NTNU

Norwegian University of Science and Technology

Prediction of large price changes in the energy market using extreme value statistics

Minxian Le

Master of Science in Physics and Mathematics
Submission date: July 2011
Supervisor: Arvid Næss, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

Preface

In the spring semester of 2011, I have been working with this master thesis of testing two methods, the average conditional exceedance rate (ACER) method and peaks over threshold (POT) method. This is mainly done on real life data from the energy market. During this work, I would very much like to thank my supervisor, professor Arvid Næss, for his support, encouragement and many useful advices and talks. Thank you Arvid. I would also like to express my appreciation to the people behind the code for the ACER method, and associate professor Sjur Westgaard for providing the energy data sets. Lastly, I would like to thank my boyfriend for his limitless support while working on this thesis.

Minxian Le, NTNU, Trondheim 7th of July, 2011.

Abstract

In this project we have first and foremost been comparing the performance of the ACER method with the POT method in the prediction of extreme values from the heavy tailed distributions; especially for data from the energy markets. The energy market is an exciting dynamic market where small singularities can make large differences in the price. Therefore it is very important and challenging to analyse and make predictions in this market. We have also analysed a dataset which is not from the energy market, to compare and see the main differences between the two markets. We have also taken in consideration of removing the return value for the dates of maturity to see whether this will have any influence on the results.

The main concept of the POT method is to find a threshold, u, and let the excesses be distributed by the Generalised Pareto Distribution. Whilst for the ACER method, we assume a specific shape of the tail, which in this project was of the kind Fréchet. We have done this analysis for five different data sets where two of them have been considered with and without their expiration dates. We have also filtrated the data sets with an AR-GARCH filter, and then used the POT and ACER on the residuals from the process. We have found out that both methods are not greatly influenced by the filtration, but we see the tendency of the POT method predicting a heavier tail than the ACER method. Further on, we can say that there are no significant large effects of removing the return values for the dates of maturity. Lastly, the data sets from the energy market prove themselves much more heavy tailed than for the data set from Norsk Hydro.

Contents

1 Introduction

Extreme Value Theory (EVT) is a branch within statistics that provides a framework in the study of behaviour in the tail regions of a distribution, being extremely small or large observations. EVT allows us to use the extremes to extrapolate the tails of a distribution to parts of it that have not yet been observed in the empirical observations. In this way, we can assess risk for highly unusual events, like for example 100-year winds, floods or extreme development in the financial markets. Since we are dealing with the extremes of events, the observations will be sparse and therefore we must value the observations we have. Classical EVT works often with block maxima or minima and fit them to the generalised extreme value family. This method does not fully make use of the observations and is a wasteful approach. The general idea of the peaks over threshold (POT) method is to work with the threshold excess over a certain threshold u. This method makes better use of the data as we work with the distribution of the threshold excesses rather than with the distribution of block maxima/minima. The average conditional exceedance rate (ACER) method produce a class of functions, the ACER functions, that is constructed in a way that they manage to capture the sub asymptotic behaviour of the extreme value distribution.

In this master thesis, we have worked on comparing the two methods ACER (average conditional exceedance rate) and POT (peaks over threshold). These two methods have been used to estimate return levels and tail quantiles for data sets mainly from the energy market. We have also filtrated the data sets with an AR-GARCH filter, and used the POT and ACER method to estimate in-sample and out-of-sample predictions on the residuals. There are both positive and negative aspects when using both methods. With the POT method the negative is the procedure for choosing a good threshold value, whereas a bad choice that can lead to ad hoc results and conclusions. With the ACER method, we have to assume a certain tail behaviour that are not necessary correctly assumed. Another weakness with the POT method is that it assumes asymptotic characteristics in the data sets, and there is no way to verify these assumptions. The positive aspect with the POT method is the easily understood theory behind the analysis, while with the ACER method we have the fact that this method makes a better use of the given data during the analysis. Although there are both positive and negative sides with the two methods, we have made an conclusion with regards to the data sets used in this thesis. We have also analysed one data set from Norsk Hydro to compare another kind of market to the ones from the energy markets.

In this master thesis, we have focused on the estimation of the positive tail. It does not necessary mean that the positive tail quantile is more important than the negative tail quantile. There are people interested in the extremely large increases as decreases in the energy market. The treatment of the negative tail is analogue to the ones of the positive tail, therefore we will concentrate on the positive tail in this thesis.

2 Theory

The basic idea of extreme value theory is to build a model which can understand the statistical behaviour of

$$M_n = \max\{X_1, ..., X_n\} \tag{1}$$

where $X_1, ..., X_n$ is a sequence of independent random variables having a common distribution function F. In theory, we can derive the distribution of M_n exactly for all n by using the assumption of independency:

$$\text{Prob}\{M_n \leq \eta\} = \text{Prob}\{X_1 \leq \eta, ..., X_n \leq \eta\} = \text{Prob}\{X_1 \leq \eta\} \cdots \text{Prob}\{X_n \leq \eta\} = \{F(\eta)\}^n \tag{2}$$

This is not useful when we do not know the distribution F. Also note that small errors in the estimation of $F(\eta)$ will lead to substantial errors in $\{F(\eta)\}^n$. So even though the possibility for estimating F is present, it is not always the best solution to do so.

2.1 The Generalised Extreme Value Distribution

The basic idea is to look at the behaviour of $\{F(\eta)\}^n$ when $n \to \infty$. The difficulty of the distribution of M_n degenerating to a point mass on the upper end-point of F when $\{F(\eta)\}^n \to 0$ as $n \to \infty$, can be solved by doing a linear renormalisation of the variable M_n

$$M_n^* = \frac{M_n - b_n}{a_n} \tag{3}$$

It can then be shown by *the extremal types theorem* [4] that if there exist sequences of constants $\{a_n > 0\}$ and $\{b_n\}$ such that

$$\text{P}\{M_n^* \leq \eta\} \to G(\eta) \text{ as } n \to \infty \tag{4}$$

for a non-degenerate distribution function G, then G belongs to one of the following families:

$$\text{I} : G(\eta) = \exp\Big\{ - \exp\Big[- \Big(\frac{\eta - b}{a}\Big)\Big]\Big\}, \quad -\infty < \eta < \infty; \tag{5}$$

$$\text{II} : G(\eta) = \begin{cases} 0, \ \eta \leq b, \\ \exp\Big\{ - (\frac{\eta - b}{a})^{-\alpha}\Big\}, \ \eta > b; \end{cases} \tag{6}$$

$$\text{III} : G(\eta) = \begin{cases} \exp\Big\{ - \Big[- \Big(\frac{\eta - b}{a}\Big)\Big]^{\alpha}\Big\}, \ \eta < b, \\ 1, \ \eta \geq b, \end{cases} \tag{7}$$

for parameters $a, b, \alpha > 0$. The distributions I, II and III are respectively the Gumbel, the Fréchet and the Weibull distribution. These three extreme value distributions can be written as the generalised extreme value (GEV) distribution

$$G(\eta) = \exp\Big\{ -\Big[1 + \xi\Big(\frac{\eta - \mu}{\sigma}\Big)\Big]^{-1/\xi}\Big\} \tag{8}$$

defined on $\{\eta : 1+\xi(\eta-\mu)/\sigma > 0\}$, where $-\infty < \mu < \infty, \sigma > 0$ and $-\infty < \xi < \infty$. Note that μ is the location parameter, σ is the scale parameter and ξ is the shape parameter. The cases $\xi > 0$ and $\xi < 0$ corresponds respectively to the Fréchet and Weibull. When $\xi = 0$ is interpreted as the limit of equation (8) as ξ approaches zero, and we have the Gumbel family with the distribution function

$$G(\eta) = \exp\Big[- \exp\Big\{ -\Big(\frac{\eta - \mu}{\sigma}\Big)\Big\}\Big], \quad -\infty < z < \infty \tag{9}$$

2.2 Kurtosis and skewness

Kurtosis is a way to measure how an empirical data set of observations differ from the normal Gaussian distribution. Another way to describe it is that it is a measure of the peakness of the probability distribution. The higher kurtosis, more of the variance is a result of infrequent extreme deviations. In this project we have worked with excess kurtosis. The difference between excess kurtosis and kurtosis is a shift in the scale of measurement. The Gaussian distribution has kurtosis 0 when dealing with excess kurtosis, while it has kurtosis 3 when dealing with kurtosis. Mathematically, the sample excess kurtosis is defined as

$$\text{Excess kurtosis} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2)^2} - 3 \tag{10}$$

where $x_1, ..., x_n$ are the observations.

Skewness is a measure of the asymmetry of the probability distribution. A negative skew-value indicates that the tail on the left side of the probability distribution density function is heavier than the right side. A positive skew-value indicates that the tail on the right side is heavier than the left side. The skew-value of zero means that the observations are relatively evenly distributed on both sides of the mean. Note that this does not necessarily indicate a symmetric distribution. Since we are considering extremely large positive changes, we will consider only the right side tail. The sample skewness is defined as

$$\text{Skew} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2)^{3/2}} \tag{11}$$

where $x_1, ..., x_n$ are the observations. Both the kurtosis and skewness will tell us how the empirical observations from the data sets behave in comparison to the Gaussian distribution.

2.3 Stationary processes

In the process of analysing the data, we must take in consideration if the data are stationary or not. Much of our analysis with the peaks over threshold (POT) method rest on the assumption that the observations are stationary. It is important that the data are stationary if we are going to obtain consistent estimators. One way to tell whether a data set is stationary is to plot the time series against time. If the graph crosses the mean of the sample many times, there are a good chance that the data is stationary [2]. But this definition is very vague. One can ask oneself what *many times* means. Therefore, there are another way to look for stationarity. We can look at the development of the standard deviation, σ, as the time progress. The sample standard deviation is defined as

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{12}$$

where we have the samples $x_1, ..., x_n$ and $\bar{x}$ is the mean of the samples. Then we can look at the plot and see if there are any noticeable trends, blocks or other systematic trends. We would like to have as less systematic trends as possible. We can also use methods like Augmented Dickey-Fuller or Phillips-Perror test to look at the stationarity of the data [1]. But we can not trust the results of the tests blindly.

2.4 Peaks over threshold (POT)

The general idea of the peaks over threshold (POT) method is that if a block maxima have approximately the distribution G from the expression

$$G(\eta) = \exp\Big\{ -\Big[1 + \xi\Big(\frac{\eta - \mu}{\sigma}\Big)\Big]^{-1/\xi}\Big\} \tag{13}$$

for some $\mu, \sigma > 0$ and ξ. Then all the threshold excess can be accounted to be distributed within the generalized Pareto distribution.

2.4.1 Generalized Pareto Family

Let $X_1, X_2, ...$ be a sequence of independent random variables with common distribution function F, and let

$$M_n = \max\{X_1, ..., X_n\}$$

Denote an arbitrary term in the sequence by X, and suppose that F satisfies (4), so that for large n,

$$\Pr\{M_n \leq \eta\} \approx G(\eta)$$

where $G(\eta)$ is equal to equation (13). Then, for large enough u, the distribution function of $(X - u)$, conditional on $X > u$, is approximately

$$H(y) = 1 - \Big(1 + \frac{\xi y}{\tilde{\sigma}}\Big)^{-1/\xi} \tag{14}$$

defined on $\{y : y > 0 \text{ and } (1 + \xi y/\tilde{\sigma}) > 0\}$, where $\tilde{\sigma} = \sigma + \xi(u - \mu)$.

In another words, this means that if a block maxima have approximating distribution G, then the points above the threshold given by u have a corresponding approximate distribution within the generalized Pareto family.

Thus the description of the stochastic behaviour of extreme events given by the conditional probability can be approximated by

$$\Pr\{X - u > y | X > u\} \approx H(y) = 1 - \Big(1 + \frac{\xi y}{\tilde{\sigma}}\Big)^{-1/\xi} \tag{15}$$

where u is the threshold and the modified scale parameter $\tilde{\sigma} = \sigma + \xi(u - \mu)$. The shape parameter ξ is equal to that of the corresponding GEV distribution and it is invariant to block size n. If $\xi < 0$ the distribution of threshold excesses has an upper limit of $u - \tilde{\sigma}/\xi$, if else the upper limit is unbounded.

2.4.2 Threshold selection

If we define the observations $x_1, ..., x_n$, and proposing a threshold u, we then can write the exceedances as $\{x_i : x_i > u\}$. We label these exceedances by $x_{(1)}, ..., x_{(k)}$, and define the threshold excesses even more specific as $y_j = x_{(j)} - u$ for $j = 1, ..., k$. Taken in consideration of the previous section, then the y's may be regarded as independent realisations of a random variable whose distribution can be approximated by a member of the generalised Pareto family. The intricate of the threshold choice is the inevitable give and take relationship between bias and variance. Too low threshold is likely to violate the asymptotic assumptions of the model, leading to bias. While too high threshold choice can give large variance because we have very few excesses with which the model can be estimated. Much of the data must then be discarded, and we get poor estimations of the parameters from the GPD. Thus, we must find the lowest threshold for which the threshold excesses fits the GPD.

There are several methods for dealing with this problem of choosing the threshold, but we are only going to look at two different methods. One is an exploratory technique carried out prior to model estimation, while the other is an assessment of the stability of parameter estimates based on the fitting of models across a range of different thresholds.

The first method is based on the mean of the GPD and lead to a mean residual life plot which can give a direction of what level of threshold choice we should choose. The first moment of the GPD can be expressed as

$$\mathrm{E}[X] = \frac{\tilde{\sigma}}{1-\xi} = \frac{\sigma + \xi(u-\mu)}{1-\xi}, \ \xi < 1 \tag{16}$$

which is linear in u. We obtain the mean residual life plot if we plot the mean empirical exceedance of the threshold against the threshold,

$$\left\{\left(u, \frac{1}{n_u}\sum_{i=1}^{n_u}(x_{(i)} - u)\right) : u < x_{\max}\right\} \tag{17}$$

where $x_{(1)}, ..., x_{(n_u)}$ is the ordered observations over threshold u. Confidence intervals can be added based on the approximation normality of sample means. The interpretation of the plot is the fact that it is approximated linear in the area where the GPD is a good fit. There can in fact be several different areas where we see linearity, so the choice of picking one can be difficult not knowing which one to choose. It is also hard to distinguish the fluctuations that are within the linear area from bigger and irrelevant fluctuations outside the area. The confidence intervals will grow as the larger the threshold becomes.

The threshold selection based on the second method is given by looking at the estimated shape ξ and scale μ parameter. The shape parameter should be the same when the GPD is applied. If we plot the estimated shape parameter against threshold, we can look for areas where it is constant. As for the scale parameter $\tilde{\sigma} = \sigma + \xi(u - \mu)$, we must modify it: $\sigma^* = \tilde{\sigma} - \xi u$. The modified scale parameter is constant with respect to u. This, with the first method, can help us to decide the threshold choice. Here we would want to choose the smallest u for the area where the parameter estimates are constant and where the mean residual life plot is linear. This gives us the opportunity to obtain as much observations as possible taking account of the threshold.

2.4.3 Dependence and correlation

As mentioned before, the POT method is derived under the assumption of iid observations. One way to check for *linear* dependence is done with the use of the autocorrelation function (ACF). For the data sets we have in this project, the ACF show that there are minimal linear dependence in the data. This also indicates that there are low dependence in general. One way dealing with dependence sequences is declustering. That is identifying clusters of dependent observations by an empirical rule and only keeping each cluster maxima. This have be done with the data sets, and the conclusion is that there are no significant difference in the results.

In this project, the package named POT in R has been used to perform the peaks over threshold analysis [8].

2.4.4 Model checking

To evaluate the quality of a fitted generalised Pareto model, we can use probability plots, quantile plots, return level plots and density plots. Assuming a threshold u and their respectively excesses $y_{(1)} \leq ... \leq y_{(k)}$, we get the probability plot by

$$\{(i/(k+1)), \hat{H}(y_{(i)}); i = 1, ..., k\} \tag{18}$$

where the estimated model

$$\hat{H}(y) = 1 - (1 + \frac{\hat{\xi} y}{\hat{\sigma}})^{-1/\hat{\xi}} \tag{19}$$

where $\hat{\sigma}$ and $\hat{\xi} \neq 0$ are the estimated scale and shape parameters. While the quantile plot is constructed by

$$\{\hat{H}^{-1}(i/(k+1)), y_{(i)}; i = 1, ..., k\} \tag{20}$$

where

$$\hat{H}^{-1}(y) = u + \frac{\hat{\sigma}}{\hat{\xi}}[y^{-\hat{\xi}} - 1] \tag{21}$$

If the generalised Pareto model is a reasonable fit for modeling excesses of u, then the probability plot and the quantile plots should show points that are approximately linear.

The return level plot consists of points $\{(m, \hat{x}_m)\}$ for large m values. $\hat{x}_m$ is the estimated m-observation return level and is defined as

$$\hat{x}_m = u + \frac{\hat{\sigma}}{\hat{\xi}}[(m\hat{\zeta}_u)^{\hat{\xi}} - 1] \tag{22}$$

where $\hat{\zeta}_u \approx \Pr\{X > u\}$. For the case of ξ being equal to zero, we have the following expression for the estimated m-observation return level

$$\hat{x}_m = u + \sigma \log(m\hat{\zeta}_u) \tag{23}$$

In the world of financial modelling, we often work with the term *value-at-risk* (VaR). This is defined as extreme quantiles of the daily returns, and the generalised Pareto threshold model provides a direct method for the estimation of this parameter. Therefore the return level plot can be seen as a graph of value-at-risk against risk on a convenient scale.

At last, we have the density function which can be compared to a histogram of the threshold exceedances. Note that all of this above is based on the assumption that $\hat{\xi} \neq 0$. The probability plot is the only one that has been used in the model checking because the rest of the plots can often be regarded as ambiguous.

2.5 Average conditional exceedance rates (ACER)

Until now, we have assumed independence while considering the problems with the POT method. ACER is a method developed under the assumption that the data can be dependent and non-stationary. Unlike the previous definition (2), we must express the probability by

$$\begin{aligned}\text{Prob}(M_n \leq \eta) &= \text{Prob}(X_1 \leq \eta, ..., X_n \leq \eta) \\ &= \text{Prob}(X_n \leq \eta | X_{n-1} \leq \eta, .., X_1 \leq \eta)\text{Prob}(X_1 \leq \eta, .., X_{n-1} \leq \eta) \\ &= \prod_{j=2}^{n} \text{Prob}(X_j \leq \eta | X_1 \leq \eta, ..., X_{j-1} \leq \eta)\text{Prob}(X_1 \leq \eta) \qquad (24)\end{aligned}$$

Using a k-step memory, we can make the following assumptions for k=1:

$$\text{Prob}\{X_j \leq \eta | X_1 \leq \eta, ..., X_{j-1} \leq \eta\} \approx \text{Prob}\{X_j \leq \eta | X_{j-1} \leq \eta\} \qquad (25)$$

and

$$\text{Prob}\{X_j \leq \eta | X_1 \leq \eta, ..., X_{j-1} \leq \eta\} \approx \text{Prob}\{X_j \leq \eta | X_{j-2} \leq \eta, X_{j-1} \leq \eta\} \qquad (26)$$

for $2 \leq j \leq n$ on (25) and for $3 \leq j \leq n$ on (26), and so forth. Then we can generally express the equation (24) as

$$\begin{aligned}\text{Prob}(M_n \leq \eta) = P(\eta) &\approx \prod_{j=k}^{n} \text{Prob}(X_j \leq \eta | X_{j-k+1} \leq \eta, ..., X_{j-1} \leq \eta) \\ &\quad \cdot \text{Prob}(X_{k-1} \leq \eta | X_1 \leq \eta, ..., X_{k-2} \leq \eta) \\ &\quad \cdot \text{Prob}(X_2 \leq \eta | X_1 \leq \eta)\text{Prob}(X_1 \leq \eta) \\ &= \prod_{j=k}^{n} (1 - \alpha_{kj}(\eta))(1 - \alpha_{k-1,k-1}(\eta)) \cdots (1 - \alpha_{11}(\eta)) \quad (27)\end{aligned}$$

where

$$\alpha_{kj}(\eta) = \text{Prob}(X_j > \eta | X_{j-k} \leq \eta, ..., X_{j-k+1} \leq \eta), \text{ for } 2 \leq k \leq j \qquad (28)$$

which denotes the exceedance probability conditioned on the $k-1$ previous non-exceedances. Note that in the case of which $k = 1$, we have $\alpha_{1j}(\eta) = \text{Prob}(X_j > \eta)$. Using the relations $(1 + x) \approx e^x$ if $|x| << 1$ we can derive the equation (27) to be

$$P(\eta) \approx P_2(\eta) = \exp(-\sum_{j=2}^{n} \alpha_{2j}(\eta) - \alpha_{11}(\eta)) \qquad (29)$$

Preceding with the conditioning on two and three the previous observations, we get

$$P(\eta) \approx P_3(\eta) = \exp(-\sum_{j=3}^{n} \alpha_{3j}(\eta) - \alpha_{22}(\eta) - \alpha_{11}(\eta)) \quad (30)$$

$$P(\eta) \approx P_4(\eta) = \exp(-\sum_{j=4}^{n} \alpha_{4j}(\eta) - \alpha_{33}(\eta) - \alpha_{22}(\eta) - \alpha_{11}(\eta)) \quad (31)$$

Since we are considering the most practical applications, when $n >> 1$ and $k \geq 2$, we get the simplified expression

$$P_k(\eta) \approx \exp(-\sum_{j=k}^{n} \alpha_{kj}(\eta)) \quad (32)$$

As mentioned previously, we can interpret $\alpha_{kj}(\eta)$ to be the probability of the j'th observation exceeding the level of threshold η conditioned on the $k-1$ previous observations being under the threshold. Therefore we interpret $\sum_{j=k}^{n} \alpha_{kj}(\eta)$ as the expected number of independent exceedances with level of threshold η and conditioned on the $k-1$ previous non-exceeding observations.

The ACER function is defined as

$$\epsilon_k(\eta) = \frac{1}{n-k+1} \sum_{j=k}^{n} \alpha_{kj}(\eta), \; k = 1, 2, ... \quad (33)$$

In order to estimate the ACER functions, we can start by defining

$$A_{kj}(\eta) = I\{X_j > \eta, X_{j-1} \leq \eta, ..., X_{j-k+1} \leq \eta\}, \; j = k, ..., n, \; k = 2, 3, ... \quad (34)$$

$$B_{kj}(\eta) = I\{X_{j-1} \leq \eta, ..., X_{j-k+1} \leq \eta\}, \; j = k, ..., n, \; k = 2, 3, ... \quad (35)$$

where $I\{\cdot\}$ is the indicator function. We can then use these definitions to express the ACER function as

$$\epsilon_k(\eta) = \lim_{n\to\infty} \frac{\sum_{j=k}^{n} A_{kj}(\eta)}{\sum_{j=k}^{n} B_{jk}(\eta)} \quad (36)$$

when we know for fact that $\alpha_{kj}(\eta) = \mathrm{E}\{A_{kj}(\eta)\}/\mathrm{E}\{B_{kj}(\eta)\}$ for $j = k, ..., n, k = 2, 3, ...$

The sample estimation of $\epsilon_k(\eta)$ can then be estimated by

$$\hat{\epsilon_k}(\eta) = \frac{1}{R} \sum_{r=1}^{R} \hat{\epsilon_k}^{(r)}(\eta) \quad (37)$$

where R is the number of realisations or samples and

$$\hat{\epsilon_k}^{(r)}(\eta) = \frac{\sum_{j=k}^{n} A_{kj}^{(r)}(\eta)}{\sum_{j=k}^{n} B_{jk}^{(r)}(\eta)} \tag{38}$$

for the r'th realisation. Then we use that $\lim_{\eta\to\infty}\sum_{j=k}^{n} B_{jk}(\eta) = n-k+1$, and get

$$\hat{\epsilon_k}^{(r)}(\eta) = \frac{\sum_{j=k}^{n} A_{kj}^{(r)}(\eta)}{n-k+1} \tag{39}$$

To construct the confidence intervals of the ACER function, we need to define the expression for the estimated standard deviation

$$\hat{s_k}(\eta)^2 = \frac{1}{R-1}\sum_{r=1}^{R}(\hat{\epsilon_k}^{(r)} - \hat{\epsilon_k})^2 \tag{40}$$

The equation (40) leads to a fairly good approximation of the 95% confidence interval where the limits are set by

$$\mathrm{CI}^{\pm} = [\hat{\epsilon_k}(\eta) - 1.96\hat{s}_k(\eta)/\sqrt{R}\ ,\ \hat{\epsilon_k}(\eta) + 1.96\hat{s}_k(\eta)/\sqrt{R}] \tag{41}$$

Since we are working with data with a heavy-tailed distribution, we know that the underlying asymptotic extreme value distribution here is of type Fréchet. This gives the tail behaviour denoted by the equation (42) [7],

$$\epsilon_k(\eta) \approx q_k(\eta)[1 + \xi_k(a_k(\eta - b_k)^{c_k})]^{-1/\xi_k},\ \eta \geq \eta_1 \tag{42}$$

where the function $q_k(\eta)$ is a weakly varying function compared to $[1 + \xi_k(a_k(\eta - b_k)^{c_k})]^{-1/\xi_k}$ so it can be considered as a constant q (especially in the tail region), and $a_k > 0$, b_k, $c_k > 0$ and $\xi_k > 0$ are all constants dependent on k. Note that when $c_k = 1$ and $q = 1$ it corresponds to the asymptotic limit, and then we have the Generalised Extreme Value distribution. The parameters c_k and q are also the sub-asymptotic parameters.

We can write equation (42) in an alternative form by introducing a new function $d_k(\eta)$:

$$\epsilon_k(\eta) \approx [1 + \xi_k(a_k(\eta - b_k)^{c_k} + d_k(\eta))]^{-1/\xi_k},\ \eta \geq \eta_1 \tag{43}$$

where the function $d_k(\eta)$ is weakly varying compared to $a_k(\eta - b_k)^{c_k}$. This written form leads to easier estimation procedures and therefore can be preferred. Expression (42) can be simplified by setting $\gamma = 1/\xi$, $\tilde{a} = a\xi$ and suppressing the k's, and we get the following

$$\epsilon(\eta) \approx q[1 + \tilde{a}(\eta - b)^c]^{-\gamma} \tag{44}$$

Taking the logarithm on both sides of equation (44) and moving the right side of the expression to the left we get the basis of the mean square error which can be minimised,

$$
\begin{aligned}
F(\tilde{a}, b, c, q, \gamma) &= \sum_{j=1}^{n} w_j |\log\hat{\epsilon}(\eta_j) - \log q + \gamma \log[1 + \tilde{a}(\eta_j - b)^c]|^2 && (45) \\
&= \sum_{j=1}^{n} w_j (y_i - \log q + \gamma x_i)^2 && (46)
\end{aligned}
$$

where $y_j = \log\hat{\epsilon}(\eta_j)$, $x_j = \log(1 + \tilde{a}(\eta_j - b)^c)$, and $w_j = (\log\text{CI}^+(\eta_j) - \log\text{CI}^-(\eta_j))^{-2}$ are weighted factors that emphasise more on the reliable data points. $\text{CI}^{\pm}$ denotes the upper and lower limit of the 95% confidence interval. It is possible to put even more emphasis or less emphasis on the different data points by replacing the exponent -2 with another number.

It is not always a trivial matter to minimise a function with five parameters. But in this case, the optimisation will be much easier if we use

$$
\gamma^*(\tilde{a}, b, c) = -\frac{\sum_{j=1}^{n} w_j (x_j - \bar{x})(y_j - \bar{y})}{\sum_{j=1}^{n} w_j (x_j - \bar{x})^2} \tag{47}
$$

and

$$
\log q^*(\tilde{a}, b, c) = \bar{y} + \gamma^*(\tilde{a}, b, c)\bar{x} \tag{48}
$$

as estimators for γ and $\log q$. Now we can minimise the function in equation (46) with respect to three parameters rather than five. We can also use optimisation methods like for example Levenberg-Marquardt least square optimisation method to estimate the five parameters $\tilde{a}, b, c, q, \gamma$.

2.6 AR-GARCH model

In general, an auto regressive conditional heteroskedasticity (ARCH) model is used to characterise an observed time series. They will model a time series much better in consideration of great variance in the volatility. ARCH models assume the variance of the current error term to be functions of the error terms of the previous period. If an auto regressive moving average model (ARMA) is assumed for the error variance, the model is a generalised autoregressice conditional heteroskedasticity (GARCH) model.

Therefore in this thesis, we use a AR-GARCH model to prefilter the time series as done by McNeil and Frey [6]. The use of an AR term and a GARCH term comes from the fact that we would like to catch the strong patterns in the observations in the sense of seasonality and significant volatility clustering in the electricity data. An AR(k)-GARCH(p,q) is an autoregressive model of order k with GARCH noise of order p,q can be written in general as

$$r_t = a_0 + \sum_{i=1}^{k} a_i r_{t-i} + \epsilon_t \tag{49}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2 \tag{50}$$

where $\epsilon_t = \sigma_t \eta_t$ and $\eta_t \sim IID(0,1)$. Further on, η_t is usually Gaussian- or Student's t-distributed, scaled to variance 1 and with υ degree of freedom. We have used a GARCH(1,1) model to characterise the conditional volatility as a function of previous volatilities and return. While we have used an AR(5) part to modify the weekly seasonality. We would like the AR-GARCH model in this thesis to look as follows:

$$\begin{aligned} r_t &= a_0 + \sum_{i=1}^{5} a_i r_{t-i} + \epsilon_t \\ \sigma_t^2 &= \phi_0 + \phi_1 \epsilon_{t-1}^2 + \phi_2 \sigma_{t-1}^2 \end{aligned} \tag{51}$$

where σ_t^2 is the conditional variance of ϵ_t. We would like to include the AR(5) term because we have 5 observations in a week, representing the weekly seasonality. Although we would like our model to look like the equation (51), we have situations where some of the desired parameters are not significant. This makes changes to our desired model depending on the data set analysed. The parameters not mentioned in our model can be significant in some of the cases, but their role in the model will be unmentionable small in the cases where they are not significant.

We can also estimate the conditional tail quantiles, $\alpha_{t,p}$, by this general expression:

$$\alpha_{t,p} = a_0 + \sum_{i=1}^{k} a_i r_{t-i} + \sigma_t \alpha_p \tag{52}$$

where $a_0+\sum_{i=1}^{k} a_i r_{t-i}$ and σ_t is the conditional mean and volatility from the AR-GARCH model, and α_p is the unconditional tail quantiles.

2.7 Tail-quantile estimation

The tail-quantile estimation is a way to evaluate a models out-of-sample performance. This has been done by first dividing the observations in two equal parts; one estimation period and one out-of-sample performance comparison. We roll one observation forward for the estimation period till we reach the end of the dataset. In other words, we re-estimate the models each day with the new observations. The standard residuals of the estimation periods are then saved and are the basis for our out-of-sample predictions. We make tail-quantile estimation for both the POT method and the ACER method to compare the two methods. In this way, we have the out-of-sample performance comparison part to compare the estimation results with. The code for the tail-quantile estimation is seen in appendix A.

3 Analysis of data

When we are doing the POT analysis we work with the log return of the data. To put it simple, that is the logarithm of data today divided by yesterday. The main advantage by doing this is that the logarithmic return (also known as the continuously compounded return) is symmetric, while the arithmetic return is not: positive and negative percent arithmetic returns are not equal. The analysis have also been done with a net simple return instead of a logarithmic return. The net simple return is expressed as

$$\frac{X_t - X_{t-1}}{X_{t-1}} \tag{53}$$

We did not find any large and significant differences between doing the analysis on the EL ICE data set with the net simple returns and the log returns, and therefore the analysis described in the thesis have been of the logarithmic kind.

The tail marker for the ACER method is set after inspecting the empirical ACER functions plotted against scaled exceedances. We have also considered the fact that the predicted return value can or can not be sensitive to the choice of the tail marker. Therefore we have checked the degree of sensitivity for the data sets by adjusting the tail marker and looking for any considerable alteration in the predicted parameters. This we do to verify the robustness of the predictions.

Also note that for both methods, we are checking the extrapolated ACER function and POT fitted GPD with their respectively confidence intervals down to a 10^{-6} level, the level of interest. We do this because we want to highlight and stress the consequences by predicting at a lower, but realistic, level.

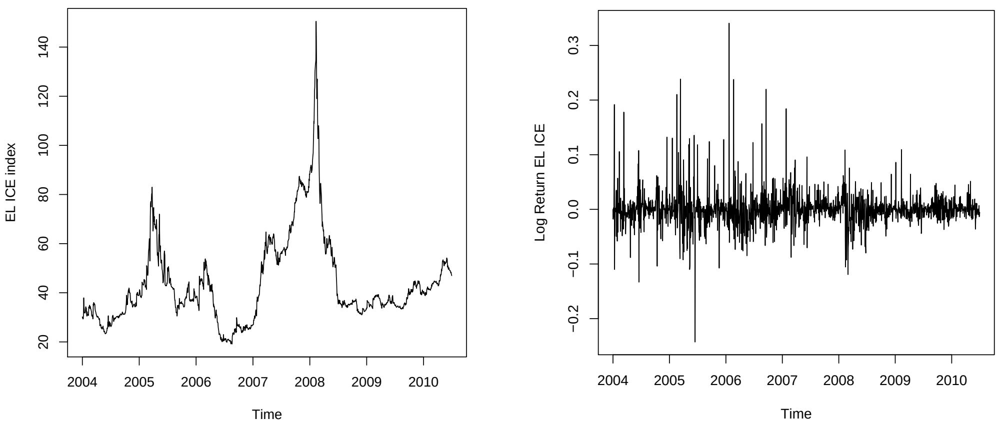
When working with the energy market and pricing futures, we have to remove the return value for the dates of maturity. This return does not denote the difference in price within a certain contract, but instead representing the price difference between two different contracts. For example, the date of maturity for the EL ICE market is the last Thursday in each month. Unfortunately, because of the lack of feedback from the large energy markets, we have only got hold of the dates of expiration for two of the data sets; the EL ICE and the Coal ICE data set. Hopefully, this will give us some kind of impression about how large an influence the removal of the expiration dates will have on the results.

For the mean residual life plot and the plot of the estimated modified scale and shape parameter, we have made sure that the x-axis never spans over an interval where there are less than approximately 1% of the total amount of observations. Because after that point, it is not longer necessary to investigate the stability and progression of the threshold value.

In the theory section, we have described the ACER parameters in equation (46). We see that some of the parameter is defined as greater than zero, but because of how the ACER method have been programmed [5] these parameters are going to be negative in the mentionings for the different data sets in the following sections.

3.1 EL ICE with expiration dates

EL ICE describes the development on the electricity market in United Kingdom. We look at monthly contracts which is liquid and available. The electricity futures contract is a deliverable contract where the buyer is obliged to make or take delivery of electricity to/from National Grid Transco in accordance with ICE Futures Europe regulations. Further on, we look at the base load, which is the period the electricity is delivered from 12PM to 12PM the next day. We have observations 5 days a week from 14.09.2004 to 31.01.2011 except for holidays (1627 observations; over 6 years worth of observations).

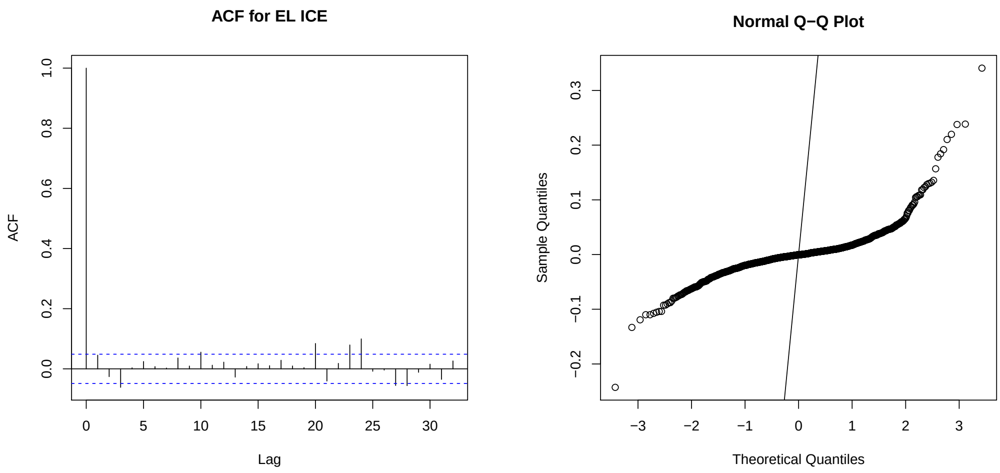


(a) The index of EL ICE from period 14.09.2004-31.01.2011. (b) The log return of the EL ICE observations.

Figure 1: Index plot and log return plot of the observations for EL ICE

In figure 1(a) and 1(b) we see the index and the log returns of the data set EL ICE. There are presumably some incident causing the dramatic outcome after passing day number 1000 (approximately after 4 years). This can be seen clearly both in the index plot and plot for the log returns. The plot for the autocorrelation function in figure 2(a) shows nearly close to zero autocorrelation in the data set, except for marginally larger values between lag 20 and lag 25. We can also see how different our data set varies from the normal distribution in figure 2(b). This can be underlined by the excess kurtosis value for the log returns of 18.45. It is also worth mentioning that the skewness is 1.85, giving us a slightly heavier right tail than the left tail.

To help us deciding a suitable threshold value u with the POT method, we look at



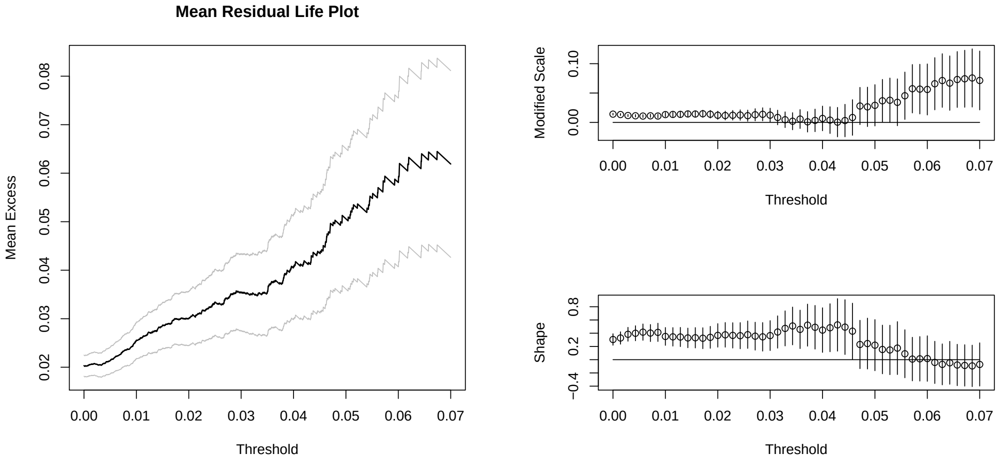
(a) The autocorrelation function.

(b) QQ-plot with the normal distribution represented as the line.

Figure 2: The autocorrelation function and the QQ-plot.

the mean residual plot in figure 3(a) and the estimated modified scale and shape parameter against different choices of threshold values in figure 3(b). We see that for the mean residual plot shows linearity from 0 to 0.03, whilst figure 3(b) shows a closely constant modified shape when the threshold is between 0 and 0.03, and the same applies for the shape parameter. We see that there is no reason for not choosing the threshold 0.03. This gives a modified scale $\sigma = 0.02332$ and a shape $\xi = 0.36328$. We have 145 observations above this threshold and this corresponds to 8.91% of the total amount of observations. To check the quality of this model, we plot the probability plot in figure 4(a) and we see that linearity is fairly present. This is just an indication of the model's quality, we can not rely fully on these plots to tell us whether a model is good or not. We can look at the stability in figure 4(b) where the standard deviation has been plotted against time (months). It does not seem very stationary by looking at the plot. The standard deviation has a high degree of variation and many peaks.

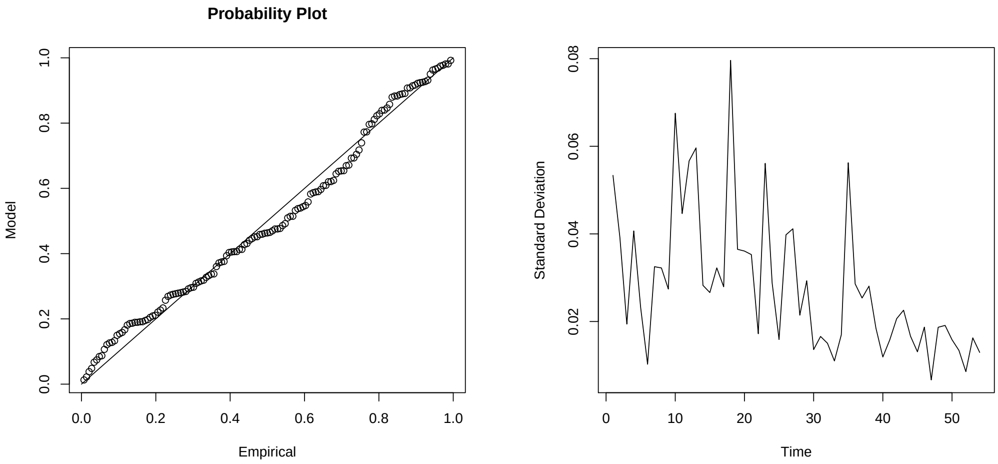
Continuing with the ACER analysis, we see from figure 5(a), we have the first four empirical ACER functions $k = 1, 2, 3, 4$. They will probably converge, so we usually do not need more ACER functions than that. From this plot we also see that the first ACER function is sufficient enough to use when estimating the parameters in equation (46). The lower ACER function we use, the more of the data is utilised and the estimates will have better accuracy. The parameters of the ACER function are $\tilde{a} = -5.0008$, $b = 0$, $c = 0.7040$, $\gamma = -5.8701$ and $q = 0.6978$. The return level is 2.26715 with the 95% confidence intervals CI = [1.92249,2.64841] for the 10^{-6} level of interest. We have defined $\gamma = 1/\xi$ in an earlier section, but due to how the ACER code is developed we have that $\gamma = -1/\xi$ in this context.



(a) The mean residual plot.

(b) The estimated modified shape and scale parameter against the threshold.

Figure 3: The plots to help on the choice of threshold.

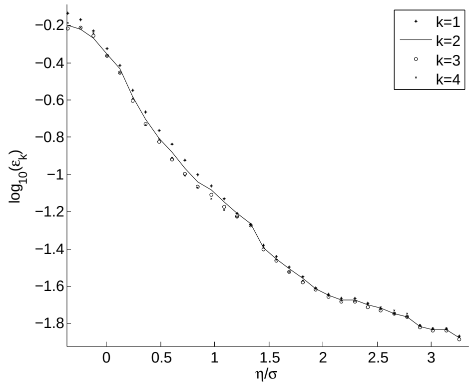


(a) The probability plot.

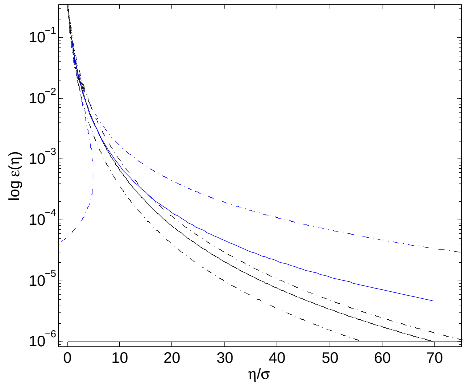
(b) The development of the standard deviation with respect to time (months).

Figure 4: Probability plot for threshold excess model fitted to EL ELICE data and checking for stationarity.

Now we filtrate this data set with an AR-GARCH filter. The one used here is an AR(5) and GARCH(1,1) model with the normal distribution as the conditioning distribution. We would like to prefer that all of the parameters used in the filter are significant down to a 0.05 level. The conditional distribution is chosen by looking at the significance of the parameters too, and are not necessary the normal distribution in every case. We see



(a) The empirical ACER functions plotted against scaled exceedances.



(b) The extrapolated ACER (black) function and POT (blue) fitted GPD with their confidence intervals respectively.

Figure 5: The EL ICE data analysis by the POT and ACER methods.

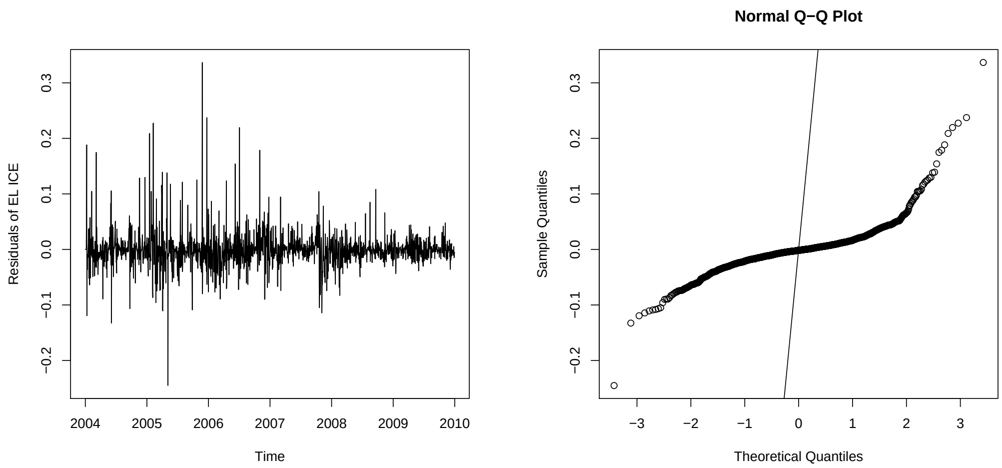
AR-GARCH parameters	Estimates (std.error)
a_0	1.056e-03 (6.223e-04).
a_1	4.456e-02 (3.650e-02)..
a_2	3.365e-03 (3.263e-02)..
a_3	-6.932e-02 (2.967e-02)*
a_4	9.748e-03 (3.148e-02)..
a_5	7.357e-02 (2.990e-02)*
ϕ_0	2.710e-07 (6.572e-07)..
ϕ_1	1.750e-02 (1.032e-02).
ϕ_2	9.824e-01 (9.995e-03)***

Table 1: Estimates of the AR-GARCH parameters for the full EL ICE data set with the normal distribution as the conditional distribution with the respectively standard errors in parenthesis. Signif. codes: ***=0.001, **=0.01, *=0.05, .=0.1 and ..=1

from table 1 that the parameters who satisfy the 0.05 level significant are a_3, a_5 and ϕ_2. The residuals of the EL ICE observations after filtration is plotted in figure 6(a). We see that this figure resembles the log return plot in figure 1(b) with larger variations in the first half of the observation period. From the QQ-plot in figure 6(b), the residuals seem to be as different from the normal distribution as the log returns in the previous case. It can be mentioned that the excess kurtosis is 18.04 and the skewness is 1.79. Both values are very similar to the excess kurtosis and skewness of the log returns. The autocorrelation function seen in figure 3.1 shows as little correlation as for the log returns of the EL ICE data set mentioned previously.

The threshold choice is done as above, and we see from the mean residual plot in figure 8(a) that there is a linear area from 0 to 0.04. From the figure 8(b) we can see that the estimated modified scale and shape parameters are constant in the region around 0-0.03. There are areas after the threshold value 0.04 that are linear and constant in these two plots, but it is not likely that the threshold value is going to be set as high as that taken in consideration of the fact that we would like to have as many observations above the threshold as possible. After careful study, the choice fell on the threshold value, $u = 0.02$. This decision leads to 224 observations above the threshold and that represents 13.77% of the total amount of the observations. Then the value of the shape and scale parameter are 0.01851(0.002114) and 0.40269(0.097156) with their respectively standard errors in parenthesis.

To check this following model, we use the probability plot that was mentioned before in a previous section. We see that there are some deviation from the linear line in the probability plot in figure 9(a), but these deviations are not serious enough for us to reconsider our fitted model. For the stationarity check in figure 9(b) we see that there is some kind of consistency in the pattern of the standard deviation with exception of some few tops which are higher than the others.



(a) The log return of the residuals of the EL ICE observations.

(b) QQ-plot with the normal distribution represented as the line.

Figure 6: The log return plot of the residuals and the QQ-plot for the EL ICE data set.

As done before, we now continue with the ACER analysis. From figure 10(a) we see the first four empirical ACER functions $k = 1, 2, 3, 4$. It is clearly that they converge, so we use the first ACER function to estimate the parameters in the ACER analysis. From figure 10(b) we find the extrapolated ACER function in black and the POT fitted GPD in blue with their confidence intervals respectively. We see in figure 5(b) that there are larger difference between the extrapolated ACER function and the POT fitted GPD than

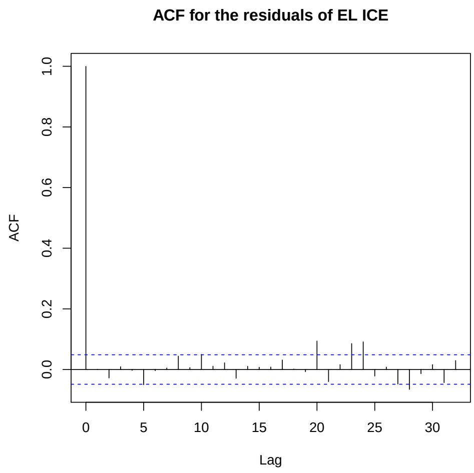
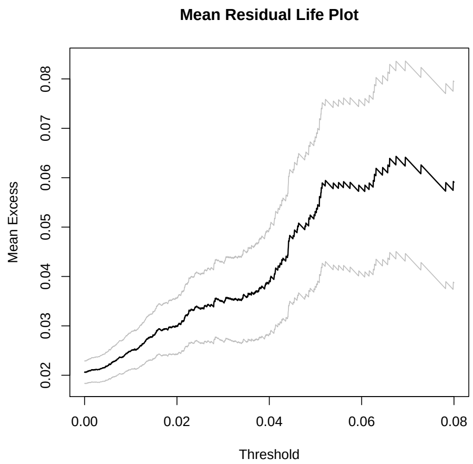
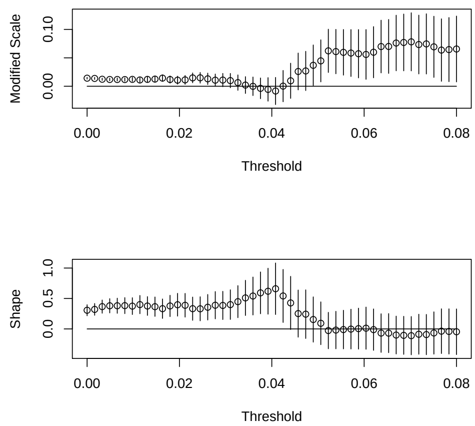


Figure 7: The autocorrelation function.



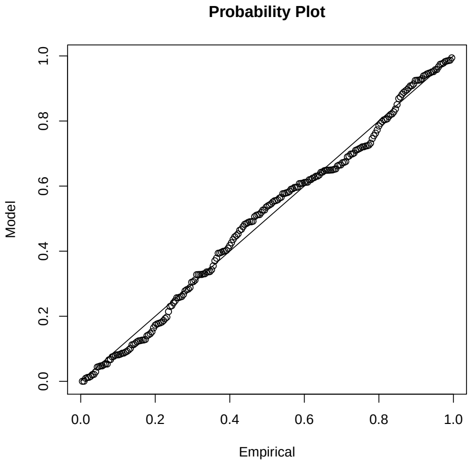
(a) The mean residual plot.



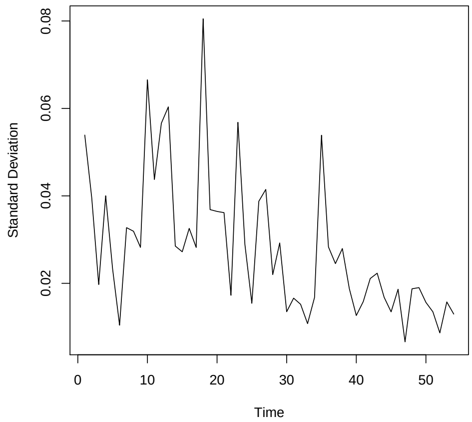
(b) The estimated modified shape and scale parameter against the threshold.

Figure 8: The plots to help on the choice of threshold.

the previous case with the log returns. We also notice that the confidence intervals of the extrapolated ACER function are larger than for the previous case. The parameters of the ACER functions are $\tilde{a} = -5.0101$, $b = 0$, $c = 0.6919$, $\gamma = -5.8372$ and $q = 0.7099$. The return level is 2.35922 with the 95% confidence intervals CI = [0.976009, 5.08484] for the 10^{-6} level of interest.

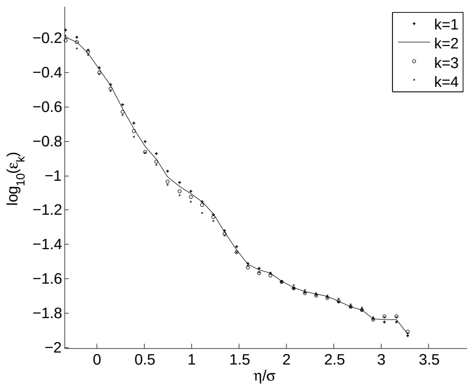


(a) The probability plot.

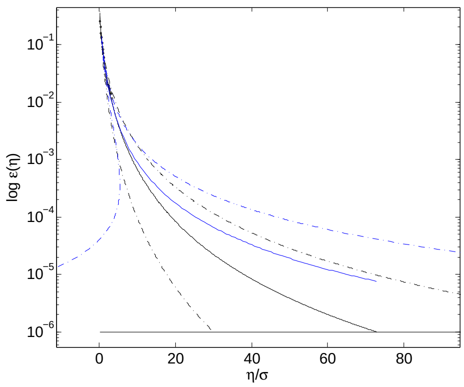


(b) The development of the standard deviation with respect to time (months).

Figure 9: Probability plot for threshold excess model fitted to the residuals of the EL ICE data and checking for stationarity.



(a) The empirical ACER functions plotted against scaled exceedances.

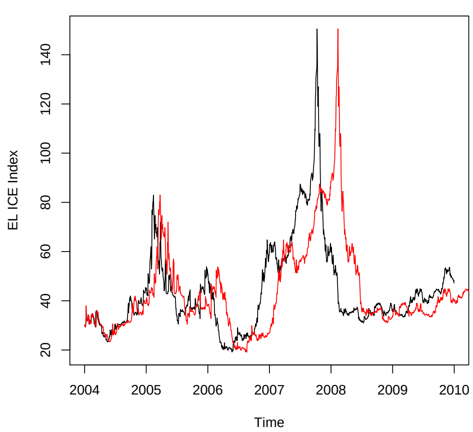


(b) The extrapolated ACER (black) function and POT (blue) fitted GPD with their confidence intervals respectively.

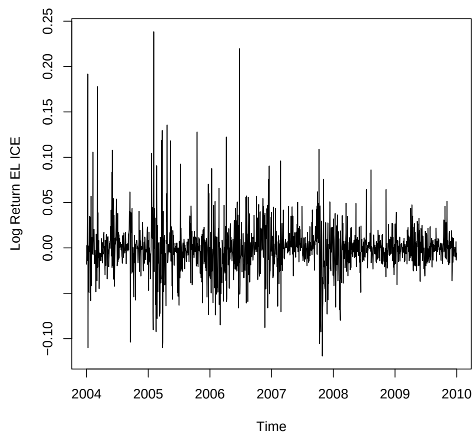
Figure 10: The residuals of the EL ICE data analysis by the POT and ACER methods.

3.2 EL ICE without expiration dates

The full dataset EL ICE without the returns from the last Thursday in every month consists of 1520 observations. This slightly smaller dataset is analysed as before, and we see in figure 11(a) the index plotted against time. Both the full dataset and the minimised one is plotted there as red and black respectively. We can see that there is a great increase around the end of year 2007 and the start of year 2008 where the highest value of the index is actually over three times the average mean of the whole dataset. The log return of both sets are also plotted in 11(b).



(a) The index of EL ICE from period 14.09.2004-28.01.2011. Red:the full dataset. Black:the dataset without the last Thursday in every month.

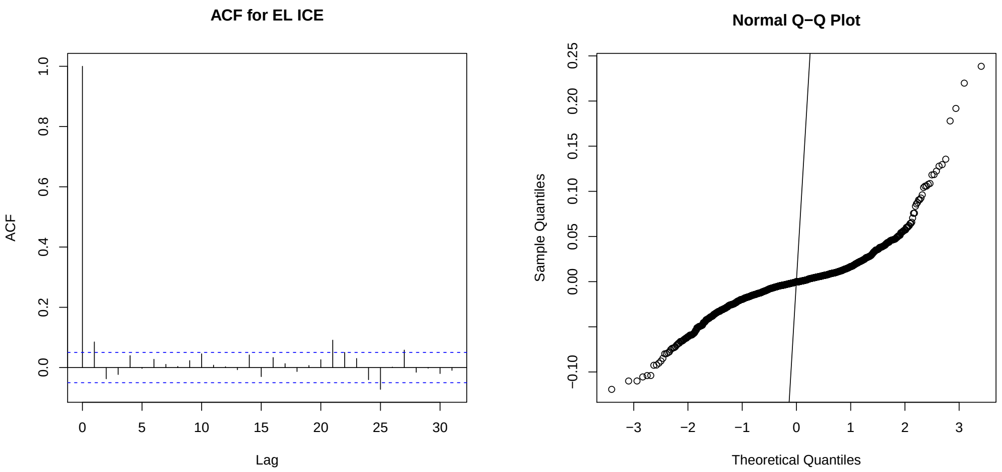


(b) The log return of the EL ICE observations.

Figure 11: Index plot and log return plot of the observations for the minimised EL ICE dataset.

The plot for the autocorrelation function in figure 12(a) shows nearly close to zero autocorrelation in the data set, so there are none concern about autocorrelation being present. We can also see how different our data set varies from the normal distribution in figure 12(b). It can be mentioned that the excess kurtosis value of 10.87, and that the skewness is 1.27 giving us a heavier right tail than left tail. The excess kurtosis is somewhat lower here than for the full data set, 18.45 in comparison to 10.87, but the skewness has not changed a lot. This previous mentioned change in the excess kurtosis can be because the data set now has less observations than the full data set; 107 less observations to be exact (6.6 %).

To help us deciding a suitable threshold value u with the POT method, we look at the mean residual plot in figure 13(a) and the estimated modified scale and shape parameter against different choices of threshold values in figure 13(b). We see that for the



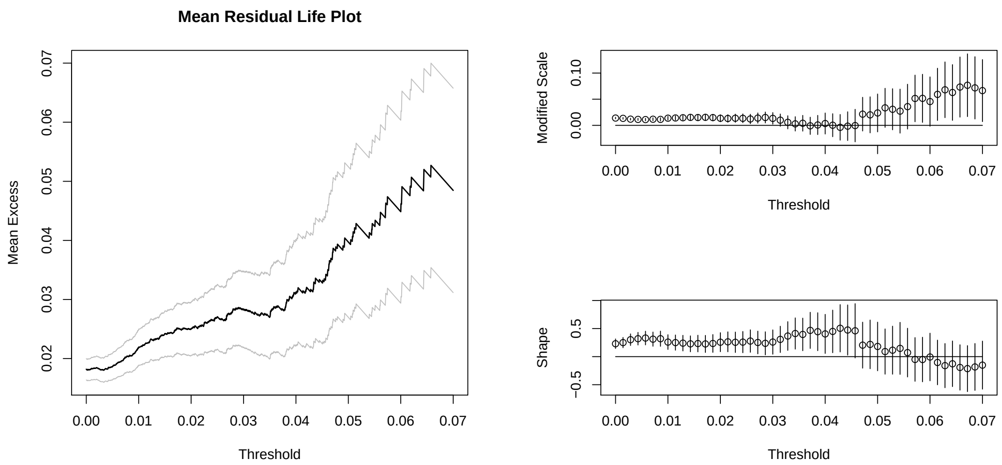
(a) The autocorrelation function.

(b) QQ-plot with the normal distribution represented as the line.

Figure 12: The autocorrelation function and the QQ-plot.

mean residual plot shows linearity from 0 to 0.03, whilst figure 13(b) shows a closely constant modified shape when the threshold is between 0 and 0.03, and the same applies for the shape parameter. Then we have the interval where a suitable threshold can be picked, but we do also need to see more closely how the estimates of the modified scale and shape parameter change with the threshold choice in table. We see that there is no reason for not choosing the threshold 0.03. This gives a modified scale $\sigma = 0.07105$ and a shape $\xi = 0.59635$. We have 170 observations above this threshold and this corresponds to 10.45% of the total amount of observations. To check the quality of this model, we plot the probability plot in figure 14(a) and we see that linearity is fairly present. We can look at the stability in figure 14(b) where the standard deviation has been plotted against time (months). It does not look very stationary, but we can see the tendency of stationary parts.

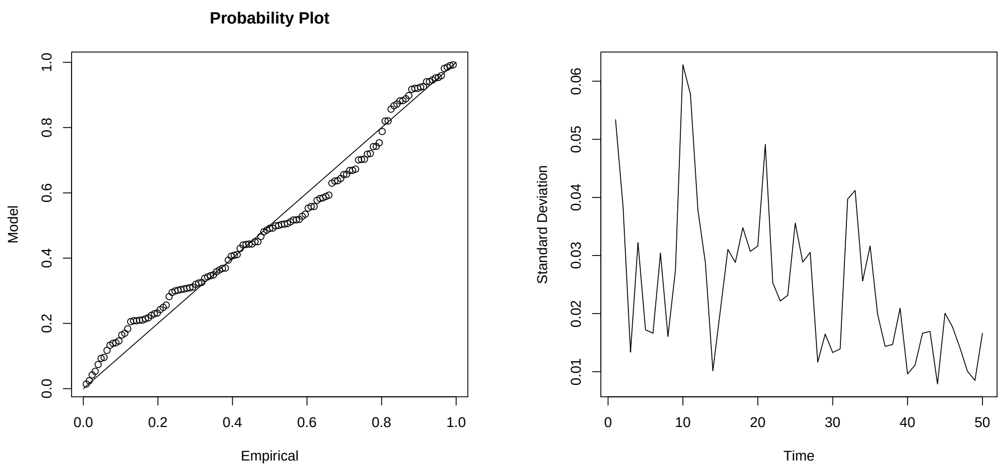
Continuing with the ACER analysis, we see from figure 15(a), we have the first four empirical ACER functions $k = 1, 2, 3, 4$. They will probably converge, so we usually do not need more ACER functions than that. From this plot we also see that the first ACER function is sufficient enough to use when estimating the parameters in the ACER method. The lower ACER function we use, the more of the data is utilised and the estimates will have better accuracy. The extrapolated ACER function and the POT fitted GPD can be seen in figure 15(b), and we see that they lie very close to each other and therefore will produce similar estimations. The parameters of the ACER function are $\tilde{a} = -5.0014$, $b = 0$, $c = 0.7564$, $\gamma = -6.9465$ and $q = 0.6533$. The return level for the ACER method is 1.23701 with the 95% confidence intervals CI = [1.01605 , 1.36627] for the 10^{-6} level of interest.



(a) The mean residual plot.

(b) The estimated modified shape and scale parameter against the threshold.

Figure 13: The plots to help on the choice of threshold.

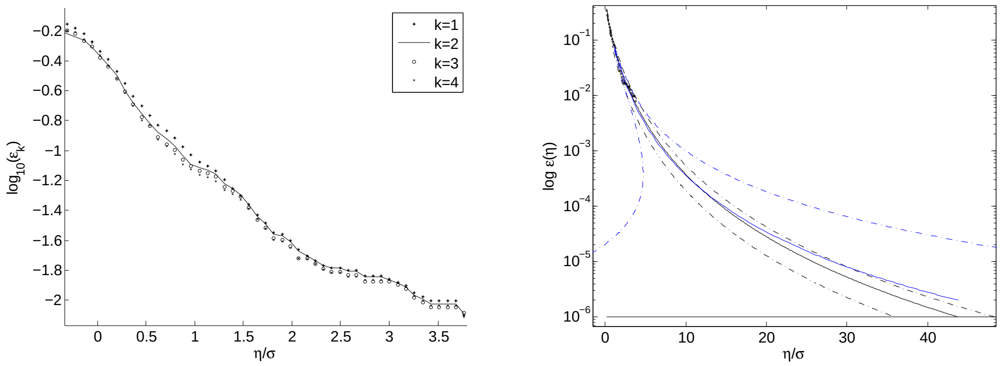


(a) The probability plot.

(b) The development of the standard deviation with respect to time (months).

Figure 14: Probability plot for threshold excess model fitted to the residuals of the EL ICE data and checking for stationarity.

While doing the filtration of this data set without the expiration dates with an AR-GARCH filter, we conclude that the best suitable model in this case consist of an AR(1) part and a GARCH(1,1) part with a normal distribution as the conditioning distribution. Seen in table 2 all the parameters are significant except for a_0 which is not significant no matter kind of conditional distribution being used. The residuals are plotted in figure



(a) The empirical ACER functions plotted against scaled exceedances.

(b) The extrapolated ACER (black) function and POT (blue) fitted GPD with their confidence intervals respectively.

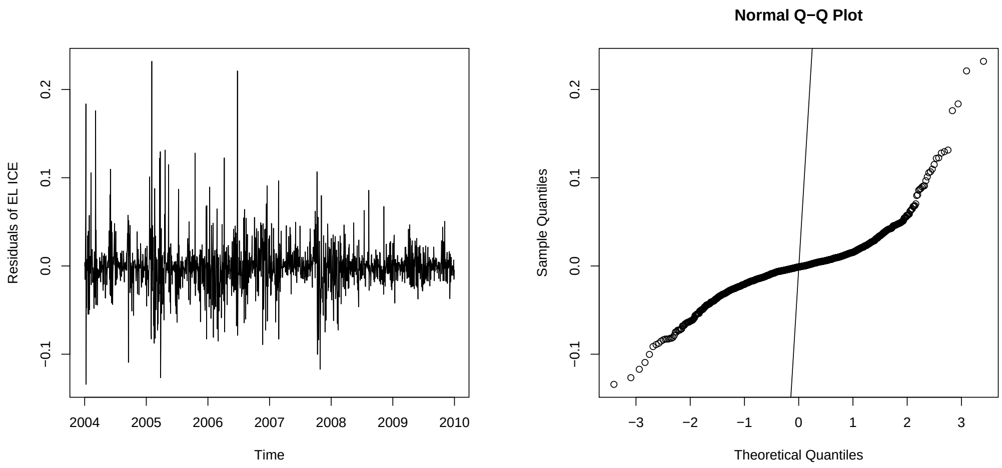
Figure 15: The EL ICE data analysis by the POT and ACER methods.

16(a), and we notice that as before it looks the same as the log returns but these residuals have definitely have different characteristics because of the filtration. In figure 16(b) we see that the residuals are quite different from the normal distribution, as we have seen in the case of the log returns. The excess kurtosis is 10.70 and the skewness is 1.19, and are similar to the ones from the log return series. The autocorrelation function in figure 3.2 shows no sign of large effects of autocorrelation.

AR-GARCH parameters	Estimates (std.error)
a_0	7.791e-04 (5.180e-04)..
a_1	1.221e-01 (3.363e-02)***
ϕ_0	3.632e-05 (7.597e-06)***
ϕ_1	2.515e-01 (4.499e-02)***
ϕ_2	7.495e-01 (3.540e-02)***

Table 2: Estimates of the AR-GARCH parameters for the reduced EL ICE data set with the normal distribution as the conditional distribution with the respectively standard errors in parenthesis. Signif. codes: ***=0.001, **=0.01, *=0.05, .=0.1 and ..=1

The analysis for finding the threshold value u is done as before. From the mean residual plot in figure 18(a) we see that there is a linear region between 0 and 0.04, and from figure 18(b) the estimated modified shape parameter and the scale parameter are constant in the same region 0-0.04. Therefore, there is no reason for not choosing the threshold value of 0.02 in this case also. This leads to 195 observations above the threshold and that equals 12.83% of the total amount of observations. The estimated modified shape parameter is then 0.01871(0.002133) and the estimated shape parameter is 0.26604(0.091514),



(a) The log return of the residuals of the reduced EL ICE observations.

(b) QQ-plot with the normal distribution represented as the line.

Figure 16: The log return plot of the residuals and the QQ-plot for the EL ICE data set without the expiration dates.

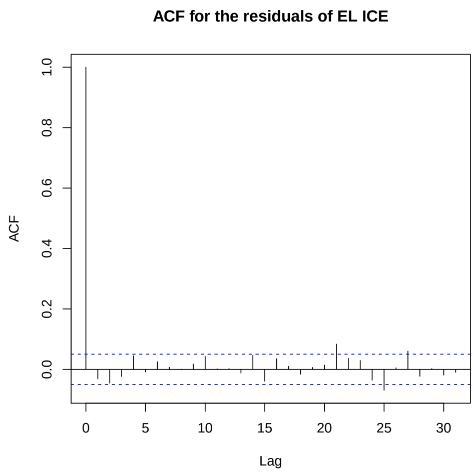
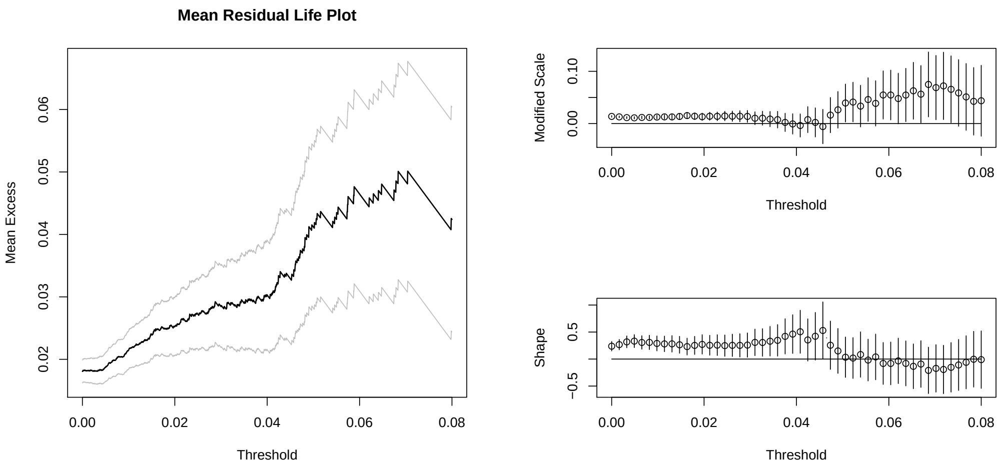


Figure 17: The autocorrelation function.

with the estimated standard errors in parenthesis respectively.

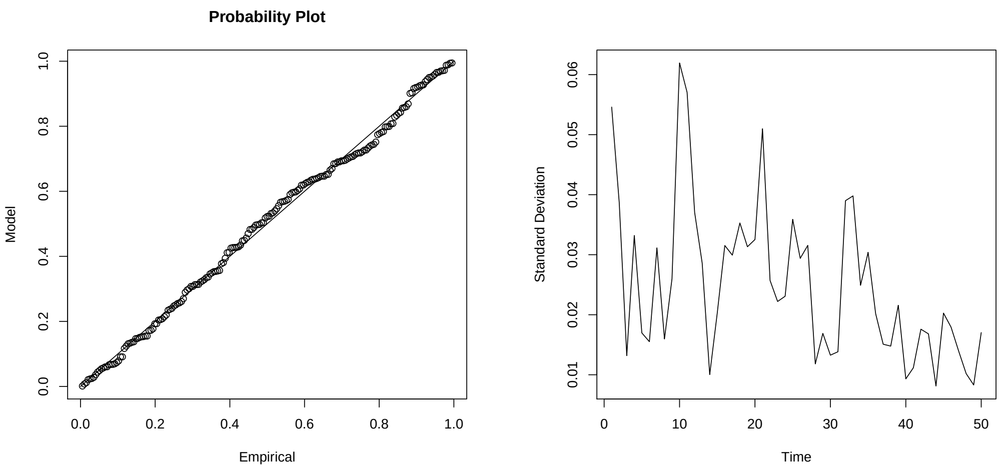
From figure 19(a) we see estimated model plotted against the empirical observations, and there is some deviation from the linear line around the area in the end. But overall, it is a relatively good fit. The stationarity check in figure 19(b) shows does not particularly show consistency in the standard deviation with regards to time. These results are very similar to the situation of the full data set with the expiration dates.



(a) The mean residual plot.

(b) The estimated modified shape and scale parameter against the threshold.

Figure 18: The plots to help on the choice of threshold.

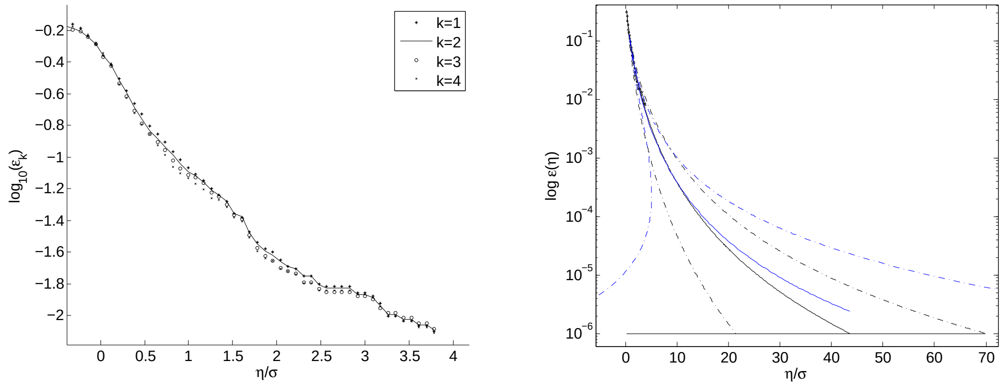


(a) The probability plot.

(b) The development of the standard deviation with respect to time (months).

Figure 19: Probability plot for threshold excess model fitted to the residuals of the reduced EL ICE data and checking for stationarity.

For the ACER analysis, we see in figure 20(b) that the first four empirical ACER functions converge in the tail. Therefore we use the first ACER function to estimate the parameters in the ACER analysis. From the extrapolated ACER function in black and the POT fitted GPD in blue with confidence intervals in figure 20(b) we see that again this plot is very much the same as the previous plot done with the log returns. This can be supported



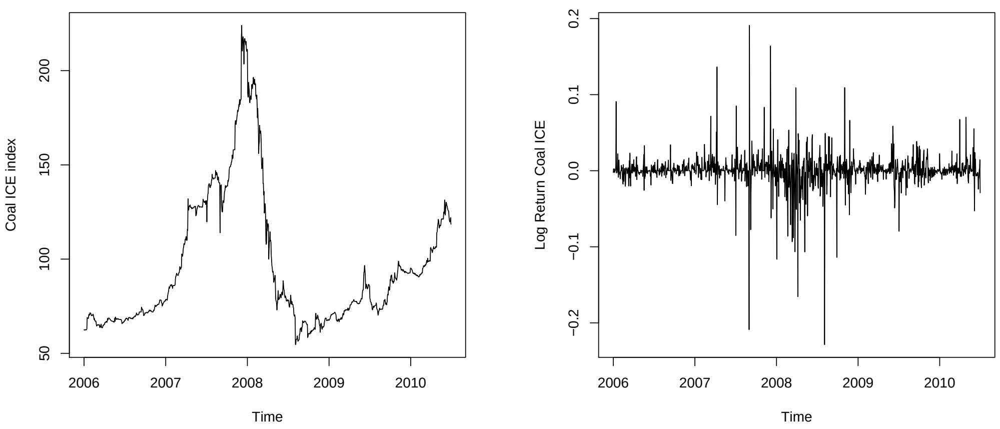
(a) The empirical ACER functions plotted against scaled exceedances.

(b) The extrapolated ACER (black) function and POT (blue) fitted GPD with their confidence intervals respectively.

Figure 20: The residuals of the reduced EL ICE data analysis by the POT and ACER methods.

by looking at specific return levels as we have done in the later section with comparisons and further discussion. The parameters of the ACER function are $\tilde{a} = -5.0081$, $b = 0$, $c = 0.7587$, $\gamma = -6.9326$ and $q = 0.6151$. The return level is 1.22702 with the 95% confidence intervals CI = [0.601943, 1.96883] for the 10^{-6} level of interest.

3.3 Coal ICE with expiration dates



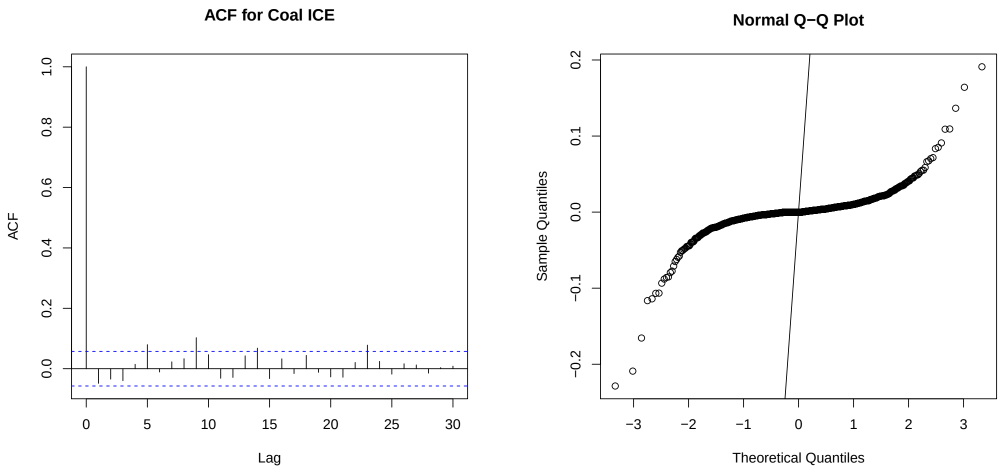
(a) The index of Coal ICE from period 17.07.2006-27.01.2011.

(b) The log return of the EL ICE observations.

Figure 21: Index plot and log return plot of the observations for Coal ICE

The Coal ICE data set is Rotterdam monthly coal futures and spans over 4.5 years approximately with 1168 observations. The index plot for the Coal ICE data set is seen in figure 21(a). We see that there are a large increase in the index from year 2007 to 2008, and after that there is a large fall in the index. We have seen that kind of behaviour before, for example in the data set of EL ICE, and are likely caused by the financial crisis in the 2007-2008 period. The log return of this data set can be seen in figure 21(b). We see that there are larger values around the area mentioned before. The autocorrelation function seen in figure 22(a) shows low presence of autocorrelation with the exception of around lag 9. Even though an autocorrelation value of approximately 0.1 is larger than the rest of the autocorrelation-values, it is not a large value seen in a bigger perspective. We then treat this data set with no concern about autocorrelation being present. The figure in 22(b) shows how large a difference this data set is from the normal distribution. This is supported by the fact that the excess kurtosis is 30.37 and the skewness is -1.26 of the log returns, which leads to an idea of this being a distribution with heavier tail than the normal distribution and with a slightly heavier left tail than the right tail.

To find a suitable threshold value u using the POT method, we can look at the mean residual plot in figure 23(a) and the estimated modified scale and shape parameter against different threshold values in figure 23(b). The mean residual plot show linearity in the area 0-0.02. But it is harder to see the linearity the further the threshold increases. The areas where the estimated modified scale parameter is constant are 0-0.02 and 0.02-0.04, and for the estimated shape parameter are the areas 0.01-0.02 and 0.02-0.03. When considering the threshold within these areas and looking at the change in the parameters with their standard errors respectively, we choose our threshold value to be 0.015. As mentioned before, we do not have many observations in this case in comparison to the last case where we had 1627 observations in comparison to the 1168 observations with Coal ICE. With the threshold value of 0.015, we have 121 observations above the threshold. That represents 10.36 % of the total observations. The estimated modified scale parameter is 0.01273(0.001902) and the estimated shape parameter is 0.36214(0.125198) with their respectively standard errors in parenthesis.



(a) The autocorrelation function.

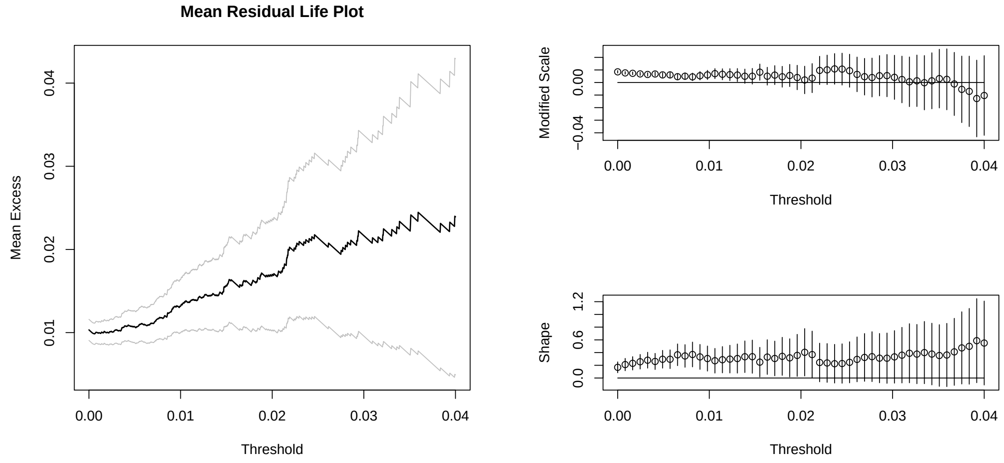
(b) QQ-plot with the normal distribution represented as the line.

Figure 22: The autocorrelation function and the QQ-plot.

To check the model performance and quality, we look at the probability plot in figure 24(a). Except for a part in the middle, the empirical data and the model align to a fairly linear line. The stationarity in figure 24(b) shows that the data set is not stationary. The standard deviation, and thus the volatility, changes a lot with respect to time.

The plots for the ACER analysis can be seen in the figures 25(a) and 25(b). We see that the ACER functions plotted against the scaled exceedances shows that the first four ACER functions converge, $k = 1, 2, 3, 4$, and thus we use the first ACER function in the further analysis. We see that the POT fitted GPD predicts a heavier tail than the extrapolated ACER function. The parameters of the ACER function are $\tilde{a} = -5.0011$, $b = 0$, $c = 0.6306$, $\gamma = -7.4101$ and $q = 0.9892$. The return level is 1.144 with the 95% confidence intervals CI = [0.903619,1.28721] for the 10^{-6} level of interest.

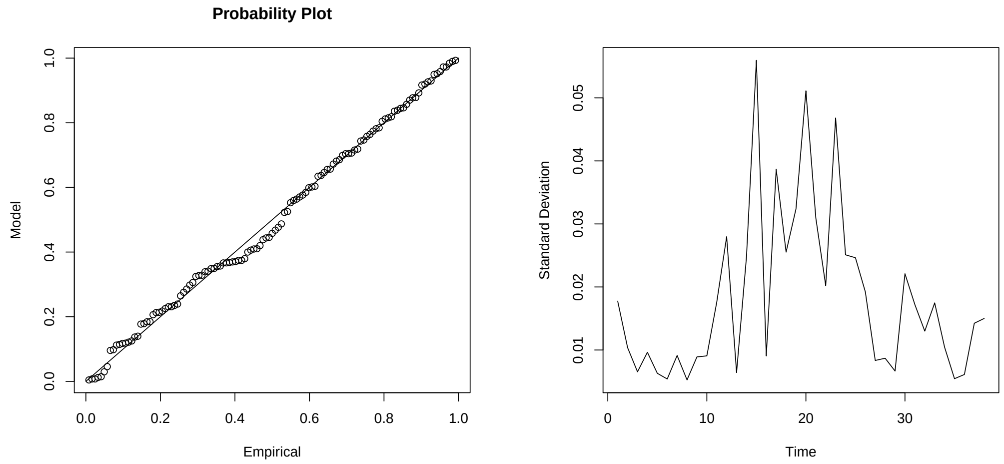
Now we are wondering whether if filtrating the observations through an AR-GARCH filter will give a model who will capture the volatility better and therefore be a more suitable model for predicting large price changes in the future. Therefore we now filtrate the Coal ICE data set through a GARCH(1,1) filter with the conditional distribution



(a) The mean residual plot.

(b) The estimated modified shape and scale parameter against the threshold.

Figure 23: The plots to help on the choice of threshold.

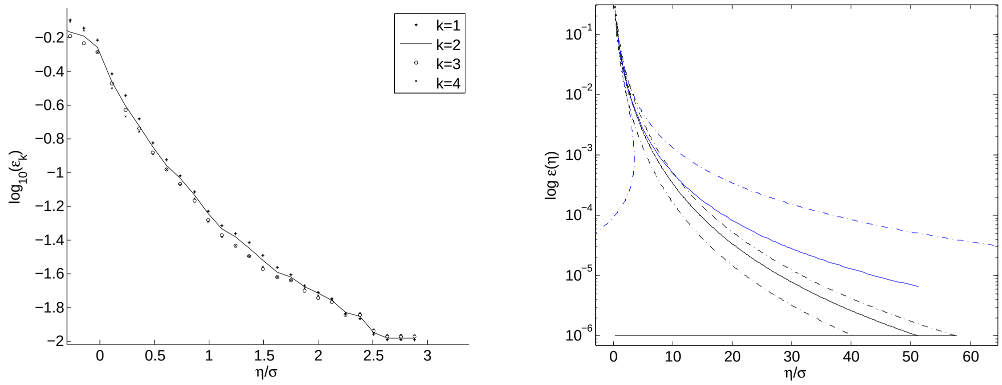


(a) The probability plot.

(b) The development of the standard deviation with respect to time (months).

Figure 24: Probability plot for threshold excess model fitted to Coal ICE data and checking for stationarity.

as the normal distribution. The method for finding the parameters in AR-GARCH is as before described. None of the parameters are significant for higher AR levels, so it is simpler for us to choose the simplest model and that is with an AR(0) part. The log return of the standard residuals of the Coal ICE observations can be seen in figure 26(a). We see that in comparison to the log return of the data set, they are almost the same. The QQ-plot in figure 26(b) is far from normal distributed, as in the case of the log returns. This information can also be supplied with the excess kurtosis of 30.72 and skewness of -1.15. In figure 3.3 we see that the autocorrelation is negligible.



(a) The empirical ACER functions plotted against scaled exceedances.

(b) The extrapolated ACER (black) function and POT (blue) fitted GPD with their confidence intervals respectively.

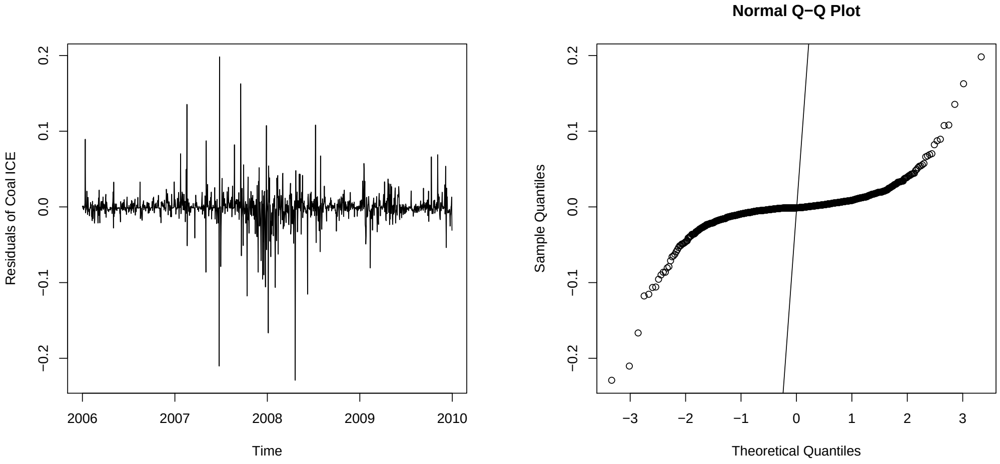
Figure 25: The Coal ICE data analysis by the POT and ACER methods.

AR-GARCH parameters	Estimates (std.error)
a_0	1.310e-03 (4.372e-04)**
ϕ_0	1.087e-06 (3.315e-07)**
ϕ_1	4.016e-02 (5.190e-03)***
ϕ_2	9.653e-01 (3.095e-03)***

Table 3: Estimates of the AR-GARCH parameters for the full Coal ICE data set with the normal distribution as the conditional distribution with the respectively standard errors in parenthesis. Signif. codes: ***=0.001, **=0.01, *=0.05, .=0.1 and ..=1

To decide upon the threshold, we look at the mean residual plot in figure 28(a) and the plot of the estimated modified scale and shape parameter against the threshold in figure 28(b). In the mean residual plot there are areas of linearity in 0-0.02 and 0.02-0.03. In the plot of the estimated modified scale and shape parameter against the threshold, we see the estimated parameters are constant from 0 to 0.02 for both parameters. After careful deliberation, the threshold value is chosen to be $u = 0.01$ with 169 observations above the threshold (14.47%). This gives us the estimated modified scale parameter 0.01078(0.001365) and the estimated shape parameter 0.38854(0.107468) with their respectively standard deviations in parenthesis.

Now we perform a quality check on the model. For the probability plot seen in fig-



(a) The log return of the residuals of the Coal ICE observations.

(b) QQ-plot with the normal distribution represented as the line.

Figure 26: The log return plot of the residuals and the QQ-plot for the residuals of the Coal ICE data set.

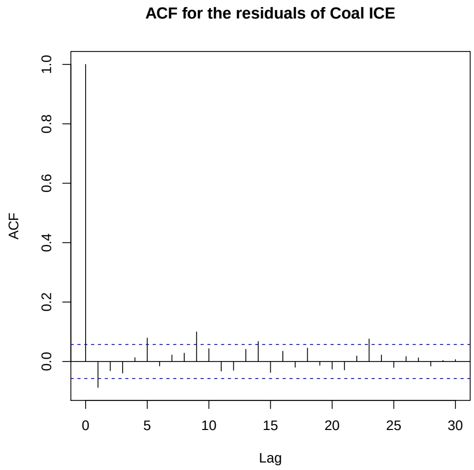
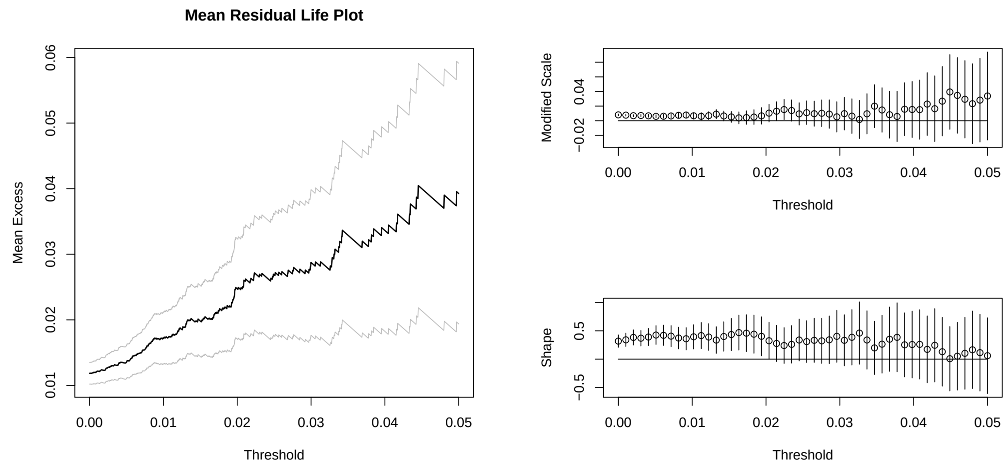


Figure 27: The autocorrelation function.

ure 29(a), we see that the empirical observations do correspond very well with the model we have set up. From figure 29(b) we check for stationarity. The residuals seem to be more stationary in three different parts in the plot, but the figure does not show consistency over the whole area as for the case with the Coal ICE data set before filtration.

Continuing with the ACER analysis, we see from figure 30(a) that the four first ACER function converge so we only use the first ACER function to estimate the parameters

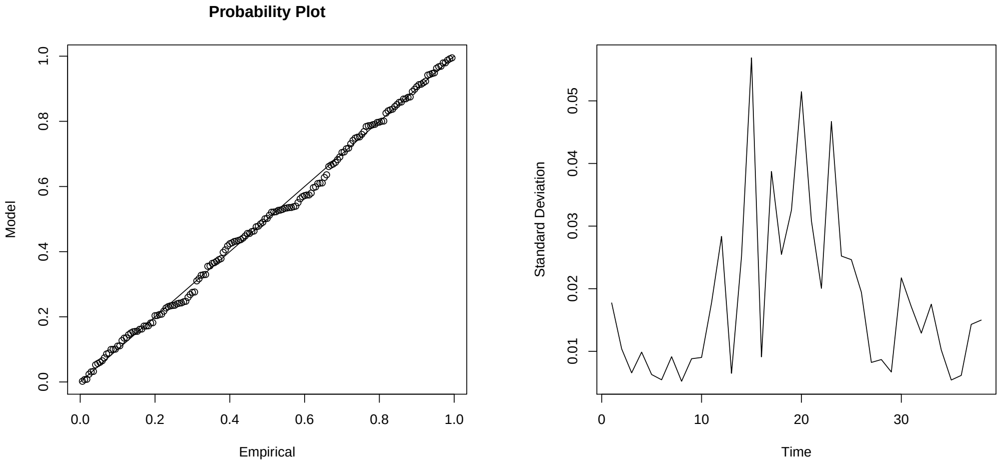


(a) The mean residual plot.

(b) The estimated modified shape and scale parameter against the threshold.

Figure 28: The plots to help on the choice of threshold.

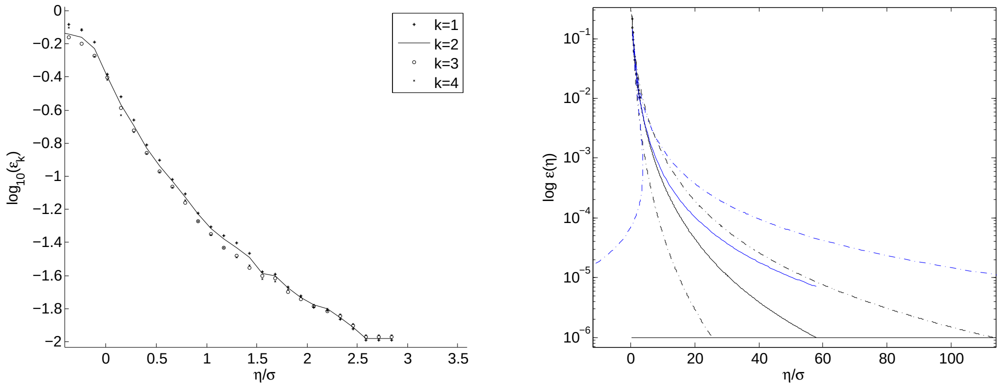
in the analysis. From figure 30(b) we see that now the figure looks much the same as the one with the log returns. We can also notice that the confidence intervals in this last plot is larger than the previous plot. The parameters of the ACER functions are $\tilde{a} = -5.0462$, $b = 0$, $c = 0.6268$, $\gamma = -7.0580$ and $q = 0.8096$. The return level is 1.29243 with the 95% confidence intervals CI $= [0.565158, 2.52772]$ for the 10^{-6} level of interest.



(a) The probability plot.

(b) The development of the standard deviation with respect to time (months).

Figure 29: Probability plot for threshold excess model fitted to the residuals of the Coal ICE data and checking for stationarity.



(a) The empirical ACER functions plotted against scaled exceedances.

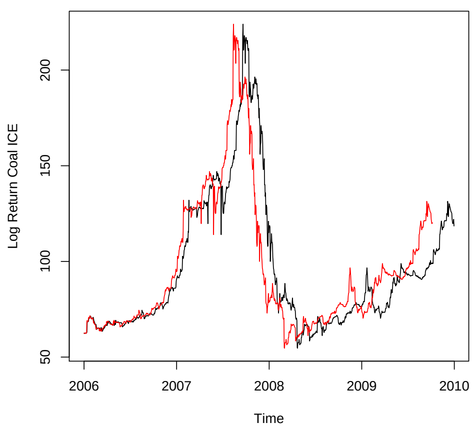
(b) The extrapolated ACER (black) function and POT (blue) fitted GPD with their confidence intervals respectively.

Figure 30: The residuals of the Coal ICE data analysis by the POT and ACER methods.

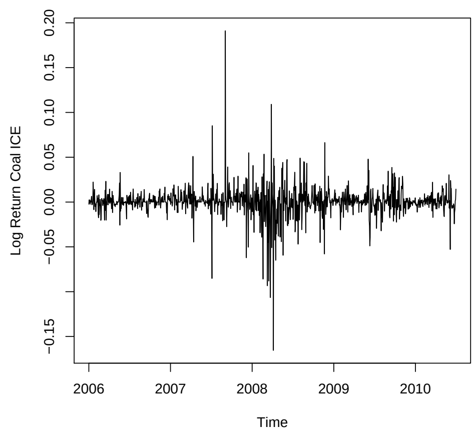
3.4 Coal ICE without expiration dates

Now we have done the analysis with the full Coal ICE data set, and we want to find out if removing the expiration dates will have an effect on the results in the analysis. If there are differences, we would also like to find out what differences there are and how large a role do they play in the complete analysis.

The Coal ICE data set has date of expiration on the last Friday in every month. The data set without the dates of expiration consists of 1101 observations. It can be mentioned that the maximum and minimum log return observation is 0.191 and -0.1655 for the reduced data set, while being 0.191 and -0.2286 for the data set with the dates of expiration. We can conclude with the fact that the maximum has not changes and therefore is not influenced by the removing the dates of expiration. But the largest minimum is caused by the expiration date of the contract. One can also compare the log return figure 31(b) with the log return figure 21(b) in last section. Even though the index plot for this reduced data set in figure 31(a) look quite similar to the one from the previous section with the full data set for Coal ICE, we can see that the log return plots show some large changes from the previous case with the full Coal ICE data set. Here we see more clearly that the some of the largest values, both negative and positive, have disappeared by removing the dates of expiration. This is more visible in this case of Coal ICE in comparison to the case of EL ICE. But there are still very large log return values around the area around year 2007-2008.



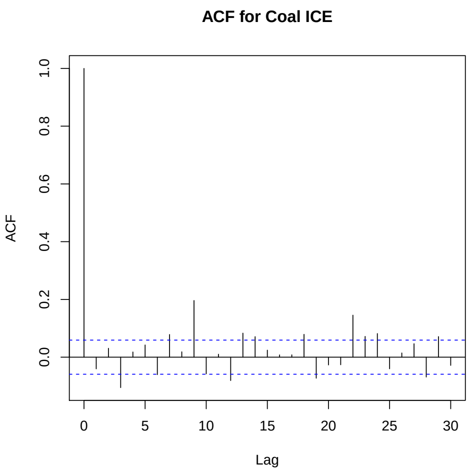
(a) The Coal ICE data set. Red:the full data set. Black:the data set without the last Friday in every month.



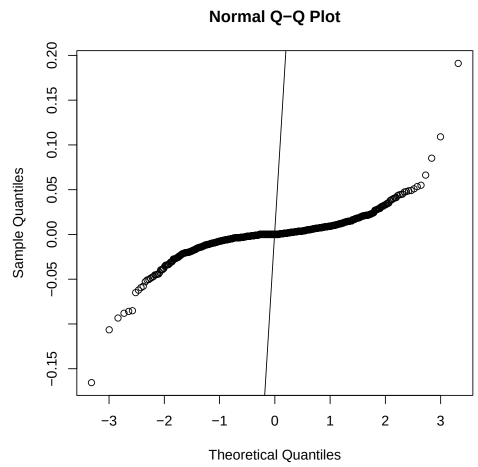
(b) The log return of the reduced Coal ICE observations.

Figure 31: Index plot and log return plot of the observations for the reduced Coal ICE data set.

It seems that the autocorrelation function in figure 32(a) shows a larger correlation in this reduced data set than in the full data set seen in figure 22(a). The largest autocorrelation value in this case is around twice as large than the largest value in the previous section, 0.2 in comparison to 0.1. But although there are some larger values for the autocorrelation around lag 10 and lag 25, there are no need to worry because the values are still quite small when looking at the full picture. The excess kurtosis is 32.56 and the skewness is 0.76. This can be supported by the QQ-plot where the normal distribution is represented as the black line in figure 32(b).



(a) The autocorrelation function.

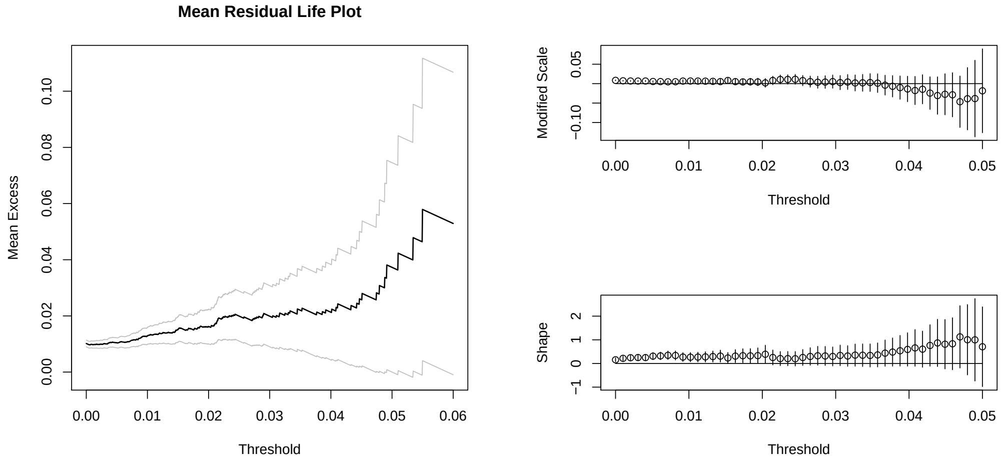


(b) QQ-plot with the normal distribution represented as the line.

Figure 32: The autocorrelation function and the QQ-plot.

Now we try to decide upon a suitable threshold value, u, with the POT method. First, we look at the mean residual plot in figure 33(a) and notice the areas where there are linearity are around 0-0.02 and 0.02-0.04. Then we look at the plot for the estimated modified scale and shape parameter against different choices of threshold values in figure 33(b). For the estimated modified scale parameter we can see areas where it is constant from around 0 to 0.03, and for the estimated shape parameter from approximately around 0 to 0.035. After looking at how the estimated parameters behave around the before mentioned areas, we can choose a suitable threshold value. In this case, we have chosen $u = 0.01$ which gives us 161 observations above threshold and that represents 14.62 % of the total observation. When there are so few observations in the first place, then we have to try and conserve as much data as possible. This choice of threshold gives the estimated modified scale parameter $\sigma = 0.009165(0.001114)$ and the estimated shape parameter $\xi = 0.281176(0.098886)$.

To check the quality of the model we have now formed, we can consider a probability plot as in figure 34(a). We see that the empirical observations and the model align pretty well with the linear line. This is the best alignment till now, but we must not



(a) The mean residual plot.

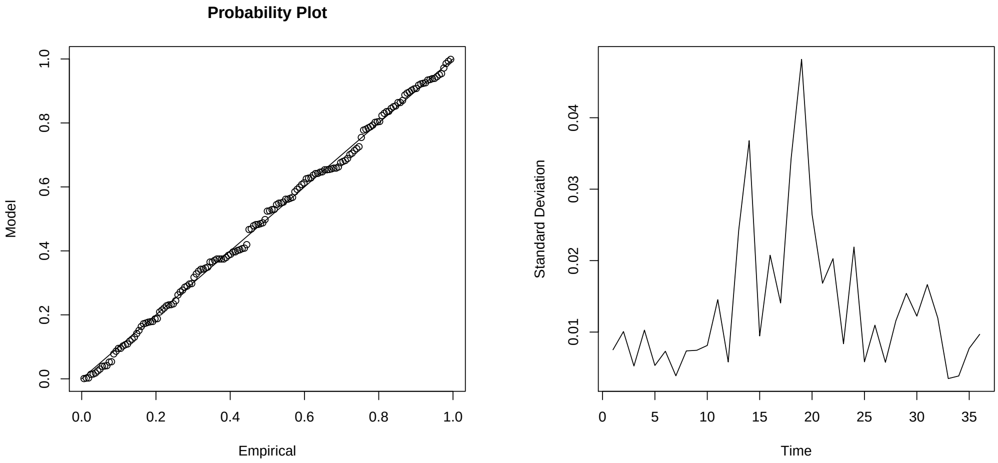
(b) The estimated modified shape and scale parameter against the threshold.

Figure 33: The plots to help on the choice of threshold.

rely everything on this one method for checking the quality of a model. We see in figure 34(b) that the data set does not likely seem to be stationary. There are large differences in the standard deviation between the part in the middle and the parts around.

Continuing with the ACER analysis, we first look at the figure 35(a) and see that the four first ACER functions converge in the middle. Using the first ACER function in the further analyse and estimation, we get the extrapolated ACER function and the GPD fitted POT in figure 35(b). We see that the aforementioned predicts a heavier tail than the extrapolated ACER function. The parameters of the ACER function are $\tilde{a} = -1.0216$, $b = 0.0018$, $c = 0.5610$, $\gamma = -26.7932$ and $q = 0.8798$. The return level is 0.469174 with the 95% confidence intervals CI = [0.375111 , 0.531734] for the 10^{-6} level of interest.

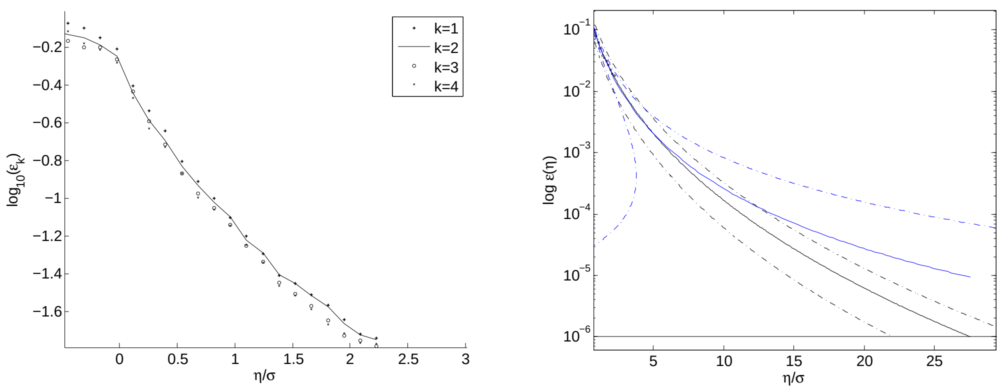
Now we filtrate the data set through an AR-GARCH filter, as done in the previous section for the EL ICE data set. We find out by looking at number of significant parameters in the model which model we should use on this data set. The model used in this case consists an AR(4) part and a GARCH(1,1) part with a normal distribution as the conditioning distribution. We see from table 4 that every parameter are significant down to a 0.05 level except for a_2 and a_3. Let us now do the same analysis for the residuals after filtration seen in figure 36(a). We notice that we used an AR(0)-GARCH(1,1) filter in the previous case when the data set had the dates of expiration, while this case we have an AR(4)-GARCH(1,1) filter. In figure 36(b) gives us an impression of how different this data set of residuals is from the normal distribution. The excess kurtosis is 24.70 and the skewness is -0.14. The autocorrelation function seen in figure 3.4 shows larger values than we have ever seen before in this thesis, but there is still no valid concern about autocorrelation at these levels. Declustering is one method to get rid of the autocorrelation, if that should be needed.



(a) The probability plot.

(b) The development of the standard deviation with respect to time (months).

Figure 34: Probability plot for threshold excess model fitted to Coal ICE data and checking for stationarity.



(a) The empirical ACER functions plotted against scaled exceedances.

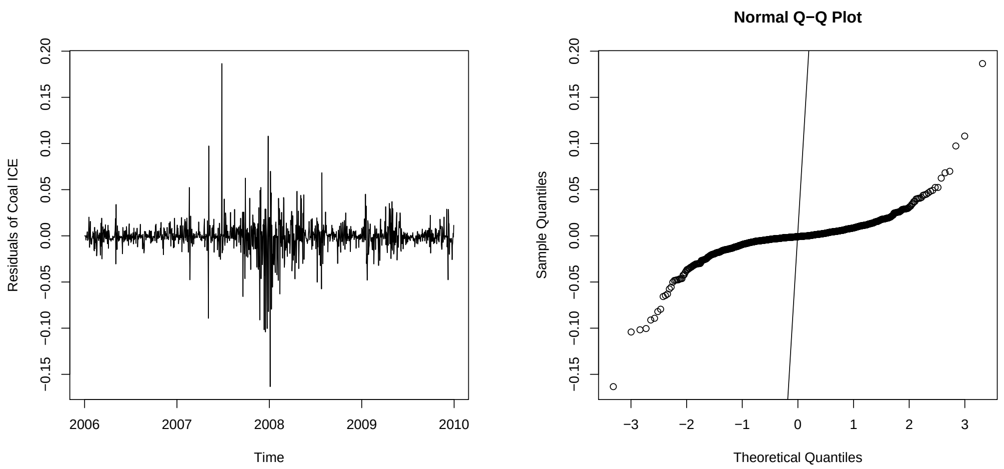
(b) The extrapolated ACER (black) function and POT (blue) fitted GPD with their confidence intervals respectively.

Figure 35: The EL COAL data analysis by the POT and ACER methods.

Finding the threshold value, u, has been done as before. From the mean residual plot in figure 38(a) we see that there seems to be linearity in the area from 0 to 0.02. While from the plot for the estimated modified shape and scale parameter seen in figure 38(b) shows

AR-GARCH parameters	Estimates (std.error)
a_0	1.340e-03 (3.426e-04)***
a_1	1.266e-01 (4.333e-02)**
a_2	4.706e-02 (4.728e-02)..
a_3	7.432e-02 (4.204e-02).
a_4	-1.610e-01 (3.049e-02)***
ϕ_0	4.234e-06 (1.284e-06)***
ϕ_1	3.750e-01 (4.260e-02)***
ϕ_2	7.561e-01 (1.889e-02)***

Table 4: Estimates of the AR-GARCH parameters for the reduced Coal ICE data set with the normal distribution as the conditional distribution with the respectively standard errors in parenthesis. Signif. codes: ***=0.001, **=0.01, *=0.05, .=0.1 and ..=1



(a) The log return of the residuals of the reduced Coal ICE observations.

(b) QQ-plot with the normal distribution represented as the line.

Figure 36: The log return plot of the residuals and the QQ-plot for the reduced Coal ICE data set without the expiration dates.

that the parameters are approximately constant also in the area from 0 to 0.02. After trying several possibilities for the threshold value we set the u to be 0.01. This gives 154 observations above the threshold, and that represents 13.99% of the total amount. The modified scale parameter is then estimated to be 0.008543(0.001108) and the shape parameter is 0.346196(0.109587) with their standard errors in parenthesis.
To perform a quality check on this model we can consider the probability plot in 39(a). The points seem to align very well with the linear line. We check for stationarity in figure 39(b), and conclude as before that there seems to be three parts where there are stationarity but overall the residuals do not seem to be stationary.

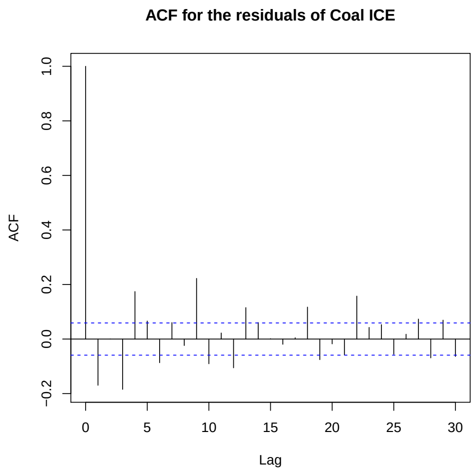
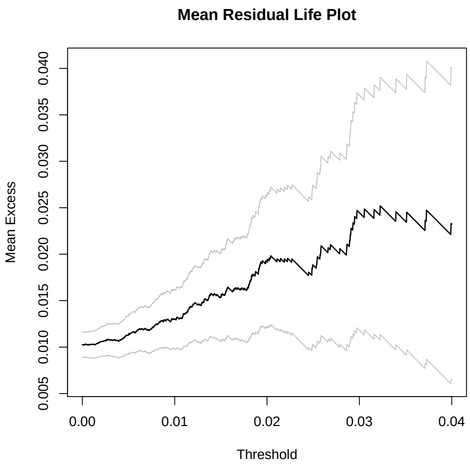
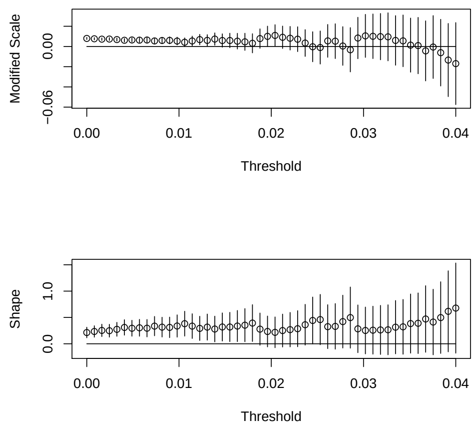


Figure 37: The autocorrelation function.



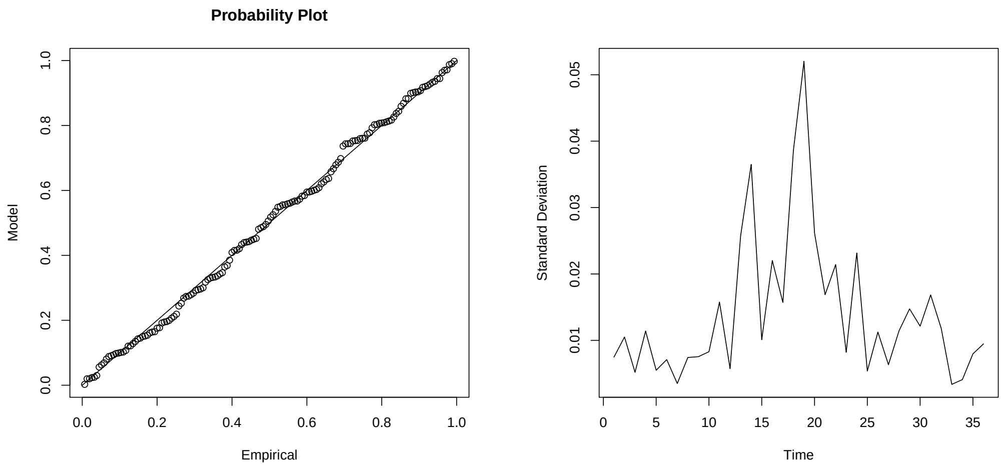
(a) The mean residual plot.



(b) The estimated modified shape and scale parameter against the threshold.

Figure 38: The plots to help on the choice of threshold.

The ACER analysis shows that the first four ACER functions converge in the middle of the plot in figure 40(a), so we use the first ACER function to estimate the parameters in the further ACER analysis. The extrapolated ACER function and the POT fitted GPD can be seen in figure 40(b). We see the POT fitted GPD estimates a much heavier tail than the extrapolated ACER function. Although the plot seem very much alike the one in figure 35(b), we see that the extrapolated ACER function is closer to the POT fitted GPD in the previous case. The confidence intervals of the extrapolated ACER function

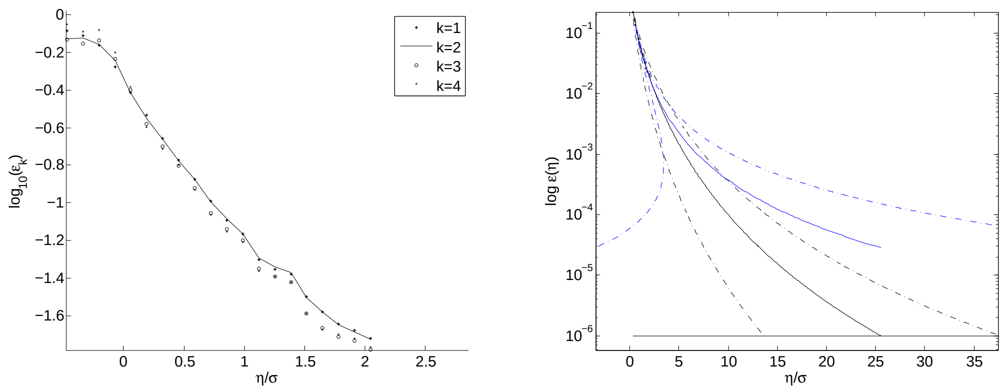


(a) The probability plot.

(b) The development of the deviation with respect to time (months).

Figure 39: Probability plot for threshold excess model fitted to the residuals of the reduced Coal ICE data and checking for stationarity.

are also larger than the before. The parameters of the ACER function are $\tilde{a} = -5.0215$, $b = 0$, $c = 0.7394$, $\gamma = -10.0474$ and $q = 0.6373$. The return level is 0.452489 with the 95% confidence intervals CI $= [0.241724,\ 0.667962]$ for the 10^{-6} level of interest.

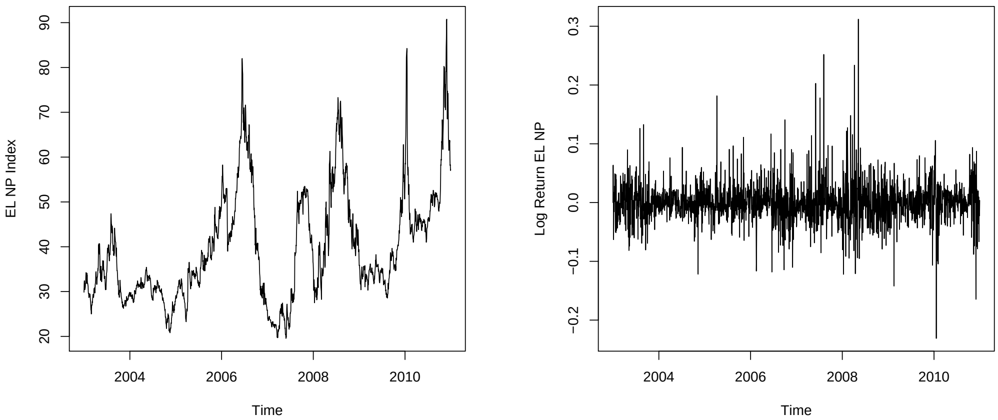


(a) The empirical ACER functions plotted against scaled exceedances.

(b) The extrapolated ACER (black) function and POT (blue) fitted GPD with their confidence intervals respectively.

Figure 40: The residuals of the reduced Coal ICE data analysis by the POT and ACER methods.

3.5 EL NP



(a) The index of EL NP from period 07.04.2003-31.01.2011.

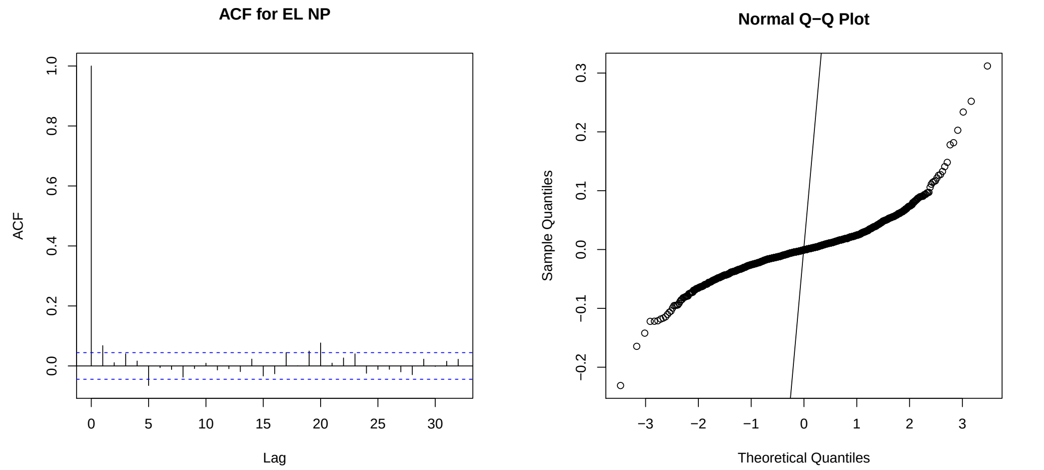
(b) The log return of the EL ICE observations.

Figure 41: Index plot and log return plot of the observations for EL NP

The EL NP data set is NordPool monthly electricity futures from 07.04.2003 to 31.01.2011. This represents 8 years with 1953 observations. The index plot of the EL NP data set can be seen in figure 41(a), and it is clearly that there are a few large fluctuations in the index prices. The log return of the EL ICE index, seen in figure 41(b), shows also some large values. The autocorrelation function in figure 42(a) shows no sign of large effects of correlation. This data set is also very different distributed in comparison to the normal distribution, and that is seen in figure 42(b). The excess kurtosis of 9.48 is not very large in comparison to the other data sets and skewness of 0.89, but this is only some ways of considering the difference between two distributions.

To find a suitable threshold value u using the POT method, we can look at the mean residual plot in figure 43(a) and the estimated modified scale and shape parameter against different threshold values in figure 43(b). The mean residual plot show linearity in the areas 0-0.02 and 0.025-0.05. The areas where the estimated modified scale parameter is constant are 0-0.04, and for the estimated shape parameter are the areas 0-0.04 and 0.045-0.07. When considering the threshold within these areas and looking at the change in the parameters with their standard errors respectively, we choose our threshold value to be 0.03. Then we have 246 observations above the threshold. That represents 12.6% of the total observations. The estimated modified scale parameter is 0.02403(0.002197) and the estimated shape parameter is 0.15538(0.066732) with their respectively standard errors in parenthesis.

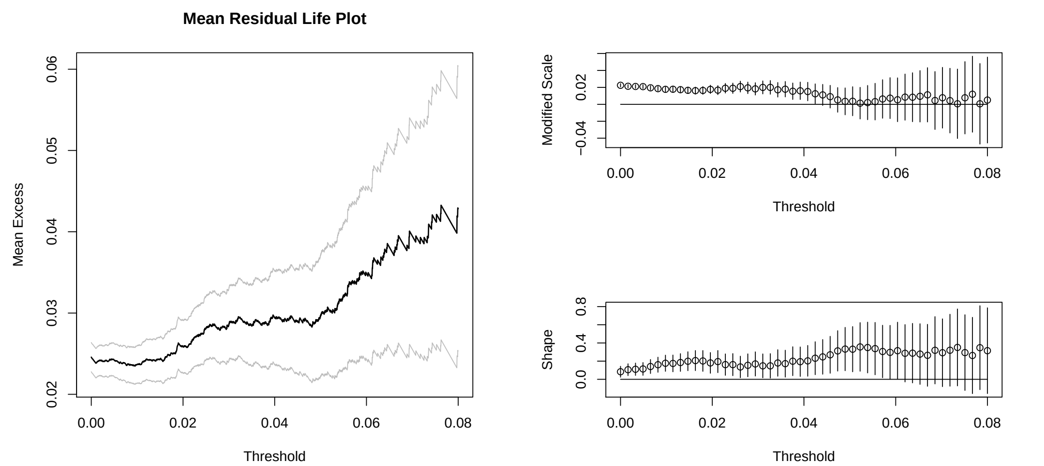
The diagnostic tools to check the quality of the fitted model can be the probability plot seen in figure 44(a). We see that the fit is quite good except for a small area in the middle. The stationarity plot in figure 44(b) shows trends of stationarity, but it is hard to say for sure whether the data set is stationary or not just from the plot.



(a) The autocorrelation function.

(b) QQ-plot with the normal distribution represented as the line.

Figure 42: The autocorrelation function and the QQ-plot.

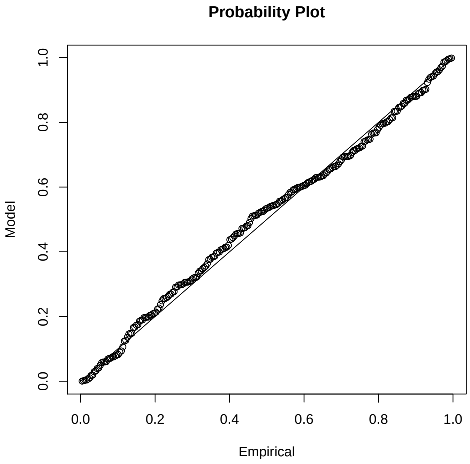


(a) The mean residual plot.

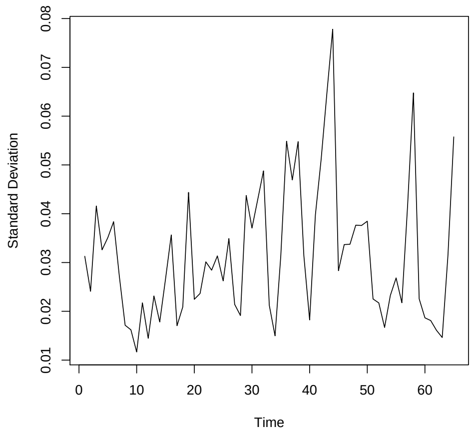
(b) The estimated modified shape and scale parameter against the threshold.

Figure 43: The plots to help on the choice of threshold.

The plots for the ACER analysis can be seen in the figures 45(a) and 45(b). We see that the ACER functions plotted against the scaled exceedances shows that the first four



(a) The probability plot.



(b) The development of the standard deviation with respect to time (months).

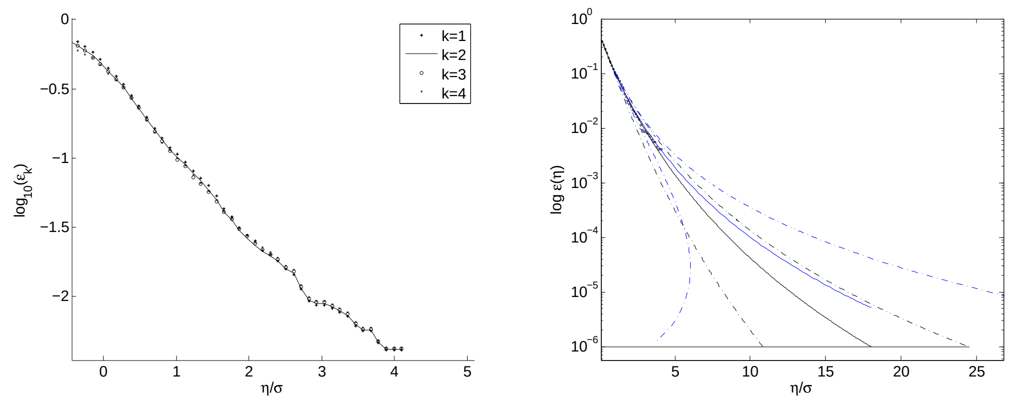
Figure 44: Probability plot for threshold excess model fitted to EL NP data and checking for stationarity.

ACER functions converge fully all the way, $k = 1, 2, 3, 4$, and thus we use the first ACER functions in the further analysis. We see that the POT fitted GPD predicts a slightly heavier tail than the extrapolated ACER function in figure 45(b), but the POT method gives a much larger confidence interval than the ACER method. The parameters of the ACER function are $\tilde{a} = -5.0023$, $b = 0$, $c = 0.9773$, $\gamma = -9.2976$ and $q = 0.5301$. The return level is 0.618708 with the 95% confidence intervals CI = [0.372341,0.841372] for the 10^{-6} level of interest.

Now we take the data set and filtrate it through an AR-GARCH filter. In this case we used an AR(0) part and a GARCH(1,1) part. One fact that has to be pointed out in this case is that every parameter except a_0 are significant whether we choose the normal distribution, Student's t distribution or QMLE (which is also based on the normal distribution) as the conditional distribution. The normal distribution has been chosen as the conditional distribution though, because it gives the least amount of parameters and therefore leads to an easier model. In figure 46(b) we see how the residuals differs from the normal distribution. The residuals from the log returns of the data set EL NP is plotted in figure 46(a). The excess kurtosis of the standard residuals is 9.48 and the skewness is 0.89, and remains quite similar to the ones from the log returns. The autocorrelation function can be seen in figure 3.5 and we see that there are no large values for the autocorrelation function.

To decide upon the threshold, we look at the mean residual plot in figure 48(a) and the plot of the estimated modified scale and shape parameter against the threshold in figure

48(b). In the mean residual plot there are areas of linearity in 0-0.02, 0.03-0.05 and 0.05-0.08. The estimated modified scale parameter and the estimated shape parameter are approximately constant around 0-0.04. After some deliberaton and further investigation, the threshold value is chosen to be $u = 0.03$ with 238 observations above the threshold (12.19%). This gives us the estimated modified scale parameter 0.02457(0.002257) and the estimated shape parameter 0.14625(0.066048) with their respectively standard deviation in parenthesis. Doing the same quality check as before, we see in figure 49(a) the probability plot where the there is a pretty good fit besides some minor deviations. It can almost seem to be some kind of a trend that the observations are above the linear line in the first half and below the line in the last half of the plot. From figure 49(b) we see that the standard residuals seem to be stationary except for a few larger peaks.



(a) The empirical ACER functions plotted against scaled exceedances.

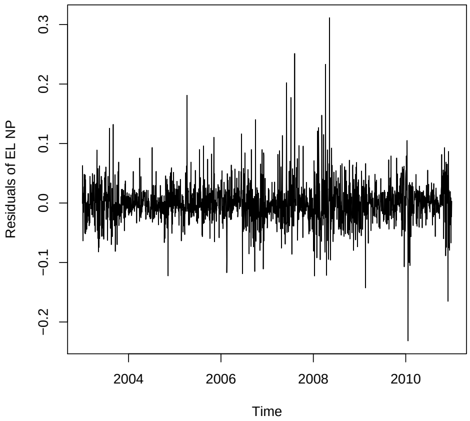
(b) The extrapolated ACER (black) function and POT (blue) fitted GPD with their confidence intervals respectively.

Figure 45: The EL NP data analysis by the POT and ACER methods.

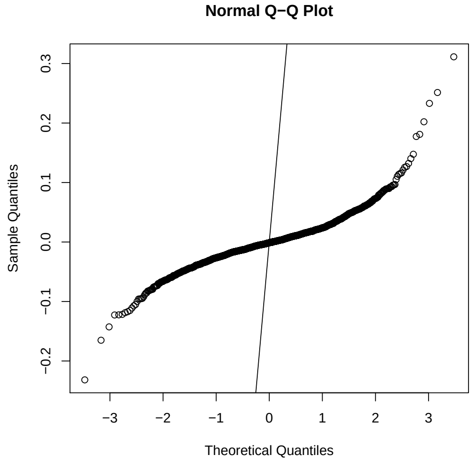
AR-GARCH parameters	Estimates (std.error)
a_0	5.937e-04 (6.420e-04)..
ϕ_0	1.824e-05 (4.275e-06)***
ϕ_1	8.080e-02 (1.057e-02)***
ϕ_2	9.100e-01 (9.879e-03)***

Table 5: Estimates of the AR-GARCH parameters for the EL NP data set with the normal distribution as the conditional distribution with the respectively standard errors in parenthesis. Signif. codes: ***=0.001, **=0.01, *=0.05, .=0.1 and ..=1

Continuing with the ACER analysis, we see from figure 50(a) that the four first ACER function fully converge so we only use the first ACER function to estimate the parameters in the analysis. The plot with the extrapolated ACER function and the POT fitted GPD



(a) The log return of the residuals of the EL NP observations.



(b) QQ-plot with the normal distribution represented as the line.

Figure 46: The log return plot of the residuals and the QQ-plot for the residuals of the EL NP data set.

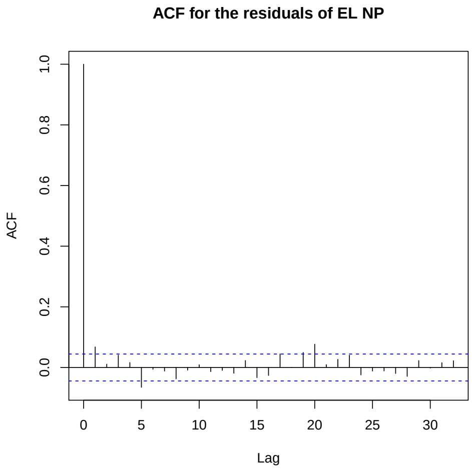
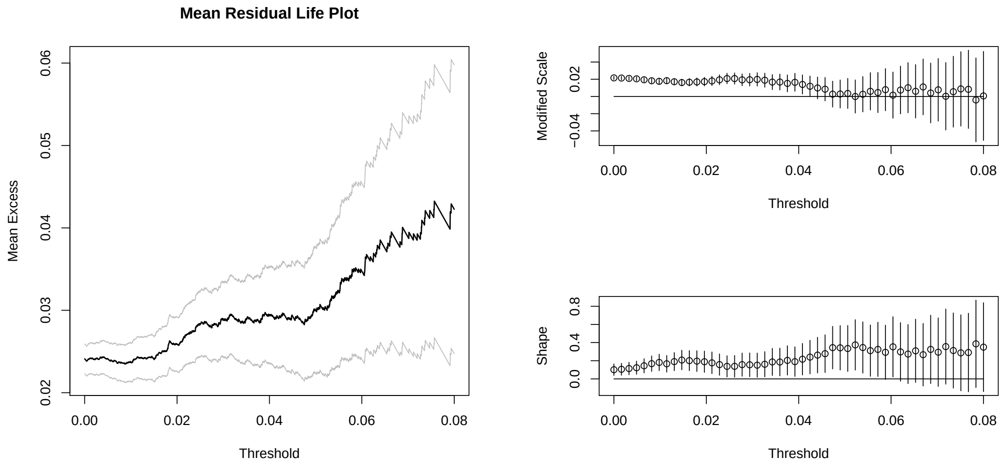


Figure 47: The autocorrelation function.

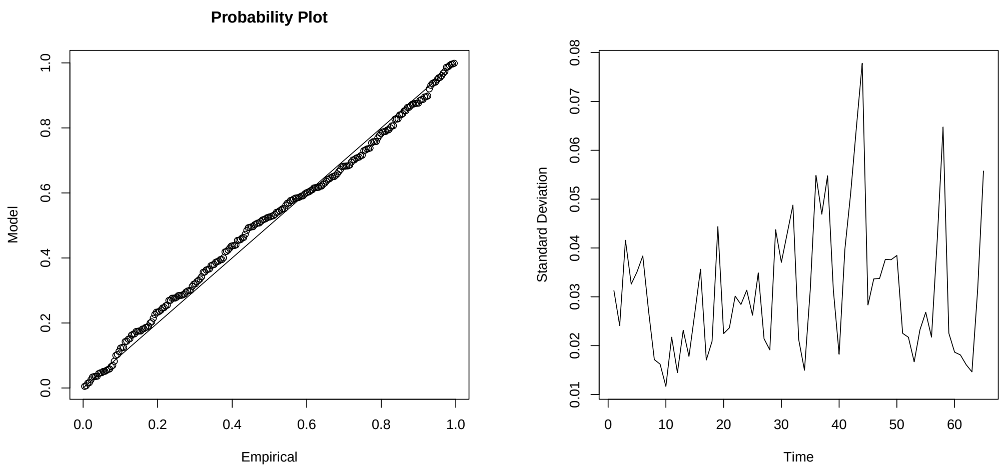
seen in figure 50(b) is almost exactly the same as the one for the log returns with the POT fitted GPD predicting a slightly heavier tail than the extrapolated ACER function. For all the other data sets, we have seen that the confidence interval is larger for the residuals than the ones for the log returns, but for in this case the confidence intervals looks quite alike. The parameters of the ACER functions are $\tilde{a} = -5.1112$, $b = 0$, $c = 0.9780$, $\gamma = -9.2806$ and $q = 0.5140$. The return level is 0.618314 with the 95% confidence intervals CI $= [0.401198, 0.847774]$ for the 10^{-6} level of interest.



(a) The mean residual plot.

(b) The estimated modified shape and scale parameter against the threshold.

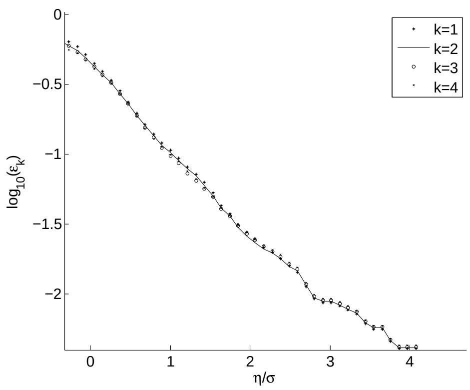
Figure 48: The plots to help on the choice of threshold.



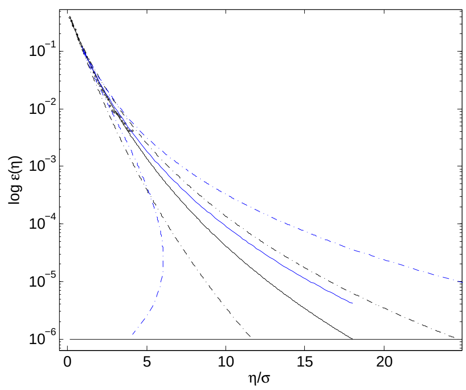
(a) The probability plot.

(b) The development of the deviation with respect to time (months).

Figure 49: Probability plot for threshold excess model fitted to the residuals of the EL NP data and checking for stationarity.



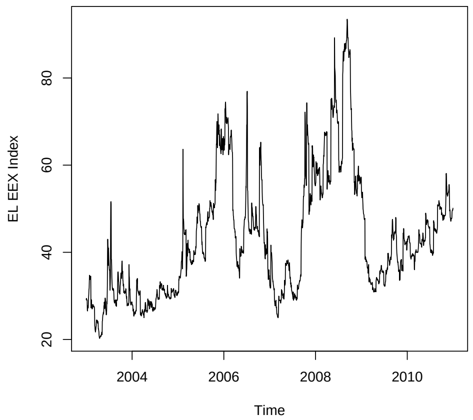
(a) The empirical ACER functions plotted against scaled exceedances.



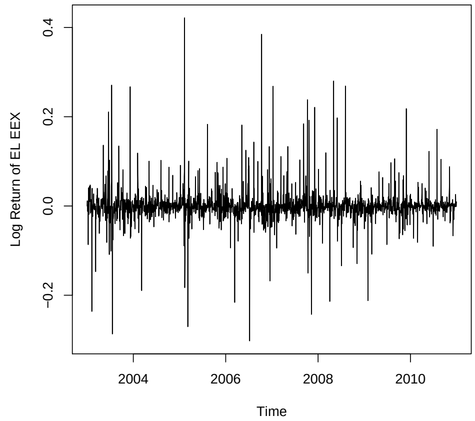
(b) The extrapolated ACER (black) function and POT (blue) fitted GPD with their confidence intervals respectively.

Figure 50: The residuals of the EL NP data analysis by the POT and ACER methods.

3.6 EL EEX



(a) The index of EL EEX from period 23.01.2003-31.01.2011.



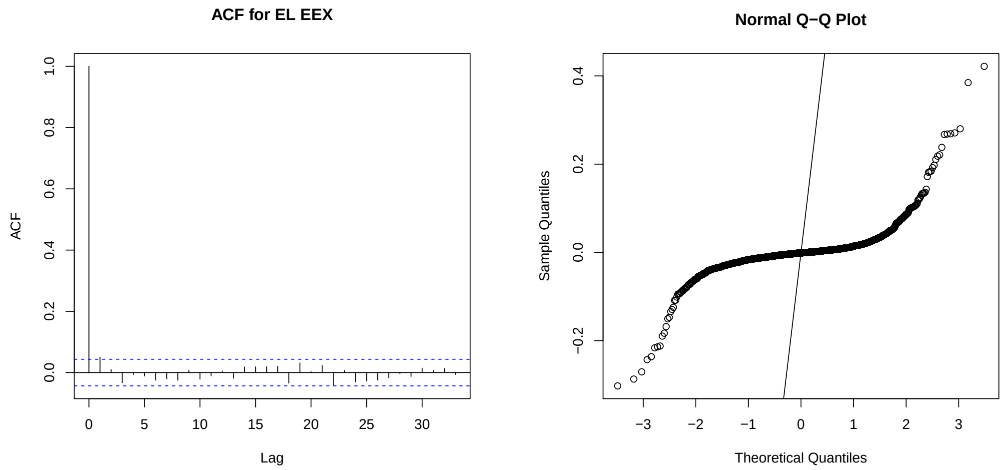
(b) The log return of the EL ICE observations.

Figure 51: Index plot and log return plot of the observations for EL EEX

The EL EEX data set is monthly futures from Germany and spans over approximately 8 years with 2033 observations. The index plot for the EL EEX data set can be seen in figure 51(a). We see that there are some fluctuations and great variations in the data set. Thus the log return plot in figure 51(b) has some larger values than the rest. The autocorrelation function seen in figure 52(a) shows extremely small autocorrelation values for the EL EEX data set. This data set has the excess kurtosis of 28.53 and the skewness of 1.59, and in figure 52(b) can we see the difference in this data set compared to the normal distribution.

Now we find a suitable threshold value by the POT method. We first consider at the mean residual plot in figure 53(a) and look for areas for where the plot is linear. It seems to consist of two linear parts; one from 0 to 0.05 and the other from 0.05 and to the end of the plot in this figure. For the plot with the estimated modified scale parameter and the estimated shape parameter, we see areas where the parameters are constant around 0-0.02 for the first mentioned and 0.02-0.05 for the last mentioned. The threshold value of 0.02 seem to fit in this case. This threshold value gives 228 observations above, which represents 11.21% of the total amount of observations. This leads to an estimated modified scale parameter 0.02636(0.00215) and an estimated shape parameter 0.43557(0.10401) with their respectively standard errors in parenthesis. The probability plot seen in figure 54(a) gives us the indication that the fitted model is pretty good. For the stationarity check, we see in figure 54(b) that the EL EEX data set is quite stationary.

The plots for the ACER analysis can be seen in the figures 55(a) and 55(b). We see that the ACER functions plotted against the scaled exceedances shows that the first four ACER functions converge, $k = 1, 2, 3, 4$, and thus we use the first ACER function in the further analysis. We see that the POT fitted GPD predicts a heavier tail than the extrapolated ACER function. The parameters of the ACER function are $\tilde{a} = -1.0105$, $b = 0.0036$, $c = 0.5$, $\gamma = -12.4279$ and $q = 0.5178$. The return level is 3.47428 with the 95% confidence intervals CI = [1.7167,5.35344] for the 10^{-6} level of interest.



(a) The autocorrelation function.

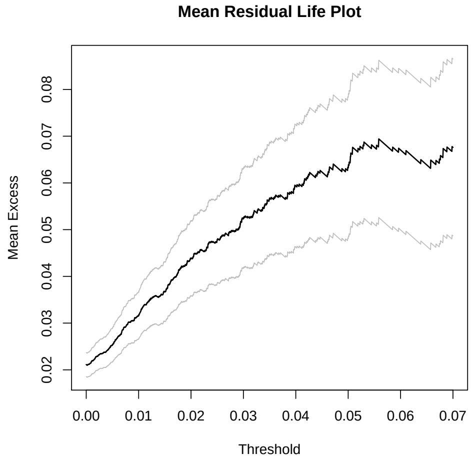
(b) QQ-plot with the normal distribution represented as the line.

Figure 52: The autocorrelation function and the QQ-plot.

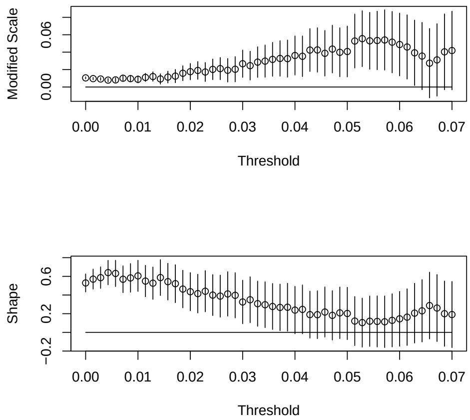
Doing the filtration as before with an AR-GARCH filter, we see that the residuals in figure 56(a). In this case we used an AR(1)-GARCH(1,1) filter, and the student t-distribution as the conditioning distribution. We see from table 6 that every parameter is then significant down to at least a 0.01 level. Note that the parameter v is the degrees of freedom in the Student's t distribution with mean equal 0 and variance equal 1.

In figure 56(b), we see the difference between the standard residuals and the normal distribution. Additional information are the excess kurtosis of 28.38 and the skewness of 1.54. Even though these values are quite similar to the values for the original unfiltered data set, but we can have two very different distributions despite the similar excess kurtosis and skewness value as mentioned before. The autocorrelation function in figure 57(a) shows very small values, even in comparison to the other data sets analysed after doing the filtration.

The threshold value, u, can be decided by looking at the mean residual plot in figure 58(a) and the plot of the estimated modified scale and shape parameter in figure 58(b). The mean residual plot has two areas of linearity, from 0 to 0.03 and from 0.03 to 0.06. The estimated modified scale is approximately constant from 0 to 0.02 and 0.02 to 0.04

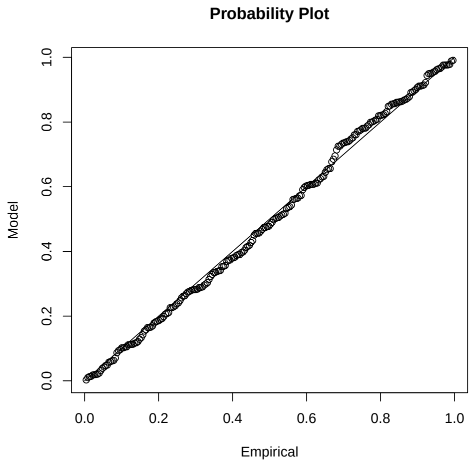


(a) The mean residual plot.

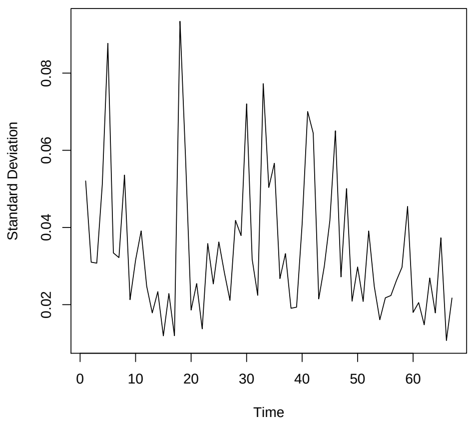


(b) The estimated modified shape and scale parameter against the threshold.

Figure 53: The plots to help on the choice of threshold.



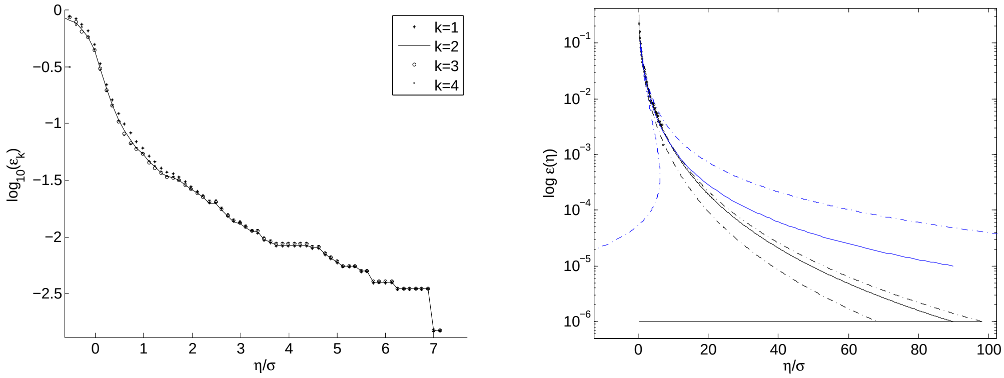
(a) The probability plot.



(b) The development of the standard deviation with respect to time (months).

Figure 54: Probability plot for threshold excess model fitted to EL EEX data and checking for stationarity.

while the estimated shape parameter seems to be constant from 0.02 to 0.03 and from 0.03 to 0.05. This leads to the choice of $u = 0.02$ as the threshold value. This gives 243 observations above the threshold and that represents 11.95% of the total amount of observations. This choice gives an estimated modified scale parameter of 0.02393(0.002821) and an estimated shape parameter of 0.47373(0.103904) with their respective standard



(a) The empirical ACER functions plotted against scaled exceedances.

(b) The extrapolated ACER (black) function and POT (blue) fitted GPD with their confidence intervals respectively.

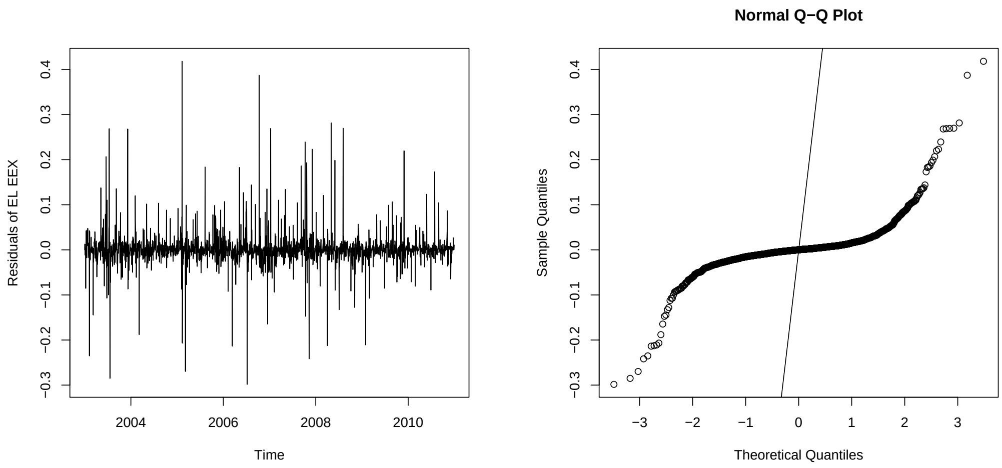
Figure 55: The EL EEX data analysis by the POT and ACER methods.

errors in parenthesis. Performing a quality check on the model, we see that the probability plot in figure 59(a) indicates a good fit. The stationarity of the standard residuals can be checked in figure 59(b), and we clearly see trends in the change of the standard deviation and that can indicate that the residuals are stationary.

AR-GARCH parameters	Estimates (std.error)
a_0	-0.0011688 (0.0003019)***
a_1	0.0594403 (0.0175041)***
ϕ_0	0.0010863 (0.0003829)**
ϕ_1	1.0000000 (0.3467539)**
ϕ_2	0.5579110 (0.0542804)***
υ	2.0673956 (0.0242457)***

Table 6: Estimates of the AR-GARCH parameters for the EL EEX data set with the Student's t distribution as the conditional distribution with the respectively standard errors in parenthesis. Signif. codes: ***=0.001, **=0.01, *=0.05, .=0.1 and ..=1

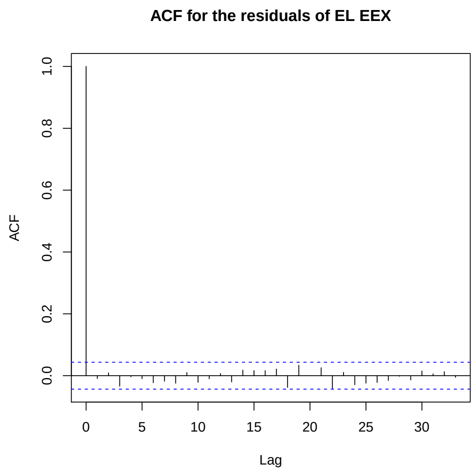
Continuing with the ACER analysis, we see from figure 60(a) that the first four ACER function fully converge all the way so we only use the first ACER function to estimate the parameters in the analysis. From figure 60(b) we see that now the POT fitted GPD is estimating a heavier tail than the extrapolated ACER function, and with very much larger confidence intervals as the previous plot for the log returns. The parameters of the ACER functions are $\tilde{a} = -1.1600$, $b = 0.0044$, $c = 0.5000$, $\gamma = -11.2456$ and $q = 0.5495$. The return level is 3.73041 with the 95% confidence intervals CI = [1.90838, 5.49226] for the 10^{-6} level of interest.



(a) The log return of the residuals of the EL EEX observations.

(b) QQ-plot with the normal distribution represented as the line.

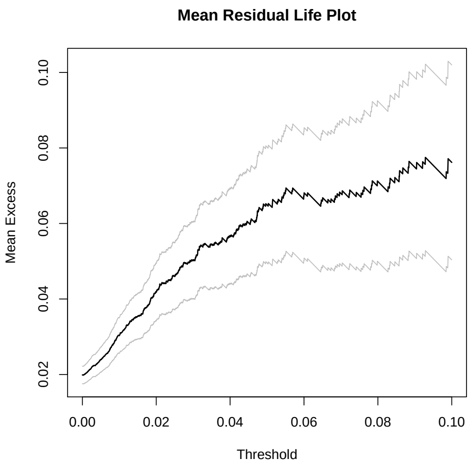
Figure 56: The log return plot of the residuals and the QQ-plot for the residuals of the EL EEX data set.



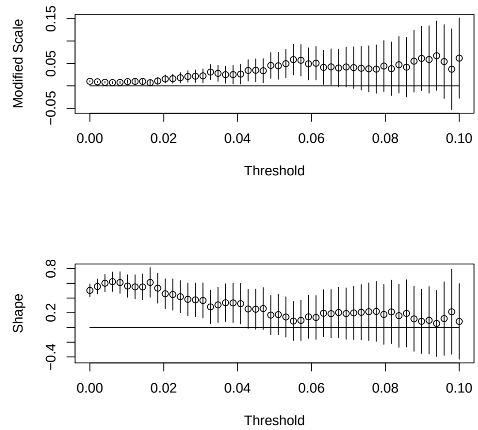
(a)

Figure 57: The autocorrelation function.

Last, but not least, we have an alternative way of viewing the in-sample quantiles we get from after filtration. In figure 61 we see the in-sample quantile plot on the log returns of EL EEX with quantiles from the POT method. The 95%, 99%, 99.5% and 99.9% quantiles are made from the standard residuals after the AR-GARCH filtration. We can see that the different colours denoting the quantiles gives us the impression of where the standard residuals predicts 95 percent of the observations should be within and so

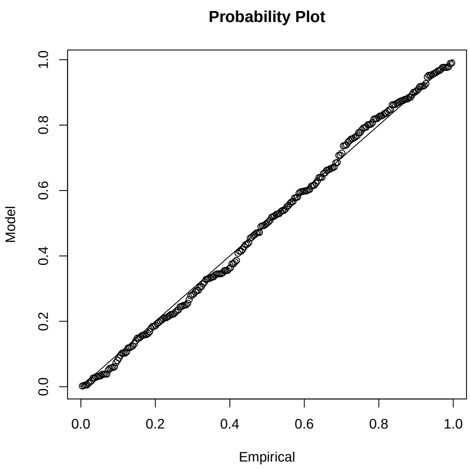


(a) The mean residual plot.

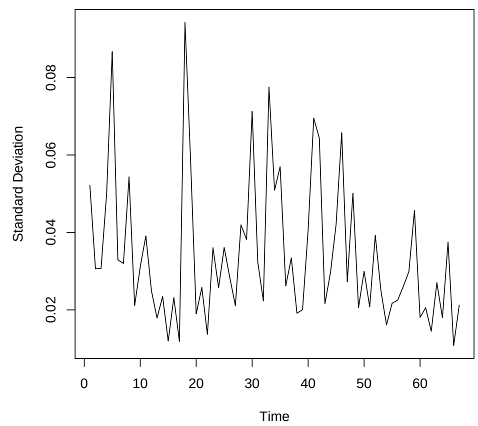


(b) The estimated modified shape and scale parameter against the threshold.

Figure 58: The plots to help on the choice of threshold.



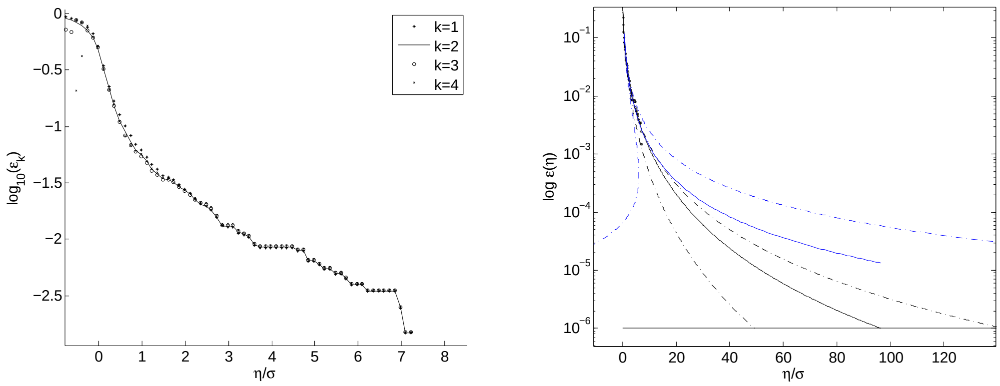
(a) The probability plot.



(b) The development of the standard deviation with respect to time (months).

Figure 59: Probability plot for threshold excess model fitted to the residuals of the EL EEX data and checking for stationarity.

forth for the other percentiles. We notice that the green 99.5% quantile and the yellow 99.9% quantile are very close to each other, so it is very hard to seperate them from one another. This is also done for the quantiles from the ACER method and this gives us a plot that resembles the one with the quantiles from the POT method with marginal differences. The quantile plot is an illustrative way of showing the quantiles, but it is



(a) The empirical ACER functions plotted against scaled exceedances.

(b) The extrapolated ACER (black) function and POT (blue) fitted GPD with their confidence intervals respectively.

Figure 60: The residuals of the EL EEX data analysis by the POT and ACER methods.

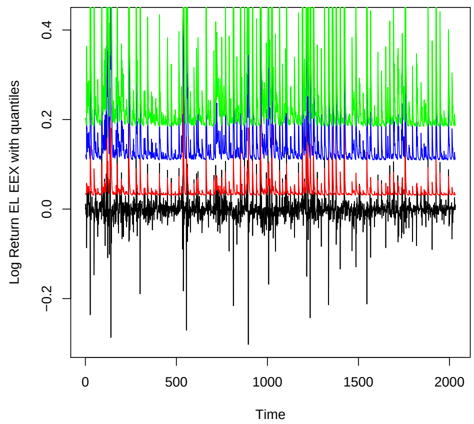
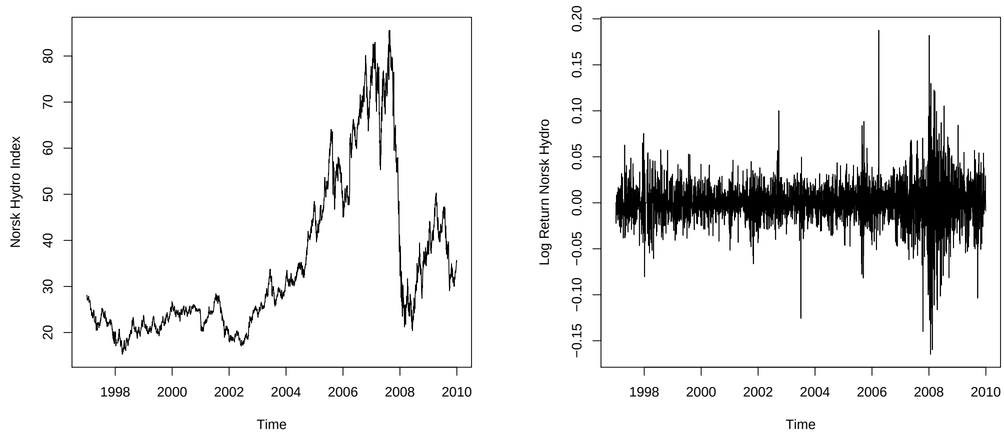


Figure 61: The in-sample quantile plot on the log returns of EL EEX with quantiles from the POT method. Red: 95%. Blue: 99%. Green: 99.5%. Yellow: 99.9%.

not very informative so thus we just add this example on the data set EL EEX as a demonstration.

3.7 Norsk Hydro

To compare the data sets we have from the different energy markets in form of electricity and coal data, we can analyse another kind of a market. Norsk Hydro is a Norwegian company working with aluminium and renewable energy. We are looking at the value of the company and therefore this represents a different kind of a market than the energy market (even though Norsk Hydro is operating within the field of energy production). We have observations from 1997 to 2010 for Norsk Hydro, consisting of 3276 observations. This is definitely the data set with the most observations. From the plot of the index of Norsk Hydro in figure 62(a), we see that the company has been experiencing a steady upwards growth in the index since about 2003 and a heavy steep downwards fall in the last years. This takes form of large values in the log return plots in figure 62(b).



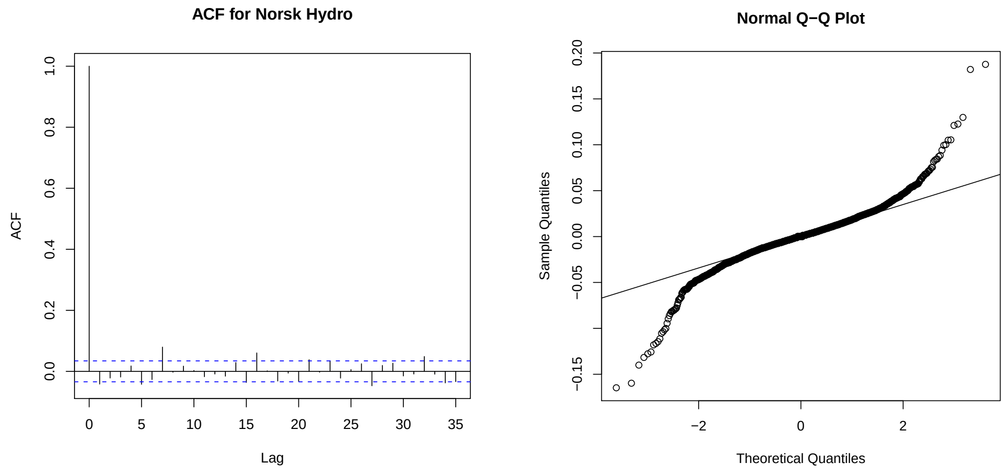
(a) The index of Norsk Hydro from period 17.09.1997-29.09.2010.

(b) The log return of Norsk Hydro observations.

Figure 62: Index plot and log return plot of the observations for the Norsk Hydro data.

It can be seen from the autocorrelation plot 63(a) that there should be no concern about linear dependence in the time series, but we do see that for some very few observations the autocorrelation is marginally significant different from zero. We can also see from the normal quantile plot 63(b) that the tails of the observations differs somewhat in comparison with the normal distribution. This can also be confirmed by the excess kurtosis of 7.43. This value is not large in comparison to some of the energy market data sets. The skewness is -0.13 and is closely to the normal distribution.

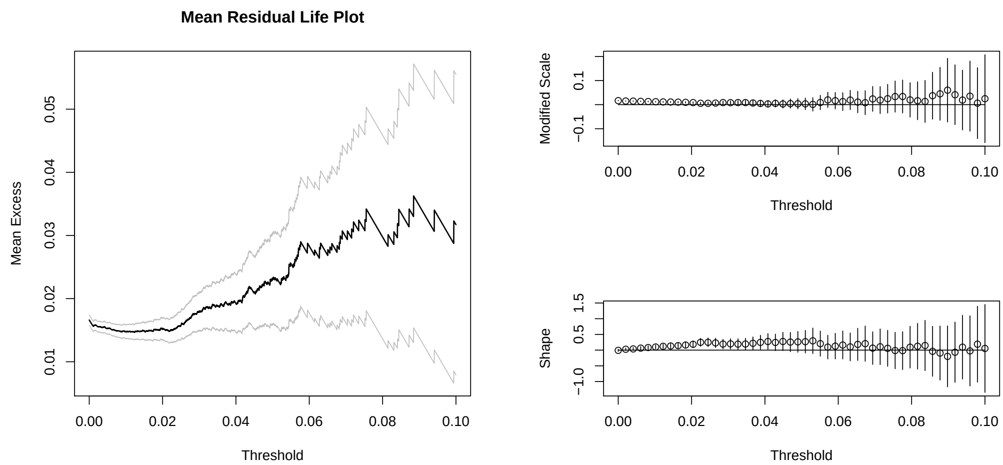
To decide the threshold value, we look at the mean residual plot 64(a) and look for parts where there is linearity in the mean excess. We see that there is an approximately linear part 0-0.02 and from 0.02-0.05. From the plot with the estimated scale and shape parameter in figure 64(b), the estimated shape parameter is most constant when the threshold is 0.02-0.05. We want to choose an u that gives us most observations as pos-



(a) The autocorrelation function for Norsk Hydro.

(b) The observations against the normal QQ-plot.

Figure 63: The autocorrelation function and the normal quantile plot for Norsk Hydro.

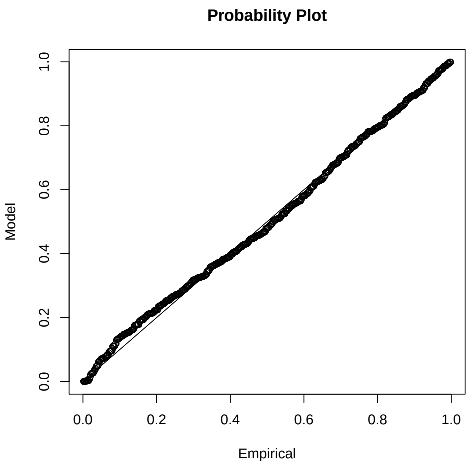


(a) The mean residual plot for Norsk Hydro.

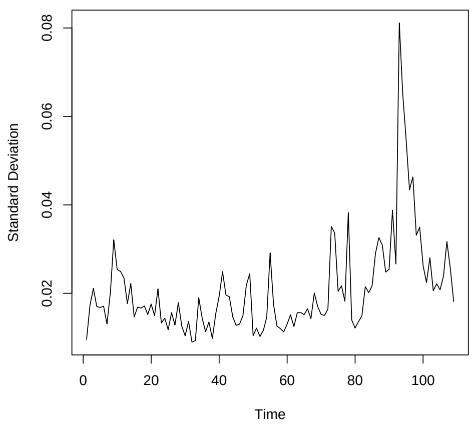
(b) The estimated shape and scale parameter against the threshold.

Figure 64: The plots to help decide the threshold value.

sible, while at the same time the mean residual plot being approximately linear and the estimated shape parameter being approximately constant in this threshold value. This can be achieved with $u = 0.02$, and we have 468 of 3276 observations above this threshold. We must mention that the number of observations above the threshold decline rapidly as the threshold increases by 0.01 each time. This implies that only few of the observations that can be regarded as extreme, and the lest is pretty 'normal' in comparison. This is also confirmed by the low excess kurtosis number. From the probability plot 65(a) we see that the linearity is pretty good in this case, but this does not give real cause for concern about the quality of the model. The standard deviation plotted against time (months) in figure 65(b) show the values of the standard deviation to be quite stable except for the very end of the period. This can indicate non-stationarity, but one can ask oneself how big of a deal will the non-stationarity towards the very end of the period have to say in the larger perspective.



(a) The probability plot.



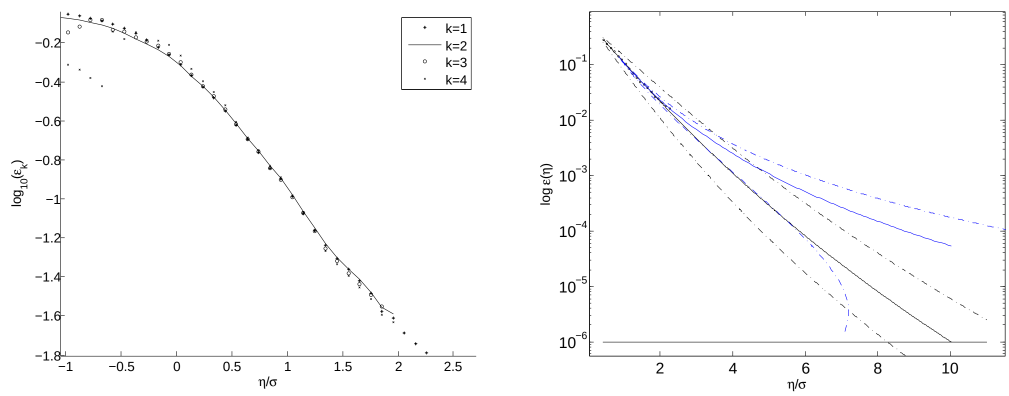
(b) The development of the standard deviation with respect to time (months).

Figure 65: Probability plot for threshold excess model fitted to the Norsk Hydro data and checking for stationarity.

For this data set we have used the first ACER function to estimate the parameters in equation (46). From figure 66(b) we see that the POT fitted GPD still predict a heavier tail than the extrapolated ACER function. This can be caused by the fact the POT method assumes stationarity, while the observations can be of the non-stationarity kind. The much heavier tail predicted by the POT method can also come of the fact that the ACER method puts weights on the observations, while the POT method does no such thing. One other aspect is that the line of the graph looks almost linear, and therefore we must really question our decision of the data being Fréchet rather being Gumbel where $\xi = 0$. The ACER parameters are $\tilde{a} = -4.9909$, $b = 0.0053$, $c = 1.0503$, $\gamma = -17.8025$ and $q = 0.3962$. The return level is 0.234155 with the 95% confidence intervals CI = [0.212583, 0.257145] for the 10^{-6} level of interest.

Now we do the AR-GARCH filtration for the Norsk Hydro data set. After careful deliberation, we choose an AR(5) part and a GARCH(1,1) part. The conditional distribution is set to the Student's t distribution. From table 7 we see that all the parameters are significant down to a 0.05 level except for a_1, a_2 and a_4. Note that the parameter υ is the degree of freedom in the Student's t distribution with mean equal 0 and variance equal 1.

In figure 67(a) we see the residuals of the log returns of the data set Norsk Hydro. The residuals have much heavier tails than the normal distribution, seen in figure 67(b), as expected. The excess kurtosis is 7.62 and the skewness is -0.19. The autocorrelation function in figure 68(a) shows no validated concern of autocorrelation in the standard residuals.



(a) The empirical ACER functions plotted against scaled exceedances.

(b) The extrapolated ACER (black) function and POT (blue) fitted GPD with their confidence intervals respectively.

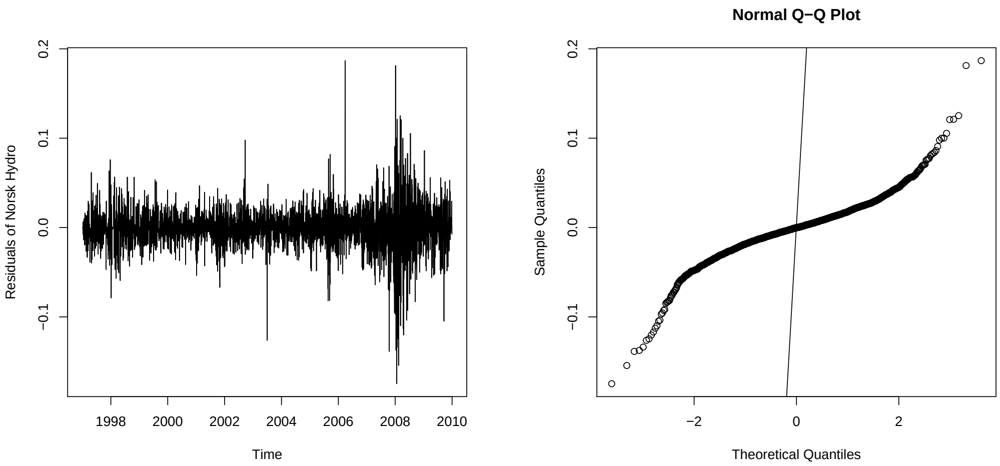
Figure 66: The Norsk Hydro data analysed with ACER plotted against the POT analysis.

To decide upon the choice of threshold, we look at the mean residual plot in figure 69(a) and the plot of the estimated modified scale and shape parameter against the threshold in figure 69(b). In the mean residual plot there is a pretty large area of linearity in the ares 0.01-0.02 and 0.02-0.045. The estimated modified scale parameter and the estimated shape parameter is approximately constant around areas 0.01-0.02 and 0.025-0.035. After some delibraton and further investigation of stability, the threshold value is chosen to be $u = 0.02$ with 450 observations above the threshold (13.74%). This gives us the estimated modified scale parameter 0.01188(0.0008402) and the estimated shape parameter 0.19986(0.0546044) with their respectively standard deviation in parenthesis.

Doing the same quality check of the model as before, we see in figure 70(a) the probability plot. It looks pretty good aligned with the linear line. From figure 70(b) we see that the standard residuals are stationary except for a large spike around the 90th month. With the log returns being quite stationary, it is expected that the standard residuals will be more stationary than before filtration.

AR-GARCH parameters	Estimates (std.error)
a_0	6.752e-04 (2.919e-04)*
a_1	1.446e-02 (1.789e-02)..
a_2	-5.057e-03 (1.786e-02)..
a_3	-6.452e-02 (1.774e-02)***
a_4	6.822e-03 (1.763e-02)..
a_5	-3.475e-02 (1.755e-02)*
ϕ_0	5.996e-06 (1.736e-06)***
ϕ_1	7.924e-02 (1.157e-02)***
ϕ_2	9.094e-01 (1.271e-02)***
υ	7.745e+00 (9.394e-01)***

Table 7: Estimates of the AR-GARCH parameters for the Norsk Hydro data set with the Student's t distribution as the conditional distribution with the respectively standard errors in parenthesis. Signif. codes: ***=0.001, **=0.01, *=0.05, .=0.1 and ..=1

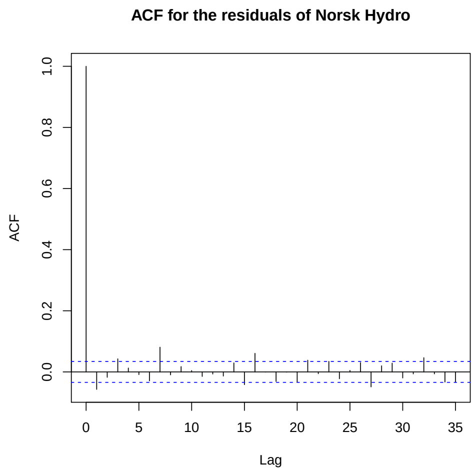


(a) The log return of the residuals of the Norsk Hydro observations.

(b) QQ-plot with the normal distribution represented as the line.

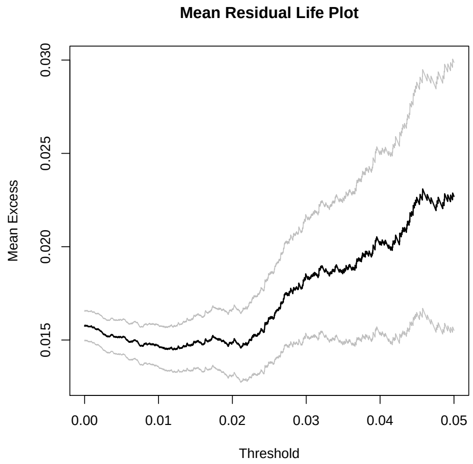
Figure 67: The log return plot of the observations and the QQ-plot for the residuals of the Norsk Hydro data set.

Continuing with the ACER analysis, we see from figure 71(a) that the four first ACER function fully converge so we only use the first ACER function to estimate the parameters in the analysis. From figure 71(b) we see that the POT fitted GPD is estimating a heavier tail than the extrapolated ACER function. This figure looks very much the same as the one for the log returns, except for a small change in where the functions intersect with the x-axis. The parameters of the ACER functions are $\tilde{a} = -5.0320$, $b = 0.0022$, $c = 1.0952$,

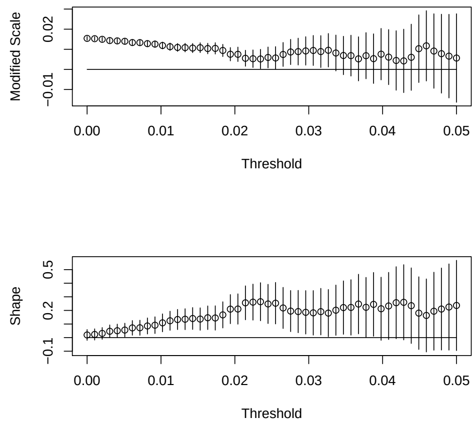


(a)

Figure 68: The autocorrelation function.



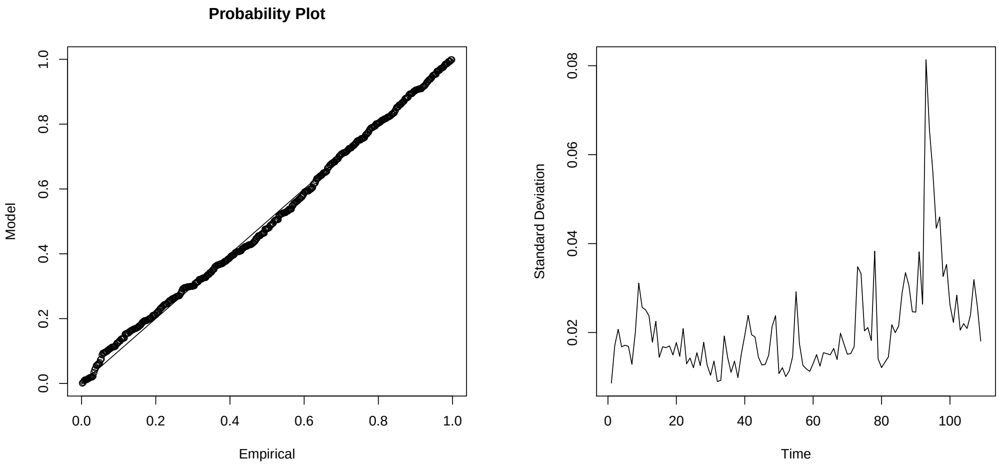
(a) The mean residual plot.



(b) The estimated modified shape and scale parameter against the threshold.

Figure 69: The plots to help on the choice of threshold.

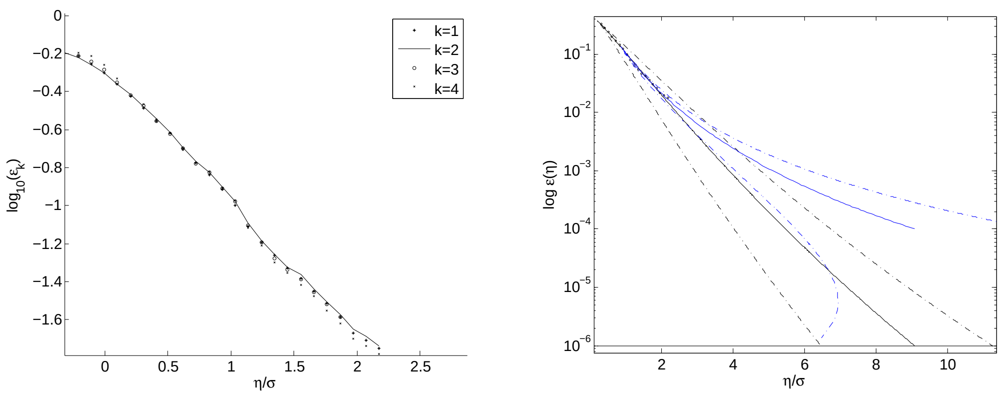
$\gamma = -20.1581$ and $q = 0.4435$. The return level is 0.212439 with the 95% confidence intervals CI = [0.151133, 0.263334] for the 10^{-6} level of interest.



(a) The probability plot.

(b) The development of the standard deviation with respect to time (months).

Figure 70: Diagnostic plots for threshold excess model fitted to the residuals of the Norsk Hydro data and checking for stationarity.



(a) The empirical ACER functions plotted against scaled exceedances.

(b) The extrapolated ACER (black) function and POT (blue) fitted GPD with their confidence intervals respectively.

Figure 71: The residuals of the Norsk Hydro data analysis by the POT and ACER methods.

4 Comparison and further discussion

In this section, we are going to put the performance of the individual data sets into perspective and compare them in form of in-sample and out-of-sample predictions.

There are really three questions that need to be answered. The first question of answering is: How is the energy market different from other markets? Here we have made the same analysis on a data set from Norsk Hydro to compare with the ones from the energy market. The second is whether the removal of the dates of expiration have something to say for the outcome of the precision of the predictions. As mentioned before, we have only got hold of the dates of expiration for two of the data sets; the EL ICE and the Coal ICE data set. That should be enough to give us an impression of how large a role the dates of expiration have for the predictions. The last, but not least, question when using the POT and ACER method for analysing: Is the one better than the other in producing better fits and predictions? Further on, we would like to emphasise the main difference between the two methods.

Before beginning the comparison for the data sets, we can say something generally about the weaknesses and strengths of the POT and ACER method.

A weakness in the POT method is the difficulties during the process of finding a suitable and well performing threshold value u. It can be difficult to look for linearity and areas where the threshold is constant in the mean residual plot and the estimated scale and shape against the threshold plot. If the plots are enlarged, we can see more irregularities and thus making us doubt the previous choice of the threshold value before the enlargement. If we compress the plots, the curves can appear to be smoother than they really are. So the arbitrariness of the choice of the threshold interval to be plotted can make a big difference in the decision of the thresholds, and a badly suited threshold value can lead to ad hoc results.

Another important weakness in the POT method is that it assumes asymptotic characteristics in the data. It can be very hard to verify these assumptions. How large must the threshold and the size of the data set be to satisfy the asymptotic assumptions? It is extremely hard to say something specific about this. The POT method assumes not only asymptotic characteristics, but it also assumes independence and stationarity among the data. We can try to detect stationarity in the data sets by plotting the standard deviation plotted against time as seen in the last section. This problem can be dealt by using an AR-GARCH model which can pick up the heteroskedasticity, similar to what Byström had done in his report [3]. The ACER method catches the phases better than the POT method, so the ACER method will not be so sensitive to stationarity. Although the theory behind the peaks over threshold method is easily understood and executed in real life, these weaknesses play a large role in the process of analysing data. The POT method has also been explored more than the ACER method, and thus giving us more knowledge and empirical research about this method.

The weakness in the ACER method is the fact that we have to assume a certain tail behaviour that are not necessary correctly assumed. The theory behind the ACER method is more complicated than the POT method, and thus more difficult and time demanding to execute in real life. The ACER method does make a much better use of the given data during the analysis than the POT method by the fact that this class of functions is built in a way that manages to capture the sub asymptotic behaviour of the extreme value distribution to some extent, and thus providing us with more accurate predictions in theory. We see it in the plots of the extrapolated ACER function and POT fitted GPD with their confidence intervals respectively in the previous section that the confidence intervals of the ACER function is much smaller than the confidence intervals of the POT fitted GPD.

For the POT method, it is also observed that the confidence intervals increase as the threshold increases. This is quite logical because as the threshold increases, the variance will also increase due to the fewer exceedance observations. Thus leading to a larger confidence interval. It is also observed that with increasing threshold value, the confidence interval also increases whereas the return level is stable.

Dataset	Excess Kurtosis	Skewness	u
EL ICE	18.45	1.85	0.03
EL ICE reduced	10.87	1.27	0.03
EL COAL	30.37	-1.26	0.015
EL COAL reduced	32.56	0.76	0.01
EL NP	9.48	0.89	0.03
EL EEX	28.53	1.59	0.02
Norsk Hydro	7.43	-0.13	0.02

Table 8: A summary of the excess kurtosis, skewness and the threshold selection by the POT method for the different data sets BEFORE filtration.

Dataset	Excess Kurtosis	Skewness	u
EL ICE res.	18.04	1.79	0.02
EL ICE reduced res.	10.70	1.19	0.02
EL COAL res.	30.72	-1.15	0.01
EL COAL reduced res.	24.70	-0.14	0.01
EL NP res.	9.48	0.89	0.03
EL EEX res.	28.38	1.54	0.02
Norsk Hydro res.	7.62	-0.19	0.02

Table 9: A summary of the excess kurtosis, skewness and the threshold selection by the POT method for the different data sets AFTER filtration.

To answer our first question, we can take look at table 8 and 9. We see that for both the log returns and the residuals after filtration it is the data sets from the energy market who has the largest excess kurtosis and skewness in comparison to the data set from Norsk Hydro, even though for some of the data sets from the energy market the difference is not very large. While for the index of the company Norsk Hydro may change for larger influences, for example the company going public. We have also noticed when looking at the plot with the extrapolated ACER function and the POT fitted GPD, the data sets from the energy market estimates a much heavier tail than the data set from Norsk Hydro.

We have also done Ljung-Box test on the return series to test for the serial correlation before and after filtration, and on the squared returns to test for GARCH effects. This is denoted by $Q(.)$ and $Q^2(.)$ respectively. We have chosen $Q(1), Q(3), Q(5)$ and $Q(8)$ in the comparison. In table 10 and 11 shows the Ljung-Box test on the log return series and on the squared returns, and we see that almost every statistic shows significance. In table 12 and 13 we see that almost all of the statistics have lost their significance after filtration. One exception is the Ljung-Box statistics on the residuals of the EL NP data set. This can be caused by the fact that we used an AR(0) part in the filtration, but then again we have a data set which refutes this hypothesis since we also used an AR(0) part on the Coal ICE data set and all of it's statistics are non-significant after filtration. These results shows that most of the serial correlation and GARCH effects disappears after doing an AR-GARCH filtration.

Dataset	$Q(1)$	$Q(3)$	$Q(5)$	$Q(8)$
EL ICE	3.3761	10.6645*	11.659*	13.9176
EL ICE reduced	11.0139*	13.9404*	16.3573*	17.6996*
COAL ICE	2.7675	6.0286	13.7156*	15.7633*
COAL ICE reduced	2.524	3.8899*	13.8877*	14.2276*
EL NP	8.9842*	12.454*	21.3321*	24.4064*
EL EEX	5.1827*	7.6993	8.0396	11.3968
Norsk Hydro	5.6995*	8.411*	15.3449*	38.5298*

Table 10: Ljung-Box statistics on the log return series. The star, *, denotes significance of level 0.05.

Dataset	$Q^2(1)$	$Q^2(3)$	$Q^2(5)$	$Q^2(8)$
EL ICE	5.5237*	7.5167	9.3241	10.6607
EL ICE reduced	27.2289*	55.341*	61.5301*	72.5855*
COAL ICE	40.4645*	41.0475*	46.5356*	48.4927*
COAL ICE reduced	5.3423*	10.8524*	23.1096*	46.7419*
EL NP	1.6045	2.7321	36.1181*	42.5802*
EL EEX	6.9813*	8.4339*	12.9718*	13.4546
Norsk Hydro	129.8152*	489.0329*	835.6988*	1258.251*

Table 11: Ljung-Box statistics on the log return series. The star, *, denotes significance of level 0.05.

Dataset	$Q(1)$	$Q(3)$	$Q(5)$	$Q(8)$
EL ICE res.	0.0778	0.157	0.895	6.0297
EL ICE reduced res.	0.1661	0.3645	2.8595	4.152
EL COAL res.	2.524	3.8899	13.8877*	14.2276
EL COAL reduced res.	1.1035	1.5822	1.7303	2.0274
EL NP res.	7.7629*	16.943*	20.5361*	22.6323*
EL EEX res.	0.0774	0.4431	0.539	1.4331
Norsk Hydro res.	0.1791	1.5207	2.7518	3.7993

Table 12: Ljung-Box statistics on the residuals. The star, *, denotes significance of level 0.05.

Dataset	$Q^2(1)$	$Q^2(3)$	$Q^2(5)$	$Q^2(8)$
EL ICE res.	2.1872	2.9097	2.9554	3.2047
EL ICE reduced res.	0.0243	1.0583	2.0834	3.401
EL COAL res.	6.0737*	6.2836	6.347	6.6808
EL COAL reduced res.	0.0264	0.2427	0.6192	1.0057
EL NP res.	0.2866	2.2531	5.1283	7.0594
EL EEX res.	0.9638	2.6436	3.5905	5.639
Norsk Hydro res.	1.1926	4.439	5.1989	5.9658

Table 13: Ljung-Box statistics on the residuals. The star, *, denotes significance of level 0.05.

4.1 EL ICE

Let us first look at the data set EL ICE with the dates of expiration. For the in-sample evaluation of the estimated tail quantiles in table 14, we see that both the POT and ACER methods before and after filtration make good estimations in the lower probabilities 0.995 and 0.999. We also notice that the number of exceedances for the ACER method before and after filtration have not changed at all. This can imply that the ACER method works better than the POT method despite serial correlations and GARCH effects. But this can just only apply for this data set, so we have to be hesitant to say anything certain before looking at all of the data sets.

Probability	Expected	POT (deviation)	ACER (deviation)	POT res.	ACER res.
0.95	81	80(-1)	78(-3)	80(-1)	78(-3)
0.99	16	20(+4)	18(+2)	19(+3)	18(+2)
0.995	8	9(+1)	9(+1)	9(+1)	9(+1)
0.999	2	1(-1)	1(-1)	1(-1)	1(-1)

Table 14: In-sample evaluation (number of exceedances) of estimated tail quantiles at different probabilities for the FULL EL ICE data set and the residuals of the EL ICE data set with the number of deviations from the expected tail quantiles respectively in brackets.

The in-sample evaluation done on the reduced of the EL ICE data set can be seen in table 15. Without the dates of expiration, the number of exceedances for the reduced EL ICE data set seem to be marginally better estimated by the ACER method than the POT method. This applies both for the log return series and the residuals.

Probability	Expected	POT (deviation)	ACER (deviation)	POT res.	ACER res.
0.95	76	77(+1)	77(+1)	80(+4)	80(+4)
0.99	15	19(+4)	17(+2)	18(+3)	15(0)
0.995	8	10(+2)	8(0)	9(+1)	9(+1)
0.999	2	2(0)	2(0)	2(0)	2(0)

Table 15: In-sample evaluation (number of exceedances) of estimated tail quantiles at different probabilities for the REDUCED EL ICE data set and the residuals of the EL ICE data set with the number of deviations from the expected tail quantiles respectively in brackets.

When looking at the return levels for the full EL ICE data set in table 16, we see that there are more changes in the return levels with the POT method than with the ACER method. This means that serial correlation and GARCH effects influence the POT method more than the ACER method. The ACER return levels for both the log returns and the residuals are slightly lower than the POT return levels. All these features mentioned for the full EL ICE data set applies also for the reduced EL ICE data set seen in table 17. It must also be mentioned that the return levels for the reduced EL ICE data set are smaller than the return levels for the full data set. For the 100 year return level, the one for the reduced EL ICE data set is around half of the ones for the full EL ICE data set. That makes sense because the residuals are smaller than the log returns in the first place.

Years	return level POT	return level ACER	return level POT2	return level ACER2
5	0.3216422	0.314226	0.3516183	0.314106
10	0.4235323	0.398456	0.4731911	0.400104
25	0.6043147	0.536854	0.6959477	0.542203
50	0.7871463	0.666254	0.9283848	0.675775
80	0.9400655	0.768326	1.127237	0.781536
100	1.022331	0.821322	1.235660	0.83657

Table 16: Out-of-sample return level table for the FULL EL ICE data set. Second and third column for the data set before filtration and fourth and fifth column for the data set after filtration (denoted by the suffix 2).

Years	return level POT	return level ACER	return level POT2	return level ACER2
5	0.2210260	0.227203	0.2221483	0.224997
10	0.2748316	0.279398	0.2773259	0.276838
25	0.3626046	0.361438	0.3677753	0.358306
50	0.4443755	0.434915	0.4524423	0.431258
80	0.5088281	0.491119	0.5194078	0.487051
100	0.5422981	0.519773	0.5542542	0.515493

Table 17: Out-of-sample return level table for the REDUCED EL ICE data set. Second and third column for the data set before filtration and fourth and fifth column for the data set after filtration (denoted by the suffix 2).

4.2 Coal ICE

When looking at the Coal ICE data set with it's dates of expiration, we notice that the in-sample evaluation in table 18 indicates that the POT method on the log returns does the best estimation of the number of exceedances. But overall we see that both method do a good job in estimating the tail quantile, especially at the 0.99 level.

Probability	Expected	POT	ACER	POT res.	ACER res.
0.95	58	57(-1)	57(-1)	57(-1)	57(-1)
0.99	12	12(0)	12(0)	12(0)	12(0)
0.995	6	6(0)	8(+2)	7(+1)	7(+1)
0.999	1	1(0)	2(+1)	1(0)	2(+1)

Table 18: In-sample evaluation (number of exceedances) of estimated tail quantiles at different probabilities for the FULL Coal ICE data set and the residuals of the Coal ICE data set with the number of deviations from the expected tail quantiles respectively in brackets.

The in-sample evaluation for the reduced Coal ICE data set in table 19 shows that none of the two methods stand out in making better estimations. Both methods do a good job in estimation the number of exceedances for the log returns and the residuals. The largest deviance from the expected number of exceedances is in the 0.95 level.

Probability	Expected	POT	ACER	POT res.	ACER res.
0.95	55	52(-3)	50(-5)	52(-3)	51(-4)
0.99	11	11(0)	9(-2)	10(-1)	11(0)
0.995	6	4(-2)	4(-2)	5(-1)	6(0)
0.999	1	1(0)	2(+1)	1(0)	2(+1)

Table 19: In-sample evaluation (number of exceedances) of estimated tail quantiles at different probabilities for the REDUCED Coal ICE data set and the residuals of the Coal ICE data set with the number of deviations from the expected tail quantiles respectively in brackets.

Looking at the return levels for the full Coal ICE data set in table 20, we see again that the ACER return levels for the log returns and the residuals are more alike than the POT return levels. Notice also that the ACER method predicts lower return levels than the POT method. This also applies for the reduced Coal ICE data set seen in table 21. The properties mentioned for the EL ICE data set in the previous section are also relevant for this data set.

Years	return level POT	return level ACER	return level POT2	return level ACER2
5	0.1962524	0.175571	0.2043003	0.183066
10	0.2579979	0.221191	0.2729264	0.232937
25	0.3674506	0.295265	0.3972244	0.31478
50	0.478044	0.363648	0.5254766	0.391105
80	0.5704833	0.417062	0.6343111	0.451155
100	0.620193	0.444628	0.693367	0.482276

Table 20: Out-of-sample return level table for the FULL Coal ICE data set. Second and third column for the data set before filtration and fourth and fifth column for the data set after filtration (denoted by the suffix 2).

Years	return level POT	return level ACER	return level POT2	return level ACER2
5	0.1294768	0.116613	0.1377790	0.107754
10	0.1640432	0.140667	0.1791255	0.129501
25	0.2222268	0.176621	0.2514717	0.162371
50	0.2781221	0.207136	0.3236521	0.190685
80	0.3231679	0.229535	0.3834346	0.211735
100	0.3468711	0.240669	0.4154083	0.222287

Table 21: Out-of-sample return level table for the REDUCED Coal ICE data set. Second and third column for the data set before filtration and fourth and fifth column for the data set after filtration (denoted by the suffix 2).

4.3 EL NP

For the EL NP data set, the in-sample evaluation of the estimated tail quantiles can be seen in table 22. We see that the estimations made are approximately equally good with either the POT method or the ACER method. Both methods predicts number of exceedances who are at most 3 away from the expected number.

We see in table 23 that in this case the POT return levels are similar whether it is for the log returns or the residuals. This characteristic has applied for the ACER return levels in the previous cases, but in this case it includes the POT return levels also. The POT method estimates a slightly heavier tail than the ACER method, as for almost all of the other data sets.

Probability	Expected	POT	ACER	POT res.	ACER res.
0.95	98	99(+1)	97(-1)	99(+1)	97(-1)
0.99	20	17(-3)	17(-3)	17(-3)	17(-3)
0.995	10	9(-1)	11(+1)	9(-1)	11(+1)
0.999	2	3(+1)	4(+2)	4(+2)	4(+2)

Table 22: In-sample evaluation (number of exceedances) of estimated tail quantiles at different probabilities for the EL NP data set and the residuals of the EL NP data set with the number of deviations from the expected tail quantiles respectively in brackets.

Years	return level POT	return level ACER	return level POT2	return level ACER2
5	0.2135263	0.193818	0.2111951	0.193234
10	0.2519824	0.224481	0.2484531	0.223902
25	0.3096101	0.268853	0.3038745	0.268284
50	0.3589922	0.305575	0.3510201	0.305017
80	0.395634	0.332142	0.3858182	0.331593
100	0.4139898	0.34525	0.4031956	0.344707

Table 23: Out-of-sample return level table for the EL NP data set. Second and third column for the data set before filtration and fourth and fifth column for the data set after filtration (denoted by the suffix 2).

It is widely known that significant in sample evidence of predictability does not guarantee significant out of sample predictability. But we conclude that results of in sample tests of predictability will typically be more credible than results of out of sample tests. In the tables for the return values, there are for mostly out-of-sample estimations. The out-of-sample predictions are much more insecure than the in-sample predictions, because they have to predict somewhere in the future with only a small observation sample as a basis. In table 24 we have used half of the data set EL NP to do the tail-quantile forecasting as explained in the theory part of this thesis. This way of doing out-of-sample predictions makes us capable of predicting quantiles for which we have real observations to compare with. We see that both the POT and ACER method does a good job in predicting the tail-quantiles, even though the POT method beats the ACER method marginally in the 0.99 and the 0.999 probability of prediction.

Note that the out-of-sample tail-quantile forecasting is done on the last three data sets; the EL NP, EL EEX and the Norsk Hydro data set. This is because the EL ICE and the Coal ICE data set are not well-behaved enough for doing the forecast.

Probability	Expected	prediction POT (deviation)	prediction ACER (deviation)
0.95	49	49(0)	49(0)
0.99	10	10(0)	8(-2)
0.995	5	6(+1)	6(+1)
0.999	1	1(0)	0(-1)

Table 24: Out-of-sample tail-quantile forecasting with EL NP.

4.4 EL EEX

For the in-sample evaluation with the data set EL EEX, we see in table 25 that neither of the two methods before or after filtration can be able to predict any closer to the expected number at the lowest level of 0.999. For the rest of the probability levels, both methods do a fairly good job in predicting the number of exceedances with the largest deviation of 2 exceedances.

For the return levels seen in table 26, we see as before that the ACER return levels for the log returns and the residuals are more similar than for the POT return levels. One other feature is that the POT method estimates a heavier tail than the ACER method. All this is familiar because of the previous data sets have the same features.

The out-of-sample tail-quantile seen in table 27 gives none of the two methods the benefit of being the best to estimate the number of exceedances. The ACER predictions are marginally better than the ones from the POT method, and it can be mentioned that the ACER predictions are spot on in the 0.99 and 0.995 probability level while the POT predictions are only one exceedance from the expected number at level 0.995. For the 0.95 probability level, we see that both methods do a fairly bad job at estimating the number of exceedance with a deviation of 11 with the POT method and 10 with the ACER method.

Probability	Expected	POT	ACER	POT res.	ACER res.
0.95	102	100(-2)	100(-2)	101(-1)	101(-1)
0.99	20	22(+2)	18(-2)	22(+2)	18(-2)
0.995	10	12(+2)	11(+1)	12(+2)	11(+1)
0.999	2	0(-2)	0(-2)	0(-2)	0(-2)

Table 25: In-sample evaluation (number of exceedances) of estimated tail quantiles at different probabilities for the EL EEX data set and the residuals of the EL EEX data set with the number of deviations from the expected tail quantiles respectively in brackets.

Years	return level POT	return level ACER	return level POT2	return level ACER2
5	0.4842896	0.456578	0.5149872	0.468373
10	0.6692529	0.594039	0.727018	0.61166
25	1.017388	0.818847	1.138758	0.848235
50	1.390237	1.02665	1.593245	1.06904
80	1.715273	1.18862	1.99819	1.24238
100	1.894493	1.272	2.224388	1.33201

Table 26: Out-of-sample return level table for the EL EEX data set. Second and third column for the data set before filtration and fourth and fifth column for the data set after filtration (denoted by the suffix 2).

Probability	Expected	prediction POT (deviation)	prediction ACER (deviation)
0.95	51	40(-11)	41(-10)
0.99	10	11(+1)	10(0)
0.995	5	6(+1)	5(0)
0.999	1	0(-1)	0(-1)

Table 27: Out-of-sample tail-quantile forecasting with EL EEX.

4.5 Norsk Hydro

Last, but not least, we look at the Norsk Hydro data set. This is the only one which is not from the energy market, and is used to look for eventual differences between the energy market and other markets. The Norsk Hydro data set is the one with the most observations, and therefore gives us the possibility to estimate the in-sample evaluation of the estimated tail quantiles down to even lower probabilities. The in-sample evaluation in table 28 shows that the POT method is clearly the better than the ACER method to estimate number of exceedances. This is the case with both the log returns and the residuals. We see that the largest deviation from the expected number of exceedances with the POT method is 4, while the largest deviation with the ACER method is 10. The largest deviances seem also to occur with the ACER method.

Probability	Expected	POT	ACER	POT res.	ACER res.
0.95	164	160(-4)	154(-10)	162(-2)	154(-10)
0.99	33	32(-1)	36(+3)	31(-2)	37(+4)
0.995	16	16(0)	26(+10)	18(+2)	24(+8)
0.999	3	5(+2)	10(+7)	5(+2)	9(+6)
0.9995	2	2(0)	7(+5)	2(0)	6(+4)
0.9999	0	0(0)	3(+3)	0(0)	2(+2)

Table 28: In-sample evaluation (number of exceedances) of estimated tail quantiles at different probabilities for the Norsk Hydro data set and the residuals of the Norsk Hydro data set with the number of deviations from the expected tail quantiles respectively in brackets.

We might have presumed that doing the AR-GARCH filtration will give better estimates with the POT method, because the residuals after filtration are more independent and stationary than the log returns. We have not always seen better estimates with the POT method after filtration though. On the contrary the changes have been better, worse or not changes at all in comparison to the log returns.

For the return levels seen in table 29 we see that many of the previously mentioned features also apply the Norsk Hydro data set, like the ACER return levels being more similar for the return levels of the residuals and the POT method predicting a heavier tail than the ACER method.

The out-of-sample tail-quantile seen in table 30, we see the estimates are not specially good in the two first probability levels for the POT method and the three first probability levels for the ACER method. The deviations are 5 to 8, and that represents a great deal since the number of exceedances is small too. For example makes the deviation of 6 up for almost half of the number of exceedance of 14. Rest of the estimates are either spot on or with a deviation of 1.

Years	return level POT	return level ACER	return level POT2	return level ACER2
5	0.1260668	0.0939669	0.1270485	0.0947087
10	0.1491534	0.104341	0.1517855	0.105617
25	0.1843875	0.118355	0.1902147	0.120451
50	0.2151299	0.129202	0.2243374	0.132003
80	0.2382403	0.136684	0.2503156	0.140008
100	0.2499078	0.140274	0.2635301	0.143858

Table 29: Out-of-sample return level table for the Norsk Hydro data set. Second and third column for the data set before filtration and fourth and fifth column for the data set after filtration (denoted by the suffix 2).

Probability	Expected	prediction POT (deviation)	prediction ACER (deviation)
0.95	82	90(+8)	89(+7)
0.99	16	21(+5)	23(+7)
0.995	8	9(+1)	14(+6)
0.999	2	1(-1)	2(0)

Table 30: Out-of-sample tail-quantile forecasting with EL EEX.

5 Conclusion and further work

5.1 Conclusion

First, we answer the question about whether the dates of expiration influence the results. It looks as though this does not have anything to say for the in-sample evaluation of the estimated tail-quantiles other than resulting in lower return level values. We must remember that we have only done this analysis with the data sets EL ICE and Coal ICE. After analysing several more data sets, then we can make more permanent and sure statements.

The data set from Norsk Hydro seems to be more well behaved than the other data sets from the energy market. This is supported by the lower excess kurtosis and skewness value, and the fact that the plot with the extrapolated ACER function and the POT fitted GPD shows a less heavier tails than the ones from the energy market.

For our last question, we must first consider the differences between the peaks over threshold (POT) method and average conditional exceedance rates (ACER) method. The main difference between the methods are that the POT method takes asymptotic assumptions in consideration during the analysis, and it also demands independence and stationarity in the data. While for the ACER method, we have to assume a certain shape/behaviour in the tail region. The length of the confidence intervals for the POT method is much larger than for the ACER method in the estimated return levels. For the POT method the length of the confidence intervals also increase with increasing threshold value. One other difficult aspect is the threshold selection when using the POT method. Different choice of threshold value can influence the estimated return values to a high degree, while altering the tail marker slightly in the ACER method does not greatly affect the estimated return levels.

One method that has been used to say something about the performance of the two methods is an in-sample evaluation of the estimated tail quantiles with different probabilities, and comparing these numbers to the expected tail quantiles. This has lead to a very ambiguous conclusion. Both the POT method and the ACER method have done a relatively good job in predicting the tail quantiles. An other method that have been used is an out-of-sample return level estimation. This has lead to the conclusion that the POT method often fit a heavier tail than the ACER method. The ACER return levels are also more similar before and after filtration, both with the log returns and the residuals. The last method that has been applied to the last three data sets, is an out-of-sample tail-quantile forecast. This shows that both methods perform equally well in predicting number of exceedances. So overall, there are no firm evidence that one method performs consistently better than the other.

5.2 Further work

Further work with this topic can be to make a program that ables us to take away certain parts in the AR-GARCH filter. This makes us able to do the same kind of analysis without so-called noise parts which can be non-significant but still influence the model. One other aspect that can be interesting is to make the out-of-sample tail-quantile estimates more than just one day into the future.

It would also be very interesting to follow the development of the analysis if we had even more observations. In this thesis the data set with the most observations is the data set for Norsk Hydro with it's 3276 days of observations. We can for example use a higher sampling frequency by using hourly observations instead of daily ones, or even more drastic; using tick data sets. This will lead to data sets with more dependence and correlation, and then maybe the POT method will be outperformed by the ACER method.

We could also analyse more data sets from the energy market, to see if there are any differences between coal, oil and electricity observations. Further on, we could have had more data sets from other markets besides the energy market to examine the differences in the markets even more.

References

[1] Anders Løland and Kjersti Aas. Statistisk analyse av energipriser trondheim, 2009.

[2] Walter Beckert. ADF handout, 2010. Available from: http://www.ems.bbk.ac.uk/for_students/bsc_FinEcon/fin_economEMEC007U/adf.pdf.

[3] Hans N. E. Byström. Extreme value theory and extremely large electricity price changes. *International Review of Economics and Finance*, 14(1):41 – 55, 2005.

[4] Stuart Coles. *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag London, 2001.

[5] Oleh Karpa. Acer method, user's guide, 2010.

[6] Alexander J. McNeil and Rüdiger Frey. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7(3-4):271 – 300, 2000.

[7] Arvid Naess. Estimation of extreme values of time series with heavy tails. (N2-2010), 2010.

[8] M. A. Ribatet. A User's Guide to the POT Package (version 1.0), August 2006. Available from: http://cran.r-project.org/.

A Appendix A: Code

```
vectorize = function(data){
newdata = as.numeric(data$close)
newdata = newdata[length(newdata):1]
return(newdata)
}

logreturn = function(data){
logdata = rep(0,length(data))
logdata[1] = 0
for(i in 2:length(data)){
logdata[i] = log(data[i]/data[i-1])
}
return(logdata)
}

stat = function(data){
x = floor(length(data)/30
newarray = rep(0,x)
for(i in 0:x-1){
newarray[i+1] = sd(data[((i*30)+1):(30*(i+1))])
}
plot(newarray,type="l",xlab="Time",ylab="Standard Deviation")
return()
}


skew = function(x){
sk = (sum((x-mean(x))^3)/length(x))/((var(x)^(3/2)))
return(sk)
}
kurt = function(x){
ku = (sum((x-mean(x))^4)/length(x))/((var(x)^2)) - 3
return(ku)
}

confi = function(model){
pro = 0:100000/100000
x = matrix(0,nrow=99999,ncol=2)
for(i in 1:99999){
x[i,] = gpd.firl(model,prob=pro[i+1])
```

```
}
return(x)
}

MODEL CHECKING FOR POT:
plots = function(data,u,sigma,xi){ #n=length(y)
n = sum(data>u)
y = rep(0,n)
H = rep(0,n)
Hinv = rep(0,n)
counter = 1
for(i in 1:length(data)){
if(data[i]>u){
y[counter] = data[i]-u
H[counter] = 1-(1+((xi*y[counter])/sigma))^(-1/xi)
# Hinv[counter] = u + ((sigma/xi)*(y[counter]^(-xi)-1))
counter = counter + 1
}
}
H2 = sort(H)
y2 = sort(y)
#Probability plot
plot((1:n)/(n+1),H2,xlab="Empirical",ylab="Model",
main="Probability Plot")
lines((0:10)/10,(0:10)/10,type="l")
#Quantile plot
#qq(u+((sigma/xi)*((((1:n)/(n+1))^(-xi))-1)),yx,xlab="Model",
ylab="Empirical",main="Quantile Plot")
##qq(MODELL)
#Density plot
#hist(y,breaks=100,xlab="x",ylab="f(x)",main="Density plot")
#lines((1:x0)/100,(1/sigma)*(1+(xi*(((1:x0)/100)-u))/sigma)
^(-(1/xi)-1),type="l")
return(H2)
}



outSaver <- function(y){ #y=dataset

n = length(y)
parameters = matrix(0,ncol=4,nrow=n/2)

for(i in 1:(n/2)){
```

```
model = garchFit(formula = ~garch(1,1),data=y[i:((n/2)+i-1)],cond.dist='QMLE',
trace=FALSE)
resnavn = paste('residuals',i,'.txt',sep='')
signavn = paste('volatility',i,'.txt',sep='')
write.matrix(model@residuals,file=resnavn)
write.matrix(model@sigma.t,file=signavn)
parameters[i,] = as.numeric(model@fit$coef[1:4])
print(i)
}

write.matrix(parameters,file='arparameters.txt')

}


outPOT <- function(y){

n = length(y)
tsparam = read.table('arparameters.txt')
quantiles = matrix(0,ncol=4,nrow=(n/2))
probs = c(0.95,0.99,0.995,0.999)

for(i in 1:(n/2)){
resnavn = paste('residuals',i,'.txt',sep='')
signavn = paste('volatility',i,'.txt',sep='')
resi = read.table(resnavn)
vol = read.table(signavn)
stdres = resi/vol
gpdmod = fitgpd(stdres$V1,1)
uquan = qgpd(p=probs,loc=1,scale=as.numeric(gpdmod$param[1]),
shape=as.numeric(gpdmod$param[2]),lambda=1-gpdmod$pat)
tspar = as.numeric(tsparam[i,])
quantiles[i,] = tspar[1] + sqrt(tspar[2] + tspar[3]*(as.numeric(resi$V1[n/2]))^2 +
tspar[4]*(as.numeric(vol$V1[n/2]))^2)*uquan
}
for(i in 1:4){
print('Number of exceedances:')
print(sum(y[(n/2+1):length(y)]>quantiles[,i]))
}
return(quantiles)
}

outACER <- function(y,aq){
```

```
n = length(y)
tsparam = read.table('arparameters.txt')
quantiles = matrix(0,ncol=4,nrow=n/2)

for(i in 1:(n/2)){
resnavn = paste('residuals',i,'.txt',sep='')
signavn = paste('volatility',i,'.txt',sep='')
resi = read.table(resnavn)
vol = read.table(signavn)
stdres = resi/vol

tspar = as.numeric(tsparam[i,])
quantiles[i,] = tspar[1]  + sqrt(tspar[2] + tspar[3]*(as.numeric(resi$V1[n/2]))^2 +
tspar[4]*(as.numeric(vol$V1[n/2]))^2)*(as.numeric(aq[i,]))
}
for(i in 1:4){
print('Number of exceedances:')
print(sum(y[(n/2+1):length(y)]>quantiles[,i]))
}
return(quantiles)
}
```