Mathematical Finance MSc Dissertation MTH778P, 2018/19

Analytic signal processing for economic time series

A study on applying analytic signal processing to identify extreme events in the Nord Pool market

Bart Kitson, ID 181023708

Supervisor: Dr. Wolfram Just

A thesis presented for the degree of

Master in Sciences in *Financial Computing*

School of Mathematical Sciences

Queen Mary University of London

Declaration of original work

This declaration is made on September 4, 2019.

Student's Declaration: I, Bart Kitson, hereby declare that the work in this thesis is my original work. I have not copied from any other students' work, work of mine submitted elsewhere, or from any other sources except where due reference or acknowledgement is made explicitly in the text, nor has any part been written for me by another person.

Referenced text has been flagged by:

1. Using italic fonts, **and**
2. using quotation marks "…", **and**
3. explicitly mentioning the source in the text.

Acknowledgements

I would like to thank my supervisor Dr Wolfram Just for his time and invaluable advice during the writing and research of this project.

Abstract

This thesis has two parts, a theoretical analysis of the analytic signal and an applied evaluation of this theory to the Nord Pool financial time series data.

The first part details the basics of the analytic signal and how it is used to calculate an instantaneous amplitude and frequency. Signals of a cosine and sum of two cosines, where theoretical solutions exist, are analysed in detail. The relationship between the analytic signal and the Fourier transform is described and how any time signal can be represented in frequency space. The theoretical section concludes by detailing the steps required to produce an analytic signal numerically so discrete times series can be analysed. Several test signals such as linearly varying amplitude, linearly varying frequency and a spike represented by a Gaussian are analysed. The instantaneous amplitude and frequency from the numerically calculated analytic signal are in excellent agreement with the input signals. A phase slip from a Gaussian spike is evident.

The second part looks at the market dynamics of the Nord Pool market to show why extreme price spikes can occur. Seven years of hourly Nord Pool data from the January 1999 is analysed using the techniques detailed in the first section in particular focusing on periods around extreme price spikes. The difference between averaged hourly data and data containing a spike are compared and clearly show the effect of a spike. Finally, the comparison of the results between using return data versus price data are presented and show that phase slips are easier to identify with price data.

Contents

1 Introduction

Signal processing has its roots in electrical engineering and is a field of study that focuses on analysing, modifying, recognising and transforming signals [1]. The signals are generated from a diverse range of sources such as audio, image, video, radar, sonar and financial. The output of signal processing can be equally diverse and covers areas such as compression, filtering, recognition as well as information extraction [2]. In the last 30 years there have been significant developments in digital computing and data storage and almost every industry uses some form of signal analysis.

As there are many applications for signal processing and there are equally many mathematical techniques which have been developed to analyse them. The cornerstone of signal analysis is the Fourier series and Fourier transforms. In this analysis any signal which exhibits periodic properties can be simply expressed as the sum of cosines and sines waves [3].

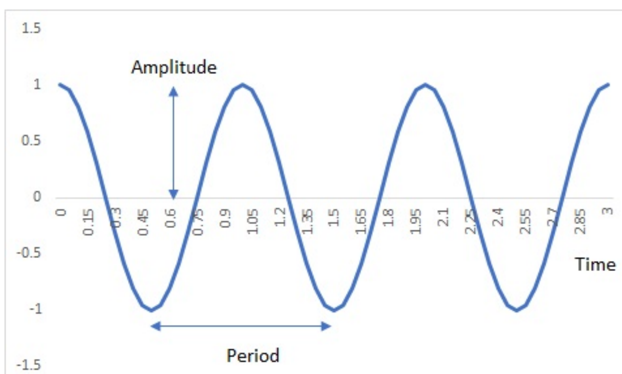


Figure 1: Amplitude and Period of a wave

The basic parameters to quantify a wave are amplitude, frequency and phase. These three measures are independent. The amplitude is a defined as the maximum magnitude displacement from the equilibrium or central value and is always positive. The SI units of amplitude is the meter but the units can vary depending upon the nature of the signal such as decibels for sound or price of an asset in a financial time series. Frequency can be defined as the number repetitive occurrences per unit of time, the SI unit is Hertz which is 1/seconds. Alternatively frequency can be defined as the inverse of the time period of the wave, where period is the time elapsed between two successive similar points (amplitude and direction) on the wave. The amplitude and time period of a typical wave is shown in Figure 1.

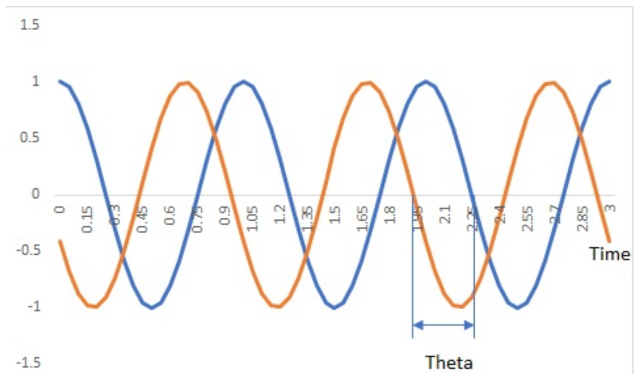


Figure 2: Illustration of phase between two waves $\pm A(t)$

The phase of a wave is defined as an offset of the wave from a given point in time and essentially specifies the starting point. The phase shift describes the timing difference between two waves shown in Figure 2. For sine and cosine waves the units are radians and spans a range of one complete cycle or 2π radians. Consider a cosine wave with frequency w and phase shift θ

$$cos(wt + \theta) = cos(\theta)cos(wt) - sin(\theta)sin(wt) \tag{1}$$

Equation (1) shows how the phase shift is a fixed allocation between a cosine and sine wave of the same frequency. For a phase of zero it is simply a cosine and with a shift of $-\pi/2$ then the wave a sine wave.

Economic time series are often affected by the impact of external recurring factors. For instance, the price of a commodity such as gas, electricity or water may be affected by daily, weekly, or monthly seasonal effects. Therefore the time series of the price may have a periodic component [4]. One way to analyse such a component or to remove seasonal effects from a time series is by using signal processing techniques. By splitting the signal into time dependent amplitudes and phase, seasonal effects can be separated from other stationary properties of the time series and therefore the data itself can be analysed [5].

Using the price of the Nord Pool electricity market from the 1st January 1999 to the 26th January 2007 this project aims to apply analytic signal processing to identify extreme events in the price dynamics of the market and to explore whether features of the phase can be used as a precursor for these extreme events.

2 Analytic Signal

2.1 Definition of an Analytic signal

An Analytic signal $z(t)$ of a real-valued signal $u(t)$ is a complex-valued function where the real part is the original function and the imaginary part is its Hilbert transform $\hat{u}(t)$.

$$z(t) = u(t) + i\hat{u}(t) \tag{2}$$

A Hilbert transform is defined as [6]:

$$H(u(t)) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{u(\tau)}{t - \tau} d\tau \tag{3}$$

The Hilbert transform has the effect of producing a signal with a phase shirt of $\pi/2$ to the original signal. Hence the Hilbert transform of a cosine is a sine [7].

$$H(\pm cos(wt)) = \pm cos(wt - \pi/2) = \pm sin(wt) \tag{4}$$

Hence for positive frequencies:

$$H(cos(wt) = sin(wt) \tag{5}$$

2.2 Instantaneous Amplitude and Frequency

A real valued function may be expressed as a cosine with an instantaneous amplitude A(t) and phase $\phi(t)$:

$$u(t) = A(t)cos(\phi(t)) \tag{6}$$

where $\phi(t)$ can be expressed in terms of an instantaneous frequency w(t) and a constant phase shift ϕ_0:

$$\phi(t) = \phi_0 + \int_0^t w(\tau)d\tau \tag{7}$$

Equation 6 has no unique solution as at each point in time there is a single value of u(t) and two unknowns, the instantaneous frequency and instantaneous amplitude. One way to produce unique values of amplitude and frequency is to define the analytic signal by:

$$z(t) = A(t)e^{(i\phi(t))} \tag{8}$$

and expanding using Euler's formula:

$$z(t) = A(t)(cos(\phi(t)) + isin(\phi(t)) \tag{9}$$

comparing Equation 9 to Equation 2 defines the instantaneous amplitude and instantaneous frequency by:

$$A(t) = (u(t)^2 + \hat{u}(t)^2)^{0.5} \tag{10}$$

$$\phi(t) = atan2(\hat{u}(t)/u(t)) \tag{11}$$

And differentiating Equation 7 defines the instantaneous frequency:

$$w(t) = d\phi/dt \tag{12}$$

Note Equation 11 uses atan2 which gives a singularly defined value for the inverse tan function in the range $-\pi$ to π [8]. In terms of the standard tan^{-1} function $(-\pi/2, \pi/2)$ then it is defined as:

$$atan2(x,y) = \begin{cases} arctan(y/x), & \text{if } x > 0 \\ \pi/2 - arctan(x/y), & \text{if } y > 0 \\ -\pi/2 - arctan(x/y), & \text{if } y < 0 \\ arctan(x/y) \pm \pi, & \text{if } x < 0 \\ undefined, & \text{if } y = 0 \text{ and } x = 0 \end{cases} \tag{13}$$

2.3 Examples of Analytic signal

2.3.1 Analytic signal Example 1

Consider the trivial example when the signal u(t) is a simple cosine wave with amplitude a_1 and constant frequency w_1, then from Equations 2 and 5:

$$u(t) = a_1 cos(w_1 t) \tag{14}$$

$$z(t) = a_1 cos(w_1 t) + i a_1 sin(w_1 t) \tag{15}$$

From Equation 10 we get:

$$A(t) = (a_1^2 cos^2(w_1 t) + a_1^2 sin^2(w_1 t))^{0.5} = a_1 \tag{16}$$

$$\phi(t) = tan^{-1}(sin(w_1 t)/cos(w_1 t) = tan^{-1}(tan(w_1 t)) = w_1 t \tag{17}$$

and from Equation 12:

$$w(t) = d\phi/dt = w_1 \tag{18}$$

hence:

$$u(t) = A(t)cos(\phi(t)) = a_1 cos(w_1 t) \tag{19}$$

The analytic signal is the same as the input signal, the instantaneous amplitude $A(t) = a_1$ and the instantaneous frequency $w(t) = w_1$.

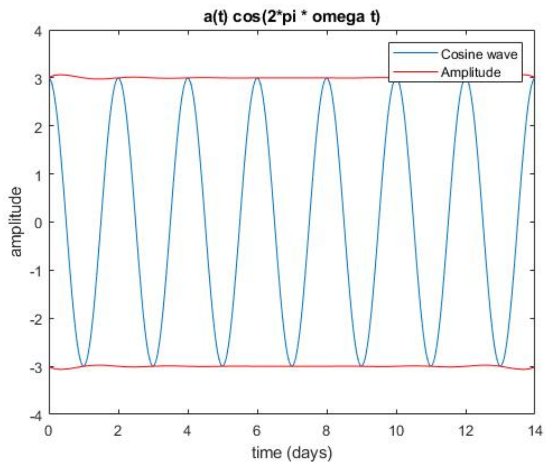


Figure 3: Simple Cosine with the instantaneous amplitude $\pm A(t)$

The chosen units for time are days as this is the most intuitive timescale when considering the Nord Pool data later. Figure 3 shows an example of a simple cosine with frequency (w_1) of $0.5 days^{-1}$ and amplitude (a_1) of 3. The instantaneous amplitude envelope is denoted by $\pm A(t)$.

2.3.2 Analytic signal Example 2

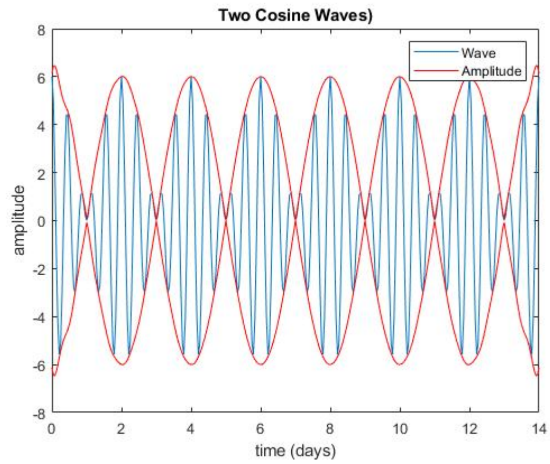


Figure 4: Sum of two cosines of different frequencies.

Consider the more complex case where the signal is the sum of two cosines with different frequencies (Figure 4). The resultant signal produces repetitive beats containing higher frequency oscillations.

$$u = a_1 cos(w_1 t) + a_2 cos(w_2 t) \tag{20}$$

Again using Equations 2 and 5 the analytic signal is:

$$z(t) = (a_1 cos(w_1 t) + a_2 cos(w_2 t)) + i(a_1 sin(w_1 t) + a_2 sin(w_2 t)) \tag{21}$$

From Equation 10:

$$A(t) = (a_1^2 + a_2^2 + 2a_1 a_2 cos(w_1 t)cos(w_2 t) + 2a_1 a_2 sin(w_1 t)sin(w_2 t))^{0.5} \tag{22}$$

using $cos(A - B) = cos(A)cos(b) + sin(A)sin(b)$ and $cos^2(x) = (1 + cos(2x))/2$ then if $a_1 = a_2$.

$$A(t) = |2cos((w_1 - w_2)/2)| \tag{23}$$

Also giving ϕ as:

$$\phi = tan^{-1}(\frac{a_1 sin(w_1 t) + a_2 sin(w_2 t)}{a_1 cos(w_1 t) + a_2 cos(w_2 t)} \tag{24}$$

and using:

$$sin(A) + sin(B) = 2sin((A + B)/2)cos((A - B)/2) \tag{25}$$

and

$$cos(A) + cos(B) = 2cos((A + B)/2)cos((A - B)/2) \tag{26}$$

and if $a_1 = a_2$ the we get:

$$\phi(t) = tan^{-1}(\frac{2a_1 sin((w_1 + w_2)/2)cos((w_1 - w_2)/2)}{2a_1 cos((w_1 + w_2)/2)cos((w_1 - w_2)/2)}) \tag{27}$$

which reduces to:

$$\phi(t) = tan^{-1}(tan(\frac{w_1 + w_2}{2} t)) = (w_1 + w_2)t/2 \tag{28}$$

and therefore:

$$w(t) = (w_1 + w_2)/2 \tag{29}$$

Equations 23 and 29 give the interesting result that the instantaneous amplitude A(t) has a frequency of half the difference of the frequency of the two input waves and the instantaneous frequency w(t) is simply the average frequency of the two input waves. The instantaneous amplitude is shown in Figure 4 and captures the amplitude of the beats.

More complex signals such as three cosines do not easily decompose and the instantaneous amplitude and frequency must be calculated numerically.

3 Fourier

3.1 Why use Fourier transform?

Before looking at the details of the Fourier transform it is worth noting the key characteristics and how it relates to the analytic signal. A signal in time space $u(t)$ may be expressed in frequency space. This amounts to representing the time signal as a sum of cosines with different amplitudes (a_i), frequencies (w_i) and phase (θ_i).

$$u(t) = \sum a_i * cos(w_i t + \phi_i) \tag{30}$$

In this form the analytic signal can be easily calculated as the Hilbert Transform of cosine is sine.

$$z(t) = \sum a_i * cos(w_i t + \phi_i) + i \sum a_i * sin(w_i t + \phi_i) \tag{31}$$

The Fourier transform is a technique which can decompose any signal into the frequency domain [3] to enable a_i and ϕ_i in Equation 30 to be calculated [9].

3.2 Fourier transform explained

The definition of a periodic signal is:

$$u(t) = u(t + nT) \text{ where } n = 0, \pm 1, \pm 2, ... \tag{32}$$

Fourier transform splits a function in the time domain into the frequency domain.

$$\hat{f}(w) = \int_{-\infty}^{\infty} u(t) e^{-2\pi i x w} dt \tag{33}$$

The Fourier transform works because the underlying cosines and sines are orthogonal (product of different frequencies integrate to zero). This property can be seen by considering the

Fourier transform of simple cosine with a phase shift. Using:

$$u(t) = acos(wt + \phi) \tag{34}$$

and the fact that:

$$u(t) = a(cos(/phi)cos(wt) - sin(/phi)sin(wt) \tag{35}$$

$$e^{ix} = cos(x) + isin(x) \tag{36}$$

and

$$\int cos(w_i t)cos(w_j t) = 0 \text{ unless } i = j \tag{37}$$

$$\int sin(w_i t)cos(w_j t) = 0 \text{ for all } i, j \tag{38}$$

$$\int sin(w_i t)sin(w_j t) = 0 \text{ unless } i = j \tag{39}$$

Then we get:

$$f(u(zi)) = \begin{cases} acos(\phi) + iasin(\phi), & \text{if } zi = w \\ 0 + 0i, & \text{otherwise} \end{cases} \tag{40}$$

So any signal can be decomposed into frequencies, shown in Figure 5 [10].

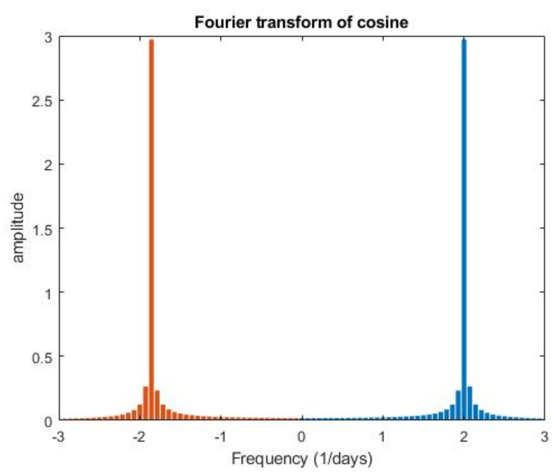


Figure 5: Fourier transform of cosine wave

3.2.1 Example of Fourier transform

In addition to the simple cosine there are many functions where there is a theoretical solution to the Fourier transform. Part of this project considers spikes or extreme values in Nord Pool data. Hence one function of particular interest is the Gaussian shown in Figure 6 [11].

$$u(t) = (1/((2\pi a))e^{(-1/2*((t-t_s)/a)^2)} \tag{41}$$

From Equation 41 it can be seen that as a tends to zero this becomes an infinite spike or dirac delta function.

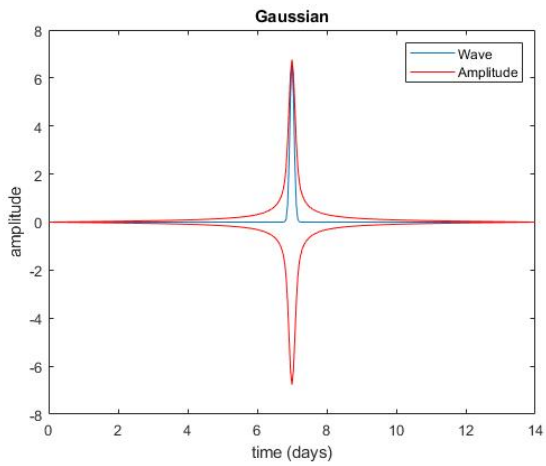


Figure 6: Gaussian centred around t=7 days.

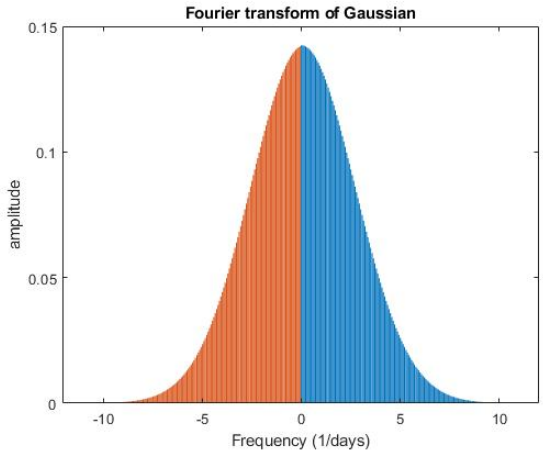


Figure 7: Amplitude of FT of Gaussian

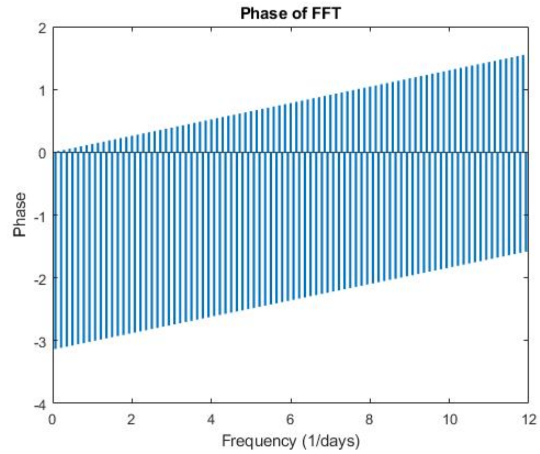


Figure 8: Phase of FT of Gaussian

The Fourier transform of a Gaussian can be derived as follows:

Proof.

$$
\begin{aligned}
H(w) &= \frac{1}{\sqrt{2\pi * a^2}} * \int_{-\infty}^{\infty} e^{-\frac{t^2}{2a^2}} e^{2\pi i w t} dt \\
&= \frac{1}{\sqrt{2\pi * a^2}} * \int_{-\infty}^{\infty} e^{-\frac{1}{2a^2}(t^2 - 4\pi a^2 t w)} dt \\
&= \frac{1}{\sqrt{2\pi * a^2}} * \int_{-\infty}^{\infty} e^{-\frac{1}{2a^2}(t^2 - 2(2\pi a^2 w)t + (2\pi i a^2 w)^2 - (2\pi i a^2 w)^2)} dt \\
&= \frac{1}{\sqrt{2\pi * a^2}} * (\int_{-\infty}^{\infty} e^{-\frac{1}{2a^2}(t - 2\pi a^2 w)^2} dt) e^{\frac{1}{2a^2}(2\pi a^2 w)^2} \\
&= \frac{1}{\sqrt{2\pi * a^2}} * (\sqrt{2\pi * a^2}) * e^{-2\pi^2 a^2 w^2} \\
&= e^{-2\pi^2 a^2 w^2}
\end{aligned}
$$

□

Therefore [12]:

$$u(w) = e^{-2\pi^2 a^2 w^2} \tag{42}$$

Equation 42 shows that a Gaussian in time space is also a Gaussian in frequency space as shown in Figure 7. Note also that as a goes to 0 then in time space the spike tends to a dirac delta function, whereas in frequency space u(w) goes to 1 and is independent of frequency (w). This means that there is an equal contribution to the signal from all frequencies.

Note that for the examples shown (Figures 5, 7) the Fourier transform is symmetrical about zero frequency. In fact this is always true if the signal is real-valued i.e. it has Hermitian symmetry (f*(x)= f(-x)) [13]. Therefore you only need to consider positive frequencies and use double the amplitude.

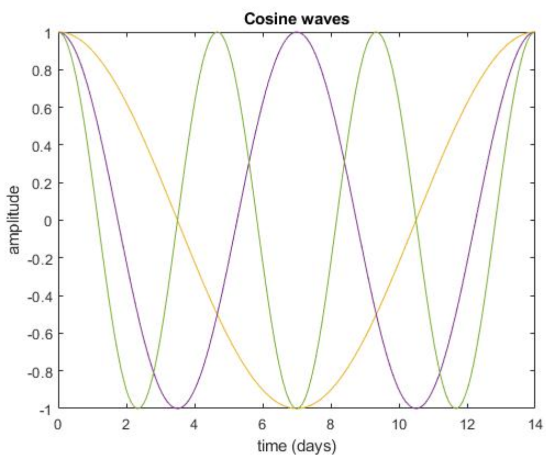


Figure 9: Example of cosine waves with integer period

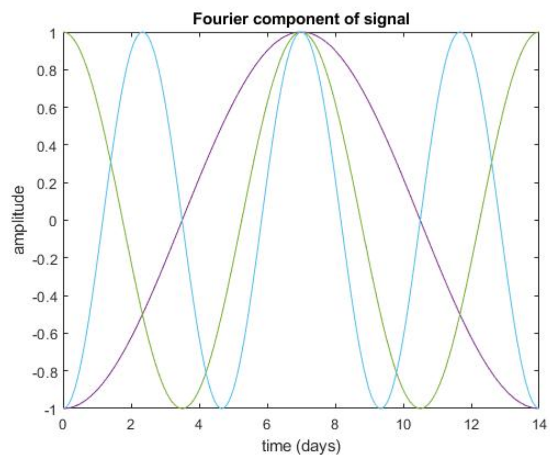


Figure 10: Contribution of first few frequencies to Gaussian

In order to gain an intuition of the Fourier transform and specifically the output coefficients a_i and θ_i it is worth looking at the transform of the Gaussian in more detail. The basic building blocks are cosines with amplitude 1 which have integer number of oscillations, the first three are depicted in Figure 9. The Fourier transform fits a weight (a_i) and a phase (θ_i) to each one of these cosine waves. As the spike occurs at t=7 days the θ_i will align each cosine wave so they have a maximum at t=7 (Figure 9 and 10), the a_i are fitted to get the height of the spike and ensure the signal is zero everywhere else.

3.3 Discrete Fourier transform

The theory for the Fourier transform was derived in continuous time over an infinite time span. In practice most data sets are of finite length and consist of observations at discrete points in time, often at regular intervals. Hence, to analyse data in this form the Discrete Fourier transform must be used. The DFT is a modification to the continuous time Fourier transform described in Equation 33 and converts a discrete signal U_n observed at N equal intervals of time into N equal intervals of frequency $\hat{U}_i$ [14]:

$$U = u_0, u_1, ..., u_{N-2} \tag{43}$$

into:

$$\hat{U}_i = \hat{u}_0, \hat{u}_1, ..., \hat{u}_{N-2} \tag{44}$$

The equal interval of frequency is given by:

$$w_i = i/N \text{ where } i = 1, 2, ...N-1 \tag{45}$$

The DFT is defined by:

$$\hat{u}_i = \frac{1}{N}\sum_{N=0}^{N-1} U_n e^{-(2\pi w_i n)i} \tag{46}$$

Expanding Equation 46 with Eulers formula:

$$\hat{u}_i = \frac{1}{N}\sum_{N=0}^{N-1} U_n(cos(2\pi w_i n) - isin(2\pi w_i n)) \tag{47}$$

which may be written as:

$$\hat{u}_i = b_i + ic_i \tag{48}$$

where:

$$b_i = \frac{1}{N}\sum_{N=0}^{N-1} U_n cos(2\pi w_i n) \tag{49}$$

and

$$c_i = \frac{1}{N}\sum_{N=0}^{N-1} U_n sin(2\pi w_i n) \tag{50}$$

Let the discrete time signal U_n be the sum of cosines with frequency w_i and phase θ_i:

$$u(t) = \sum_{i=1}^{N/2} a_i cos(2\pi w_i + \theta_i) \tag{51}$$

Using the identity $cos(A+B) = cosAcosB - sinAsinB$:

$$U_n = \sum_{N=0}^{N-1}(a_i cos(\theta_i)cos(2\pi w_i n)) + \sum_{N=0}^{N-1}(a_i sin(\theta_i)sin(2\pi w_i n)) \tag{52}$$

Sub U_n in Equation 52 into Equation 47:

$$\begin{aligned}\hat{U}_i = \frac{1}{N}\sum_{N=0}^{N-1}(\sum_{N=0}^{N-1}(a_i cos(\theta_i)cos(2\pi w_i n)) + \sum_{N=0}^{N-1}(a_i sin(\theta_i)sin(2\pi w_i n)))cos(2\pi w_i n) \\ - \frac{1}{N}\sum_{N=0}^{N-1}(\sum_{N=0}^{N-1}(a_i cos(\theta_i)cos(2\pi w_i n)) + \sum_{N=0}^{N-1}(a_i sin(\theta_i)sin(2\pi w_i n)))sin(2\pi w_i n)\end{aligned} \tag{53}$$

Using Orthogonality conditions:

$$\begin{aligned}\frac{1}{N}\sum_{n=0}^{N-1} cos(w_i n)cos(w_j n) &= 1 \text{ if } i = j \\ &= 0 \text{ otherwise}\end{aligned} \tag{54}$$

$$\frac{1}{N}\sum_{n=0}^{N-1} cos(w_i n)sin(w_j n) = 0 \text{ for all i,j} \tag{55}$$

$$\begin{aligned}\frac{1}{N}\sum_{n=0}^{N-1} sin(w_i n)sin(w_j n) &= 1 \text{ if } i = j \\ &= 0 \text{ otherwise}\end{aligned} \tag{56}$$

Sub Equation 54-56 into Equation 53:

$$\hat{U}_i = a_i cos(\theta_i) + isin(\theta_i) \tag{57}$$

And comparing Equation 57 to 48 you get:

$$a_i = (b_i^2 + c_i^2)^{0.5} \tag{58}$$

and

$$\theta_i = tan^{-1}(c_i/b_i) \tag{59}$$

4 Numerical Calculations of Analytic signal

The analytic signal can now be calculated for any signal with the following steps:

1. Compute Fourier coefficients using Equations 49 and 50 to get b_i and c_i

2. Compute a_i and ϕ_i from Equations 58 and 59.

3. Calculate u(t) and $\hat{u}(t)$ from Equations 2 and 5

4. Calculate A(t) and $\phi(t)$ from Equations 10 and 11

5. Calculate w(t). In order to calculate instantaneous frequency from Equation 12 the phi must be "unwrapped". This is a process to remove the discontinuities in phi(t) by adding 2π when the change is greater than π [15].

The data set for Nord Pool is recorded hourly so to keep timescales consistent in the examples we consider time in days with smallest increment being 1/24. Hence for 2 weeks this would be 336 samples (2*7*24)

4.1 Numerical Calculation of Analytic signal Example 1

Starting with the trivial example described in section 2.3.1 and following the 5 steps outlined above.

$$u(t) = a_1 cos(2 * pi * w_1 * t) \tag{60}$$

and setting $a_1 = 3$ and $w_1 = 0.5$

A Discrete Fourier transform will transform the signal into frequency space, note that the frequency buckets are, given by Equation 45, $n/total_{time}$ i.e $1/14$, $2/14$ etc.

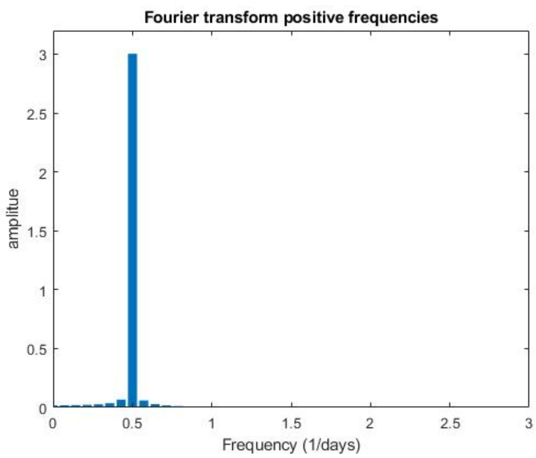


Figure 11: Amplitude from the FT of cosine (step 1)

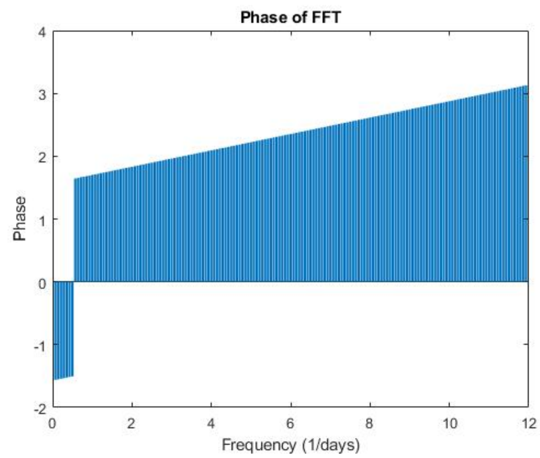


Figure 12: Phase from the FT of cosine (step 1)

Step 1: Figure 11 shows the expected result that the amplitude of the DFT is 3 at frequency of 0.5 and zero everywhere else. Figure 12 shows the phase of zero at frequency of 0.5, the other phase numbers for other frequencies are not important as they have zero amplitude

Step 2:

$$u(t) = 3cos(2 * \pi * 0.5 * t + 0) \tag{61}$$

Step 3:

$$z(t) = 3cos(2 * \pi * 0.5 * t) + i3sin(2 * \pi * 0.5 * t) \tag{62}$$

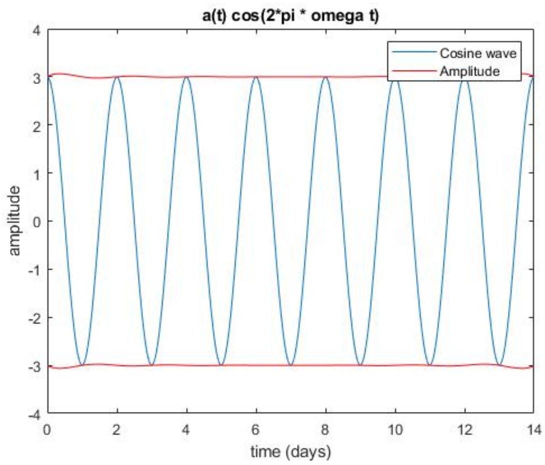


Figure 13: Signal (cosine) and A(t) evaluated numerically

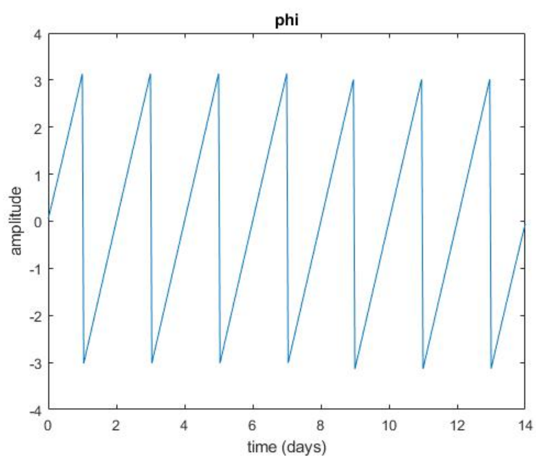


Figure 14: phi(t) for cosine

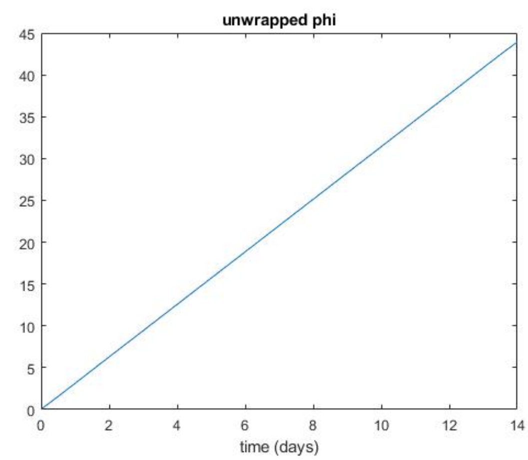


Figure 15: Unwrapped phi

Step 4:

$$A(t) = (9cos^2(2\pi 0.5t) + 9sin^2(2\pi 0.5t)^{0.5} \tag{63}$$

$$\phi(t) = atan2(3sin(2 * \pi * 0.5)/3cos(2 * \pi * 0.5)) \tag{64}$$

A(t) and $\phi(t)$ are evaluated numerically for each discrete t, clearly in this simple case A(t)=3 and $\phi(t) = 0.5t$

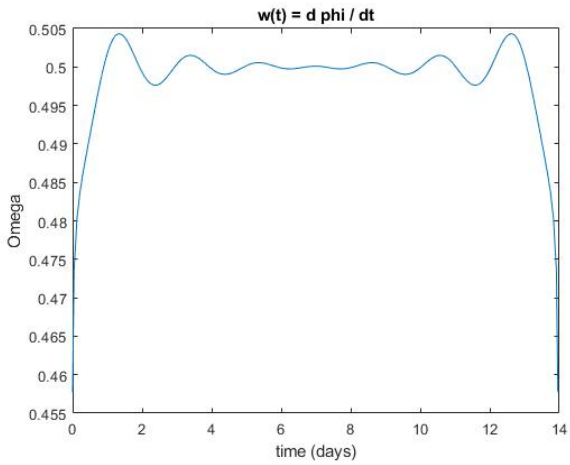


Figure 16: Instantaneous frequency of cosine signal

Step 5: $\phi(t)$ is unwrapped (Figures 14 and 15). The unwrapped phi (ϕ_{un}) can now be differentiated numerically to produce w(t) as shown in Figure 16:

$$w(t) = 1/(2\pi) * (\phi_{un}(t+1)\breve{}\phi_{un}(t))/dt \tag{65}$$

Where dt is $1/24$ in this case. Again this could be done analytically for this simple case.

4.2 Numerical Calculation of Analytic Signal Example 2

For the remaining examples the same 5 steps are followed, however, only the salient data will be shown. Using the example of two cosines described in section 2.3.2 shown in Figure 4:

$$u = a_1 cos(w_1 t) + a_2 cos(w_2 t) \tag{66}$$

Where $a_1 = a_2 = 3$ and $w_1 = 2.0$ and $w_2 = 2.5$

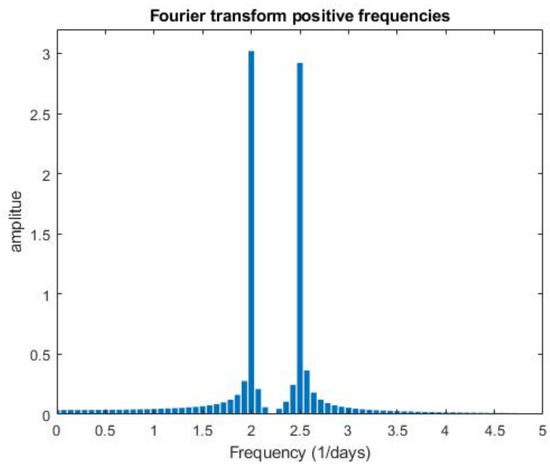


Figure 17: Amplitude of FT for sum of two cosines

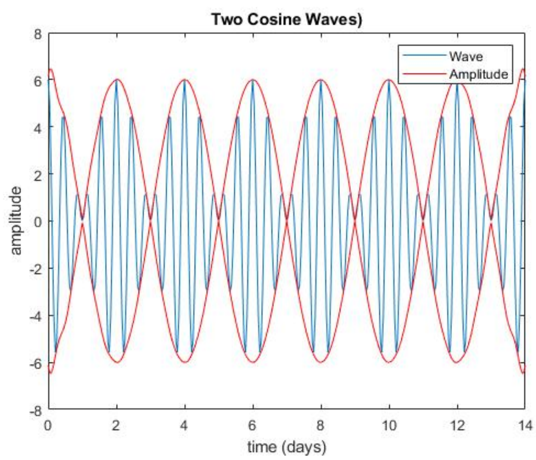


Figure 18: Signal and instantaneous amplitude for sum of two cosines

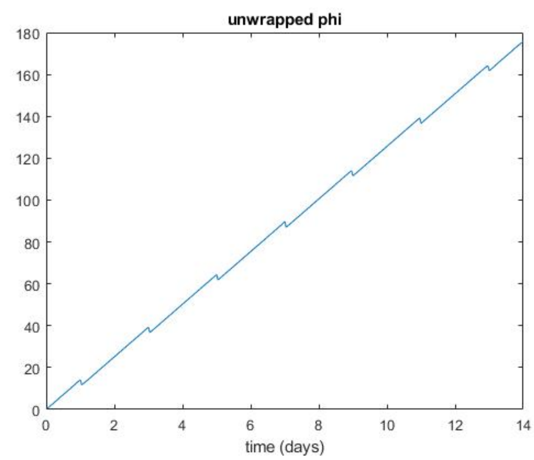


Figure 19: The unwrapped phi (using Matlab function)

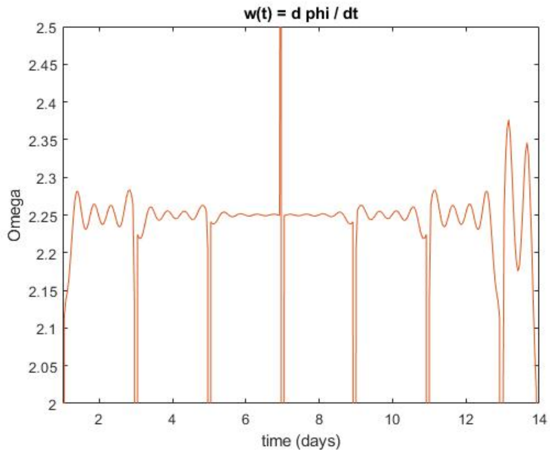


Figure 20: Instantaneous frequency obtained by numerically differentiating unwrapped phi

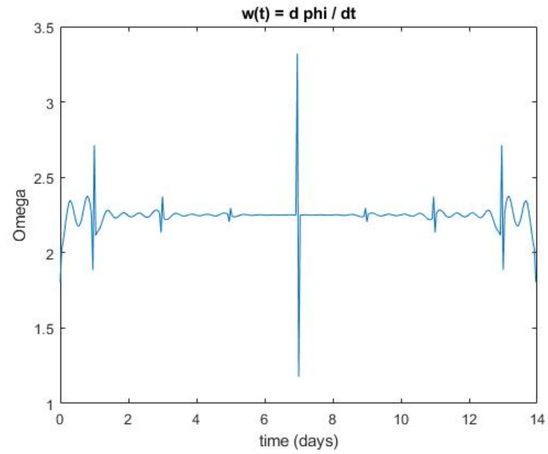


Figure 21: Instantaneous frequency obtained by adjusting Matlab unwrapped phi

Figure 17 shows the result from the FT on the signal, it correctly identifies the two frequencies at 2 and 2.5 with amplitude of 3. Figure 18 shows that he instantaneous amplitude agrees with the theoretical values derived in section 2.3.2. This is not surprising given the Fourier transform had correctly identified the two frequencies. Figure 19 shows the unwrapped phi (using Matlab) function displays some “kinks”. Figure 20 shows the numerical differentiation of this unwrapped phi and gives some numerical spikes around the theoretical value of 2.25 (average of two signal frequencies) derived in section 2.3.2. These spikes are caused from the phi calculation when both u and $\hat{u}$ are zero and the atan2 definition (Equation 11) shows that the function value is undefined at these points. These spikes can be removed by ensuring phi has not changed by more than π in a single time step as shown in see Figure 21. The procedure involves subtracting integer values of π from $\phi(t+1) - \phi(t)$ to ensure value of this difference lies between 0 and π.

So far the signals have been simple cosines (or sum of two cosines) where a theoretical solution exists. The excellent agreement with theoretical values is not surprising as the DFT has correctly identified the frequencies in the signal and so the theoretical and numerical solutions are the same. The next three examples look at some test signals where a theoretical solution is not available. The output can however be compared to the properties of the input signal.

4.3 Numerical Calculation of Analytic Signal Example 3

Consider the more complex case of cosine with time varying amplitude:

$$a(t) = 1 + 0.01t \tag{67}$$

Using a fixed frequency of 0.5, the input signal becomes:

$$u(t) = (1 + 0.01t)cos(2 * \pi * 0.5 * t) \tag{68}$$

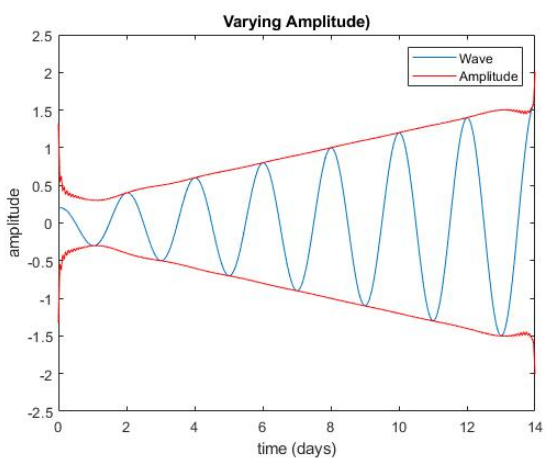


Figure 22: Signal of Cosine with linearly varying amplitude

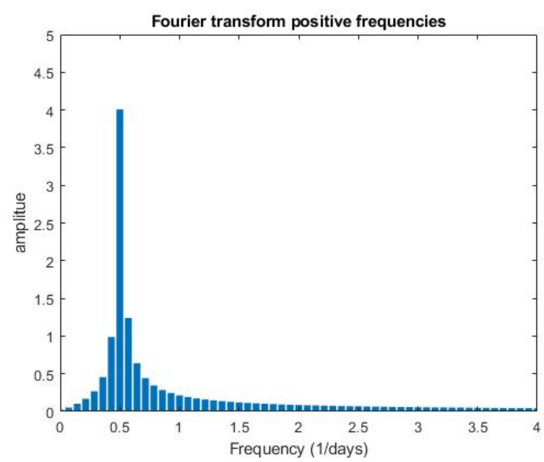


Figure 23: Amplitude from DFT of Cosine with linearly varying amplitude

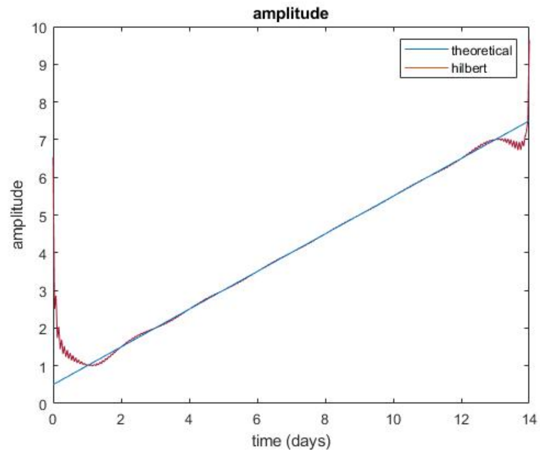


Figure 24: Instantaneous Frequency matches the signal value of 0.5

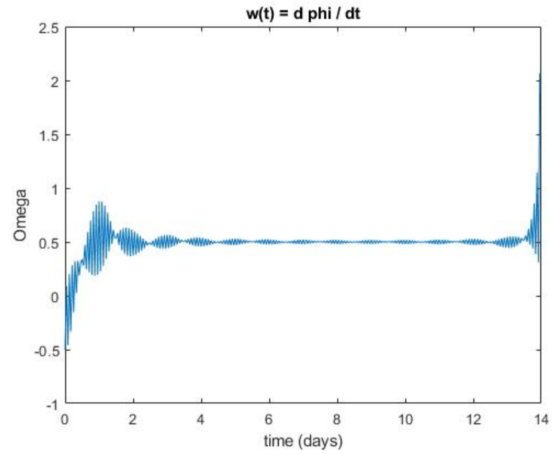


Figure 25: Instantaneous Amplitude and the theoretical amplitude

The resultant signal is shown in Figure 22. The DFT of this signal, Figure 23, shows non-zero amplitudes over a range of frequencies around the predominant input frequency of 0.5. Figure 25 shows that the Instantaneous Frequency matches the signal value of 0.5 with some small instability at the start and end time values. Figure 24 shows the instantaneous amplitude matches the time varying amplitude of the input signal again with small amount of numerical instability at the edges.

4.4 Numerical Calculation of Analytic Signal Example 4

Next input signal to consider is a cosine with time varying frequency:

$$w(t) = 0.2 + 0.1t \tag{69}$$

Hence w(t) varies linearly from 0.2 to 1.6 over the 14 day period.

From Equation 6:

$$u(t) = cos(2 * \pi * (0.2t + 0.1 * 0.5 * t^2)) \tag{70}$$

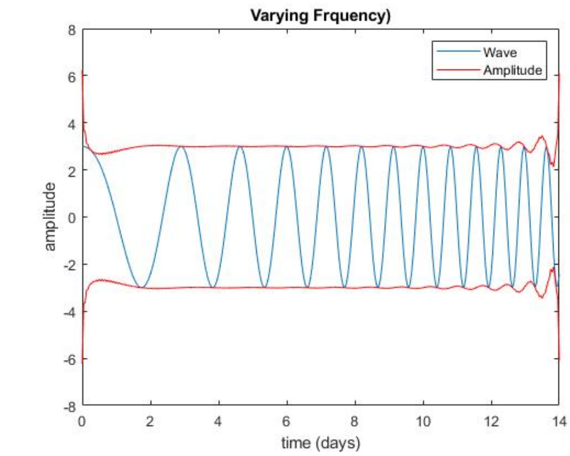


Figure 26: Signal of Cosine with linearly increasing frequency

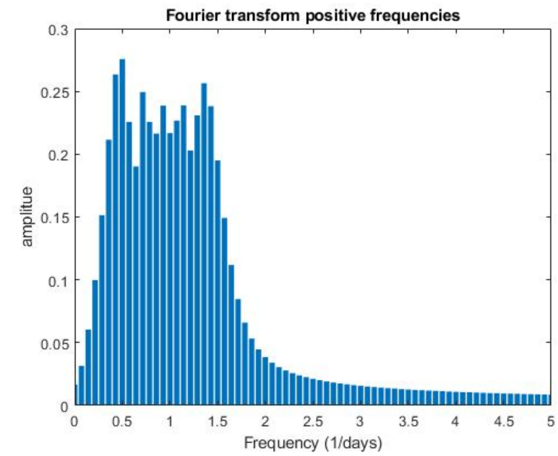


Figure 27: DFT for linearly increasing frequency

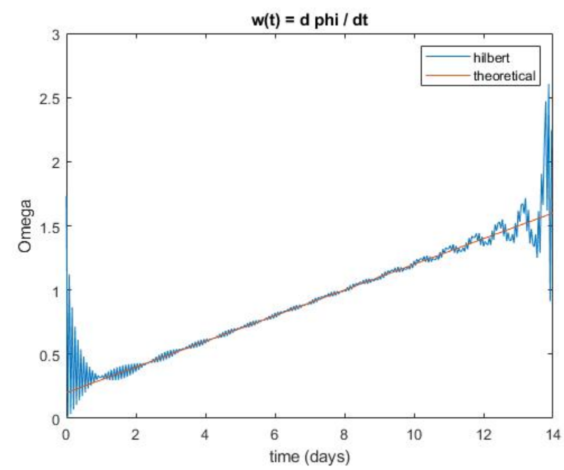


Figure 28: Instantaneous Frequency with theoretical frequency from input signal

The input signal is shown is Figure 26. Figure 27 shows the amplitude from the DFT, this doesn't appear as a smooth function. The highest amplitude are in the frequency buckets ranging from 0.2 to 1.6 which is in line with the input signal. Figure 26 also shows that the instantaneous amplitude is around 1 which again in good agreement with the input signal. There is small area of numerical instability at the start and end time values. Figure 28 shows the instantaneous frequency matching the input signal frequency with the exception of the edge values.

4.5 Numerical Calculation of Analytic signal Example 5

Finally, the project is considering spikes or extreme values in Nord Pool data and so it is worth analyzing the analytic signal from a spike. The simplest spike would be a Dirac Delta function although this is not possible numerically so the Gaussian approximation described in section 3.2.1 Equation 41 is used. We chose $t_s = 7$ days with value of $a = 1/(12)$.

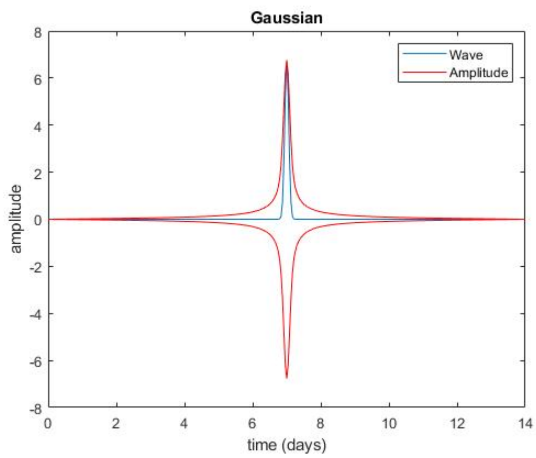


Figure 29: Gaussian centred around t=7 days

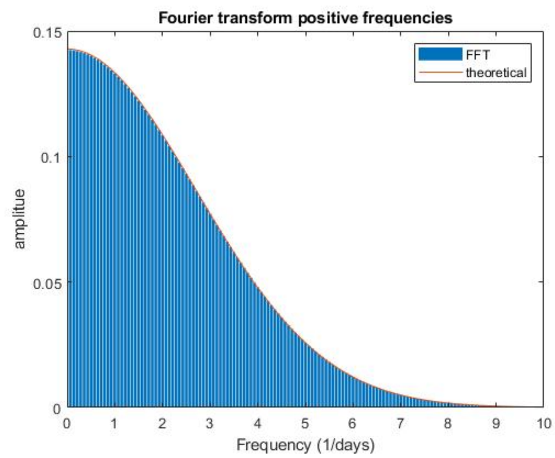
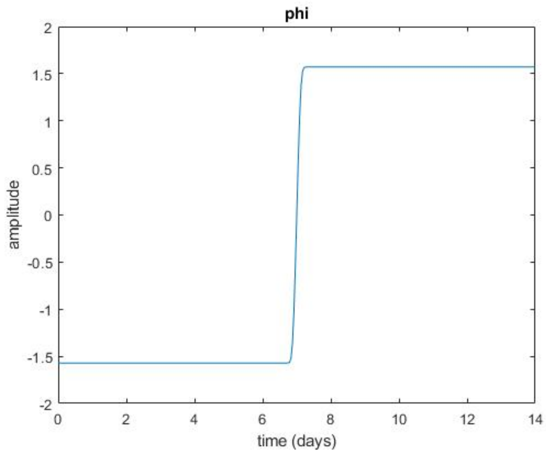


Figure 30: DFT of Gaussian.



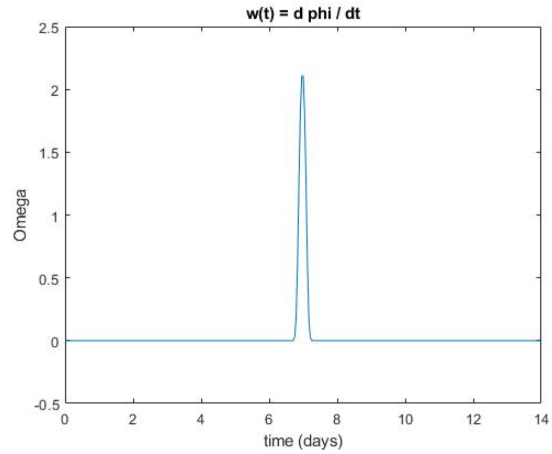


Figure 31: Instantaneous phase of Gaussian from DFT

Figure 32: Instantaneous frequency from Gaussian Signal

The instantaneous amplitude A(t) captures the spike well although is somewhat wider than the signal (Figure 29). The instantaneous phase (Figure 31) shows the phase shift in the region of the spike. This gives rise to a spike in the instantaneous frequency shown in Figure 32. This would be expected as a very high frequency is required to produce a spike or extreme change in value. This theoretical characteristic phase “slip” or instantaneous frequency spike can be compared to real price data containing spikes obtained from the Nord Pool market.

5 Nord Pool

5.1 Nord Pool Market

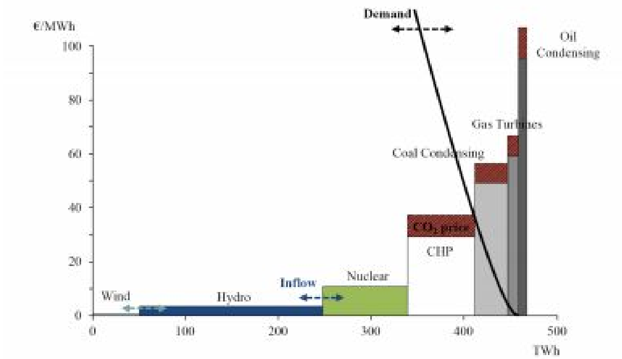


Figure 33: Power production price curve in Nordic electricity market. [16]

Figure 33 shows the different methods of power production for the Nord Pool market. On the x-axis is the total annual production and the y-axis shows the production cost. The width of each block shows the typical capacity for that method of power production. The wind and hydro production is very cheap however their output is variable and depends upon parameters which cannot be controlled, namely amount of wind and rain. Other forms of power such as nuclear, combined Heat and Power (uses coal), coal gas and oil are controllable but are more costly. Spikes in prices occur when demand starts to exceed supply and this is usually due to low supply. There are two main reasons for low supply. The first is weather related, if there is not enough wind / rain then this can reduce the amount produced from wind and hydro [17]. The second is an outage at a power facility through planned maintenance or unforeseen failure. The reduction in cheap supply causes the demand to be met by the more expensive production which causes a spike in the price [18].

5.2 Analysis of Nord Pool Data

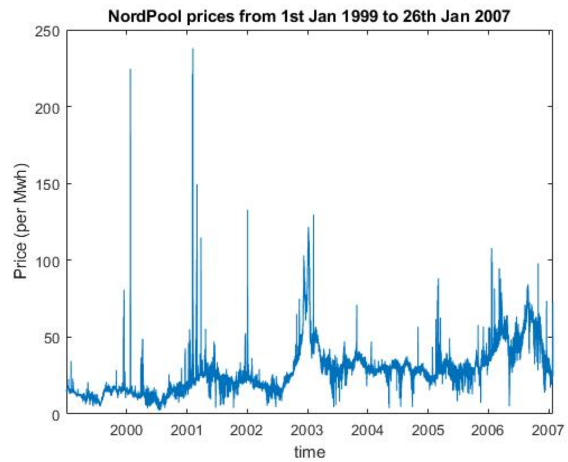


Figure 34: Hourly Nord Pool prices from 1st Jan 1999 0:00 to 26th Jan 2007 23:00

The data set analysed was just over 8 years of hourly Nord Pool prices in Eur per Megawatt Hour (EUR MWh) from 1st Jan 1999 0:00 to 26th Jan 2007 23:00 which contained 70,752 data points shown in Figure 34. The average price is 27.48 and there are several extreme price spikes (both up and down) with a maximum price of 238.01 and a minimum of 2.33.

5.2.1 Analysis of Average Nord Pool data

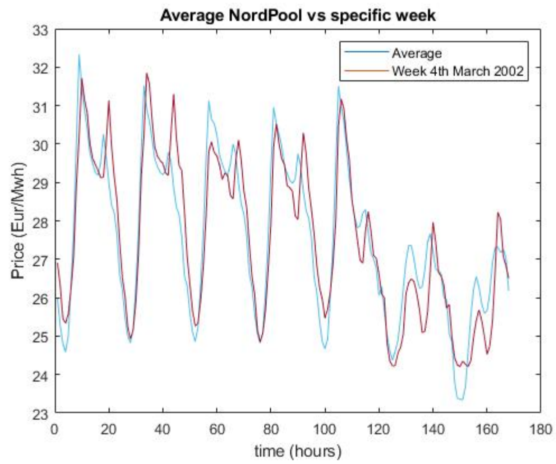


Figure 35: Average Nord Pool price for given hour and day of week using hourly data from 1 Jan 1999 0:00 to 26th Jan 20087 23:00 and typical week of 4th Mar 2002 (averaged price matched)

For each day at a given time there 421 individual prices. Figure 35 shows the average price per hour / day of the week over the dataset along with a typical week (4th Mar 2002 price is adjusted so the average weekly price matches). The price evolution over the week follows a periodic pattern which is consistent with demand. The prices reach a peak at around 8am each workday as industry begins and there is peak domestic demand. The prices stay high throughout the day with another mini peak at 5pm. After 5pm the prices decrease and reach a low around 3am as industrial and domestic demand fall. The prices are typically much lower at the weekend as there is limited industrial demand.

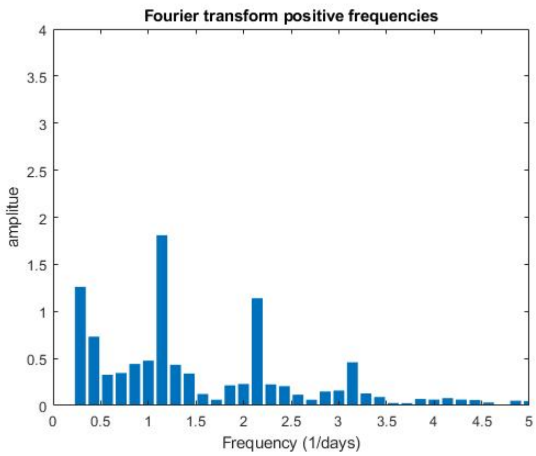


Figure 36: Positive frequencies of DFT

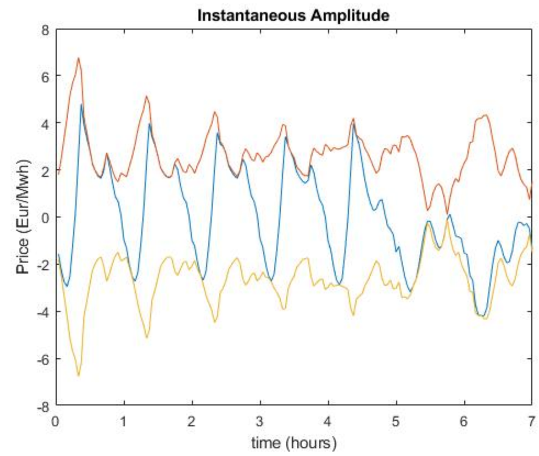


Figure 37: Instantaneous amplitude of signal of demeaned average data starting on Monday

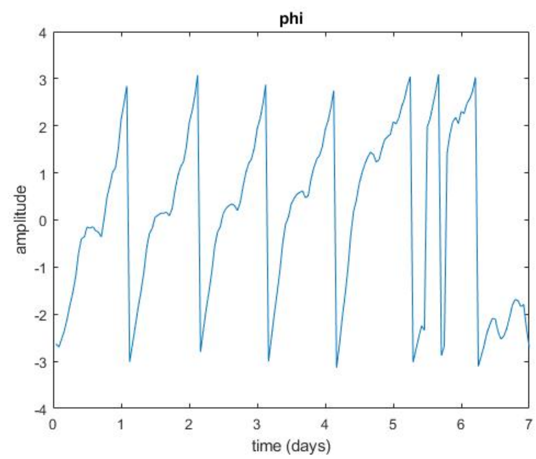


Figure 39: Instantaneous phase from DFT

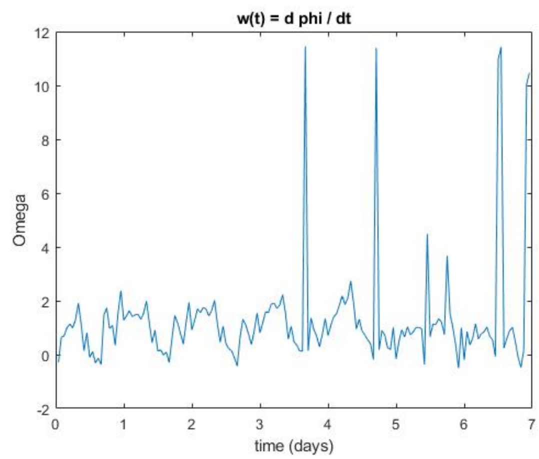


Figure 40: Instantaneous amplitude of signal unwrapped

Using the step 1 to 5 detailed in section 4.1 the instantaneous amplitude and frequency can be derived from the average data in shown in Figure 35. The outputs are shown in Figures 36 to 40. The Fourier analysis (figure 36) shows amplitude peaks around frequencies of

1,1/2 and 1/3 of a day. This is in line with the daily periodicity the data exhibits during the weekdays. Note the instantaneous amplitude (figure 37) is relatively stable during the weekdays with the daily fluctuations being explained by the periodicity. At the weekend the periodicity breaks down and the fluctuation are explained by changes in the instantaneous amplitude. The instantaneous phase (Figure 38) shows similar behaviour as phi is relatively well behaved during the weekdays and becomes unstable at weekends. Figure 39 also shows the week end data to be very noisy.

5.2.2 Analysis of typical Price Spike

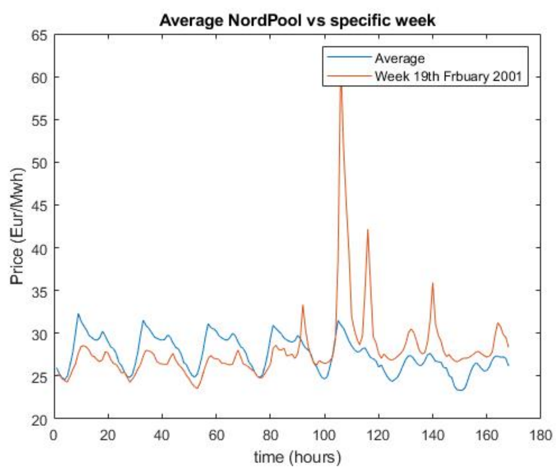


Figure 41: Hourly Prices for week of 19 Feb 2001 and average price.

Figure 41 shows the hourly price evolution for the week of 19th Feb 2001 with the average hourly price shown for comparison using left hand scale to visually match prices at the start of the week. The prices match the average price data from Monday 19th Feb through to Thursday morning 22nd Feb. Then on Thursday afternoon the price spikes at 6pm and deceases overnight albeit at elevated levels versus average prices. On Friday at 8am as the demand rises the prices spikes over 100 percent with another large spike associated with the

5pm peak demand. This volatile behaviour continues through the Saturday and to a lesser extent on the Sunday with the prices at elevated levels versus the average.

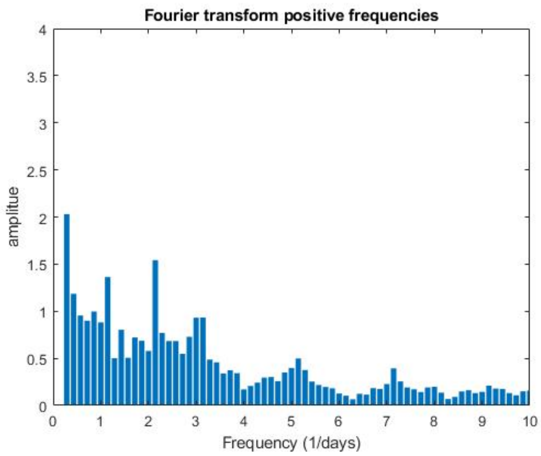


Figure 42: Positive frequencies of DFT

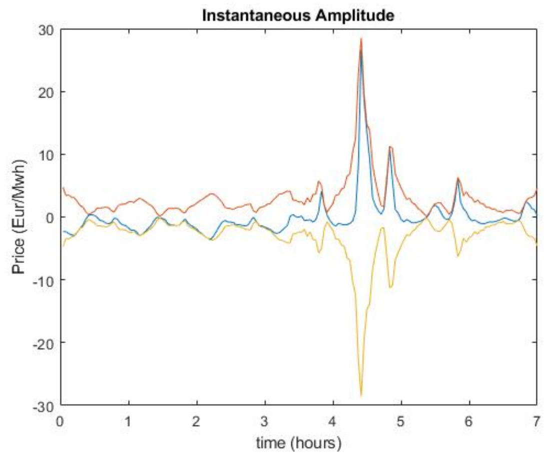


Figure 43: Instantaneous amplitude of the signal

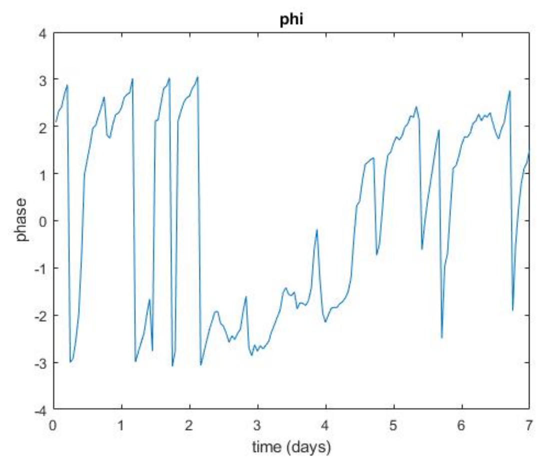


Figure 44: Phase from DFT of average data

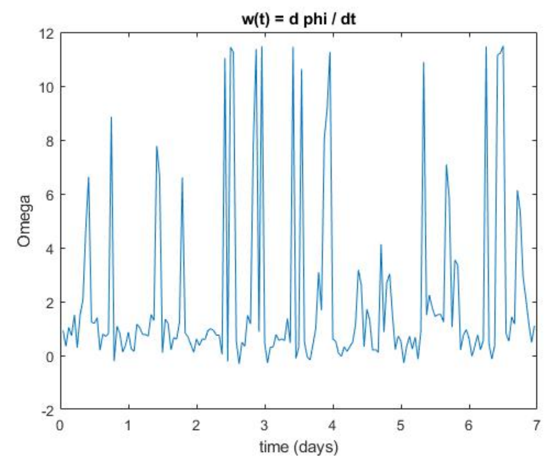


Figure 45: Instantaneous frequency of signal

Again following steps 1 – 5 to create the instantaneous amplitude and frequency for the data of the week of 19th Feb 2001, the output is shown in Figures 42 to 45. It is interesting to compare the differences between average data and data where there is an extreme price move, in this way the effect of a price spike can be observed as opposed to features we are seen in typical weeks. The DFT shown is Figure 42 has amplitudes at a range of frequencies with the spikes at 1,1/2 and 1/3 of days much less pronounced compared to the average data shown in figure 36. The instantaneous amplitude in Figure 43 increases to accommodate the spike. There is a significant difference in the instantaneous phase (Figure 39 and 44). The average data shows a regular pattern during the weeks days where the phase with the spike shows that as early as day 2 there is significant disruption to the regular pattern. The instantaneous frequency shows similar differences (Figure 40 compared to Figure 45) as the number of frequency spikes increases to several per day in the presence of a spike.

5.3 Comparison between Price and Log Return Data

The Nord Pool market has ninety occurrences of upwards and downward spikes over the eight year dataset. To analyse the specific spikes a dataset is taken of three hundred points centering at the spike that is being analysed. To analyse the spikes two approaches are used, firstly a Hilbert transform is done on the original data and the phase is found. The second approach is using a lognormal return, a lognormal return is calculated using:

$$u_{log}(t) = log(\frac{u(t)}{u(t-1)}) \tag{71}$$

and is used to get rid of the natural exponential increase shown in the market.

What is also analysed is commonly refered to as a rolling dataset, to model the rolling dataset the data leading up to but not including the spike is used a the same hilbert transform is used. The reason this is done is to analyse if the spikes are predictable before the spikes happen using the analysis preset in this project. In this section an example of each of the three different types of peak are displayed which include a single positive spike, a single negative spike, and a collection of spikes or a multiple peak example.

5.3.1 Single Spike Increase

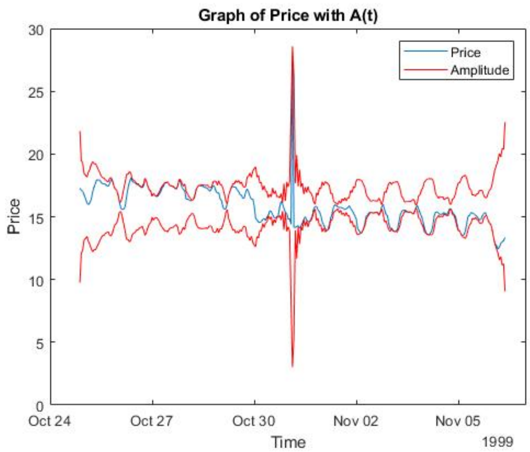


Figure 47: Price of Nord Pool with A(t)

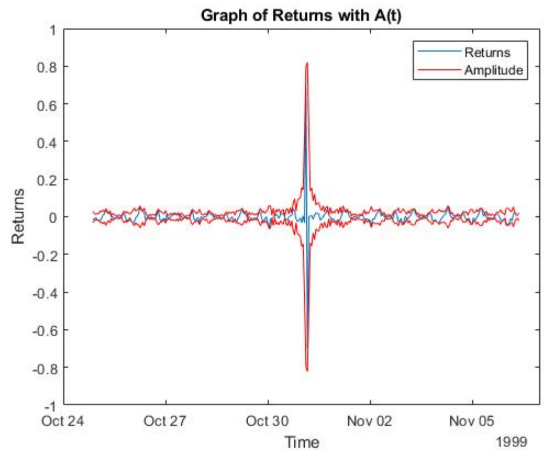


Figure 48: Returns of Nord Pool with A(t)

Figure 47 shows a single upwards spike with its amplitude, this spike from looking just at the data points shows no precursor for the spike happening and also shows no obvious change in the data after the occurrence of the spike. Figure 48 shows the lognormal returns of a single upwards spike and its amplitude, it can be seen it differs from the single upwards spike as it also have a negative spike, this is due to the return data having to reset back to the zero after the initial spike. Again there is no clear precursor for a spike and no change after the spikes occurrence.

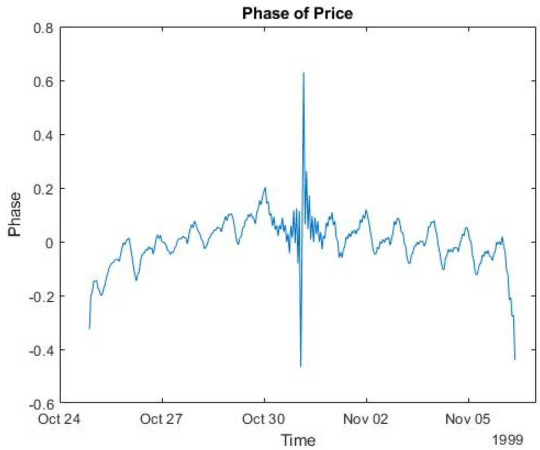


Figure 49: Phase of price

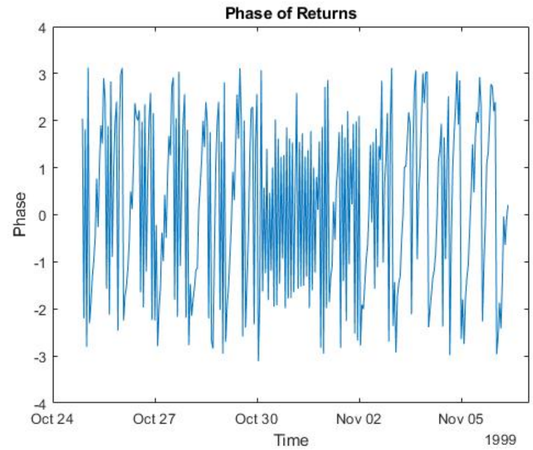


Figure 50: Phase of returns

Figure 49 shows the phase of the price data, in this graph one can see a definite phase slip as there is a positive and negative peak around the area that figure 47 shows a peak,this agrees with the theoretical Gaussian phase slip shown in section 4.5 and Figures 31 and 32. Figure 50 shows the phase of the lognormal return data, here it can be seen that the phase bunches up around the spike indicating a phase slip, but no clear phase slip is seen. Both graphs show the start of their phase slips occurring before the actual peak happens, indicating it may be possible for the Hilbert transform to be used to predict extreme data.

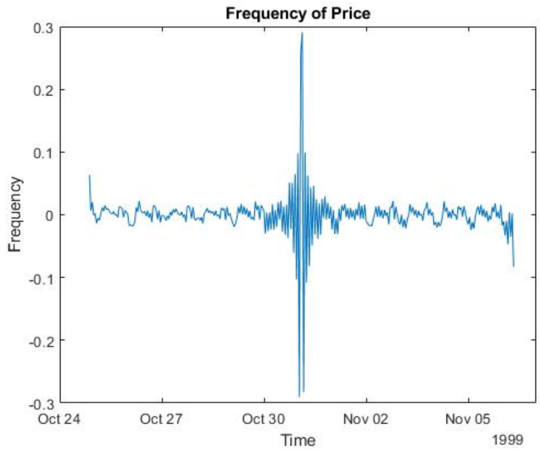


Figure 51: Instantaneous frequency of price

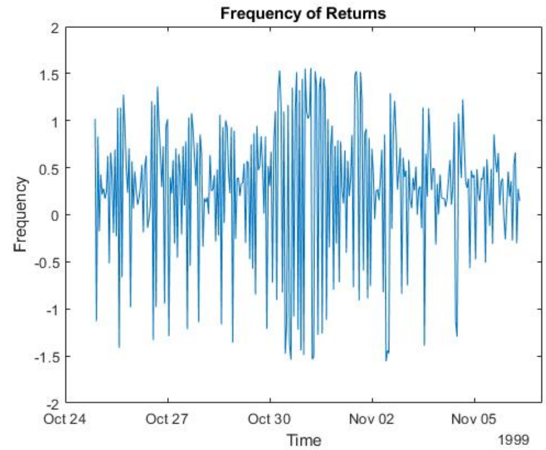


Figure 52: Instantaneous frequency of returns

Figure 51 shows the instantaneous frequency of the price, here we can see a clear spike, Figure 52 shows the instantaneous frequency of the returns, as above the slip is less obvious but in this case clearly observable by the gap in the graph. Again the beginning of each indication of a spike happens just before the spike as shown above.

5.3.2 Period just before Spike increase

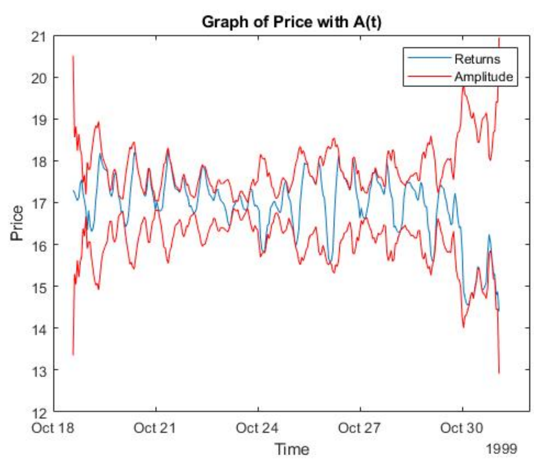


Figure 53: Price of Nord Pool with A(t)

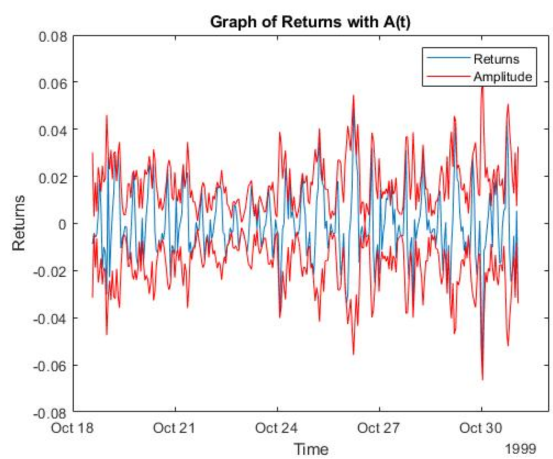


Figure 54: Returns of Nord Pool with A(t)

Figure 53 and 54 shows a model of the rolling Hilbert, both the price and returns are as expected and don't show an indication of the spike to come.

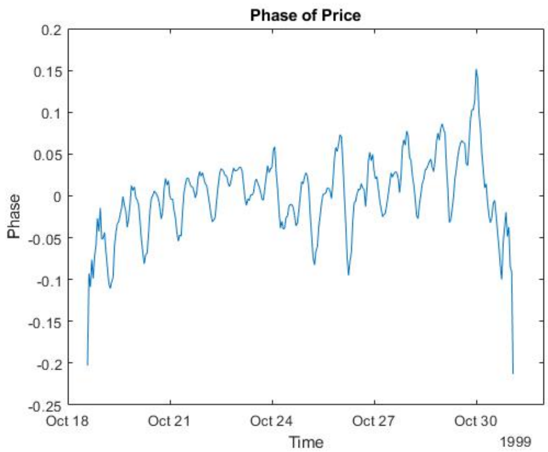


Figure 55: Phase of price

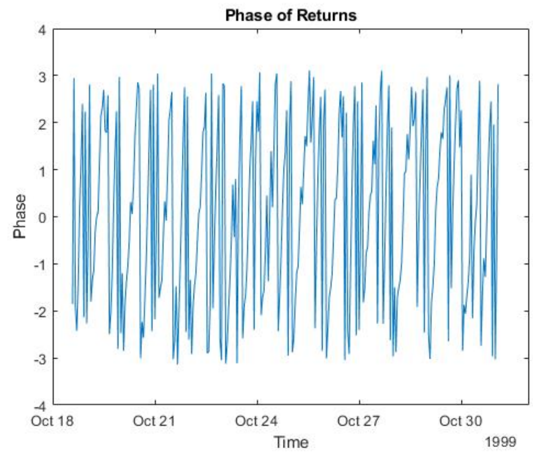


Figure 56: Phase of returns

Figure 55 and 56 show the phase, the phase of returns doesn't change at all during the data set showing no indication that a spike may occur. The phase of price has a large phase drop just before the end of the data set and the occurrence of a spike. This could either show that a spike is about to occur or just be due to the fact that a Hilbert transform is less accurate around the edge data, as seen in section 4.

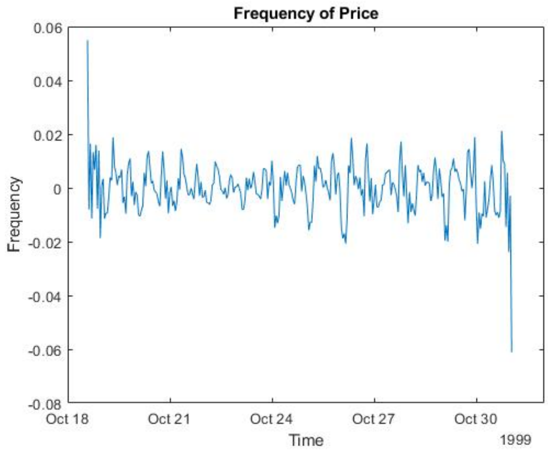


Figure 57: Instantaneous frequency of price

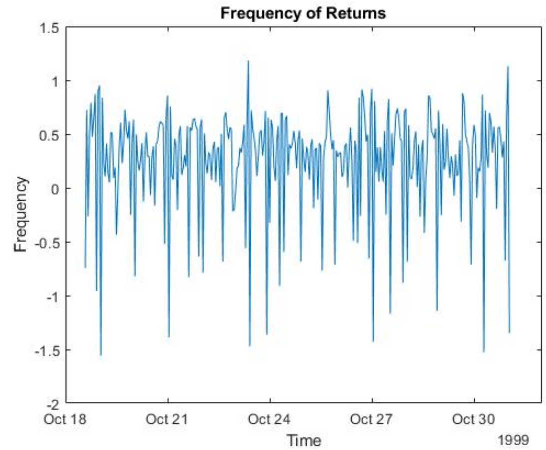


Figure 58: Instantaneous frequency of returns

Figure 57 and 58 show the instantaneous frequency, the frequency of returns doesn't change at all during the data set showing no indication that a spike may occur. The frequency of price has a large frequency drop just before the end of the data set and the occurrence of a spike as above.

5.3.3 Single Spike Decrease

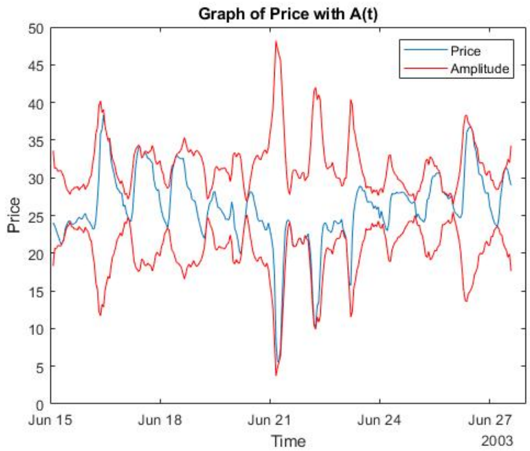


Figure 59: Price of Nord Pool with A(t)

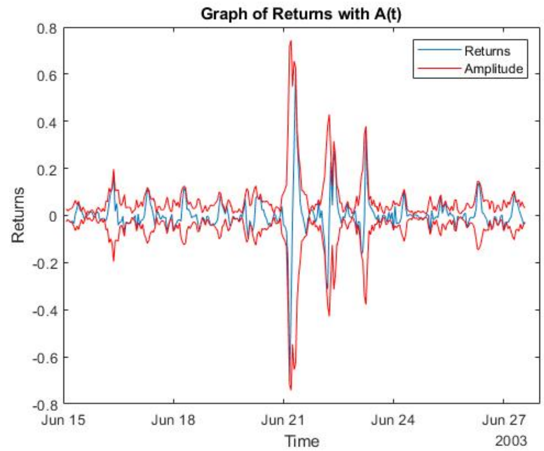


Figure 60: Returns of Nord Pool with A(t)

Figure 59 shows a single downwards spike with its amplitude, this spike from looking just at the data points shows no precursor for the spike happening and also shows no obvious change in the data after the occurrence of the spike. Figure 48 shows the log normal returns of a single downward spike and its amplitude, it can be seen it differs from the single downward spike as it also have a negative spike, this is due to the same reason as in 5.3.1.

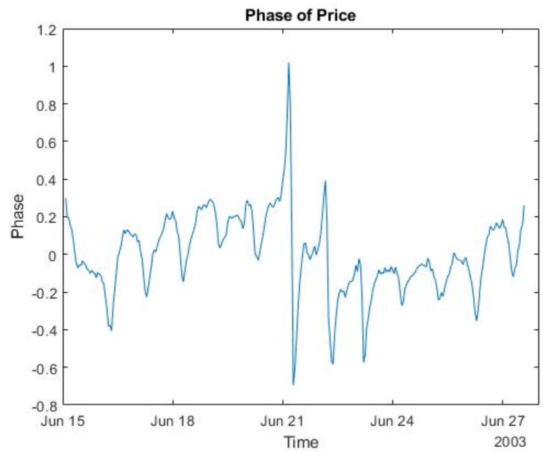


Figure 61: Phase of price

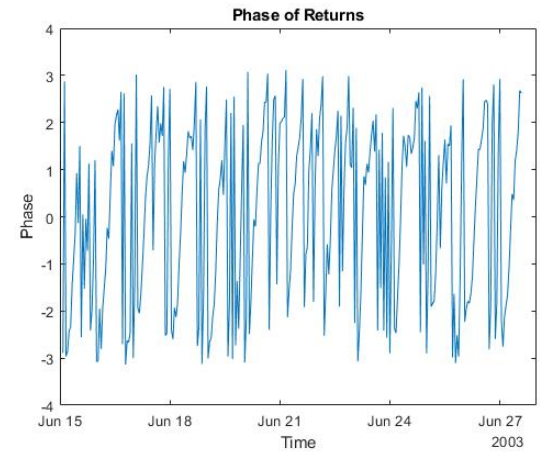


Figure 62: Phase of returns

Figure 61 shows the phase of the price data, in this graph one can see a definite phase slip as there is a positive and negative peak around the area that Figure 59 shows a peak. Figure 62 shows the phase of the lognormal return data, here it can be seen that there is no indication of a phase slip being shown. As above in figure 61 the start of the phase slip occurs before the spike.

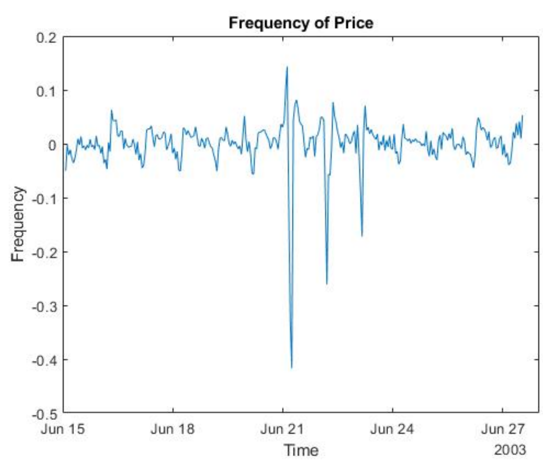


Figure 63: Instantaneous frequency of price

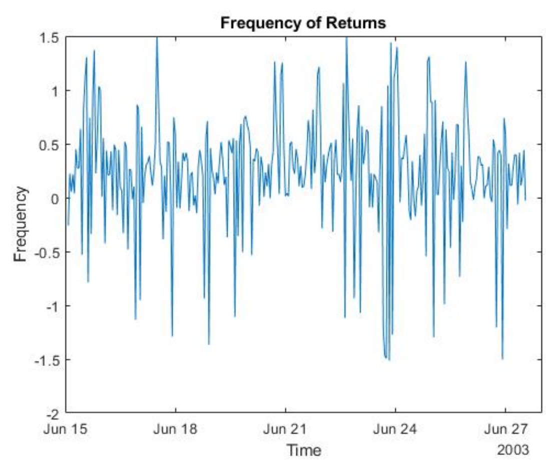


Figure 64: Instantaneous frequency of returns

Figure 63 and 64 show the instantaneous frequency of the price and returns, here the price data again shows a large frequency change around the data point, the return data indicates a spike by the fact that is is completely positive around the spike, but it is not a clear indication of a spike.

5.3.4 Period just before spike decrease

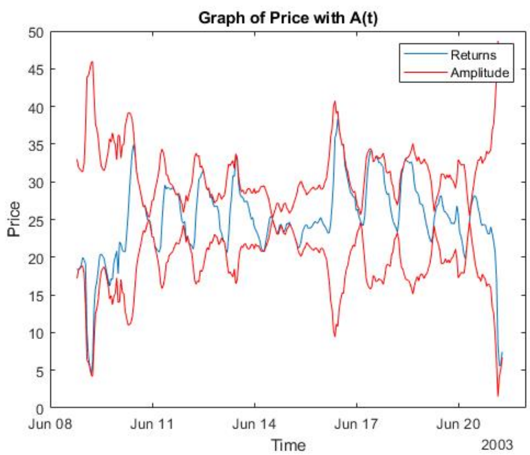


Figure 65: Price of Nord Pool with A(t)

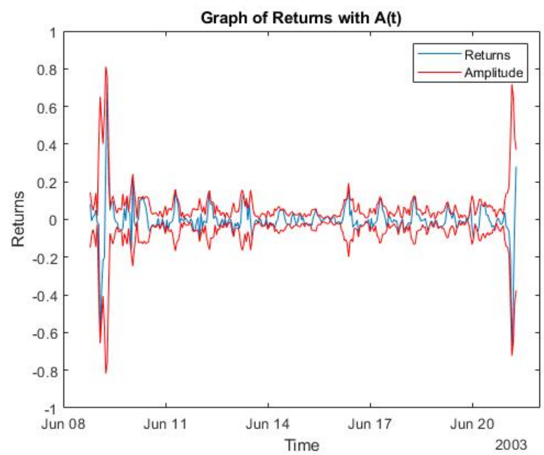


Figure 66: Returns of Nord Pool with A(t)

Figure 65 and 66 shows a model of the rolling Hilbert, both the price and returns are as expected and don't show an indication of the spike to come.

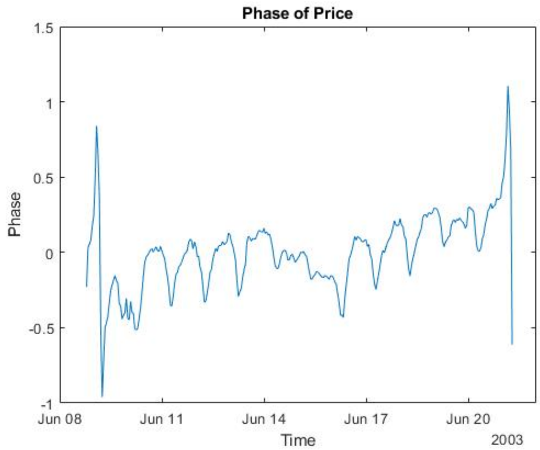


Figure 67: Phase of price

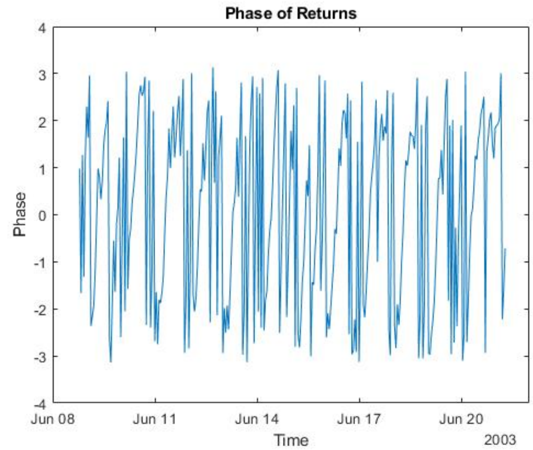


Figure 68: Phase of returns

Figure 67 and 68 show the phase. The phase of returns doesn't change at all during the data set showing no indication that a spike may occur. The phase of price again has a large phase drop just before the end of the data set and the occurrence of a spike.

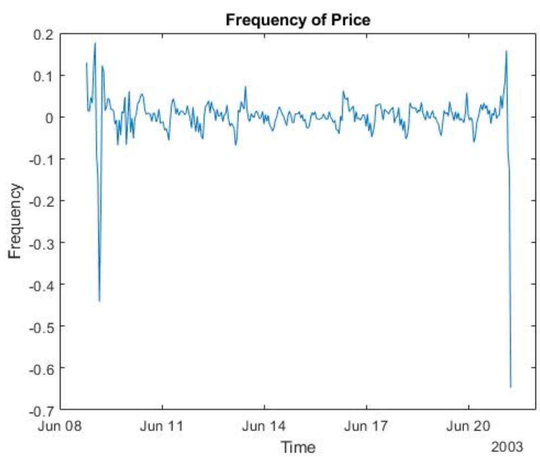


Figure 69: Instantaneous frequency of price

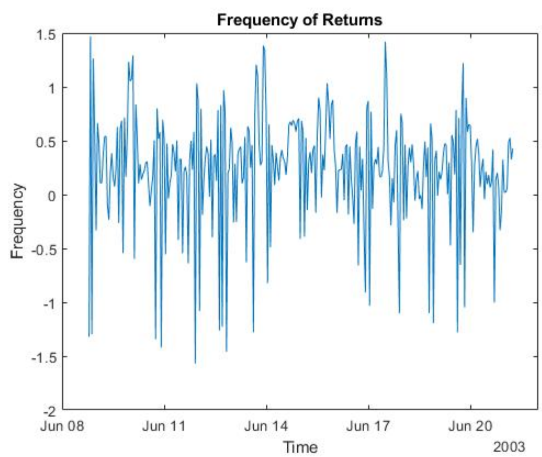


Figure 70: Instantaneous frequency of returns

Figure 69 and 70 show the instantaneous frequency. The frequency of returns doesn't change at all during the data set showing no indication that a spike may occur. The frequency of price has a large frequency drop just before the end of the data set and the occurrence of a spike as above.

5.3.5 Multiple spikes

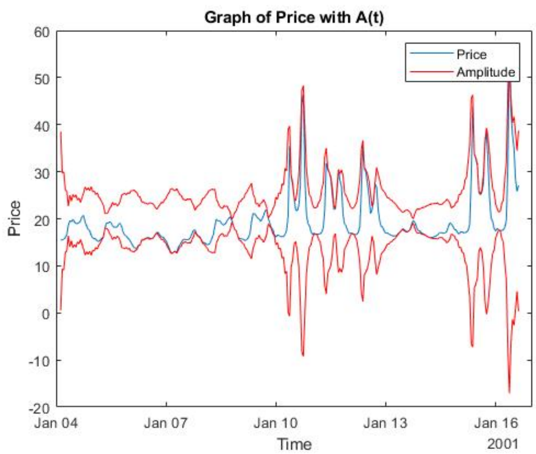


Figure 71: Price of Nord Pool with A(t)

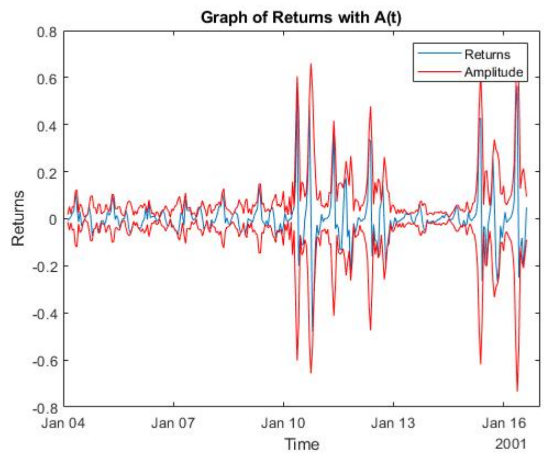


Figure 72: Returns of Nord Pool with A(t)

Figures 71 and 72 shows a multi peak, this is interesting for two reasons firstly, so we can see if multiple phase slips occur when a Hilbert transform is used. The second, is that there is no peak during the weekend, January 14th and 15th in the figure. This is due to the electricity demands on the weekend, and further backs up the theory off amount of electricity available due to demand being a cause for the spikes (section 5.1).

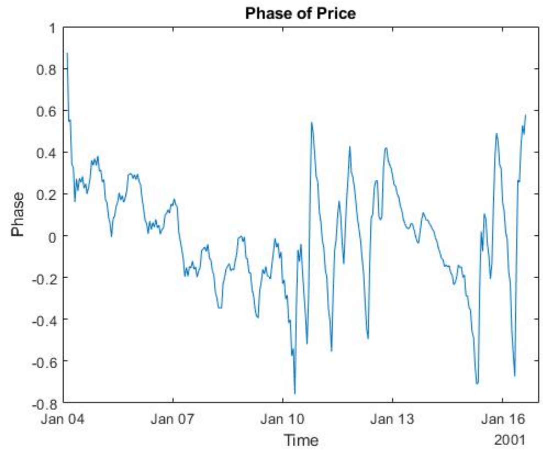


Figure 73: Phase of price

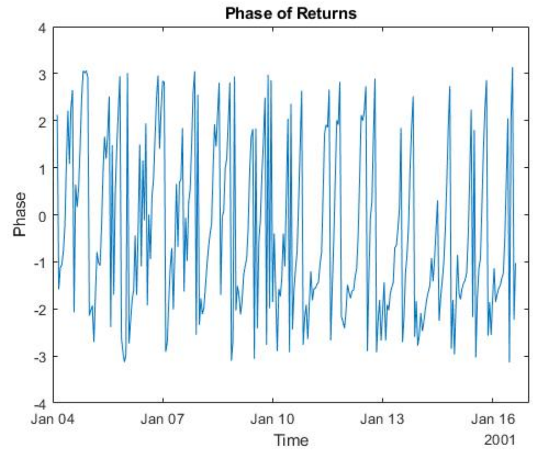


Figure 74: Phase of returns

Figure 73 shows a phase slip for every occurrence of a peak, again preceding the peak, this is interesting as it shows that the phase slip is a perfect indicator for a spike in the data. Figure 74 shows no evidence of a spike.

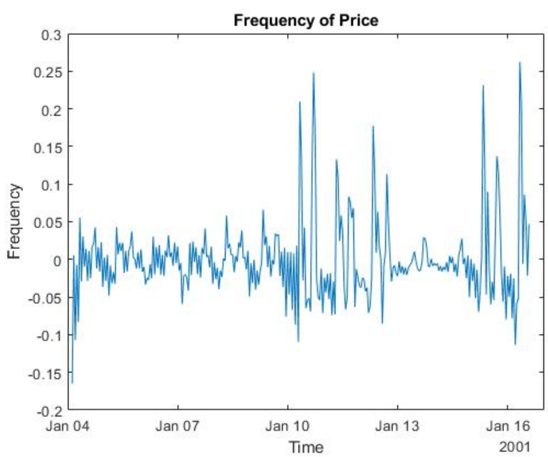


Figure 75: Instantaneous frequency of price

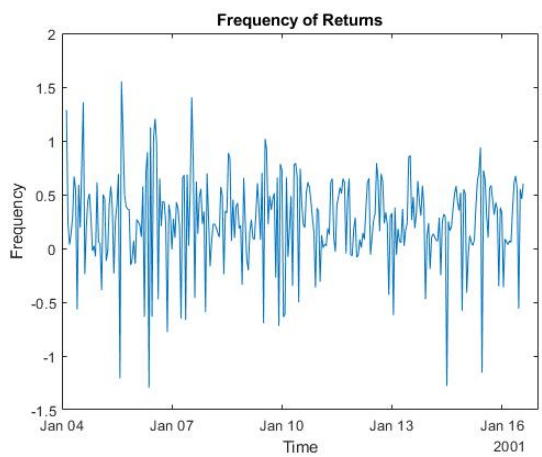


Figure 76: Instantaneous frequency of returns

Figures 75 and 76 agree with the points made on the phase.

5.3.6 Period just before multiple spikes

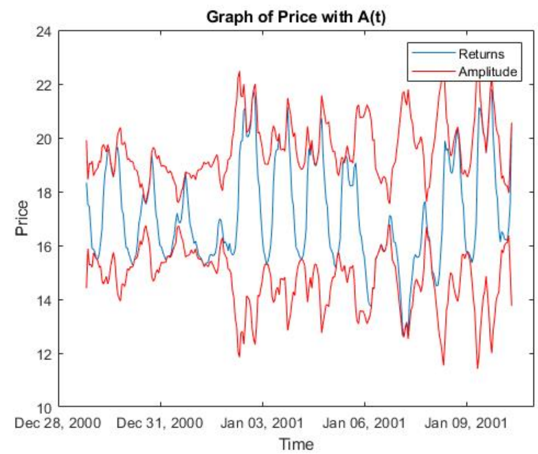


Figure 77: Price of Nord Pool with A(t)

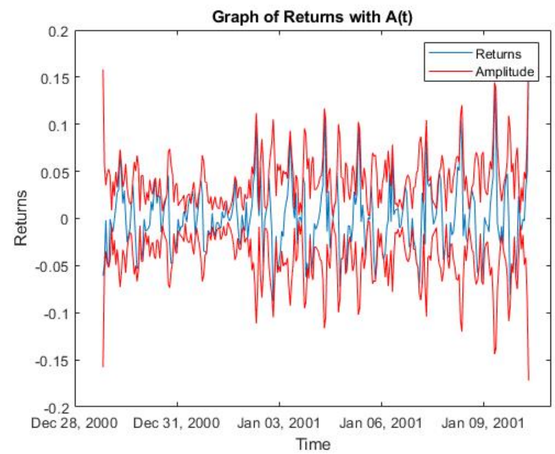


Figure 78: Returns of Nord Pool with A(t)

Figure 77 and 78 shows a model of the rolling Hilbert, both the price and returns are as expected and don't show an indication of the spike to come.

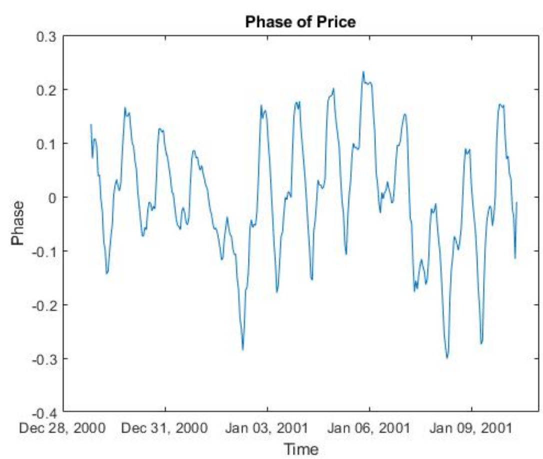


Figure 79: Phase of price

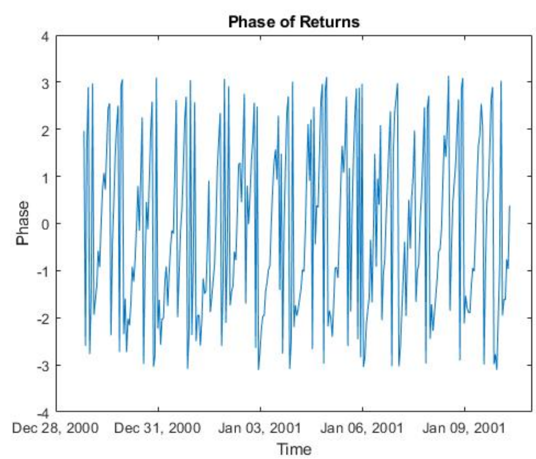


Figure 80: Phase of returns

Figure 79 and 80 show the phase. Here neither the returns or the price show any indication of a phase slip. This is interesting, as in section 5.3.2 and 5.3.4 the edge of the graph has a large phase drop, but in this example there isn't a drop. Therefore more research needs to be done in determining what caused the phase drops above.

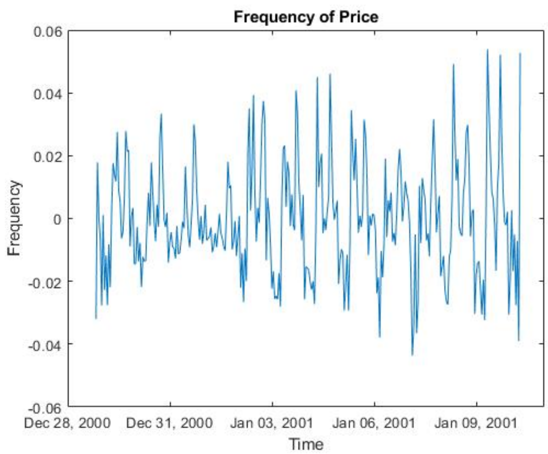


Figure 81: Instantaneous frequency of price

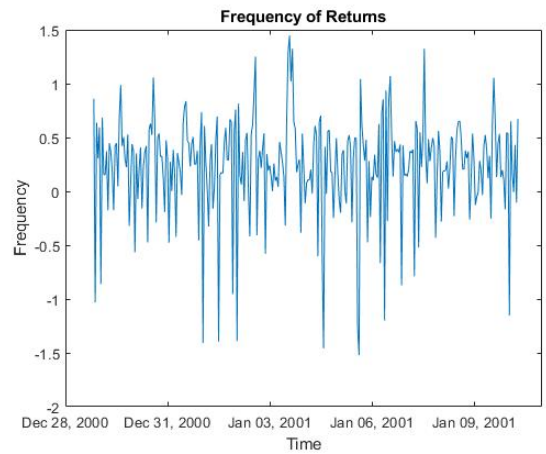


Figure 82: Instantaneous frequency of returns

Figure 81 shows a similar result to Figure 79, Figure 82 has a small frequency slip in the middle of the data due to unknown causes.

5.3.7 No Spike data

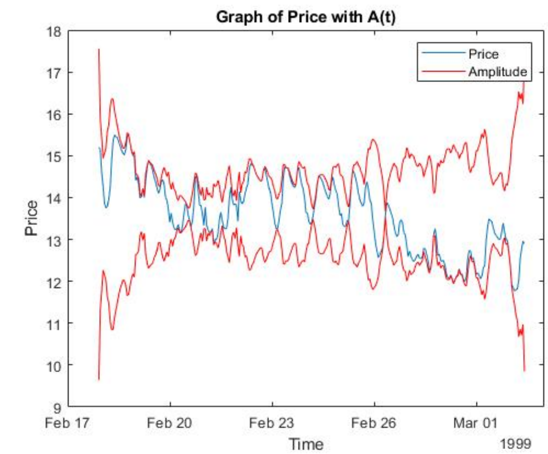


Figure 83: Price of Nord Pool with A(t)

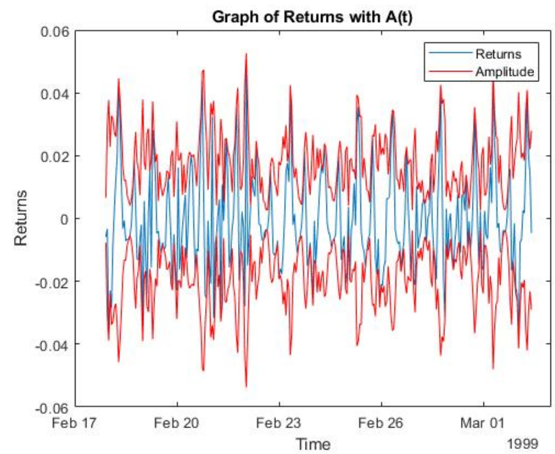


Figure 84: Returns of Nord Pool with A(t)

Figures 83 and 84 show data with no spikes. As expected it follows the normal trend of the average data.

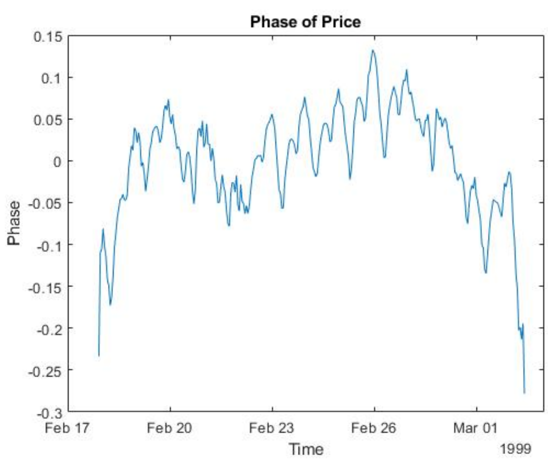


Figure 85: Phase of price

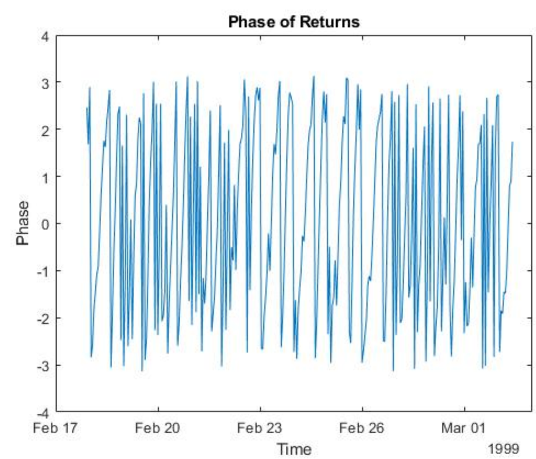


Figure 86: Phase of returns

Figure 85 again shows a sharp drop off in the phase as the end of the graph, meaning it is more likely that the occurrence in the data above is due to the inaccuracy of the Hilbert transform. Figure 86 is as expected.

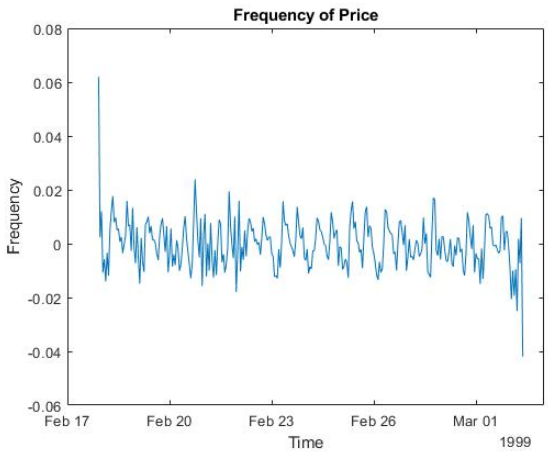


Figure 87: Instantaneous frequency of price

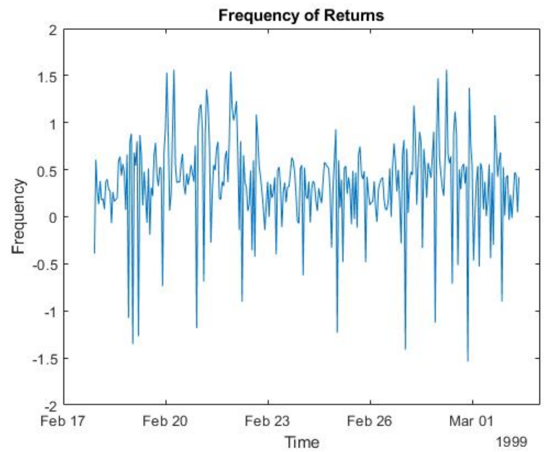


Figure 88: Instantaneous frequency of returns

Figure 87 again shows the sharp drop, Figure 88 is again as expected.

5.4 Discussion

From Figures 47 to 88 it can be seen that the phase slip can be clearly seen when using a Hilbert transform on the price data, it is less obvious, but still sometimes visible when looking at the lognormal returns of the data. From this it is concluded that the lognormal return is not as accurate at showing a phase slip than the normal price. In the single peaks, when looking at the simulated rolling Hilbert's, it can be seen there is a sharp drop at the right hand edge of the graph, this is also shown in the no peak data. This could therefore be said to be due to the Hilbert being inaccurate around the edge data. But this doesn't show up on a multi peak data set, so more research would needed to be completed to say whether these results are anomalies, or if they can be used to predict a spike.

6 Conclusion

The first part of this thesis showed how the instantaneous amplitude and instantaneous frequency could be derived using the analytic signal. In was then shown that theoretical solutions exist for the analytical signal in the simple cases of the signal being a cosine and also the sum of two cosines. In the case of the signal being the sum of two cosines it gave the interesting result that the instantaneous amplitude had a frequency of half the difference of the two cosines used to construct the signal and the instantaneous frequency was the average of the two input frequencies. It also showed the analytic signal could be constructed numerically by representing the input signal as a sum of cosines obtained from a Discrete Fourier Transform. By using simple test signals of a cosine with linearly varying frequency and linearly varying amplitude it was shown that the analytic signal technique did an excellent job in identifying the instantaneous amplitude and instantaneous frequency. Finally, the test signal of a Gaussian was used to represent a price spike which show the slip in phase and jump in instantaneous frequency around the time of the spike. The second part was to analyse the Nord Pool data set across a eight year period from the 1st January 1999 to the 26th January 2007 and see if the occurrence of extreme events (spikes in price) were observable and predictable by the phase when using Hilbert transforms on the data. It was found that spikes produced a phase slip in the data as expected due to initial research into Gaussian equations, these phase slips were also found to occur prior to the spike itself. This was followed up by looking into a rolling data set (modeled by using data prior to a spike), to see if any precursor was found. The data showed the possibility that there was movement in the phase prior to the spike but more analysis needs to be done to conclude whether these were anomalies or in fact phase drops before a spike. Future work could be done on analyzing the data just before the peak to see if the phase drop happening just before a peak occurs due to the peak or other reasons. One theory being the quality of the results of the Hilbert around edge values and this could lead on to using a window smoothing on the Fourier transform and proceeding with the same research.

References

[1] Sengupta, N. Sahidullah, M. Saha, G. (August 2016) "Lung sound classification using cepstral-based statistical features". Computers in Biology and Medicine. 75 (1): 118–129.

[2] Alan, V. Schafer, O. Schafer, R. (1989) Discrete-Time Signal Processing. Prentice Hall.

[3] Gray, J. (2015) The Real and the Complex: A History of Analysis in the 19th Century, Springer International Publishing.

[4] Kendall, M. Hill, A. (1953) The Analysis of Economic Time-Series-Part I: Prices. Journal of the Royal Statistical Society. Series A (General), 116(1), 11-34. doi:10.2307/2980947

[5] Ladiray, D. Palate, J. Mazzi, G. Proietti, T. (2018) Seasonal Adjustment of Daily Data. 16th Conference of the International Association of Official Statisticians (IAOS). OECD Headquarters, Paris, France, 19-21 September 2018

[6] Kress, R. (1989) Linear Integral Equations, New York: Springer-Verlag, p. 91. ISBN 3-540-50616-0.

[7] King, F. (2009) Hilbert Transforms, 2, Cambridge: Cambridge University Press, p. 453, ISBN 978-0-521-51720-1

[8] Bernard,P. (1997) On Instantaneous Amplitude and Phase of Signals. IEEE Transactions on Signal Processing, Institute of Electrical and Electronics Engineers, 45, pp.552 - 560. ffhal-01684083f

[9] Karan, A. (2013) Fourier series And Fourier Transform. IOSR Journal of Mathematics ISSN: 2278-5728. Volume 4, Issue 6 , PP 73-76

[10] Taneja, H.C. (2008) "Chapter 18: Fourier integrals and Fourier transforms", Advanced Engineering Mathematics, Vol. 2, New Delhi, India: I. K. International Pvt Ltd, ISBN 978-8189866563.

[11] Casella, G. Berger, R. (2001) Statistical Inference. 2nd ed. Duxbury.

[12] Melbourne, I. (2019) Fourier Transforms, MA934 Numerical Method, Warwick University.

[13] Kulkarni, S. (2017) The Minimum Reader on Fourier Transforms, Caltech, February 8, 2017.

[14] Strang, G. (1994). Wavelets. American Scientist. 82 (3): 250–255. JSTOR 29775194.

[15] McKilliam, R. Quinn, B. Clarkson, I. Moran, B. (2010) "Frequency Estimation by Phase Unwrapping," IEEE Transactions on Signal Processing, vol. 58, no. 6, pp. 2953-2963.

[16] Knapik, O. (2017) Modeling and forecasting electricity price jumps in the Nord Pool power market.

[17] Breseden, H. (2016) The Nord Pool Market Model.

[18] Kerr, S. (2019) The Nord Pool Funds. The Hedge Fund journal, 34.