutorials/futures-getting-started#lesson4. |
| [36] | | National Hurricane Center, "NHC Data Archive," [Online]. Available: https://www.nhc.noaa.gov/data/#hurdat. |
| [37] | | EIA, "3:2:1 Crackspread Explained," 4 2 2013. [Online]. Available: https://www.eia.gov/todayinenergy/includes/crackspread\_explain.php. |
| [38] | | EIA, "Hurricane Harvey caused U.S. Gulf Coast refinery runs to drop, gasoline prices to rise," 11 9 2017. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=32852. |
| [39] | | K. Parasuraman, "Hurricane Florence — Building a Simple Storm Track Prediction Model," 24 September 2018. [Online]. Available: https://towardsdatascience.com/hurricane-florence-building-a-simple-storm-track-prediction-model-1e1c404eb045. |
| [40] | | N. W. Service, "Tropical Definitions," [Online]. Available: https://www.weather.gov/mob/tropical\_definitions#:~:text=A%20tropical%20storm%20is%20a,(34%20to%2063%20knots).&text=A%20hurricane%20is%20a%20tropical,(64%20knots%20or%20greater).. |
| [41] | | N. D. Archive, "https://www.nhc.noaa.gov/data/#hurdat," [Online]. Available: https://www.nhc.noaa.gov/data/#hurdat. |