Queen Mary
University of London

Specific Learning Difficulties Cover Note

Student ID: 181087641

Advice for assessors and examiners

Guidelines for markers assessing coursework and examinations of students diagnosed with Specific Learning Difficulties (SpLDs) –
As far as the learning outcomes for the module allow, examiners are asked to mark exam scripts sympathetically, ignoring the types of errors that students with SpLDs make and to focus on content and the student's understanding of the subject. Specific learning difficulties such as Attention Deficit Disorders, dyslexia and or dyspraxia may affect student performance in the following ways:

- The candidate's spelling, grammar and punctuation may be less accurate than expected
- The candidate's organisation of ideas may be confused, affecting the overall structure of written work
- The candidate's proof reading may be weak with some errors undetected, particularly homophones and homonyms which can avoid spell checkers

Under examination conditions, these difficulties are likely to be exacerbated. Errors are likely to become more marked towards the end of scripts.

Useful approaches can include:

- Reading the passage quickly for content
- Including positive/constructive comments amongst the feedback so that students can work with specialist study skills tutors on developing new coping strategies
- Using clear English and when correcting; explain what is wrong and give examples
- Using non-red coloured pens for comments/corrections

Colleagues in schools are asked to ensure that students with specific learning difficulties access the support provided by the Disability and Dyslexia Service.

For more information regarding marking guidelines see DDS webpage http://www.dds.qmul.ac.uk/staffinfo/index.html and the Institutional Marking Practices for Dyslexic Students

Disability and Dyslexia Service
Room 3.06, Francis Bancroft Building
Queen Mary, University of London
Mile End Road
London
E1 4NS

T: +44 (0) 20 7882 2756
F: +44 (0) 20 7882 5223
E: dds@qmul.ac.uk
W: www.dds.qmul.ac.uk

Alteration or misuse of this document will result in disciplinary action

Time Series Analysis of Electricity Spot Prices

Supervisor: Dr Wolfram Just

Nirmit Dhanani (ID: 181087641)

School of Mathematical Sciences

August 2021

Abstract

This dissertation aims to take a deep-dive into the electricity market and conduct data analysis on our subject dataset. Additionally, the dissertation is also there to act as a quick primer for the electricity market and to discuss the future going forward. One of the key aims is investigate the correlation between prices over a long period of time, specifically between the hourly prices and daily prices. We do these by calculating the mean, variance and the ACF for our dataset. Our results were that, forecasting electricity prices is an extremely complex endeavour due to the multiple seasonalities present in this commodity. We found that certain hours had low correlations (12AM) such that, prices now have low impact on future prices, whereas for other hours such as 7AM, the there was significant positive correlation, indicating, that the prices now, do have an impact on future prices. Finally, we explored different forecasting methods, such as Facebook's Prophet, LSTM (a type of recurrent neural network), and GARCH. We found that they each had their own merits but LSTM was a viable method for forecasting. Finally, we briefly touched on the impact of renewable energy and it's volatility inducing presence, which will be an key focus for forecasting going forward.

Acknowledgements

I just wanted to make it know that I was deeply touched by how supportive and helpful Dr. Wolfram Just was (and is), so I sincerely wanted to this opportunity to thank you them for their support. This project would not have been possible without them. Furthermore, I would also like to express my gratitude to Dr William Weimin Yoo, as their module teaching was instrumental in helping deepen my understanding of Time Series analysis. I would also like to thank the rest of the Mathematical Sciences faculty, as given the issues I have had to deal with, I very much appreciate how kind and accommodating everyone has been. Finally, I would also like to thank my mother, who has been with me through the thick and thin, and especially supported me in the recent months due to my issues.

Contents

1 Introduction

Have you ever wondered if there was a way to predict stock market prices? Financial time series analysis is one way to tackle this question and it is commonly used across the world of finance and statistics. Time series analysis has been around but it only started gaining in (public) popularity after the attention brought to it by the 2003 Nobel Prize winners: Professors Robert Engle and Clive Granger [1]. Engle developed a new statistical model for time series data called Autoregressive Conditional Heteroskedasticity model (ARCH) [19] which is a way to estimate risk by modelling volatility that resembles the real markets. Another reason for rise in popularity has been the rise of "Quants" the so called Alchemists of Wallstreet [2]. One of the least known but best performing hedge funds is Renaissance Technologies, better known as RenTech which was founded by the mathematician Jim Simons (reference). RenTech's flagship fund, known as the Medallion fund is said to the most successful systematic hedge fund of all time. The fund reported a return of 76% in 2020, and in 2008 when S&P 500 was down -37%, the fund was up 82%. It has annualised return (end of 2019) of 66% and 39% after fees [15]. Quantitative hedge funds like these, market makers like Citadel and high frequency trading firms have all played a hand in the rise of mathematics being applied to the financial markets. Additionally, given the technological advances both in terms of software and infrastructure, has meant that we now have more data than we know what to do with hence this is the perfect environment to use time series analysis to better understand this data in order to help us predict future events.

On the topic of uncertainty, COVID-19 further highlights the volatility of the markets as the financial models weren't able to predict or fully take into account for an pandemic since one hasn't occurred in over a 100 years and data is lacking. Recently, we have seen negative overnight electricity prices in Germany and UK due to the rise of renewable energy production [6]. As you can see, it can be very useful to have data and be able to accurately predict future prices during random extreme events, for example, the supply-side economic shock caused via Suez Canal blockage [45] and the impact on commodities. This is where time series analysis can come into play and help mitigate future price risk by predicting future prices (to an adequate confidence level).
On the topic of volatility, the California Duck Curve [37] is an interesting phenomenon to talk about. This term was coined in 2013 report made by the California Independent System Operator (CAISO) [9], which stated that, as more and more solar panels (renewable energy generators) are added onto the system, it will drive the net demand down significantly during mid-day, which means there will be a significant ramp-up required in energy generation to meet the demand for rush-hour and evening. This is a risk and expensive for the market participants. Additionally, due to the fluctuations in energy demand, this also has a significant impact on electricity prices, especially during Spring and Summer seasons, hence increased volatility.
As a result, we now have an ever-increasing volatility problem to address due

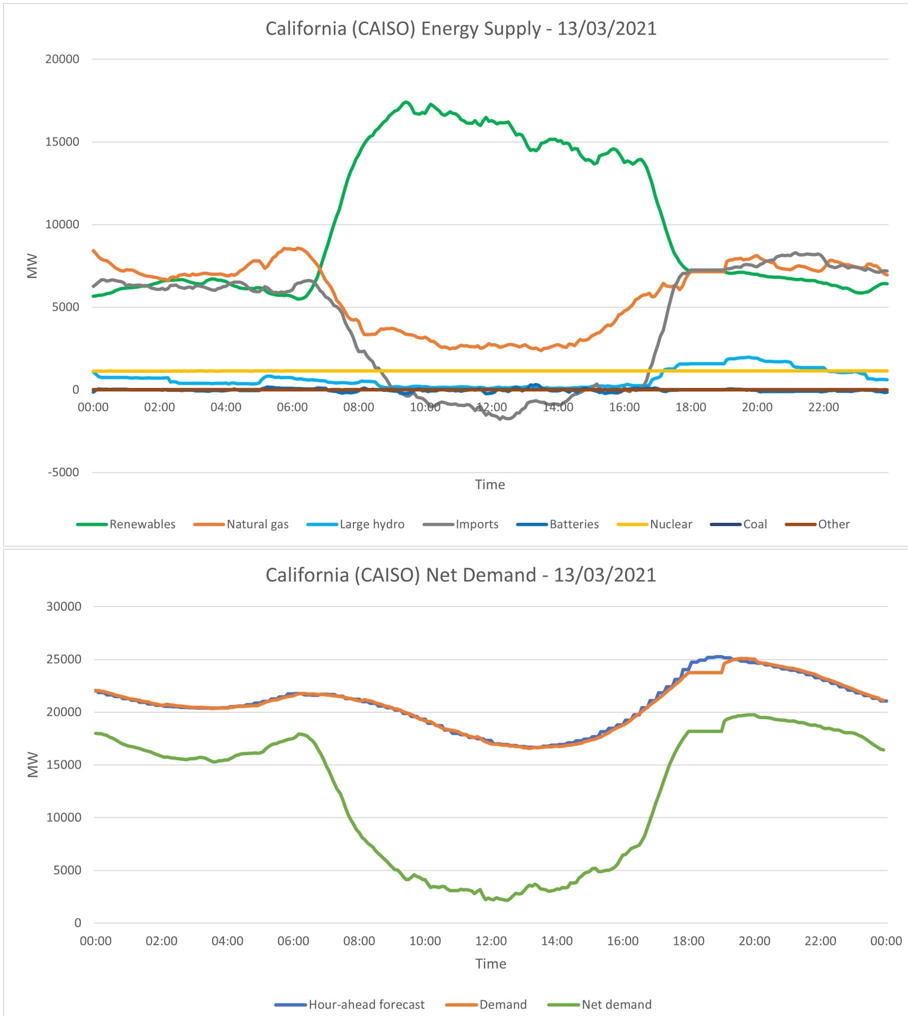

Figure 1: The first plot tells us the total energy supply breakdown on 13/03/2021. The bottom plot tells us the net demand (demand less renewables) and you can see somewhat of a duck shape hence the name. Data Source: CAISO [9].

to the energy generation from renewable sources. California and other tropical places have the duck curve problem, whereas for us in Europe, we face another, which is negative prices [6] due to the efficiency of wind farms (as well as hydroelectricity but this depends on the country). The key challenge here being the storage of electricity and distribution (network). Part of the reason for this has also been due to Governments subsidising renewable energy projects, in order to increase adoption, hence this has led to the inevitable rise in renewable energy at an exponential level.

For this dissertation we will be investigating a large dataset that consists of electricity prices in EUR per MWh (Mega Watts per hour), with a sampling rate of one hour, from 1st of January 1999 (01-01-1999) to 26th of January 2007 (26-01-2007). The market that we obtained this data from is the Nord Pool Spot market, which is the leading (day-ahead) power market in Europe, based in Norway.

To condense this all down, the contributions from this dissertation will be:

1. Comparison of Hourly and Daily: mean, variance, and ACF values to infer interesting results.
2. Exploring different forecasting methods suitable for predicting electricity prices.
3. We will also briefly touch upon the future of electricity price forecasting with context to global changes towards the end.

To give more structure, in this dissertation we shall first gain an overview of the electricity markets so we have some context to the analysis we will conduct later. We will also touch upon the relevant theory concepts behind time series analysis such as stationarity, seasonality, logarithmic returns, autocorrelation function (ACF) and forecasting. The aim of this project is to explore the dataset to identify any interesting autocorrelation patterns but also explore how one can forecast electricity prices using appropriate models and packages such as GARCH and Prophet [21]. First, we shall look at the mean and variance values of the hourly and daily returns of the dataset, then we will conduct similar analysis except with ACF values (plots). Finally, we will briefly explore forecasting to see if we can build suitable forecasting models and how well they fit our dataset. We will not be doing strict model comparison in this project but we will explore some models and do brief comparison using the Root Mean Square Error (RMSE) metric.

I have decided not to discuss or explore the ARIMA model. The reason why I did not include SARIMA or ARIMA model comparisons to LSTM or Prophet models is due to the data. Our dataset has hourly sampling rate with a weekly (and daily to an extent) seasonality which leads to our m value to be 168, which means the computing power required for an output is significantly longer than the latter models mentioned. In order to get rid of this seasonality I would have had to either calculate average daily, weekly or monthly prices in order to make 12 month forecasts. However, consolidating the data, would lead to accuracy loss which I wanted to minimise. Furthermore, I wanted to explore if the LSTM and Prophet models were viable in forecasting over a period of time with granular data, as they seem to be easier to utilise compared to SARIMA (which requires a lot of fine tuning and good statistical knowledge). Plus, SARIMA forecasting has been explored extensively before and I wanted to explore some newer modelling methods (although I do understand LSTM is not exactly new). Some good alternatives are TBATS [17] and more complex machine learning models such as Long short-term memory (LSTM) [25]. As a last note, we will also explore ways to mitigate potential problems with forecasting electricity prices in the future due to ongoing energy problems such the increase in renewable energy uptake.

2 Background

Everyone has most likely heard of the stock market, which we can commonly refer to as the Equities market. Other common markets that exist:

- The Fixed Income (Bond) Market
- Foreign Exchange (FX) Market
- Derivatives Market
- Commodities Market

Now to keep things relevant, I will focus solely on the commodities market, as that is where our electricity market lies. In strict economic terms, a commodity is a homogeneous good or a good that has substantial fungibility, meaning individual units or parts are indistinguishable from each other [29]. As a result, a commodity usually ends being a natural resource like corn, sugar, or coffee.

2.1 Electricity Markets

The European Electricity Market has undergone significant changes since over the past few decades. The key two primary changes have been the deregulation of the market(s) [30] and more recently, reducing greenhouse emissions in aid of the drive to a carbon neutral economy. Electricity used to be produced and sold by state-owned companies, essentially creating a monopolistic market. This was around 1980s era and coincidentally around this period, we also saw a rise in popularity of the free markets idea. This results in historically publicly and owned and operated industries such as Railway become deregulated. The argument being that deregulation would lead to competition, which would give rise to innovation, which can then lead to better long-term benefits for the consumer as well as industry in general. It was this belief that initially gave momentum for the electricity markets to be deregulated. The Nord Pool market in this project was established in 1992 [42], with the British market being established (as well as reformed) in 1990 as a wholesale market. The wholesale market model implemented in Britain in 1990 was abandoned in 1997, with a New Electricity Trading Arrangements (NETA) being implemented in March 2001. The old wholesale market design (prior to 1997) was marketed to other countries such as Ukraine, Brazil, Gujarat (large state in India) and even supported by the World Bank [48].

Electricity is a commodity which stands out amongst it's peers in the sense that it is a commodity that requires some of the more complex infrastructure and trading requirements. This is due to the fact that electricity, as it stands with the current technology, is very difficult to store. This has a significant impact on its market and pricing i.e. the supply and demand must be balanced at all times. Additionally, you need a network to transfer electricity, as a result a global electricity grid or market is currently not feasible.

2.2 The Grid

The process of delivering and pricing the electricity produced is quite complex. It begins from the variety of different electricity generators such as a Power Plant, and more recently renewable energy sources such as Wind Farms. From here, step-up transformers increase the voltage of the generated electricity as it most economically efficient. The explanation being the inverse relationship between voltage and current i.e. high voltage will lead to low current which is great for long distance electricity transportation (as less energy is lost via heat from the low current due to the low resistance). The electricity is then transported across the grid to wherever it has been allocated, when it reaches its destination, step-down transformers then decrease voltage to appropriate levels depending on the users e.g. 230V for homes or offices and 33,000V for large factories [18].

2.3 The Sub-Markets

Given the complexity of electricity and complications it brings, electricity is traded in a variety of different market types. The key here is that generally, the markets represent different time horizons (e.g. daily, monthly or annual). I will briefly walk your through what each of these markets does and their function.

Balancing Market

The balancing market exists to balance the (excess) supply and demand, usually adjusted by a regulator. This is to ensure there is continuous flow of energy(electricity) given the volatility of the market especially nowadays with the introduction of renewable energy sources [46]. The exact mechanisms and nuances vary from market to market (country). Hence, this market exists closer to real-time to accommodate for fluctuations.

Intra-day Market
The Intra-day markets would be classified under the spot market umbrella given they are more instantaneous in nature (real time). The time horizon for this market is the same day (can bid for electricity up to 15 minutes before delivery time) in other words, the purpose of this market is to provide liquidity to cope with any sudden rise in demand during the day.

Day-ahead Market

The Day-ahead market is a market that allows for bidding on electricity prices for one day ahead, hence the name. This is usually the most common type of market within all the electricity markets around the world, as it allows the players (producers and consumers) to submit their bids to ensure electricity delivery in a manageable time frame (up to 24 hours in advance for Nord Pool). As a side note, this market also falls within the spot market umbrella.

Futures Market

The forwards and futures market(s) allow the qualified market players to hedge their positions (bids) against potential losses to minimise risk i.e. risk management. The time horizon within these markets ranges to any time that exceeds the day-ahead market time horizon e.g. week-ahead, month-ahead or year-ahead contracts. As a result of the time horizon, these contracts are usually traded extensively before the expiry (delivery) date. To give a recent example of why this may be important and potential impacts is the Western Texas Intermediate (WTI) Crude Oil went negative for the first time ever at -$37 in April 2020 [35]. The reason for this was that certain contracts (May delivery) were about to expire which meant certain people would be receiving 1000+ physical barrels as per contract, except they did not have any storage capacity, which caused them to actually have to pay people in order to not have to take the delivery. Given the time horizon of these contracts, the prices are based primarily on forecasts (e.g. weather, statistical analysis) as opposed to purely supply and demand.
Something else to take note is that Bilateral trading still does take place, another name for this is over-the-counter (OTC) trading. This is because OTC trading involves private contracts (agreement) between two parties i.e. party A will delivery x MW of electricity to party B on 22nd of July 2022. Outside of the auction based markets such as Nord Pool, OTC trading makes up the remaining amount of electricity trading volume.

Power Pool

As we are dealing with the Nord Pool market, it is important to explain what power pools are. Power pools act as a central point which collect bids by power generators on the price they would be willing to sell their electricity at and retailers at the prices they would like to buy at. The pool collects all of these bids creating a supply curve, then using an estimated demand (consumption) curve, they set the market clearing price (MCP) as seen in Figure 2. Power pools are usually established by the national governments (the UK Power pool is ran by National Grid) as a way to ensure there is enough market competition for the benefit of consumers. Power pools received some criticism over their opaque operations (i.e. how they would estimate the demand and market participants were limited).

Power Exchange

A Power Exchange is an entity which operates an market where electricity is actively traded. The market participants are generally the power generators, distributors and (large) consumers. Overall, an power exchange is more balanced when compared to a power pool, as it operates via a two-sided auction instead of one-sided. Ironically, Nord Pool is a Power Exchange despite what it's name may otherwise suggest.

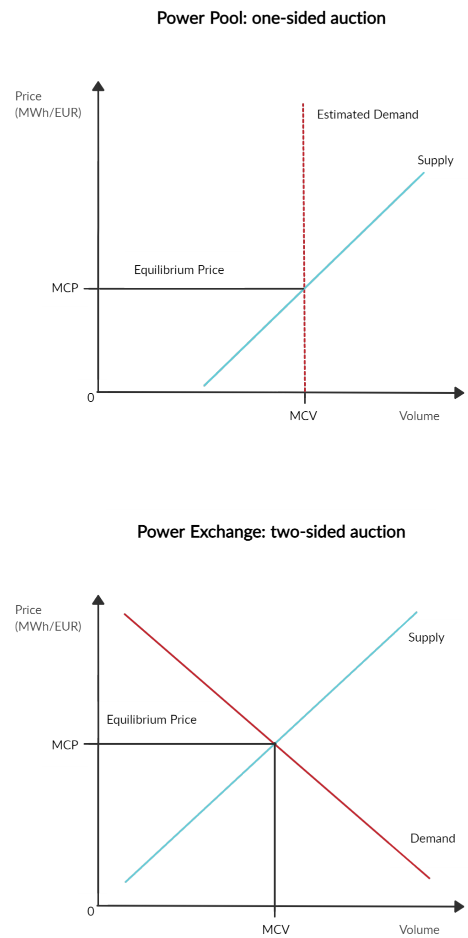

Figure 2: This figure allows us to visualise what the demand and supply curves may look like for different market structures and where the equilibrium price will be. The illustrations were created by me on Creatly.com with references taken from [51]

2.4 Nord Pool

The Nord Pool Spot market is essentially a blind auction where all power generation bids and consumption offers must come in by a deadline. None of the pool market participants know about each others' bids and offers. Then, a centralised market algorithm decides which bids and offers to use and allocates a best fit schedule. The market price is based on merit order. Which means that the consumption offers are ranked in decreasing order, supply is ranked in increasing order which leads to a market equilibrium price that leads to maximum social welfare. The Nord Pool spot market (Elspot) is a day-ahead market itself [23] and the market that our dataset is from.

2.5 Time Series

In simple terms, a time series is an x amount of observations plotted against time t. Some common examples would be, global temperature for the past 50 years or S&P 500 price to date (since inception) as shown below in Figure 3. For this project, we will be looking at a discrete time series which is a time series where observations are gathered at specific, equally placed times i.e. on a daily, monthly, quarterly or yearly basis.

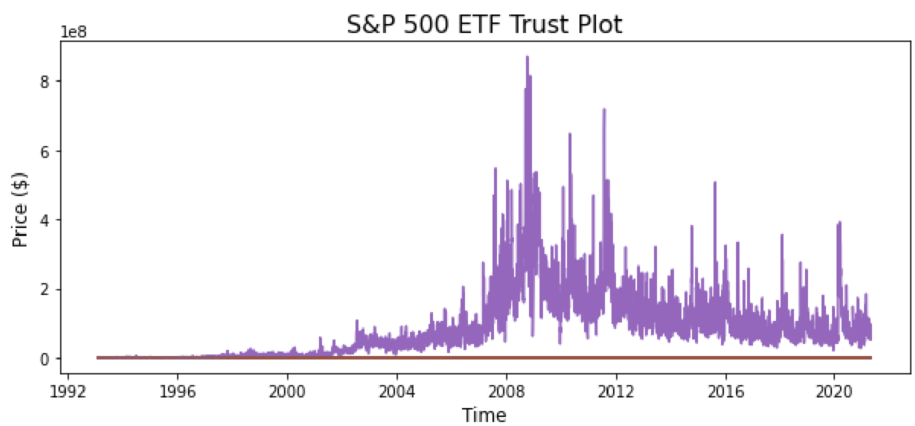

Figure 3: This figure shows the price of S&P 500 Index from it's launch till 01/04/2021. The figure was generated in Python via Yahoo Finance API with data provided by Nasdaq [16].

2.6 Stationarity

Stationarity is a fundamentally important part of time series analysis because most analysis operates under the assumption that the data is or (can be made) stationary. A stationary data set (or time series) is one where the mean and variance are constant over time. The strict definition of stationarity is as follows:

Definition 1. [54]
We define a time series model for the observed data X_t to be a specification of all the joint distribution of the random vectors $(X_1, X_2, ..., X_n)$, $n = 1, 2, ..., \infty$.

For future reference, we can let $X_t = (X_1, X_2, ..., X_n)$ be our time series of length n.
The issue here is that this definition is too strict for practical uses and thus we define a weaker version of stationarity shown below.

Definition 2. [54, 8]
We can say that $\{X_t\}$ is *weakly stationary* if:

1. $\mu(t)$ is independent of t

2. $\gamma(t+h, t)$ is independent of t for each h

One thing to note is that γ refers to the *autocovariance function* which I will explain below. To explain the criteria for weak stationarity, there are two points. First, criterion being that the stochastic process moment does not depend on time t i.e. (μ) must be constant and be independent (i.e. unaffected) of t. Second criterion being, that the Autocovariance function (γ) for t and t plus a lag h is independent of time t. In other words, the autocovariance of X_t is constant.

2.7 Autocorrelation Function

If we have a time series that is stationary, then we can define the **autocorrelation function (ACF)** of X_t at lag h as

$$\rho(h) = Corr((X_{t+h}), X_t) = \frac{\gamma(h)}{\gamma(0)} = \frac{Cov(X_{t+h}, X_t)}{Var(X_t)} \tag{1}$$

Similarly, we can also define the **autocovariance function (ACVF) at time lag** h as

$$\gamma(h) = Cov((X_{t+h}), X_t) \tag{2}$$

Now this is of importance because if one wants see if a series is predictable, we find the covariance which is a measure of linear dependence. Correlation is a function of covariance, in other words, the distinction is that correlation values are standardised (between -1 and 1) when compared to the covariance values.

Hence, comparing two different correlation values becomes easier and more useful, as the values are standardised (the classic apples to oranges argument).

This leads us to the ACF, which tells us the impact of our current dataset to future values. To bring it back to our project, the ACF will tell us the impact current price will have on future prices. This is done via comparing lagged versions of the dataset to itself e.g. correlation between the current price, compared to the price yesterday, 1 month ago and 1 year ago. Where you may get different values such as 0.9 for yesterday indicating strong correlation and 0.1 for the 1 year price indicating low correlation.

When we compute ACF either via RStudio or similar software, we generally tend to use a 95% confidence interval (CI) to judge the significance of values. This CI threshold is also used when trying to building time-series models such as ARIMA.

2.8 Returns and Log Returns

Given we are focusing more financial time series analysis, returns are a very important concept for us. This is because utilising returns, percentage returns and log returns gives us a much clearer picture as opposed to using the normal prices. The key benefit to using returns versus say, normal price for analysis is the added benefit of normalisation, which means instead of comparing values at different scales, we can compare all values via a common scale [43].
Definition 3. Return [10]

$$R_t = \frac{(P_t) - (P_{t-1})}{P_t} \tag{3}$$

Where R_t is the return of an asset at time t, and P_t is the price at time t.
The reason for using log-returns is even more evident in that can help make your dataset more stationary (Figure 4 to Figure 6), which satisfies the dataset criteria required by most time series forecasting models, including artificial intelligence (AI) and machine learning (ML) models (algorithms) [4]. The three benefits being values are time-additive, normally distributed and approximate raw-log equality.

$$R_t = \log(1 + R_t) = \log(\frac{P_t}{P_{t-1}}) = \log(P_t) - \log(P_{t-1}) \tag{4}$$

$$\sum_{i=1} \log(1 + R_t) = \log(P_n) - \log(P_{t_0}) \tag{5}$$

Definition 4. Approximate raw-log equality [10]

$$\log(1 + R) \approx R, R << 1 \tag{6}$$

where << refers to a value being "much less than" as opposed to "less than" we normally see.
Definition 4. is important as the log-returns we will be dealing with for our dataset are very small hence this is of great benefit to us. The benefit being that our return values will be small due to hourly changes (being small), and even less so for daily returns hence this law helps us preserve raw values.
Now I shall briefly analyse our dataset to showcase how our plots changes with standard prices and returns.

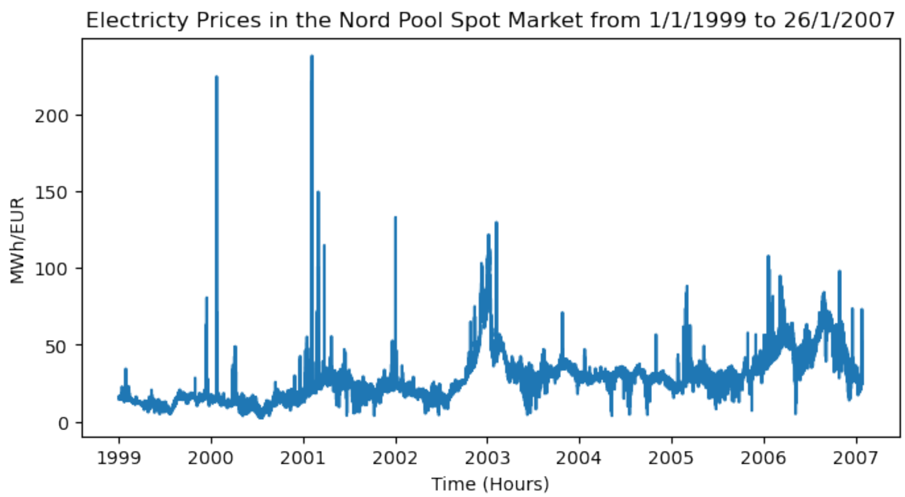

Figure 4: Electricity prices from 1/1/1999 to 26/1/2007 within the Nord Pool Spot Market. Data source: Nord Pool.

We can see from the Figure 4 that there is somewhat an upward trend to the price however there are significant periods of volatility which are not entirely cyclical. Thus we can conclude the data is not stationary. To further note, the extreme price spikes from 2000 to 2003 occurred between January to February, however from 2004 to 2006, this was not the case. We do notice a seasonal pattern however, it may not be clear right now hence I have plotted two different weeks so that we can see patterns during a normal week compared to a volatile week.

As we see in Figure 5 on the non-volatile week, there is a clear daily as well as weekly seasonality present. For example, Monday to Friday has high prices due to higher demand but then as the weekend approaches, prices fall due to low demand (as people might be pursuing manual activities and offices will be empty). From the volatile week, we can clearly notice that it was an outlier period as prices settled relatively quickly. A common reason might be an especially cold winter day.

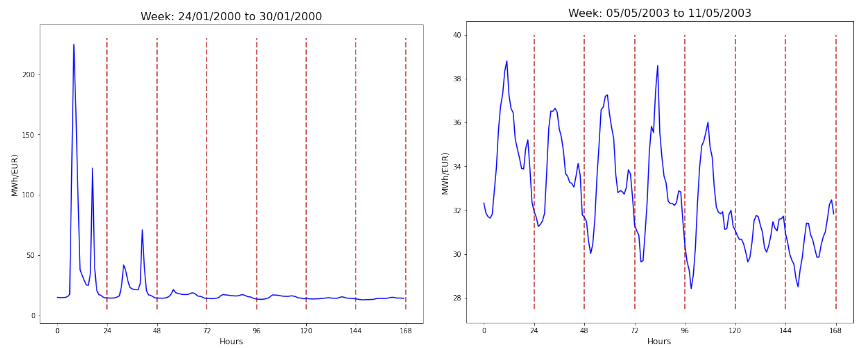

Figure 5: Two randomly selected weeks showcasing price fluctuations from an abnormal to a normal period. The left plot represents week beginning 24/1/2000 and right represents week beginning 5/5/2003.

In order to make our data stationary, so that we can calculate the logarithmic returns, either via Excel, RStudio or Python. For this, part I utilised RStudio for speed and efficiency purposes.

In Figure 6 we can see our transformed data (logarithmic returns). We can see that there are clear periods of volatility or non-stationary behaviour, however, our data now seems to be stationary, compared to the previous state. Ideally, you prefer to have more cyclical and simple seasonal patterns, as can be seen in Figure 5 for week beginning 5th May 2003. However, in our case there is still a lot of volatility which classic ARMA or ARIMA models may find difficult to adapt to. Conveniently, this is where Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models come in to play, in order to help us model our volatility accurately.

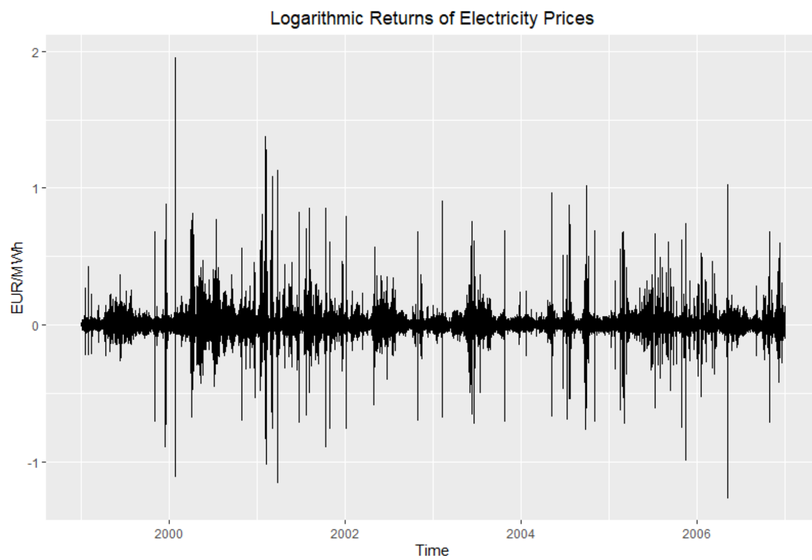

Figure 6: Logarithmic returns on hourly prices from 1/1/1999 to 26/1/2007 within the Nord Pool Spot Market. Data source: Nord Pool.

2.9 GARCH Model

The GARCH model is a variation of the ARCH (autoregressive conditional heteroscedasticity) model. To explain these words, the autoregressive part refers to the AR model, where the values in a time series are affected by some previous period. Heteroscedasticity refers to our volatility and the conditional part refers to the fact that our volatility (variance) is not constant or equally distributed. Both of these models have become important to time series analysis, especially more so in a financial context. The models were developed and first coined by economist Robert. F Engle in 1982 [19] who then went to win the Nobel Memorial Peace Prize for Economics in 2003 jointly with British economist Clive Granger for their work on these models and processes [20].
Another interesting thing to note is that we can combine any ARMA models with any (G)ARCH models. We also include ARCH models within GARCH models e.g. a GARCH(1,0) model is an ARCH (1) model. Let us first define an ARCH(p) model and then we will define the GARCH(p,q) model.

We let a_t be our ARCH(p) process and ϵ_t be the Gaussian white noise with unit variance. Now we have

ARCH Model [55, 19]

$$a_t = \sigma_t \epsilon_t \tag{7}$$

$$\sigma_t = \sqrt{\omega + \sum_{i=1}^{p} \alpha_i a_{t-i}^2} \tag{8}$$

GARCH Model [55, 19]

$$a_t = \sigma_t \epsilon_t \tag{9}$$

$$\sigma_t = \sqrt{\omega + \sum_{i=1}^{p} \alpha_i a_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2} \tag{10}$$

What the ARCH model says, is that the value of time series today, is affected by (or is a function of) the random error today and the time series value in a previous time period. The GARCH model is the same except that for the value of our time series today, is also affected by the volatility of the time series at a previous time period. The reason for developing the GARCH model, was due to the "bursty" nature of the ARCH model, whereas the GARCH model modelled volatility so that it lingered for a while (think of volatility in ARCH as a steep slope versus a slightly more gradual slope in the GARCH model).

To summarise, the GARCH model is a good fit if you want to effectively analyse volatile datasets or commodities such as electricity or metals. This is important nowadays given the increased volatility in electricity prices due to the rise of renewable energy [44], as we discussed earlier (California Duck Curve).

2.10 Facebook Prophet

Facebook Prophet, know nowadays as just Prophet, was created by the core data science team at Facebook in order help everyone forecast at scale effectively. Prophet is an open source forecasting tool available to everyone in Python [21] and R [22]. It was written using Stan [12] which means it does analyse and fit models quickly due to the code. Due to the large quantity available at Facebook, there was a large demand for forecasts to be done. However, as we know, data can be complex to work with and the average person (even at Facebook) will not have the knowledge to create adequate time series models for forecasting purposes. For example, some data may have multiple seasonalities which is difficult to account for, and there may even be special days to consider depending on the country. Hence arises the question: "How can we enable everyone to be able to forecast with a certain degree of accuracy, to meet this forecast demand?". Prophet i.e. partial automation was the answer and this is what lead Taylor and Letham to develop Prophet at Facebook. As a side note, Robert Hyndman who created the R forecast library [27], also helped build the foundations of this tool.
As we can see (Figure 7), Facebook utilises an analyst-in-the-loop model. This framework assumes that the analyst or person, has no statistics background, hence the automated section. However on the flip side, the person can use their

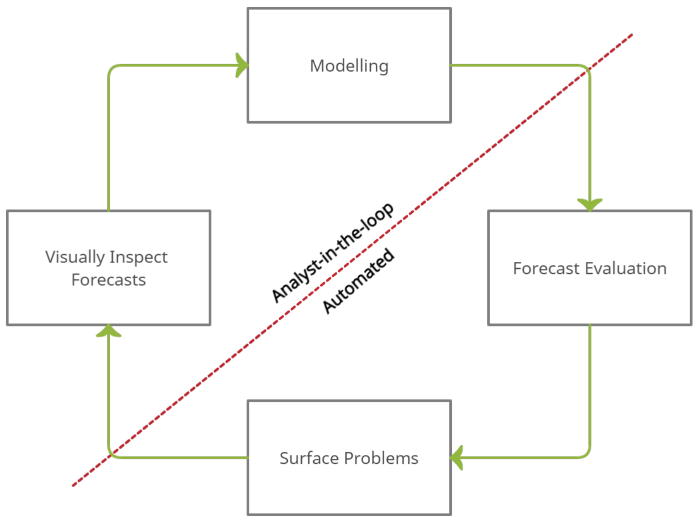

Figure 7: The analyst-in-the-loop model for forecasting at scale effectively. Recreated on Creatly.com but referenced from [47].

knowledge, expertise and intuition to input changes in order to tweak the model to their desired outcome.

The model used by Prophet is a simple decomposable time series model, more specifically, it is based on the generalised additive model (GAM) [24]. Which means that each component is additive but they may not always be linear. The model is as follows:

$$y(t) = b(t) + s(t) + h(t) + \epsilon_t \tag{11}$$

I will break down each components further by explaining what they represent.

For this equation, b_t represents the piecewise defined linear trend in the time series. One thing to note about Prophet is that it implements assumes growth rate is constant when the trend is linear and growth rate decreasing with time t in a non-linear model. This has more to do with population growth, capacity and saturation of markets i.e. a human element. The exact models can be found in their white paper, we will not be discussing them in detail in this project.

s_t represents the seasonality which is how frequently something (or event) occurs in regular cycles within the time series e.g. on an hourly, daily, or weekly basis. The seasonality model in the Prophet library relies on Fourier to provide a reliable and flexible model that can account for multiple seasonalities within one dataset (as is the case with Electricity).

h_t represents the holidays. This is a very unique feature as it given the additive nature of the model, one can add in particular holiday dates which can help capture and forecast that trend as opposed to ignoring it. Dummy variables are used to take into account these particular holiday and event inputs from the user.

ϵ_t represents the i.i.d (white) noise. The unique thing about Prophet model is that ϵ_t is widened as much as possible i.e. it is present whereas the other components such as h_t may not always be present.

2.11 Long Short Term Memory

Given the complexity of this particular topic, I will be keeping this as brief and short as possible with the help of diagrams.

Long-Term Short Memory (LSTM) is a derivative and a more evolved version of a Recurrent Neural Network (RNN). RNNs are a type of advanced (machine learning) algorithm developed to handle sequential data, some common uses include speech recognition, financial markets data and DNA sequence analysis. Compared to vanilla neural networks, RNNs retain past information. However, the drawback here is that classic RNNs are unable to retain long term information (time series), meaning that RNNs can forget what they have seen in the past in longer sequences thus having short-term memory [25].

The issue lies with gradients, which you can think of as an input that determines how much the output of a function or cell changes. Classic RNNs face the issue of exploding and vanishing gradients, both of which we do not want. Exploding gradient leads to higher importance being assigned to certain weights which can impact your model's training. Vanishing gradients are issue I explained above, essentially the gradients become so small, that they have no impact on learning. This is where LSTMs come into play with the sigmoid function which takes the gradient inputs and squishes them so that they stay between 0 and 1.
LSTMs contain 3 gates, which are gatekeepers of information i.e. which information to keep and reject. We have the input, forget and output gates. Let each "X" stand for multiplication, with the closest sigmoid (σ) function to the subject "X", be the gates. The forget gate is the first "X" function we see on the cell state line (hence the sigmoid function leading into it also forms a part of the gate) and the input gate is the second "X" function. Finally, the output gate is the the third "X" function (counting from left to right).

The forget gate decides which information should be discarded or kept. The current input and previous hidden state information are combined and put through the first sigmoid function. Values come out between 0 and 1, the closer to 0, the more you forget, the closer to 1, the more information you retain. This value now gets multiplied by the previous cell state value and thus we have a forget

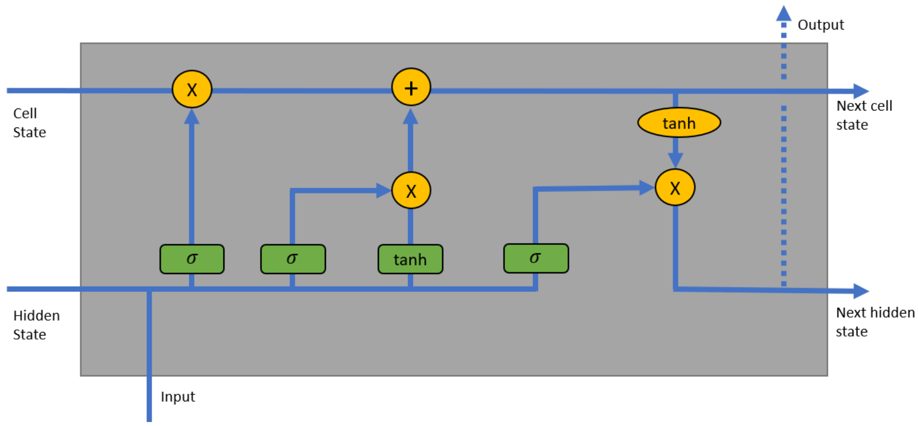

Figure 8: LSTM cell reproduced in Microsoft PowerPoint with reference from Christopher Olah's blog [41].

factor value.

We follow a similar process for the input gate, except now, we also multiply the output of the *tanh* function, with the sigmoid output, all of which assigns a weighting to the importance of information. The *tanh* function exists to regulate the network, as values are squished between -1 and 1. Now we add the outputs from the input gate and forget gate to get a new cell state value.

For the output gate, we take out new cell state value and multiply similar to the two previous gate, by the old hidden state and input values via a sigmoid function. After this, we get our next hidden state value and this process continues on.

LSTMs are ideal for time-series analysis and forecasting given the fact that they can retain long-term information. For this project, we will implement LSTM model in Python with coding guidance taken from a helpful LSTM time-series application blog, written by Jason Brownlee [28]. We will be utilising Python Keras library in order to implement and run the algorithm on our large dataset.

2.12 RMSE

A brief section on why we shall use RMSE to evaluate the effectiveness of the models. Generally, people tend to use either RMSE or MAE (Mean Absolute Error) in order to judge the performance of models in academia. There have been many debates as to which one is better [13, 53].

RMSE is defined [38] as follows:

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_t - y_t)^2}{n}} \tag{12}$$

where y_t represents our observed (actual) values at time period t, $\hat{y}_t$ our forecast (prediction) values at time period t and, n is our number of forecast values.
The reasoning for choosing RMSE is as follows, as mentioned in [13]:

1. RMSE avoids using absolute values (compared to MAE) which is extremely undesirable for mathematical computations.

2. RMSE is calculated from our sample average whereas MAE uses the median. Hence if there is skewness, (as there is for our hourly returns), MAE will be affected.

3. RMSE gives a larger weighting to larger errors i.e. is more punishing compared other commonly used metrics such as MAPE (Mean Absolute Percentage Error) and MAE.

As we would like to be as accurate as possible, RMSE seems to be the ideal error valuation metric to use.

On a final note, we will define spikes as any value greater than 2σ away from the average value of that particular plot. Although in general, it is any value that stands out in contrast to the rest of the data.

3 Analysis

3.1 Mean

Here we analyse the mean and variance values to see if we notice any interesting patterns. To define daily returns, it is comparing the average price (return) at 12AM, to the price at 12AM on the previous day e.g. the price at 12AM on 5/1/1997 is compared to the price at 12AM on 4/1/1997 and this might give us a return of 0.005% (as an example). I have created a separate table with all the daily prices as well as return values, which will allow me to compute and compare the ACF plots with ease.

Mean					
Time	Hourly	Daily	Time	Hourly	Daily
12am	-0.01342	0.000179	12pm	-0.01743	0.000212
1am	-0.03579	0.000170	1pm	-0.01329	0.000201
2am	-0.02568	0.000170	2pm	-0.01061	0.000194
3am	-0.01788	0.000174	3pm	-0.00547	0.000192
4am	0.00415	0.000180	4pm	0.00586	0.000199
5am	0.04123	0.000196	5pm	0.01827	0.000206
6am	0.04870	0.000210	6pm	-0.00015	0.000205
7am	0.06142	0.000237	7pm	-0.01634	0.000192
8am	0.04965	0.000245	8pm	-0.01630	0.000184
9am	0.00467	0.000235	9pm	-0.00428	0.000176
10am	0.00218	0.000226	10pm	-0.01391	0.000171
11am	-0.00782	0.000218	11pm	-0.03762	0.000164

The mean of the hourly dataset is 0.00000669257, we can essentially take this to be **0** given how extremely small the value is. The mean for daily dataset is 0.00019735571 which is larger than the hourly mean value. An interesting observation is just how much larger the individual mean values for the hourly dataset are compared to the complete dataset mean. Whereas, for the daily returns dataset, the mean is close the daily return values. The highest hourly return value is 0.06142 at 7am, whereas the highest daily return value is 0.00245 at 8am, however 7am is second largest at 0.00237. The results here logically make sense since the early morning hours will tend to have an overall increase in demand, which can lead to price increases i.e. rush hour. Similarly, the two lowest respective values for the hourly and daily datasets are -0.03579 at 1am and 0.000164 at 11pm.

The results here are within normal expectations, for example at 1am most people would be sleep and energy usage would be low and likewise with 11pm where most people would be either asleep or about to go to sleep. However, one clear thing to note is that the hourly value is negative which means, on average, the prices fall at 2am, whereas with the lowest daily return value is still positive at 11pm, and even at 2am (0.000194). As a whole, the daily return values are all positive. A potential explanation could be of human nature i.e. Nordic population grew between 1999 and 2007 [39], technological advances (increases consumption) and business costs increasing whilst needing to maintain profits [40]. Thus, we see prices increase over time.

To further help illustrate any trend, we can look at Figure 9 below. Here we can see that there is some clear seasonality, as the day goes on, which is to be expected given the average cycle a person goes throughout the entire day. Interestingly, there is already a positive hourly return value at 4am, which then goes on to spike even further during the usual rush hour period (6-10AM). The values then decrease, only to rise back up during the evening rush hour, then declining further. One oddity is that the return values fall sharply after 5pm, only to spike at 9pm, and then continuing to decline. One reason could be that most people return from work to home at 6pm, maybe 7pm, eat, and then finally do some leisure activities around 8-9PM, which then leads to the small increase in the hourly return value (although it is still negative).

Now, if we look at the daily return values, the curve is relatively stable, however it follows the shape of the hourly mean curve to some extent, which goes to show some weak seasonality. From the 'Hourly vs Daily Mean' plot, we can see that as a whole, the daily mean return values are extremely small i.e. close to zero, whereas the hourly return values are more extreme. This is interesting because, if the daily return mean values do not change significantly and stay constant, it can be used as a good reference point to forecast future prices at those specific hours (so we then focus on modelling the volatility i.e. variance).

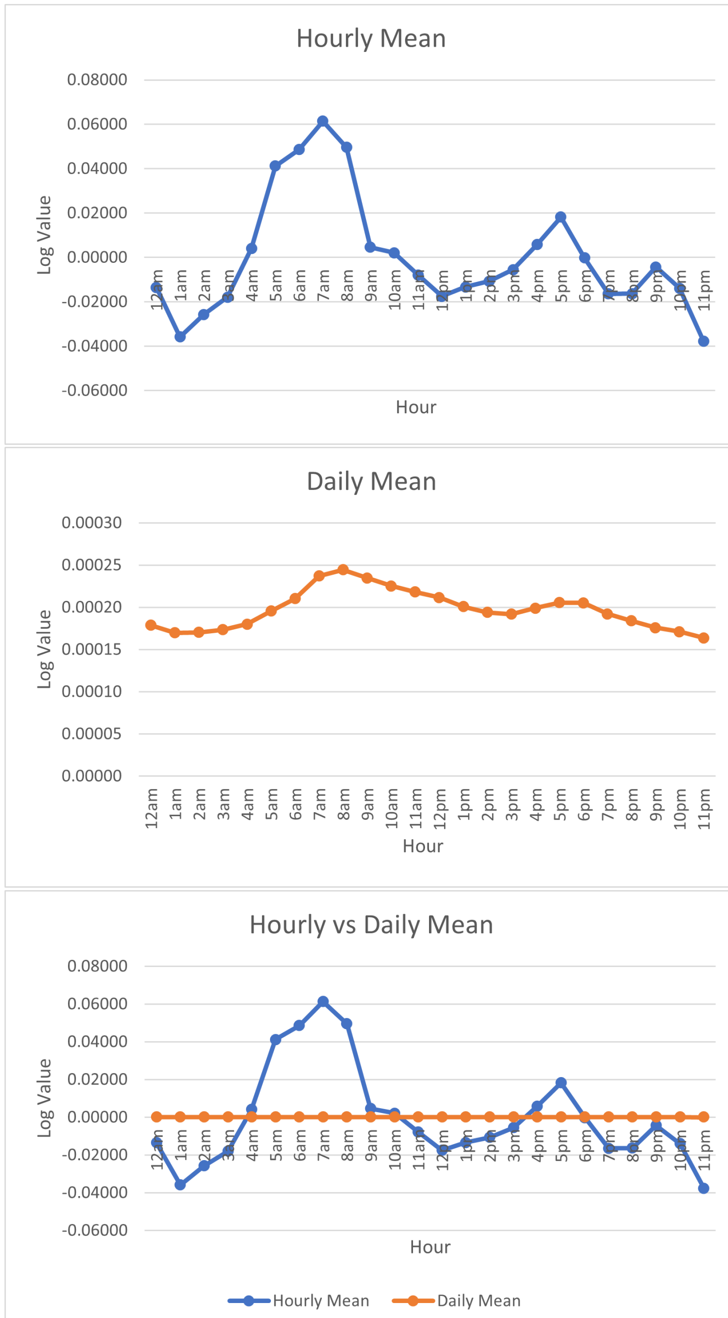

Figure 9: Hourly and Daily mean values plotted separately as well as together.

3.2 Variance

Variance					
Time	Hourly	Daily	Time	Hourly	Daily
12am	0.00321	0.00483	12pm	0.00102	0.01045
1am	0.00293	0.01077	1pm	0.00047	0.01063
2am	0.00301	0.01851	2pm	0.00037	0.01053
3am	0.00333	0.02029	3pm	0.00105	0.01059
4am	0.00272	0.02282	4pm	0.00186	0.01256
5am	0.00492	0.02401	5pm	0.00273	0.01491
6am	0.00522	0.02933	6pm	0.00325	0.01049
7am	0.00935	0.03922	7pm	0.00321	0.00609
8am	0.00766	0.04833	8pm	0.00167	0.00366
9am	0.00364	0.03116	9pm	0.00058	0.00281
10am	0.00198	0.02126	10pm	0.00068	0.00213
11am	0.00186	0.01374	11pm	0.00109	0.00239

Now let us discuss variance, to give some context, variance represents the volatility of prices i.e. how close the return values to the mean. A large variance would mean that the return values frequently deviate away from the mean, whereas a smaller values suggests the return values closer to the mean, with less chances of outlier values being present.

The variance of the hourly returns dataset is 0.00349182865 or 0.00349 to keep it consistent to the table values. The variance of the daily returns dataset is 0.01589. Comparing the two, we can clearly see that the daily returns variance is significantly higher than that of the hourly returns dataset. This is interesting, as looking at the mean values, you may expect the variance of the daily returns dataset to be lower. Taking into account the low variance of the hourly returns dataset, we infer that the price of the previous hour has a significant impact on being able to predict what happens next. In other words, if we are following a particular time period of the day such as night (12am-3am) or rush hour, we can be sure that the values will follow a similar trend, hence less likelihood of outliers and thus low variance. The daily returns variance can be explained by the fact, that what may exactly happen the next day or more (e.g. sudden weather changes), can be very difficult to predict, thus there is a greater chance of outlier values being present, hence the higher variance.

The highest hourly variance value is 0.00935 at 7am, which coincides with the 7am hourly mean value being the highest, and makes sense. For the daily returns dataset, the highest variance value came to 0.04833 at 8am, with the 7am value being second highest at 0.03922. Similar to the mean value analysis, the high variance makes sense given the rush hour period, and the value correlate well with the hourly return dataset. The lowest hourly variance value came to 0.00037 at 2pm, whereas the lowest daily variance came to 0.00213 at 10pm. e.g. we can say with 95% confidence that if the price at 12am on 20/2/2001 was £35.25 MWh, then the price at 12am on 21/2/2001 will be £36.16.

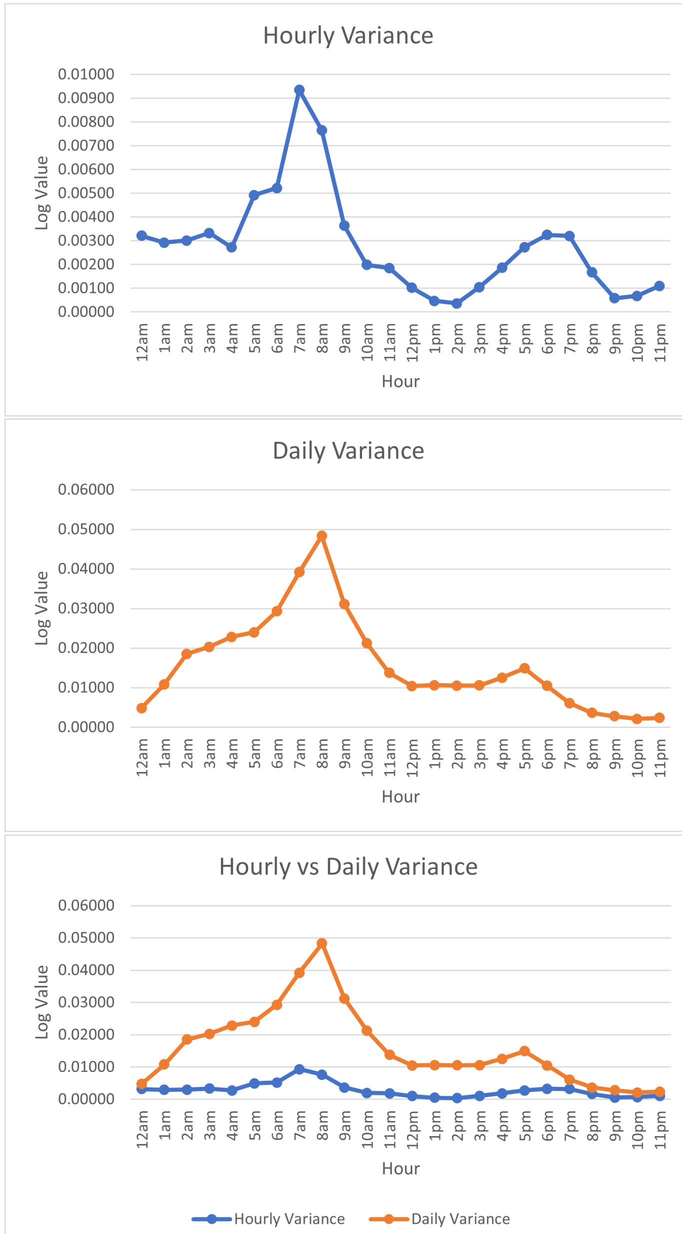

Figure 10: Hourly and Daily variance values plotted separately as well as together.

3.3 Autocorrelation

In this section, we will compare the ACF plots between the hourly and daily returns at each hour of the day. I will also caution that the lag scale may appear deceiving given our data set, for example, a lag of 1 in the below graphs represents 365.25 days. For example, a lag of 0.02 is equivalent to 7 days or one week. Finally, the left hand side of the plots will be the hourly returns and daily returns for the right hand side. I have used lag of 500 for each plot, to try more information, and where required we will also look at smaller lags.

Let us begin with 12AM and move in chronological order for simplicity.

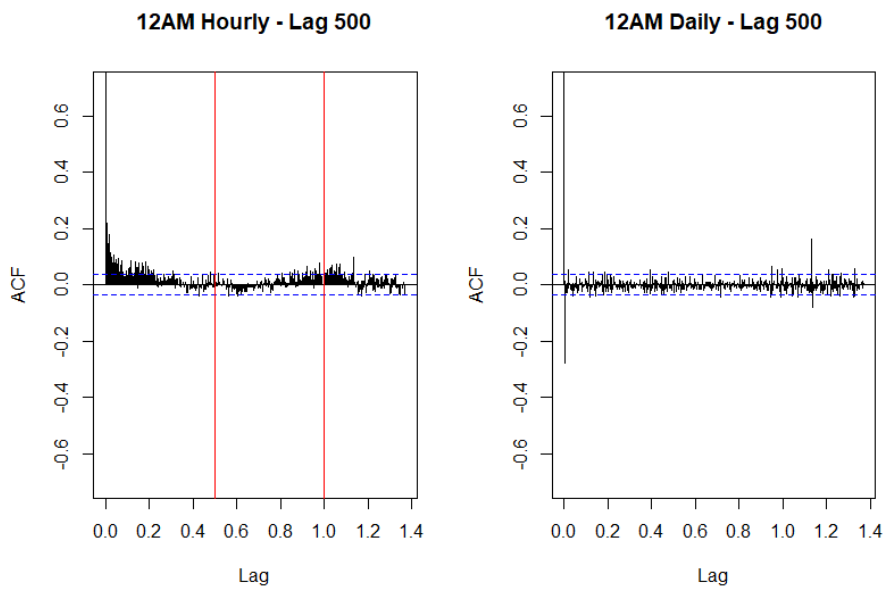

Figure 11: ACF plots for Hourly and Daily returns at 12AM computed using RStudio.

To my point about using lag of 500, if for example we did a lag of 50, for the hourly returns, we would not have noticed anything. However, once we look further, we can clearly see there is some seasonality to the hourly returns (cycle repeats after lag 1 i.e. 1 year). Comparing this to the Daily returns plot, we do not notice any similarities, only that the lags are not nearly as positively correlated as the hourly returns.

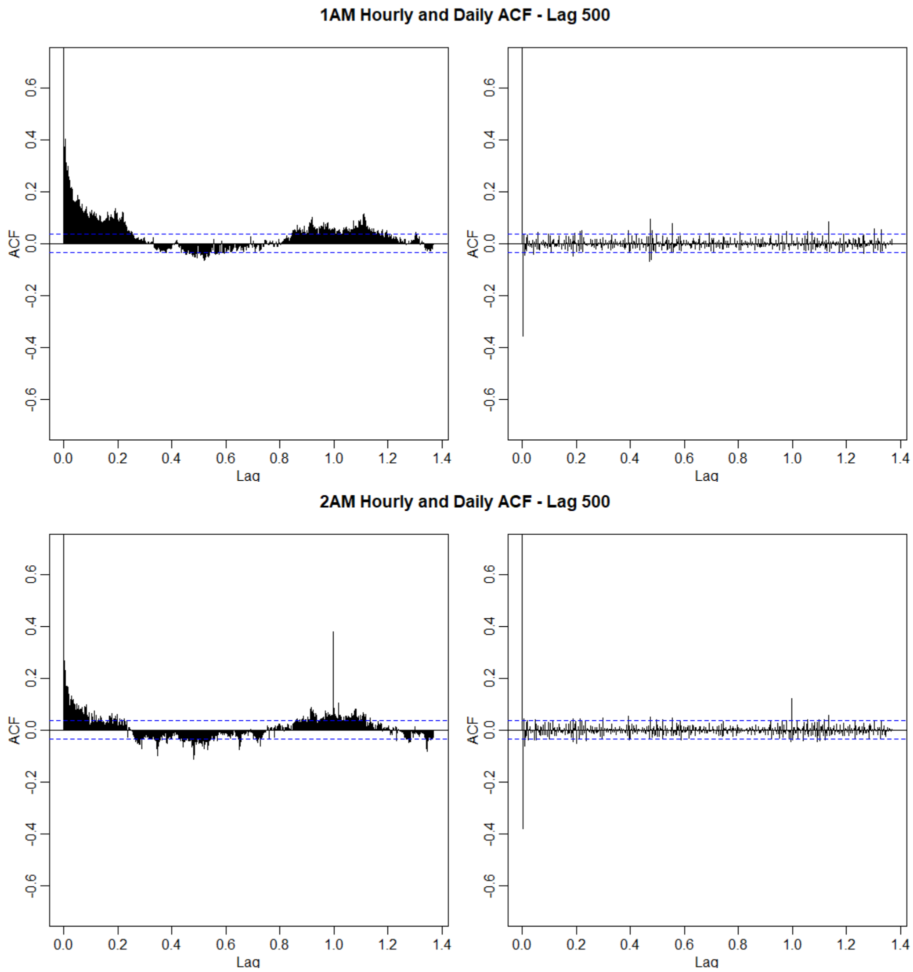

Figure 12: ACF plots for Hourly and Daily returns at 1AM and 2AM respectively.

Similar to the 12am plots, we notice that both the 1am and 2am plots have seasonality i.e. cyclical at the mark of lag 1, which is equivalent to 365 days. One interesting thing to note is that at lag 1 in both the hourly and daily plots for 2am, we see a significant rise in our correlation value, this could potentially be helpful in building long-term forecasts for 2am. We do notice a spike in the 2AM ACF at a lag of 365 (as in Figure 12). The change from positive to negative correlations around lag 100 seems to coincide with the Daylight Savings Time change within Europe, which (as it is commonly known) happens during the last Sunday in March.

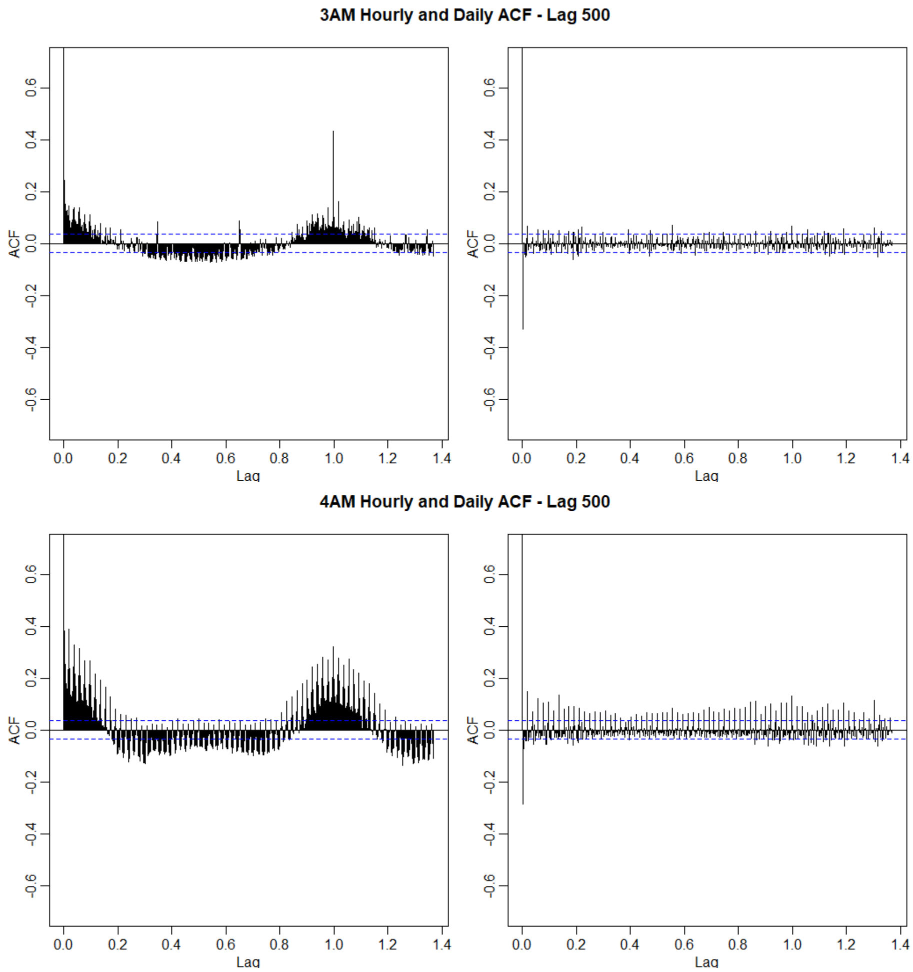

Figure 13: ACF plots for Hourly and Daily returns at 3AM and 4AM respectively.

The 3am and 4am plots differ slightly to the 12am, 1am and 2am plots, but largely follow a similar trend (periodicity). One interesting comparison at 4am is that the correlations do not reach their highest negative values at lag 0.5, whereas at 12am to 3am, they do. Comparing the hourly to daily differences again, we notice 3am follows a similar pattern to 1am and 2am respectively, hence no new information to decipher. However, the daily 4am is interesting in that, we now regularly see significant positive correlation values occur at regular intervals. This could be attributed back to our mean and variance (Figures 9 & 10) figures, where we start to see an increasing trend from 4am onward.

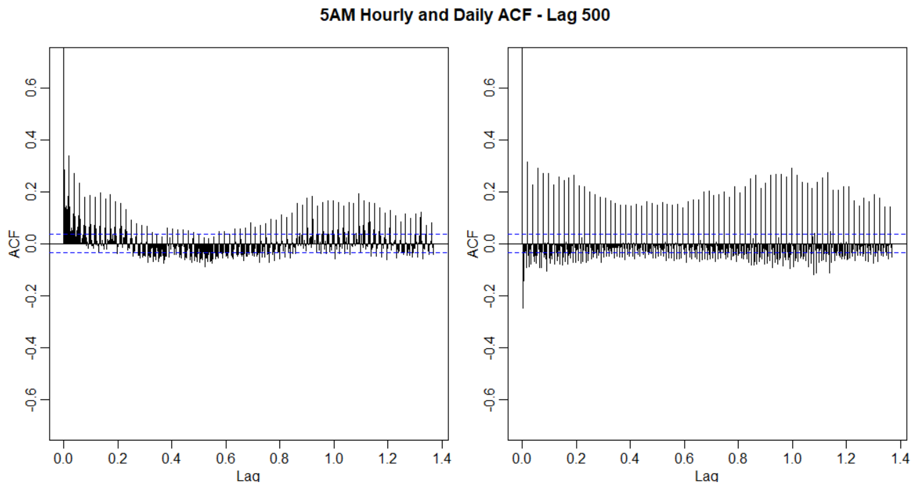

Figure 14: ACF: Hourly and Daily values at 5AM computed using RStudio.

The 5am plot is interesting, as now for the hourly returns, we tend to see that whilst there is some seasonality and significant correlation values are present, it is not as strong as previous hours analysed. For the daily returns, we can see that there are significant positive spikes, but negative correlations are more frequent. Next, we shall analyse 6am in-depth to see if we can notice any deeper patterns.

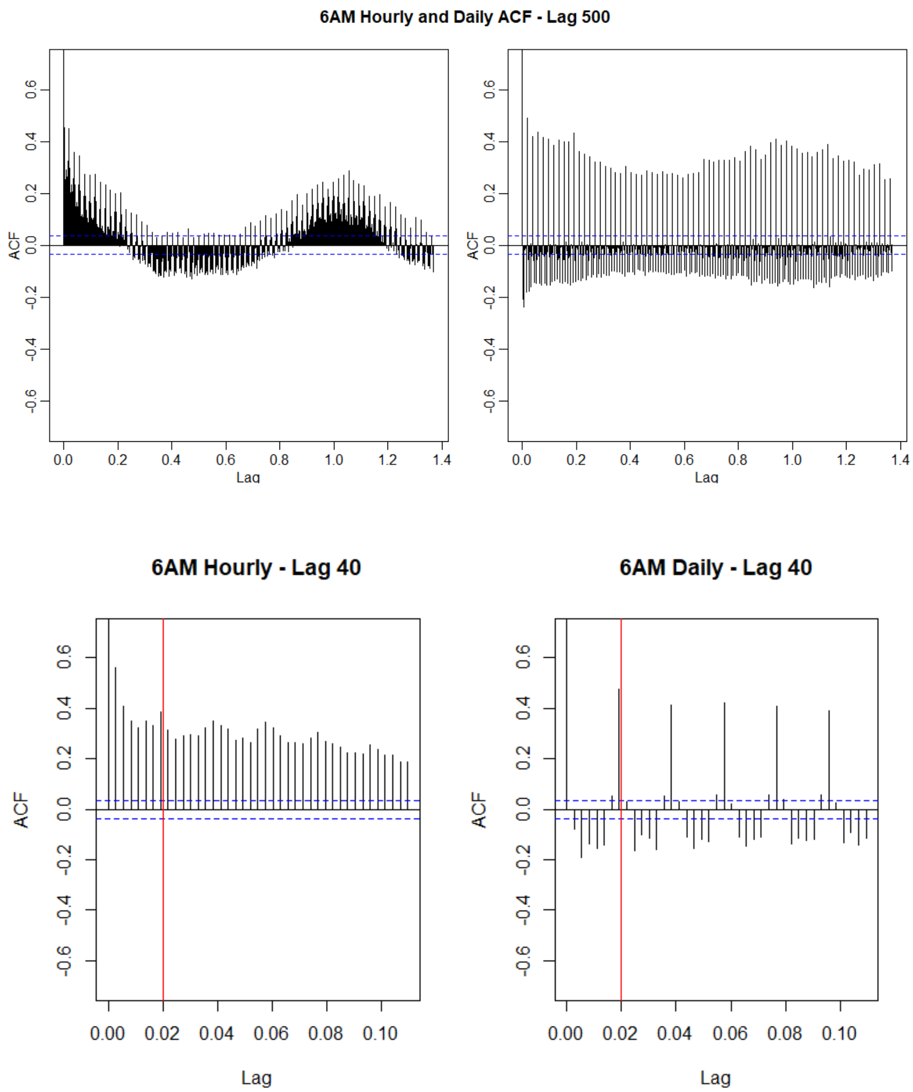

Figure 15: ACF: Hourly and Daily values at 6AM computed using RStudio.

As a reminder, lag 0.02 is essentially 1 week or 7 days hence we can see that there is weekly seasonality with the hourly and daily returns. The weekly seasonality for daily returns has been present for majority of the daily returns at all hours, however, I have left most of these plots out of this main analysis section, in the interest of keeping things concise.

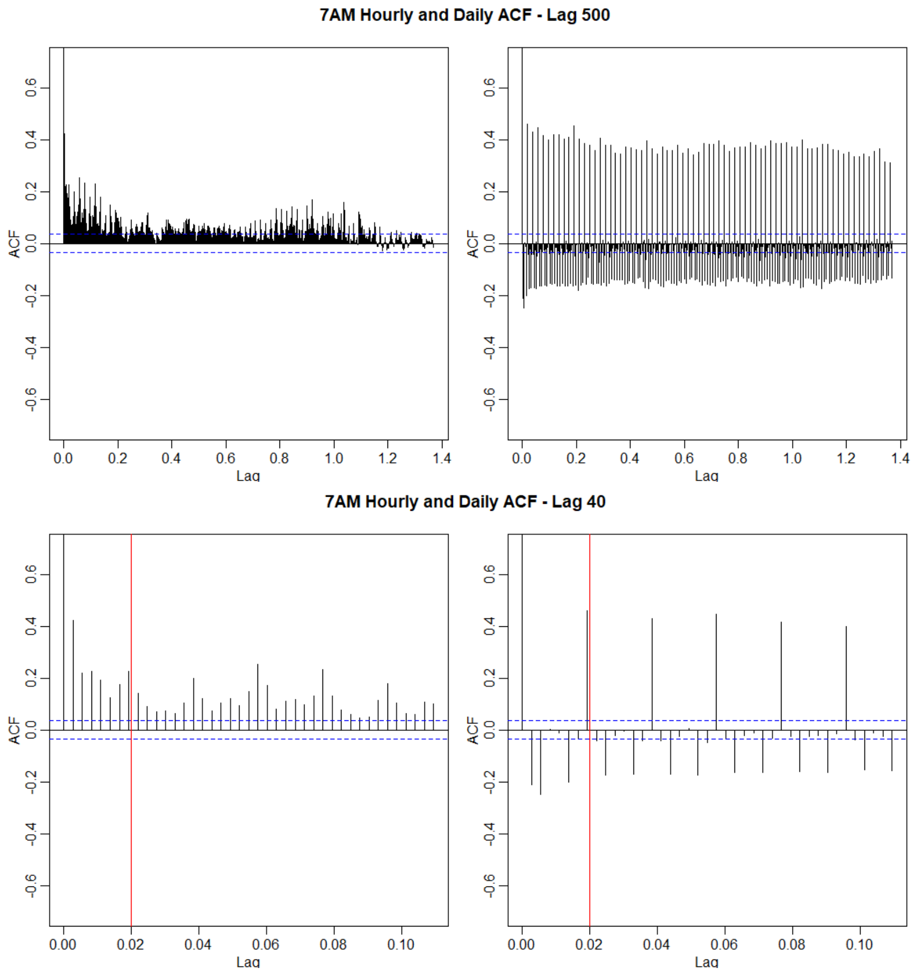

Figure 16: ACF: Hourly and Daily values at 7AM computed using RStudio.

Similar to the 6am plots, we can see some strong weekly seasonality in the lag - 40 plot, further highlighted by the vertical red lines. Now one thing that is different about the 7am hourly returns is just how positively correlated all of the values are up to lag 1, after that, there is some negative correlation. Comparing this to daily returns, you may expect, mostly positive correlation however, see both negative and positive correlations. One reason for this could be the weekday versus weekend effect i.e. the demand for electricity at 7am on Saturday, or Sunday will not be nearly the same as Monday hence negative correlation. We also find that the positive correlation spike only occurs once per week, the rest are negative or non-significant if positive. We can observe an almost inverse pattern if we compare the Hourly and Daily Lag 40 plots. For

example, in the Hourly plot, we see positive spike on day 7 (lag 0.02), then the correlation value goes down over the week, slowly rising until hitting the peak again at day 14 (lag 0.04).
Now, if we look at the Daily plot, we see a extreme positive spike at lag 0.02, which then turns negative, until lag 0.04, when there is another extreme positive spike. Overall, this suggests strong weekly seasonality (we can see a similar pattern in Figure 15), one where our lag of 0.02 can be used in Time Series models for accurate forecasts. Another interesting observation is that, whilst the hourly return plots look quite different when compared to that of 6am, the daily return plots are almost identical i.e. correlation values hover around 0.4 as well as -0.2.

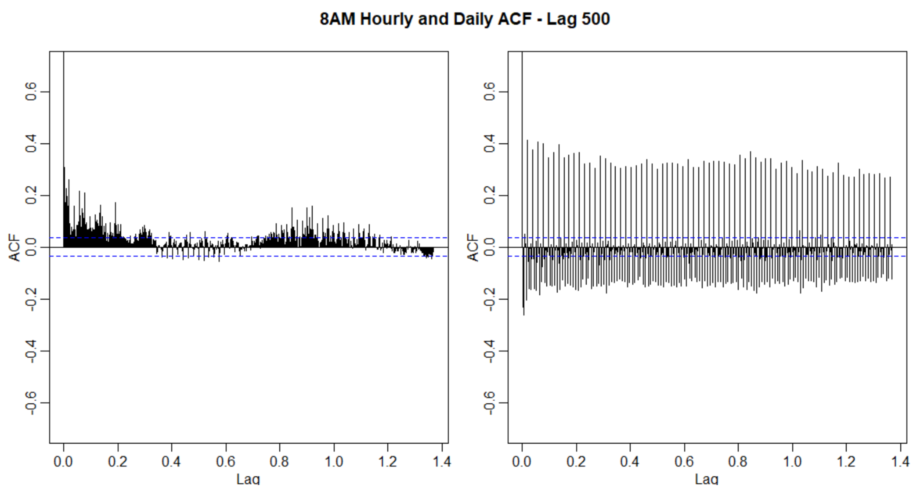

Figure 17: ACF: Hourly and Daily values at 8AM computed using RStudio.

At 8am we do start to see a more osculating pattern occur between lag 0.35 (approximately day 127), and 0.7 (day 255). This suggests some quarterly and/or weather seasonality i.e. summer periods when electricity load might be lower than usual. As for the Daily ACF plot, we can see that is still similar to that of the previous plots for 6am, and 7am.

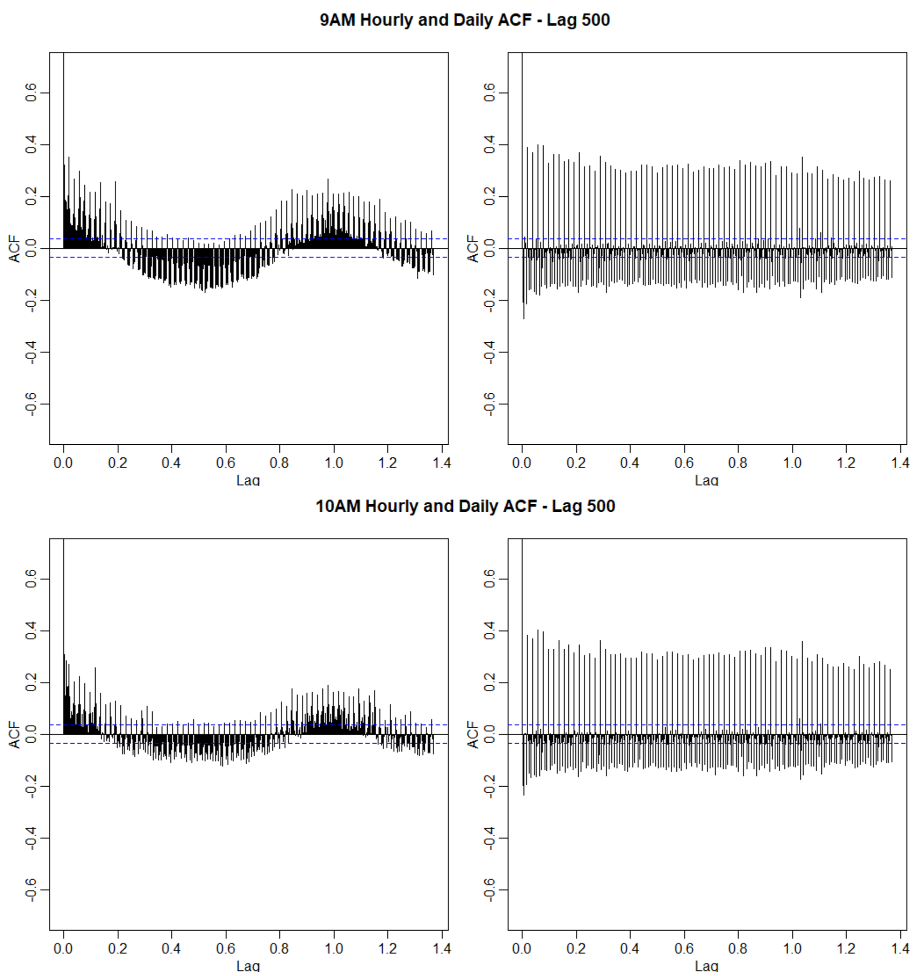

Figure 18: ACF: Hourly and Daily values at 9AM and 10AM computed using RStudio.

Here, for the hourly plots at both 9am and 10am, we can see a clear periodic trend emerge, with the notable lags (turning points) being 0.2, 0.8 and 1. Similar to previous interpretations, this suggests annual seasonality. Comparing both to their respective daily plots, we notice a similar pattern to that seen during previous hours analysed, in that we see some weekly seasonality as well as the frequency of positive correlations being lower than negative and the correlation values staying within the interval range of [-0.2, 0.4]. 11am follows a similar pattern to that of 10am and 9am. The primary difference being the hourly plot, which shows that between lag 0.2 and 0.8, the negative correlations are less significant (majority fall within the 95% confidence interval) compared to 10am. The 12pm plots exhibit similar patterns to the ones we saw for 7am

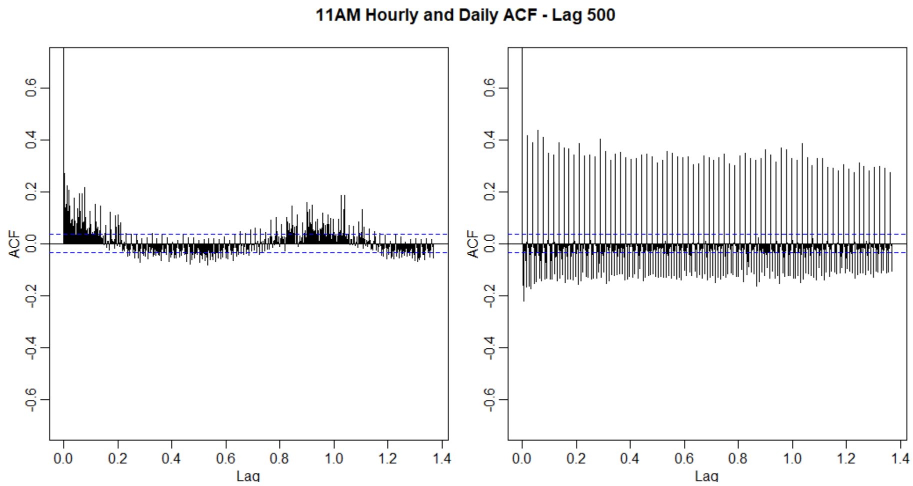

Figure 19: ACF: Hourly and Daily values at 11AM computed using RStudio.

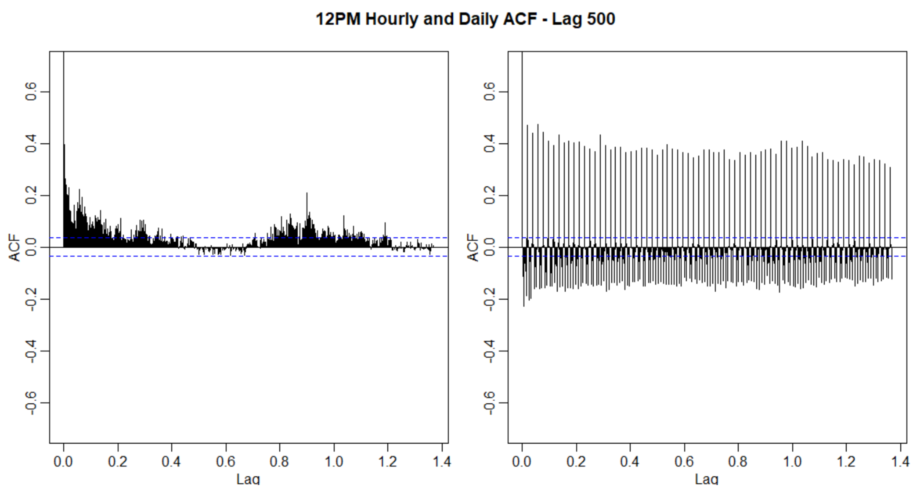

Figure 20: ACF: Hourly and Daily values at 12PM computed using RStudio.

and 8am respectively. For example, for the hourly plot, we notice a significant positive correlations with some negatives between lag 0.5 and 0.7. The hourly plot is similar in that, the correlations tend to lie within the range of correlation values [-0.2, 0.4].

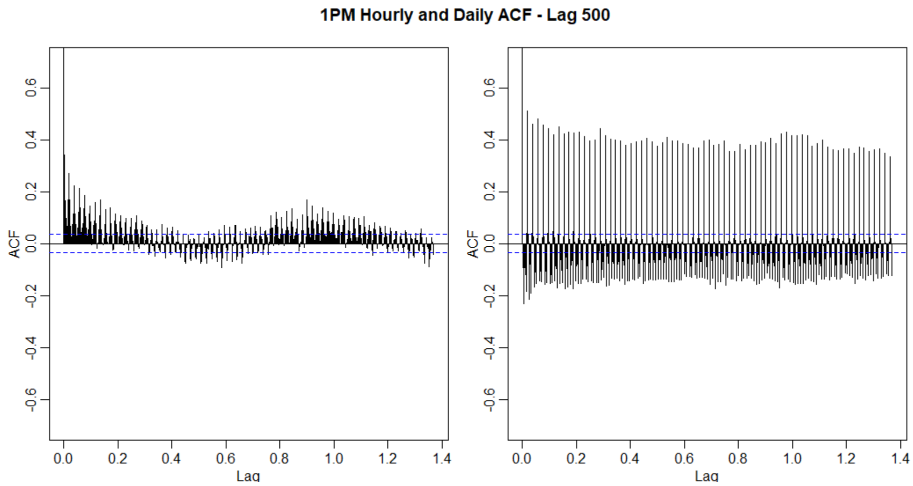

Figure 21: ACF: Hourly and Daily values at 1PM computed using RStudio.

The hourly plot for 1pm is interesting compared to previous hours, as it has some clear oscillations, which usually clumped together in groups and undergo a seasonal trend (yearly). Comparing to hourly plot, we can see this has some impact, as the negative correlations, compared to previous hours, has increased overall (i.e. now these values are for the better part, above the 95% CI).

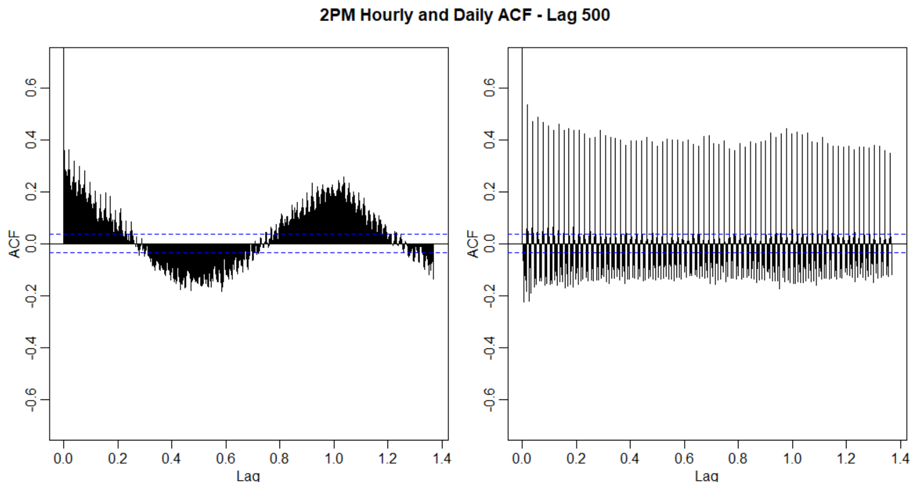

Figure 22: ACF: Hourly and Daily values at 2PM computed using RStudio.

We see that the 2pm plot seems like a robust upgrade from 1pm, as the periodic (cyclical) pattern is much more evident now, with clear turning points shown (at lag 0.25 and 0.75 respectively). The only thing that has changed for the daily plot now, is that the negative correlations have increased in value but also stay around similar values. We can tell from the daily plot that previous returns greatly influence (or correlate) the current returns and this is ever present. The same can be said for the hourly plot as all of the correlation values exceed the 95% Confidence Interval (CI), which means they are significant, and do have an impact. My analysis for 3pm and 4pm is largely the same as what I concluded for 2pm plots. Another possible explanation is renewable energy generation (solar panels), hence why we see some negative correlations. However, I feel this affect may not be as significant as one might expect nowadays given our dataset is between 1999-2007, when renewable energy was not as prevalent as it is today.

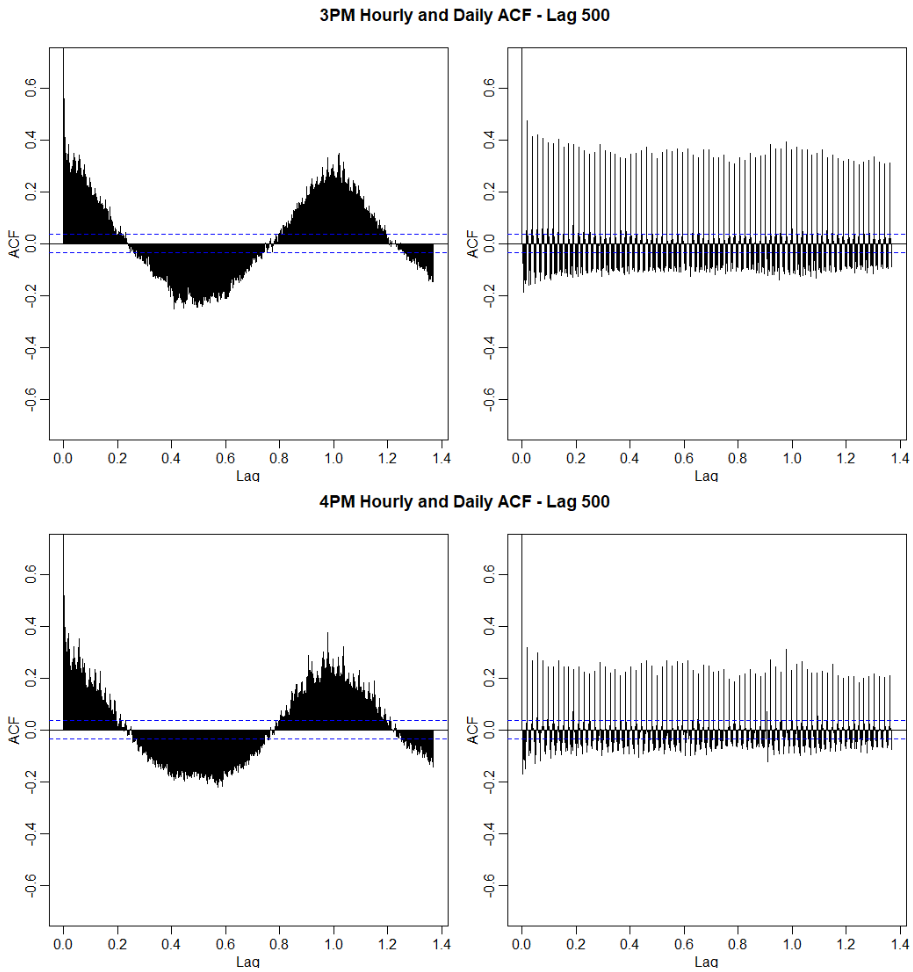

Figure 23: ACF: Hourly and Daily values at 3PM and 4PM computed using RStudio.

5pm is interesting as it seems to show annual seasonality, however, the cycles or periods within are non-standard when we compare to 3pm ACF which has smooth peaks and troughs (almost like normal distribution during it's cycles). Hourly and daily correlation values are the same to an extent, shows that both have the same level of impact on future prices. Compared to 3pm, the values are also lower.

Now comparing 5pm to 6pm yields interesting comparisons. Both have positive correlations up until lag 36 (1 month). Afterwards, the values oscillate in opposite directions, as you can see between lag 100 and lag 250 (between 0.3 and

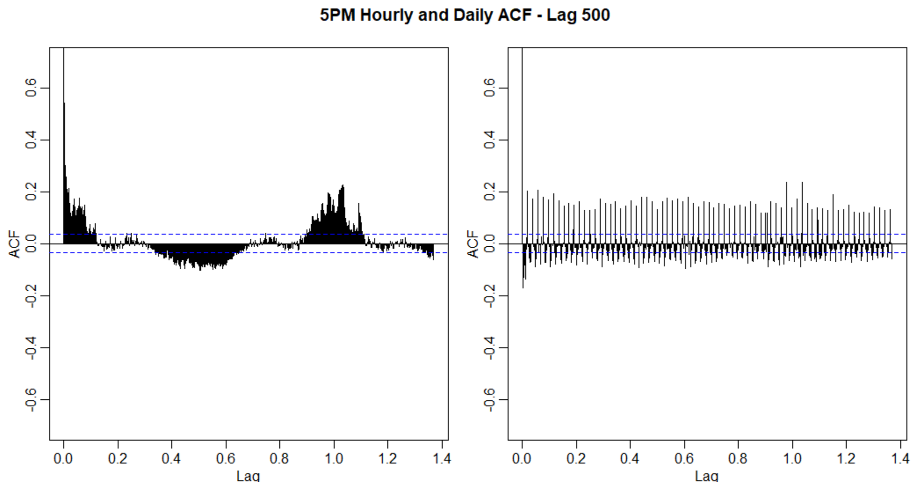

Figure 24: ACF: Hourly and Daily values at 5PM computed using RStudio.

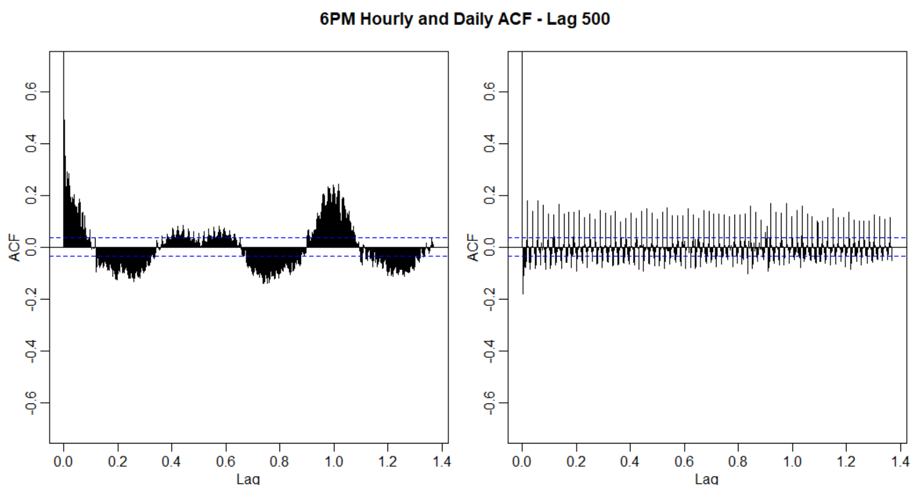

Figure 25: ACF: Hourly and Daily values at 6PM computed using RStudio.

0.7 lag values), the 5PM ACF values are negatively correlated whereas 6PM ACF values are all positive. The 7PM ACF values follow a similar pattern to the 6PM ACF values. Comparing 7PM to 8PM, mirrors much of what the comparison between 6PM and 5PM. It should be noted that 5PM and 8PM ACF values are different, for example, the negative values around lag 183 (lag 0.5 in Figure 22), are more significant and higher than 5PM where the negative correlation values persist over a longer period of time. Comparing the Daily

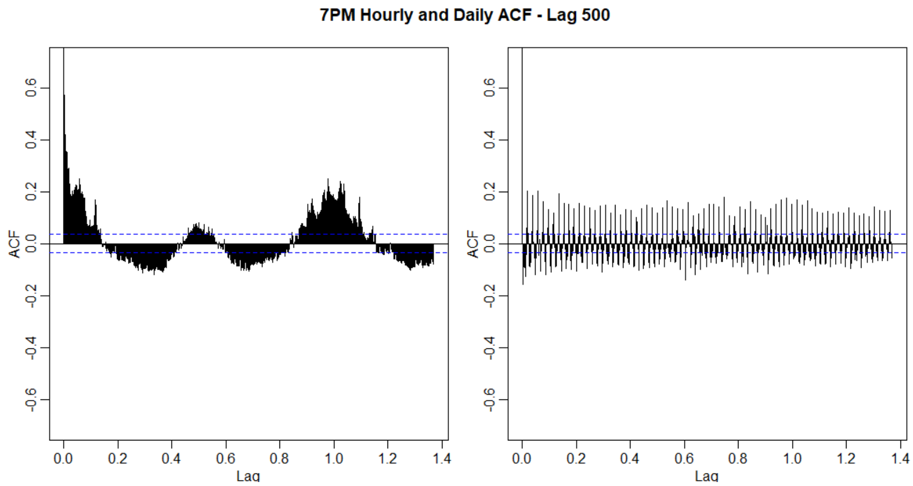

Figure 26: ACF: Hourly and Daily values at 7PM computed using RStudio.

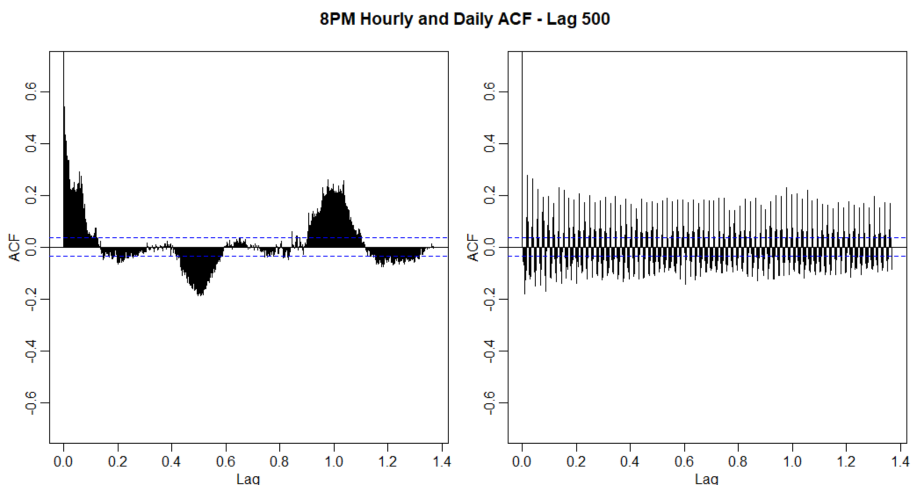

Figure 27: ACF: Hourly and Daily values at 8PM computed using RStudio.

ACF values, we notice that the 8PM values are higher than 7PM, aside from that, there are no major distinctions to be made.
The 9PM ACF shows interesting oscillations, which seem to suggest a sort of seasonal pattern - a monthly seasonality. In a high level sense, the 9PM Hourly and Daily ACF values are similar to that of 8PM, this makes sense given they are both evening time periods.

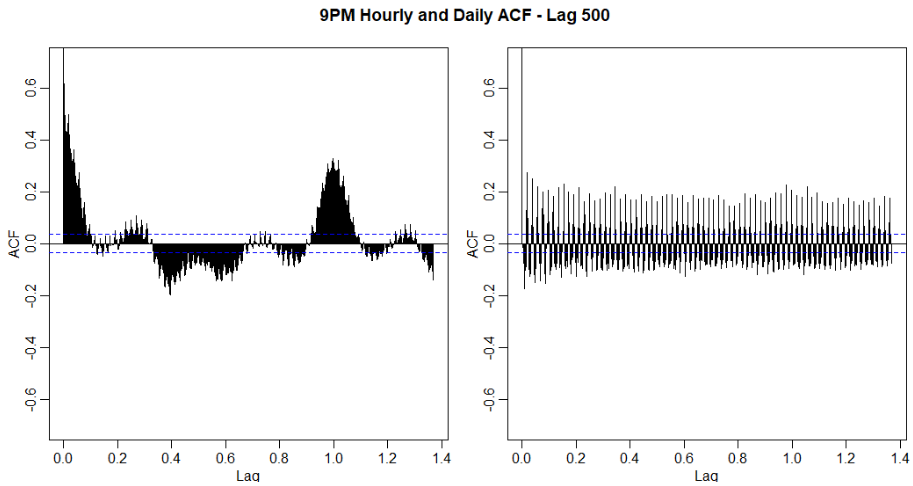

Figure 28: ACF: Hourly and Daily values at 9PM computed using RStudio.

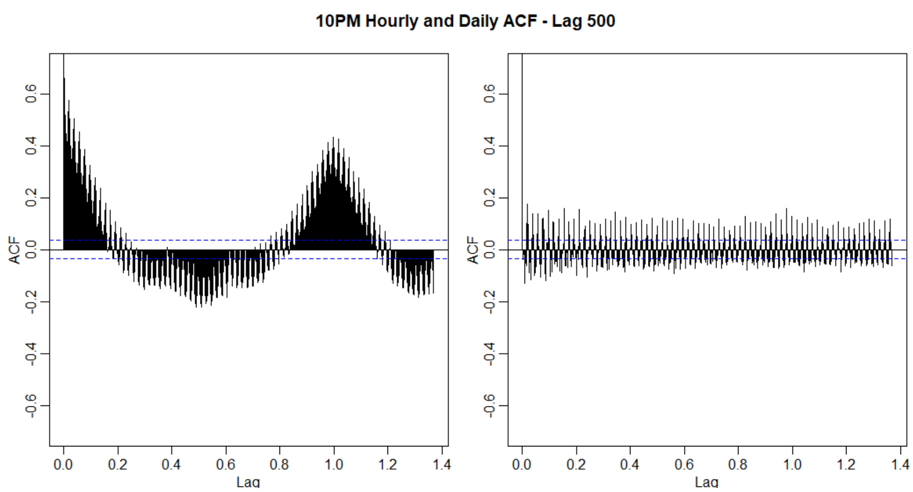

Figure 29: ACF: Hourly and Daily values at 10PM computed using RStudio.

Finally, comparing our 10PM and 11PM ACF figures, we can see that the 10PM ACF values are much more extreme and significant compared to 11PM. A distinction to notice is that at 11PM, the values are positively skewed for a longer period of time compared to 10PM. We can also see that the Daily ACF values are smaller for 11PM when compared to 10PM. This means that 11PM values can be predicted with more confidence than 10PM, and this makes logical sense given 11PM is late night time period and majority of the population will be

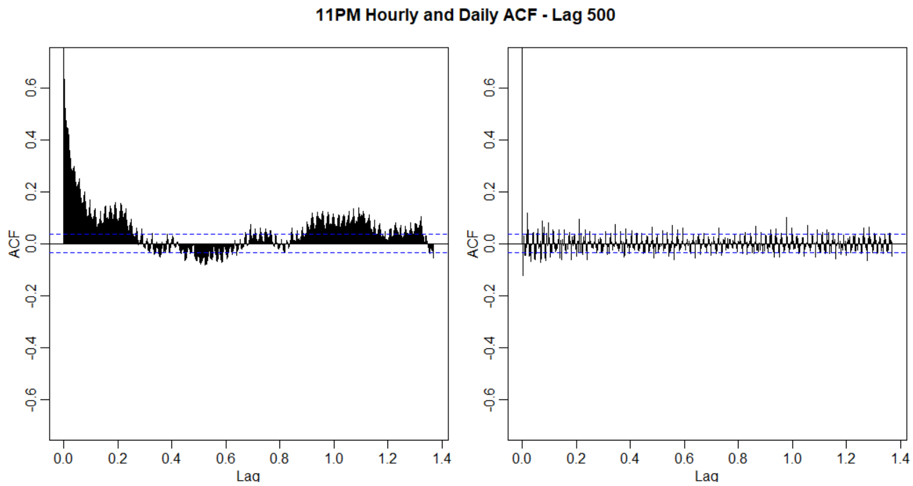

Figure 30: ACF: Hourly and Daily values at 11PM computed using RStudio.

asleep or will be soon (compared to 10PM). To further back it up, we can look at the 12AM as well as 1PM ACFs and notice similar patterns.

Summary

To summarise subsections 4.1 to 4.2, we notice a few trends, looking at the daily variance plot, we can see over the hours, the correlations increase and decrease in value, in an almost identical fashion to our hourly variance as well as hourly mean plots in Figures 9 & 10. This can be interpreted in two ways, one is that as the hourly returns increase, so does the variance, leading us to infer that a higher mean (or price) indicates an increased level of volatility (variance). All in all, both the Daily & Hourly Mean plots, as well as the Daily & Hourly Variance plots show us a similar trend as the day goes on. For example, the highest values occur in the morning hours of 5AM to 9AM, with the lowest values occurring during the night hours between 9PM and 3AM.

For section 4.3, we notice that Daily ACF values, from 5AM until 4PM, stay within the interval of [-0.2,0.4], with the positive value being above 0.2. However, after 4PM onward, the ACF values fall to 0.2 or below with the interval now changing to [-0.1,0.2], indicating that these hours are less affected by the prices today when compared to hours between 5AM to 4PM, where they are more likely to be impacted.

3.4 GARCH

Here we used Python and various libraries in order to fit a GARCH model and forecast over periods to see the fit. Through trial and error, we settled on GARCH(5,0) model which is essentially an ARCH(5) model. This was because our parameters (variables) had significant p values as opposed to other tried models such as GARCH (1,1), GARCH(2,1), GARCH (2,0) and GARCH(3,1).

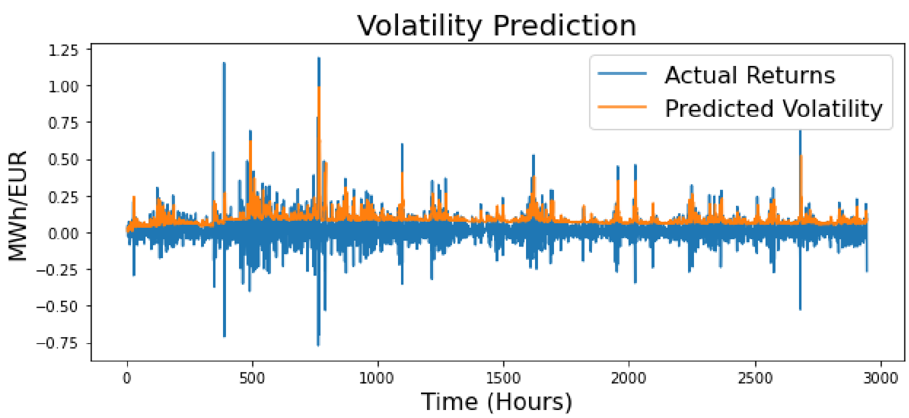

Figure 31: Created in Python with code modified and referenced from Ritvik Khakar's Github page [31].

Mean Model

	coef	std err	t	P>|t|	95.0% Conf. Int.
mu	7.3504e-03	1.457e-03	5.045	4.532e-07	[4.495e-03,1.021e-02]

Volatility Model

	coef	std err	t	P>|t|	95.0% Conf. Int.
omega	2.5551e-03	6.815e-04	3.749	1.775e-04	[1.219e-03,3.891e-03]
alpha[1]	0.2255	6.694e-02	3.369	7.554e-04	[9.429e-02, 0.357]
alpha[2]	0.5123	0.214	2.392	1.675e-02	[9.258e-02, 0.932]
alpha[3]	0.0351	1.898e-02	1.852	6.404e-02	[-2.051e-03,7.235e-02]
alpha[4]	0.0537	2.752e-02	1.950	5.123e-02	[-2.863e-04, 0.108]
alpha[5]	0.0880	3.125e-02	2.816	4.865e-03	[2.674e-02, 0.149]

Figure 32: Model test results.

It is important to emphasise here that we are measuring volatility (variance) as opposed to actual returns or prices (this is our data). Here we can see that the GARCH model is effective at being able predict the volatility spikes that occur. Thus, we can utilise this method to gauge the future volatility. For example, if the GARCH model predict increased volatility, then it might be worth it as a energy supplier to buy future energy contracts at set prices to avoid having to pay excessive prices in the intra-day or day-ahead markets.

Interestingly, given the usefulness of the GARCH model, people have combined it with the ARMA model in order to build an overall effective ARMA-GARCH model for forecasting [33]. This is useful as we can combine the ARMA process to capture general mean trends and then combine this with variance trends (i.e. our spikes), which can lead to more accurate predictions.

3.5 Facebook Prophet

As mentioned before, Prophet is a useful method (or tool) to quickly analyse business data with ease and allow you forecast with minimal statistical knowledge requirements.

With reference to Figure 33 let us breakdown it down step by step. For our trend plot, we can see it aligns well with Figure 4 as we constant price fluctuations, with an upwards trend in 2003, followed by a decline and a gradual increase in price leading to 2007.

This holiday plot is quite interesting. I gathered up all the relevant public holidays within the member countries of Nord Pool within our dataset time period (1999-2007), aggregated these holidays and made it a parameter for analysis. For disclosure, the countries I included are as follows: Sweden, Norway, Finland, Denmark, and Germany. Now that we see price spikes relevant to holidays plot, it is fairly self explanatory i.e. public holidays near the start of the year (New Years Eve) have some of the largest variations in price. We also have a Spring/Summer period of holidays but the Autumn/Winter period has relatively few spikes.

Now we get to the interesting weekly trend plot, which is self explanatory and further strengthen our belief discussed earlier that prices are higher during weekdays and lower during the weekend due to demand.

For the yearly plot, we can see that there is seasonal effect (e.g. seasons of winter and summer). Interestingly, I would have expected prices to trend higher during the winter months as opposed to lower as winters are typically have longer dark periods of the day, hence more street lights and other utilities being used. One reason I can think behind this is that winter demand is relatively stable whereas summer demand is lower and fluctuates, hence to combat this, utility companies raise prices to keep profitable [3].

Finally, for our daily plot, we can see a clear cyclical pattern, where the prices are lower during the nigher and higher during the day, with the peak being during rush hour.

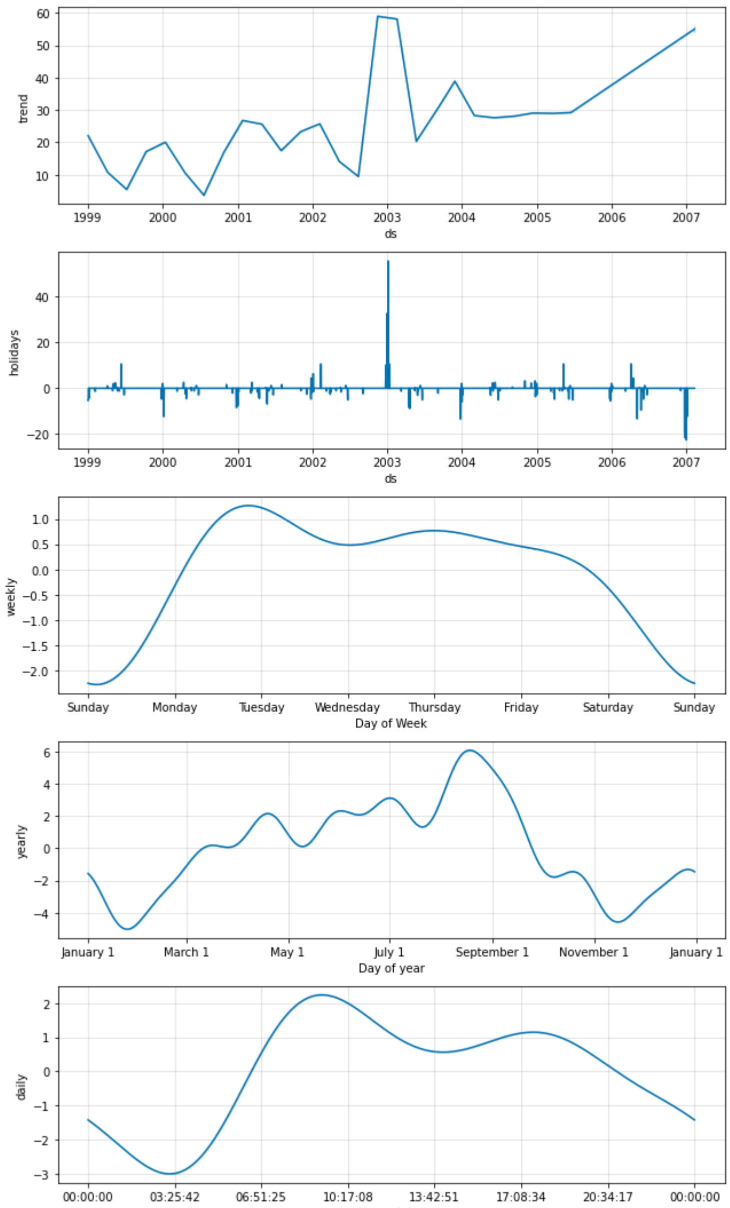

Figure 33: This figure shows us the seasonal and trend breakdown of our model.

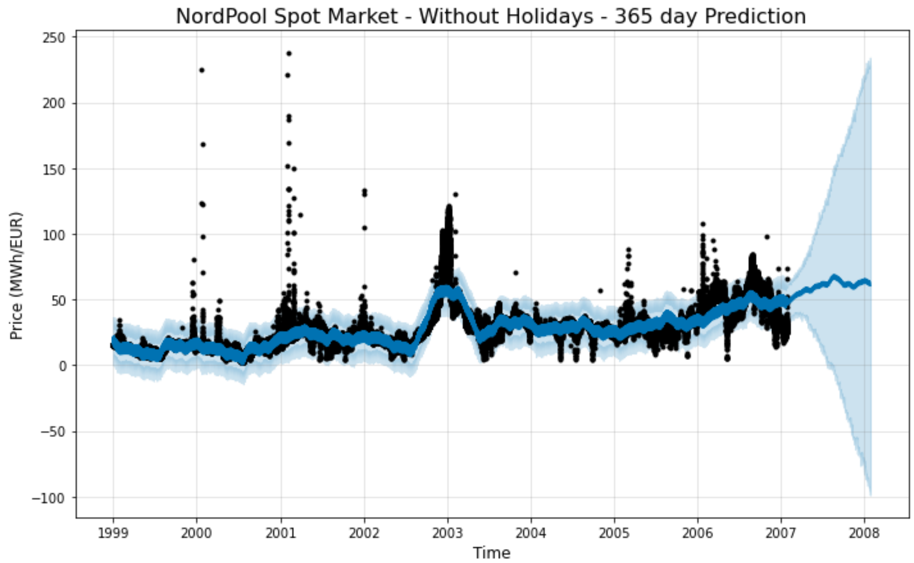

Figure 34: Prophet model without Nordic holidays as a parameter.

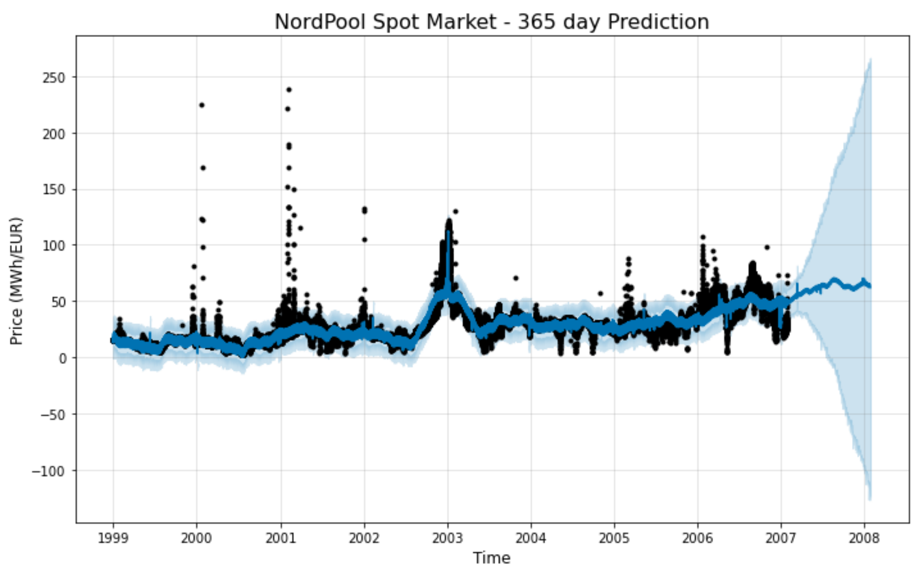

Figure 35: Prophet model with Nordic holidays included as a parameter.

For this forecasting section, I initially created a model without the holidays parameter, and it the model itself was not able to track the price spikes, which meant the 365 day forecast could be improved. As a result, I decided to create a holidays parameter (as explained earlier), hence we can clearly see the successful result in the bottom plot, where our new model was able to track some of the spikes. The key thing I would want to highlight here is the fact that, forecasting over longer period of time is not ideal as we can in Figure 35. We can see that for the first 3 months, the confidence interval is relatively normal in size, however as we forecast further, this gap widens to an extreme amount. For context, in both models, the confidence interval was set to 95%, hence the light blue area surrounding the blue line is the 95% confidence interval space.

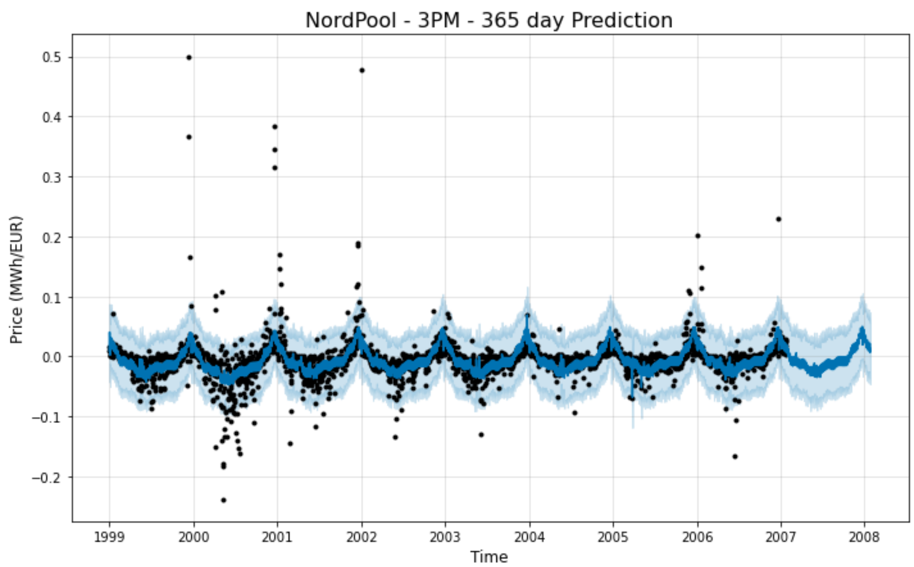

Figure 36: 3PM hourly returns forecast.

For curiosity's sake, I created a model with data from only 3PM returns, as the 3PM returns were cyclical (as we could see from our 3PM ACF in Figure 23). We can see here that Prophet actually does a good job of forecasting out long-term returns (prices). The drawback here being that the seasonality (or how cyclical the data is) must be simple, which in our case for electricity prices, is not always the case.

3.6 LSTM

Figure 37: Regression LSTM model. Blue is actual data, orange is training data, and green is testing data. The Y-axis is price in MWh/EUR and the X-axis is Time (Hours). Created via Python code with references from Jason's Machine Learning blog [28] and Tamara Louie's time series seminar [34].

With our LSTM model, I was surprised at just how accurate the model was despite only using 67% of the data, as the training set. For this mode, we ended up with an RMSE score of 2.46 on the training section, and 1.98 on the test section for our pure dataset. However, when I used the same mode except the input was our logarithmic returns, the RMSE came back much smaller, with both training and testing at 0.05. This is not a direct comparison but we did a similar method using Prophet, where we trained on 1975 days worth of data and then forecast prices every 30 days up until 26/1/2007 to test accuracy. Our RMSE for that model was 6.86. Hence it is safe to say that LSTM has an edge here.

To close off the section, as we know of the problem with long-term accurate forecasting, one way LSTMs could end up being vastly superior and very useful, is due to their scalability. They can be used for short-term models, which will have high confidence interval predictions and can work out in our favour, as shown in these papers [32, 5].

3.7 Distributions

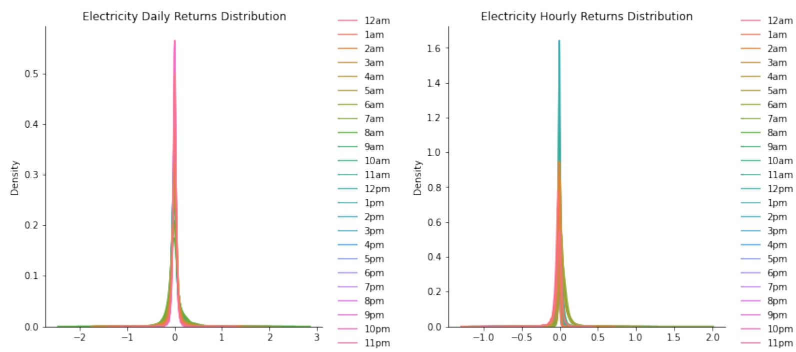

Figure 38: Two plots created to compare the distribution of returns between hourly and daily.

From the below figure, we can clearly tell that one of them follows a more Gaussian distribution than the other. This makes logical sense given that we remove part of the volatility that comes with hourly returns dependencies (e.g. price at 8AM will heavily impact price at 9AM), as a daily return value is the average of 24 data points. If we were to partially ignore this, it could be said that we should use the daily return values in order to make more reliable (read: stable) electricity price forecasts as they follow a more Gaussian distribution, compared to the hourly return values.

4 Conclusion

A lot has been discussed in this dissertation both on the data analysis as well as literature side. I initially set out to learn more about the energy (electricity) industry, learn about the regulatory challenges as well as aiming to de-mystify the industry. As, to the average consumer, an electricity bill can look daunting with the prices, calculations and other miscellaneous information. Finally, I also wanted to explore this industry in a more mathematical sense i.e. what are effective methods of accurately forecasting electricity prices? What are the challenges we face?

After crunching through 70,000 or so values, utilising Excel, Python and RStudio, we now can reach for the conclusion. Initially, we analysed the mean and variance values between the hourly and daily returns, of each hour. The highest hourly and daily returns were 7AM and 8AM respectively, which makes logical sense, with the lowest values being 1AM and 11PM. However, I noticed that the hourly mean return became negative at 6PM instead of a later time like 8PM. The daily returns were all positive, which was interesting as it does comply with our observation that electricity prices have been trending up over the years. We had similar observations with regards to our variance analysis.

We found out that there are a number of effective ways to forecast electricity prices. However, the best methods remain unknown (assuming the large operators keep their models proprietary as online information is sparse) or form a mixture of statistical and machine learning approaches. The reason for this is due to the challenges with the data and the nature of electricity itself. We had a large dataset spanning over 8 years with a sampling rate of 1 hour, as a result we had multiple seasonalities to account for. This is not something that can be easily done with an ARIMA model. In order, for me to have implement an accurate ARIMA model, I would have had to use my daily average values and potentially cut them down further into weekly or monthly averages. This was something I wanted to avoid as a significant part of this dissertation was to see if there are any differences between hourly and daily prices and forecasting (ACF).

We found that there are differences between the hourly and daily ACFs. One stark difference is that, the daily ACF values are constant and follow a weekly seasonality (which is logical given electricity prices do fluctuate in a cyclical pattern over a week). The hourly ACF values are more volatile but this due to the ACF values essentially being influenced by the hours prior and after them to an extent. As a result, the hourly ACF plots show us more long term seasonalities (monthly and yearly).

We also observed, GARCH volatility forecasting, noting that it is a great method to model something with high volatility, such as electricity prices. Finally, we tried out some more non-traditional forecasting methods such as Facebook Prophet and LSTM (neural networks) and decided to get their RMSE outputs. Facebook's Prophet is not ideal for long or medium term forecasting of electricity prices as it fails address the volatility, similar to ARMA models. However, the good thing is that Prophet model is additive, hence we can always add more parameters. Additionally, I tested out the method on our 3PM dataset which has one of the lowest mean and Prophet was able to handle it very well, looking at the confidence intervals. On that note, the LSTM method was a success, both in terms of how easy it was to implement and it's forecasting ability as evidenced by Figure 37.

Finally, we looked at the hourly and daily returns distribution plots, where it was evident that the hourly returns were more skewed hence, if we wanted more stable forecasts, it was best to use the daily return values given the more even (Gaussian) distribution.

In conclusion, there is a lot more research to be done and I hope to carry on this project further.

5 Next steps

Given my mention of wanting to continue working on this project. I thought it would be useful to discuss next steps and how this project could be further improved.

I felt that the inclusion of new data would have been interesting to contrast and compare, as the markets have changed when we compare the 1999 to 2007 period, to 2021. For example, considering additional data like electricity consumption, production (by source such as solar, wind, nuclear, geothermal, hydropower), temperature, predicted temperature, and other metrics such as population growth. It makes sense to take into account more recent data such as from 2010-2020 given the rise in renewable energy generators. Another interesting area to explore would be to see just how much more volatile electricity prices have gotten from 1999-2009 to 2010-2020 (when this year ends). This would have been achievable by comparing similar, large datasets, and comparing GARCH forecasts as well as other variance-focused analysis.

For the ACF analysis, it may have been good to compare additional lag values to get a comprehensive picture such as lag values of 14, 50, 100 and 250.

For a more specific comparison of models project, it would have been a better if I had reduced the dataset to monthly averages then computed forecasts comprising of SARIMA, TBATS, Holt-Winters, ARMA-GARCH and LSTM methods. The use of monthly data is to suit the SARIMA method as it cannot handle multiple seasonalities well.
As we saw in Figure 36, with my 3PM forecast using Prophet, one alternative method of forecasting prices could actually be utilising different methods for different parts of the day, maybe even specific hours or in 15 minute intervals. I believe combining the outputs from all of this models to then make an informed decision may lead to the best outcome. On that note, I have come upon the topic of probabilistic price forecasting for electricity [50, 7, 26, 14, 52] which is fascinating, as it accepts the idea of uncertainty in the real world. For prices, this would be great as we could get p-values for certain extreme events when we known things might happen (i.e. around December or January) and as a result, we could buy derivatives or futures contracts, in order to mitigate risk.

For the distribution section, I could have calculated the skewness and kurtosis first, and then made a judgement based off of those values. From that, we could have yielded extra information such as which hour, be it hourly or daily, follows a GBM (Geometric Brownian Motion) or which hour has the highest and lowest skewness.

Another topic I wanted to explore was the impact of Bitcoin [49] and cryptocurrency mining, on electricity prices in Europe. The ideal price or currency to compare and contrast prices with would be Bitcoin, since it has been around

2009, with a volatile price history. For example it would be fascinating to do, correlation/covariance test for 2017 (first extreme price spike for Bitcoin) prices, and then compare to 2020 and 2021 data as well. There is a great website where electricity consumption data is available, which is run by the University of Cambridge (Cambridge Bitcoin Electricity Consumption Index) however, the initial model was developed by Marc Bevand in 2017 [36]. The reason I bring up Bitcoin mining is due to China's relentless crack down on this currency. Last year the Bitcoin mining accounted for 1.1% or 86 TWh (Terawatt hours) of electricity in China [11] and this is significant, as if you recall earlier, electricity has a storage problem and even a small percentage load difference can have impact on prices and profit of the market participants.

One thing is for certain from all this - forecasting electricity prices accurately is only going get more challenging from now on, and therefore it is a challenge we must welcome.

References

[1] The sveriges riksbank prize in economic sciences in memory of alfred nobel 2003.

[2] Quants: The alchemists of wall street, 2010.

[3] U.S. Energy Information Administration. Electrcity explained, 2021. https://www.eia.gov/energyexplained/electricity/prices-and-factors-affecting-prices.php.

[4] Nesreen K Ahmed, Amir F Atiya, Neamat El Gayar, and Hisham El-Shishiny. An empirical comparison of machine learning models for time series forecasting. *Econometric Reviews*, 29(5-6):594–621, 2010.

[5] Andrés M Alonso, F Javier Nogales, and Carlos Ruiz. A single scalable lstm model for short-term forecasting of disaggregated electricity loads. *arXiv preprint arXiv:1910.06640*, 2019.

[6] Sören Amelang. Negative electricity prices: lockdown's demand slump exposes inflexibility of german power. https://energypost.eu/negative-electricity-prices-lockdowns-demand-slump-exposes-inflexibility-of-german-power/.

[7] François Bouffard and Francisco D Galiana. An electricity market with a probabilistic spinning reserve criterion. *IEEE Transactions on Power Systems*, 19(1):300–307, 2004.

[8] Peter J Brockwell, Peter J Brockwell, Richard A Davis, and Richard A Davis. *Introduction to time series and forecasting*. Springer, 2016.

[9] The California Independent System Operator (CAISO). Data page, 2021. https://www.caiso.com/todaysoutlook/Pages/supply.html.

[10] John Y Campbell, Andrew W Lo, and A Craig MacKinlay. *The econometrics of financial markets*. princeton University press, 2012.

[11] Lefteris Karagiannopoulos Carlos Torres Diaz. Bit late for bitcoin: How china's crackdown is reducing more emissions than whole countries emit, 2021. https://www.rystadenergy.com/newsevents/news/press-releases/bit-late-for-bitcoin-how-chinas-crackdown-is-reducing-more-emissions-than-whole-countries-emit/.

[12] Bob Carpenter, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. Stan: A probabilistic programming language. *Journal of statistical software*, 76(1):1–32, 2017.

[13] Tianfeng Chai and Roland R Draxler. Root mean square error (rmse) or mean absolute error (mae)?–arguments against avoiding rmse in the literature. *Geoscientific model development*, 7(3):1247–1250, 2014.

[14] Shing-Chow Chan, Kai Man Tsui, HC Wu, Yunhe Hou, Yik-Chung Wu, and Felix F Wu. Load/price forecasting and managing demand response for smart grids: Methodologies and challenges. *IEEE signal processing magazine*, 29(5):68–85, 2012.

[15] Bradford Cornell. Medallion fund: The ultimate counterexample? *The Journal of Portfolio Management*, 46(4):156–159, 2020.

[16] accessed via Yahoo Finance Data provided by Nasdaq. Yahoo sp 500 etf, 2021. https://finance.yahoo.com/quote/SPY/.

[17] Alysha M De Livera, Rob J Hyndman, and Ralph D Snyder. Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American statistical association*, 106(496):1513–1527, 2011.

[18] Western Power Distribution. The electric journey. https://www.westernpower.co.uk/downloads/4921.

[19] Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the econometric society*, pages 987–1007, 1982.

[20] Robert F Engle and Clive WJ Granger. Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, pages 251–276, 1987.

[21] Sean J Facebook(Taylor and Benjamin) Letham. Prophet official github page. https://facebook.github.io/prophet/.

[22] Sean J Facebook(Taylor and Benjamin) Letham. Prophet r. https://cran.r-project.org/web/packages/prophet/index.html.

[23] Nord Pool Group. Day-ahead market, 2021. https://www.nordpoolgroup.com/the-power-market/Day-ahead-market/.

[24] Trevor J Hastie. *Generalized additive models*. Routledge, 2017.

[25] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[26] Tao Hong and Shu Fan. Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting*, 32(3):914–938, 2016.

[27] Rob J Hyndman and George Athanasopoulos. *Forecasting: principles and practice*. OTexts, 2018.

[28] PhD Jason Brownlee. Time series prediction with lstm recurrent neural networks in python with keras, 2016. https://machinelearningmastery.com/time-series-prediction-lstm-recurrent-neural-networks-python-keras/.

[29] Robert C.Kelly Jason Fernando. Commodity. https://www.investopedia.com/terms/c/commodity.asp.

[30] Paul L Joskow. Lessons learned from electricity market liberalization. *The Energy Journal*, 29(Special Issue# 2), 2008.

[31] Ritvik Kharkar. Garch stock modeling. https://github.com/ritvikmath/Time-Series-Analysis/blob/master/GARCH Stock Modeling.ipynb, 2020.

[32] Weicong Kong, Zhao Yang Dong, Youwei Jia, David J Hill, Yan Xu, and Yuan Zhang. Short-term residential load forecasting based on lstm recurrent neural network. *IEEE Transactions on Smart Grid*, 10(1):841–851, 2017.

[33] Shiqing Ling and Michael McAleer. Asymptotic theory for a vector arma-garch model. *Econometric theory*, 19(2):280–310, 2003.

[34] Tamara Louie. Applying statistical modeling machine learning to perform time-series forecasting, 2016. https://github.com/tklouie/PyData$_L A_2 018/blob/master/PyData_L A_2 018_T utorial.ipynb.$

[35] Jacob Manoukian. What do negative crude oil prices even mean?, 2020. https://www.jpmorgan.com/wealth-management/wealth-partners/insights/what-do-negative-crude-oil-prices-even-mean.

[36] University of Cambridge Mark Bevand. Cambridge bitcoin electricity consumption index, 2021. https://cbeci.org/.

[37] Kassia Micek. California duck curve 'alive and well' as renewable, minimum net load records set, 2021. https://www.spglobal.com/platts/en/market-insights/latest-news/electric-power/032621-california-duck-curve-alive-and-well-as-renewable-min-net-load-records-set.

[38] James Moody. The electric journey. 2019. https://towardsdatascience.com/what-does-rmse-really-mean-806b65f2e48.

[39] Kjell Nilsson. Nordregio news 2 2017: 20 years of regional development. 2017. http://norden.diva-portal.org/smash/record.jsf?pid=diva2%3A1106252dswid=-4147.

[40] Ofgem. Understanding trends in energy prices. https://www.ofgem.gov.uk/gas/retail-market/retail-market-monitoring/understanding-trends-energy-prices.

[41] Christopher Olah. Understanding lstms. https://colah.github.io/posts/2015-08-Understanding-LSTMs.

[42] Nord Pool. Nord pool history page. https://www.nordpoolgroup.com/About-us/History/.

[43] quantivity. Why log returns. ://quantivity.wordpress.com/2011/02/21/why-log-returns/.

[44] Tuomas Rintamäki, Afzal S Siddiqui, and Ahti Salo. Does renewable energy generation decrease the volatility of electricity prices? an analysis of denmark and germany. *Energy Economics*, 62:270–282, 2017.

[45] Mary-Ann Russon. The cost of the suez canal blockage. https://www.bbc.co.uk/news/business-56559073.

[46] Engie SA. Understanding the capacity market. https://www.engie.co.uk/wp-content/uploads/2016/07/capacitymarketguide.pdf.

[47] Sean J Taylor and Benjamin Letham. Forecasting at scale. *The American Statistician*, 72(1):37–45, 2018.

[48] Steve Thomas. The history of wholesale electricity market design in great britain, 2001. http://www.psiru.org/reports/electricity-wholesale-market-uk-1990-2001.html.

[49] Fran Velde et al. Bitcoin: A primer. 2013.

[50] Can Wan, Zhao Xu, Yelei Wang, Zhao Yang Dong, and Kit Po Wong. A hybrid approach for probabilistic forecasting of electricity price. *IEEE Transactions on Smart Grid*, 5(1):463–470, 2014.

[51] Rafal Weron. *Modeling and forecasting electricity loads and prices: A statistical approach*, volume 403. John Wiley & Sons, 2007.

[52] Rafał Weron. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International journal of forecasting*, 30(4):1030–1081, 2014.

[53] Cort J Willmott and Kenji Matsuura. Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Climate research*, 30(1):79–82, 2005.

[54] Dr William Weimin Yoo. Mth6139 - time series 2020/1 lecture notes. https://qmplus.qmul.ac.uk/pluginfile.php/2541202/mod$_{r}esource/content/2/notes.pdf$.

[55] Eric Zivot. Garch, 2011. https://faculty.washington.edu/ezivot/econ589/ch18-garch.pdf.