

Lappeenranta University of Technology
Faculty of Technology
Department of Mathematics and Physics

Analysis of outliers in electricity spot prices with example of New England and New Zealand markets

The topic of this Thesis was approved
by the Department of Mathematics and Physics

October 2, 2008

Supervisors: Prof. Ph.D. Heikki Haario and Ph.D. Tuomo Kauranne.
Examiners: Prof. Ph.D. Heikki Haario and Ph.D. Tuomo Kauranne.

Lappeenranta, October 21, 2008

Matylda Jabłońska
Punkkerikatu 5 D 57
53850 Lappeenranta
matylda.jablonska [at] lut.fi

Abstract

Lappeenranta University of Technology
Department of Mathematics and Physics

Matylda Jabłońska

**Analysis of outliers in electricity spot prices with example
of New England and New Zealand markets**

Master's thesis

2008

63 pages, 57 figures, 16 tables

Supervisors: Prof. Ph.D. Heikki Haario and Ph.D. Tuomo Kauranne.

Key words: time series, electricity spot price, spike, Discrete Fourier Transform

Electricity spot prices have always been a demanding data set for time series analysis, mostly because of the non-storability of electricity. This feature, making electric power unlike the other commodities, causes outstanding price spikes. Moreover, the last several years in financial world seem to show that 'spiky' behaviour of time series is no longer an exception, but rather a regular phenomenon. The purpose of this paper is to seek patterns and relations within electricity price outliers and verify how they affect the overall statistics of the data. For the study techniques like classical Box-Jenkins approach, series DFT smoothing and GARCH models are used. The results obtained for two geographically different price series show that patterns in outliers' occurrence are not straightforward. Additionally, there seems to be no rule that would predict the appearance of a spike from volatility, while the reverse effect is quite prominent. It is concluded that spikes cannot be predicted based only on the price series; probably some geographical and meteorological variables need to be included in modeling.

Acknowledgements

This study would not have been carried out if not for Scholarship from Lappeenranta University of Technology and Grant from Fortum Oy company.

I would like to express my gratitude for valuable supervising support from Ph.D. Tuomo Kauranne and Piort Ptak, who stated directions of this work.

Szczególne podziękowania składam moim najbliższym – rodzicom, bratu i przyjaciołom – którzy zawsze mnie wspierali i pomagali w podejmowaniu najważniejszych życiowych decyzji.

.....

Contents

1	Introduction	2
2	Theoretical background	3
2.1	New England and New Zealand electricity markets	3
2.2	Classical time series approach	6
2.2.1	Basic models - ARMA	6
2.2.2	Preparing Box-Jenkins models	7
2.3	ARCH/GARCH modeling	7
3	Statistical analysis of NEPool and Otahuhu data	8
3.1	General information and basic statistics	8
3.2	Normality	10
3.3	Inner dependencies	12
3.4	Discrete Fourier transform smoothing	16
3.5	Week days analysis	21
4	Analysis of outliers	30
4.1	Spikes with respect to original price series	30
4.2	Spikes with respect to price series smoothed by DFT	37
4.3	Outliers in return series	45
4.4	Predictability of spikes based on seasons and price volatility	50
4.4.1	Outliers vs. year seasons	50
4.4.2	Outliers vs. price volatility	55
4.5	What do electricity suppliers really earn on spikes?	58
5	Conclusion and future work	60
	References	62

1 Introduction

Electricity spot prices have always been a demanding data set for time series analysis. One of the main features that differentiate electricity from other stocks and commodities traded on stock exchanges is that it cannot be stored in warehouses. Therefore, most of techniques for stock management do not apply to power exchange. The limits on delivery emerge from supply grid capacity constraints. If transmission is not limiting electricity trading, the electricity delivery takes place normally and prices are reasonably stable. If there appears a congestion in some region, and thereby the marginal congestion cost becomes active (see Hadsell and Shawky [13]), electricity is supplied only to those consumers who pay more. The other crucial feature of electricity prices is their high overall volatility. These issues have been widely studied for years.

Nowadays there are plenty of methods for price and price return forecasting; one of the most common ways to do that is the classical time series approach. These kinds of analyses are very important in every branch of industry including electricity pricing. Different corporations try to find models explaining electricity price behaviour. Since it is hard to perfectly represent a given phenomenon in a way that it would be faultless with predictions, every modeling process consists of attempts to find a compromise between proper representation of historical data and reasonable forecasting ability. One could ask why to try any forecasting at all, if it is so difficult to do it properly. In fact, the answer is not straightforward. But the more attempts we make to predict something, the higher the probability that we will succeed some day. Training time series give more practice in considering different approaches of modeling.

In case of electricity prices, many researchers try to focus on sources of the high variability of prices. Hinz [10] stated that prediction of sudden and significant changes in electricity prices may be formed based on proper statistical analysis and forecasting of electricity demand. Different papers cover various forecasting approaches and methods' comparison. For example Conejo *et al.* [11] show that time series analysis outperforms neural networks and wavelet techniques in generating day-ahead predictions. There have also been studies carried out on specific features like mean-reversion of electricity prices (see Huisman *et al.* [15]). Moreover, methods like regime-switching models are being estimated more and more often (see Karakatsani and Bunn [18], Kanamura and Ohashi [17]). It is also discussed that the transition probabilities in reality are not constant in the model within the whole time horizon. One of the most important issues in electricity price analysis and forecasting is to be able to predict occurrence of spikes. Kanamura and Ohashi [16] proposed a structural model, which is able to predict spikes up to some level as resulting mostly from demand seasonality.

So far nobody has succeeded in creating a perfect tool for electricity price prediction, since there is a high level of randomness in these kinds of series. However, some patterns can be identified. Therefore, the purpose of this study is to investigate two electricity

price time series: New England Pool (NEPool) and Otahuhu (one of New Zealand Nodes). Both data sets were found on the Internet. The original series were different in time horizon. However, for the purpose of this study, exactly the same scope was taken for the analysis. The approach taken in this study employs not only techniques of signal smoothing and classical time series procedures, but also performs an extensive spike investigation. We try to find dependencies and patterns within the outliers' occurrence by verifying their autocorrelation and correlation with price volatility changes. On the other hand, an analysis of the data with removed outliers is also carried out. This paper can act as a basis for building a dual model of electricity spot prices suggested by Ptak *et al.* [19].

The structure of this Thesis is as follows. The next section briefly goes through the theoretical background for the problem: specificity of electricity price data, classical Box-Jenkins time series approach and definitions of example heteroscedastic models. Section 3 covers statistical analysis of electricity price and price return series for both original and DFT-smoothed data. Section 4 moves on to investigation of outliers specifically. Finally, section 5 concludes and gives proposals for future work.

2 Theoretical background

2.1 New England and New Zealand electricity markets

New England and New Zealand electricity markets are of a slightly different character. NEPool is a not-for-profit company stating the hour-ahead and day-ahead system prices for regional electricity trading. Their role is to state the prices such that electric power supply and demand match. The New Zealand market works as a combination of state-owned, trust-owned and public companies. "The main participants are seven generator/retailers who trade at 244 nodes across the transmission grid. The generators offer their plant at grid injection points and retailers bid for electricity offtake at grid exit points" [22]. The data sets are also different from geographical point of view. New England is a part of continent with ocean shore just on one side of the region. New Zealand is an island country surrounded by seas and, therefore, exposed more to oceanic weather changes.

One of the crucial aspects in the New Zealand grid is that most of electricity production takes place in the South of the country (for the Southern Island electricity supply grid see Figure 2), whereas the highest demand is mostly in the inhabited and developed regions in the North (for the Northern Island grid see Figure 1).

Figure 3 presents a map of New England Pool with an example day-ahead price situation on the market. The print screen comes from the NEPool web page [20], where the data is refreshed every 5 minutes.

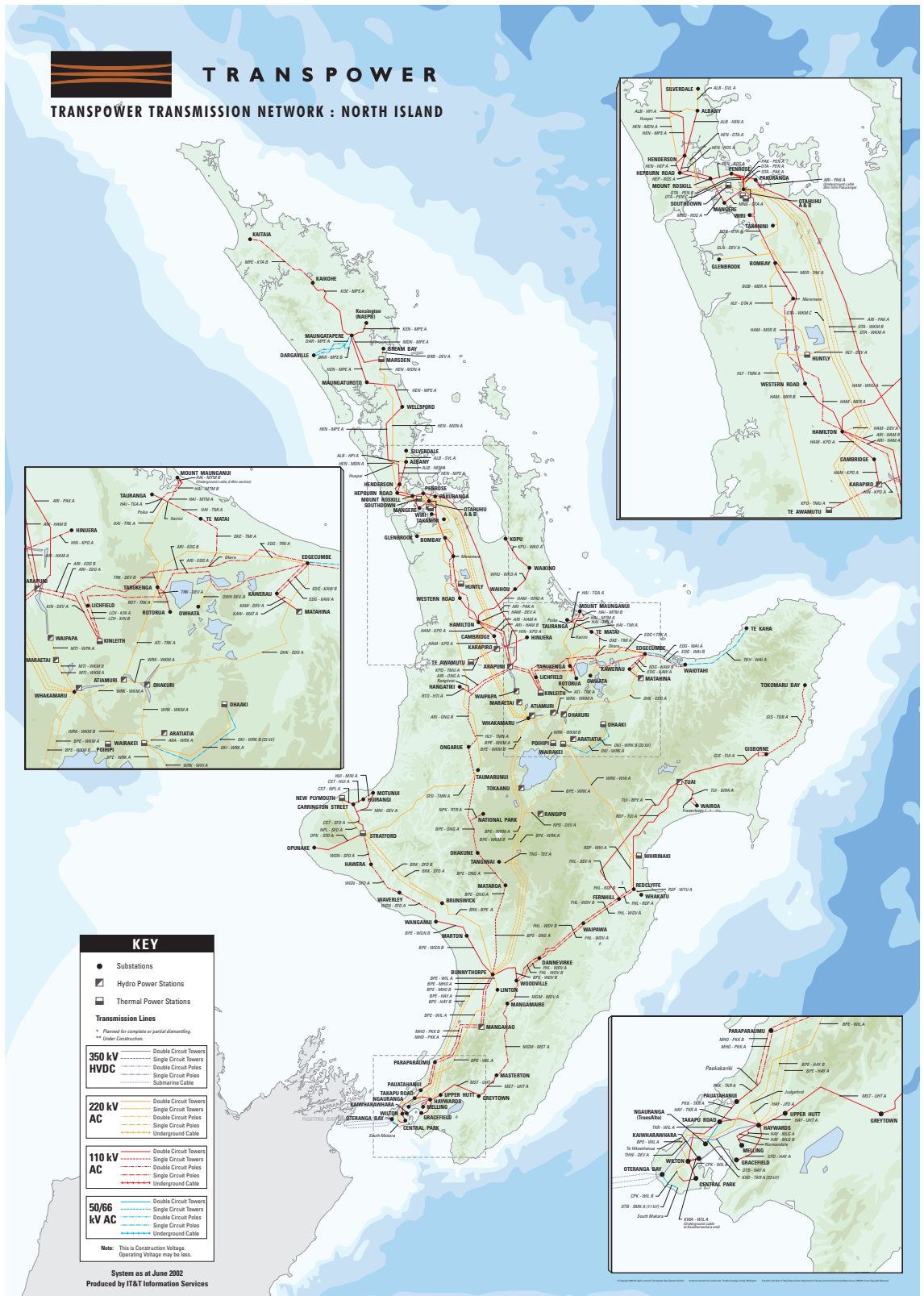


Figure 1: Northern New Zealand supply grid [21].



Figure 2: Southern New Zealand supply grid [21].

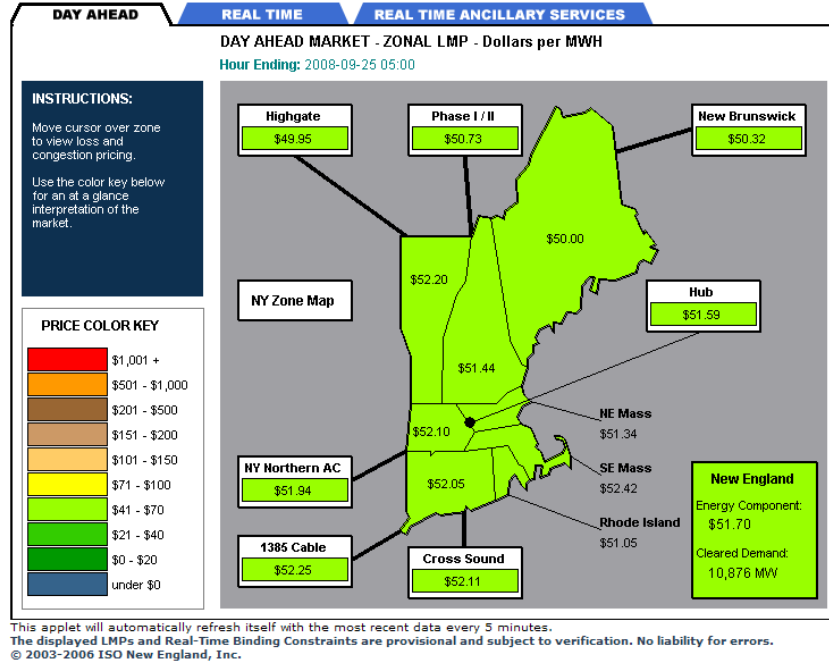


Figure 3: NEPool map with example day ahead prices [20].

The data set analyzed in this study comes from Otahuhu – a Node from Auckland region in the very north of the New Zealand Northern Island.

2.2 Classical time series approach

2.2.1 Basic models - ARMA

A time series is a sequence of observations based on a regular timely basis, *e.g.* hourly, daily, monthly, annually, *etc.* The classical time series analysis (see Box *et al.* [6]), partially utilized in this study, covers fitting autoregressive (AR) and moving average (MA) models. Basically, it considers analyzing the data to find dependencies between current and historical observations. These models can also be extended by associated heteroscedastic models. The first proposed ones were: autoregressive conditional heteroscedasticity known as ARCH (see Engle [2]) and generalized autoregressive conditional heteroscedasticity, namely GARCH (see Bollerslev [3]). A wide overview of modern variations of these models was made by Tsay [12].

The most common models are autoregressive (AR) and moving average (MA). The former represents a current observation in terms of lagged past realizations of a given process. An autoregressive model of order r , *i.e.* $AR(r)$, is introduced by the following definition:

- $x_t = C + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_n x_{t-r} + u_t$
- $u_t \sim N(0, \sigma^2)$ – white noise

The moving average models, on the other hand, state that a given observation is not related to the previous process realizations but to the historical values of process noise. A moving average model of order m , *i.e.* $MA(m)$, is introduced by the following definition:

- $x_t = C + \psi_1 u_{t-1} + \psi_2 u_{t-2} + \dots + \psi_n u_{t-m} + u_t$
- $u_t \sim N(0, \sigma^2)$

However, the AR and MA models may also be combined together to create the autoregressive moving average models ($ARMA(r, m)$), which join the properties of previously presented ones.

- $x_t = C + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_n x_{t-r} + \psi_1 u_{t-1} + \psi_2 u_{t-2} + \dots + \psi_n u_{t-m} + u_t$
- $u_t \sim N(0, \sigma^2)$

The main assumption for this approach is that the residuals of models mentioned above are white noise – normally distributed random numbers. Therefore, the r lags of series observations and m lags of white noise are complete to fit a model such that its residuals are purely random. Moreover, both $AR(r)$ and $MA(m)$ are special cases of $ARMA(r, m)$ model, *i.e.* $ARMA(r, 0)$ and $ARMA(0, m)$ respectively.

2.2.2 Preparing Box-Jenkins models

Each attempt to fit an ARMA model to a given series consists of a full set of pre-analysis and fitting steps. There are certain requirements concerning the data, such that they make it possible to find a reasonable and well working ARMA model.

The first prerequisite is that the series is stationary, *i.e.* the mean value and standard deviation remain constant in the series over time. There are certain statistical tests making it possible to verify hypotheses whether a series is stationary or has a unit root. If data appear to be non-stationary, the easiest way is to create an integrated series (a series of differences). Basically, the matter is to eliminate trend from the data. There also happens to exist strong seasonality in the observations, which is why seasonal differencing might be necessary.

If the series is stationary, the next step is to analyze the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the series. Based on that a decision is made to choose a proper order of ARMA (r, m) model.

Then the process moves to parameter estimation for the chosen model. Finally, a forecast is prepared. However, ARMA models need to be monitored in an on-going manner so that amendments can be carried out, if necessary.

2.3 ARCH/GARCH modeling

Not all time series can be explained by ARMA models. Sometimes they reveal some non-stationarity in terms of volatility, *i.e.* the series variance is not constant and it depends on its historical values.

An autoregressive conditional heteroscedasticity (ARCH) model (see Engle [2]) represents the variance of the current error term as a function of the previous time period error terms' variances. ARCH simply describes the error variance by the square of a previous period's error. These types of models are widely used for time series that have a feature of so-called variance clustering, which means noticeable periods of higher and lower disturbances in the series. In general, an ARCH(q) model is represented as follows:

- $u_t = \sigma_t z_t$
- $\sigma_t^2 = K + \alpha_1 u_{t-1}^2 + \dots + \alpha_q u_{t-q}^2$,

where u_t is the corresponding ARMA(r, m) model residual series, $z_t \sim N(0, 1)$ and σ_t^2 are the variance estimates for time points t .

The model is a generalized autoregressive conditional heteroscedasticity (GARCH) (see Bollerslev [3]), if an autoregressive moving average model (ARMA-type model) is stated to represent the error variance. In that case, the GARCH(p, q) model (where p stands for the order of the GARCH terms σ_t^2 and q stands for the order of the ARCH terms u_t) is given by:

- $u_t = \sigma_t z_t$
- $\sigma_t^2 = K + \alpha_1 u_{t-1}^2 + \dots + \alpha_q u_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_p \sigma_{t-p}^2$

The models presented above are the most popular ones for explaining heteroscedasticity in time series. Usually, GARCH(1, 1) is sufficient as a compromise between simplicity of a model and its satisfactory fit to the empirical data. One of the best arguments supporting this choice is Albert Einstein's statement that the model should be "as simple as possible – but not more simple than that".

3 Statistical analysis of NEPool and Otahuhu data

The purpose of this section is to investigate the general statistical features of the given two series: New England Pool and Otahuhu (a node of New Zealand) electricity prices.

3.1 General information and basic statistics

The original data set consists of 2551 daily observations of NEPool electricity prices (7 days a week) from 03 Jan 2001 to 28 Dec 2007. The New Zealand set covers a longer period with every half an hour observations, but we use only day average prices for the same time interval as NEPool. Moreover, there were 4 days missing within this period for Otahuhu, therefore, the lacking values were replaced by linearly interpolated magnitudes. We also raise some doubts about quality of some observations, since the prices vary from 0.01 to over 500 New Zealand dollars. To avoid values close to zero the Otahuhu data are increased by 10. This operation does not change the overall character of the series.

According to the financial theory, we analyze both the prices and price logarithmic returns. The return series are created as follows:

$$r_t = \ln \frac{P_t}{P_{t-1}} \quad (1)$$

where

- r_t is return for moment t ,
- P_t is the asset's price at moment t
- P_{t-1} is the price at moment $t - 1$.

Moreover, the character of equation (1) supports our decision about adding a constant series to Otahuhu data. If there was for example a jump between prices from 0.01 to 10 dollars, the log-return ($\log(\frac{10}{0.01}) = \log(1000) \approx 6.91$) would not be naturally higher than between values like 10.01 and 20 dollars ($\log(\frac{20}{10.01}) = \log(1.998) \approx 0.692$). Therefore, without such a regularization term it would be ten times as high.

The first information on a time series usually comes after following a graphical representation. Therefore, we plot both prices and returns for NEPool in Figure 4 and for Otahuhu in Figure 5.

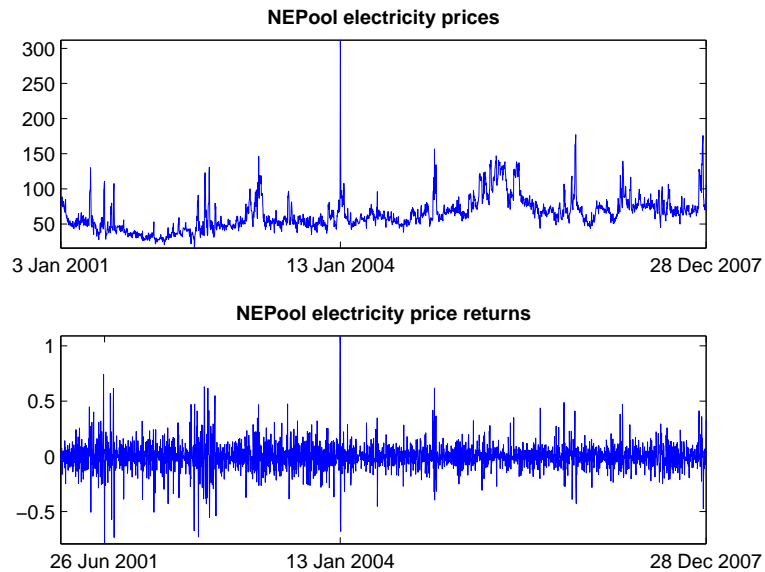


Figure 4: NEPool electricity prices and price log-returns.

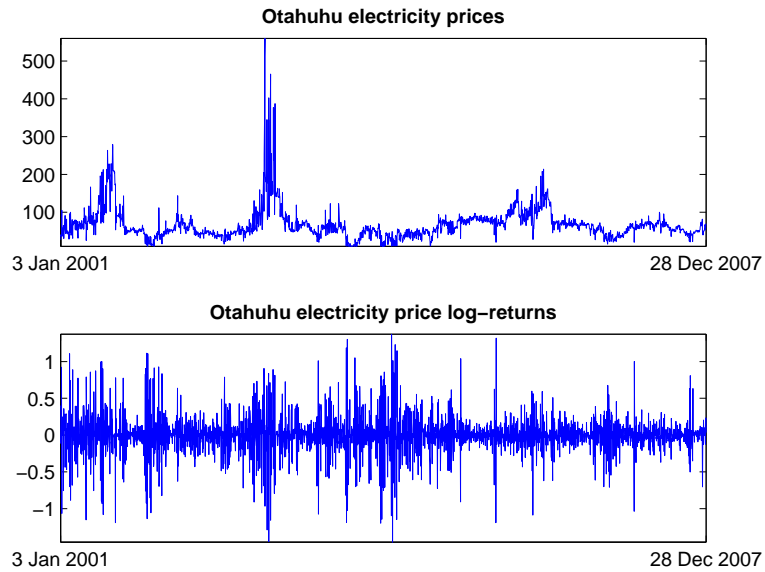


Figure 5: Otahuhu electricity prices and price log-returns.

Values of the most important distribution parameters are collected in Table 1. The NEPool prices vary from 15.8538 to 311.7500, while the Otahuhu data – from 10.01 to 560.22. This shows a huge spread of magnitudes over the given 7 years. On the other hand, the returns seem to be of a relatively small range when compared to the prices, but this is a result of logarithmic operation.

Table 1: Basic statistics for NEPool and Otahuhu electricity prices and price log-returns.

	NEPool prices	NEPool returns	Otahuhu prices	Otahuhu returns
count	2551	2550	2551	2550
mean	64.3845	$1.0134 \cdot 10^{-4}$	67.1442	$3.6908 \cdot 10^{-4}$
std	23.5171	0.1235	41.7196	0.2686
max	311.75	1.0901	560.22	1.3725
min	15.8538	-0.7911	10.01	-1.4543

3.2 Normality

The next step is to verify the type of distribution for both series. In finance it is often the case that the data are required to have normal distribution. Therefore, let us investigate the NEPool's and Otahuhu's character. As before, we start from a graphical representation, but now we plot normalized histograms of both series against theoretical normal probability density functions (PDF) (see Figure 6 and 7).

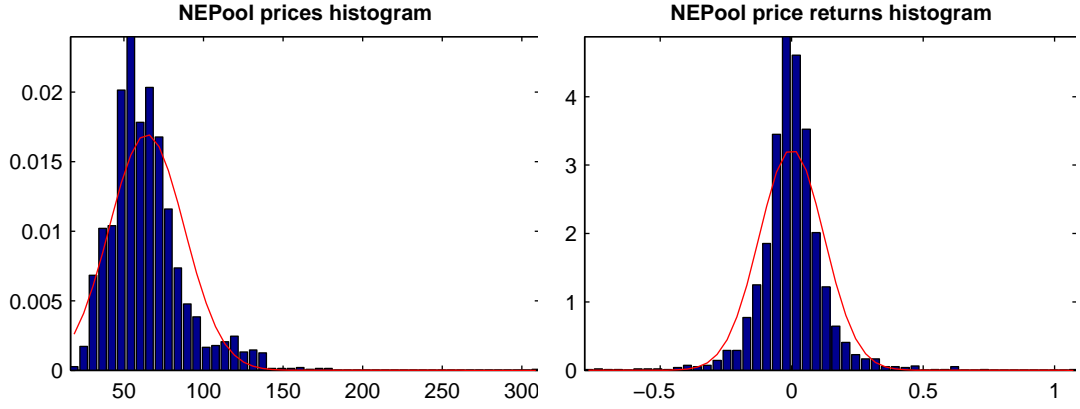


Figure 6: Normalized histograms for NEPool electricity prices (left panel) and price log-returns (right panel).

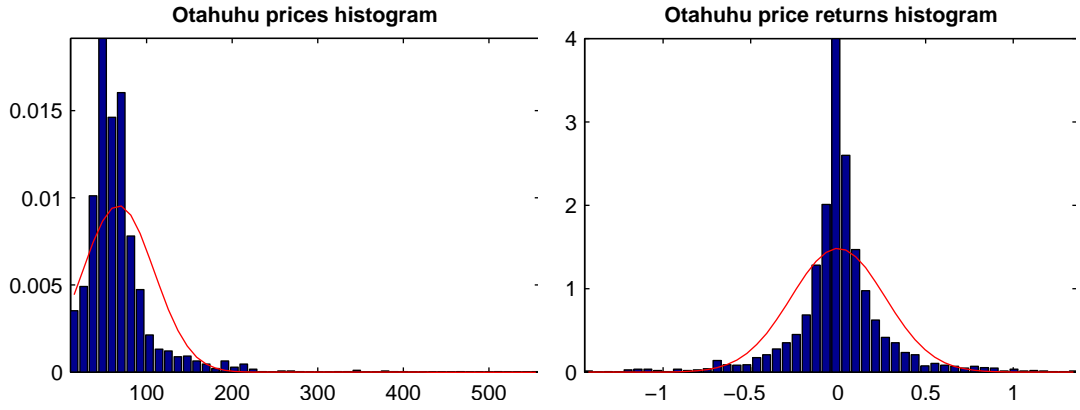


Figure 7: Normalized histograms for Otahuhu electricity prices (left panel) and price log-returns (right panel).

Secondly, we compute two most common parameters used for comparing a given probability distribution with the normal one – skewness and kurtosis. The results can be found in Table 2. Knowing that the model values should be 0 for skewness and 3 for kurtosis, we can easily see that neither prices nor log-returns follow the normal distribution.

The last step is to perform a formal statistical test for verifying normality of a given distribution. Here we choose the Lilliefors test with statistic calculated as follows:

$$L = \max_x |scdf(x) - cdf(x)|$$

where *scdf* is the empirical cumulative density function (CDF) estimated from the sample and *cdf* is the normal CDF with mean and standard deviation equal to the mean and standard deviation of the sample. In Table 2 the result can be seen – the null hypothesis was rejected for both series with 5% significance level.

Table 2: Basic statistics for NEPool and Otahuhu electricity prices and price log-returns.

	NEPool prices	NEPool returns	Otahuhu prices	Otahuhu returns
skewness	1.5561	0.1985	3.7735	-0.1318
kurtosis	9.5035	11.1252	29.0714	8.4626
Lilliefors test H_0	rejected	rejected	rejected	rejected

Summarizing this subsection, we may state that neither given NEPool and Otahuhu prices nor their returns follow the normal distribution.

3.3 Inner dependencies

Here we move to an analysis of other features of the data. Figures 4 and 5 show that the series are not stationary, which simply means that neither their mean values nor their standard deviations remain constant over time. Therefore, we should perform a formal test.

Let us assume that we have a process

$$y_t = \phi y_{t-1} + u_t \quad (2)$$

where y_t and u_t are the given time series and model residuals respectively. Then the Dickey-Fuller [1] (DF) test examines the null hypothesis $\phi = 1$ (the process has a unit root, *i.e.* its current realization appears to be an infinite sum of past disturbances with some starting value y_0 ; see Brooks [8]) versus the one-side alternative $\phi < 1$ (the process is stationary). The test statistics look as follows

$$DF = \frac{1 - \hat{\phi}}{\hat{SE}(1 - \hat{\phi})} \quad (3)$$

and follow a non-standard distribution, critical values of which were derived from experimental simulations.

A similar test is the Phillips-Perron [5] test. However, this one relaxes assumptions about lack of autocorrelation in the error term. Its critical values are the same as for Dickey-Fuller [1] test.

Even though the presented tests work well in obvious cases, there has been some criticism of them. A problem appears when the process has the ϕ value close to the non-stationarity boundary, *i.e.* $\phi = 0.95$. Such a process is by definition still stationary for DF and PP tests. It has been proven that these tests often do not distinguish whether $\phi = 1$ or $\phi = 0.95$, especially if the sample is of a small size. Therefore, a different test was developed with the opposite null hypothesis. The Kwiatkowski-Phillips-Schmidt-Shin [6] test (KPSS) states $H_0 : y_t \sim I(0)$ against $H_1 : y_t \sim I(1)$. Its statistics looks as follows

$$KPSS = \frac{\sum_{i=1}^n \hat{S}_i^2}{n^2 \hat{s}^2} \quad (4)$$

where n is the sample size, $\hat{S}_i^2 = \sum_{j=1}^i \psi_j$ (sum of residuals ψ_t from original series regressed on trend and constant) and s^2 is a sample long-run variance.

The confirmatory analysis (DF/PP joint with KPSS) gives a better view on whether obtained stationarity/non-stationarity results are robust (see Brooks [8]). The most desirable outcomes are when H_0 is rejected by DF/PP and accepted by KPSS or exactly opposite – accepted by DF/PP and rejected by KPSS. If H_0 is accepted or rejected in both tests simultaneously, the results are conflicting and one cannot say unequivocally which one is right.

Table 3 collects outputs from all test for all price and return series with 5% significance level. We obtain one conflict – for NEPool prices. Otherwise, the series are stationary.

Table 3: H_0 decisions of DF, PP and KPSS tests for NEPool and Otahuhu prices and price returns.

	DF	PP	KPSS
NEPool prices	rejected	rejected	rejected
NEPool returns	rejected	rejected	accepted
Otahuhu prices	rejected	rejected	accepted
Otahuhu returns	rejected	rejected	accepted

In econometrics stationarity is one of the most important conditions for time series modeling. Therefore, bearing in mind the graphical representation of the prices and the conflict we obtained, the analyses cover the log-returns series in parallel.

Now we move to plotting autocorrelation functions (ACF) and partial autocorrelation functions (PACF) for both series. As we can see in Figure 8, the ACF of NEPool prices seems to die out slowly, whereas the PACF plot reveals a very significant spike at lag 1. These two facts lead us to use an ARMA(1,0) model for the process estimation. Analogically, we plot ACF and PACF for Otahuhu prices and discover a similar characteristic (see Figure 9). ARMA(1,0) model would be relevant here as well.

Figure 10 demonstrates the ACF and PACF plots for the NEPool price returns. When compared to the prices' PACF, there are no spikes comparably springing aside at any lag for neither ACF nor PACF of the returns. However, there are still a few above the significance level and these are, in particular, the second lags for both functions.

Plots of ACF and PACF for Otahuhu returns in Figure 11 show the most significant values at first spikes for both functions. Thus, ARMA(1,1) models could be applicable for NEPool and New Zealand returns.

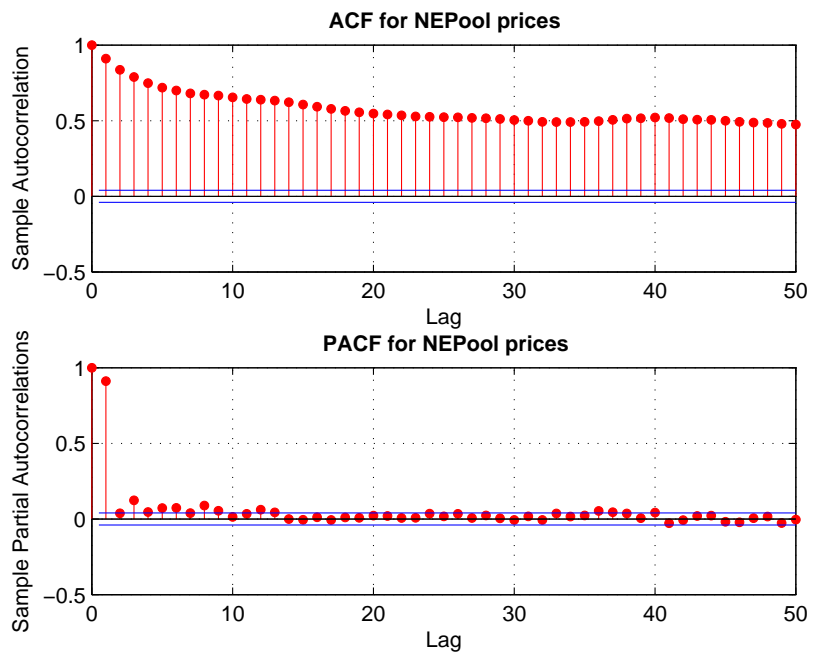


Figure 8: ACF and PACF for NEPool electricity prices.

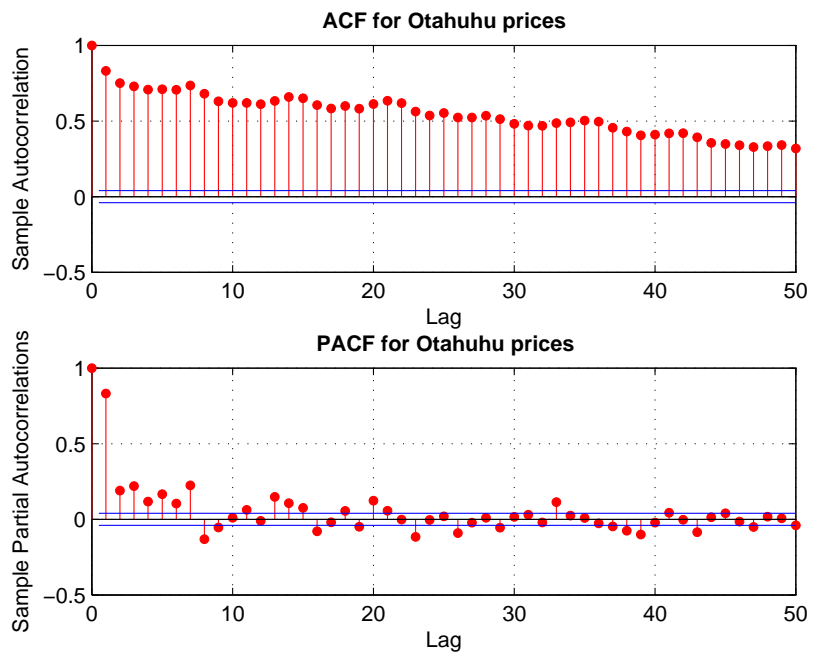


Figure 9: ACF and PACF for Otahuhu electricity prices.

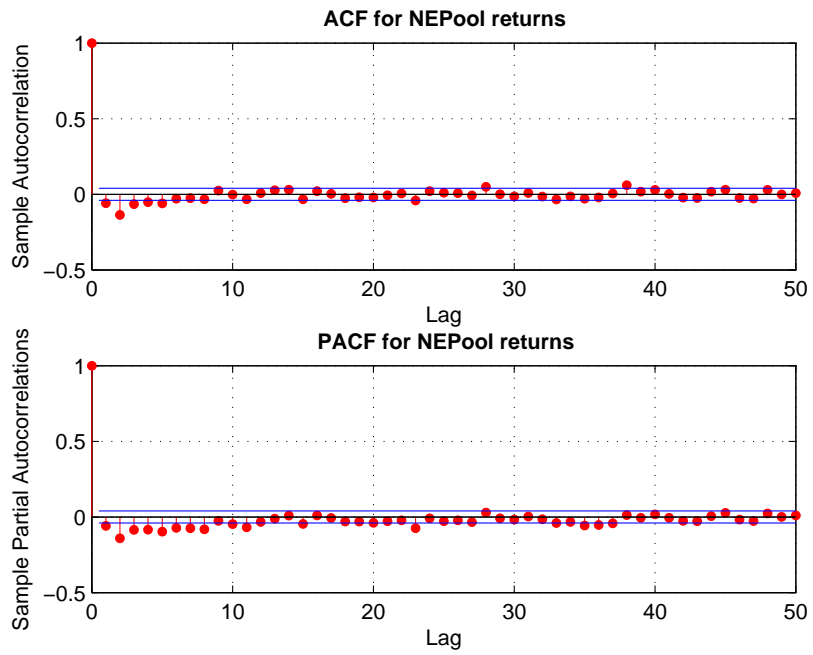


Figure 10: ACF and PACF for NEPool electricity price returns.

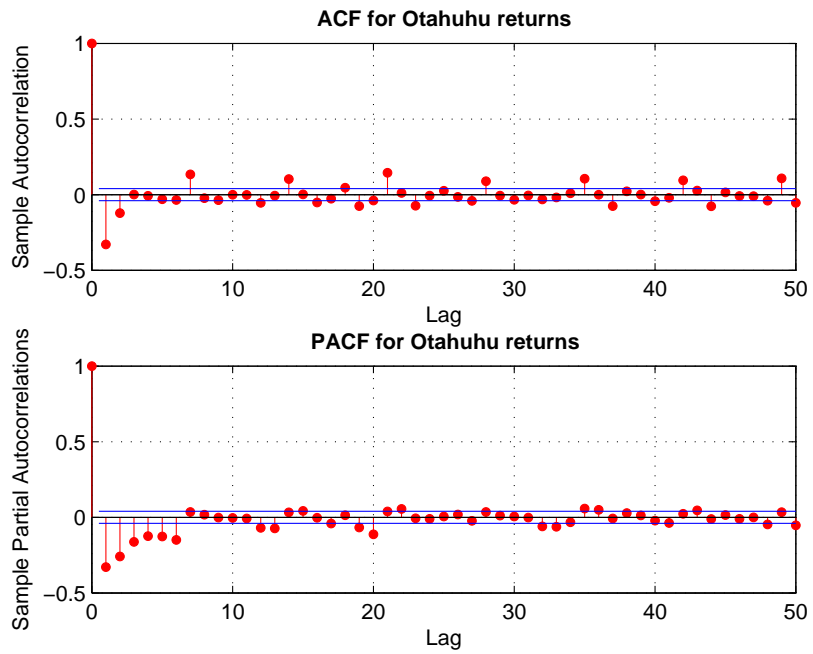


Figure 11: ACF and PACF for Otahuhu electricity price returns.

Moreover, plots of returns in Figures 4 and 5 demonstrate so-called variance clustering. Thus, we can separate periods of higher and lower level of disturbances. Therefore, the last step in this subsection is to test for an ARCH/GARCH effect in both series. Here we use Engle's test with statistics $T(R^2)$, where T represents the number of squared residuals considered in the regression and R^2 is the sample multiple correlation coefficient. The test rejected the null hypothesis in both cases, which means there exists heteroscedasticity in price and return series for both regions.

This subsection showed very important results from modeling point of view. One can expect that for estimation of given processes ARMA and ARCH/GARCH type of models are needed.

3.4 Discrete Fourier transform smoothing

It is a really rare situation that a time series is not noisy. This is why different techniques of smoothing signals have been developed. The one chosen for this study is discrete Fourier transform, which was widely described by Bracewell [7]. The general idea is based on transforming a sequence of complex numbers into another by the following formula:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} kn} \quad k = 0, \dots, N-1 \quad (5)$$

where $e^{\frac{2\pi i}{N}}$ is a primitive N -th root of unity, X_k is the transformed series and x_n is the original sequence. The easiest way to interpret this equation is that computed numbers X_k stand for the amplitude and phase of sinusoidal components of the original series. An inverse operator (inverse discrete Fourier transform) is

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k e^{\frac{2\pi i}{N} kn} \quad n = 0, \dots, N-1 \quad (6)$$

which restores the sum of sinusoidal components.

The general idea is to verify which of the frequencies are most significant in the process description. Then the smoothed signal is reconstructed with use of only the most crucial components. In numerical methods, a fast Fourier transform algorithm is employed to obtain the DFT representation.

The first step is to compute and plot the DFT representation of NEPool and Otahuhu prices. Since X_k is a sequence of complex numbers, we plot and analyze norm of the numbers understood as the classical complex number module

$$|X| = \sqrt{(Re(X))^2 + (Im(X))^2}.$$

Figure 12 presents norms of FFT for NEPool and Otahuhu data series, respectively.

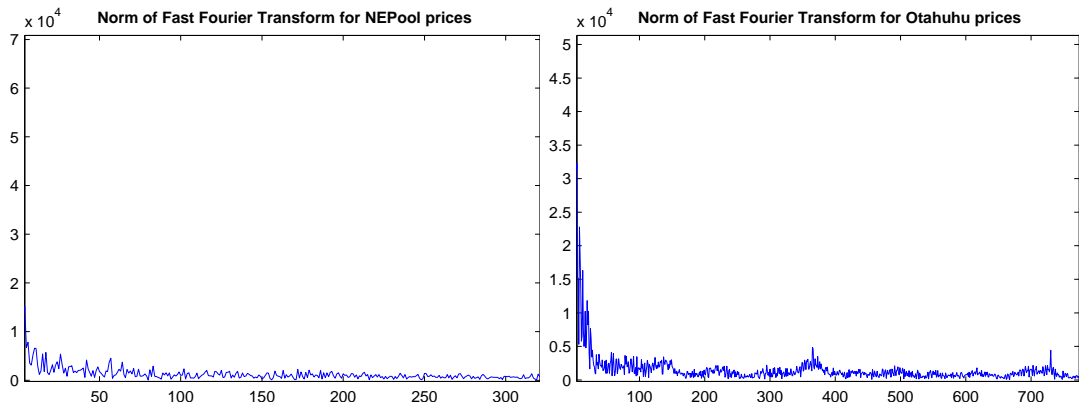


Figure 12: Norm of FFT for NEPool (left panel) and Otahuhu (right panel) prices.

The magnitudes of components decrease gradually, however, we need to decide which interval to choose for further analysis. Here the 60th element seems to be a boundary of significance for NEPool and 30th for Otahuhu. One could think that for Otahuhu also components like 365th and 730th should be included, but after the reconstruction process they just create a high frequency wave on the main signal. Therefore, in X_k we replace all not crucial components by zeros. Then the IDFT can be computed to retrieve the main signal from the original data. The results of this operation for New England and Otahuhu series are presented in Figure 13 and Figure 14.

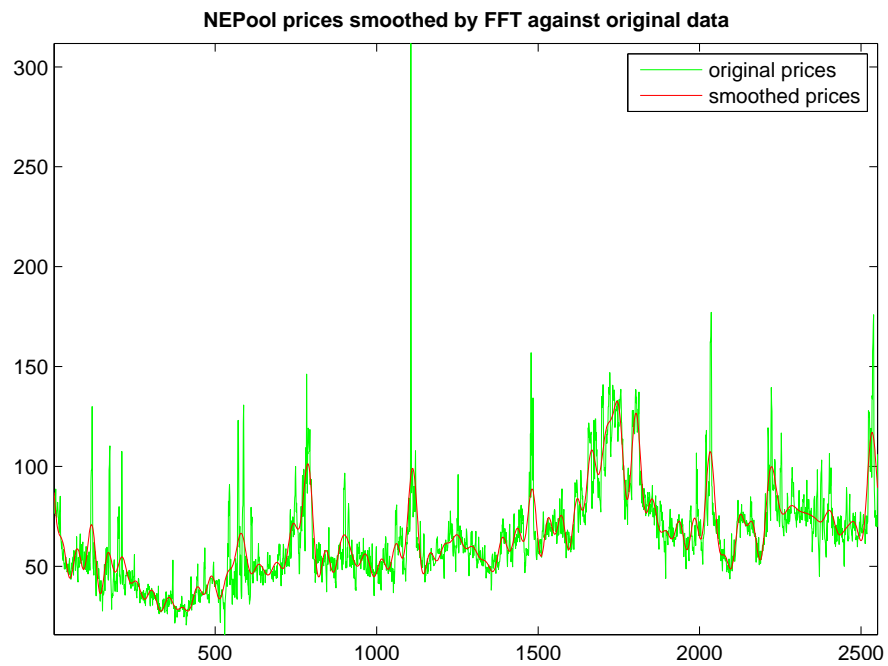


Figure 13: NEPool prices smoothed by FFT against original data.

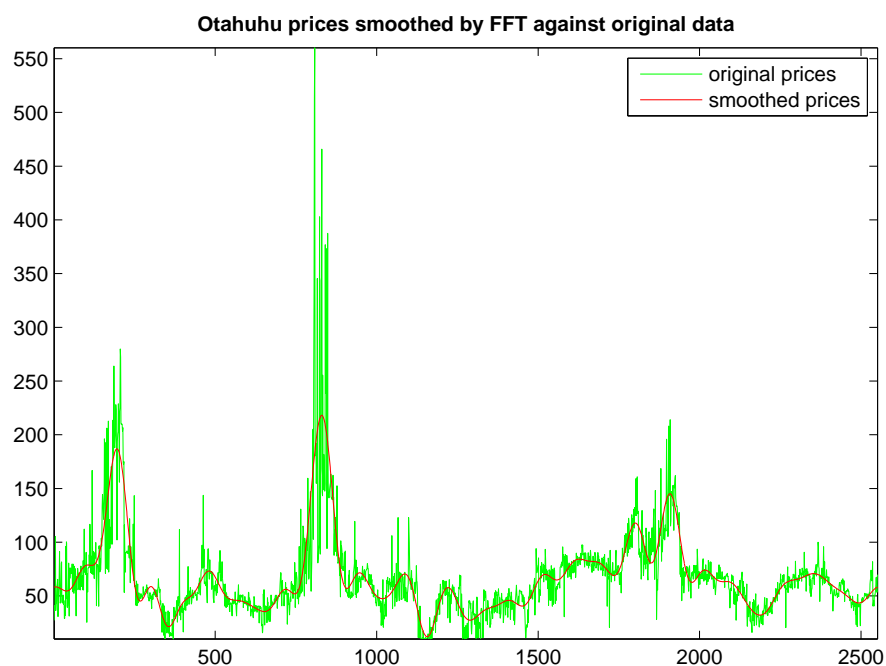


Figure 14: Otahuhu prices smoothed by FFT against original data.

For both data sets the smoothed paths clearly follow the primary series; they do not, however, explain numerous spikes.

Next we verify how the smoothing influenced the price returns, see Figures 15 and 16 for NEPool and Otahuhu respectively.

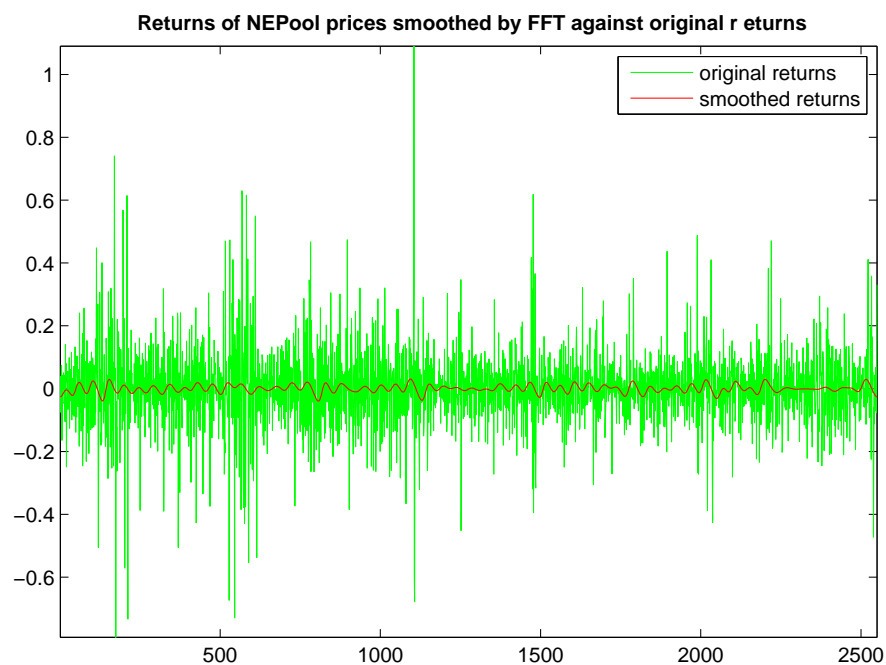


Figure 15: Returns of NEPool prices smoothed by FFT against original data.

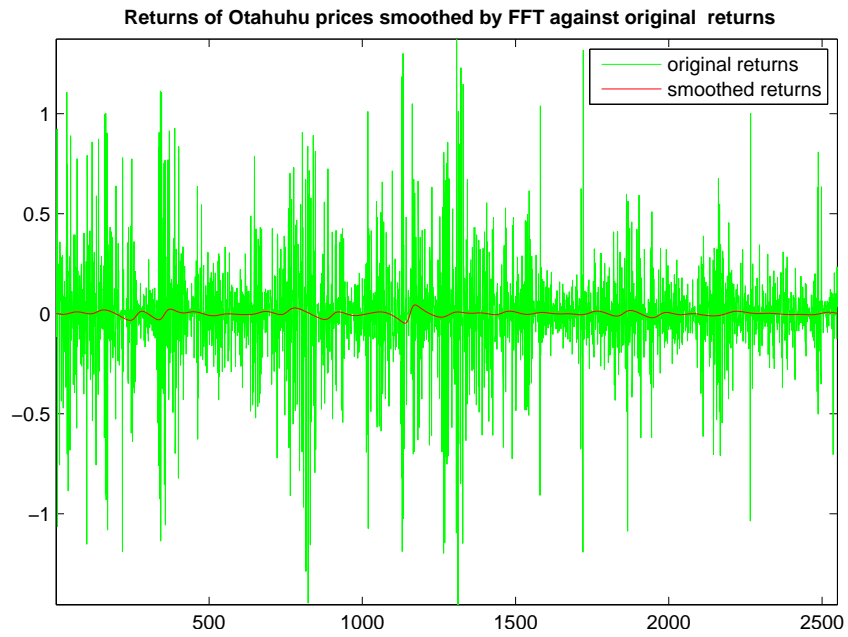


Figure 16: Returns of Otahuhu prices smoothed by FFT against original data.

As we can see, returns of the smoothed prices do not explain much of the original log-return series. Moreover, the fairly regular look of the smoothed returns wave may indicate significant autocorrelation of the process. ACFs and PACFs for both smoothed prices and returns are plotted in Figures 17, 18 and 19, 20 for NEPool and Otahuhu respectively.

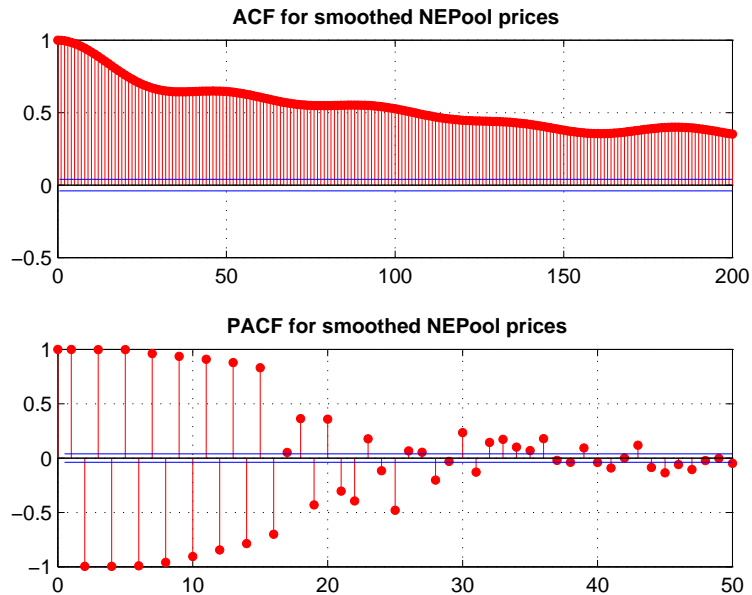


Figure 17: ACF and PACF for smoothed NEPool prices.

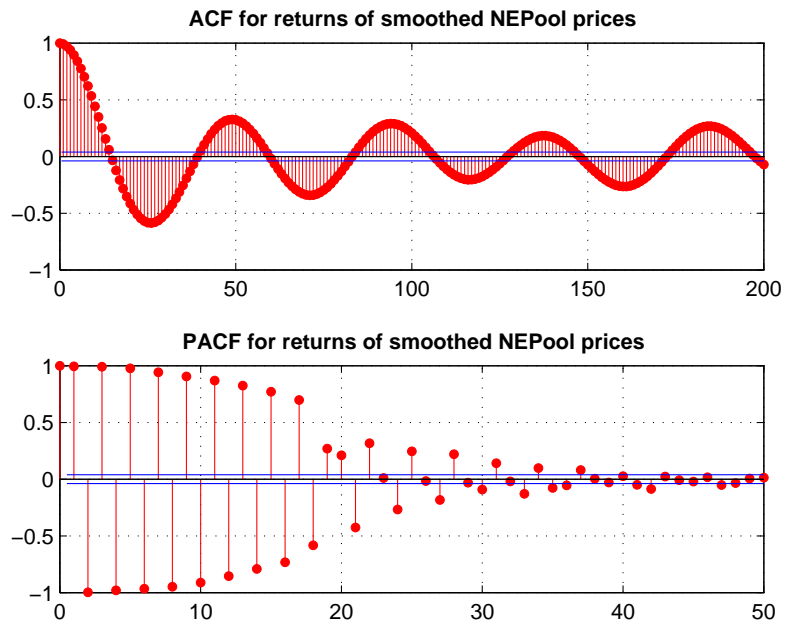


Figure 18: ACF and PACF for returns of smoothed NEPool prices.

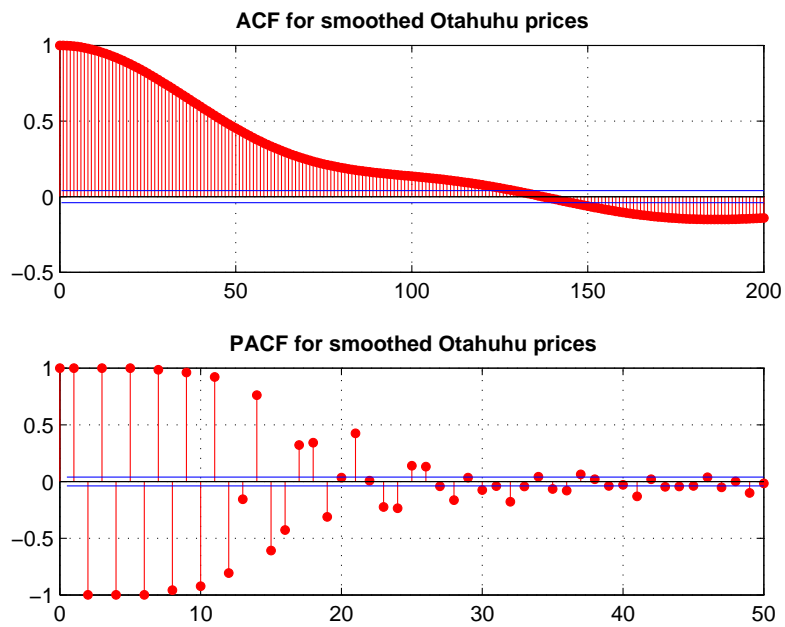


Figure 19: ACF and PACF for smoothed Otahuhu prices.

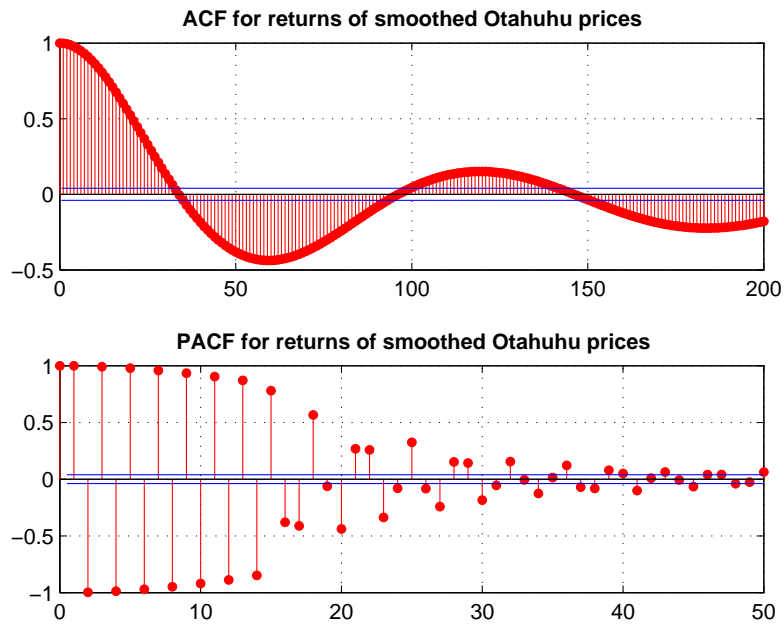


Figure 20: ACF and PACF for returns of smoothed Otahuhu prices.

The plots could lead to a conclusion that smoothing of the prices results in revealing high autocorrelation and seasonality from the original series, but this is only an effect of a phase nature of Fourier transform. Moreover, DFT does not eliminate ARCH/GARCH effect from the price series. As the Engle's test states, there still remains heteroscedasticity in the processes.

3.5 Week days analysis

Since the data set consists of daily prices, it gives an interesting base for dummy analysis. It is well known that electricity demand is highly dependant on days of week. On the other hand, demand is a crucial factor steering prices. Therefore, how are prices related to week days? Figures 21 and 22 present simple plots of original and DFT smoothed prices for separated week days – from Monday to Sunday – for NEPool and Otahuhu respectively. The general path of the process seems to be of a similar character for all week days.

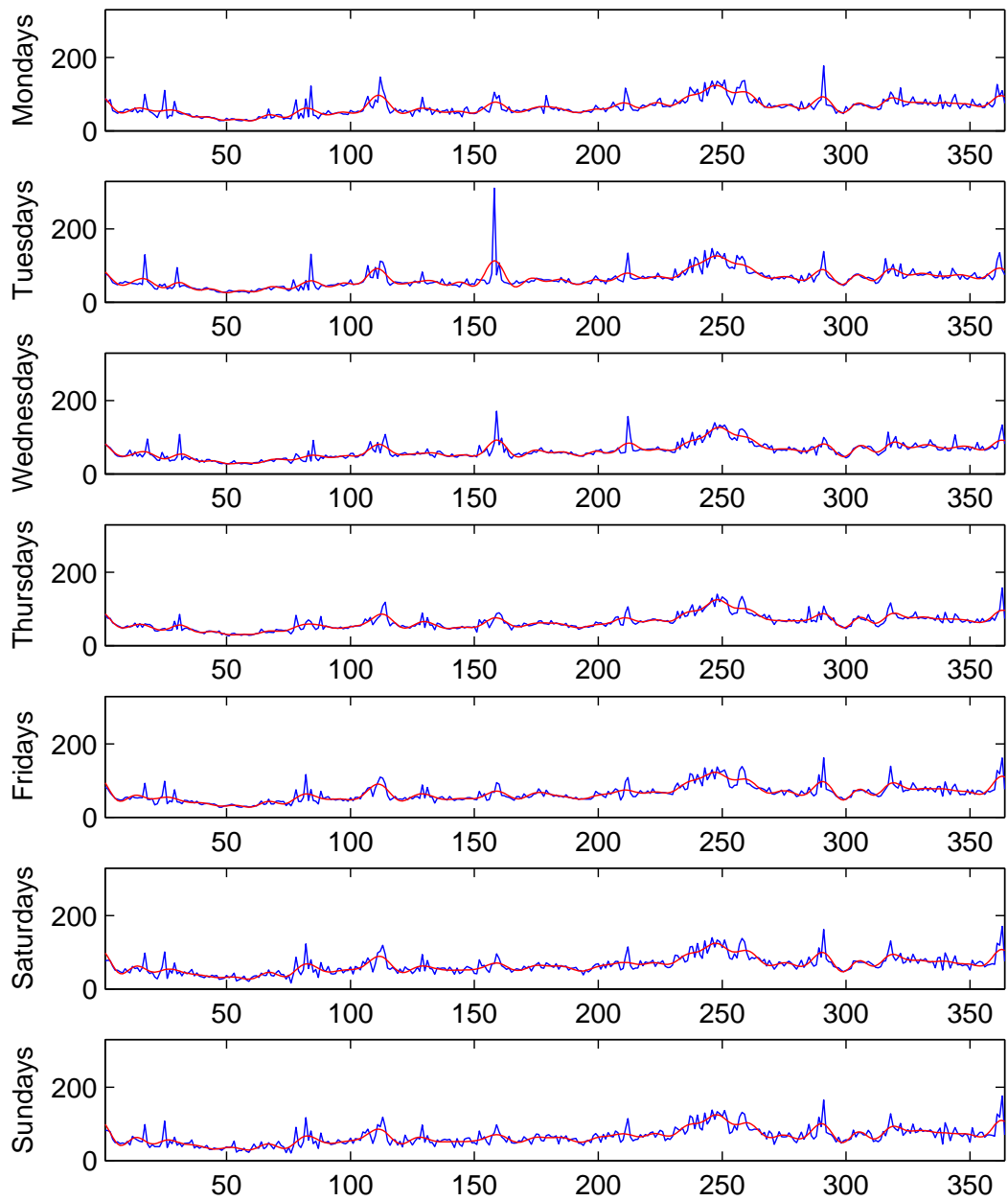


Figure 21: NEPool electricity original (blue) and DFT smoothed (red) prices split with respect to days of the week.

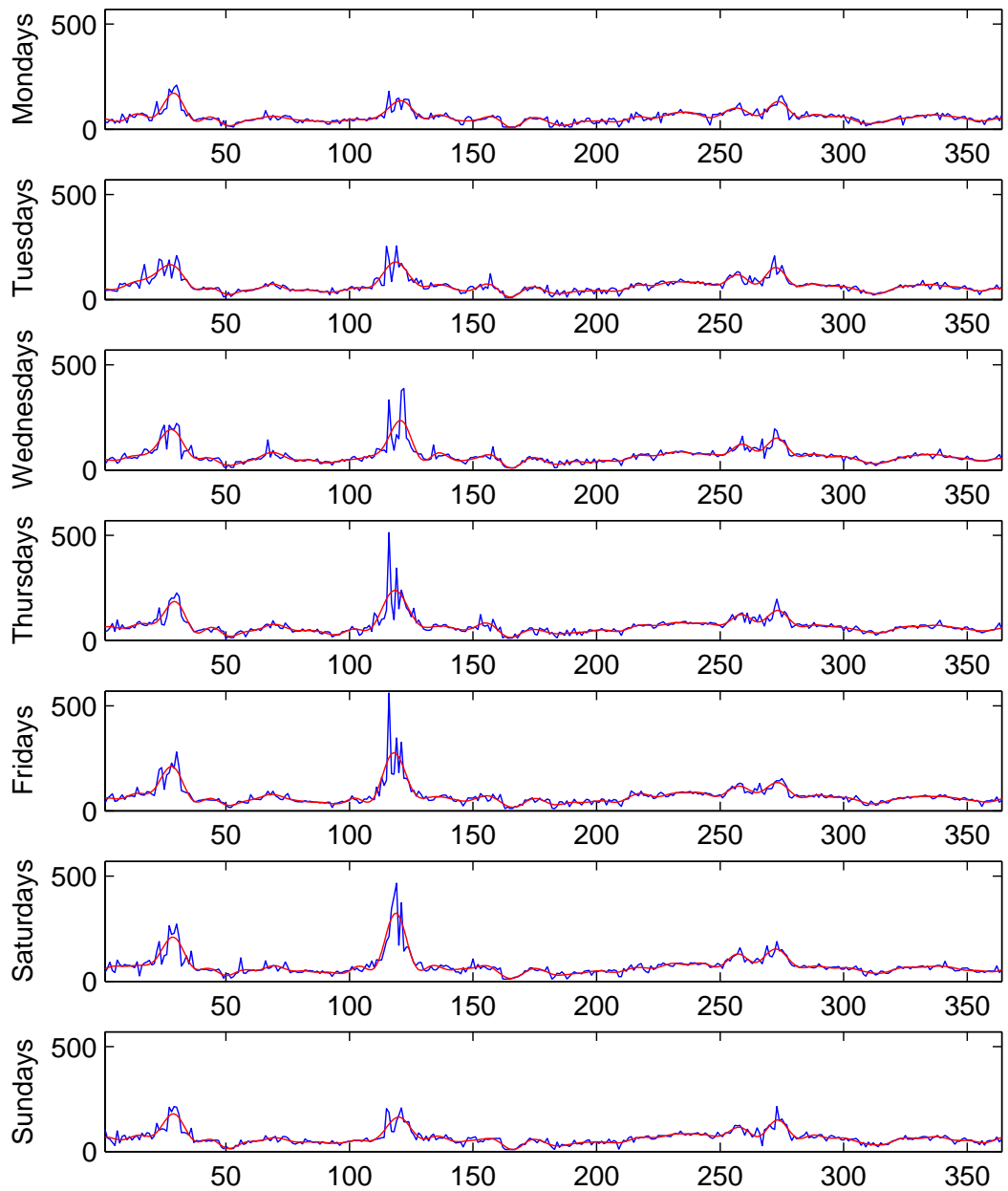


Figure 22: Otahuhu electricity original (blue) and DFT smoothed (red) prices split with respect to days of the week.

Now let us compare the mean values of prices for different week days over the whole period. Results collected in Table 4 show that for NEPool Wednesdays have in average

the lowest prices, while Sundays get the highest. On the other hand, the Otahuhu prices are on average the lowest on Mondays and the highest on Saturdays. Moreover, the New Zealand series have relatively higher volatility than NEPool, while having comparable mean values.

Table 4: Basic statistics for NEPool and Otahuhu electricity prices split with respect to days of the week.

	NEPool mean	NEPool st dev	Otahuhu mean	Otahuhu st dev
Monday	64.7422	23.3647	57.8311	30.3384
Tuesday	64.1553	26.4285	66.3546	35.9977
Wednesday	63.1298	22.4910	68.6161	43.7648
Thursday	63.9913	21.0027	68.6207	44.1576
Friday	64.7699	23.2371	70.3705	47.9148
Saturday	64.7821	23.8649	73.1128	50.5082
Sunday	65.1243	24.0286	65.0864	33.7501

The differences between days are relatively small and standard deviations remain similar within NEPool and Otahuhu data. A graphical representation of the mean values together with upper and lower limits is included in Figure 23 for NEPool (left panel) and for Otahuhu (right panel).

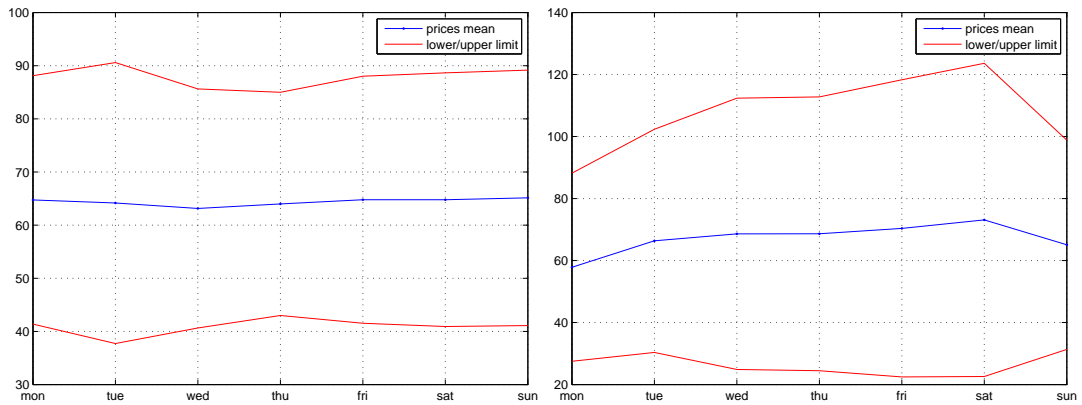


Figure 23: NEPool and Otahuhu electricity prices averages with lower and upper limits split with respect to days of the week.

Analogically, an analysis of price log-returns can be carried out. Figure 24 presents seven NEPool series of weekly data with regard to week-days. We can see that Mondays have the highest volatility, Saturdays and Sundays present the most uniform realizations of the returns with the lowest magnitudes of disturbances, while days from Tuesday to Friday are moderately volatile, but reveal most visible spikes in the series.

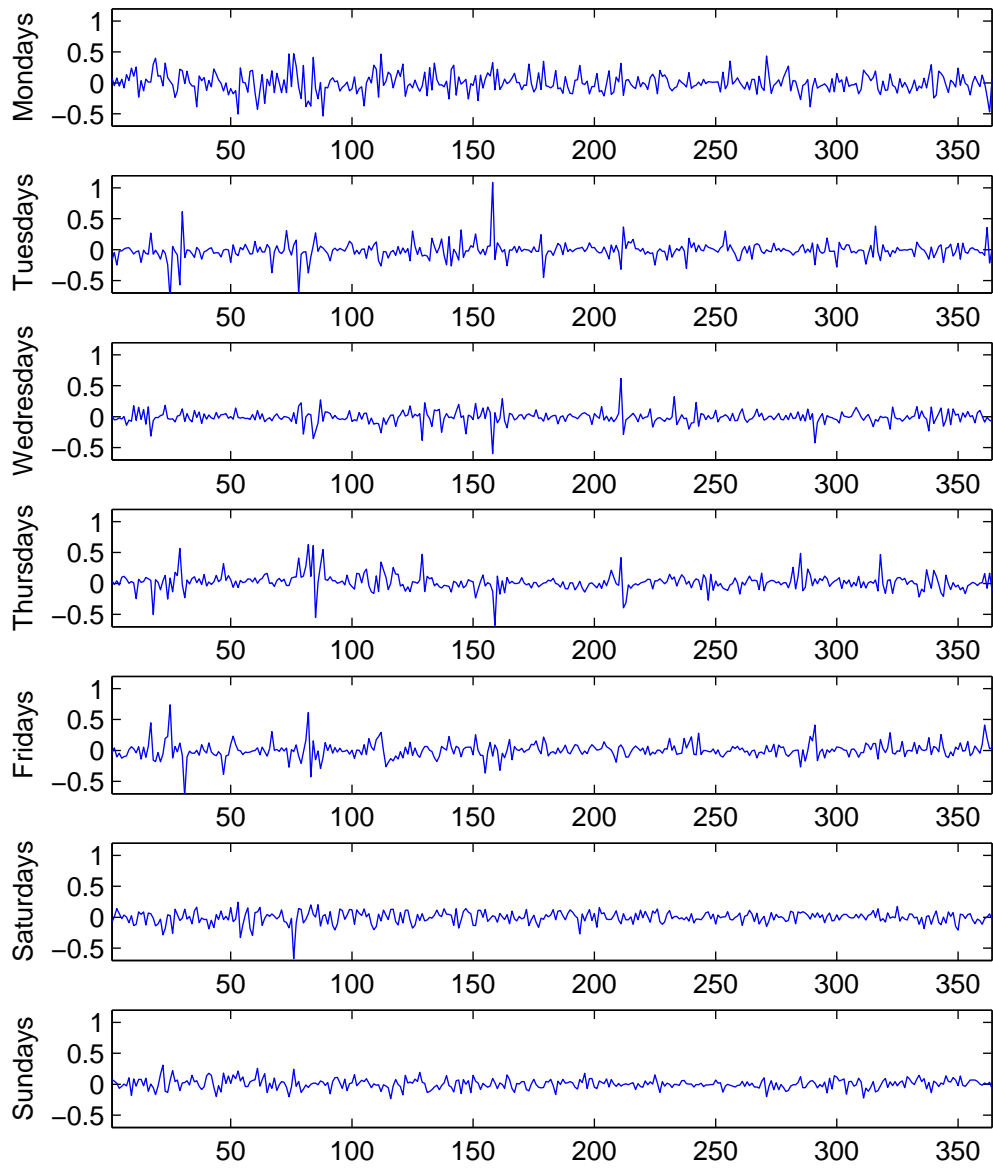


Figure 24: NEPOOL electricity price returns split with respect to days of the week.

We present an analogical plot for Otahuhu in Figure 25. Notice that all 7 series look comparably volatile in the left halves of the plots. In the second half of the analysed period we can distinguish Mondays and Tuesdays as days with higher disturbances and the remaining ones as with lower returns.

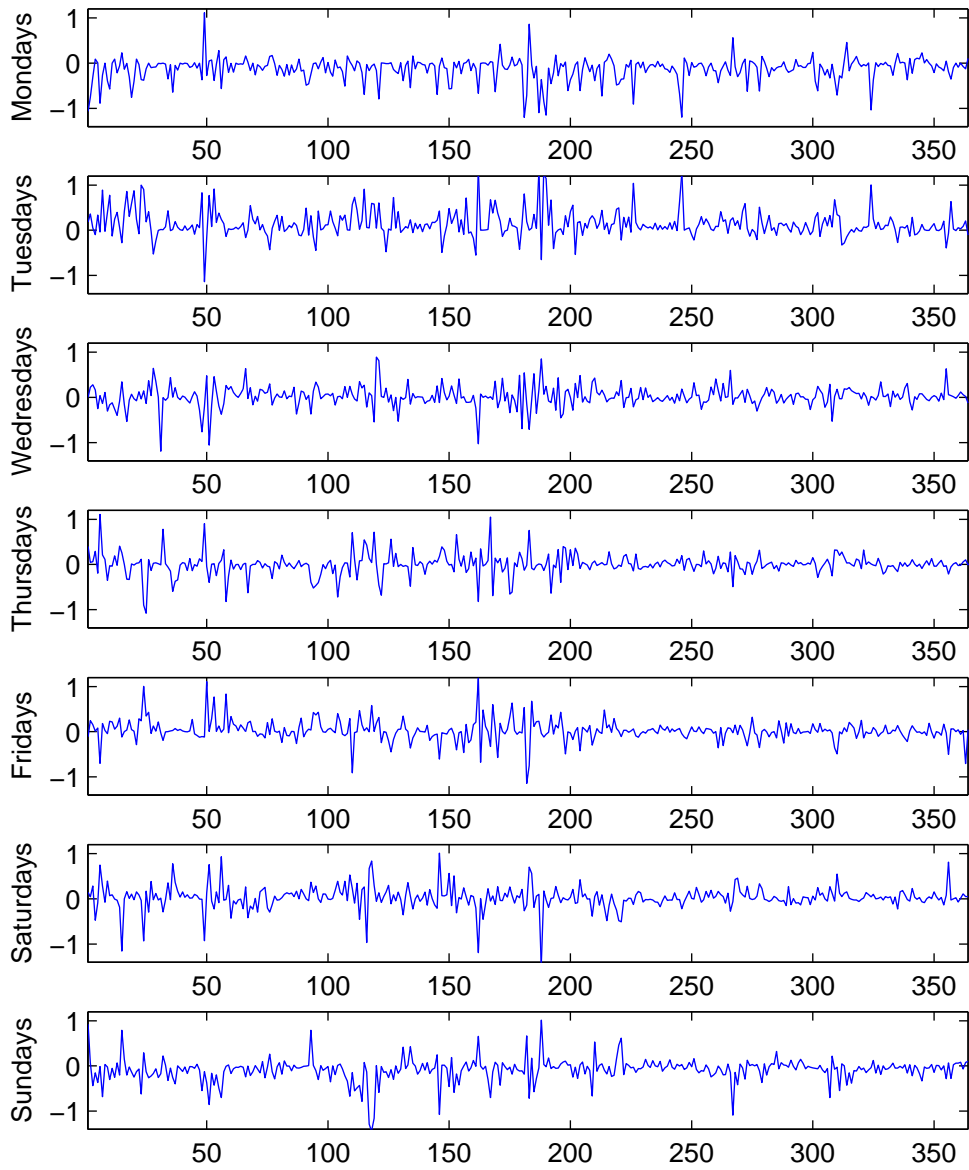


Figure 25: Otahuhu electricity price returns split with respect to days of the week.

The basic statistics of the NEPool series collected in Table 5 lead us to the conclusion that negative returns occurring from Monday to Wednesday may be a reason for averagely lowest prices on Wednesdays. On the other hand, mostly positive returns on the other days create the highest prices on Sundays. For Otahuhu, the average negative returns of Mondays and Sundays work on the lowest prices on Mondays, while the other 5 days

with positive mean values of returns relate to the highest prices on Saturdays.

Table 5: Basic statistics for NEPool electricity price returns split with respect to days of the week.

	NEPool mean	NEPool st dev	Otahuhu mean	Otahuhu st dev
Monday	-0.0020	0.1575	-0.1288	0.2617
Tuesday	-0.0154	0.1414	0.1422	0.2996
Wednesday	-0.0090	0.1045	0.0194	0.2363
Thursday	0.0212	0.1357	0.0011	0.2423
Friday	0.0039	0.1272	0.0245	0.2346
Saturday	-0.0040	0.0965	0.0301	0.2537
Sunday	0.0058	0.0807	-0.086	0.2601

Similarly to prices, we collect the mean values with upper and lower limits in Figure 26 for the NEPool returns (left panel) and for Otahuhu returns (right panel).

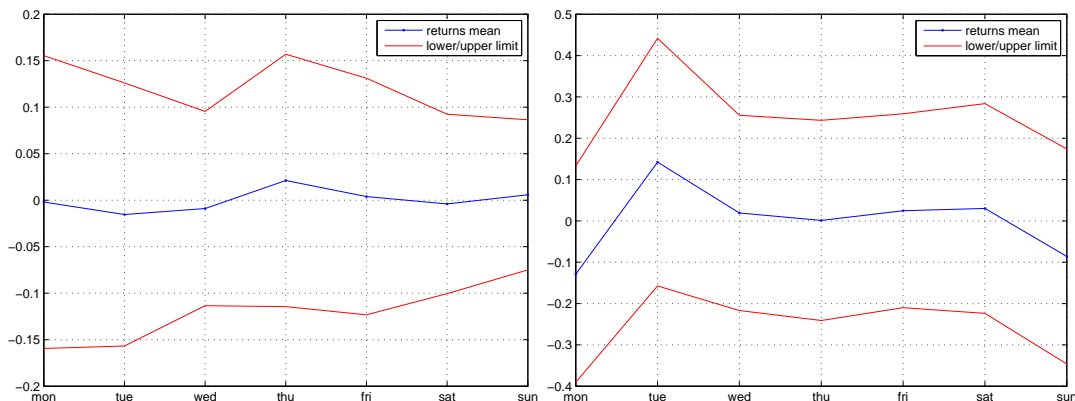


Figure 26: NEPool and Otahuhu electricity price returns average with lower and upper limits split with respect to days of the week.

Finally, we plot the autocorrelation functions of all split NEPool series: 7 for prices (Figure 27, left panel) and 7 for returns (Figure 27, right panel). An interesting observation is that even though prices show a high autocorrelation with respect to days of the week, the returns seem to be uncorrelated from this point of view. Simply, log-returns of Mondays do not explain the other Mondays results, Tuesdays do not explain Tuesdays *etc.* Otahuhu weekly ACFs reveal similar features (see Figure 28).

The formal statistical test of Lilliefors shows that the series repartition by week days does not lead to normally distributed data. For NEPool only the Sunday returns and for Otahuhu Wednesday and Saturday returns have the null hypothesis accepted.

Finally, we verify existence of ARCH/GARCH effect in all 28 series. As before, we use the Engle's test for this purpose. As a result with 5% significance level we obtain that all the week day price series reveal heteroscedasticity.

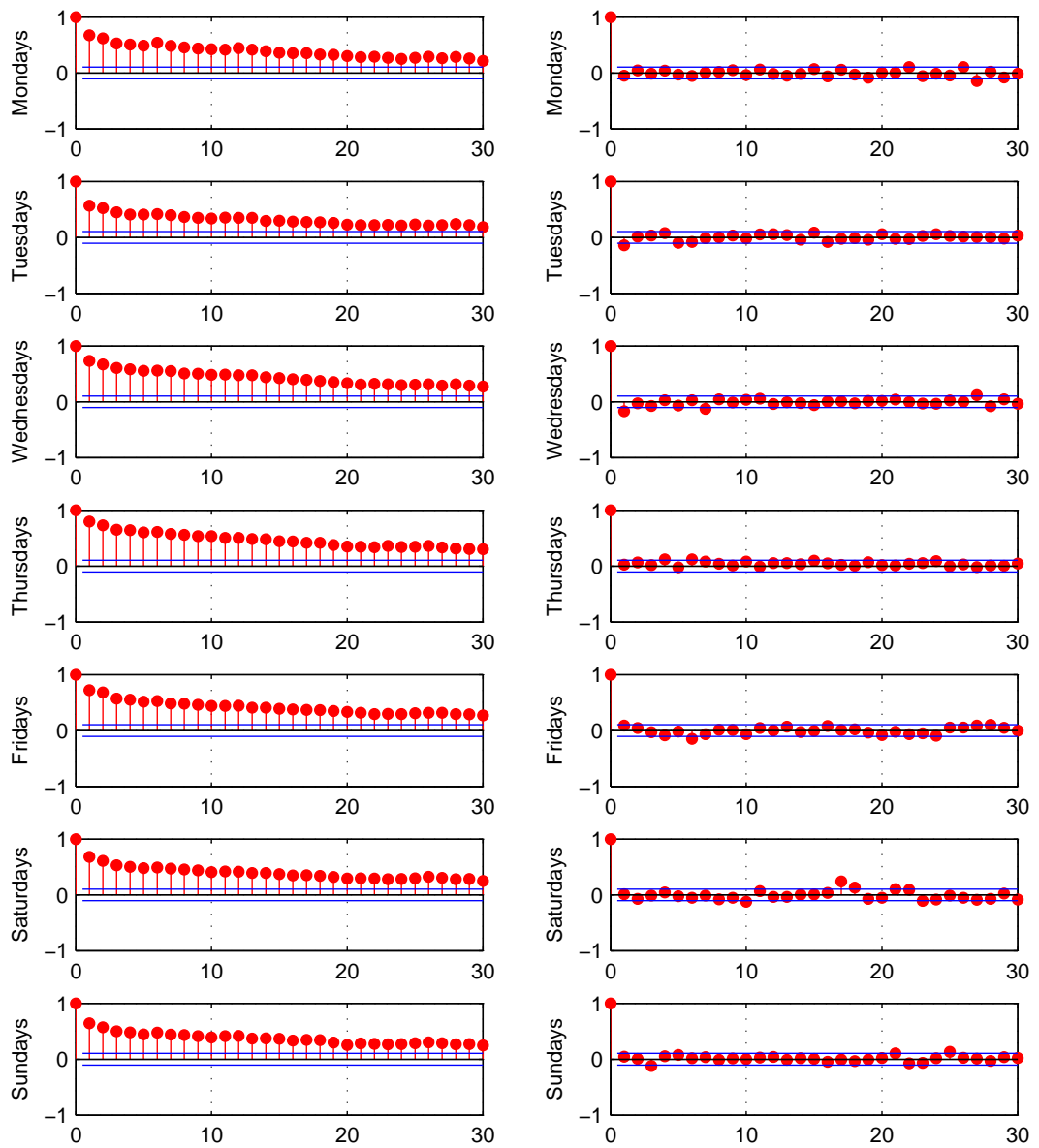


Figure 27: ACFs for NEPool electricity prices (left panel) and price returns (right panel) split with respect to days of the week.

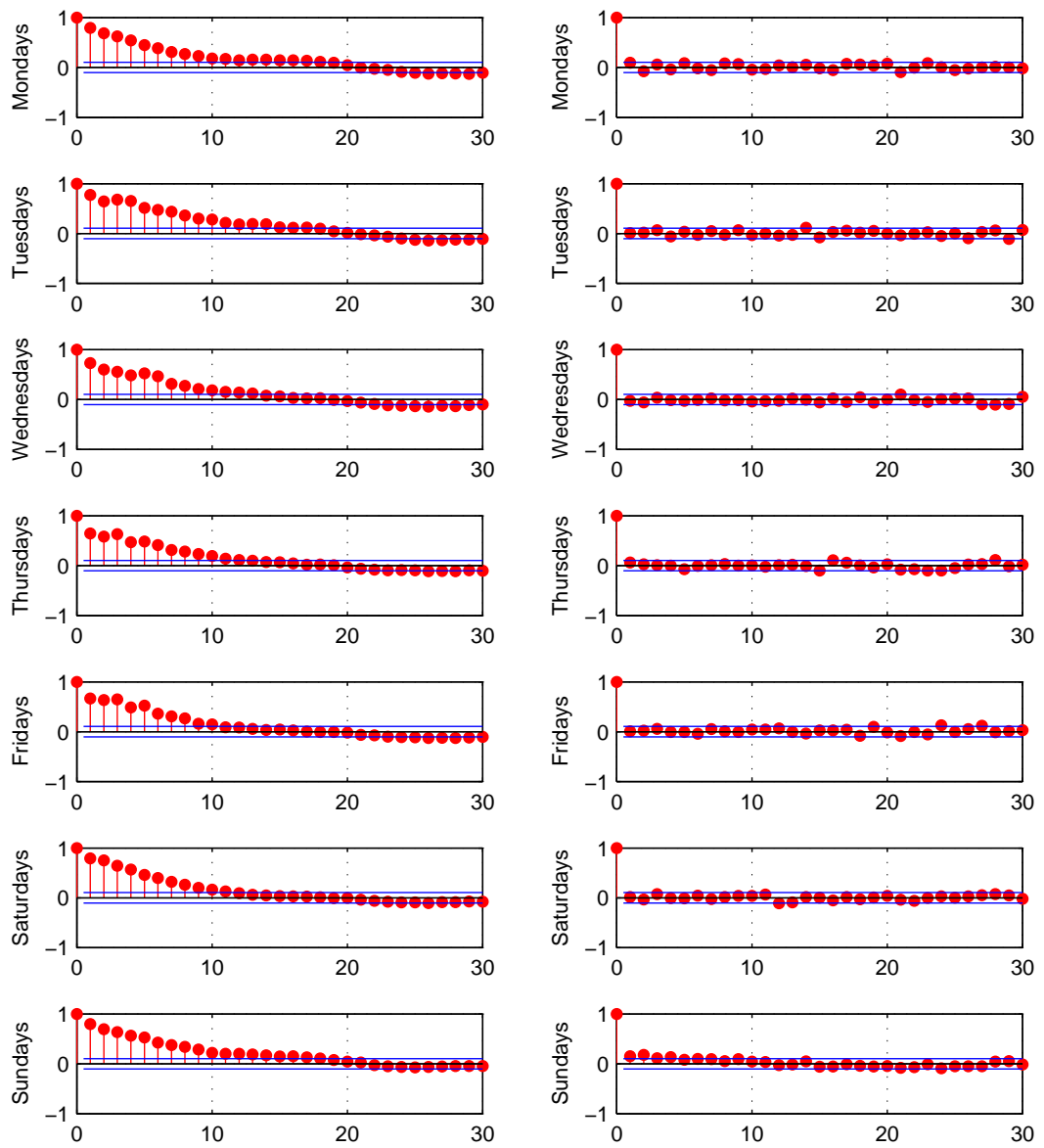


Figure 28: ACFs for Otahuhu electricity prices (left panel) and price returns (right panel) split with respect to weekdays.

4 Analysis of outliers

In this section we analyze the NEPool and Otahuhu figures from a slightly different point of view – not modeling them as a time series, but studying the magnitudes and frequency of so-called spikes.

An outlier (in this paper called also a spike or jump) can be interpreted in many ways. Intuitively, it is just a visible outlier when compared to values in its neighborhood. For the purpose of this study, we define a spike as a process realization which exceeds the mean value (μ) in a specified window W by more than doubled standard deviation (σ) of the window. In particular, if $x(i)$ is a price in moment i and W is the window size, then we compare value of $x(i)$ with its neighborhoods parameters: $\mu + 2\sigma$ and $\mu - 2\sigma$, considering $\frac{W}{2}$ observations on the left from price $x(i)$ and $\frac{W}{2}$ observations on the right. We will, however, carry out the analysis from two points of view: the original series and the smoothed one.

4.1 Spikes with respect to original price series

Here we compare each spot price with parameters computed for specified windows within the original data. For each chosen window size we construct a vector of spikes magnitudes. An outlier is calculated as a difference between the price and local (window) mean value; all elements not accepted as spikes are equalized to zero. For each window size we analyze a total number of spikes, average magnitudes of spikes and their average frequency.

Since the original sample is of size 2551, we choose window sizes as 700, 600, 500, 400, 300, 200 and 100. The obtained NEPool and Otahuhu spike vectors are presented in Figures 29 and 30. It is worth mentioning that the original New England and New Zealand series have both visible upward and downward spikes. We can see a visibly similar character of the spikes distribution, even though the width of windows chosen for analysis range from 100 to 700. Moreover, the spikes seem to reveal clustering – we can distinguish between periods of higher and lower frequency of spikes occurrence. Thus it could be worth to review price series with removed spikes.

The first intuitive verification of spikes inner dependencies is calculating the ACF and PACF for them. Any significant values at lags 7 or 30 would indicate on seasonality. However, the outliers appear not to reveal this feature. The most significant values are at lags 1-4, rather due to the character of spike vector definition, *i.e.* that non-spikes are equal to zero and they make most of the values.

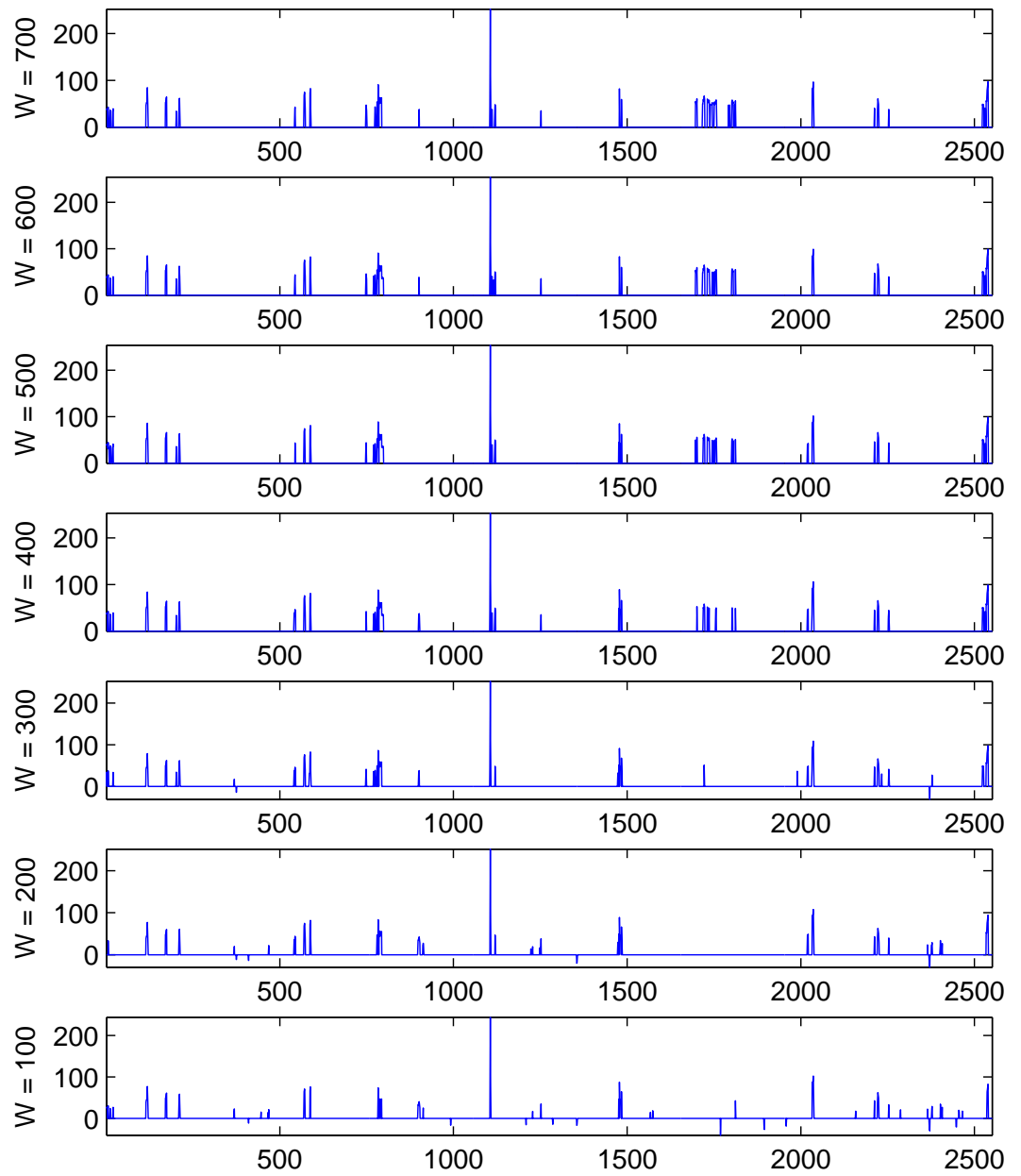


Figure 29: NEPool spike vectors for different window sizes.

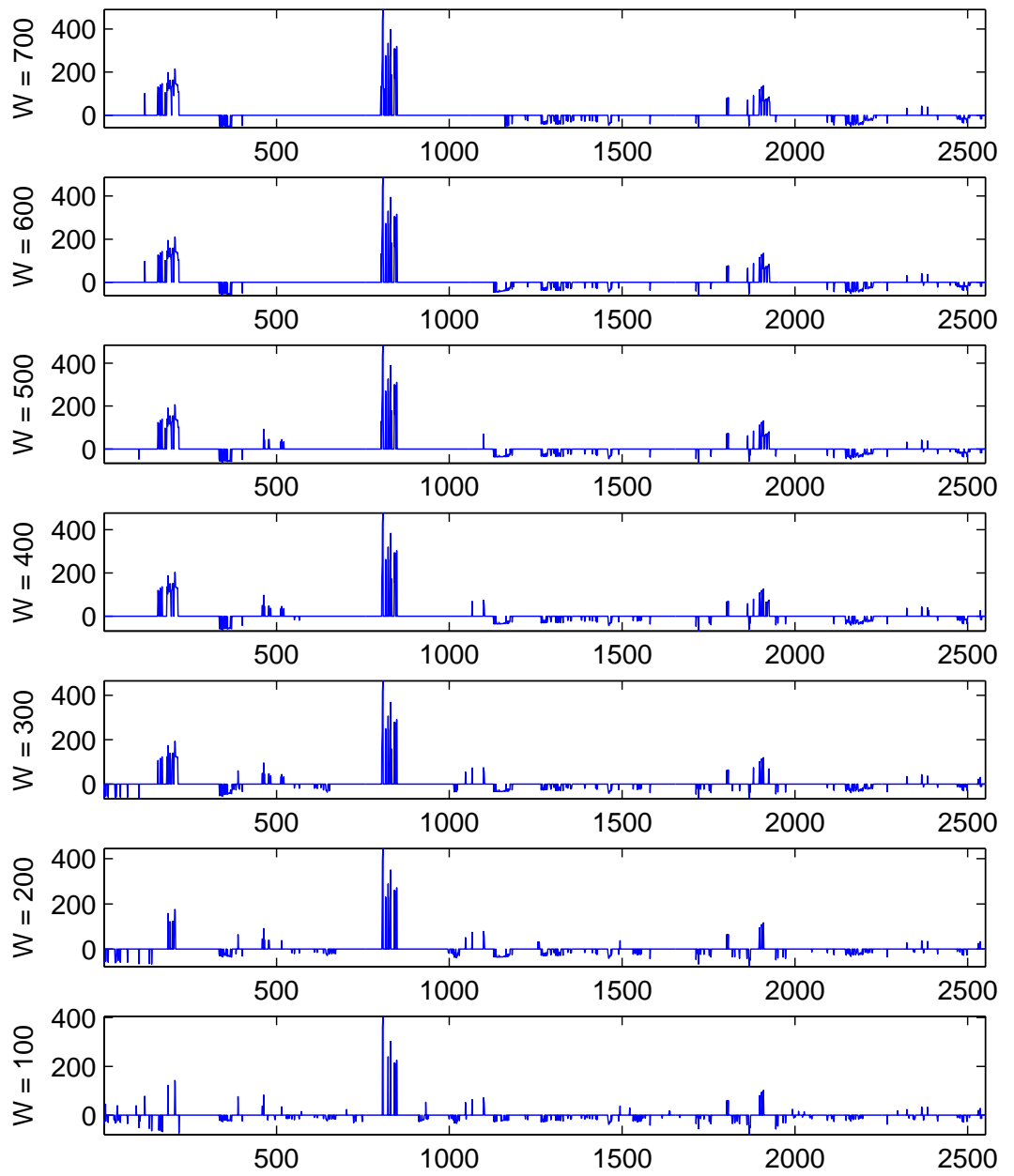


Figure 30: Otahuhu spike vectors for different window sizes.

The basic statistics for the NEPool and Otahuhu spike vectors are collected in Tables 6 and 7 respectively.

Table 6: Basic statistics of NEPool spike vectors for different window sizes.

window	count	mean value	frequency (per 100 days)
700	85	57.4704	4.5921
600	92	55.2749	4.7155
500	83	55.8887	4.0468
400	90	53.0616	4.1841
300	76	54.6654	3.3763
200	86	49.8506	3.6580
100	69	47.9029	2.8152

Table 7: Basic statistics of Otahuhu spike vectors for different window sizes.

window	count	mean value	frequency (per 100 days)
700	99	208.4502	3.8808
600	91	229.6558	3.5672
500	98	208.8678	3.8416
400	91	205.4020	3.5672
300	77	222.9197	3.0184
200	52	272.3904	2.0384
100	49	214.6862	1.9208

A promising observation is that despite different window sizes chosen, the average magnitudes of spikes are comparable. Also their frequency calculated per 100 days is of the same order. This can mean that it may be possible to predict at least spikes occurrences and magnitudes.

Now let us analyze the price series with spikes subtracted from them. In Figures 31 and 32, such series are presented for NEPool and Otahuhu respectively. We can see that the prices without spikes are more stationary than with them. Table 8 collects decisions for H_0 of DF, PP and KPSS tests. Moreover, we can notice that formal results are ambiguous in verifying price stationarity for all window sizes. On the other hand, all return series are clearly stationary (DF/PP H_0 rejected for all and KPSS H_0 accepted for all).

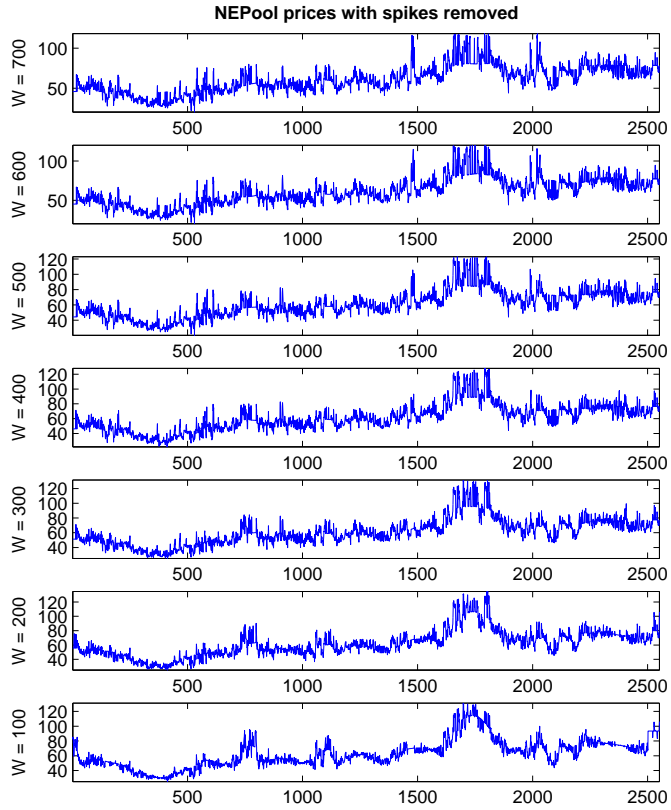


Figure 31: NEPool price series with removed spikes.

Table 8: H_0 decisions of DF, PP and KPSS tests for NEPool and Otahuhu prices with spikes removed.

	NEPool			Otahuhu		
	DF	PP	KPSS	DF	PP	KPSS
W = 700	rejected	rejected	rejected	rejected	rejected	rejected
W = 600	rejected	rejected	rejected	rejected	rejected	rejected
W = 500	rejected	rejected	rejected	rejected	rejected	rejected
W = 400	rejected	rejected	rejected	rejected	rejected	rejected
W = 300	rejected	rejected	rejected	rejected	rejected	rejected
W = 200	rejected	rejected	rejected	rejected	rejected	accepted
W = 100	rejected	rejected	rejected	rejected	rejected	accepted

Unfortunately, in spite of clustering character of the spikes, clustering in prices is not completely eliminated. The Engle's test still indicates that all created series reveal heteroscedasticity. Finally, we can see that especially in Otahuhu case, the smallest windows do not eliminate much of the highest spikes.

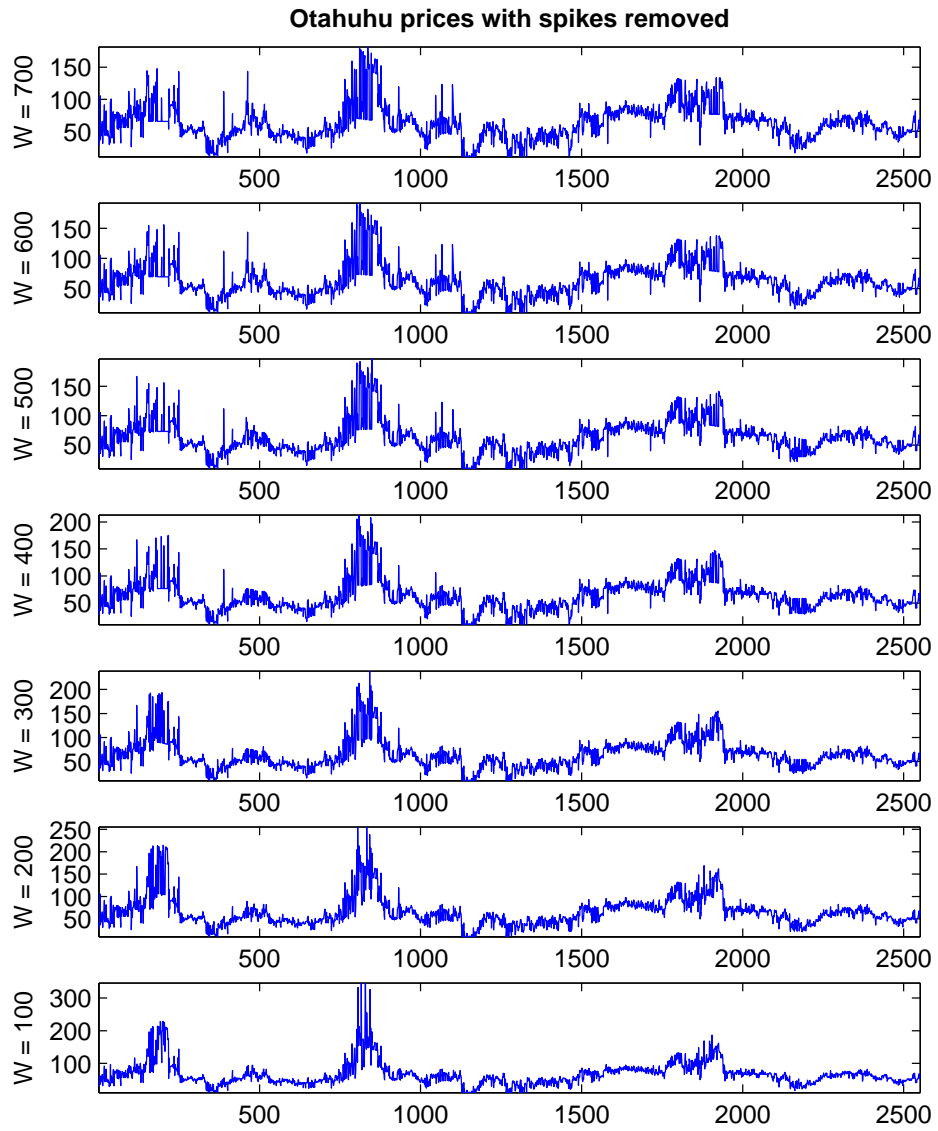


Figure 32: Otahuhu price series with removed spikes.

For prices with outliers removed, we can study the returns as well. In this case we are mostly interested in the shape of histograms and the fact how close the returns are to be accepted as normally distributed. Therefore, let us plot the NEPool new returns histograms for different window sizes (see Figure 33) against histogram of the original New England price returns (Figure 33, bottom right panel).

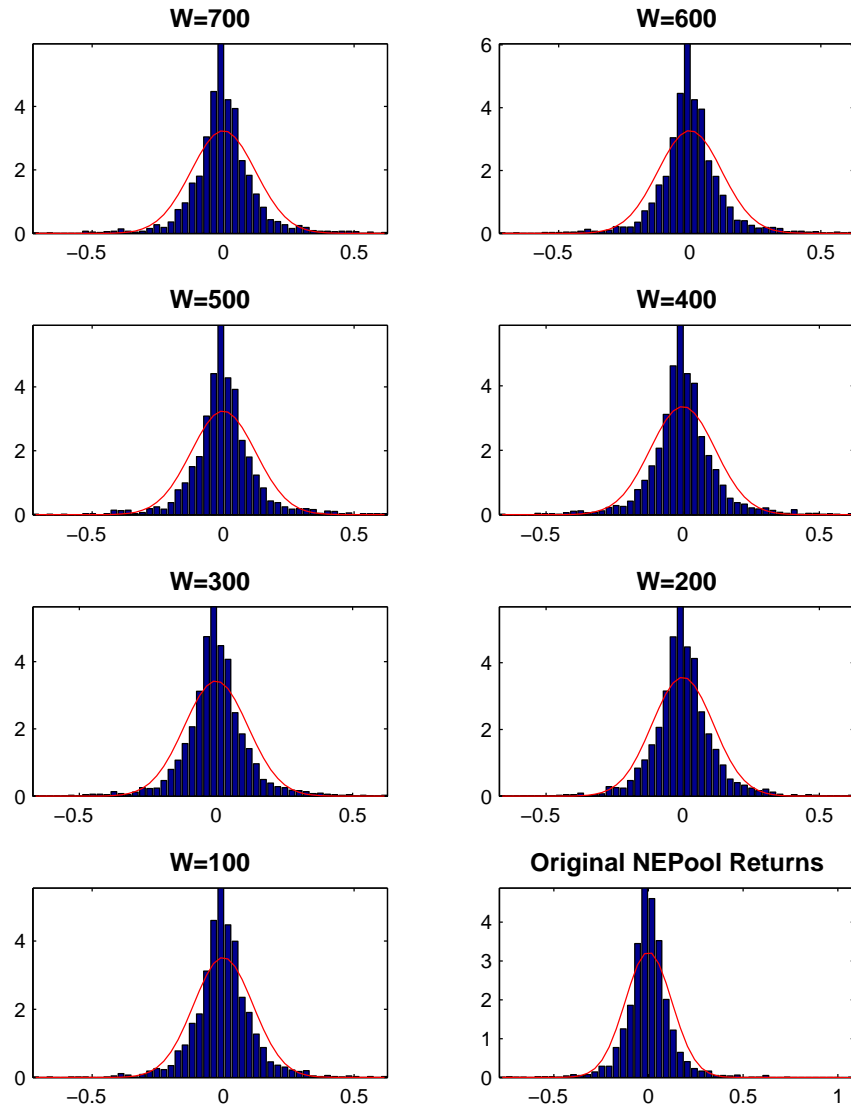


Figure 33: Histograms of NEPool price returns series with removed spikes against original returns.

We can see, that all histograms look alike, which means that removing spikes did not change much of the data character. Similarly, we plot the histograms for Otahuhu data as presented in Figure 34, with the original returns' histogram for comparison in the bottom right panel.

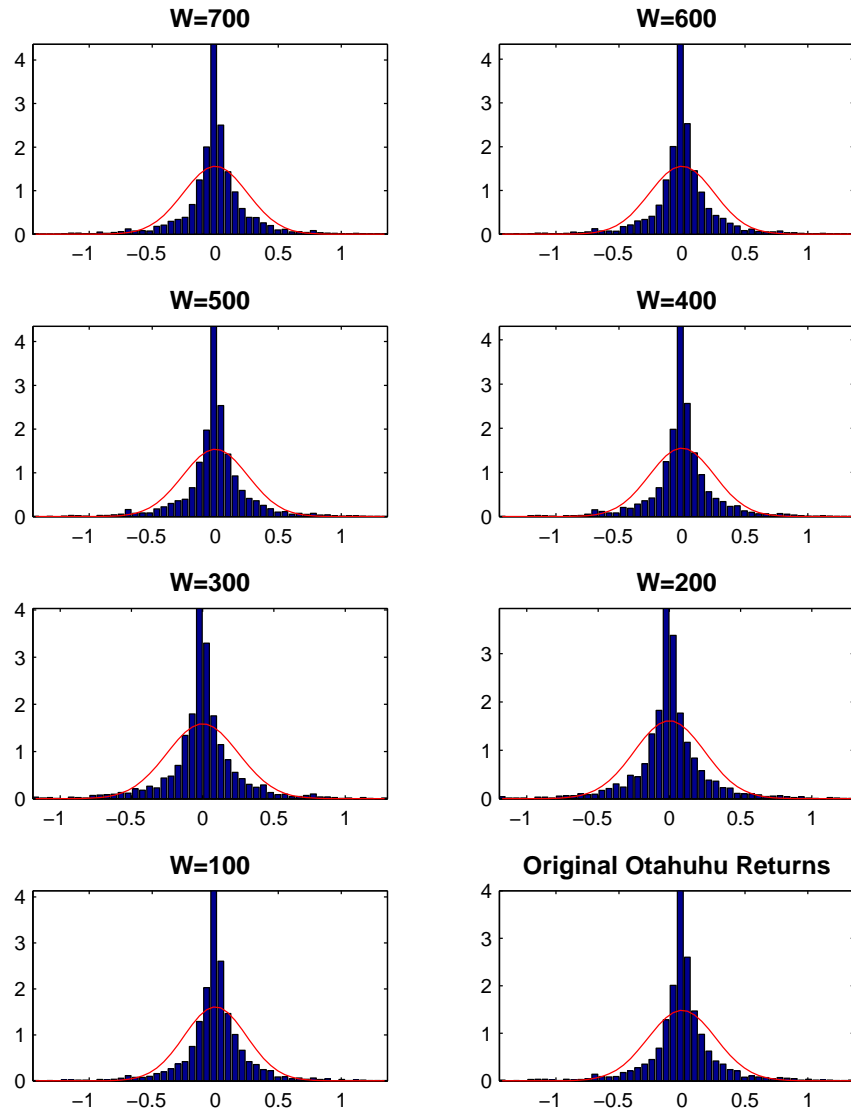


Figure 34: Histograms of Otahuhu price returns series with removed spikes against original returns.

To sum up we can conclude that Otahuhu as well as NEPool returns remain comparable for different window size. However, their shapes are far from normal distribution.

4.2 Spikes with respect to price series smoothed by DFT

As we saw in the previous section, considering spikes as outliers with regard to the original price series did not give any positive results neither for NEPool nor for Otahuhu. Another

way to distinguish outliers is to compare the prices with smoothed ones. Here we process similarly to the previous analysis with the same window sizes but now comparing each price $x(i)$ with values $\mu_s + 2\sigma_s$ and $\mu_s - 2\sigma_s$, where μ_s is the mean value and σ_s is the standard deviation of smoothed series in a given window.

Figure 35 presents spikes obtained for New England data and Figure 36 – for New Zealand.

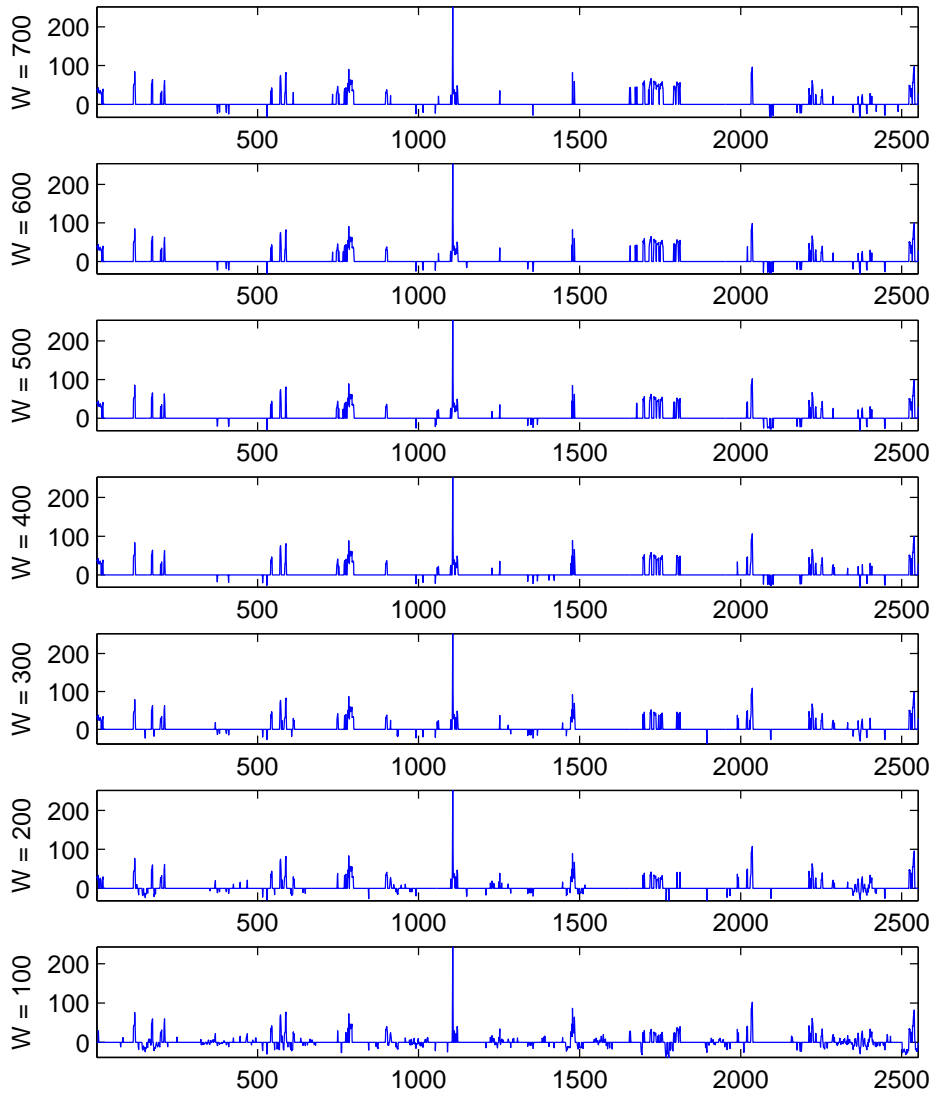


Figure 35: NEPool spike vectors for different window sizes.

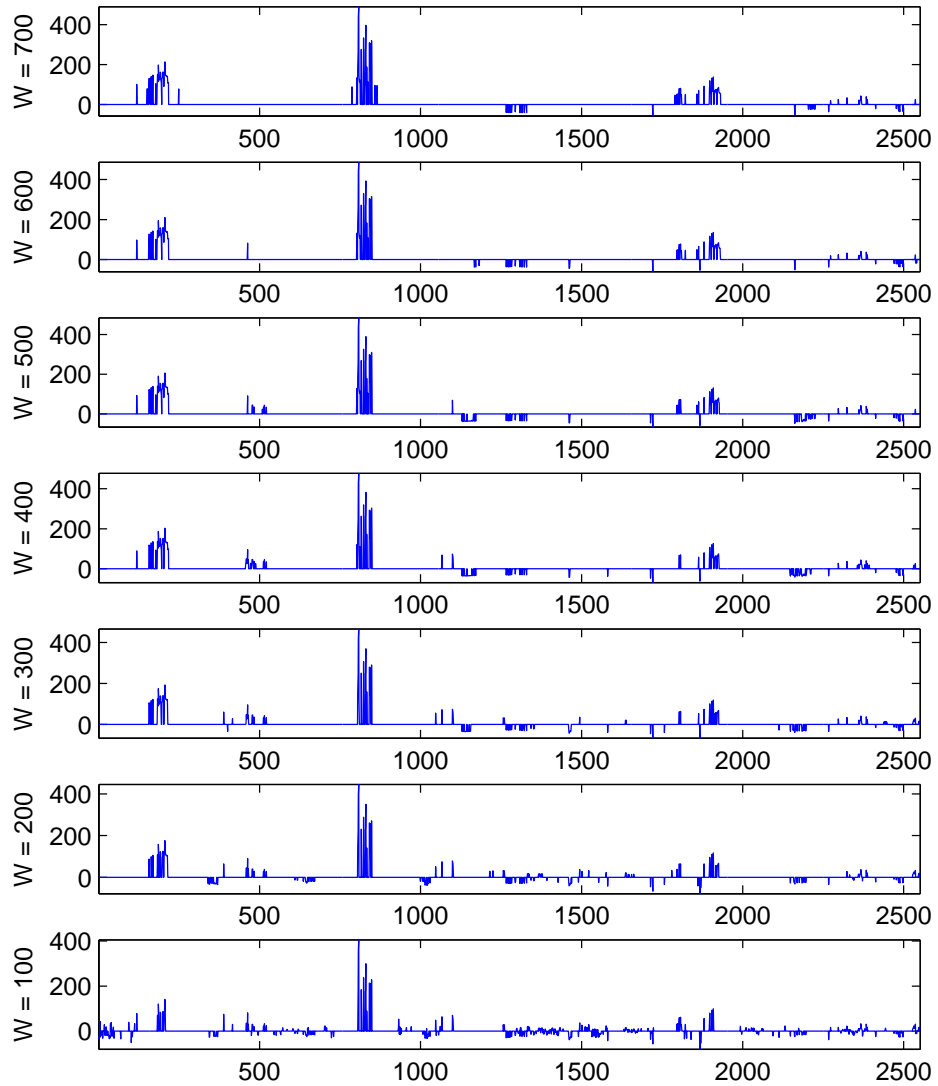


Figure 36: Otahuhu spike vectors for different window sizes.

Analogically, we can observe clustering character of the spikes for both cases, which again means that removing spikes from original series may change the overall nature of the data. However, firstly, we compute the basic statistics for NEPool (Table 9) and Otahuhu (Table 10). The values are comparable for different window sizes. For the smallest width (100), however, for both series we notice the lowest mean values (for Otahuhu almost twice lower than for window size 200) and highest standard deviations (for Otahuhu more than 1.6 times higher when compared to the other windows).

Table 9: Basic statistics of NEPool spike vectors for different window sizes.

window	count	mean value	frequency (per 100 days)
700	224	48.8817	8.7809
600	225	49.1919	8.8201
500	216	49.0352	8.4673
400	213	47.2891	8.3497
300	218	44.2099	8.5457
200	242	41.7792	9.4865
100	323	37.7015	12.6617

Table 10: Basic statistics of Otahuhu spike vectors for different window sizes.

window	count	mean value	frequency (per 100 days)
700	142	126.1144	5.5664
600	134	126.9550	5.2528
500	128	137.2787	5.0176
400	137	124.7104	5.3704
300	133	109.5491	5.2136
200	132	100.8475	5.1744
100	234	54.0570	9.1729

Next, we plot the new series with spikes removed; see Figure 37 for NEPool and Figure 38 for Otahuhu. Similarly, we can see that the smallest window eliminates more volatility from the prices. This, however, does not clear away ARCH effect in neither series. The Engle’s test rejects the null hypothesis which means that in spite of removing spikes, there is still some heteroscedasticity in the prices and returns.

Moreover, we again obtain conflicting results when it comes to assessing stationarity (see Table 11). Lack of unit root is confirmed only for the smallest window sizes. However, all tests give unambiguous results for return series – they are stationary for all window sizes.

Table 11: H_0 decisions of DF, PP and KPSS tests for NEPool and Otahuhu prices with spikes removed.

	NEPool			Otahuhu		
	DF	PP	KPSS	DF	PP	KPSS
W = 700	rejected	rejected	rejected	rejected	rejected	rejected
W = 600	rejected	rejected	rejected	rejected	rejected	rejected
W = 500	rejected	rejected	rejected	rejected	rejected	rejected
W = 400	rejected	rejected	rejected	rejected	rejected	rejected
W = 300	rejected	rejected	rejected	rejected	rejected	accepted
W = 200	rejected	rejected	rejected	rejected	rejected	accepted
W = 100	accepted	accepted	rejected	rejected	rejected	accepted

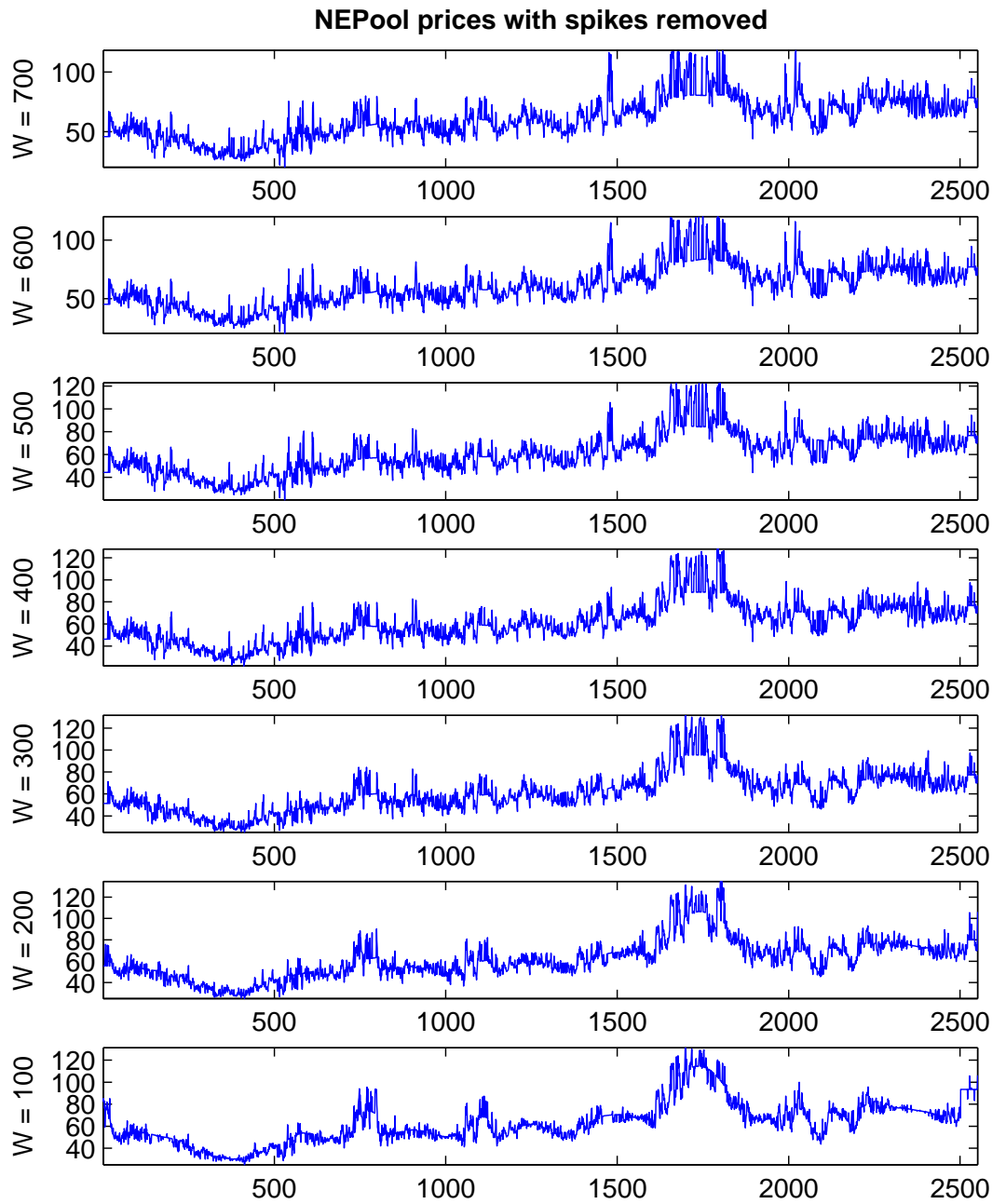


Figure 37: NEPool price series with removed spikes.

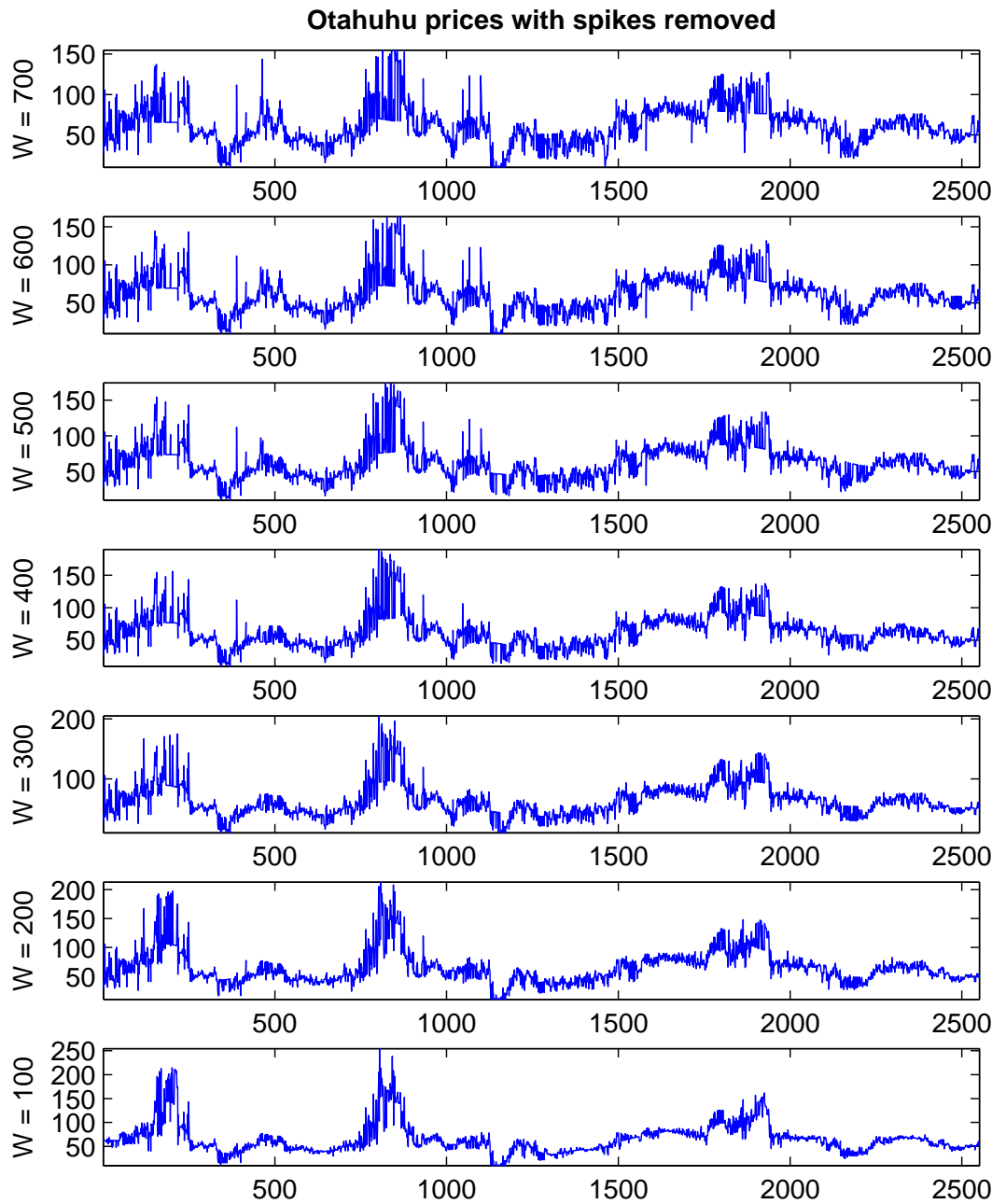


Figure 38: Otahuhu price series with removed spikes.

Finally, we again verify what influence can be observed on the prices when we remove spikes from the series. Plots of histograms compared to the original returns histograms are presented in Figures 39 and 40 for NEPool and Otahuhu respectively.

An effect of removing spikes is mostly visible for the smallest window sizes. The general shape of histograms becomes much closer to theoretical normal PDFs only for NEPool series. On the other hand, they gain a significant concentration of zero or close to zero values.

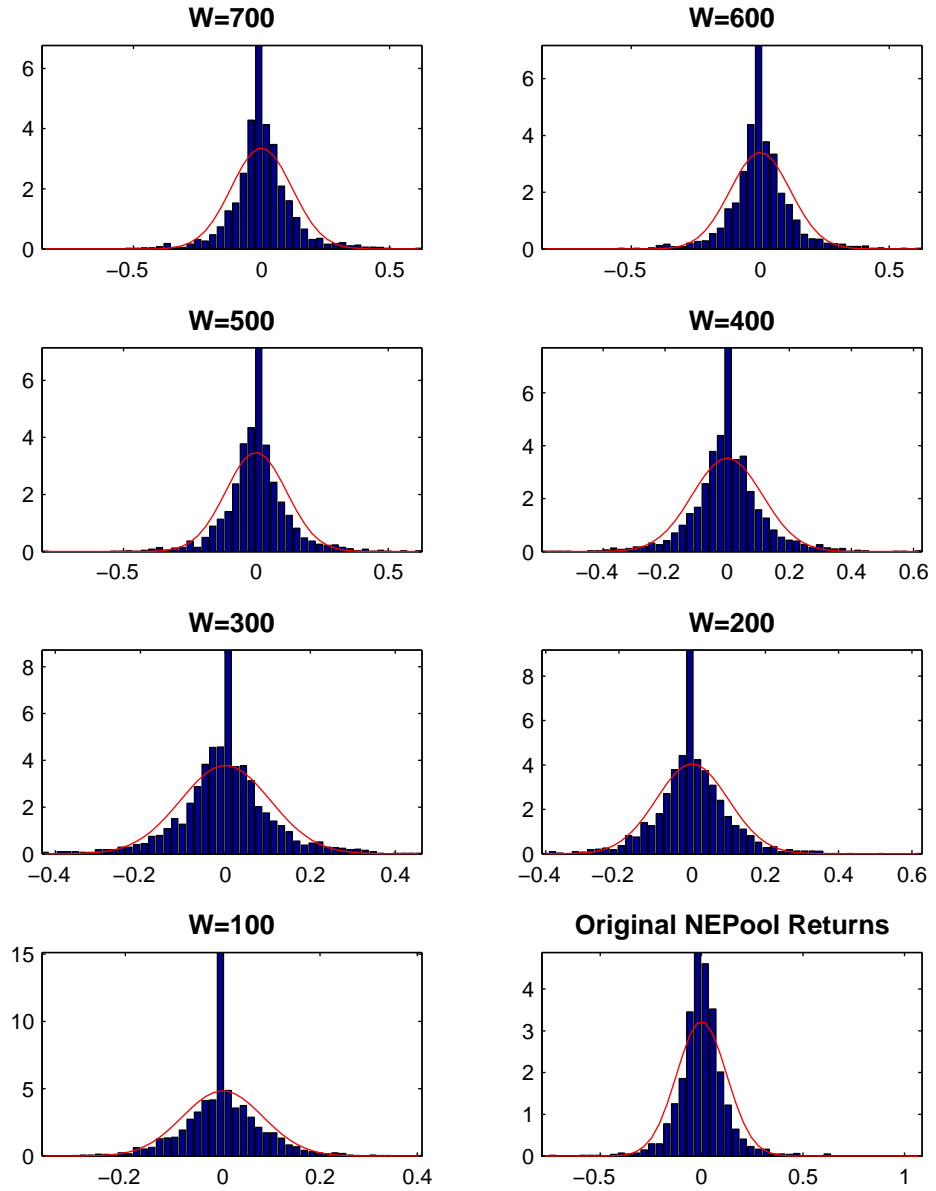


Figure 39: Histograms of NEPool price returns series with removed spikes against original returns.

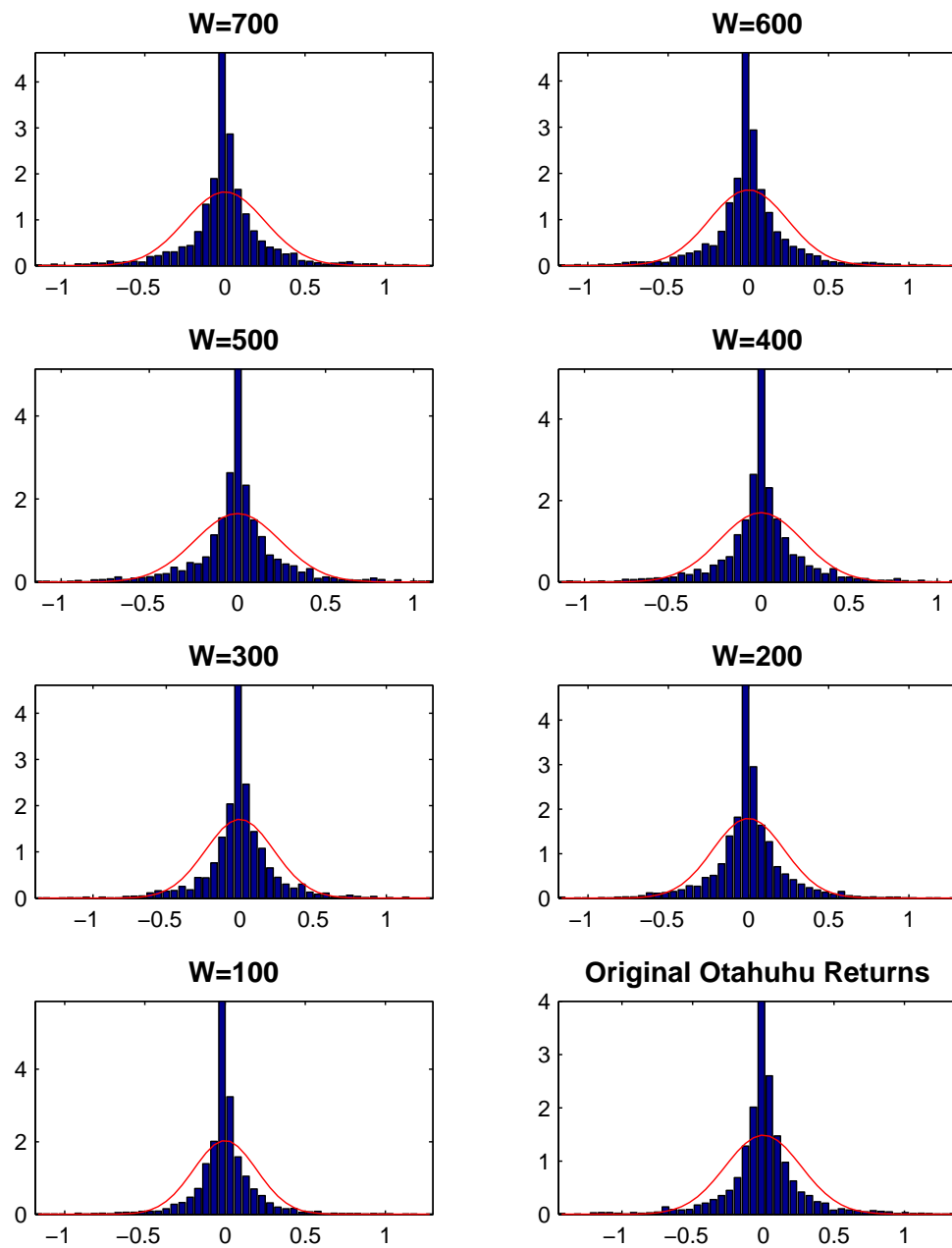


Figure 40: Histograms of Otahuhu price returns series with removed spikes against original returns.

4.3 Outliers in return series

Not only prices have spikes. The return series reveal considerable outliers as well. Therefore, we analyze them in this section. In this case we do not perform any smoothing, thus analysis of spikes can be performed only with respect to the original return series. However, there is still an option of doing it either from the all data point of view or by window analysis analogical to the one from previous section.

An outlier is defined in the same way as in the price analysis. A return $r(i)$ is considered a spike if its magnitude crosses either $\mu + 2\sigma$ or $\mu - 2\sigma$, where μ is mean value of returns (whole series or in a given window) and σ stands for the returns standard deviation.

Let us start from the general viewpoint, which means specifying spikes with respect to the whole time horizon. In upper panel of Figure 41 we plot the NEPool return series with red lines meaning the $\mu \pm 2\sigma$ outlier acceptance/rejection boundaries. The bottom panel shows obtained spike vector with 1st of each January and July marked on the X axis.

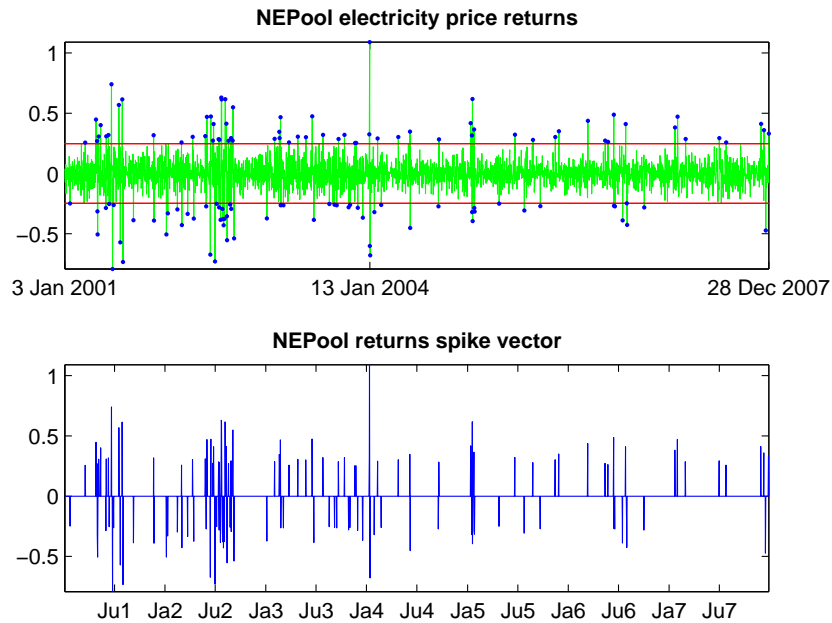


Figure 41: NEPool price returns with outlier acceptance/rejection boundaries.

We can see that similarly to price spikes, return outliers show a clearly clustering character. Figure 42 presents analogical plots for Otahuhu series. These ones also do not appear uniformly within the time horizon.

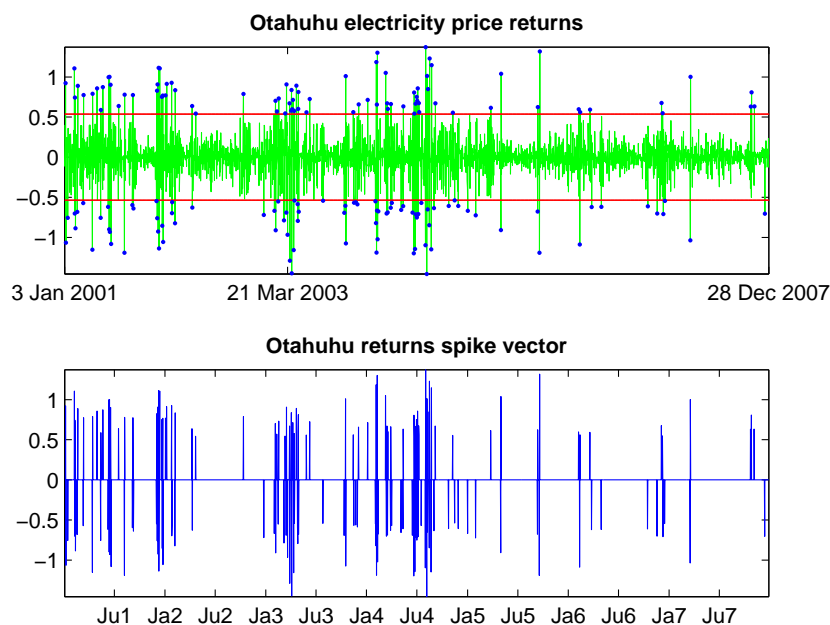


Figure 42: Otahuhu price returns with outlier acceptance/rejection boundaries.

Moreover, one can say that some of the spike clusters look analogical in both series, however, with half year shift between the data sets. This is easily explainable by geographical difference between the regions. *I.e.* spikes appearing in NEPool returns around the 1st of July 2002 would correspond to similarly looking outliers in Otahuhu data around 1st of January 2002. In the same way NEPool spikes after 01-07-2002 and around 01-07-2006 look alike Otahuhu ones after 01-01-2003 and after 01-01-2007 respectively. We conclude that it results from price and return volatility typically higher in summer months. There is also a year with higher density of spikes for both series: from Jan 2003 to Jan 2004 for New England and from Jul 2003 to Jul 2004 for New Zealand.

Table 12 collects basic statistics for both spike vectors. We notice that Otahuhu return series could be considered as the more spiky one. It not only has more outliers, but they are of higher absolute magnitudes and higher standard deviation. The NEPool return spikes make around 5.12% of all observations, whereas the New Zealand vector reaches more than 6.78% of all data.

Table 12: Basic statistics for NEPool and Otahuhu return spike vectors.

	NEPool return spikes	Otahuhu return spikes
count	132	173
percent of all	5.1765%	6.7843%
mean	0.0154	-0.0247
st dev	0.4003	0.8218

Considering the overall character of outliers occurrence, we investigate possible seasonality in it. Therefore, we plot ACFs and PACFs for NEPool (Figure 43) and Otahuhu (Figure 44) spike vectors. The results do not show any significant values at lags which could contain seasonality, *i.e.* 7th or 30th.

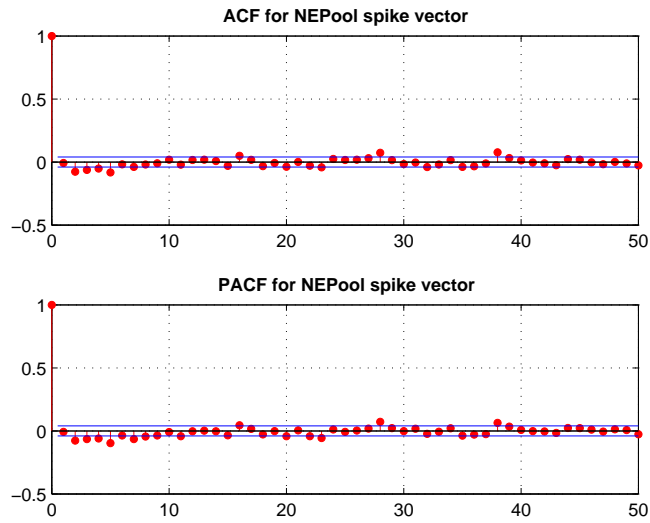


Figure 43: ACF (upper panel) and PACF (bottom panel) of spike vector for NEPool returns.

Even though we do not find any seasonal patterns in the spikes, we can take one more step in the analysis. Similarly to prices, outliers are removed from the return series. However, in this case we replace each of them by a normally distributed number with mean and standard deviation equal to the original returns' parameters. Figures 45 and 46 present obtained new return series together with new autocorrelation and partial autocorrelation functions.

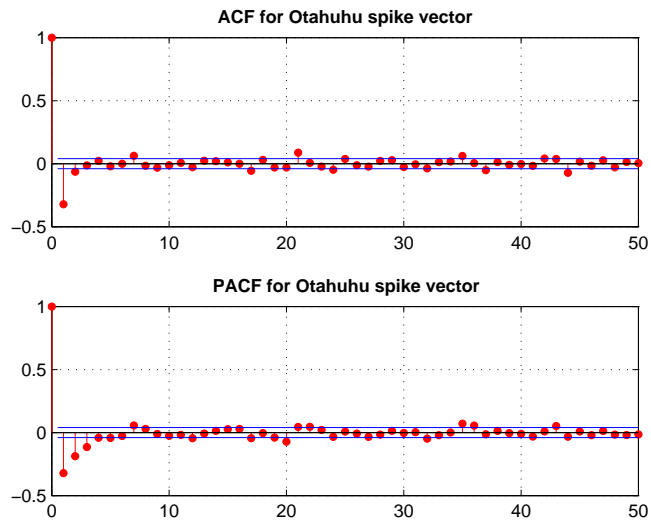


Figure 44: ACF (upper panel) and PACF (bottom panel) of spike vector for Otahuhu returns.

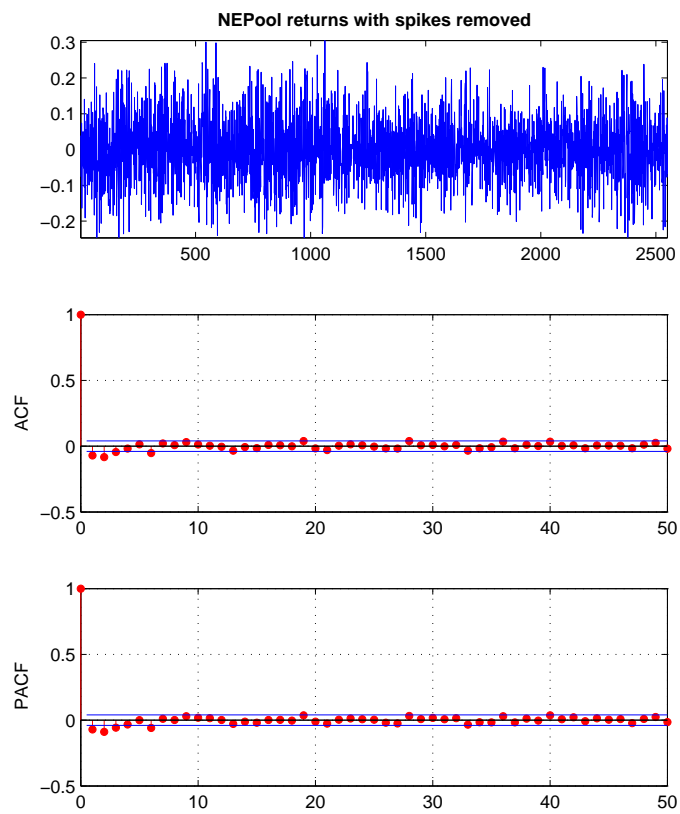


Figure 45: NEPool returns with spikes removed and their ACF and PACF.

On the other hand we can see that removing outliers from Otahuhu returns slightly emphasized significance of both ACF and PACF values at every 7th lag, which indicates that there exist some weekly patterns. This feature is not visible for NEPool.

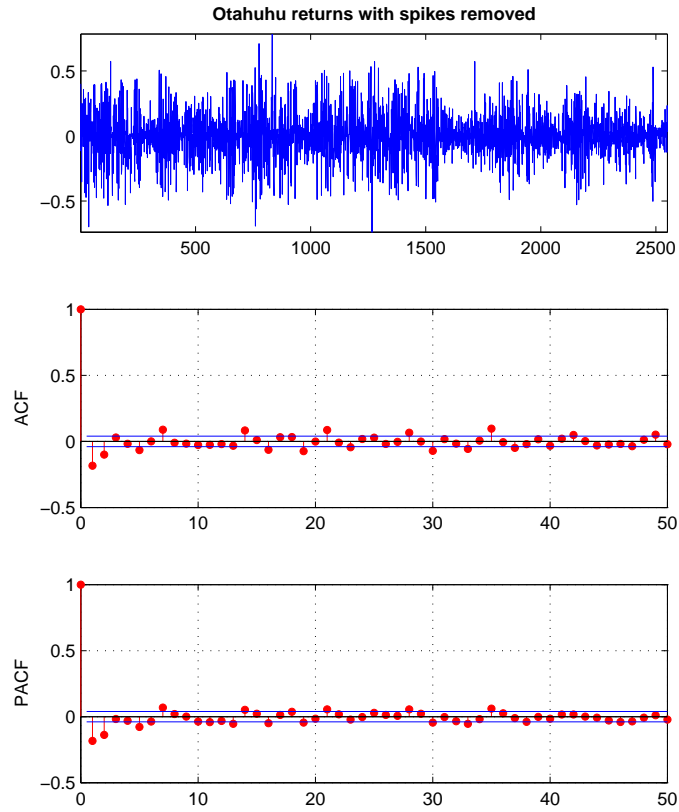


Figure 46: Otahuhu returns with spikes removed and their ACF and PACF.

Finally, we investigate distribution of returns without spikes. Figure 48 presents normalized histograms for both modified return series against theoretical normal PDFs. Even though the formal normality test still rejects the null hypothesis (see Table 13), we can see that especially NEPool histogram is much closer to normal distribution when compared to Figure 6.

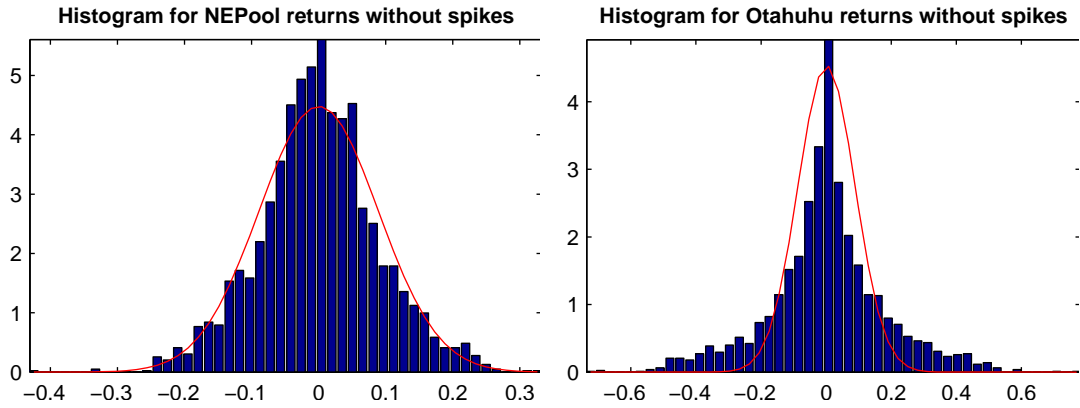


Figure 47: Histograms for NEPool (left panel) and Otahuhu (right panel) returns with spikes removed.

Table 13 presents also values for skewness and kurtosis of modified returns, which again shows improvement with respect to the original series (see Table 2). Since, in particular for NEPool, the parameters are reasonably close to the model ones (0 for skewness and 3 for kurtosis) we perform one more formal normality test. Jarque-Bera test verifies two-sided goodness of fit with its statistic

$$JB = \frac{n}{6} \left(s^2 + \frac{(k - 3)^2}{4} \right) \quad (7)$$

where n is the sample size, s – the sample skewness and k – the sample kurtosis. The values of test statistic come from Chi-square distribution with 2 degrees of freedom.

Table 13: Normality statistics for NEPool and Otahuhu price returns with spikes removed.

	NEPool	Otahuhu
skewness	0.0146	-0.0449
kurtosis	3.3608	4.4885
Lilliefors test H_0	rejected	rejected
Jarque-Bera test H_0	rejected	rejected

After this analysis we may conclude that highlighting spikes in returns themselves had a more positive effect from normality point of view, than eliminating spikes from prices.

4.4 Predictability of spikes based on seasons and price volatility

4.4.1 Outliers vs. year seasons

For this part of work only the highest spikes from both series are chosen. The aim of this analysis is to investigate whether it is possible to find some characteristics of spikes based on the year seasons when they occur.

Analysis of NEPool 21 highest outliers shows that more than half of them appeared in summer time. The rest occurred in either spring/fall or winter. In Figure 48 the chosen values are presented with 10 time steps backwards and 10 time steps ahead. It is difficult to split the magnitudes with respect to seasons. However, difference between the highest and lowest summer spike is significantly smaller than in case of two other season groups. Therefore, it can be stated that summer outliers in New England are more frequent but less volatile in magnitudes.

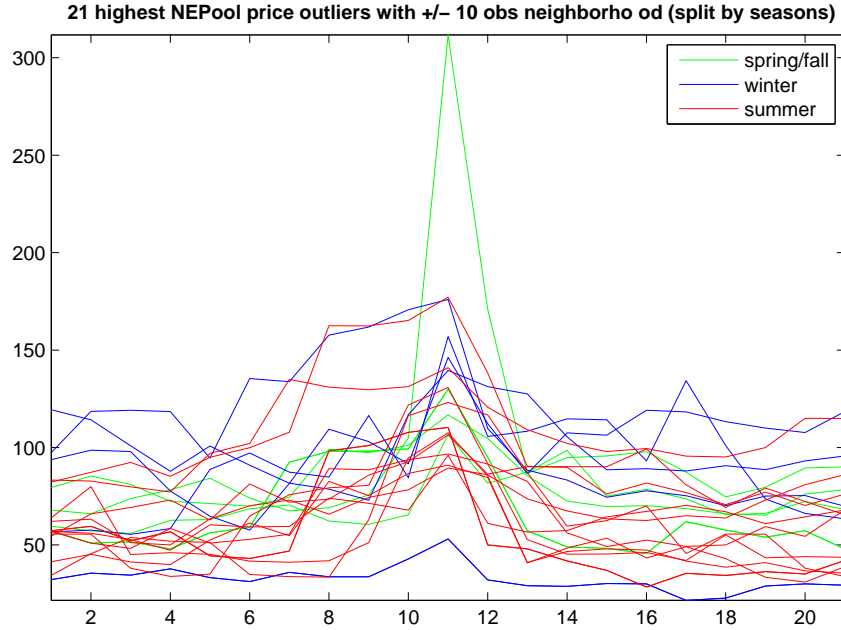


Figure 48: NEPool price 21 highest spikes split by season of occurrence.

We can also analyze observations around the spikes. Table 14 presents the chosen 21 NEPool outliers (column 1), difference between prices 10 time steps to the right and 10 time steps to the left from the given spike (column 2) and, finally, calculated difference as a percentage value of the outlier (column 3). It appears that the maximal positive and negative differences are of the same order – roughly 26.8% and -26.4%. On the other hand, the average value for positive differences is 13.2% while the negative one reaches only approximately -9.29%.

In case of New Zealand data the analysis of 34 highest outliers gives a slightly different outcome when compared to NEPool. Summer spikes appear to be the least frequent, though also have the smallest spread of magnitudes. Moreover, most winter spikes are within a similar range as summer ones. Outliers occurring in spring and fall show the highest disturbances and volatility; they are often preceded or followed by other spikes within distance of 10 observations backwards or ahead (see Figure 49).

Table 14: NEPool 21 highest spikes split with respect to season of occurrence.

spring/fall spikes	difference between +/- 10 obs	difference as percent of spike
130.0000	-9.3289	7.1761
311.7500	22.2064	7.1231
106.7500	8.6557	8.1084
116.7600	-1.0900	0.9335
winter spikes	difference between +/- 10 obs	difference as percent of spike
53.1030	-2.8800	5.4234
146.2000	24.6800	16.8810
156.9048	6.6971	4.2683
139.5300	-23.6800	16.9713
175.9701	-27.6300	15.7015
summer spikes	difference between +/- 10 obs	difference as percent of spike
110.2951	-15.1800	13.7631
107.5000	-28.3812	26.4011
90.9204	-2.5709	2.8277
123.0858	-0.5300	0.4306
130.8000	-11.9261	9.1179
79.5300	10.4160	13.0969
96.6669	11.1000	11.4827
96.0000	11.0000	11.4583
141.0000	31.6484	22.4457
177.1800	-6.1300	3.4598
106.5300	10.9974	10.3233
89.5066	23.9900	26.8025

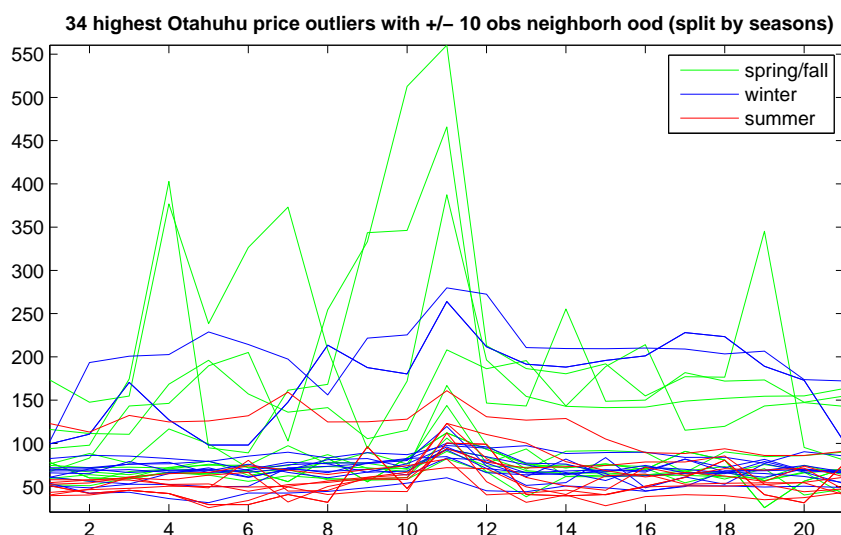


Figure 49: Otahuhu price 34 highest spikes split by season of occurrence.

Analogically to Table 14, Table 15 presents outcome of spikes' neighborhood analysis. Similarly to NEPool, Otahuhu differences between prices 10 days before and after a spike

occurrence are in 50 percent positive and in 50 percent negative. The level of maximal positive and negative differences are of similar order as well. Moreover, the average values for them are very close for Otahuhu (9.35% and -9.6%) while being comparable with NEPool results as well.

Table 15: Otahuhu 34 highest spikes split with respect to season of occurrence.

spring/fall spikes	difference between +/- 10 obs	difference as percent of spike
112.1900	-7.0400	6.2751
166.8100	13.7600	8.2489
143.6900	11.4100	7.9407
560.2200	13.5100	2.4116
465.8800	60.6400	13.0162
387.4500	-9.8400	2.5397
106.0700	-12.4000	11.6904
93.4000	-33.3100	35.6638
42.9400	-5.1900	12.0866
208.1100	26.7700	12.8634
82.3200	8.4400	10.2527
90.7600	-1.5100	1.6637
winter spikes	difference between +/- 10 obs	difference as percent of spike
263.9200	2.7000	1.0230
279.8400	70.1200	25.0572
92.3900	-6.3100	6.8297
60.2900	-4.8700	8.0776
119.5100	2.4000	2.0082
97.5400	2.9300	3.0039
40.4500	-5.1100	12.6329
93.5900	-5.5000	5.8767
84.6500	8.2600	9.7578
82.9400	-0.0500	0.0603
51.2600	8.7200	17.0113
49.4300	-12.0000	24.2768
100.0300	-3.0200	3.0191
95.8300	-1.1200	1.1687
summer spikes	difference between +/- 10 obs	difference as percent of spike
100.3600	25.0600	24.9701
111.8900	3.7700	3.3694
71.8300	-2.5700	3.5779
123.0500	6.3900	5.1930
123.0100	3.4100	2.7721
94.0500	-7.2500	7.7087
160.9400	-32.1500	19.9764
81.9300	8.1700	9.9719

Presented groups of spike neighborhoods can also be averaged over the seasons so

that it is possible to compare their general shape. For these results see Figures 50 and 51 for NEPool and Otahuhu respectively.

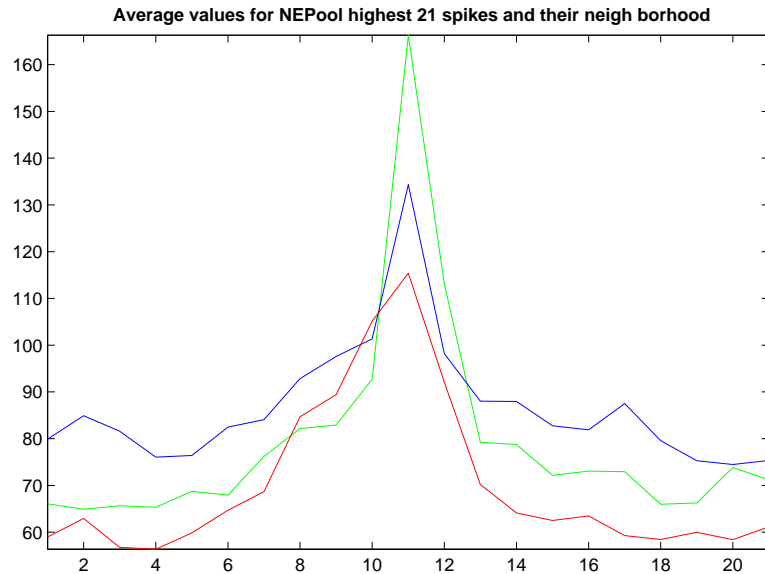


Figure 50: Average values for NEPool 21 highest outliers and their neighborhood.

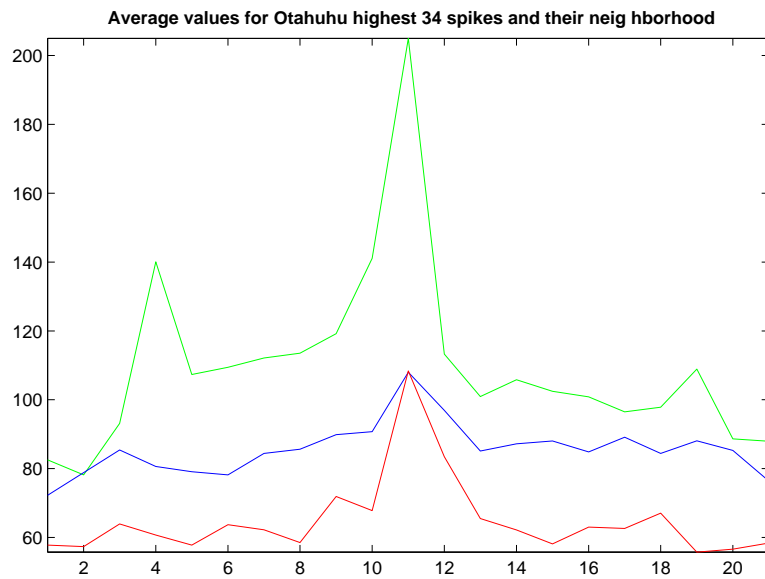


Figure 51: Average values for Otahuhu 34 highest outliers and their neighborhood.

The analysis shows that there are both common and distinct features in spikes and their neighborhood in both price series, when split by year seasons. Differences between seasons when the outliers occur may stem from geographical differences between the regions. New Zealand is a country surrounded by seas, thus the weather in spring and fall in that area might be more suddenly changing and causing sudden jumps of electricity demand and, therefore, prices. This, however, would require verification with meteorological data.

4.4.2 Outliers vs. price volatility

While comparing Figures 4 and 5 with Figures 29 and 30 we may notice that each time a spike appears, price volatility increases around the given spike time point. Therefore, we might be interested in a brief analysis whether a spike occurrence is usually preceded by increase in variance. If so, it would mean that it could be somehow possible to expect a spike when the price volatility is rising.

As it was verified in Section 3.3 both price series reveal heteroscedasticity, *i.e.* not constant variance over the time horizon. The ACF/PACF analysis suggested fitting ARMA(1,0) models to both price series. Combining this fact with heteroscedasticity we get the following approach for modeling:

1. fit ARMA(1, 0) models
2. find residuals for the models
3. estimate conditional sigmas for the series

Let us perform the first step. We obtain the regression coefficients $(\phi_1, C)_{ne} = (0.9117, 5.6925)$ and $(\phi_1, C)_{ot} = (0.8321, 11.2907)$ for NEPool and Otahuhu respectively. The second stage leads us to a general view of the models' residuals. Figures 52 and 53 present the original data series plotted against fitted ARMA(1,0) models together with residuals.

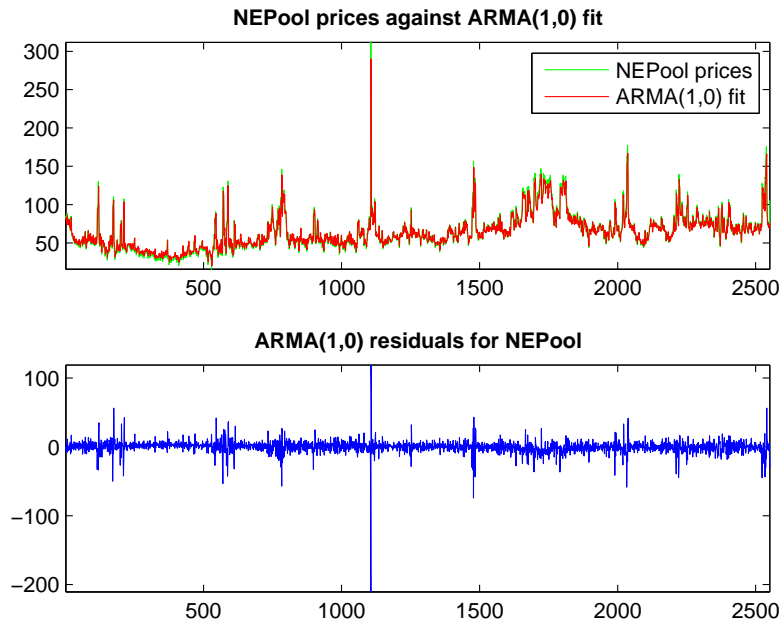


Figure 52: NEPool prices against ARMA(1,0) fit (upper panel) and model residuals (bottom panel).

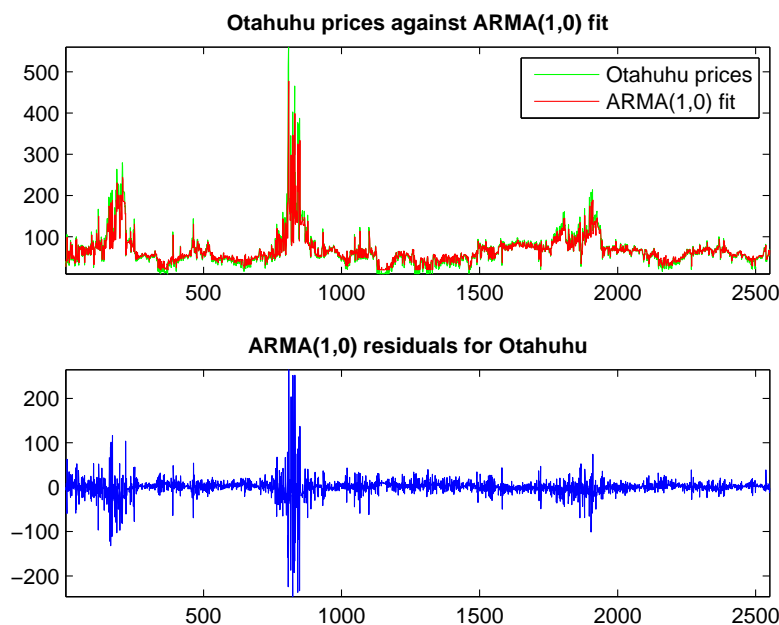


Figure 53: Otahuhu prices against ARMA(1,0) fit (upper panel) and model residuals (bottom panel).

We can notice clear heteroscedasticity in the residuals. Therefore, as mentioned before, we fit GARCH models. Consistent with simplicity approach, we choose GARCH(1, 1) for both cases. The obtained estimates are presented in Table 16.

Table 16: GARCH estimates for NEPool and Otahuhu ARMA(1,0) residuals.

	NEPool estimates	Otahuhu estimates
K	8.8250	3.2981
α_1	0.4285	0.1400
β_1	0.5621	0.8573

Now we are interested in comparison of the series variance estimates with spike occurrence. From the MATLAB `garchfit` procedure we obtain estimates for the volatility in prices. Having those as series, we may plot them against the spikes presented in Figures 29 and 30 (Section 4.1). For the analysis we choose the outliers obtained for the smallest window size. Results are plotted in Figures 54 and 55 for NEPool and Otahuhu respectively. Indeed, there is an increase in variance visible around most of the spikes for both series. However, only for some cases there happens an increase of volatility before the outlier occurs.

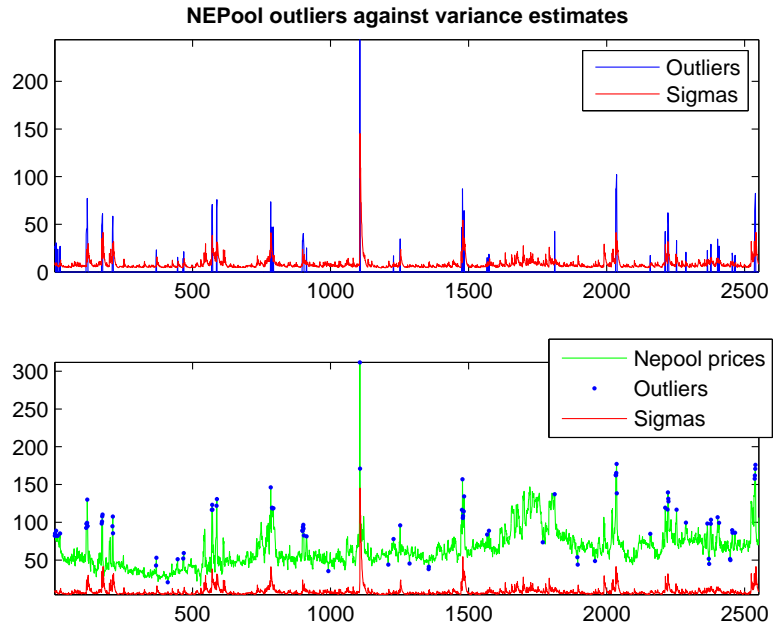


Figure 54: NEPool outliers against price variance estimates.

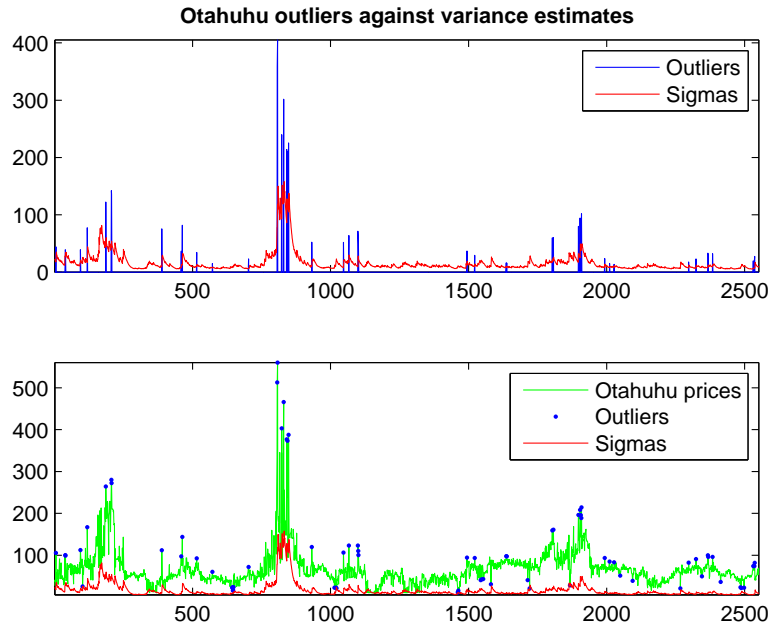


Figure 55: Otahuhu outliers against price variance estimates.

Since graphical comparison of spikes and variance estimates shows clear relation between them, correlation investigation may provide more information in this matter. Correlation coefficient between NEPool spikes and variance series equals 0.3096 which is accepted as significant. However, if we calculate the correlation between the spike series

and lagged (1 lag) volatility estimates, the coefficient rises up to 0.4696. Similarly, for Otahuhu series correlation we obtain coefficient equal to 0.2349, which is also significant, but when proceeding with the same lagging operation the number increases up to 0.3144 for 1 lag and 0.3276 for 2 lags. This simply means that it is more probable that variance will increase after the spike than before. On the other hand, some general conclusion can be made:

- If a high spike occurs, the variance will jump and then gradually come back to its previous level.
- If a spike occurs after a long stable period, there will appear a similar (in magnitude) spike within 3-6 days after the first one.
- If variance rises twice or more, there will most probably either be a long period with persistently high variance or a spike will appear.

4.5 What do electricity suppliers really earn on spikes?

We mentioned before that the last several years in financial world seem to show that 'spiky' behaviour of time series is no longer an exception, but rather a regular phenomenon. Therefore, we might suppose that *e.g.* in electricity branch of industry outliers make a significantly high income for the suppliers.

Within the original data set we do not have any information about the real sales of electricity within either of analysed regions. Thus we simply calculate the revenues as integral of the series. Figure 56 presents plots of NEPool revenues cumulated within the 7 years of time horizon based on both all prices and the spikes only. It appears that the revenues calculated only from spiky time points made nearly 10% of income from 3 Jan 2001 to 28 Dec 2007. Of course, these revenues are just intuitive numbers not including electricity production expenses or the real sales. Most likely the percentage relation between the outlier and total income would be even higher, since we know that jumps in prices appear when there is a congestion and the supply is maximal possible.

For comparison, we also choose the month with the highest outlier in the NEPool data set, *i.e.* January 2004, and calculate the cumulated revenues from this period only. It appears that within one month spikes made almost up to 50% of total income.

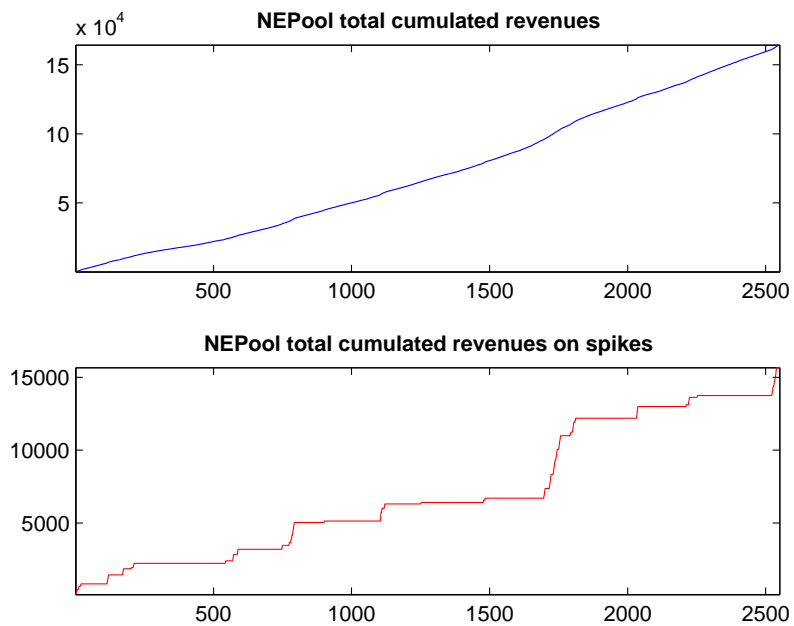


Figure 56: NEPool total cumulative revenues (upper panel) and total revenues from outliers only.

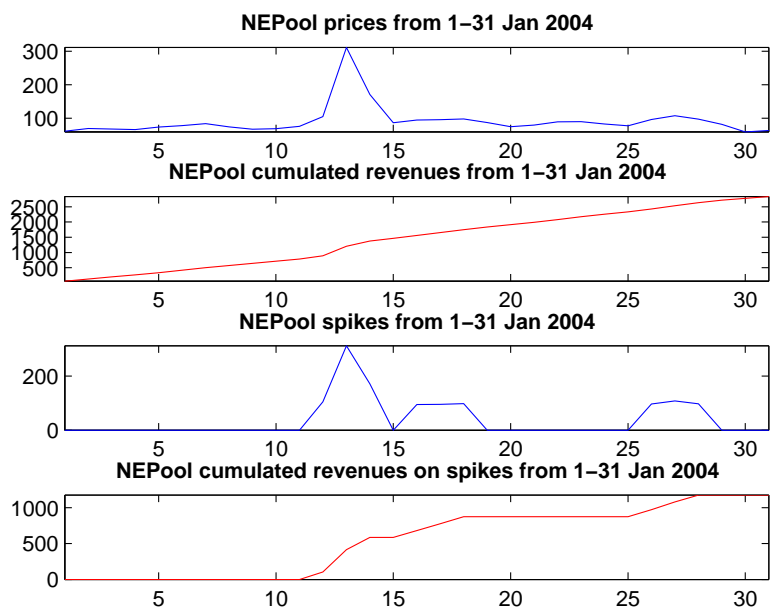


Figure 57: NEPool revenues from 1-31 Jan 2004: total (second upper panel) and spike only (bottom panel).

This short analysis is just of a visualizing character with some basic assumptions made for calculations, but we can clearly see how important studying outliers in electricity spot

prices is. They make a very significant percent of total revenues and, if electric power producers and suppliers are able to predict them, they could also predict some level of that uncertain part of their income.

5 Conclusion and future work

In this study an extensive analysis of two electricity price time series (NEPool and Otahuhu) was performed. Both series contain daily data and cover exactly the same period, from 03 Jan 2001 to 28 Dec 2007. The data sets seem to reveal significantly different characters, which perhaps stems from their different geographical origins.

The analysis started from basic statistical investigation of both electricity prices and price logarithmic returns. Then it went through investigation of series normality and internal dependencies. Since the data series reveal both non-normality and heteroscedasticity, price smoothing was employed. However, the mentioned features still remain significant in the smoothed data. Moreover, data split by week days keep the characteristics as well.

The most crucial part of the study covered analysis of outliers in both price and return series. The idea was to separate those observations, whose magnitudes significantly differ from their neighborhood. There were different window sizes defined, within which the data was averaged. Then each observation of the series was compared with its surrounding windows' means and standard deviations. The outliers were defined with respect to the original as well as the smoothed data. It appears that spikes are of clustering character, but removing them from the series does not eliminate heteroscedasticity from prices and does not lead to obtaining normality in the data. Similar results were obtained for return outliers. Even though NEPool returns without spikes became much closer to normally distributed, statistical tests still rejected the normality hypotheses. Moreover, the Otahuhu return series with outliers removed still remained as heavy-tailed distributed.

The final step of outlier analysis moved on to verify whether they were predictable based on year seasons of occurrence or change in volatility before they appear. When it comes to seasons, NEPool and Otahuhu series reveal different characteristics, which most probably stems from different geographical locations of the regions. Regarding the analysis of outliers with respect to series volatility, it appears that in most cases spikes are not preceded by increase in price variance. However, there is always increase of volatility after the jump occurs.

We come to a conclusion that being able to forecast spikes might be of very high desire among electricity producers and suppliers. The results of this study show that outliers may not be possible to be predicted based only on the price series. It is very likely that some geographical and meteorological aspects need to be considered as well, especially for the time points when the outliers appeared. Moreover, analysis of demand-supply

relations should be investigated, since spikes appear when there is an issue of supply grid constraints.

It has been concluded by Escribano *et al.* [9] that electricity prices seem to act in two separate regimes: with and without congestions in the supply grid. Therefore, as Ptak *et al.* [19] proposed, a dual model may be an example solution for electricity price modeling. Within such, one part would be explaining regular behaviour of prices and returns when there are no delivery constraints within the supply grid. The second part could stand for expressing the price behaviour in case of congestions. The analysis presented in this thesis gives a decent basis for development of the dual model.

References

- [1] Dickey, D. A and Fuller, W. A. (1979). Distribution of the estimates for autoregressive time series with a unit root. *Journal of Empirical Finance* **1**: 83-106
- [2] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflations. *Econometrica* **50**: 987-1007
- [3] Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* **31**: 307-327
- [4] Phillips, P. C. B. (1987). Time series regression with a unit root. *Econometrica* **55**: 863-898
- [5] Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? *Journal of Econometrics* **54**: 159-178.
- [6] Box, G. E. P, Jenkins, G. M. and Reinsel, G. C. (1994). *Time Series Analysis: Forecasting and Control*, 3rd edition. Prentice Hall, Engewood Cliffs, NJ.
- [7] Bracewell, R. N. (2000). *The Fourier Transform and its Applications*, 3rd edition. McGraw-Hill Companies, Inc., United States
- [8] Brooks, C. (2002) *Introductory econometrics for finance*. Cambridge University Press, United Kingdom
- [9] Escribano, A., Peña, J. I. and Villaplana, P. (2002). Modeling electricity prices: international evidence. Working paper 02-27, Economics Series 08. Universidad Carlos III de Madrid.
- [10] Hinz, J. (2003). Modelling day-ahead electricity prices. *Applied Mathematical Finance* **10**: 149-161
- [11] Conejo, A. J., Contreras, J., Espínola, R. and Plazas, M. A. (2005). Forecasting electricity prices for a day-ahead pool-based electric energy market. *International Journal of Forecasting* **21**: 435- 462
- [12] Tsay, R. S. (2005). *Analysis of Financial Time Series*, 2nd edition. Wiley-Interscience, Hoboken, New Jersey, United States.
- [13] Hadsell, L. and Shawky, H. A. (2006). Electricity Price Volatility and the Marginal Cost of Congestion: An Empirical Study of Peak Hours on the NYISO Market, 2001-2004. *The Energy Journal* **27**: 2
- [14] Haldrup, N. and Nielsen, M. Ø. (2006). A regime switching long memory model for electricity prices. *Journal of Econometrics* **135**: 349-376
- [15] Huisman, R.,Huurman, C. and Mahieu, R. (2007). Hourly electricity prices in day-ahead markets. *Energy Economics* **29**: 240-248
- [16] Kanamura, T. and Ohashi, K. (2007). A structural model for electricity prices with spikes: Measurement of spike risk and optimal policies for hydropower plant operation. *Energy Economics* **29**: 1010-1032
- [17] Kanamura, T. and Ohashi, K. (2008). On transition probabilities of regime switching in electricity prices. *Energy Economics* **30**: 1158-1172

- [18] Karakatsani, N. V. and Bunn, D. W. (2008). Intra-day and regime-switching dynamics in electricity price formation. *Energy Economics* **30**: 1776–1797
- [19] Ptak, P., Jabłońska, M., Habimana, D. and Kauranne, T. (2008) Reliability of ARMA and GARCH models of electricity spot market prices. *Proceedings of European Symposium on Time Series Prediction 2008*. Porvoo, Finland
- [20] <http://www.nepool.com>
- [21] <http://www.gine.org>
- [22] <http://www.wikipedia.org>