Intraday Patterns in Natural Gas Futures: Extracting Signals from High-Frequency Trading Data

Jung Heon Song, Marcos López de Prado, Horst D. Simon, Kesheng Wu

***Abstract*—High Frequency Trading is pervasive across all electronic financial markets. As algorithms replace an increasing number of tasks previously performed by humans, cascading effects similar to the Flash Crash of May 6th 2010 become more likely. In this work, we demonstrate how signal-processing tools could reveal interesting patterns from high frequency trading using natural gas futures as an example. We focus on Fourier analysis and correlation between weather forecasts and natural gas prices. From the Fourier analysis of Natural Gas futures market, we see strong evidences of High Frequency Trading in the market. The Fourier components corresponding to high frequencies (1) are becoming more prominent in the recent years and (2) are much stronger than could be expected from the structure of the market. Additionally, significant amount of trading activities occur in the first second of every minute, which is a tell-tale sign of algorithmic tradings triggered by clock. To illustrate the potential of cascading events, we study how weather forecasts drive natural gas prices. We show that after separating data according to seasons, the temperature forecast is strongly cointegrated with natural gas price. This splitting of data is necessary because in different seasons the natural gas demand depends on temperature through different mechanisms. We also show that the variations in temperature forecasts contribute to a significant percentage of the average daily price fluctuations, which confirms the expectation that the variations in temperature dominates the volatility of natural gas.**

***Index Terms*—Time series analysis, non-uniform FFT, co-integration**

I. Introduction

Advances in computing power, paired with legislative changes in both the United States and Europe, made high frequency trading (HFT) economically viable (de Prado, 2011; Hasbrouck and Saar, 2013). It has been estimated that HFT firms account for over 60% of the volume in the US equity markets, quickly approaching 50% of the volume in futures markets (Boehmer et al., 2014; Iati, 2009; Commodities Future Trading Commission (CFTC), 2010). However, what followed this surging interest was a growing concern about HFT: rapid algorithmic trading may amplify the volatility inherent in the market, leading to cascading events such as the May 6, 2010 flash crash (Easley et al., 2011; Menkveld, 2013; Smith, 2010; Vuorenmaa, 2013). In this paper, we explore a number of signal processing tools to understand the trading of natural gas (NG) futures and the impact of temperature forecast on the volatility of trading.

NG futures have remained one of the most heavily traded energy contracts for many years. In 2012, the number of futures contracts traded totaled more than 50 million, surpassed only by Light Crude NYMEX and Brent Crude ICG (Wu et al., 2013). Though many factors affect the NG futures price, one of the most important is the expected temperature. In cold weather, natural gas is used to heat up buildings, and in hot weather, generating electricity to power air conditioners. Since the weather forecast is inherently noisy, studying the relationship between temperature forecast and NG futures price could be instructive in understanding the potential of cascading events.

In the US, the most widely available temperature forecast is provided by the National Oceanic and Atmospheric Administration (NOAA), a scientific agency within the United States Department of Commerce. It provides two sets of global weather forecasts, one from its own forecast model and the other as an ensemble average of proprietary and third-party models. Its own weather prediction model is named the Global Forecast System (GFS) (Kanamitsu et al., 1991). GFS produces forecasts of the entire planet 4 times a day with a forecasting

All authors were affiliated with Lawrence Berkeley National Lab when this work was performed. Additional affiliations are noted separately.

University of Minnesota

Guggenheim Partners

horizon of 16 days. The ensemble, on the other hand, is a collection of forecasts from 20 other models that run concurrently with the GFS on the same time scales (Hamill et al., 2013). This ensemble is known as the Global Ensemble Forecast System (GEFS), and was initiated by the National Centers for Environmental Prediction (NCEP) to address the uncertain nature of Numerical Weather Prediction (NWP). An extreme example that necessitates the use of the GEFS is the following: a butterfly flapping her wings can have a cascading effect leading to wind gusts thousands of miles away. Such sensitivity illustrates that a small perturbation can overtime lead to a noticeable disparity between the observed and the predicted data. The GEFS employs Bred Vector and Ensemble Tranform to generate initial perturbations and stochastic total tendency perturbation scheme to account for increase uncertainty from growing instabilities (Wei et al., 2006).

In this paper, we study the GEFS model to present evidence that the forecast temperature and the NG futures price are cointegrated. Because temperatures and natural gas demands are intrinsically related in multiple ways, this cointegration relationship can only be detected after we have split the data according to seasons. Within each season, the relationship between natural gas demands and temperature is well-defined by the expected consumption and easier to detect.

The rest of the paper is organized as follows. Section II discusses related work in cointegration and frequency analysis. Section III contains a more detailed description of the data used in this study. Section IV analyzes NG futures prices through Fourier analysis, and shows the first set of evidences of automated trading in the market. This section also contains an alternative frequency analysis method known as Lomb-Scargle Periodogram. Section V extends the observations seen in the previous section and studies the signals in the volumes of trading activities. Section VI presents results that support GEFS is an error correcting model with respect to NG futures price, and shows the impact of error in temperature forecast. We provide a brief summary in Section VII.

II. Related Work

In the research literature, the volatility of temperature is widely acknowledged as a dominant factor that impact the natural gas price (Bower and Bower, 1985; Elkhafif, 1996; Considine, 2000; Mu, 2007). The standard tools used to analyze volatility are autoregressive conditional heteroskedasticity (ARCH) and generalized autoregressive conditional heteroskedasticity (GARCH) (Engle, 2001). With these tools, researchers have identified a number of interesting features about the volatility of natural gas prices. For example, Pindyck (2004b) and Pindyck (2004a) found a positive time trend in volatility for natural gas, although not of great economic importance, and that shock induced volatility are short-lived. Linn and Zhu (2004) showed higher natural gas price volatility after the release of natural gas storage report. Mu (2007) incorporated these results with temperature shocks to assess their impact on natural gas price dynamics. Furthermore, Mu (2007) demonstrated temperature as the "more direct and purely exogenous measure of demand shocks," and provided the first documented evidence to support the Samuelson hypothesis, i.e., commodity futures volatility declines with contract horizon (Samuelson, 1965).

Despite the importance of temperature forecast on the natural gas futures price, we have yet to come across studies that analyze the possible interplay between forecast models, and their respective roles in natural gas futures price. Another motivation for this work is that the financial market has significantly changed since the publication of the aforementioned studies. Legislative actions that promote high frequency trading strategies have transformed the market in such a way that it is imperative to reevaluate the market behavior by taking into account of the presence of HFT algorithms. As of this writing, a significant amount of work has been done to evaluate the behavior of high-frequency trading on stocks (Brogaard, 2010; Easley et al., 2011; de Prado, 2011; Hasbrouck and Saar, 2013; Vuorenmaa, 2013), but not much has been done on futures trading activities. Therefore, we focus our study on futures trading in this work.

As the standard tools of ARCH and GARCH have been thoroughly explored elsewhere, our work makes use of a few alternative techniques. These tools allow us to study the data and understand the propagation of volatility in different ways than reported thus far. Next, we briefly review the three main techniques we plan to use, Fourier Transform, Lomb-Scargle periodogram, and cointegration between prices and temperature forecasts.

A. Non-Uniform Fast Fourier Transform

It is well-known that the financial markets exhibits complex dynamics in different time scales (Müller et al., 1993; Podobnik et al., 2012; Schoeffel, 2011). One possible way of studying such dynamics is the Fourier transform, which decomposes a function in time into a summation of a number of simple oscillations, each of which can be described by a frequency and its amplitude. These frequencies and their amplitudes are collectively known as the Fourier spectrum of the original function.

Given a function over time $f(t)$, the Fourier transform is

$$g(k) = \frac{1}{2\pi} \int f(t) e^{2\pi i k t} dt.$$

Function $g(k)$ is the Fourier transform of $f(t)$, which is also known as the Fourier coefficient. This is the continuous form of the more commonly used discrete Fourier transform that replaces the integral with a summation (Bloomfield, 2004).

In many applications, the function $f(t)$ is sampled uniformly in a time period, in which case, the summations used to compute the Fourier coefficients share a significant number of common expressions. These common expression can be utilized to drastically reduce the amount of computations needed, leading to a faster way of computing the Fourier coefficients known as the Fast Fourier Transform (FFT) (Welch, 1967; Walker, 1996).

The usefulness of FFT is recognized by many researchers (Bloomfield, 2004; Carr and Madan, 1999; Davidson et al., 1997; Praetz, 1979). However, it has not been widely used to analyze trading activities because trading records are not uniform time series. Trades in real markets happen at unpredictable time points. The most common way to apply FFT on a non-uniform time series is to turn the non-uniform time series into a uniform one. Some common techniques for this transformation include resampling the time series, and binning the original data. Most of the futures contracts are only actively traded during a few hours on each week day. Resampling such a time series with interval less than a day would leave daily gaps, and daily time series have weekly gaps. To study high-frequency features, we would like to have time series with resolution in minutes or seconds. However, these time series would be limited to be with in a trading day. Such a short time series could not be used to study the interactions between the human scale activities in days and weeks with the high-frequency components that are measured in minutes or seconds.

There are alternative techniques specifically proposed to regularize financial time series, which can be considered as redefining "time," for example, autoregressive conditional duration (Engle and Russell, 1998), stochastic conditional duration model (Bauwens and Veredas, 2004), and volume time (Easley and O'Hara, 1987). However, these techniques are not generally accepted by the research community yet, and a significant amount of additional work may be needed before they are understood enough to be accepted.

In this work, we apply a much more direct approach of performing a Fourier transfrom on the irregular time series with a technique named Non-uniform Fast Fourier Transform, (NUFFT) (Dutt and Rokhlin, 1993; Greengard and Lee, 2004).

Given a sample of the function $f(t)$ at some irregular time points t_i, NUFFT computes the Fourier coefficients g at a set of regularly spaced frequencies. For example, if the function f is defined in a time period of one year, the frequencies used to compute the Fourier coefficients g could be k times per year, where k is an integer. The formula for computing a $g(k)$ involves all N input data points of $f(t)$, therefore, a straightforward implementation of this non-uniform Fourier transform for computing N Fourier coefficients would require $O(N^2)$ time. Dutt and Rokhlin (1993) rearranged the computations of these Fourier coefficients into a number of matrix-vector multiplications involving a Toeplitz matrix. Since each multiplication with the Toeplitz matrix takes $O(NlogN)$ time, and the number of such matrix-vector multiplications is independent of the input data values, the overall computation time is $O(NlogN)$. In this notation, the computational complexity of NUFFT is the same as the FFT on regular time series.

Greengard and Lee (2004) developed a strategy to implement the NUFFT algorithm in an efficient software data structure and have made their software implementation available to the public [1]. Our study uses their software implementation of NUFFT.

There have been many different studies of HFT data on different aspects of the data. However, we believe that ours is the first comprehensive study of the Fourier components of HFT data without resampling, binning or otherwise transforming the

[1] The NUFFT software is available at http://www.cims.nyu.edu/cmcl/nufft/nufft.html.

raw trading data. We expect this study to provide valuable time-domain information to broaden our understanding of and offer new insights into high-frequency trading.

B. Lomb-Scargle periodogram

Another common technique for dealing with unevenly sampled data is to interpolate data values to a regular grid. Vaníček (1969), Lomb (1976) and others have devised a least-squares spectral analysis method to accommodate such data structures (Babu and Stoica, 2010). This technique is known as the Lomb-Scargle periodogram. It evaluates data, and associated sinusoidal waves at the times, t_n, that the data are obtained (Scargle, 1982). The Lomb-normalized periodogram for N data is defined as follows (Thomson and Emery, 2014):

$$P(w) = \frac{\left[\sum_{i=1}^{N}(x_n-\overline{x})\cos[w(t_n-\tau)]\right]^2}{2\sigma^2\sum_{i=1}^{N}\cos^2[w(t_n-\tau)]} + \frac{\left[\sum_{i=1}^{N}(x_n-\overline{x})\sin[w(t_n-\tau)]\right]^2}{2\sigma^2\sum_{i=1}^{N}\sin^2[w(t_n-\tau)]}$$

where

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

the sample mean,

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i-\overline{x})^2$$

the sample variance, and the time offset

$$\tau = \frac{1}{2w}\tan^{-1}\left(\frac{\sum_{i=1}^{N}\sin(2wt)}{\sum_{i=1}^{N}\cos(2wt)}\right).$$

One major advantage the Lomb-Scargle periodogram is that it can cope with low frequency noise in records with large time gap. The NG trades occure very frequently, and therefore will not have large time gap in the time series. Given that our initial motivation is to study impact of high-frequency trading, the low frequency noise might not be of great interest. However, it might still worthwhile to compare Lomb-Scargle periodogram with NUFFT[2].

[2]The authors gratefully acknowledge the suggestion from an anonymous referee for this comparison.

C. Cointegration

After extracting the periodic patterns from a time series, we often expect the remaining features in the time series to be explained by some random processes. Next we review a technique that allows us to compare two random processes. The intent is to use this technique for comparing the temperature forecasts with natural gas prices.

Standard statistical methods, least squares, for example, are mostly designed for stationary processes, but financial time series, such as trading prices of natural gas futures, cannot usually be modeled as stationary random processes. When dealing with non-stationary variables, standard regression analysis fails, leading to spurious regressions that hint at the existence of relationships even when there are none. In many cases, non-stationary processes could be turned into stationary ones by taking differences. Given a time series x_t and lag operator L, a difference time series can be expressed as $y_t = (1-L)x_t = x_t - x_{t-1}$. Formally, a time series is integrated of order d if $(1-L)^d X_t$ is integrated of order 0. In other words, $(1-L)^d X_t$ is expected to be generated by a stationary process. Given two or more time series, each of which is integrated, if there exists some linear combination that yields a lower order of integration, the series are said to be cointegrated (Granger, 1981; Engle and Granger, 1987).

Intuitively, cointegrated time series could be regarded as following the same underlying random process. Granger (1981) and colleagues have demonstrated that cointegration is a critical measure in determining the relationship between two non-stationary processes. When analyzed with common statistical methods, integrated time series tend to have strong spurious correlation that cannot be removed by detrending the data. The two non-stationary time series must be cointegrated before we can trust the codependence is not spurious.

There are many different approaches available which analyze the relationship between time series, most notable of which include Support Vector Machine (SVM), ARIMA, ARMA, ARCH and GARCH. However, all of these methods come with their own presumed models of how the variables are related. Before we commit to any of of these models, we believe it to be prudent to determine how the temperature forecast models are cointegrated with the natural gas futures.

According to the definition of cointegration, we can test for this property by first computing a linear combination of the two input time series and then test if the resulting time series is stationary. The linear combination is typically generated through ordinary least squares and the test for stationarity is performed with one of the unit root tests (DeJong et al., 1992; Hassler and Wolters, 1994). These are based on the observation that the characteristic equations of autoregressive time series have the value 1 as their roots (i.e., unit roots) when the time series are non-stationary, whereas the same cannot be said for stationary time series. There are a number of different tests for unit roots (Dickey and Fuller, 1979; Phillips and Perron, 1988). In this work, we choose to use the augmented Dickey-Fuller test, which has been widely analyzed in the research literature (Cheung and Lai, 1995; DeJong et al., 1992).

Earlier we mentioned that there are two distinct mechanisms that the temperature can affect the demand of natural gas. When it is cold, natural gas is used to heat up buildings; therefore, lower temperatures lead to higher demand on natural gas. When it is warm, natural gas is used to generate electricity; consequently, higher temperatures lead to greater demand for natural gas. We anticipate a need to divide time series into a low temperature period and a high temperature period in order to see the interactions between temperature and nature gas futures prices clearly. This is a special feature in our application of cointegration.

III. Data Used for Analyses

The weather forecast data used in our study are retrieved from the NOAA website[3]. A database is formed for every release, and contains as many tables as forecasts, storing them in a grid structure, with latitude as rows and longitude as columns, following the GRIB protocol [4]. Data files are assembled in a way that the planet has been discretized by integer-valued latitudes and longitudes. Each database contains 65 tables, recording temperature values of every 6 hours, up to 384 hours into the future (or 16 days). Each table has columns describing longitude values, from 0 to 359, and rows for latitude, from 0 to 179. Additionally, each database contains an IndexTable from which some metadata associated with the tables, such as the URL it was extracted from, period during which the forecast is effective, time of data release, and of the download. The size of each database is roughly 40 MB, and the cumulative size of the files used in the analysis amounts to 52 GB.

The NOAA forecast models measure temperature at different levels in the air. The two levels that are most relevant to our use are one at 850 millibars (mb) of atmosphere pressure, and the other at 2 meter above the ground. The isosurface with the atmospheric pressure of 850 mb coincides with low-level jet streams, and is frequently studied because the majority of severe weather can be observed at this level. The forecast temperature at 2 meter-above-ground is the temperature reported by the mass media, and is generally considered as the temperature experienced by a human body. However, it fluctuates significantly throughout the day and is very sensitive to local topography, whereas the temperature at 850 mb is much less affected by details on the ground and is more stable throughout the day. Following the common practice in atmospheric science (Wallace and Hobbs, 2006), we consider the temperature at 850 mb as a better representation of the atmospheric condition and use it exclusively in this work.

In order to represent temperature at the national level, we select six out the nine geographic divisions within the United States recognized by the U.S. Census Bureau. For each of these geographic divisions, we pick a reference city for the computation of a national average of temperature, weighted by population. These six regions, and their corresponding reference cities, are Chicago from the East North Central, Minneapolis from the West North Central, New York City from the North East, Atlanta from the South Atlantic, Dallas from the West South Central, and Los Angeles from the Pacific. These cities are major financial centers in the U.S., and representative of each geographic region. This gives us a population-weighted national temperature average, which should therefore be associated with energy consumption.

The natural gas futures trading data used in this study contains all trading records from the beginning of 2007 (2007-01-01 00:00) to middle of 2014 (2014-06-11 18:00). In this study, we will only examine the price, volume and the time of the trading records. The time values are recorded with

[3]http://ready.arl.noaa.gov/READYcmet.php

[4]More information can be found at http://www.cpc.ncep.noaa.gov/products/wesley/wgrib.html

one-millisecond resolution.

The trading of futures contracts involves contracts of different maturity dates. As a contract reaches its maturity date, trading on it ceases while the trading operations continues with a new contract that matures in the future. This periodic change of contracts creates systematic changes in prices that needs to be removed from the cointegration test. We use the rolled prices (after removing the price gap caused by the roll) instead of the actual trading prices. The details of this is given in Section VI. Note that the Fourier analysis is applied directly on the actual trading prices, not on the rolling prices.

IV. Analysis of Prices of Trades

In this section, we seek to provide evidences of high-frequency trading in natural gas futures market. The tools we use for this purpose are two frequency domain analysis techniques known as Non-Uniform Fast Fourier Transform (NUFFT) and Lomb-Scargle Periodogram. The output from the Fourier Transform is known as the Fourier spectrum. To apprise the presence of high-frequency trading, we plan to show that the amplitudes of high-frequency components of the Fourier spectrum are significantly higher than could be expected from the structure of the market, which is discussed next.

A. Structure of Trading Operations

Natural gas futures contracts are traded at Chicago Mercantile Exchange (CME) under the code NG [5]. The electronic markets Globex and ClearPort operate six days a week Sunday – Friday, 6:00PM – 5:15PM (Eastern US Time) with a 45-minute break each day beginning at 5:15PM. Open Cry trading operates Monday – Friday 9:00AM – 2:30PM (Eastern US Time). The trading records do not distinguish which venue has conducted the trades, but the volume of trades is noticeably high when Open Cry trading is in operation.

To see how this weekly operating schedule translates to a Fourier spectrum, we define a function that is 1 when the market is open and 0 when the market is closed. Figure 1 shows the spectrum of this function. The horizontal axis in this figure is the frequency and the vertical axis is the amplitude. Both axes are in log scale.

[5]http://www.cmegroup.com/trading/energy/natural-gas/natural-gas_contract_specifications.html

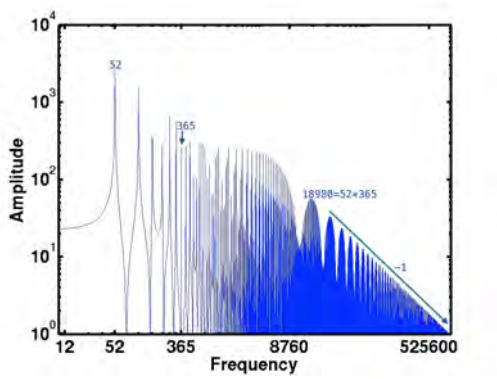


Figure 1: Fourier spectrum of the structure of NG trading activities

The frequency in this and the subsequent spectra is measured in the time unit of year. An oscillation of once per year has the frequency of 1. An oscillation of once per week will have a frequency of 52 because there are 52 weeks in a year. A periodic pattern that occurs once per day has the frequency of 365 in a normal year and 366 in a leap year. Similarly, the frequency of once per hour is 8760 (=365*24) and once per minute is 525600 (=365*24*60). In this work, we only compute the amplitudes of integer frequencies.

As in most Fourier analysis, we are interested in the peaks in the spectrum. A frequency k is a peak as long as its amplitude $g(k)$ is larger than those of the two neighboring frequencies, $g(k) > g(k-1)$ and $g(k) > g(k+1)$. However, we are generally more interested in prominent peaks. In this work, we consider a peak as prominent if its amplitude is larger than the amplitudes of a sizable neighborhood around it. For a frequency k, the size of this neighborhood is taken to be $k/10$. On a log-scale plot of a spectrum, a neighborhood of this size should be clearly visible. The corresponding peak will be visually prominent. Mathematically, $g(k)$ is a prominent peak, if the following is true

$$g(k) > g(j), \qquad \forall j \in [k-\lceil k/10 \rceil, k+\lceil k/10 \rceil], j \neq k.$$

In some analyses, it is more convenient to refer to the power of an oscillation instead of its amplitude. In which case, the spectrum is called a power spectrum. Since the power of an oscillation is effectively the square of the amplitude, the strongest peaks in a Fourier spectrum are also the strongest peaks in a power spectrum. We will not distinguish the two different types of peaks in the future discussions.

We refer to patterns in the spectrum in Figure 1 as the structural patterns of the market in later discussions. We use this spectrum to determine the relative strengths of natural frequencies such as once per day (with frequency of 365) and once per week (with frequency of 52). Here are a few observations about the structural patterns.

1) The frequency with the most prominent amplitude, i.e., the strongest peak, is 52, which is once per week. This incidentally coincides with the cycle of weekly report about stockpiles by the Department of Energy (DOE) (normally on Thursdays), but the simply trading operation's model defined above does not include such weekly activities.
2) The five frequencies with the strongest amplitudes are the following multiples of 52: 52, 104, 260, 312, and 156, which corresponds to cycles of once per week, twice per week, five times per week, six times per week, and three times per week.
3) There is a peak at the frequency of 365 (once per day). However, this peak is lower than the nearby peaks corresponding to events occurring 6 and 8 times a week. This is somewhat surprising since we would have expected the cycle of daily opening and closing of trading operations to appear more prominently in the spectrum.
4) At higher frequencies, there is a series of peaks that are harmonics of the weekly cycle and the daily cycle, i.e., at the frequency that are multiples of 18980 (=365*52).
5) The peaks at multiples of 18980 form a straight line in the log-log plot. This indicates a power law relationship between the amplitude and the frequency, $g \propto k^{\alpha}$, where g denotes the amplitudes of the peaks and k denotes the frequencies of the peaks. The exponent α is -1, which means $g \propto 1/k$. This power law relationship is characteristic of the step function that describes the market opening and closing. Thus, this power law relationship in the Fourier spectrum is exactly what is expected.

B. FFT of Prices

Next, we apply NUFFT on the prices of the NG trades between 2007 and 2013. The NUFFT algorithm is applied to the trading records from each of the seven years separately. For each year, we compute 1,000,000 frequencies[6]. The Fourier coefficients are plotted in Fig 2. As in the previous figure, the graphs are in the log-log scale and frequencies are measured by the number of cycles per year. Since the amplitudes of neighboring frequencies show little continuity, we have opted to plot the amplitude-frequency pairs as scattered points. Furthermore, to limit the number of points shown, we have elected to ignore those frequencies with small amplitudes.

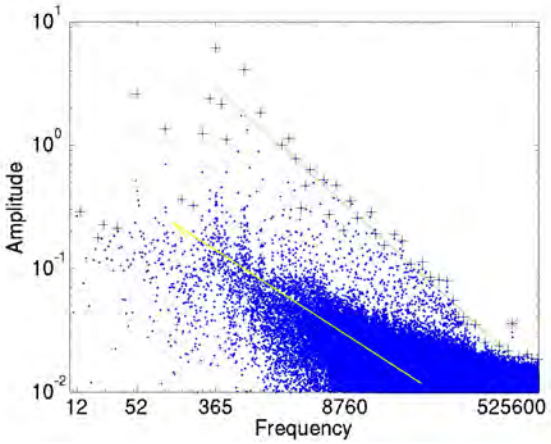
	Top 5 Frequency nodes				
Year	1st	2nd	3rd	4th	5th
2007	365	730	52	312	416
2008	366	732	52	313	418
2009	365	730	52	312	416
2010	365	730	52	312	416
2011	365	730	52	312	416
2012	366	732	52	313	418
2013	365	730	52	312	416

Table I: 5 Strongest frequency nodes in the NUFFT spectra, they correspond to cycles of once per day, twice per day, once per week, six times per week, and eight times per week.

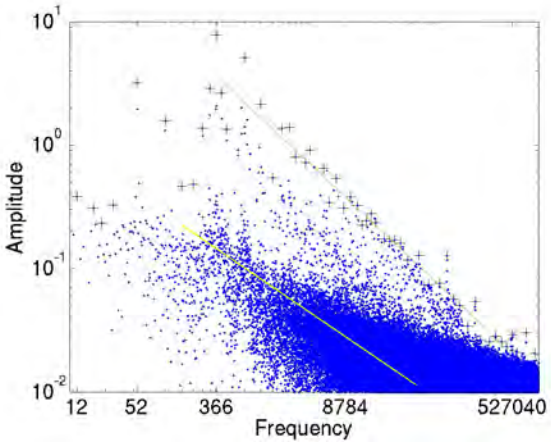
The plots in Fig 2 clearly differ from Figure 1. For example, the most prominent peak in each of the seven years shown in Fig 2 has the frequency of once per day, whereas, in Figure 1, the once per week frequency is the most prominent peak. Though this observation can be easily explained by the daily cycle in the actual trading activities, there is no theoretical guidance to explain whether the daily cycle or the weekly cycle would be the most prominent in the Fourier spectrum.

Table I shows the five most prominent peaks in the spectrum for each year. Notice that these five frequencies are remarkably stable through out the years, indicating that the underlying mechanism that generated these peaks are persistent. Since these frequencies are different from those of Figure 1, we conclude that the factors not captured in our simple model of the market operation must have created the strong daily peaks. In particular, we speculate that the daily rushes to trade immediately after the market open and right before the market close are the root cause of these once per day and twice per day cycles observed through Fourier analysis.

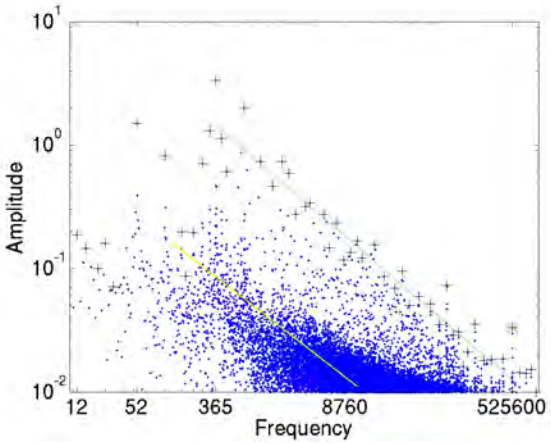
[6]The parameter to NUFFT actually requests 2,000,000 frequencies. Because the prices are real values, the Fourier coefficients are symmetric. The reported Fourier coefficients are those associated with positive frequencies.



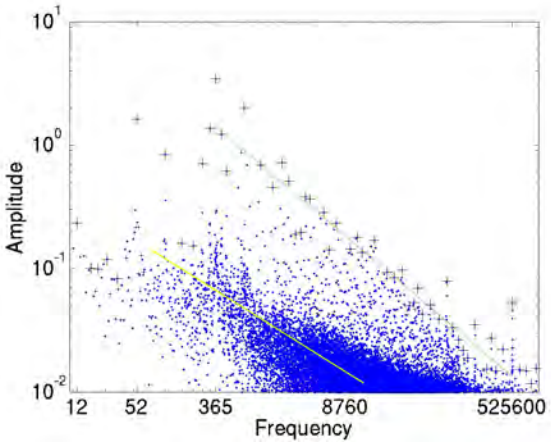
(a) 2007



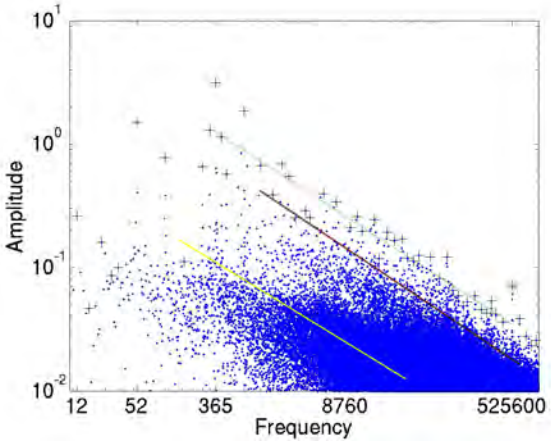
(b) 2008



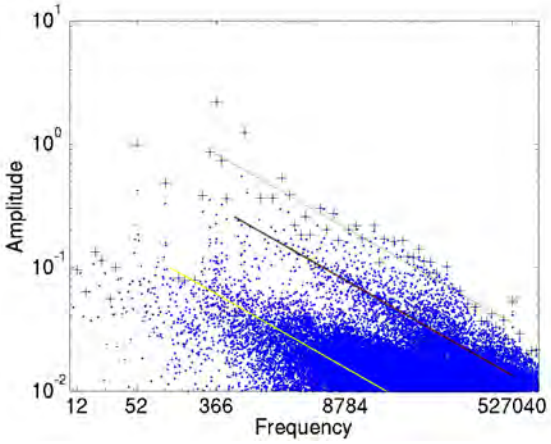
(c) 2009



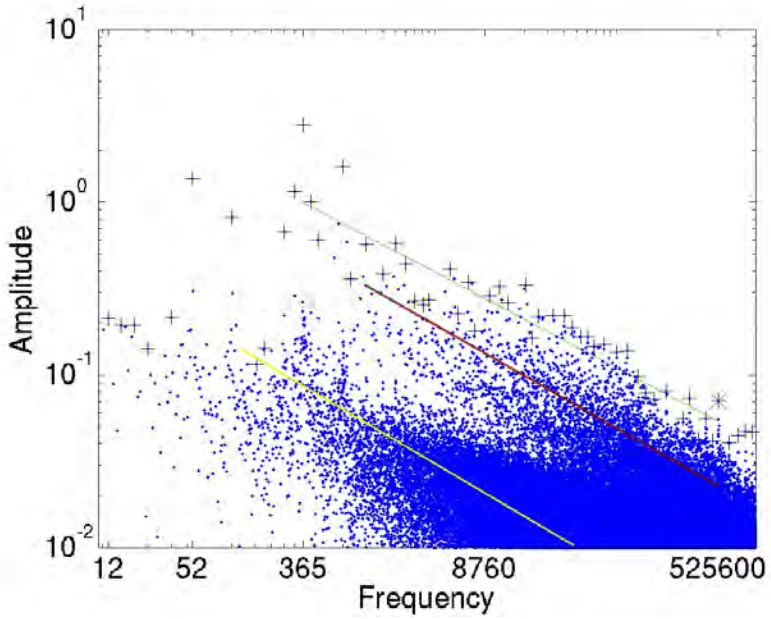
(d) 2010



(e) 2011



(f) 2012



(g) 2013

Figure 2: Log-log of the magnitude of the Fourier coefficients, 2007 - 2013

We regard the activities at frequencies of once per day and once per week as low frequencies for this discussion. Above discussion indicates that NUFFT has captured the expected features of the trading activities at low frequencies. Next, we examine the Fourier components with higher frequencies, such as once per hour (8760 cycles per year) and once per minute (525,600 cycles per year).

In Figure 1 we see a prominent power law relationship that indicates the amplitudes of peaks are inverse proportional to the higher frequencies. We observe a similar power law relationship between the amplitudes of the peaks and the higher frequencies in Fig. 2. To quantify this power law relationship, we have extracted the prominent peaks that dominate relatively large neighborhoods[7]. The resulting prominent peaks are marked with the symbol plus '+' in Fig 2. We then compute the exponent of the power law through a regression on those peaks with frequencies between once per day and once per minute. The resulting exponents are shown in Figure 3. Overall, we see that all seven exponents are significantly larger than -1 in the structural pattern. This demonstrates that the amplitudes of the high-frequency components are significantly stronger than can be expected from the structure of trading operations.

[7]More details on this is given in Song et al. (2014).

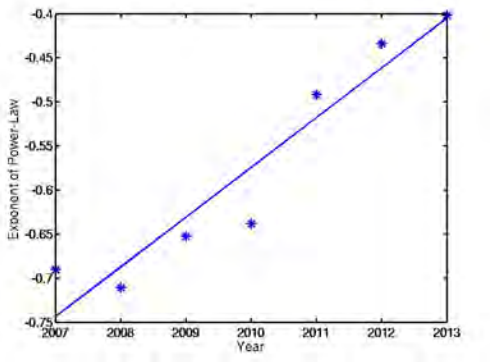


Figure 3: The exponents of the power law relationship detected at different years

Additionally, we see that the seven data points in Figure 3 are quite close to the linear regression line shown, indicating that the exponent of the power law is increasing from year to year. The implication is that the relative strength of the high frequency components in the market have been increasing over the years.

This type of power law distribution has been

Year	Frequency	Rel. Strength
2007	525600	6.7
2008	527040	5.1
2009	525600	13.7
2010	525600	20.3
2011	525600	15.6
2012	527040	15.7
2013	525600	15.4

Table II: The frequencies and relative strengths of the local peaks around once per minute in the Fourier spectrum of trading prices.

	Top 5 Frequency nodes				
Year	1st	2nd	3rd	4th	5th
2007	6	1462	9	2918	12
2008	4	1464	7	2924	213
2009	4	1452	9	21	2908
2010	8	1	1456	4	11
2011	3	1451	6	2899	204
2012	3	1459	2915	205	1251
2013	6	1462	11	2905	215

Table III: 5 Strongest frequency nodes in the Lomb-Scargle Periodogram.

observed by a number of researchers applying different techniques (Müller et al., 1993; Bouchaud et al., 2000; Gabaix et al., 2003; Gabaix, 2008; Filimonov and Sornette, 2014). Different explanations have been proposed in the literature (Müller et al., 1997; Hirshleifer and Teoh, 2003; Hommes, 2006; Lillo and Farmer, 2004). For example, Hardiman et al. (2013) modeled HFT activities as a Hawkes process. In the case of the structural pattern shown in Figure 1, we believe that the power law is caused by the opening and closing of the market. In the actual trading data, we believe that the relative strength at higher frequencies is contributed by increased amount of automated high-frequency trading activities (Hasbrouck and Saar, 2013; Smith, 2010)[8].

In all seven years shown in Fig 2, we see a noticeable peak at the frequency about once per minute (525600 in most years and 527040 in leap years). In these plots, the relative strengths of the peaks are not obvious because there are too many data points nearby. To quantify the strength of this peak, we compare its amplitude against the average amplitude of 100,000 nearby frequencies. This ratio is shown as the relative strength of the peak around once per minute frequency in Table II. In the recent five years, these peak amplitudes are at least ten times stronger than the average of 100,000 nearby frequencies. Given these once-per-minute peaks appear in all seven years, there must be persistent activities in the market with the frequency of once per minute.

In addition to the relative strengths, this table also shows the actual frequency of the peak in each year. We observe that frequencies of the peaks are exactly once per minute. This precision suggests that the origin of these peaks may be systematic in nature. A likely source of this precise action would be automated trading triggered by clocks. In the next section, we examine the trading volumes and find another tell-tale sign of precise periodic actions in the market.

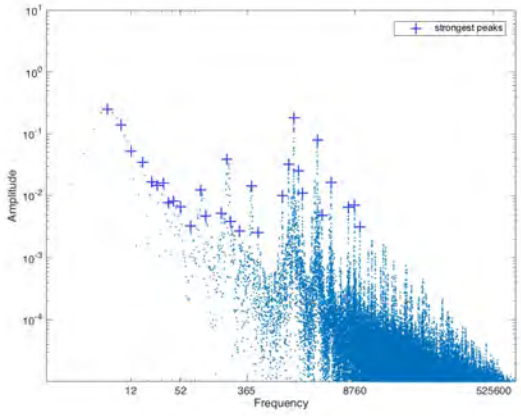
It is possible for NUFFT to produce spurious features in a spectrum. However, all the peaks we have reported so far are persistent ones that appears in all years. It is unlikely that some random spurious feature would appear as consistent patterns for 7 years. Furthermore, these prominent peaks in the Fourier spectra corespond to weekly and daily cycles of known market behaviors and therefore not likely from the spurious spectra.

Another popular technique for studying spectra of unevenly sampled time series is Lomb-Scargle Periodogram (Thomson and Emery, 2014). This method improves upon the Fourier analysis in handling low-frequency noise introduced by long gapes in the data records. Using the Lomb-Scargle Periodogram routine found in MATLAB, and proceeding in the same manner as with NUFFT, we obtain the following log-log scale of amplitude against frequency in Figure 4.

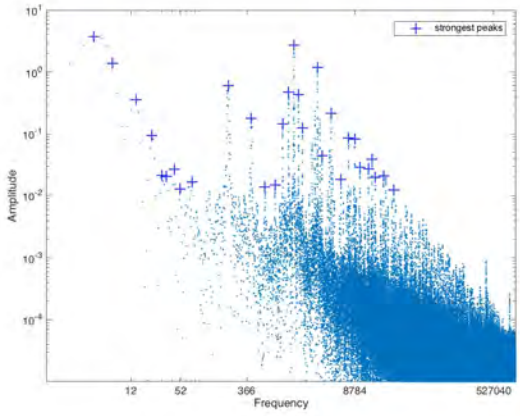
Again, we have elected to ignore frequencies with very small amplitudes, and plotted the amplitude-frequency pairs as scattered points. Table III lists 5 strongest frequency nodes in power spectra.

Figure 4 and Figure 2 have some superficial similarities. For example, both of them have the tallest peaks on the left side of the plots and the peaks with higher frequencies fall off following power law. These agreements indicate that both of them are capturing common features present in the trading records. However, there are also significant disagreements between the results of NUFFT and those of Lomb-Scargle Periodogram. For example, the frequencies of the strongest peaks in Lomb-Scargle Periodograms have a median value of 1462, see Table III, which is very close to cycles of

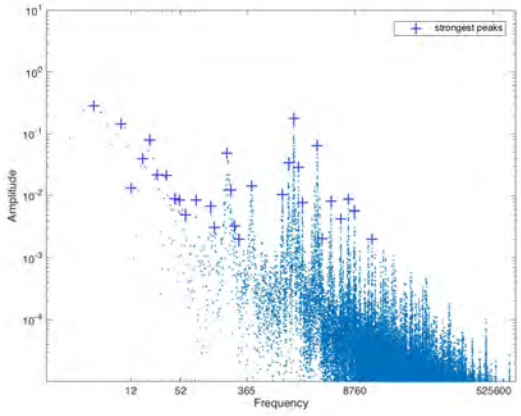
[8]Even though our observation of increasing HFT activities agrees with many others, we are also aware of contrary indicators reported by some researchers (Iati, 2009; Baron et al., 2014)



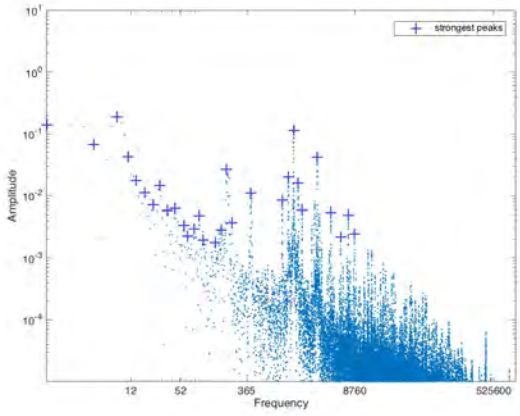
(a) 2007



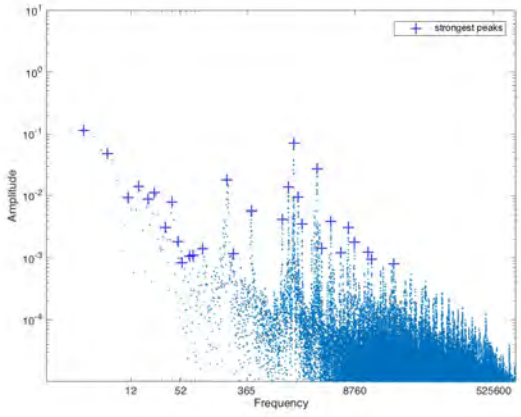
(b) 2008



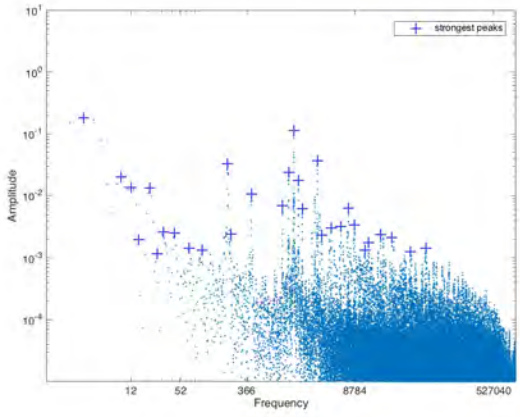
(c) 2009



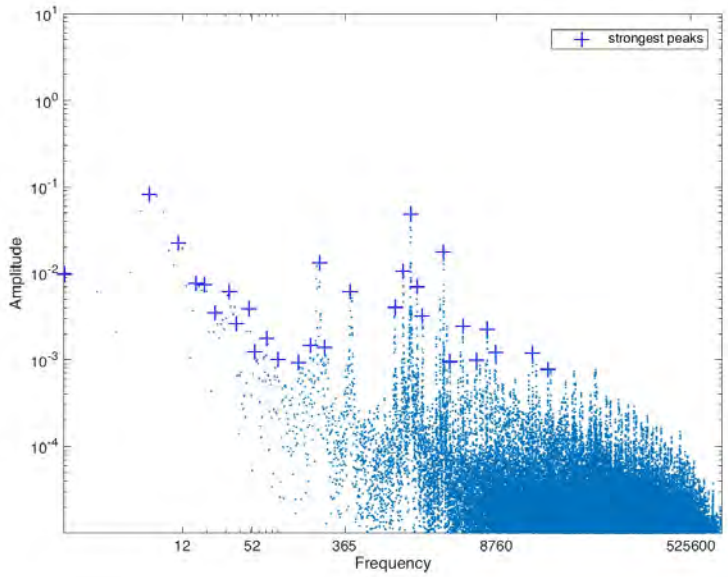
(d) 2010



(e) 2011



(f) 2012



(g) 2013

Figure 4: Log-log of the magnitude of the Lomb-Scargle Power Spectral Density, 2007 - 2013

four times per day ($4 \times 365 = 1460$), while the strongest peaks from NUFFT has a frequency of 365 (cycling once per day). It appears to the authors that the once per day cycles could be associated with daily operations of the market and daily trading volume variations, it is much hard to come up with mechanisms that have a cycle of four times per day. This suggests that NUFFT is much more likely to have captured the essence of the trading operations.

The Lomb-Scargle Periodogram is designed to cope with data gaps, however, there is no significant gap in the trading records of natural gas futures, therefore, there is no chance for it to shine. The Lomb-Scargle Periodogram is known to capture the main peaks (Jenkins and Jarvis, 1999), which explains to some extend why the high-frequency peak at once per minute observed in NUFFT results is sometimes not visible in Lomb-Scargle Periodograms. In Table I, we see the frequencies of the main peaks are consistent over the seven years of the trading data, however, the frequencies of the top peaks can vary quite a lot in Table III. Since NUFFT produces consistent peaks with frequencies matching known market features, we regard NUFFT as more effective in capturing signals from the trading data.

V. Analysis of Volumes Traded

In the previous section, we identified a persistent periodic trading pattern at the frequency of once per minute. Next, we examine the trading activities in an "average" minute. Our initial attempt to construct this "average" minute was to build a histogram of trading activities in every second of a minute for all trades in a given year. However, this produced no clear signal. Next, we describe an approach that is able to show tell-tale signs of automated algorithmic trading activities.

A. *Fractions of Tradings in a Second*

In an earlier publication, Easley et al. (2012) used a sample of E-mini S&P500 futures trades between 11/07/2010 and 11/07/2011 to show that a large fraction of trade in a minute occurs in the first second. This is a clear sign that a significant number of trading activities are triggered by clock and programmed to be executed at the beginning of a minute. Well-known examples of such time-based trading strategies include Time-Weighted Average Price (TWAP) and Volume-Weighted Average Price (VWAP), both of which recompute the average price and adjust the trading operations on a regular time interval. To see whether a similar automated trading

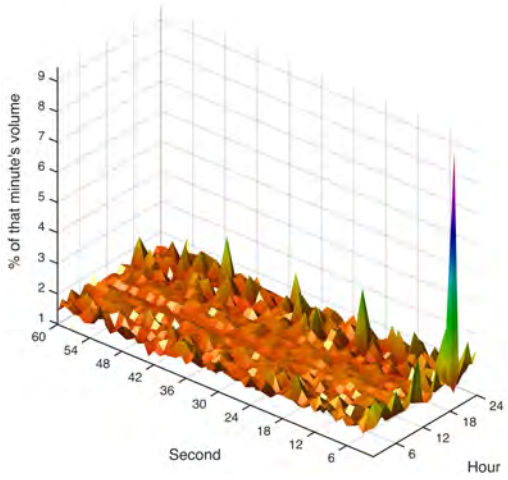


Figure 5: Fraction of trading volumes in a minute occuring at specific second within the minute from all trades from 2013.

is present in natural gas futures market, we compute the same fractions of trades in each minute and show the results in Figure 5.

If the trades happen at the uniform rate, then each second of a minute would handle $\frac{1}{60}$th of the trade volumes, about 1.67%. However, in Fig. 5, we see a number of distinct spikes. Clearly, there is something interesting in this figure, but we would like to eliminate influence from the structure of the market operations before we present a full discussion of the signals that could be identified with the volumes of trades.

B. Trades at Market Opening

From Fig. 5, we see that the tallest peak is at the first second of the 18th hour. The same is true for data from other years that are not shown in this paper. The time of this peak corresponds to the daily opening of the electronic market for NG futures at Chicago Mercantile Exchange (CME). In the 45 minutes prior to this 6PM market opening, electronic orders accumulate in a queue and most of them would be executed during the first few seconds, right after market opens. If orders arrive at the same rate during the market recess as during the normal business hours, then the number of orders to be processed during the first second after the market opens will be as much as 43% of all orders

Year	$s = 0$	$s \leq 4$	$s \leq 5$	$s \leq 6$
2007	0	3	5	6
2008	1	2	2	3
2009	7	9	9	10
2010	11	11	11	11
2011	11	11	12	12
2012	9	12	12	12
2013	10	15	17	17

Table IV: Location (second within a minute) of the largest volume fraction of trading in a minute of filtered data (segregated by hour of the day and averaged over the whole year).

processed during the 6PM hour[9]. We realize that the number of trades is different from the number of orders, and consequently, the fraction of trades happening during the first second of 6PM hour could be significantly different from 43%. However, we would still expect the number of trades processed at the opening of the market to be considerably higher than the average number of trades processed. Fig. 6 presents the number of trades processed each minute of a day.

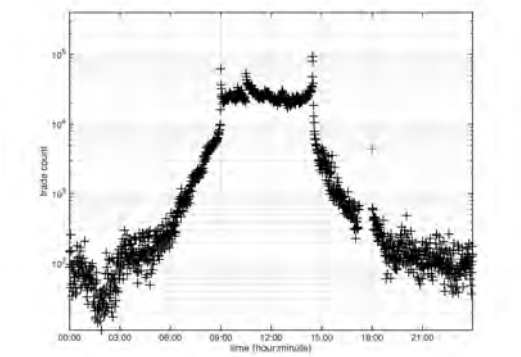
Given a sample of trades over a year, we aggregate those that have the same hour and minute, and plot the trade counts. A red plus sign in each of Figure 6 denotes trades that happened immediately after the market opening. We observe that the number of NG futures traded right after market opening is much higher than the subsequent time. As a large fraction of these trades will be handled during the first second after the market opens, we seek to eliminate this structural feature from the analysis. To do this, we have chosen to discard these trades during the first minute after market opening.

C. Presence of Algorithmic Tradings

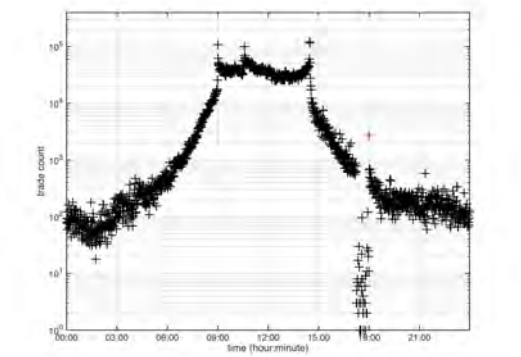
After removing the trades during the first minute of 6PM hour, we reproduce the fractions of trades per second shown earlier in Figure 7.

From Figure 7, we determine the second that has the highest point for each of the 24 hours. Table IV shows that the number of times the highest point occur in the first few seconds of a minute. From this

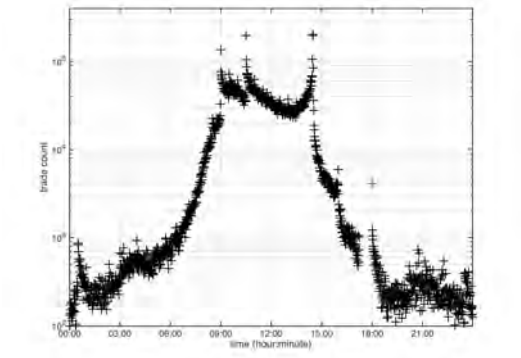
[9]Denote the arrival rate of orders to be 60α per hour, the number of orders during the recess (5:15PM-6:00PM) will be 45α and the number of orders arriving during the first second of all minutes during the 6:00PM hour is α, and the total number of orders to be processed during the 6PM hour will be $45\alpha + 60\alpha = 106\alpha$. Thus the fraction of orders needs to be processed during the first second of the 6PM hour is $(45\alpha + \alpha)/106\alpha = 46/106 \sim 43\%$.



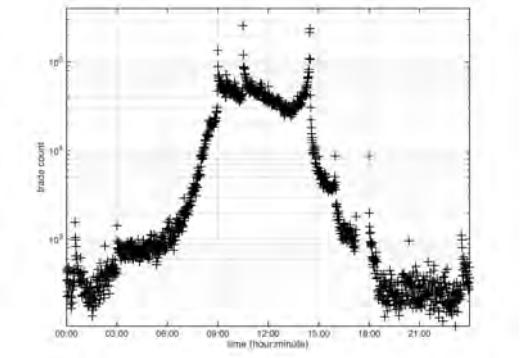
(a) Number (in log scale) of trades occuring at specific minute of a day in 2007



(b) Number (in log scale) of trades occuring at specific minute of a day in 2008



(c) Number (in log scale) of trades occuring at specific minute of a day in 2009



(d) Number (in log scale) of trades occuring at specific minute of a day in 2010

table, we see that a significant number of maximums occur in the first few seconds. For example, in 2013, 15 out of 24 maximums occurred within the first 5 seconds, whereas it is 3 in 2007. Moreover, of that 15 times, 10 of them occurred in the first second. This trend suggests that compared to 2007 and 2008, the automated trading based on clock triggers has become more active in the recent years.

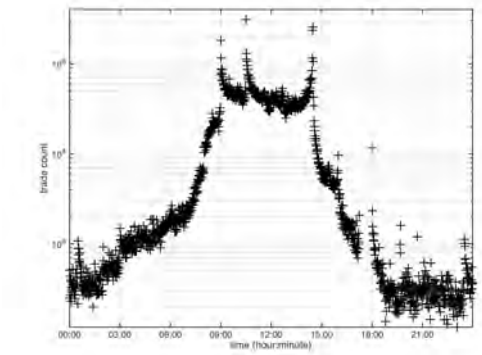
VI. Impact of Temperature

In the previous sections, we have established NG futures markets have undergone major changes in their trading activities in the past few years. We presume part of the evolution of the trading algorithms is to better utilize temperature forecasts to anticipate the demand of natural gas and other energy commodities. This will allow the prices of natural gas futures to respond more quickly to the new information, which is good in general. However, this may also increase the likelihood that automated trading algorithms would overract to the forecast errors and causing unexpected volatility and liquidity pressure.

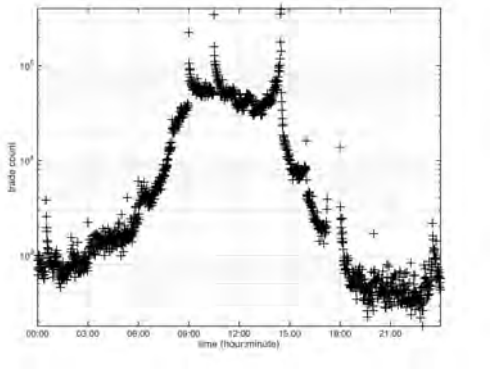
In our study, we opt to use the temperature forecast produced by the ensemble model GEFS for its stability and accuracy. Since GEFS is released every 6 hours, we divide the trading data into 6-hour intervals and set the closing price as the price of each time interval. Since we have only downloaded weather forecasts from 2013-09-17 00:00 $\sim$ 2014-06-11 12:00, we will limit the NG futures trading information to this time window for this part of the study.

A. Correlation between Temperature and Price

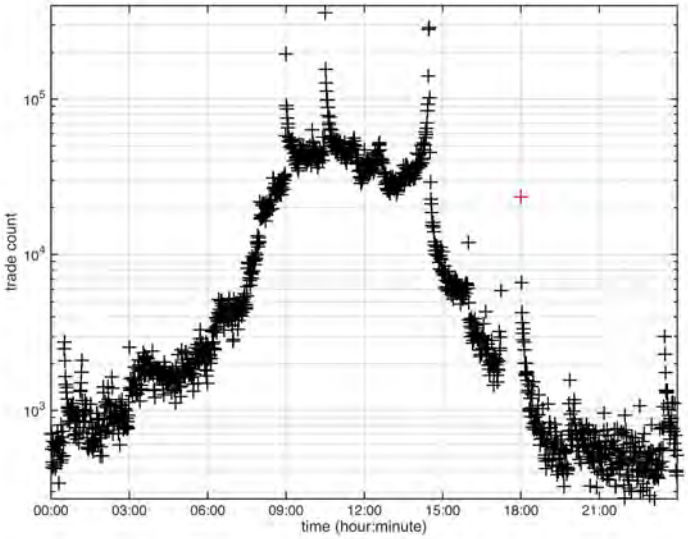
We start our exploration with the simple scatter plot of the prices versus temperatures in Figure 8. In Table V, we report the correlation coefficients between NG futures prices and the mean temperature values. Recall that a correlation coefficient



(e) Number (in log scale) of trades occuring at specific minute of a day in 2011



(f) Number (in log scale) of trades occuring at specific minute of a day in 2012



(g) Number (in log scale) of trades occuring at specific minute of a day in 2013

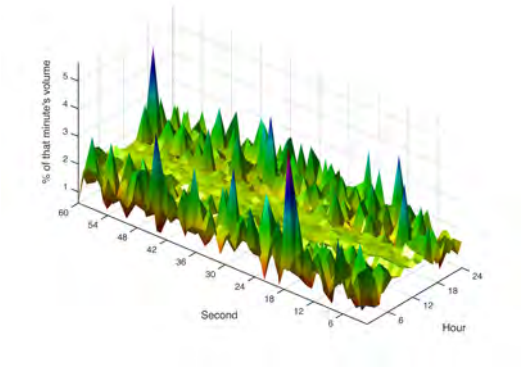
Figure 6: Concentration of NG trade for every minute. The red cross marks trades at the first minute after market opens at 18:00.

	Correlation coefficient
National Average	**-0.4374**
Chicago	-0.3711
Minneapolis	-0.3468
New York City	-0.3853
Dallas	-0.4411
Atlanta	-0.4575
Los Angeles	-0.4281

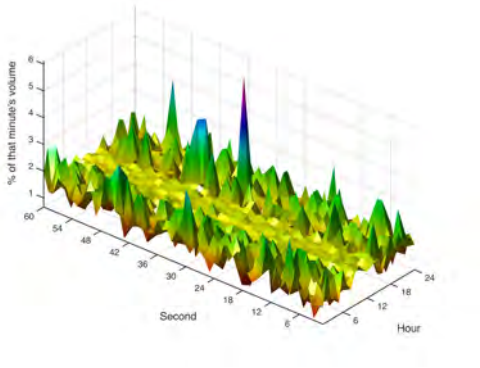
Table V: Correlation coefficients between NG futures prices and mean forecast temperature (2013-09-17 00:00 $\sim$ 2014-06-11 12:00).

near 1 indicates that two variables satisfy a strong linear relationship. Given the values in Table V are between -0.3 and -0.4, the evidence for linear relationship between price and temperature appears to be weak.

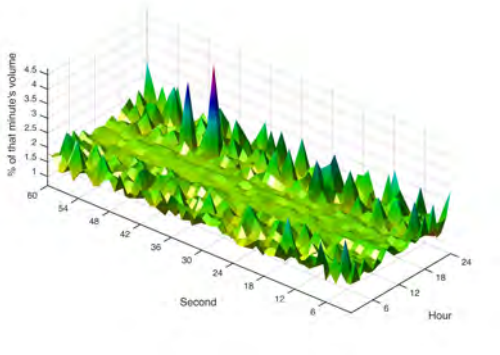
Note that the temperature reported in Figure 8 is in Kelvin, where 260K is about -13 degrees Celsius and 9 degrees Fahrenheit, and 290K is 17 degrees Celsius and 62 degrees Fahrenheit. Overall, there is a trend of higher prices when the temperature is low, because there is higher demand on natural



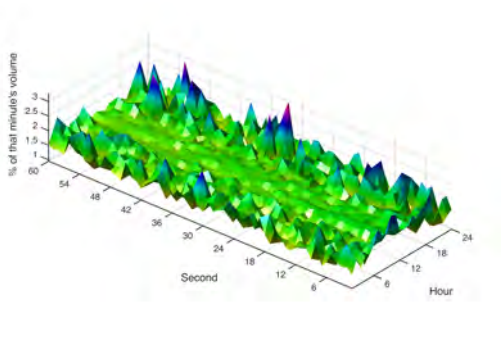
(a) Fraction of trading volumes of 2007 in a minute occuring at specific second with the first minute of 18:00 removed



(b) Fraction of trading volumes of 2008 in a minute occuring at specific second with the first minute of 18:00 removed



(c) Fraction of trading volumes of 2009 in a minute occuring at specific second with the first minute of 18:00 removed



(d) Fraction of trading volumes of 2010 in a minute occuring at specific second with the first minute of 18:00 removed

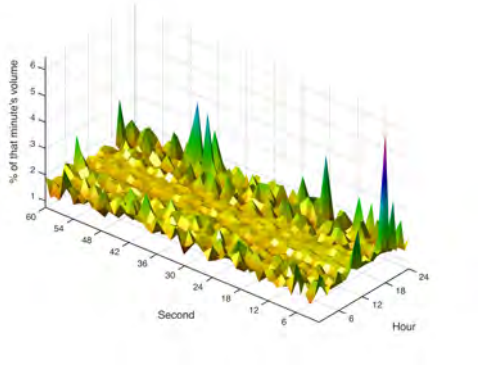
gas for heating when the outside temperature is lower. However, as discussed before, we expect higher temperature could lead to higher natural gas prices as well, therefore we believe that a more sophisticated tool is necessary to better understand the relationship between the temperature and the NG futures price. In particular, it is necessary to divide the time series into different temperature zones to allow the different influence mechanism to be studied separately.

B. Cointegration

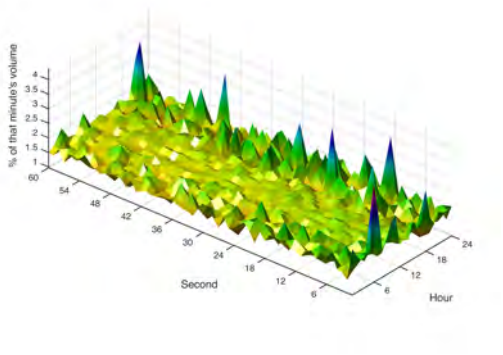
Because the NG futures periodically roll to a new contract with a longer maturity date, a time series of the actual trading prices would have periodic jumps due to these contract changes. In the Fourier analysis conducted in the previous sections, since these periodic jumps in prices were happending every month, we did not notice any obvious impact on the higher frequency operations. However, for this study, we found it necessary to remove these price jumps. A common approach of removing the price gaps caused by these contract changes is to use the *rolling price* instead of the actual prices. The rolling prices are computed as follows.

1. For each trade entry at time i with price p_i, compute the change in price, $\delta_i = p_i - p_{i-1}$.
2. Whenever there is a change in contract, compute δ_i from the second 6-hour time window and leave out the value from the first time window after the contract roll.
3. Take cumulative sum of δ_i to produce a new time series for price.

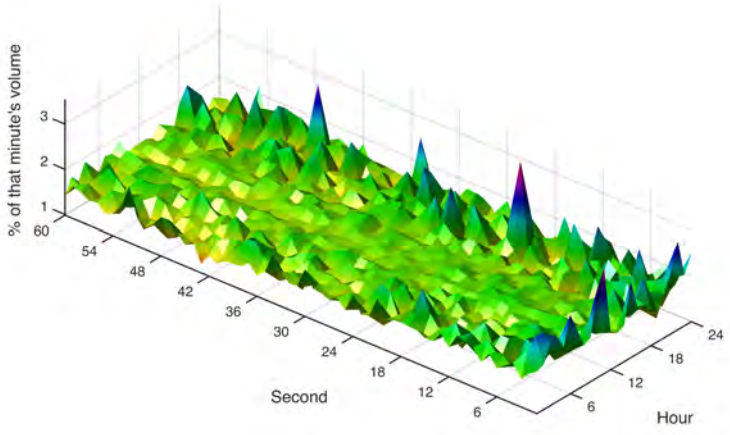
To test for cointegration, we regress the rolling price on GEFS data, and run an augmented Dickey-Fuller (ADF) test on the residuals for a unit root



(e) Fraction of trading volumes of 2011 in a minute occuring at specific second with the first minute of 18:00 removed



(f) Fraction of trading volumes of 2012 in a minute occuring at specific second with the first minute of 18:00 removed



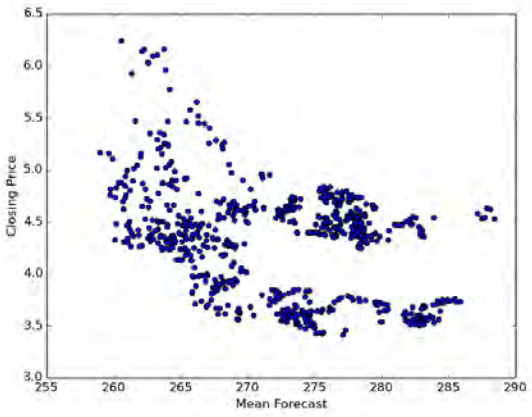
(g) Fraction of trading volumes of 2013 in a minute occuring at specific second with the first minute of 18:00 removed

Figure 7: Concentration of filtered NG trades for every hour at each second, irrespective of the minute

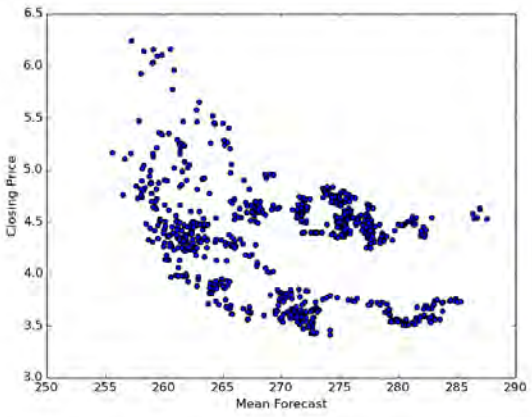
(Dickey and Fuller, 1979). If the p-value associated with this test is less than 0.05, we can conclude that the rolling price and the GEFS forecast are cointegrated with confidence level of 95%. The implication here is that some linear combination of the two data sets is stationary, which describes a particular kind of long-run equilibrium relationship, or an error correcting model. By the Engle-Granger representation theorem, we can then say that the forecasted temperature is an error correction model (ECM) of the rolling price (Engle and Granger, 1987).

City	p-value
New York	0.50278
Minneapolis	0.21831
Atlanta	0.92450
LA	0.52416
Chicago	0.17413
Dallas	0.91643
Population-weighted national average	0.61249

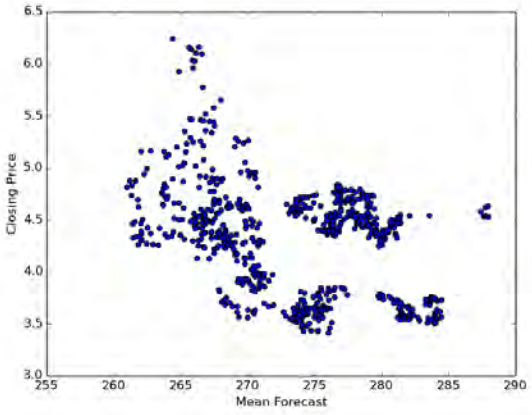
Table VI: p-value for all cities, from 2013-09-17 00:00 to 2014-06-11 12:00



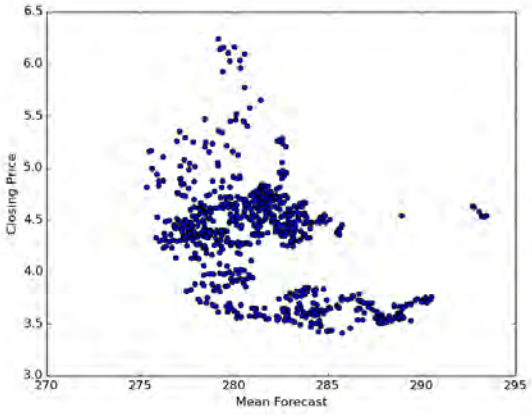
(a) Chicago



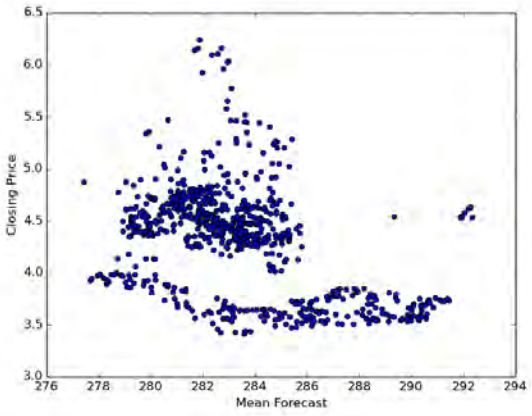
(b) Minneapolis



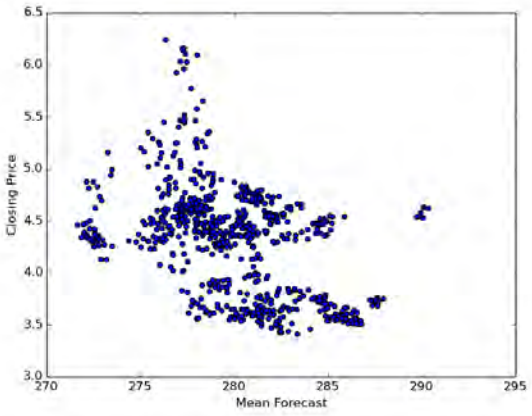
(c) New York City



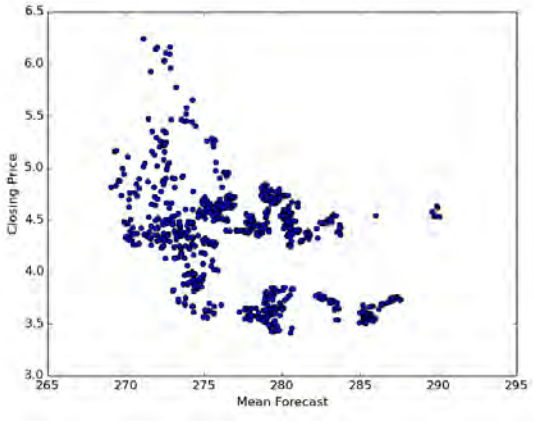
(d) Dallas



(e) Los Angeles



(f) Atlanta



(g) Population weighted national average

Figure 8: Scatter plots for cities using mean forecast temperature (2013-09-17 00:00 $\sim$ 2014-06-11 12:00).

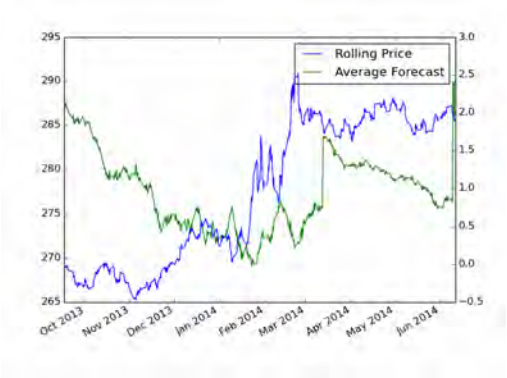


Figure 9: Rolling price and average of forecasted temperature (in Kelvin)

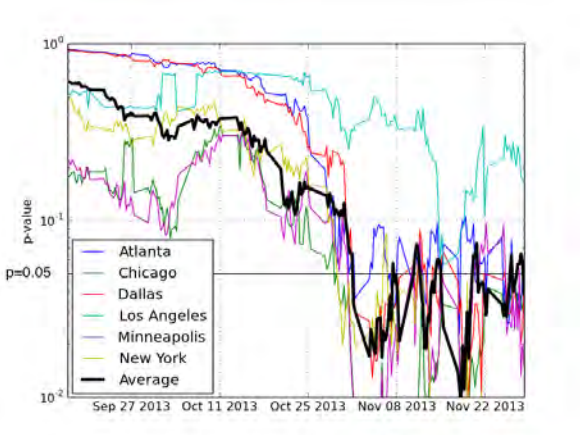


Figure 10: p-values of each case, with starting date as a point on the x-axis to 2014-06-11 12:00

Our two time series are shown in Figure 9. Table VI shows that the p-values are all much larger than 0.05, which indicates that the cointegration test fails if all data points are used. As discussed earlier, there are two different ways temperature could affect the demand, and therefore, the price of natural gas. Next, we explore how to split the time series so that the interaction between the forecast temperature and the price might be more easily captured by cointegration.

Since we do not know for sure if temperature forecast cointegrates with price, we first explore if these two variables display such a relationship in any part of the time window. There are encouraging signs that the two cointegrate at least in some time intervals. For example, if we exclude data from 2013-09-17 00:00 to 2013-10-31 18:00, the p-value of the national average drops to 0.06, and the p-value decreases as more dates are excluded. Furthermore, if we work with data starting from 2013-11-01 00:00, the resulting p-value is 0.065, and for starting date of 2013-11-01 06:00 or later, the p-value becomes less than 0.05. Figure 10 shows how p-value changes for all 7 cases if the starting dates are changed.

Three observations appear to be relevant.

1) Except for Los Angeles, p-values of the others only start to increase after 2013-11-01.
2) Since the average temperature of Los Angeles is much higher than those of other cities,

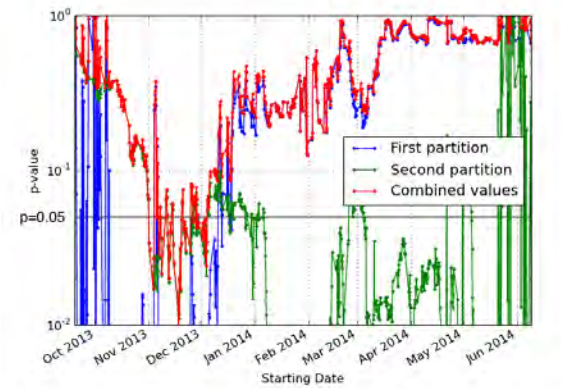


Figure 11: p-values after dividing the data into 2 parts (in log scale).

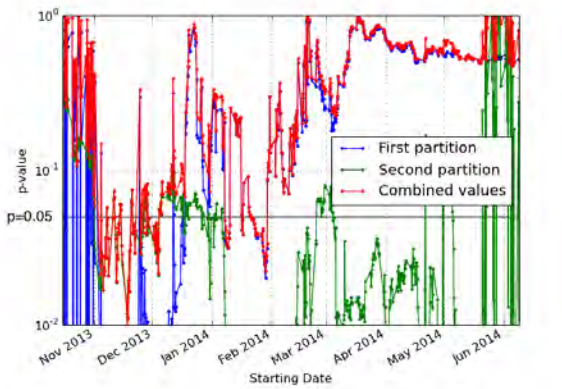


Figure 12: 2 p-values for data after 2013-10-11 06:00 (in log scale).

Splitting Date	1st p-value (p_1)	2nd p-value (p_2)
2013-11-18 06:00	0.00683	0.00896

Table VII: Optimal splitting date, and associated p-values for each segment.

natural gas is less likely to be used for heating in LA. It makes sense that relative to others, LA's p-value does not vary too much.

3) The national average's p-value becomes less than 0.05 only after October, which roughly coincides with the start of a new season. This suggests forecast data should be divided in order to take into account of seasonality.

Now we know that at least for some period of time, the price cointegrates with the forecasted temperature. Next, we explore the data to see if the time window could be split into two where the two variables cointegrate in each time window separately. We plot the p-values of these two time windows against the splitting point in Figure 11, where the *First partition* denotes p-value of data from 2013-09-17 00:00 to a point on the x-axis, and the *Second partition* for the remaining part.

From Figure 11, we could choose a date at which both p-values are less than 0.05. Since there are a number of dates satisfying this requirement, we choose a date that minimizes $p_1 + p_2$. The result of this minimization is shown in Table VII.

Clearly, we could divide the time series into two time windows and in each time window the temperature forecast cointegrates with the price. However, the two partitions do not resemble common notions of seasons. In particular, the range of the second division extends multiple seasons. Next, we explore whether spliting the time seres into three time windows could better match appropriate seasons. To do this, we first choose a split point that minimizes p_1. From Figure 11, $\min p_1$ is found in 2013-10-11 06:00. We set this to be the first split date, and proceed to find the second split point in the remaining portion of the time series. The p-values for different split dates are shown in Figure 12.

Based on Figure 12, the optimal split dates that minimize the sum of the p-values of the next two partitions is November 18 2013 as shown in Table VIII.

The three divions from Table VIII have similar problem as the two division from Table VII, i.e., the last division covers a long period of time and spans multiple seasons. Next, we choose the end point of the first division to be 2013-11-18 06:00, and then proceed to find the second split point that minimizes $p_2 + p_3$, see Figure 13.

The two split points and the p-values from the three time windows are listed in Table IX.

In this case, the last time window is very small and the largest time window still spans multiple seasons. Next, we move the first split point to be 2013-12-04 06:00 and proceed to find the second split point that minimizes $p_2 + p_3$. We record the p-values in Figure 14 and Table X.

After a number of tries, we finally arrive at two split points that reasonably match the convential notion of seasons, and within each time window, the augmented Dickey-Fuller tests give us p-values that are much less than 0.05, which indicates that

1st Splitting Date	2nd Splitting date	1st p-value, p_1	2nd p-value , p_2	3rd p-value, p_3
2013-10-11 06:00	2013-11-18 06:00	5.668355e-22	3.944314e-5	8.9599e-3

Table VIII: Optimal splitting dates for three partitions, and associated p-values for each segment.

1st Splitting Date	2nd Splitting date	1st p-value, p_1	2nd p-value , p_2	3rd p-value, p_3
2013-11-18 06:00	2014-06-04 18:00	0.00683	0.002590	0.000009

Table IX: Second optimal splitting date, and associated p-values for each segment

1st Splitting Date	2nd Splitting date	1st p-value, p_1	2nd p-value , p_2	3rd p-value, p_3
2013-12-04 06:00	2014-04-01 06:00	0.003096	0.002950	0.001817

Table X: Optimal splitting date, and associated p-values for each segment

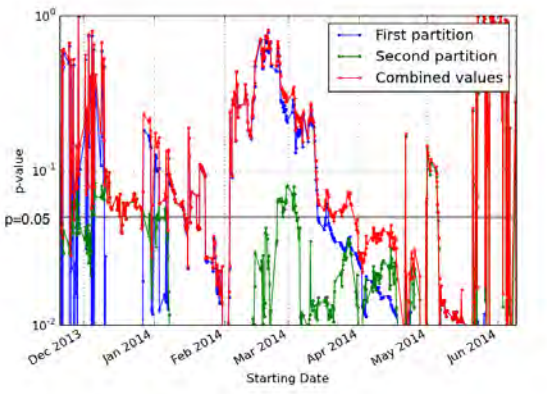


Figure 13: 2 p-values for data after 2013-11-18 06:00 (in log scale).

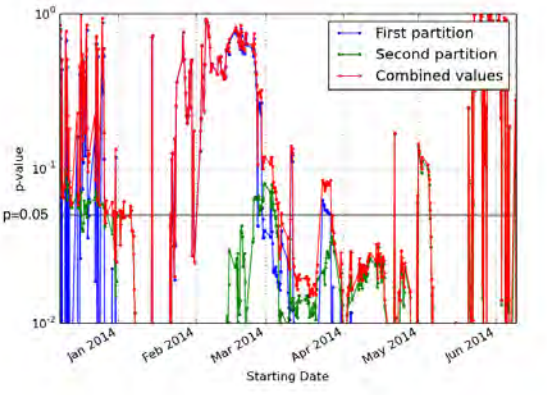


Figure 14: 2 p-values for data after 2013-12-04 06:00 (in log scale).

GEFS forecast temperature cointegrates with the NG futures price within each time window. This division of the time series into seasons is necessary because in different seasons the temperature has different impact on the demand and therefore price of natural gas. Since the temperature cointegrates with the price, by the Engle-Granger representation theorem, the former is an error correction model of the price.

C. Impact of Forecast Errors

We now discuss the accuracy of GEFS temperature forecasts. For each forecast horizon, we evaluate the standard deviation of the difference between the predicted and the observed values. As the forecasting model has time horizon of 16 days (or 384 hours), carrying out this process for each 6 hour interval yields 64 sets of standard deviation values.

We have shown from above that partitioning the data is necessary in order to take into account seasonal patterns in weather. We do the same here, and note how the standard deviations vary for different timeframe and horizon shown in Figure 15.

In the 2nd partition (corresponding to the winter session), we observe significant rise in standard deviations, because the winter storms could change the temperatures quickly. Forecasts for the next 96 hours, and onwards, are shown to carry much volatility. Predictions made for the summer are relatively stable, in that the majority of errors lie within a 1 degree interval. For GEFS during the autumn, which closely matches with the time frame of the 1st partition, we observe a moderate increase in the standard deviations, its contour closely resembling that of the winter season. Nevertheless, the maximum error of 3 degrees is only attained for the time horizon of 384 hours. Figure 15 provides concrete evidence that even the forecasting model

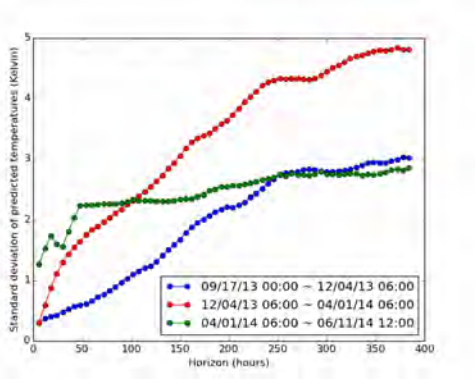


Figure 15: Standard Deviation of GEFS Forecasting Errors

known to be most reliable could have significant errors, which might contribute much volatility to the market.

To further our cointegration analysis and assess the market's adaptibility when an error in temperature is introduced. We carry out the following steps for each partition.

1) For each 6-hour time interval, we are given 64 sets of forecast temperatures. Together with rolling prices, construct an error correction model, using the constant deterministic regressor (Pedroni, 1999).
2) Based on the model created, identify the equilibrium price and temperature, and denote them by (*Equil. Price*, *Equil. Temp*).
3) Introduce the error in temperature to *Equil. Price*. This corresponds to the standard deviation of forecast temperature at each horizon, and are plotted in Figure 15.
4) Label the new temperature by *New Temp*. Using the ECM model created above, predict the next 1000 sets of values from (*Equili. Price*, *New Temp*).
5) Impose the tolerance $\epsilon = 0.01$. If the price difference of two consecutive data is less than ϵ, declare this to be the new equilibrium price and temperature.
6) Express the difference between the new and the old equilibrium price in terms of the standard deviation of daily returns.

In Tables XI $\sim$ XIII, 'Equil. Price' denotes the original equilibrium price based on the ECM created, 'Displacement' the new equilibrium price after noise is introduced to the weather forecast, and 'Convergence' the number of steps needed to attain error less than 0.01. The analysis sheds light on the role played by the forecasting model. In the spring (Table XIII), the mean of the percent difference in terms of standard deviation of daily returns is 17%. This implies that the forecast variability contributes to 17% of the standard deviation of daily returns. The mean in the fall (Table XI) is 1.23%, and 4.97% in the winter (Table XII). We note that although the percentages for the fall (Table XI) and the winter (Table XII) are small in comparison to that of the spring (Table XIII), differences between the new and old equilibrium prices are much greater.

We remark that the cointegration and ECM analysis further highlight the role GEFS actively plays in formulating errors in standard deviation of daily returns, and can benefit regulators who need to assess the risks posed by a greater presence of automated traders, and microstructural researchers who can take the discovered patterns into account when developing their own models.

VII. Summary

In this study, we demonstrate a number of signal processing tools could be used to extract useful patterns from the trading records of natural gas futures. We considered three sets of techniques, Fourier analysis, Lomb-Scargle Periodogram, and co-integration. Fourier analysis on trade data revealed the the high-frequency components are increasing over the past few years, which agree with the general rise of high-frequency trading. Fourier analysis also detects once per minute signal with increasing in magnitude in the recent data. The precision of the frequency suggests that there are automated trading based on clock. In fact, we find that many more trades happen during the first second of a minute, which again is an indication of automated trading triggered by clock. This was especially true for 2010, 2011, and 2013, and gives a convincing proof that TWAP and VWAP are likely responsible.

We use the co-integration technique to study the relationship between temperature forecast and natural gas futures price. The results indicate that the temperature forecast and natural gas futures price are co-integrated when the trading data are separated into seasons. This is necessary because the mechanism relating the temperate and price are different in different seasons. Furthermore, we analyze the role GEFS plays in determining natural gas futures price, and presented evidence that it

Horizon	Equil. Price	Displacement	Conv. steps	SD introduced	% Diff in SD of daily retu.
24	-0.1802	-0.1826	3	0.4265	0.36
48	-0.1697	-0.1866	4	0.5964	2.56
72	-0.1494	-0.1633	4	0.7697	2.10
96	-0.1463	-0.1570	2	1.0353	1.62
120	-0.1573	-0.1741	3	1.2439	2.55
144	-0.1590	-0.1810	2	1.5956	3.34
168	-0.1414	-0.1575	2	1.9551	2.44
192	-0.1231	-0.1243	2	2.1750	0.19
216	-0.1166	-0.1195	2	2.2910	0.43
240	-0.1242	-0.1259	2	2.5972	0.26
264	-0.1361	-0.1449	2	2.7846	1.34
288	-0.1420	-0.1496	2	2.8205	1.15
312	-0.1576	-0.1601	2	2.8060	0.37
336	-0.1624	-0.1661	2	2.8913	0.55
360	-0.1734	-0.1721	2	2.9393	0.19
384	-0.1766	-0.1748	2	3.0227	0.28

Table XI: ECM Analysis for 1st (fall) partition (2013-09-17 00:00 to 2013-12-04 06:00).

Horizon	Equil. Price	Displacement	Conv. steps	SD introduced	% Diff in SD of daily retu.
24	1.5461	1.5366	1	1.1110	3.49
48	1.5424	1.5213	1	1.6516	7.72
72	1.5693	1.5501	1	1.9690	7.02
96	1.6309	1.6195	1	2.2571	4.16
120	1.6273	1.6141	1	2.5415	4.83
144	1.5755	1.5605	1	2.9440	5.47
168	1.5265	1.5106	1	3.3490	5.84
192	1.5018	1.4835	1	3.5896	6.69
216	1.4673	1.4457	2	3.9473	7.88
240	1.4903	1.4744	2	4.2765	5.81
264	1.5697	1.5600	2	4.3357	3.54
288	1.6574	1.6530	2	4.3406	1.60
312	1.6959	1.7027	2	4.5546	2.48
336	1.7184	1.7281	2	4.7198	3.53
360	1.7442	1.7555	2	4.8012	4.12
384	1.7586	1.7731	2	4.8180	5.33

Table XII: ECM Analysis for 2nd (winter) partition (2013-12-04 06:00 to 2014-04-01 06:00)

Horizon	Equil. Price	Displacement	Conv. steps	SD introduced	% Diff in SD of daily retu.
24	1.9312	1.9269	2	1.6056	11.33
48	1.9287	1.9251	2	2.2411	9.56
72	1.9131	1.9109	2	2.2714	5.89
96	1.9070	1.9064	2	2.3053	1.60
120	1.9274	1.9258	2	2.3227	4.02
144	1.9646	1.9590	2	2.3203	14.51
168	1.9844	1.9770	1	2.3889	19.51
192	1.9843	1.9725	1	2.5427	31.01
216	1.9697	1.9615	1	2.5797	21.39
240	1.9515	1.9428	1	2.6741	22.69
264	1.9344	1.9269	2	2.7650	19.68
288	1.9295	1.9328	2	2.7545	14.92
312	1.9284	1.9210	2	2.7426	19.09
336	1.9320	1.9222	1	2.7269	25.74
360	1.9383	1.9269	1	2.7815	29.67
384	1.9322	1.9233	2	2.8564	23.15

Table XIII: ECM Analysis for 3rd (spring) partition (2014-04-01 06:00 to 2014-06-11 12:00)

is an error correction model. We estimated that a significant percentage of the daily volatility in NG prices is caused by weather forecasting errors. Even for the most accurate model, GEFS, temperature forecasting errors may be responsible for 20% of the daily volatility in this market.

REFERENCES

Babu, P. and Stoica, P. (2010). Spectral analysis of nonuniformly sampled data – a review. *Digital Signal Processing*, 20(2):359–378.

Baron, M., Brogaard, J., and Kirilenko, A. A. (2014). Risk and return in high frequency trading. http://ssrn.com/abstract=2433118.

Bauwens, L. and Veredas, D. (2004). The stochastic conditional duration model: a latent variable model for the analysis of financial durations. *Journal of econometrics*, 119(2):381–412.

Bloomfield, P. (2004). *Fourier analysis of time series: an introduction*. John Wiley & Sons.

Boehmer, E., Fong, K. Y. L., and Wu, J. J. (2014). International evidence on algorithmic trading. http://ssrn.com/abstract=2022034.

Bouchaud, J.-P., Potters, M., and Meyer, M. (2000). Apparent multifractality in financial time series. *The European Physical Journal B-Condensed Matter and Complex Systems*, 13(3):595–599.

Bower, R. and Bower, N. (1985). Weather normalization and natural gas regulation. *Energy Journal*, 6(2):101–115.

Brogaard, J. (2010). High frequency trading and its impact on market quality. http://heartland.org/sites/default/files/htf.pdf.

Carr, P. and Madan, D. (1999). Option valuation using the fast Fourier transform. *Journal of computational finance*, 2(4):61–73.

Cheung, Y.-W. and Lai, K. S. (1995). Lag order and critical values of the augmented dickey-fuller test. *Journal of Business & Economic Statistics*, 13(3):277–280.

Commodities Future Trading Commission (CFTC) (2010). Proposed rules. *Federal Register*, 75:33198–202.

Considine, T. (2000). The impacts of weather variations on energy demand and carbon emissions. *Resource and Energy Economics*, 22:295–314.

Davidson, R., Labys, W. C., and Lesourd, J.-B. (1997). Walvelet analysis of commodity price behavior. *Computational Economics*, 11(1-2):103–128.

de Prado, M. L. (2011). *Advances in High Frequency Strategies*. PhD thesis, Dept. of Financial Economics, Complutense University.

DeJong, D. N., Nankervis, J. C., Savin, N. E., and Whiteman, C. H. (1992). The power problems of unit root test in time series with autoregressive errors. *Journal of Econometrics*, 53(1):323–343.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366):427–431.

Dutt, A. and Rokhlin, V. (1993). Fast fourier transforms for nonequispaced data. *SIAM Journal on Scientific computing*, 14(6):1368–1393.

Easley, D., de Prado, M. L., and O'Hara, M. (2012). The volume clock: Insights into the high frequency paradigm. *The Journal of Portfolio Management*, 39(1):19–29.

Easley, D. and O'Hara, M. (1987). Price, trade size, and information in securities markets. *Journal of Financial Economics*, 19:69–90.

Easley, D., Prado, M. L. D., and O'Hara, M. (2011). The microstructure of the flash crash: Flow toxicity, liquidity crashes and the probability of informed trading. *Journal of Portfolio Management*, 37(2):118–128.

Elkhafif, M. (1996). An iterative approach for weather-correcting energy consumption data. *Energy Economics*, 18:211–220.

Engle, R. (2001). Garch 101: The use of arch/garch models in applied econometrics. *Journal of economic perspectives*, pages 157–168.

Engle, R. F. and Granger, C. W. J. (1987). Cointegration and error-correction: Representation, estimation, and testing. *Econometrica*, 55:251–276.

Engle, R. F. and Russell, J. R. (1998). Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica*, pages 1127–1162.

Filimonov, V. and Sornette, D. (2014). Power law scaling and "dragon-kings" in distributions of intraday financial drawdowns. arXiv preprint arXiv:1407.5037.

Gabaix, X. (2008). Power laws in economics and finance. Technical report, National Bureau of Economic Research.

Gabaix, X., Gopikrishnan, P., Plerou, V., and Stanley, H. E. (2003). A theory of power-law distributions in financial market fluctuations. *Nature*, 423(6937):267–270.

Granger, C. W. (1981). Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, 16(1):121–130.

Greengard, L. and Lee, J.-Y. (2004). Accelerating the nonuniform fast fourier transform. *SIAM Review*, 46:443–454.

Hamill, T. M., Bates, G. T., Whitaker, J. S., Murray, D. R., Fiorino, M., Galarneau Jr, T. J., Zhu, Y., and Lapenta, W. (2013). Noaa's second-generation global medium-range ensemble reforecast dataset. *Bulletin of the American Meteorological Society*, 94(10):1553–1565.

Hardiman, S. J., Bercot, N., and Bouchaud, J.-P. (2013). Critical reflexivity in financial markets: a hawkes process analysis. *The European Physical Journal B*, 86(10):1–9.

Hasbrouck, J. and Saar, G. (2013). Low-latency trading. *Journal of Financial Markets*, 16(4):646–679.

Hassler, U. and Wolters, J. (1994). On the power of unit root tests against fractional alternatives. *Economics Letters*, 45(1):1–5.

Hirshleifer, D. and Teoh, S. H. (2003). Herd behaviour and cascading in capital markets: A review and synthesis. *European Financial Management*, 9(1):25–66.

Hommes, C. H. (2006). Heterogeneous agent models in economics and finance. *Handbook of computational economics*, 2:1109–1186.

Iati, R. (2009). *High-frequency Trading Technology: A TABB Anthology*. TABB Group. Available at http://www.tabbgroup.com/PublicationDetail.aspx?PublicationID=498.

Jenkins, B. and Jarvis, M. J. (1999). Mesospheric winds derived from superdarn hf radar meteor echoes at halley, antarctica. *Earth Planet Science*, 51:685–689.

Kanamitsu, M., Alpert, J., Campana, K., Caplan, P., Deaven, D., Iredell, M., Katz, B., Pan, H.-L., Sela, J., and White, G. (1991). Recent changes implemented into the global forecast system at nmc. *Weather and Forecasting*, 6(3):425–435.

Lillo, F. and Farmer, J. D. (2004). The long memory of the efficient market. *Studies in Nonlinear Dynamics & Econometrics*, 8(3).

Linn, S. and Zhu, Z. (2004). Natural gas prices and the gas storage report: public news and volatility in energy futures markets. *Journal of Futures Market*, 24(3):283–313.

Lomb, N. R. (1976). Least-squares frequency analysis of unequally spaced data. *Astrophysics and space science*, 39(2):447–462.

Menkveld, A. J. (2013). High frequency trading and the new market makers. *Journal of Financial Markets*, 16(4):712–740.

Mu, X. (2007). Weather, storage, and natural gas price dynamics: fundamentals and volatility. *Energy Economics*, 29(1):46–63.

Müller, U. A., Dacorogna, M. M., Davé, R., Pictet, O. V., Olsen, R. B., and Ward, J. R. (1993). Fractals and intrinsic time: A challenge to econometricians. *Unpublished manuscript, Olsen & Associates, Zürich*.

Müller, U. A., Dacorogna, M. M., Davé, R. D., Olsen, R. B., Pictet, O. V., and von Weizsäcker, J. E. (1997). Volatilities of different time resolutions – analyzing the dynamics of market components. *Journal of Empirical Finance*, 4(2):213–239.

Pedroni, P. (1999). Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and statistics*, 61(S1):653–670.

Phillips, P. C. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2):335–346.

Pindyck, R. (2004a). Volatility and commodity price dynamics. *Journal of Futures Market*, 24(11):1029–1047.

Pindyck, R. (2004b). Volatility in natural gas and oil markets. *Journal of Energy and Development*, 30(1):1–19.

Podobnik, B., Wangy, D., and Stanley, H. E. (2012). High-frequency trading model for a complex trading hierarchy. *Quantitative Finance*, 12(4):559–566.

Praetz, P. D. (1979). Testing for a flat spectrum on efficient market price data. *The Journal of Finance*, 34(3):645–658.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial management review*, 6(2):41–49.

Scargle, J. D. (1982). Studies in astronomical time series analysis. ii-statistical aspects of spectral analysis of unevenly spaced data. *The Astrophysical Journal*, 263:835–853.

Schoeffel, L. (2011). Time scales in futures markets and applications. arXiv preprint arXiv:1110.1727.

Smith, R. (2010). Is high-frequency trading inducing changes in market microstructure and dynamics? http://ssrn.com/abstract=1632077.

Song, J. H., Lopez de Prado, M., Simon, H. D., and Wu, K. (2014). Exploring irregular time series through non-uniform fast fourier transform. http://ssrn.com/abstract=2487656.

Thomson, R. E. and Emery, W. J. (2014). *Data Analysis Methods in Physical Oceanography*. Elsevier Science.

Vaníček, P. (1969). Approximate spectral analysis by least-squares fit. *Astrophysics and Space Science*, 4(4):387–391.

Vuorenmaa, T. A. (2013). The good, the bad, and the ugly of automated high-frequency trading. *The Journal of Trading*, 8(1):58–74.

Walker, J. S. (1996). *Fast Fourier Transforms*, volume 24. CRC press.

Wallace, J. M. and Hobbs, P. V. (2006). *Atmospheric science: an introductory survey*, volume 92. Academic press.

Wei, M., Toth, Z., Wobus, R., and Yuejian, Z. (2006). Initial perturbations based on the ensemble transform (et) technique in the ncep global operational forecast system. *Tellus A*, 60(1):62–79.

Welch, P. D. (1967). The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Transactions on audio and electroacoustics*, 15(2):70–73.

Wu, K., Bethel, W., Gu, M., Leinweber, D., and Rübel, O. (2013). A big data approach to analyzing market volatility. *Algorithmic Finance*, 2(3):241–267.