Time Series Analysis of the Nord Pool Electricity Spot Market

Jared Newman

23 April 2018

Abstract

Electricity spot market prices are influenced by a variety of factors determining its equilibrium price. Some of these variables include market design, transmission congestion, fuel prices, bidding strategies, historical electricity prices, weather conditions and season, and equilibrium of supply and demand. The Nord Pool electricity spot prices from 1999 to 2007 are recorded hourly and show much volatility over time. Various features of time series analysis were discussed and implemented in this study in order to determine if a time series model is able to approximate the Nord Pool electrcity prices based on historical electricity prices. Regression, autoregressive (AR) models, autoregressive moving average (ARMA) models, autoregressive integrated moving average (ARIMA) models, autoregressive conditional heteroscedasticity (ARCH) models, and random walks are considered in this study for analysis of the hourly historical prices. Analysis showed that time series analysis of price is unable to satisfy the assumption of stationarity; and that more variables are needed in combination with a more dynamic process, such as a stochastic model, in order to adequately model the time series.

1 Introduction

The ability to successfully model a time series is one of great value as it can capture the behaviour of historical data for better prediction of the future. Furthermore, electricity spot prices, which easily form time series, are volatile because of their uncertainty and inelastic demand (Deng & Oren, 2006, p.940). Development of a time series model to approximate the behaviour of this data, however, is difficult due to the complexity of electricity spot market prices. Day-ahead forecasts of electricity prices can easily surpass the simplicity of a statistical model (Weron, 2014), so any reasonable model must match the complexity of the data. Many methods exist to form such a model, and those surveyed in this study include (1) regression, (2) autoregressive (AR) time series models, (3) autoregressive moving average (ARMA) models, (4) autoregressive integrated moving average (ARIMA) models, (5) autoregressive conditional heteroscedasticity (ARCH) models, and (6) random walks.

In this study a time series is taken from the Nord Pool spot market price of electricity in EUR/MWh for the dates spanning 1 January 1999 through 26 January 2007. The data is taken hourly without any omissions, resulting in 70,752 data points that compose a discrete and regularly spaced time series. As this data relates to spot market prices, it qualifies as a financial time series, though it relates to the prices of energy. This dataset has been analysed before by Erzgräber et al. (2008) through the Hurst exponent and long range correlations; short-term forecasts of the Nord Pool market were sought by Kristiansen (2014) using regression, myopic, and futures models; and the relationship between spot and future prices in the Nord Pool market was investigated by Botterud et al. (2009). This study builds on the literature by applying additional mathematical concepts to the dataset in order to best determine whether a time series model can be useful in predicting future trends in the electricity spot market.

This study is organised as follows. Section 2 introduces the dataset alongside the basic features and literature pertaining to time series. Section 3 applies the time series models to the data and tests it for the assumptions of time series, with an evaluation of the performance of the best model to determine whether a time series model can successfully approximate the Nord Pool electricity spot market data. Section 4 concludes the paper with remarks on possible alternative strategies for modelling the data using statistical methods.

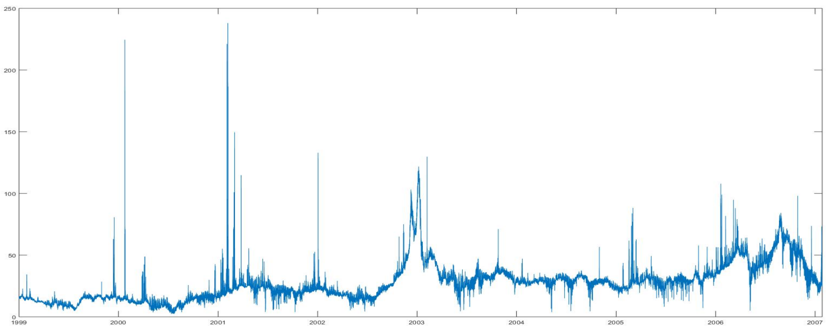
2 Dataset and time series modelling

The dataset as taken from the Nord Pool electricity spot market is affected by an equilibrium of supply and demand forces just as with any other market price. The Nord Pool operates as an exchange in which institutional participants trade power contracts for delivery to be made the following day. This makes it a 'day-ahead' market based on one-hour auctions for power contracts to be bought or sold spanning the 24 hours of the next day. At the conclusion of each auction, all orders are aggregated based on supply and demand to determine the equilibrium spot price (Erzgräber et al., 2008; Kristiansen, 2014).

Time series take many concepts utilised in technical analysis, such as a moving average, mean reversion, and seasonality; a combination of times series analysis and technical analysis have been shown to be superior to isolated analysis (Fang & Xu, 2003), but even this combination may fall short of the many factors affecting electricity prices. Some of these factors include the design of the market, transmission congestion, fuel prices, bidding strategies, historical electricity prices, weather conditions and season, and equilibrium of supply and demand as described above (Hu et al., 2009). The large number of factors affecting this series therefore leads to the hypothesis that a time series model is insufficient in capturing the complexity of the price movements.

The 70,752 data points comprising the range of dates under analysis is depicted below in Figure 1. Before discussing the results from the application of time series analysis on the data, definitions and assumptions of time series must be introduced.

Figure 1: Nord Pool electricity spot market prices, 1999-2007



A time series, which is a 'collection of observations taken sequentially in time', can be used for many applications including finance, economics, energy, and many others (Chan & Hang, 2002).

Definition 2.1. *A time series is a sequence of random variables* $\{X_t\}_{t=1,2,...}$

Time series can also be discrete or continuous, and either regularly or irregularly spaced (Chan & Hang, 2002). The ability to successfully comprehend a time-dependent series of data can lead to many insights on future behaviour. In reality, however, several important parameters and assumptions must be considered in order to successfully create a predictive model from a time series. One of the most important assumptions is stationarity, which must be tested for a valid model to be created (Nowotarski & Weron, 2018). Before defining stationarity, autocovariance is needed, which is denoted by the Greek letter γ and is calculated as:

$$\gamma_{(X_{t+\tau},X_t)} = cov(X_{t+\tau}, X_t), \tag{1}$$

for all indexes t and lags τ.

From autocovariance, stationarity can be determined.

Definition 2.2. *A time series X_t is weakly stationary if*
1. $E(X_t) = \mu_{X_t} = \mu < \infty$, that is, the expectation of X_t is finite and does not depend on t, and
2. $\gamma(X_{t+\tau}, X_t) = \gamma_\tau$, that is, for each τ the autocovariance of random variables $(X_{t+\tau}, X_t)$ does not depend on t (it is constant for a given lag τ).

Therefore, stationarity requires that there is no trend in the data over time and that the data is not correlated with itself (Horváth, 2013). Computing the logarithmic returns of the data over the time horizon typically results in a stationary time series, so this is the first step. This step is taken by transforming the data into a return function with the following formula:

$$r_\Delta(t) = ln\frac{x(t)}{x(t-\Delta)}. \tag{2}$$

These returns are useful in calculating the autocorrelation function (ACF) as described below. The logarithmic returns are first analysed for seasonality, which is the appearance of a cyclical pattern in a time series (Alonso & García-Martos, 2012). The returns are also tested for serial correlation, which is when error terms in a time series are correlated with each other over time. Significant serial correlation leads to the use of more complex models, beginning with an autoregressive (AR) model. If the original series is not stationary, it must be transformed appropriately before applying AR processes.

Autocovariance was introduced in developing the second constraint for achieving stationarity. This function is also useful in its part in the formula for autocorrelation, which determines whether ‘knowledge of the past’ has any value in predicting future observations. From the autocovariance function, the autocorrelation function (ACF) can be found as

$$\rho_X(\tau) = \frac{\gamma_X(\tau)}{\gamma_X(0)} = \mathrm{corr}(X_{t+\tau}, X_t) \quad \text{for all t,}\tau. \tag{3}$$

The ACF leads into the autoregressive (AR) processes, through which a present value in the series can be explained by a function of past values. An AR(p) process uses p past values.

Definition 2.3. *An autoregressive process of order p is written as*

$$X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \ldots + \Phi_p X_{t-p} + Z_t, \tag{4}$$

where Z_t is white noise, i.e., $Z_t \sim WN(0, \sigma^2)$ and Z_t is uncorrelated with X_s for each $s < t$.

An AR(1) process is thus an autoregressive process of order 1 given by

$$X_t = \Phi X_{t-1} + Z_t, \tag{5}$$

where $Z_t \sim WN(0, \sigma^2)$ and Φ is a constant. White noise is represented by uncorrelated, erratic behaviour, which is thus unpredictable if it represents a random variable (Zhang, 2016).

When an AR model is insufficient in modelling the series, a moving average (MA) model may be required. A moving average is a smoothing process by which data can take the average of a prespecified number of sequential values of a time series; this average therefore lags by as many values as are taken to form the first average. The MA can also be centred, and this centred moving average (CMA) takes terms both before and after the observation to find an average centred at the observation (Hyndman, 2010).

The use of either an AR model or an MA model may yet be insufficient, so their combination is the next step. An autoregressive moving average (ARMA) model takes properties of both AR and MA models by using parameters of p lags for AR and q periods for MA. An ARMA(p, q) model is denoted by

$$x_t = b_0 + b_1 x_{t-1} + \ldots + b_p x_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} \quad (6)$$
$$E(\varepsilon_t) = 0, \quad E(\varepsilon_t^2) = \sigma^2, \quad \mathrm{Cov}(\varepsilon_t, \varepsilon_s) = E(\varepsilon_t \varepsilon_s) = 0 \quad \text{for } t \neq s$$

where $b_1, b_2, \ldots b_p$ are the AR parameters and the $\theta_1, \theta_2, \ldots \theta_q$ are the MA parameters.

The complexity of ARMA models is further limited by the instability in the parameters, with small changes able to result in drastically different final parameter estimates. Additionally, there is no standard procedure for the choice of parameters.

Autoregressive integrated moving average (ARIMA) models use both parameters in the ARMA model and add an integration term (I) accounting for non-stationarity of the series (Palma, 2016). In an ARIMA(p,d,q) model, d is a nonnegative integer indicating degree of integration in the time series; it is commonly 1, as a series is stationary if $d = 0$ (Diebold, 2007; Ling et al., 2015).

If a series can be successfully transformed to satisfy the assumptions of stationarity, the residuals must be tested for autoregressive conditional heteroscedasticity (ARCH). Heteroscedasticity is the dependence of the variance of the error term on time. The squared residuals are regressed on a lagged value of the squared residual; a test for ARCH(1) uses a lag of 1 value. The residuals do not show ARCH if the coefficient on the squared lagged residual does not significantly differ from zero (Kokoszka et al., 2017). Evidence of ARCH results in incorrectly specified standard errors and the ability to predict variance of the errors over time (Fryzlewicz, 2007).

If a series cannot be transformed to satisfy the assumptions of stationarity, it may be a random walk. An ARIMA(*p,1,q*) process behaves similar to random walks (Diebold, 2007; Zhou et al., 2006).

Definition 2.4. *A random walk is a series for which a value is determined by the value in the previous period plus an unpredictable random error. A random walk is described by the following:*

$$x_t = x_{t-1} + \varepsilon_t, \quad E(\varepsilon_t) = 0, \quad E(\varepsilon_t^2) = \sigma^2, \tag{7}$$
$$Cov(\varepsilon_t, \varepsilon_s) = E(\varepsilon_t \varepsilon_s) = 0 \quad \textit{if } t \neq s$$

A random walk therefore has an error term with constant variance that is uncorrelated with previous error terms. Previously discussed regression methods cannot be used to estimate an AR(1) model when the time series is a random walk.

The use of a random walk in a non-stationary series is allowable since one of the assumptions of stationarity is a constant and finite value, and all stationary time series have a finite mean-reverting level. A mean-reverting level is the level that a series tends to approach, whether it is above or below this level and is calculated as

$$x_t = \frac{b_0}{1 - b_1} \tag{8}$$

A random walk has no mean-reverting level. This is derived from the idea that if x_t is at its mean-reverting level, $x_t = b_0/(1 - b_1)$ as shown in (8). In a random walk, $b_0 = 0$ and $b_1 = 1$, so $b_0/(1 - b_1) = 0/0$ which is undefined. A random walk has a variance of

$$x_t = (t - 1)\sigma^2$$

for any period t, which approaches infinity as t increases. This proves that a random walk cannot be a stationary time series, as stationarity requires finite variance.

A random walk can be made stationary by differencing the series. First-differencing simply takes the difference between the value of an observation and the previous observation, and is suggested in the time series literature when significant serial correlation is found in the linear trend model before implementing a more complex AR model. The first-difference results in

$$\begin{gathered} y_t = x_t - x_{t-1} = \varepsilon_t, \quad E(\varepsilon_t) = 0, \quad E(\varepsilon_t^2) = \sigma^2, \\ \mathrm{Cov}(\varepsilon_t, \varepsilon_s) = E(\varepsilon_t \varepsilon_s) = 0 \quad \text{for } t \neq s \end{gathered} \tag{9}$$

The expected error thus becomes zero and the best forecast in period

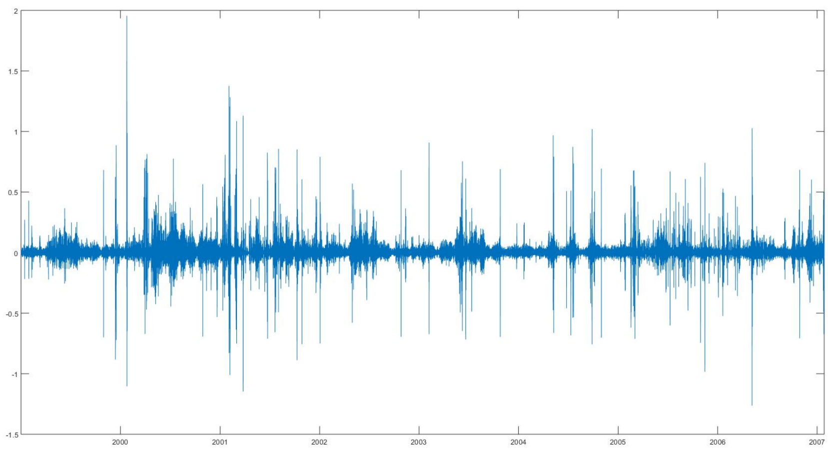
$t-1$ based on the first-difference is zero. The first-differenced random walk is stationary as it is now an AR(1) model with $b_0 = 0$ and $b_1 = 0$, which results in a mean-reverting level of $b_0/(1-b_1) = 0/1 = 0$. Additionally, the variance of this series is the variance of the error term and is $\text{Var}(\varepsilon_t) = \sigma^2$. This proves that both the mean and variance are both constant and finite in all periods, the first-differenced random walk is stationary and can be modelled with linear regression. Using an AR(1) model of the first-differenced random walk does not aid in forecasting future values but only leads to the conclusion that the original series is a random walk. Random walks can be extended to include a drift parameter; this models trend as it grows each period, on average, by the drift δ. This is a model of stochastic trend and is applicable here as stocks often follow stochastic movements (Diebold, 2007).

3 Time series analysis and evaluation of model

Analysis of the time series begins with a return to the plot shown in Figure 1. As is plainly visible, there are several spikes in the prices and a general upward trend. Simple least-square regression analysis showed that the trend is significantly different from zero. The Durbin-Watson statistic for this linear regression is 0.043, indicating strong and positive serial correlation. With positive serial correlation, estimates of the standard error are smaller than in reality (Williams, 2015). This Durbin-Watson value supports the claim that the data is not best modelled by a linear regression. Further regression analysis found that a linear regression explains more of the variation in the data than does an exponential regression.

The first attempt at making the series stationary was application of (2) to produce hourly logarithmic returns. The result of this transformation is shown below in Figure 2, which depicts these returns.

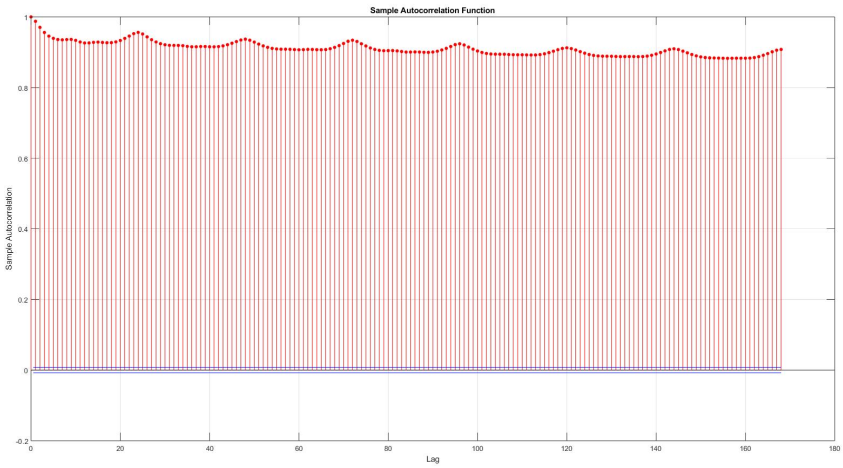
Figure 2: Hourly logarithmic returns of spot market data



Tests for stationarity additionally determine whether a series is covariance stationary, meaning that the variance does not change significantly over time (Pagan & Schwert, 1990; Xiao & Lima, 2007). The logarithmic returns in Figure 2 show signs of volatility clustering but without pattern, and spikes in the returns data show that variance does not depend on time. The Durbin-Watson statistic is invalid for models with lagged variables, and in logarithmic returns this lag is evident (Nerlove & Wallis, 1966). Instead the Ljung-Box (LB) test is employed to determine if serial correlation is present in the residuals (Lee & Park, 2016). Applied to the logarithmic returns, the LB test indicates that there is not enough evidence to reject the null hypothesis that the return residuals are not autocorrelated; in other words, the logarithmic returns are not serially correlated. This is reinforced by a space-time separation plot provided by Erzgräber et al. (2008), which only indicates weak 24-hour periodicity. The unit root test (Dickey & Fuller, 1979; Pagan & Schwert, 1990) and the KPSS test (Kwiatkowski et al., 1992; Xiao & Lima, 2007) are commonly used tests for stationarity in financial data, but are not used here as the LB test provides adequate results. Additionally, it must be noted that seasonality, which appeared possible in the original data in Figure 1, is no longer apparent in the returns data in Figure 2. An explanation for this could be the variation in months of the year during which the prices reached their annual peak. The data therefore does not require to be de-seasonalised.

Examination of the autocorrelation functions is next. It is worthwhile to first look at the ACF of the original data, as depicted in Figure 3 below. The first 168 lags are shown to visualise the lags for a full week. The first autocorrelation is at 0.9876 and does not decline rapidly, as is required in a stationary series; additionally, this function shows the high levels of autocorrelation of the data, never breaching 0.88. These ACF values are far greater than the acceptable range of autocorrelations.

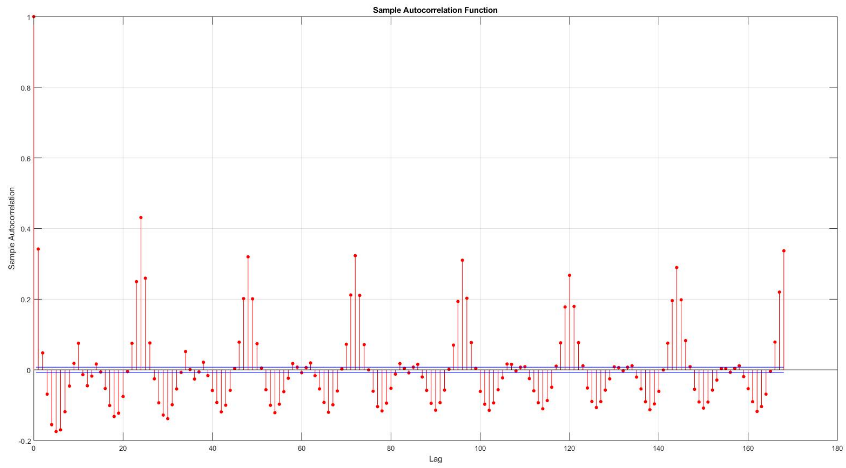
Figure 3: Autocorrelation function of price data



Acceptable ranges for ACFs are found by using a t-test on the standard error of the residual correlation of $1/\sqrt{(n)}$, where n is the number of observations (Brockwell & Davis, 2016; Diebold, 2007). The autocorrelation is divided by the standard error, which is computed as $1/\sqrt{(70,752)} = 0.00377$ for this dataset. Using 5% significance limits, the t value must be less than 0.00739, which is a strict requirement to satisfy the assumptions for no serial correlation. Due to the large number of observations, this bound is relatively small and difficult to achieve, but the severity of autocorrelations in the original series is far from this requirement.

The ACF for the logarithmic returns shows that the series retains autocorrelations higher than are within the acceptable range as shown in Figures 4 and 5 below. Figure 4 depicts the ACF and Figure 5 depicts the partial autocorrelation function (PACF); this latter function controls for the values of the series at all shorter lags by removing linear dependence between intermediate variables (McLeod & Zhang, 2006).

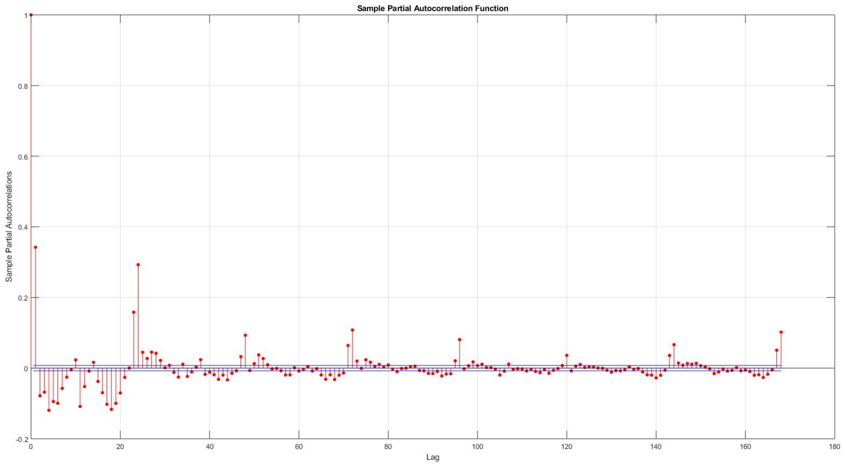
Figure 4: ACF of logarithmic returns



The ACF illustrates the higher significance of each 24-hour cycle affecting future observations, such that the highest autocorrelations appear around 24 hours, 48 hours, and so on with gradual decline. This decline is common for autocorrelation functions, but time series models that do not suffer from serial correlation do not show this severity in breaches of the acceptable range. Moving average models find that the corresponding ACF drops to zero after q values, with q representing the number of observations used in calculating the moving average. The ACF in Figure 4 shows that a moving average model does not capture the series, as its autocorrelation never remains at zero, and few values within the acceptable range are obtained.

The PACF provides insight on the most influential lagged values within the past 24 hours. The extreme difficulty in reducing the autocorrelations to the acceptable range will prove the complexity of the development

Figure 5: PACF of logarithmic returns



of a suitable time series model.

It is clear from the high autocorrelations in Figure 4, as shown in the first lag, that an AR(1) of the logarithmic returns of the data will be insufficient in modelling the series. Additionally, it can be deduced before testing further models that a large number of lags in the autocorrelation functions require integration with the AR process before it can possibly find all autocorrelations with the acceptable range. Although it may be intuitive to test various AR processes utilising combinations of lag 1, 24, 48, 72, 96, 120, and 144, none of these combinations satisfy the range given for the autocorrelation functions, and serial correlation remains evident in all cases, even at higher lags. An AR process alone is therefore insufficient to model the data, so the model is incorrectly specified.

As a 24-hour periodicity is expected in hourly data, the time series feature of a moving average (MA) is approached next. Using the logarithmic returns, several combinations of moving averages were taken on the data and the results showed no improvement to the model. Both the 24-period MA and 24-period CMA were expected to aid in significantly reducing the peaks in the data, but this could not be accomplished even with larger moving averages extending to a 1-week (168-period) MA. This is due to the times that witnessed consistently elevated prices that eventually returned to normal values, as can be seen in Figure 1 around January 2003. Furthermore, MA is more commonly seen when the ACF depicts a negative autocorrelation for the first lag, but this is not the case here (Nau, 2017). It is most likely that a moving average is not helpful in the creation of any time series model for this data.

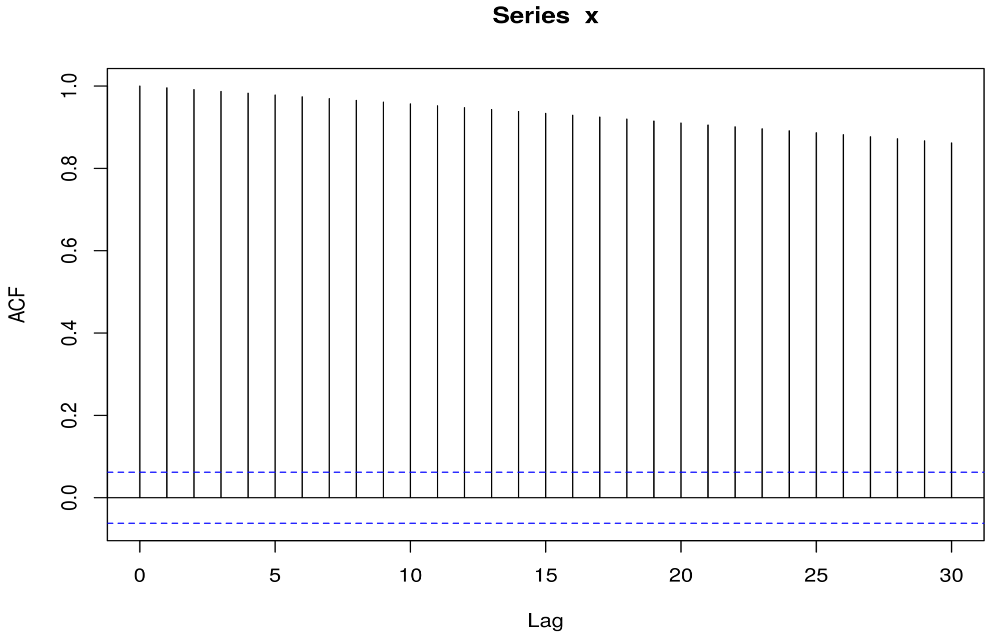
Combination of an AR model and an MA model is therefore the next step, but this assumes that the series is stationary as this is the same as an ARIMA model in which $d = 0$. Therefore, tests move forward to the ARIMA model with parameters estimated for p and q, with d fixed at 1 for all tests.

The ACF of the log returns shows that the strongest initial autocorrelations appear in the first six lags before approaching zero in a sinusoidal manner; this suggests an AR(6) process. The periodicity of the strongest autocorrelations is 24; this suggests an MA(24) process. Using an ARIMA model with parameters $(p, d, q) = (6, 1, 24)$ does not result in the autocorrelations remaining within the specified bounds of the target range. Inclusion of a seasonal component in an ARIMA model using the notation ARIMA$(6, 1, 24)(1, 0, 0)[24]$ integrates the 24^{th} lag, but this model is insufficient as well. A majority of the autocorrelations of this model remain outside the acceptable range. ARIMA that omitted MA found improvements in the results, as might be expected based on the conclusions after isolated MA tests were performed. Issues remained in an ARIMA model using only AR and I, however.

The process of increasing p in the AR component only increased the autocorrelation of lag $p + 1$ and gradually increased autocorrelations of all lags less than p. This applied to lags greater than 6 as well, and is caused by the high autocorrelations in the original time series that do not decline significantly with increasing lags, reflected as a high dependency on recent values that cannot be made independent through AR processes. The possibility of a random walk to model the series is therefore considered.

It is useful to first examine the ACF of a *simulated* random walk; this is shown in Figure 6 below. Here, the autocorrelations begin near 1 and gradually decline, though at a faster rate than the ACF of the original spot market data. This supports the time series as a random walk.

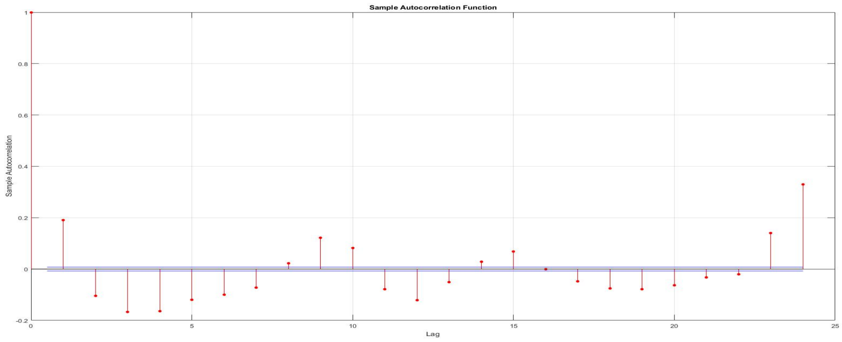
Figure 6: Sample ACF of a random walk (QuantStart, 2018)



The ACF of the first-differenced series in Figure 7 below shows that the autocorrelations do not fall within the given bounds after any number of lags. This is less supportive of the random walk model. The autocorrelations appear to follow a sinusoidal curve as with the autocorrelations of the log returns, indicating a definite pattern as opposed to the appearance of white noise as is representative of the first-difference of a series that is a random walk.

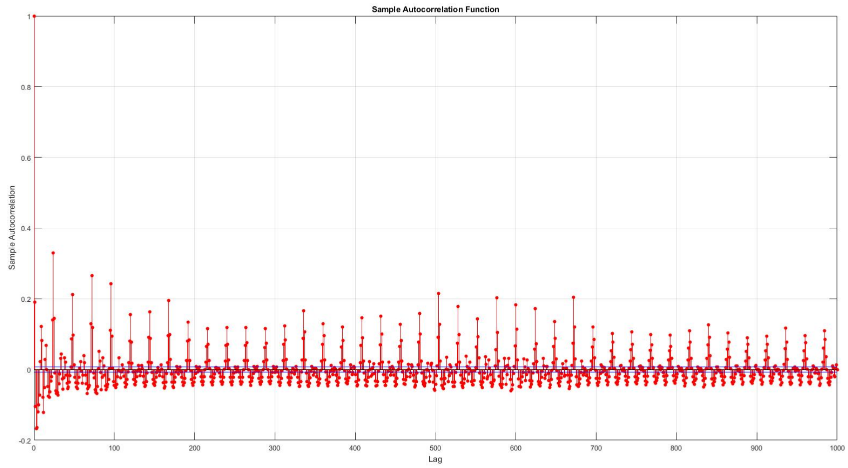
As the ACF is often analysed extending to a larger proportion of the number of observations than in Figure 7, an ACF with 1000 lags shows another pattern. In Figure 8, this ACF is depicted with 1000 lags. This figure shows the clear autocorrelations recurring every 24 periods, but that besides large absolute autocorrelations in the first several lags there is little

Figure 7: ACF of the first-difference of spot market data, 24 lags



pattern to be found. Nevertheless, the autocorrelations persist well above the bounds, which describes the high volatility of the time series.

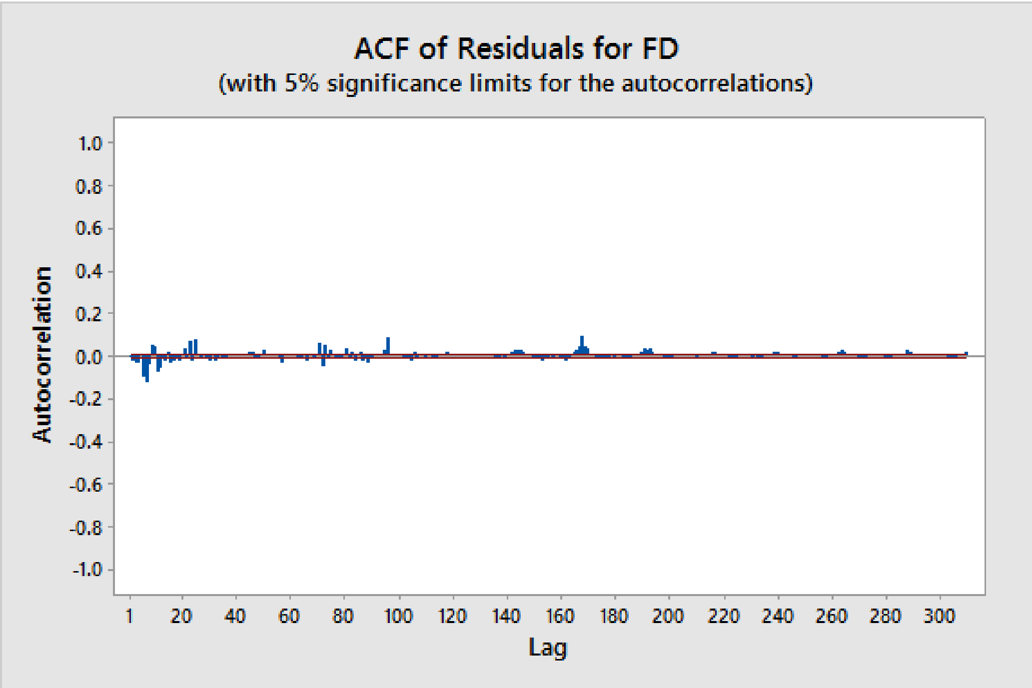
Figure 8: ACF of the first-difference of spot market data, 1000 lags



As the series cannot be said with confidence to follow a random walk, the best model is most likely one that utilises autoregressive processes in combination with a first-difference; the improvement made by a moving average is uncertain.

The model that minimised the autocorrelations appeared to be an ARIMA model utilising a first-difference, five lags in the AR component, and three 24-period seasonal lags in the AR component. This appears as the ACF in Figure 9 below. As discussed above, the addition of lags beyond five caused the subsequent lag to significantly increase and for previous lags to slightly increase. An ARIMA model with 23 AR lags finds no uniquely severe lags, but that all that remain outside the significance limits for the ACF and therefore does not improve upon this model.

Figure 9: ACF of the best model using ARIMA



This model suggests the following formula:

$$
\begin{aligned}
x_t = \quad & 0.9284x_{t-1} - 0.1536x_{t-2} - 0.1491x_{t-3} \\
& -0.1308x_{t-4} - 0.1087x_{t-5} + 0.2586x_{t-24} \\
& +0.0787x_{t-48} + 0.1828x_{t-72} + \varepsilon_{t-1}
\end{aligned}
\tag{10}
$$

The model therefore incorporates the first-difference in the highly correlated first lag (calculated by taking $1 - 0.0716$ as is the coefficient seen in Table 1 in the Appendix); the following four lags are all negative and smaller in magnitude, and the seasonal components are all positive and slightly greater in magnitude than the second through fifth lags.

Although this is the best model found using time series analysis and specifically ARIMA modelling, it does not capture the complexity of the time series and is therefore insufficient for practical use. The ACF in Figure 9 shows the consistently high autocorrelations even when one of the more complex time series models is applied, so it does not satisfy the constraints of stationarity.

4 Conclusion

The Nord Pool electricity spot market data from 1999 to 2007 shows high degrees of autocorrelation and therefore pose significant problems in modelling the series through time series analysis. The logarithmic returns of the series were not serially correlated, but retained high levels of autocorrelation in nearly all lags. Through further analysis by autoregressive (AR) processes, moving average (MA) models, and autoregressive integrated moving average (ARIMA) models, the autocorrelations were unable to be pushed to within the bounds necessary for the model to become stationary. Since stationarity was not achieved through this time series analysis, it was not tested for autoregressive conditional heteroscedasticity (ARCH). First-differencing the original series determined that the data do not sufficiently follow a random walk, as per the consistently elevated autocorrelations.

The ARIMA models were therefore relied on for the most accurate model for the series. This optimal model used a first-difference, five AR lags, and three 24-period seasonal AR lags, but did not manage to fulfill the assumptions of stationarity. It therefore cannot be used for practical purposes.

More advanced models are therefore required to suitably fit the data. As a random walk was considered in this investigation, a dynamic model such as a stochastic process may better model the series. More success has been achieved by research that utilised additional variables, however, such as Kristiansen (2014) who incorporated independent variables of inflow levels and reservoir levels in combination with a regression model that achieved roughly 7.5% mean absolute percentage error. Use of the Hurst exponent by Erzgräber et al. (2008) found the ability to capture up to 20% of the numerical variation, though they acknowledged that a single exponent will not explain the complexity of a real-world process. Ergemen et al. (2015) looked at the variable of supply elasticity and its co-integration with time for aid in future research. Based on the insufficiency in time series features applied here, it is clear that more variables must be considered besides spot market price in combination with a stochastic or other more complex model.

Appendix

Table 1. Minitab output for final ARIMA model

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
AR 1	0.0716	0.0038	19.03	0.000
AR 2	-0.1536	0.0037	-41.32	0.000
AR 3	-0.1491	0.0037	-40.05	0.000
AR 4	-0.1308	0.0037	-35.21	0.000
AR 5	-0.1087	0.0037	-29.09	0.000
SAR 24	0.2586	0.0037	69.64	0.000
SAR 48	0.0787	0.0038	20.55	0.000
SAR 72	0.1828	0.0037	49.03	0.000
Constant	0.000071	0.007686	0.01	0.993
Mean	0.00010	0.01089		

```
Number of observations:  70750
Residuals:     SS =  295702 (backforecasts excluded)
               MS =  4  DF = 70741
```

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	3119.4	3837.7	4404.9	4494.7
DF	3	15	27	39
P-Value	0.000	0.000	0.000	0.000

References

Botterud, A., Kristiansen, T., & Ilic, M. (2010). The relationship between spot and futures prices in the Nord Pool electricity market. *Energy Economics*, [online]. Available at: `https://www-sciencedirect-com.ezproxy.library.qmul.ac.uk/science/article/pii/S0140988309002278` [Accessed 19 October 2017].

Brockwell, P. & Davis, R. (2016). *Introduction to Time Series Forecasting*. Switzerland: Springer, p.395.

Deng, S.J., & Oren, S.S. (2006). Electricity derivatives and risk management. *Energy*, [online]. Available at: `https://www.sciencedirect.com/science/article/pii/S0360544205000496` [Accessed 1 March 2018].

Ergeman, Y., Haldrup, N., & Rodríguez-Caballero, C. (2015). Common long-range dependence in a panel of hourly Nord Pool electricity prices and loads. *Department of Economics and Business Economics, Aarhus University*, [online]. Available at: `ftp://ftp.econ.au.dk/creates/rp/15/rp15_58.pdf` [Accessed 18 March 2018].

Erzgräber, H., Strozzi, F., Zaldívar, J., Touchette, H., Gutiérrez, E., & Arrowsmith, D. (2008). Time series analysis and long range correlations of Nordic spot electricity market data. *Physica A*, [online]. Available at: `https://www-sciencedirect-com.ezproxy.library.qmul.ac.uk/science/article/pii/S0378437108006778` [Accessed 19 October 2017].

Han, C., Phillips, P., & Sul, D. (2012). Lag Length Selection in

Panel Autoregression. *University of Texas at Dallas*, [online]. Available at: http://www.utdallas.edu/~d.sul/papers/BIC15_Table.pdf [Accessed 19 March 2018].

Horváth, L., Kokoszka, P., & Rice, G. (2013). Testing stationarity of functional time series. *Journal of Econometrics*, [online]. Available at https://www.sciencedirect.com/science/article/pii/S0304407613002327?via%3Dihub [Accessed 1 March 2018].

Hu, L., Taylor, G., Wan, H., Irving, M. (2009). A Review of Short-term Electricity Price Forecasting Techniques in Deregulated Electricity Markets. *Proceedings of the Universities Power Engineering Conference*, [online]. Available at: http://tarjomefa.com/wp-content/uploads/2017/08/7574-English-TarjomeFa.pdf [Accessed 19 October 2017].

Hyndman, R.J. (2010). Moving averages. *International Encyclopedia of Statistical Sciece*, [online]. Available at: https://robjhyndman.com/papers/movingaverage.pdf [Accessed 8 March 2018].

Kokoszka, P., Rice, G., & Shang, H. (2017). Inference for the autocovariance of a functional time series under conditional heteroscedasticity. *Journal of Multivariate Analysis*, [online]. Available at: https://www.sciencedirect.com/science/article/pii/S0047259X17303512?via%3Dihub [Accessed 8 March 2018].

Kristiansen, T. (2012). Forecasting Nord Pool day-ahead prices with an autoregressive model. *Energy Policy*, [online]. Available at: https://www-sciencedirect-com.ezproxy.library.qmul.ac.uk/science/article/pii/S0301421512005381 [Accessed 19 October 2017].

Kristiansen, T. (2014). A time series spot price forecast model for the Nord Pool market. *International Journal of Electrical Power & Energy Systems*, [online]. Available at: https://www-sciencedirect-com.ezproxy.library.qmul.ac.uk/science/article/pii/S0142061514001094 [Accessed 19 October 2017].

Lee, T. & Park, C. (2016). Tests for serial correlation in mean and variance of a sequence of time series objects. *Journal of Statistical Computation and Simulation*, [online]. Available at: https://faculty.franklin.uga.edu/cpark/sites/faculty.franklin.uga.edu.cpark/files/LB_test_JSCS_published.pdf [Accessed 21 March 2018].

Ling, S., McAleer, M., & Tong, H. (2015). Frontiers in Time Series and Financial Econometrics: An overview. *Journal of Econometrics*, [online]. Available at: https://www.sciencedirect.com/science/article/pii/S0304407615000974?via%3Dihub [Accessed 19 February 2018].

McLeod, A.I. & Zhang, Y. (2005). Partial Autocorrelation Parameterization for Subset Autorregression. *Journal of Time Series Analysis*, Vol. 27, No. 4, [online]. Available at: https://pdfs.semanticscholar.org/11aa/f7d30ea675ca5ce7f93d731acbc48dedb2e1.pdf [Accessed 8 March 2018].

Nerlove, M. & Wallis, K. (1966). Use of the Durbin-Watson Statistic in Inappropriate Situations. *Econometrica*, Vol. 34, No. 1, [online]. Available at: http://www.jstor.org/stable/1909870?read-now=1&seq=1#page_scan_tab_contents [Accessed 21 March 2018].

Nowotarski, J. & Weron, R. (2018). Recent advances in electricity price forecasting: A review of probabilistic forecasting. *Re-*

newable and Sustainable Energy Reviews, [online]. Available at: https://www-sciencedirect-com.ezproxy.library.qmul.ac.uk/science/article/pii/S1364032117308808 [Accessed 8 March 2018].

Pagan, A. & Schwert, G. (1990). Testing for covariance stationarity in stock market data. *Economics Letters*, Vol. 33, No. 2, [online]. Available at: https://ac.els-cdn.com/016517659090163U/1-s2.0-016517659090163U-main.pdf?_tid=cf1c97d1-42f7-4262-808f-10c73cf7f2d6&acdnat=1524244298_5b46b450c39002182ac55cbaf5cfce1c. [Accessed 8 March 2018].

Palma, W. (2016). *Time Series Analysis.* John Wiley & Sons, Incorporated, p.1097.

Pettit, L. (2017). *Time Series*, lecture notes. Queen Mary, University of London.

QuantStart (2018). *White Noise and Random Walks in Time Series Analysis.* Available at: https://www.quantstart.com/articles/White-Noise-and-Random-Walks-in-Time-Series-Analysis [Accessed 21 March 2018].

Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, [online]. Available at: https://www-sciencedirect-com.ezproxy.library.qmul.ac.uk/science/article/pii/S0169207014001083 [Accessed 17 October 2017].

Xiao, Z. & Lima, L. (2007). Testing Covariance Stationarity. *Econometric Reviews*, Vol 26, No. 6, [online]. Available at: https://www.tandfonline.com/doi/pdf/10.1080/07474930701639080?needAccess=true [Accessed 19 March 2018].

Zhang, X. (2016). White noise testing and model diagnostic checking for functional time series. *Journal of Econometrics*, Vol. 194, No. 1, [online]. Available at: https://www.sciencedirect.com/science/article/pii/S0304407616300604 [Accessed 19 March 2018].

Zhou, M., Yan, Z., Ni, Y.X., Li, G., & Nie, Y. (2006). Electricity price forecasting with confidence-interval estimation through an extended ARIMA approach. *IEE*, [online]. Available at: http://digital-library.theiet.org/content/journals/10.1049/ip-gtd_20045131 [Accessed 1 March 2018].