

Oil Price Volatility

James P. Hamski *City University of New York*

Oil price volatility ...

Keywords: oil price, commodities, time series analysis, volatility clustering

Acknowledgements

I would like to thank John Kemp, Thompson Reuters energy journalist, for posing the question investigated in this research paper. In addition, I thank Tancred Lidderdale and Mason Hamilton from the U.S. Energy Information Administration for additional information pertaining to the question.

Introduction

In a commodity trading market the price level is expected to be tied to the system dynamics. Volatility, the variation in price over time, reflects uncertainty in the supply, demand, and delivery of the commodity being traded. This research paper aims to answer a simply posed question: “does oil price volatility scale with price?”; i.e., can we expect to observe larger price swings when the price is near \$100/barrel vs \$20/barrel?

If oil price volatility reflects uncertainty about supply and demand dynamics, it isn’t immediately clear whether we should expect volatility to depend on price level. Higher oil prices are associated with “tightness” in the supply market, meaning there is little excess capacity to increase production. However, factors such as storage dynamics, the ability of producers to increase production fast enough to bring more oil to market in response to high prices (“rebalancing”), and market speculation complicate this picture and mean it must be studied empirically.

If it is found that oil price volatility is dependent on price level, the relationship may follow a scaling formula. For instance, if we can expect volatility of \$1/barrel when oil is at \$20/barrel, can we expect volatility of \$5/barrel at \$100/barrel price levels via a simple linear scaling rule? Three methods are presented in this research paper to answer this question: (1) regression modeling of price and volatility, (2) viewing volatility within oil price regimes, and (3) using multivariate

Generalized Autoregressive Conditional Heteroskedasticity (GARCH) modeling.

Note that in this research paper *oil price* is used to specifically mean spot-traded crude oil. This represents only one component of the oil markets, and most of the actual oil price is determined by futures and long term delivery contracts (need cite). This research paper is concerned with understanding the energy system using pricing information. In this way, it differs from much of the published research in that it is not concerned with forecasting prices or volatility. Nor is it addressing exogeneous system elements such as equity markets or interest rates, though the literature shows that the crude oil market and larger economic indicators are intertwined (add cite). Instead, it contributes to our understanding of the system dynamics of an essential energy commodity.

```
library(broom)
library(changepoint)
library(knitr)
library(magrittr)
library(TSA)
library(xts)
```

```
load("wti_project.Rda")
```

Exploratory Data Analysis

Data Source

The data source is the West Texas Intermediate (WTI) nominal (i.e. not inflation adjusted) daily spot price record from the U.S. Energy Information Administration. The WTI series was filtered to the date range January 2, 1986 through December 30, 2016.

```
summary(wti.xts) %>% tidy()
```

```
##   Var1      Var2      Freq
## 1      Index Min. :1986-01-03
```

```
## 2          Index 1st Qu.:1993-09-01
## 3          Index Median :2001-06-11
## 4          Index Mean   :2001-06-19
## 5          Index 3rd Qu.:2009-03-31
## 6          Index Max.   :2016-12-30
## 7          wti.xts      Min.    : 10.25
## 8          wti.xts      1st Qu.: 19.38
## 9          wti.xts      Median : 28.01
## 10         wti.xts      Mean    : 42.87
## 11         wti.xts      3rd Qu.: 63.47
## 12         wti.xts      Max.    :145.31
```

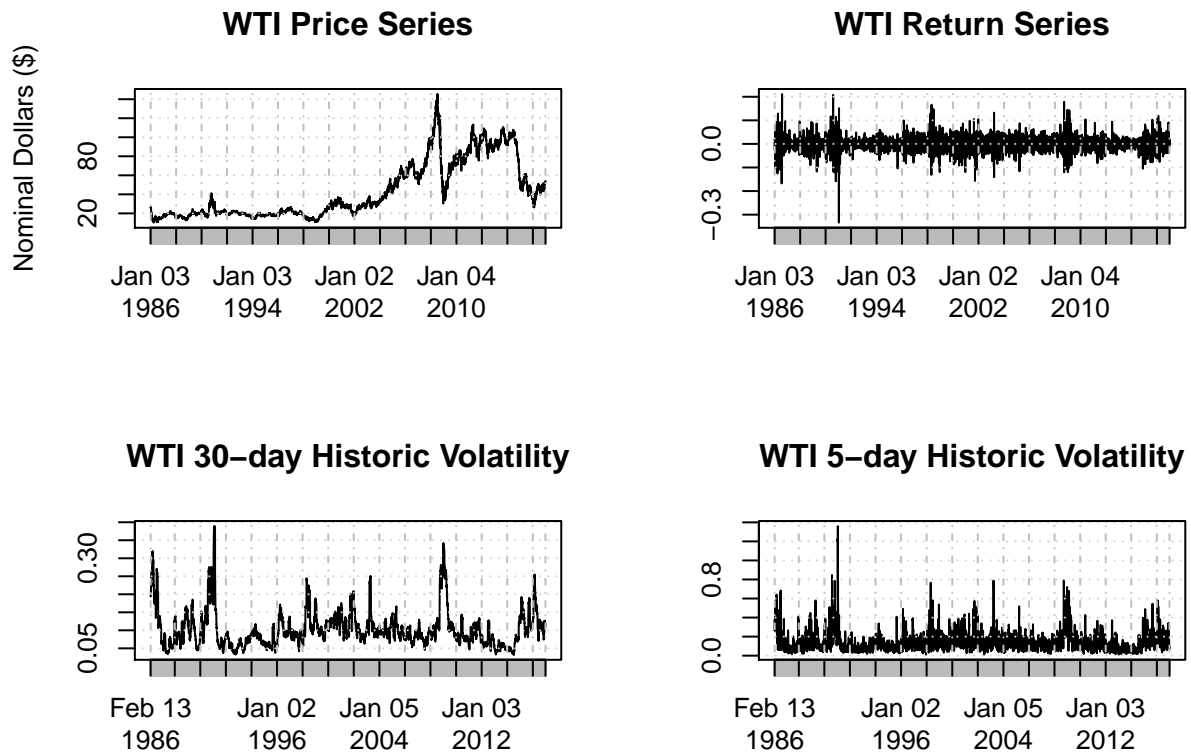
Returns and Volatility

In this research paper, volatility is characterized two ways: (1) 5-day historic volatility and (2) 30-day historic volatility. In addition, the relationship between the returns themselves and price level is investigated. Single-period returns were calculated as:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

and in some sections, their absolute values are used.

```
par(mfrow = c(2, 2))
plot.xts(wti.xts, type = "l", ylab = "Nominal Dollars ($)", main = "WTI Price Series")
plot.xts(wti.return, type = "l", main = "WTI Return Series")
plot.xts(wti.combined[,3], type = "l", main = "WTI 30-day Historic Volatility")
plot.xts(wti.combined[,4], type = "l", main = "WTI 5-day Historic Volatility")
```



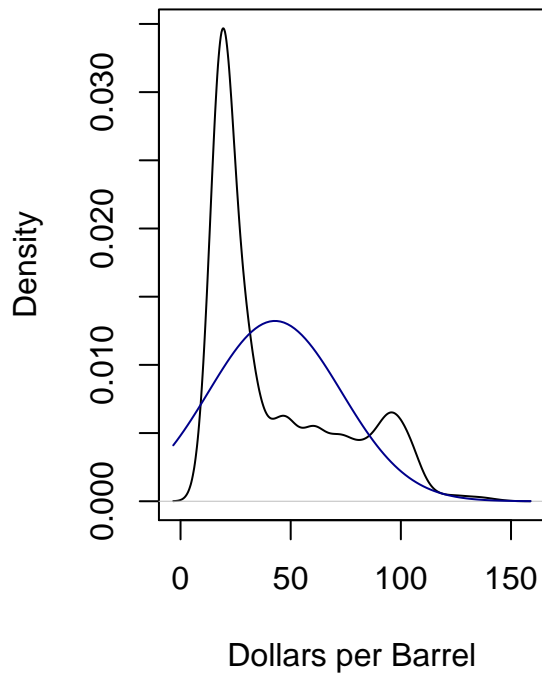
As seen in Figure 1, most of the series from 1986 through 2004 contains prices between \$10/barrel and \$40/barrel. This results in a price series with left skew and a long right tail. These distribution characteristics are common in financial time series. The return series displays the

```
par(mfrow = c(1, 2))

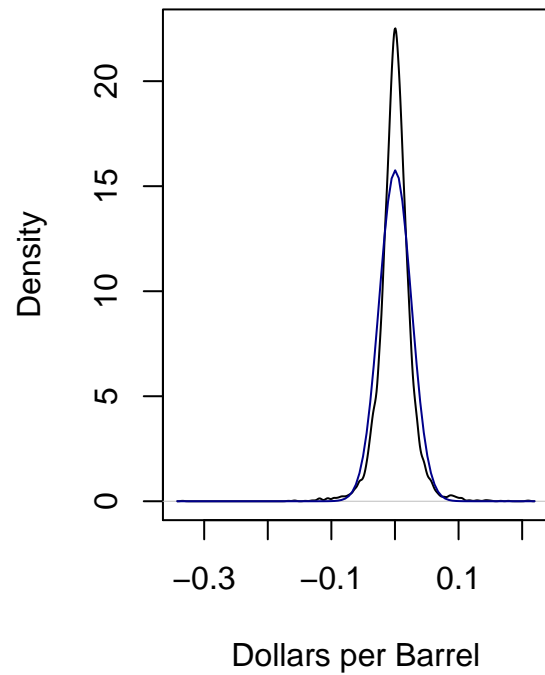
plot(density(wti.xts), main = "Oil Prices Density Plot", xlab = "Dollars per Barrel")
curve(dnorm(x, mean=mean(wti.xts), sd=sd(wti.xts)),
      col="darkblue", lwd=1, add=TRUE, yaxt="n")

plot(density(wti.return), main = "Oil Returns Density Plot", xlab = "Dollars per Barrel")
curve(dnorm(x, mean=mean(wti.return), sd=sd(wti.return)),
      col="darkblue", lwd=1, add=TRUE, yaxt="n")
```

Oil Prices Density Plot

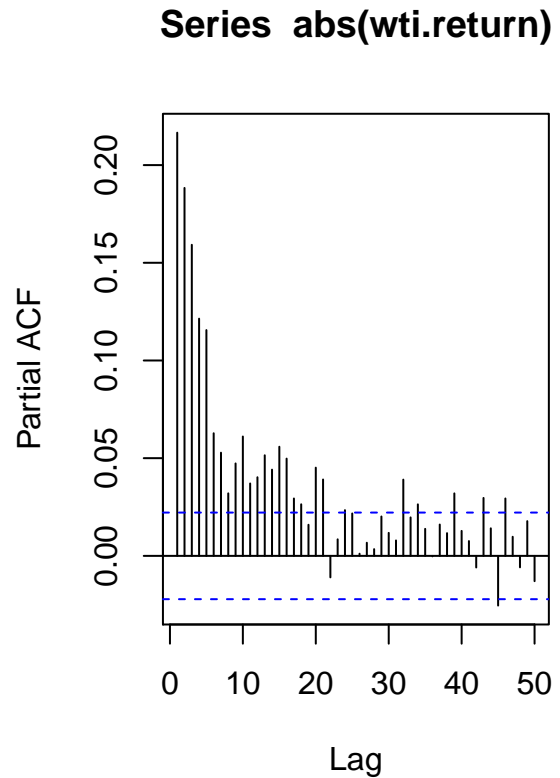
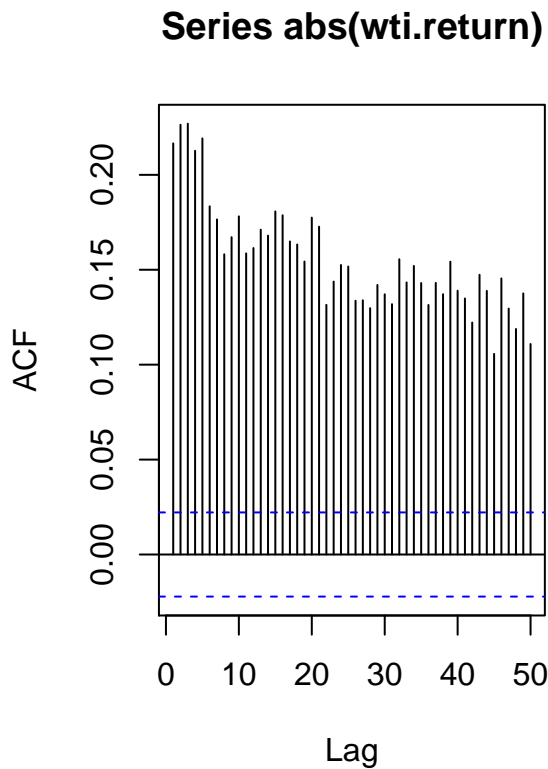


Oil Returns Density Plot



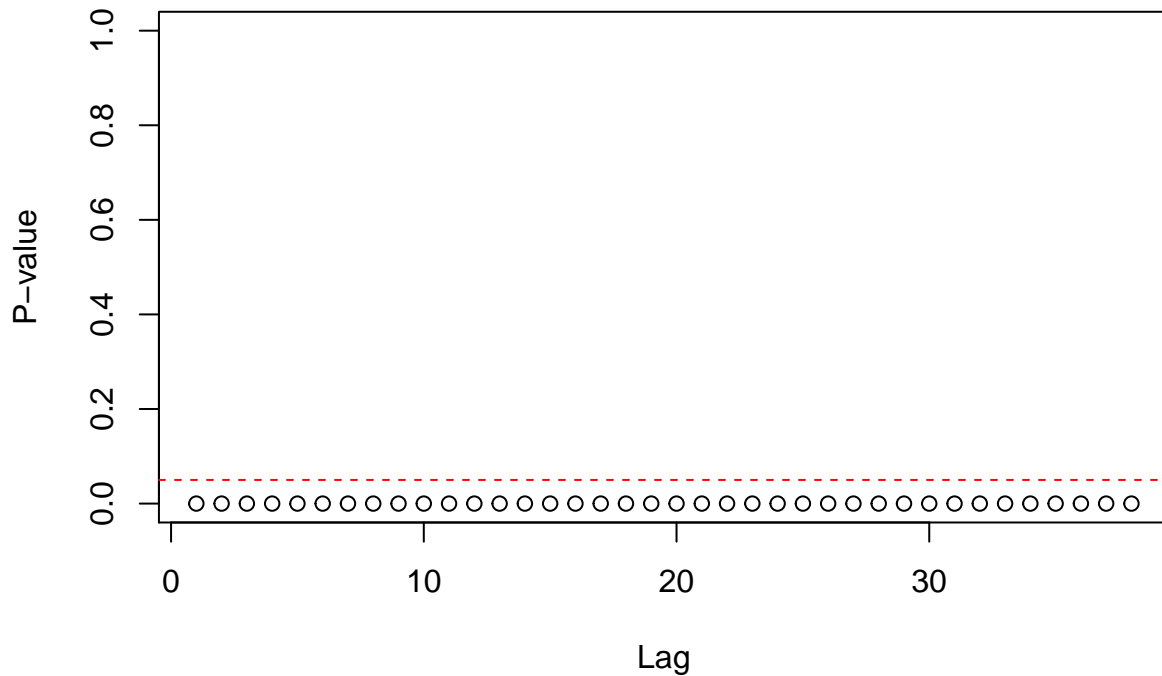
Times series exploration

```
par(mfrow = c(1, 2))  
acf(abs(wti.return), lag.max = 50)  
pacf(abs(wti.return), lag.max = 50)
```



A challenge in analyzing financial time series in general, and spot oil market prices specifically, is that the variance structure may be independent, but not identically distributed. Oil prices exhibit periods of low volatility (i.e. relatively constant prices) and periods of high volatility (i.e. changing prices). This is referred to as volatility clustering. This violates the assumption in the most frequently used time series model, the autoregressive integrated moving average (ARIMA) model.

```
McLeod.Li.test(y = wti.return)
```



Price-Volatility Regression Analysis

```
cov.results <- cov(wti.df)[1,2:4]
cor.results <- cor(wti.df)[1,2:4]

cov.cor <- cbind(cov.results, cor.results)
colnames(cov.cor) <- c("Covariance", "Correlation")
kable(cov.cor)
```

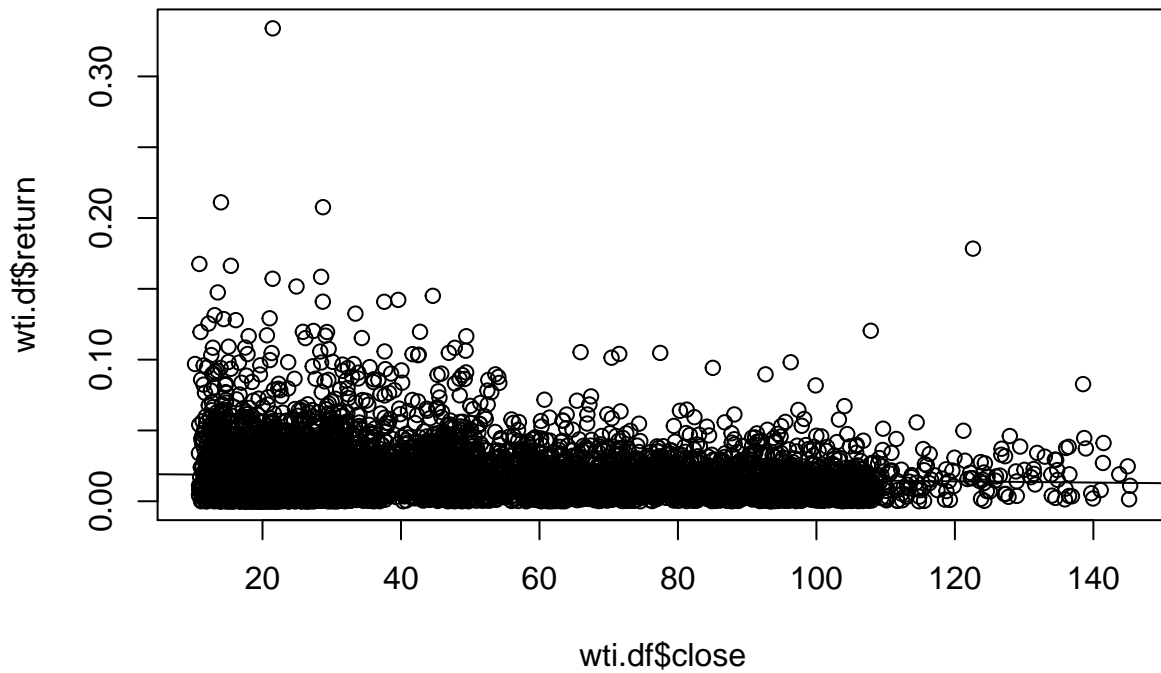
	Covariance	Correlation
return	-0.0396337	-0.0719290
30-day	-0.2933686	-0.1864200
5-day	-0.3757726	-0.1212286

The case of 30-day historic volatility indicates a negative relationship between price level and volatility. However, this result appears to be due to a cluster high volatility around \$20/barrel, creating a leverage point. Residual analysis indicates that this is not a good relationship to model

with linear regression.

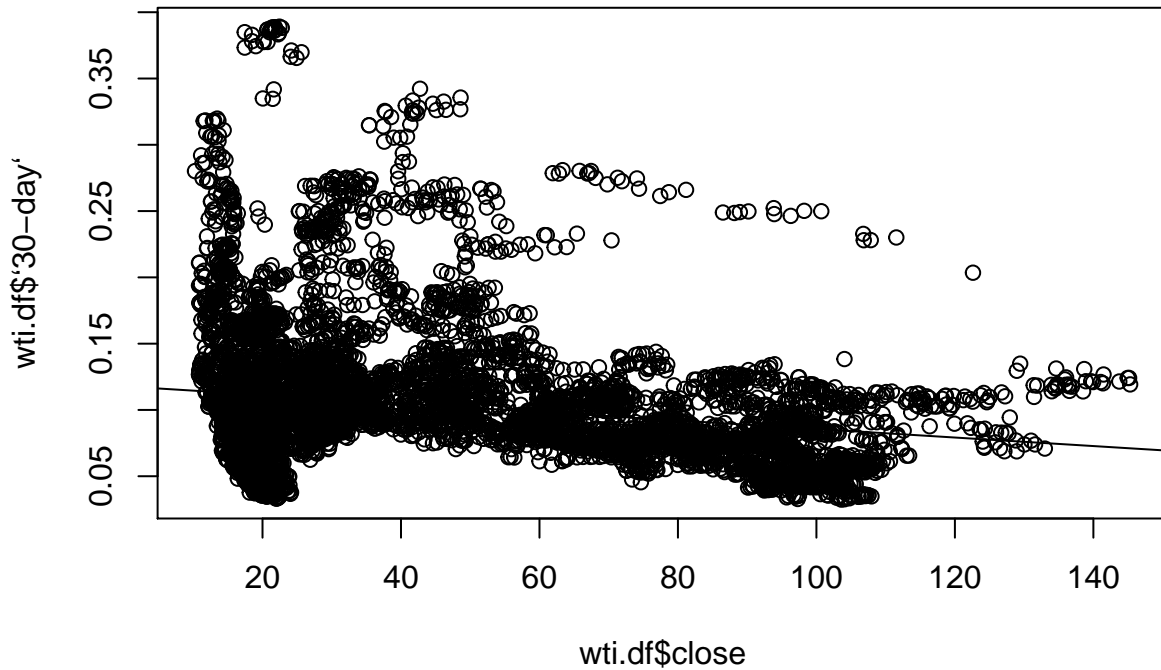
```
returns.lm <- lm(return ~ close, data = wti.combined)
TD.lm <- lm('30-day' ~ close, data = wti.combined)
FD.lm <- lm('5-day' ~ close, data = wti.combined)

plot(x = wti.df$close, y = wti.df$return) + abline(returns.lm)
```



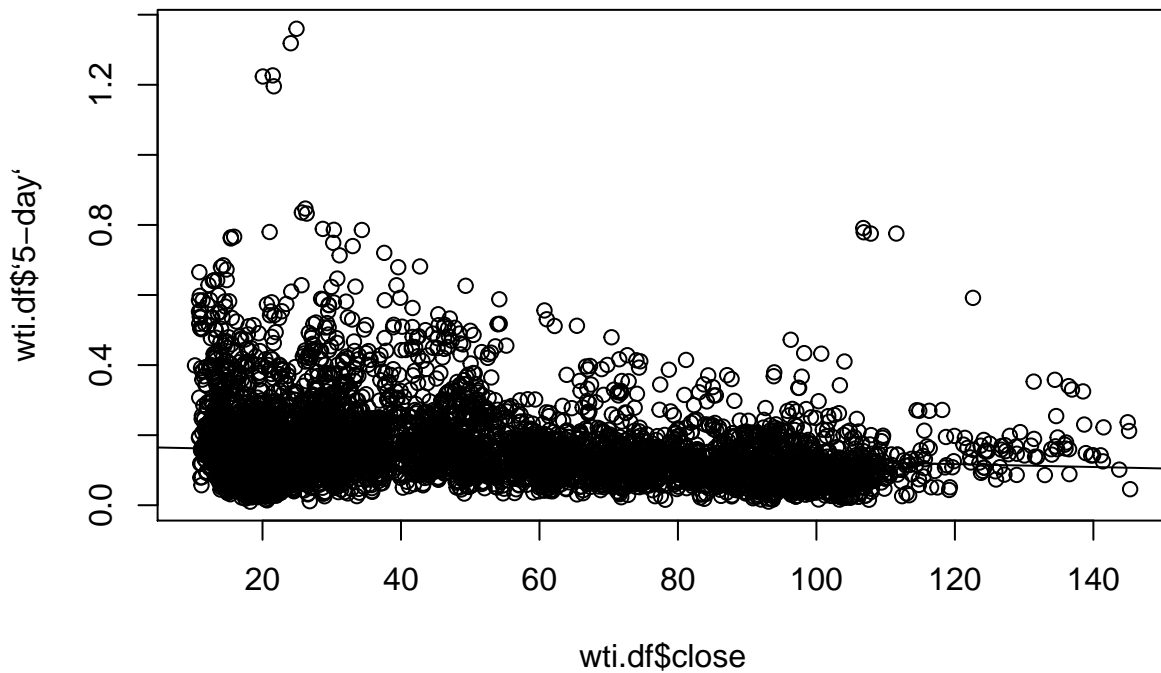
```
## numeric(0)
```

```
plot(x = wti.df$close, y = wti.df$'30-day') + abline(TD.lm)
```

```
## numeric(0)
```

```
plot(x = wti.df$close, y = wti.df$'5-day') + abline(FD.lm)
```



```
## numeric(0)
```

As seen in Figure 1, nominal oil prices have spent time as high as \$145/barrel. However, the majority of the time series is far lower, with a median price of \$28/barrel. This means that

the dataset is unbalanced and higher price levels represent a smaller portion of the dataset. In addition, oil price (and financial time series in general) exhibits volatility clustering. Therefore, it is anticipated that this simple regression model based on price and volatility is not the best possible solution to the question of characterizing the dependency of volatility on price.

Comparing Volatility across Price Regimes

Changepoint detection aims to detect the point or points where the statistical properties of a sequence of observations change (Killick and Eckley, 2014). Changepoint detection was used to detect change in the oil price mean throughout the period of record using the Pruned Exact Linear Time (PELT) algorithm (Killick et al. 2012). The time series between these changepoints represent “price regimes” (i.e. time series between changepoints) which have generally similar mean oil price compared to the entire record.

Changepoint detection proceeds by minimizing a cost function over possible locations and number of changepoints. The cost function:

References