A method to investigate the time evolution of the probability distribution of high temperature extremes in The Netherlands, applied to the extremes in the period of May 2006-April 2007

Dirk Wolters, Student of Wageningen University (850331971060)
Supervisors: dr. Leo Kroon (Meteorology and air quality)
and drs. Paul Torfs (Hydrology and quantitative water management)

June 2007
## Table of contents

Table of contents............................................................................................................. 1
1. Introduction..................................................................................................................... 3
    Motivation....................................................................................................................... 3
    Methodology.................................................................................................................. 3
    Structure of the report.................................................................................................... 3
2. Material and methods..................................................................................................... 5
    2.1 Used data.................................................................................................................. 5
    2.2 Brief overview of the procedure............................................................................... 5
    2.3 Stationarity and seasonality..................................................................................... 5
        2.3.2 Removing seasonality of the mean................................................................. 6
    2.4 Detecting non-stationarity....................................................................................... 6
        2.4.1 Moving averages............................................................................................... 6
        2.4.2 Sums of deviations........................................................................................... 6
        2.4.3 Range over standard deviation methods......................................................... 6
    2.5 Data transformations............................................................................................... 7
    2.6 Model fitting and parameter estimation.................................................................... 8
    2.7 GEV distributions.................................................................................................... 9
    2.8 POT distributions.................................................................................................... 10
        2.8.1 GPD distributions.............................................................................................. 10
        2.8.2 Poisson distributions......................................................................................... 11
        2.8.3 Poisson-GPD distributions................................................................................. 11
        2.8.4 Threshold selection............................................................................................ 12
    2.9 Method of application............................................................................................... 12
3. Results and discussion ................................................................................................. 14
    3.1 Stationarity............................................................................................................... 14
        3.1.1 Moving averages............................................................................................... 14
        3.2.2 Sums of deviations........................................................................................... 14
        3.2.3 Range over standard deviation method......................................................... 15
    3.3 Data transformations............................................................................................... 16
    3.4 Model fitting............................................................................................................. 17
        3.4.1 Non-transformed data....................................................................................... 17
        3.4.2 Transformed data............................................................................................. 18
    3.5 Poisson representations............................................................................................ 20
5. Conclusions and recommendations............................................................................... 23
References.......................................................................................................................... 24
Attachment I...................................................................................................................... 25
Attachment II..................................................................................................................... 28
Attachment III................................................................................................................... 29
Summary

The primary goal of this study was to assess the time evolution of the characteristics of the occurrence of high temperature extremes in The Netherlands, in the period from 1851 until present. This has been done by analyzing month temperatures measured at De Bilt weather station. As the primary data is seasonal in mean, variance and skewness of distribution, a transformation approach has been used, to represent all data points in a homogeneous way. Using this approach, a transformed, homogeneous time series has been obtained, in which extremal characteristics of the original time series have been conserved and unified between seasons. Generalised Extreme Value (GEV) models for the probability distribution of high temperature extremes, characterized by three parameters, are used for further analysis. A time series has been constructed of the evolution of estimates of the three parameters of this model. From this, it is concluded that the mean of high temperature extremes has increased significantly since the onset of recent climate warming. The variance in high temperature extremes might have increased also, but less significant. There seems to be no systematic change in the shape characteristics of the probability distribution of high temperature extremes.

As a second step, estimates have been made of the instantaneous probability distribution of high temperature extremes before, and after recent climate warming. Both estimates have been used to assess the sequence of high month temperature extremes that occurred between May 2006 and April 2007. It is concluded, that even incorporating linear climate warming, this sequence of extremes might be a very unlikely event in extreme value models. Non-linear effects in climate change (e.g. systematic changes in circulation patterns) might have played a role in this, fundamentally affecting the probability distribution of high temperature extremes. This conclusion however remains uncertain, as the used methods for assessing the possibility of a sequence of extremes are not exact but indicative.
1. Introduction

Motivation

In recent years, there has been much discussion (e.g. KNMI, 2006(a), KNMI, 2007, Van Oldenborgh, 2006) on the increase in occurrence and magnitude of high temperature extremes in the Netherlands, with regard to recent climate change. Several high temperature extremes that occurred have indicated, that the main impact of climate change might not be the warming of the mean itself, but the increase in high temperature events. The summer of 2003 is an example that is often referred to. Even more extreme, at least for the Netherlands, than the summer of 2003, might be the long sequence of extremely warm months, from May 2006 until present, April 2007. Different approaches have been used to assess the influence of an increase of the mean temperature and a possible increased climatic variance, on the occurrence of high temperature extremes.

Methodology

In general, in the climatology of weather observations, a month-by-month approach is used. A time series of measurements is mostly seen as being 12 separate monthly time series, being related, but incomplicable to each other. For example, regarding month temperatures, each month of course has a very different mean temperature, but also large differences in variance exist, as well as in the skewness of the distributions per month. Therefore statistical properties of such a time series are in many cases assessed by regarding observational data from single months separately. This way, with regard to temperature extremes, extremes in the same month or season can be compared, while extremes that have occurred in different months can not.

In this study, a different approach to the above is proposed. Month temperatures, from the period of January 1851 until April 2007 from De Bilt weather station, in the Netherlands, are analyzed as one time series, putting all months in a row, in chronological order. To make months comparable, transformations are used for each month. By doing so, a transformed time series is obtained, with values on a new scale, all months having a similar stochastic behavior.

On this transformed time series, an investigation will be made of the evolution of the characteristics of the stochastic distribution of high temperature extremes throughout the past 157 years. This will be done by using commonly applied extreme value distributions, using block maxima and peak-over-threshold approaches. When applying these distributions, it will also become clear, if the transformations used have worked out well.

Furthermore, the high temperature extremes in monthly average temperatures that occurred between May 2006 and April 2007 have been assessed. This 12-month period is interesting, because remarkably many high temperature extremes occurred. Some have put forward the hypothesis of a sudden, non-linear ‘climate jump’, while others have stated that these high temperature extremes can be well explained by a linear warming, combined with some increase in climatic variance. It will be tried, to investigate how likely the sequence of high temperature extremes that occurred, would be within a linear warming trend. It will be assessed, if the hypothesis of a ‘2006 climate jump’ can be invalidated using the method mentioned above. All conclusions from this study are statistical; no connection has been made to the physics of climate change.

Structure of the report

See page 1 for a table of contents. In this report, first of all, all methods used will be explained briefly. This is done in the Materials and methods section, where all methods are explained in the order of which they were applied in this study. For the non-statistician, it is necessary to read this part, to understand the rest of this report. Methods for the study of data stationarity will be explained, after which the used data transformations will be discussed, and the methods used for the analysis of extreme values. In the ‘Results and
discussion’ part, results of all used methods are shown, and the result of every method will be discussed separately. Results are shown and discussed in logical order. In the ‘Conclusion and recommendations’ part, all discussed results are taken together, and the final conclusions are drawn. Also, some recommendations are done on possible future research, making use of the conclusions of this study. After this, a list of references is printed, followed by the attachments.
2. Material and methods

2.1 Used data

The data used are month temperatures at De Bilt weather station (WMO 6260) obtained from the Dutch royal weather institute KNMI. These temperatures are defined by taking the mean of all daily mean temperatures in a month. Month temperatures from the period of January 1851 until April 2007 were used. No study has been performed on the homogeneity of the data from 1851-1900, which will be assumed agreeable for the methodology and goals of this study. The homogeneity of the rest of the period has been studied by KNMI (KNMI, date unknown), although this study has focused on the bulk on the data, and will not for certain have covered all effects on high temperature extremes. This data has been used because it is easily available, and gives a good representation climatic effects on temperature. A similar study performed on other data in the same period, for example daily maximum temperatures, should not evidently have similar results. To prevent any misunderstanding of the words used, the word ‘month temperature’ will be used as above, being the average day temperature throughout one month, and being one data point in the basic data in this study. The word ‘long-term mean’ will be a mean of month temperatures of one month throughout a certain time period, for example the long-term mean of January 1936-1985.

2.2 Brief overview of the procedure

The main objective in this study, is to obtain an estimate of the time evolution of the probability distribution of high temperature extremes throughout the dataset. This is done, by applying a moving fit of a model for extreme values to the data. Before the data could be sufficiently described by such a model, transformations had to be applied for each month. The transformations used here, are Box-Cox power transformations (see Data transformations). In order to find the best transformations, a Box-Cox transformation parameter has been fitted to a sufficiently large sample of identically distributed data points. To find such a sample, seasonality of the mean had to be removed, and a long enough stationary time period should be found. Thus, the steps performed, in logical order, will be:
1) Removal of seasonality of the mean
2) Finding a stationary period
3) Fitting of transformations to this period
4) Using these transformations on all data
5) Fitting of an extreme value model on these transformed data.

Also, attempts will be made to fit an extreme value model to non-transformed data.

2.3 Stationarity and seasonality

2.3.1 Stationarity

In this study, a sufficiently long period needs to be assumed stationary. Without the assumption of stationarity (Davis and Brockwell, 1996), identical distribution of data points can never be assumed. Stationarity in a strict sense means, that all data from a given time series are generated by the same stochastic process. In that case, the mean and the variance have one exact value, to which they converge for all samples. In processes that occur in nature, this principle of stationarity almost never holds. Many real time series are autocorrelated to some extent, violating the principle of identical distribution of all data. For environmental and ecological systems, in most cases seasonal patterns exist, resulting from weather influences. Furthermore, overall stochastic behavior of many systems eventually, or even constantly changes systematically. Still, in many cases, stationarity can be assumed.

In most cases, systems having autocorrelation or cyclic behavior are generally still considered stationary, as long as their mean and variance do not change on time scales longer than the cyclic or autocorrelation time scale. In some cases, stationarity of the mean and stationarity of the variance are seen as separate characteristics.
The data assessed in this study of course incorporate a very strong seasonality, as every month in the year has its own mean and variance. Although per se this is not a violation of stationarity on a broader time scale, the seasonality in the mean will be removed. By doing this, it becomes more convenient to address the question of non-stationarity further. The seasonality of the mean is removed by calculating the long-term averages of all month temperatures within a selected period, and subtracting these means from all month temperatures. Next to a seasonality in the mean, the data is also expected to have a seasonality in the variance. Furthermore, month temperatures are expected to be distributed quite asymmetrically (skewed) around the mean. This skewness could be significantly different for each month. The seasonality in the variance and in the skewness can not be removed in a good way by dividing by the standard deviation, or any such simple operation. Thus, they will be retained at first. Later, these seasonalities will be addressed by power transformations.

2.3.2 Removing seasonality of the mean
For all methods in this study, a seasonality of the mean is not desired, this seasonality will be removed by subtracting long-term means from all data points. The goal is to remove the seasonality only, using long-term means from a long period, this period being the same for every long-term mean. For this, the period of 1936-1985 is taken, as the characteristics of the seasonality of the mean are not expected to have changed significantly within this period. A shift of this period of some years or decades forward or backward is not expected to affect the removal of seasonality strongly. After the seasonality has been removed, all information on non-seasonal variations in the mean will stay intact. All of the following analysis in this report, will be performed on these anomalies from 1936-1985 long-term means.

2.4 Detecting non-stationarity

2.4.1 Moving averages
There are different methods for performing tests on stationarity. Most straightforward, is to construct moving averages and moving variances. A short moving average will have many fluctuations, retaining those from the original time series, while a long moving average will behave much slower. Short moving averages are influenced greatly by short-time scale noise effects. Long moving averages dampen these effects, though not being able to detect shorter non-stationary episodes, and not being able to detect the exact moment of a change in system mean accurately. By trying, a balance has to be found between short and long moving averages. One should of course be careful, not to decide upon a moving average, taking the result that fits into a desired conclusion. Moving averages provide a good intuitive view on stationarity of the mean, but can give no exact measure of it. The same holds for moving sample variances. With moving averages, a first conclusion can be drawn on which period seems stationary and long enough to perform our further analysis on.

2.4.2 Sums of deviations
Another simple procedure for addressing stationarity of the mean, consists of looking at the sum of deviations from any given mean. Within an identically distributed and stationary sample, the sum of deviations from the mean until any point should fluctuate around a constant value. If this sum in parts of the sample is systematically increasing or decreasing, the local mean must be different from the overall mean. This would indicate non-stationarity. This method, although simple, provides another intuitive view on stationarity, which is more sensitive than moving averages. This method will be used, by systematically going through different mean values, searching for all means that, for the eye, generate a more or less time-constant sum of deviations, indicating a stationary period.

2.4.3 Range over standard deviation methods
Expanding the principle of the above method, even a more or less objective quantification of stationarity can be made. This can be done by applying the Range over standard deviation (R/S) method (Van de Putte, 2006). In this method, for a certain period, the difference
between the maximum and minimum of \( S_n \) (the so-called ‘range’), is seen as a measure of stationarity. The value is generalized by dividing by the sample standard deviation in the period and by the square root of its length \( n \). The R/S statistic yields high values when non-stationarity is likely. An R/S expression with a certain length will be moved through the data, just as a moving average. Any non-stationary episodes should cause a peak in this moving R/S value. On one hand, an R/S value moving in time should be examined that uses a length that smoothes out the data enough. On the other hand, the length should be much less than the length of any stationary periods to be expected.

See table 1, for some commonly used reliability values of a conclusion of non-stationarity based on different R/S values.

**Table 1: R/S thresholds for the indication of non-stationarity (source: Van de Putte, 2006)**

<table>
<thead>
<tr>
<th>R/S value</th>
<th>1.62</th>
<th>1.75</th>
<th>2.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>reliability</td>
<td>90%</td>
<td>95%</td>
<td>99%</td>
</tr>
</tbody>
</table>

A mathematical expression for the R/S statistic is given in equation (2), in which \( sd_x \) is the standard deviation of the whole sample, and \( m \) is the whole sample length, and \( n=1,2,...,m \). For every \( n \), a sum until \( n \), of the deviations from the whole period mean, is calculated. \( S_{\text{min}} \) and \( S_{\text{max}} \), are the minimum and maximum of these sums of deviations in the sample.

\[
R / S = \frac{|S_{\text{min}}| + |S_{\text{max}}|}{sd_x \cdot \sqrt{m}}
\]  

(2)

### 2.4.4 Physical validation

As stated earlier, it will be hard to objectively address stationarity in the time series considered. Mathematical methods can not fully prove that the assumption of stationarity is correct. Extra information can be used though, to further base any conclusions on stationarity. In this particular case, it is known that the second half of the nineteenth century belonged to the little ice age, a globally cooler period, from roughly 1500 to 1900 A.D (KNMI, 2006(b)). Therefore, a non-stationary episode should have occurred at the beginning of the 20th century, as a result of a warming at the end of this cooler period. The next large-scale physical event that should have caused non-stationarity in temperatures, is the onset of recent climate warming. In Western Europe, the most significant warming is considered to have started in the second half of the 20th century presumably around 1980 (KNMI, 2003). A period that can be assumed stationary and useful for our analysis, is therefore expected to have occurred in between these non-stationary period, so roughly between 1900 and 1980.

### 2.5 Data transformations

The first transformation was the removal of the seasonal effect in the mean temperature, by subtracting 1936-1985 long-term means. Still, months are differently distributed, because they have a different variance, and their distributions might be differently skewed. The most radical approach to deal with the problem of differences in variance between months, is to just make the months equally distributed, using appropriate transformations for this. Of course, this should be done in a way that preserves the extremity of all values. In other words, extremes should, in their own months, remain as extreme as they were before the transformations, while unifying the extremity of every value over all months.

A first step in doing this, are **power transformations**. These transformations are based on the principle, that skewness of distributions changes when raised to a power, and that for every distribution, a power exists that makes it symmetrical around the mean (Peltier et al., 1998, URL 1). By applying the method of maximum likelihood (see next chapter), maximum likelihood estimates (MLEs) can be found for this power. The power transformations used in this study are Box-Cox transformations, as defined in equation (3) and (4), in which \( X \) are the values to be transformed, and \( \lambda \) is the Box-Cox power found (Peltier et al., 1998, URL 1).
\[ Y(X; \lambda) = \frac{X^\lambda - 1}{\lambda} \quad \text{if } (\lambda \neq 0) \]
\[ Y(X; \lambda) = \ln(X) \quad \text{if } (\lambda = 0) \]  

These transformations have been tried out at first, as an attempt to find a way of making the data homogeneous enough for extreme value-analysis, and were selected as the method of choice after this. Although they appear to work here, this might be different for other data.

Box-Cox transformations can only be performed on positive values. Therefore, first of all, the integer of the lowest value in the dataset plus 1 is added to all the data. This homogeneous linear transformation does of course not change any of the characteristics under investigation. For the reference period, for every month an optimal \( \lambda \) will be estimated. After this, the data in the reference period are transformed. After transformation, any skewness is removed optimally. Still, data are incomparable, because different transformations result in different means and variances for every month. Therefore, the mean and standard deviation of every transformed month in the reference period are calculated. The last part of the transformation is then a subtraction of this reference mean, and a division by this reference standard deviation. All steps described above are summarized in equation (5), in which \( X \) are the deviations from long-term mean to be transformed, \( X_{\text{min}} \) is the smallest value in the dataset, \( \lambda \) is the reference period Box-Cox power, and \( SD \) and \( M \) are the standard deviation and mean in the Box-Cox transformed reference period. \( \lambda, M \) and \( SD \) are of course different for every month.

\[ Y(X; X_{\text{min}}, \lambda, SD, M) = \left[ \frac{(X + (X_{\text{min}} + 1)^\lambda)^{\frac{1}{\lambda}} - 1}{SD} \right] / M \]  

To have some overview, in table 2, the different datasets are denoted by DATA0, DATA1 and DATA2. These terms will be used if any confusion between datasets is possible.

**Table 2: Different datasets used**

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATA0</td>
<td>month temperatures from January 1851 until April 2007</td>
</tr>
<tr>
<td>DATA1</td>
<td>deviations of DATA0 from the 1936-1985 long-term means of DATA0</td>
</tr>
<tr>
<td>DATA2</td>
<td>DATA1 transformed per month by equation (5)</td>
</tr>
</tbody>
</table>

**2.6 Model fitting and parameter estimation**

The transformed data will be analyzed with the use of extreme value distributions. The objective of these distributions, is to find likely expressions for the extremes occurring in a certain stochastically behaving system. Therefore, the distribution of extremes in measurements is analyzed. Certain standard distributions for extreme values, such as GEV, GPD or Poisson distributions are fitted to measurement data. These models are based on the principle, that the occurrence of different types of extremes can be described in most practical cases, by selecting an appropriate model, optimizing its parameters. First, the fitting of these parameters to the data is explained, and after this, information is provided on the different models used.

As an example, so-called GEV distributions contain the parameters \( \mu, \sigma \) and \( \xi \). For finding the distribution that fits the measurements best, optima in \( \mu, \sigma \) and \( \xi \) must be found, creating a distribution that deviates as little as possible from the values observed. In extreme value statistics, estimation of parameters always balances between having enough data, and having extreme enough data. The more extreme the data used, the better it has the characteristics of an extreme value distribution. On the other hand, using more extreme data directly means using less data, and less accuracy.

In this study, estimates are fitted using maximum likelihood theory (Reiss and Thomas, 2001, URL 2, URL 3). In daily language, likelihood is somewhat equal to probability. In
statistics, however, not. Likelihood does not apply to any probability, but it is used to quantify
how likely any estimated parameter, or set of parameters, is. The better a parameter
estimates a probability distribution that fits the outcomes observed, the more likely it is
considered. Maximum likelihood estimation is a way of calculating backward from a known
outcome (known from data), into parameters that would predict this outcome. Likelihood is
calculated from a likelihood function.

The likelihood function is defined as a conditional probability density function turned around.
The usual probability density function \(f(x|\theta)\), always gives a probability density at \(x\), given
the condition of the parameter value or vector \(\theta\) being true. We usually view this function as only
giving a probability density at \(x\). Of course though, in a sense, it gives also a probability
density \(\theta\), if we would consider \(\theta\) varying, and \(x\) constant. This could be expressed as the
Likelihood function \(L(\theta|x)\). The value of \(L\) becomes higher, the more likely a chosen \(\theta\) would
produce the \(x\) observed. Because \(L\) in most cases no longer ranges precisely from 0 to 1, and
because its values depend on measurements, it is not a real probability anymore, and is
therefore called likelihood.

It can be shown, that \(L(\theta|x) \propto f(x|\theta)\). The likelihood function is proportional to the
probability density function; the main difference is, that in the probability density function \(\theta\) is
a known value (or, more practical, an estimated condition) and \(f(x)\) is calculated, while in
the likelihood function \(x\) is taken constant (conditional), and \(L(\theta)\) is calculated. Maximizing \(L\),
varying \(\theta\) then gives the \(\theta\) with the highest likelihood to produce given \(x\). \(\theta\) values
corresponding to such a maximum are called Maximum likelihood estimates (MLE) of \(\theta\). In
many cases, the negative logarithm of \(L\) is considered, because it can be more convenient to
determine. In this study, all extreme value distributions are fitted using MLE parameters. This
is done numerically, as in most cases multidimensional maxima are needed.

In the MLE fitting procedure, to every parameter in the fitted model, a certain standard error
can be assigned, indicating the uncertainty of the fit per parameter. Likelihood values for
different modes can not be directly compared, unless models are concerned that have a
‘nested’ hierarchy, meaning that they must be the same model, only differing in the number
of degrees of freedom.

2.7 GEV distributions

The oldest and most basic method of selecting extremes, is the method of block maxima
(Reiss and Thomas, 2001; Gumbel, 1960; Gilloeland and Katz, 2006, URL 4, Katz et
al., 2002). In this method, extremes are selected, as being the maxima of given, fixed
intervals (blocks). For example, from month data, extremes can be selected by taking year
maxima.

According to extreme value theory, identically distributed block maxima can be modeled with
a generalized extreme value (GEV) distribution, defined by equation (6), in which \(F\) is the
cumulative distribution function (CDF) of block maximum \(x\).

\[
F(x; \mu, \sigma, \xi) = \exp \left( -\left[ 1 + \xi \left( \frac{x - \mu}{\sigma} \right) \right]^{\frac{-1}{\xi}} \right)
\]  \(6\)

In this distribution, three parameters exist: \(\mu, \sigma\) and \(\xi\), \(\mu\) is a location parameter, determining
the location of the peak of \(f\). \(\sigma\) is a scale parameter, determining the ‘widthness’ of the
distribution. \(\xi\) is a shape parameter that determines shape, especially the type of tails in the
distribution. For three ranges of \(\xi\), three types of GEV distributions can be distinguished. For
\(\xi < 0\), the GEV is called a Weibull distribution, or GEV2. For \(\xi = 0\) it is a Fréchet distribution, or
GEV1. For \(\xi > 0\), the GEV is a Gumbel distribution, or GEV0. Gumbel distributions have two
infinitely long tails. Fréchet distributions have a lower limit, and Weibull distributions have an
upper limit. See figure 1. Using a GEV, predictions can be made about the occurrence of
block maxima that are distributed identically to the block maxima that were fitted to the GEV.
In most applications, especially in climatic phenomena, yearly maxima are used. In this way, yearly cycles are eliminated, so that data points in the distribution of maxima are identically distributed. For example, in catchment hydrology, high discharges at the end of winter would make certain block maxima systematically higher than others, if blocks are chosen that are not a multiple of 12 months.

If the aim of this study would have been to address danger from heat waves (analogous to danger from high water levels in hydrology), yearly maxima of month temperatures could have been used. For study on climate change though, not only the highest temperatures in a year are of interest, but all deviations from average are. In this study, it has thus been tried to remove any yearly cycles. In the non-transformed data (DATA1), only the cycle in the mean has been removed, and still a cycle in variance is present. Since winter months have more high temperature extremes than other months, most of the yearly maxima found will be winter maxima, and high temperature extremes in the other seasons are thus taken less into account. Fitting a GEV model on non-transformed data might cause a higher uncertainty, but also might cause an invalidity of any GEV-type description, as months are differently distributed, and a combination of these distributions might not have a typical GEV shape anymore. Using Q-Q plots, the validity of the assumption of a GEV-type distribution will be visually examined. It is expected, that extremes in the transformed data (DATA2) can be accurately described by a GEV distribution. The non-transformed data might be less describable by a GEV model.

2.8 POT distributions
The approach of taking extremes from block maxima discussed above, has an important drawback. Out of every block, only one measurement is taken. In the case treated here, out of 12 months, 11 are discarded, and only one data value, the maximum, is retained. When occasionally other significant extremes occur within the block, valuable data is not used. To improve this, the approach can also be turned around in a sense. Instead of taking maxima from fixed periods, one can take all the maxima above a fixed threshold, allowing all lengths of periods to exist between them. This yields a lot of extra information, because more data can be taken into account. A distribution can be found for the values of the extremes (the excesses above the threshold), but also for the frequencies of exceedance of the threshold. The approach thus applied, is referred to as a ‘peak over threshold’ (POT) approach (Reiss and Thomas, 2001, Gilleland and Katz, 2006, URL 4, Katz et al., 2002). Some different extreme value distributions can arise. It can be proven, that when the extremes of a phenomenon can be described by a GEV distribution, also POT distributions can be applied.

2.8.1 GPD distributions
The magnitudes of exceedance of a high threshold, can in the asymptotic case be considered to be Pareto distributed. The Pareto distribution applies for many phenomena, in which some events occur very frequently, while very many occur only occasionally. For example, the sizes of human settlements can be considered Pareto distributed; there are only a few very big cities, while many small villages exist. Many situations in which an equilibrium exists between the small and the large, follow a Pareto distribution (URL 5). This also works for the distribution of extremes, in the form of threshold excesses. Small excesses occur relatively frequently compared to large excesses. See equation (7) for the CDF of the generalized Pareto distribution (GPD) family.
\[ F(x; u, \sigma, \xi) = 1 - \left[ 1 + \frac{\xi(x-u)}{\sigma} \right]^{-\frac{1}{\xi}} \] (7)

This function gives the cumulative probability for \( X \) exceeding the value of \( x \), given that is already exceeds the threshold \( u \).

Note, that the scale parameter \( \sigma \) in this equation is not the same as the GEV distribution \( \sigma \) in equation (6). It can, though, be expressed in a relation with GEV distribution \( \sigma \) and \( \mu \). The shape parameter \( \xi \) can be directly interchanged between a GEV distribution and a GPD. Therefore, from a GEV distribution, GPD parameters can be deduced, while the reverse is not possible, because \( \mu \) remains unknown.

2.8.2 Poisson distributions

When exceedances of a threshold occur independently, it can be stated that every occasion has a certain, identical probability of exceeding the threshold. This probability can be easily estimated from measured exceedances. This idea leads to the assumption of threshold exceedances being binomially distributed (URL 6). The binomial distribution has the well-known mass function \( f \) in equation (8)

\[ f(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k} \] (8)

The chance, in any experiment, of having exactly \( k \) successes is determined from the number of trials \( n \) and the chance of success per trial \( p \). Working with large binomial samples, especially when cumulative probabilities are needed, is very difficult and time consuming. In some cases with a large \( n \) and \( p \), it suits to transform a binomial distribution into a normal distribution, by using an approximate transformation. When this normal approximation does not hold, it is necessary to use so-called Poisson distributions, the distributions to which all binomial distributions with large \( n \) converge. Poisson distributions have parameter \( \lambda = np \).

Equation (9) gives the probability mass function for the probability of having \( k \) successes in \( n \) trials.

\[ f(k; \lambda) = \frac{e^{-\lambda} \lambda^k}{k!} \] (9)

In the case treated here, \( \lambda \) can be the number of months \( n \) considered, multiplied by the chance of threshold exceedance per month \( p \). \( \lambda \) is also referred to as the expected value. Cumulative probability values can be obtained from tables or from mathematical functions applied by statistical pc software.

The parameter \( \lambda \) can directly be deduced from GEV distribution parameters \( \mu, \sigma \) and \( \xi \) and threshold \( u \), using the expression in equation (10) (Gilleland and Katz, 2007).

\[ \log \lambda = -\frac{1}{\xi} \log \left( 1 + \xi \frac{(u-\mu)}{\sigma} \right) \] (10)

2.8.3 Poisson-GPD distributions

Through the mathematical point process concept, it is possible to fit an extreme value model using both the Poisson-distributed exceedances, and the GPD distributed exceedences. Because not only magnitudes of excess, but also exceedance frequencies are taken into account, through this method it becomes possible to deduce all GEV parameters, including the location parameter \( \mu \), using a POT method instead of a block maxima approach. Usually, the outcome of a point process fit is expressed in GEV parameters \( \mu, \sigma \) and \( \xi \). Through transformations (Katz et al., 2002), this GEV distribution can be converted into a GPD, a Poisson distribution, or into a GEV distribution for maxima of blocks of different lengths than the default. In this way, extremes are expressed with a similar concept as in the block maxima method, while
much more data than only block maxima can be taken into account. The point process model can be viewed as a concept that unifies block maxima and POT approaches for extremal distributions. Another possibility that arises when using a point process approach, is to account for non-stationarity and seasonality, by using non-constant thresholds. In this study, point process models will be used, represented in GEV parameters.

2.8.4 Threshold selection

In POT models, the choice of a threshold is rather non-trivial. On one hand, one needs to put the threshold sufficiently high to let the assumption of the peaks being GPD distributed, be valid. On the other hand, as much data as possible should be taken into account, making the result more reliable. The method used in this study for choosing an appropriate threshold, is to fit the models to a whole range of thresholds, investigating the response of the parameters (Gilleland and Katz, 2007). A threshold should then be chosen that is as low as possible, so that as much data as possible are incorporated, so that the standard error of the MLE is smallest. In the same time, it should not be biased by an invalidity of the GPD or Poisson distribution assumption. Thus, a higher threshold should give similar parameter values. When plotting parameter MLE values with standard error bars for a range of many different thresholds, a threshold is chosen around a range with more of less constant parameter MLEs, while having a standard error as small as possible.

It is also possible to use variable thresholds (Gilleland and Katz, 2007). In the case studied here, this is particularly interesting. Now, without the need to transform the data, the yearly cycle in variance of extremes can be taken in into account in a GEV model in some way. A certain quantile value of the measurements will be chosen for all months, to construct a vector of monthly thresholds that can be applied in POT model fits. It will also be investigated, how much difference this varying threshold makes to parameter estimates and their standard errors. Of course, the results will also be compared to results that were obtained from model fits to the the Box-Cox transformed data.

2.9 Method of application

After fitting the different models, the distribution of high temperature extremes in the reference period assumed stationary, has been estimated. It will be investigated, how much the results of the different approaches to extreme value modeling deviate from each other. Approaches that will be considered, are a GEV point process fit on the non-transformed data, with constant and monthly threshold, and a GEV point process fit to the transformed data. By plotting Q-Q plots of observed versus modeled quantiles, the characteristics of the model fits will be compared. Furthermore, plots will be made of MLEs of the transformed data model parameters, moving in time. This procedure can be compared to a moving average that uses a certain window, only not calculating averages, but GEV parameter MLEs. Now, information is obtained about possible changes in characteristics of the extremes that occur in the time series. Furthermore, a graph will be shown of the exceedance probabilities of all year maxima in the transformed dataset, being calculated by the most likely model for their particular years. It will be checked, if these exceedance probabilities are homogeneous in time, which would be an extra evidence for the validity of the methods used.

With the different extremal distributions found, furthermore, estimates can be made of the probability of the sequence of extremes in 2006 and 2007, mentioned earlier. Using the best GEV fit directly, only the year maxima of 2006 and 2007 can be assessed. This can be done by assuming a distribution as in the reference period, or by assuming a distribution after recent climate change, by estimating GEV parameters from an extrapolation of the above mentioned time series of parameter MLEs. Because having many extremes in a year will also make a high yearly maximum more likely, taking this maximum is in fact a possible choice for characterizing the whole period. One can feel though, that this would give a rather poor estimate, because all other months are not taken into account explicitly.

Another interesting way to estimate the probability of the whole period occurring, more directly investigating the whole succession of very warm months, is a Poisson representation.
This way, the frequency of exceedance of a certain threshold in the investigated period, can be related to estimated or observed exceedance frequencies, based on data from the reference period. Then, a probability for the occurrence of the whole period, taking into account all months, can directly be deduced. Large difficulties do arise though, in choosing a threshold, and in defining the period in which the number of exceedances are investigated. Both choices have a great impact on the outcome, while being very arbitrary. It would not at all be scientific to choose the threshold and number of months exactly so, that the return time of the event is maximized. Still, this is what one would do intuitively, because one tries to use a threshold and a period length that best ‘characterize’ the event under investigation. In the case treated, there seems to be no clear way to avoid this problem. The method above will be applied using a period length of 12 months (May 2006-April 2007), and comparing results for some different thresholds. Results will be indicative, but far from exact.

Lambda values and return times for the occurrence of the whole period will be estimated, for the situation before, and after climate change. This will be done using the GEV fit to the reference period, and using an extrapolation of the time series of GEV model parameters until 2006, to account for recent climate change. These lambda values will be obtained using equation (10).

In applying the above methods, mostly the statistical program R will be used. R is a software package used for many statistical applications, and many additional toolkits and functions have been written to be performed by the program. R is operated via a command line interface, permitting the use of scripts, while some toolkits have their own graphical user interface. The R toolkit extRemes, developed by UCAR, was used for finding MLE estimates for distributions. The toolkit ‘car’ was used for finding and applying Box-Cox transformations. See attachment I, for all R scripts that were constructed and used for this study.
3. Results and discussion

3.1 Stationarity

3.1.1 Moving averages

See attachment II for a time series of deviations from long-term mean temperatures. In this time series, moving averages of 2 years and of 10 years length are plotted. The three periods that are expected from climatic history, are visible in this plot. From roughly 1851 until 1910, clearly a cooler period exists, ending in a significant warming from 1890 until 1910. Then, indeed a more or less stationary period seems to exist, ending in a warming in the eighties. From these moving averages, it seems, that the period 1910-1985 might be suitable for the assumption of stationarity, and thus for the use as a reference period. No non-stationarity can be seen in this period from this figure, while before and after it, the moving averages show a large evidence for non-stationarity. Furthermore, in figure 2, moving standard deviations are shown.

3.2.2 Sums of deviations

The first temperature that shows a period of more or less constant sums of deviations, is the value of around -0.9K, see figure 3. From 1851 until around 1900, no non-stationarity can be indicated. After 1910, the mean has evidently changed. The next value showing a period in which the sum of deviations does not change very systematically, is -0.05K, see figure 4. Before 1910, and after 1990, the mean evidently differs from this -0.05K, but within this period, no severe non-stationarity can be pointed out. The next temperature that shows a more or less constant sum of deviations, is 1.1K. Because the y-axis would get very wide otherwise, figure 5 is zoomed in on the period after 1950. Before 1990, the mean of DATA1 is significantly lower than 1.1K; after 1990 it seems more or less stationary around it.

Figure 2: moving standard deviation

Figure 3: sum of deviations of DATA1 from -0.9K
3.2.3 Range over standard deviation method

In figure 6, a moving 20-year R/S statistic is plotted. As would be expected, two major peaks exist, around 1900 and around 1980, the moments of respectively the end of the little ice age, and the beginning of strong recent global warming. In figure 7, a plot of 30-year R/S values, a similar, but more pronounced pattern exists.

Values of the R/S statistic can also be used to further investigate the possible stationary and non-stationary periods that arise from figure 3, 4 and 5. The first period with an approximately constant sum of deviations, 1851-1900, yields an R/S value of 1.46, indeed raising no strong evidence of non-stationarity. 1851-1915 however, yields a R/S value of 3.18. Thus, between 1900 and 1915, mean values have to have changed. 1900-1980 yields a R/S value of 1.56. Taking 1910-1980 significantly lowers the value, yielding 1.26. 1910-1985 yields the slightly higher R/S value of 1.32. 1910-1990 yields 1.59, coming close to a R/S value that might indicate non-stationarity. Values obtained above might be interpreted using critical values in table 1. The values in figures 6 and 7 seem to indicate many non-stationary episodes, using critical values in the table. They exceed the value of 1.5 on many occasions. The 20-year or 30-year periods R/S values in these figures are likely to show some seemingly non-stationary effects though, caused by long-term weather oscillations, of which most are not known into great detail, for example NAO fluctuations. The above obtained single R/S values are taken over longer periods, and thus might be better detecting the structural non-stationarity that is of importance here.

On the basis of the figures shown, of moving average, sums of deviations and R/S values, it is concluded that the assumption of stationarity can be made on the period after 1910, and before 1985. This period will be used as a reference period for further analysis of high temperature extremes.
3.3 Data transformations

Before transformation, the data were distributed per month as in figure 8. After transformation according to equation (5), data points become distributed as in figure 9. Please note, that y-axis values in this figure are no longer temperatures, but standardized dimensionless values.
Figure 9: Distribution of transformed temperature anomalies per month

Comparing figure 8 and 9, for the eye it becomes clear how the transformations have worked. In figure 8 it can be seen, that there are large systematic differences in the distribution of month temperatures between different months. Winter months have more variation than summer months, and thus have more low and high temperature extremes. From figure 9, no systematic difference in distribution between months is visible anymore. This visually indicates that the used transformations have worked as they were hoped to. Because this type of transformations is normally not used in this particular context, this is an interesting conclusion, as it was not self-evident a priori.

3.4 Model fitting

3.4.1 Non-transformed data

See table 3 for parameter MLEs and negative log-likelihoods of all model fits performed over the reference period. First, an attempt is made to fit a point process model using non-transformed data (DATA1). A constant threshold of 0.5 is used, being selected by fitting the model parameters to a range of thresholds. Next, a variable threshold is defined. This is done by constructing vectors $V_n$ of the $n^{th}$ lowest temperature for each month. As a threshold, the $47^{th}$ vector is chosen, as these values all lie relatively close to the value of 0.5. See figure 10, in which threshold values are plotted within the distributions of different months.

| Table 3: Parameter MLEs, standard errors and negative log-likelihoods |
|-----------------------------|--------------|---------------|----------------|-------|
|                           | $\mu /se \mu$ | $\sigma /se \sigma$ | $\xi /se \xi$ | -log L |
| DATA1, threshold 0.5       | 2.169/0.086  | 0.86/0.034     | -0.294/0.043   | 166.2 |
| DATA1, variable threshold  | 2.178/0.0851 | 0.863/0.033    | -0.302/0.040   | 159.0 |
| DATA2                      | 1.381/0.055  | 0.531/0.023    | -0.311/0.054   | 38.5  |

The outcome of the point process fit on non-transformed data seems useful at first glance. Standard errors of parameter MLEs seem reasonably small. Also, it seems that the use of a variable threshold does not improve the fit significantly; parameter estimates are almost the same, and the negative log-likelihood of the fit only increases slightly.
3.4.2 Transformed data

Next, point process models are fitted to the reference period using transformed data (DATA2). Again, the threshold value of 0.5 seems close to optimal, having fitted the data to a range of thresholds. See again table 3 for the found parameter MLEs and the negative log-likelihood of the fit. Also, a ‘moving’ point process fit is made for this data. Plots of MLEs obtained by this procedure, are shown in attachment III. From this figure, an estimation can be made of the influence of climate change on the location parameter. Drawing a trendline using the method of the smallest sum of squares, it appears as an estimate, that from 1985, it has increased with 0.0432 per year. For 2006, this yields the estimated value of $\mu=2.502$.

Also, exceedance probabilities of yearly maxima throughout the whole period are plotted, see figure 11. In this figure, probabilities are defined by models that are most likely for every single year, using the parameter values that are plotted in attachment III. The average exceedance probability is 0.507, of which the expected value would be 0.5.

Q-Q plots of modeled versus observed quantiles are shown in figure 12, for both the original data, and the transformed data. For the original data, only a Q-Q plot from the constant threshold model is shown, as parameters in the variable threshold model are practically the same.
The left Q-Q plot in figure 12, of the model fit to the original data, raises concern. In the high tail of the distribution, it seems that the observations show a large deviation from the model. The lower part of the distribution seems well modeled. It seems, that in the high tail of the distribution, the data cannot be fitted into a GEV model, as it is not homogeneous enough. High temperature extremes seem to differ too much between months to be considered homogeneous enough for a GEV representation, also when a variable threshold approach is applied. The Q-Q plot of the point process fit to the transformed data (figure 12, right plot) looks much better than that of the fit to the non-transformed data. The model found shows no large systematic deviations from the transformed observed values.

Furthermore, the pattern in figure 11, seeming homogeneous in time, is an extra argument for the hypothesis that the transformations used, have completely covered the statistical non-homogeneities that existed before, also outside the 1910-1985 reference period. The same can be said for figure 13, showing that the parameters found can predict the binomial threshold exceedance behavior of the data in the reference period well. This information provides further evidence for the validity of the methods used. The fact that it seems to work for this data, does not imply though, that it would work for all other datasets of temperature. Altogether, figure 9, 11, 12 and 13 show nicely, how homogeneity is needed for GEV models to be applicable, and how methods of data transformation have provided for that, in this case.

The graph shown in attachment III, of the time evolution of the parameter MLEs of a 15-year GEV fit, has some interesting implications. The evolution of the location parameter $\mu$ is comparable to the evolution of the mean in attachment II. The effect of the warming after 1985 is clearly visible, and well beyond the scale of the standard error ranges.

The evolution of the scale parameter $\sigma$ is less distinct. Several features seem interesting. At first, there is a sudden drop in $\sigma$ a few years before 1900, accompanied by a decrease in its standard error. This might be explicable from the fact that in 1897, the weather station was moved from Utrecht to the nearby town De Bilt. Because of the 15-year width of the parameter fit, the effect should appear some 7 years before the moving of the weather station, which conforms to the figure. Throughout the 19th century, $\sigma$ seems stationary. Some increase appears from the figure, in the period after 1985. It remains doubtful if $\sigma$ really has increased significantly in this period. The pattern in attachment III could also be an apparent effect, resulting from the warming of the mean in the period of after 1985, giving large samples an extra variance that results from a changing mean within the sample period.
Figure 2 does not show an increased variance however, although the distinction should be made between variance of the bulk of the data, and variance of the extremes. It can not be proven that any of these two variances have increased, while figure 2 highly supports the idea that the variance of the bulk of the data has not changed.

The shape parameter $\xi$ in attachment III shows a very chaotic pattern. Large fluctuations are visible, and, as standard errors are large, any apparent changes seem within the standard error bounds, and thus insignificant. No conclusions on systematic changes of $\xi$ can be drawn from this graph, the only conclusion being, that $\xi$ is a parameter that is hard to estimate accurately.

### 3.5 Poisson representations

First of all, it is tested how well the data responds to a representation of the Poisson parameter $\lambda$, as it would be in the models above, using equation (10). For this, $\lambda$ values are calculated for different thresholds, and compared to $\lambda$ values that are deduced from observed exceedances of that threshold in the reference period. The modeled $\lambda$ value is deduced from the GEV parameters obtained from the point process fit to 1910-1985 data. Observed $\lambda$ values are of course from the reference period as well. A graph of this is plotted in figure 13.

![Observed and modeled values of Lambda](image)

**Figure 13:** modeled (black) and observed(grey, dotted) $\lambda$ values for different thresholds, transformed data (DATA2)

From these graphs, it appears that every threshold above 0.5 would produce a good estimate of $\lambda$. Now, estimates can be made of the chance of occurrence of the number of exceedances seen in the 12-month period in the end of the data, between May 2006 and April 2007, for different thresholds and for different GEV parameter values. Estimates will be based on the transformed data, using a $\lambda$ that has been deduced from GEV parameter estimations. For the results, see table 4. We can take into account an estimated overall warming since 1985, by shifting the location parameter $\mu$ to the value of 2.502, as estimated for 2006 earlier. Using the new location parameter in the model from table 4, the values in table 5 arise. Attachment III gives some evidence of an increased $\sigma$ parameter since 1985. This value can not be easily extrapolated from the figure. Therefore, an estimated new $\sigma$ value of 0.638 has been taken, obtained from a point process fit to the 1985-2007 period. See table 6.
Table 4: Poisson modeled return periods of apr 2006-apr 2007, reference model $\mu=1.381$, $\sigma=0.531$ and $\xi=-0.311$

<table>
<thead>
<tr>
<th>threshold</th>
<th>expected</th>
<th>realised</th>
<th>return period (y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>3.80</td>
<td>11.00</td>
<td>1670</td>
</tr>
<tr>
<td>0.75</td>
<td>2.76</td>
<td>11.00</td>
<td>3.03E+04</td>
</tr>
<tr>
<td>1.00</td>
<td>1.93</td>
<td>11.00</td>
<td>1.04E+06</td>
</tr>
<tr>
<td>1.25</td>
<td>1.29</td>
<td>10.00</td>
<td>7.98E+06</td>
</tr>
<tr>
<td>1.50</td>
<td>0.76</td>
<td>10.00</td>
<td>8.88E+08</td>
</tr>
<tr>
<td>1.75</td>
<td>0.46</td>
<td>10.00</td>
<td>2.88E+11</td>
</tr>
<tr>
<td>2.00</td>
<td>0.24</td>
<td>6.00</td>
<td>1.56E+08</td>
</tr>
<tr>
<td>2.25</td>
<td>0.10</td>
<td>6.00</td>
<td>6.35E+10</td>
</tr>
<tr>
<td>2.50</td>
<td>0.03</td>
<td>5.00</td>
<td>1.40E+12</td>
</tr>
<tr>
<td>2.75</td>
<td>0.00</td>
<td>4.00</td>
<td>2.65E+14</td>
</tr>
<tr>
<td>3.00</td>
<td>0.00</td>
<td>4.00</td>
<td>$\infty$</td>
</tr>
<tr>
<td>3.25</td>
<td>0.00</td>
<td>3.00</td>
<td>$\infty$</td>
</tr>
<tr>
<td>3.50</td>
<td>0.00</td>
<td>2.00</td>
<td>$\infty$</td>
</tr>
<tr>
<td>3.75</td>
<td>0.00</td>
<td>2.00</td>
<td>$\infty$</td>
</tr>
<tr>
<td>4.00</td>
<td>0.00</td>
<td>2.00</td>
<td>$\infty$</td>
</tr>
<tr>
<td>4.25</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 5: Poisson modeled return periods of apr 2006-apr 2007, model $\mu=2.502$, $\sigma=0.531$ and $\xi=-0.311$

<table>
<thead>
<tr>
<th>threshold</th>
<th>expected</th>
<th>realised</th>
<th>return period (y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>11.62</td>
<td>11.00</td>
<td>2</td>
</tr>
<tr>
<td>0.75</td>
<td>9.35</td>
<td>11.00</td>
<td>4</td>
</tr>
<tr>
<td>1.00</td>
<td>7.39</td>
<td>11.00</td>
<td>14</td>
</tr>
<tr>
<td>1.25</td>
<td>5.73</td>
<td>10.00</td>
<td>31</td>
</tr>
<tr>
<td>1.50</td>
<td>4.35</td>
<td>10.00</td>
<td>192</td>
</tr>
<tr>
<td>1.75</td>
<td>3.20</td>
<td>10.00</td>
<td>2001</td>
</tr>
<tr>
<td>2.00</td>
<td>2.28</td>
<td>6.00</td>
<td>112</td>
</tr>
<tr>
<td>2.25</td>
<td>1.56</td>
<td>6.00</td>
<td>879</td>
</tr>
<tr>
<td>2.50</td>
<td>1.00</td>
<td>5.00</td>
<td>1651</td>
</tr>
<tr>
<td>2.75</td>
<td>0.60</td>
<td>4.00</td>
<td>2506</td>
</tr>
<tr>
<td>3.00</td>
<td>0.32</td>
<td>4.00</td>
<td>4.35E+04</td>
</tr>
<tr>
<td>3.25</td>
<td>0.15</td>
<td>3.00</td>
<td>5.40E+04</td>
</tr>
<tr>
<td>3.50</td>
<td>0.05</td>
<td>2.00</td>
<td>4.27E+04</td>
</tr>
<tr>
<td>3.75</td>
<td>0.01</td>
<td>2.00</td>
<td>5.09E+06</td>
</tr>
<tr>
<td>4.00</td>
<td>0.00</td>
<td>2.00</td>
<td>2.71E+11</td>
</tr>
<tr>
<td>4.25</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 6: Poisson modeled return periods of apr 2006-apr 2007, model μ=2.502, σ=0.638 and ξ=-0.311

<table>
<thead>
<tr>
<th>threshold</th>
<th>expected</th>
<th>realised</th>
<th>return period (y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>8.62</td>
<td>11.00</td>
<td>6</td>
</tr>
<tr>
<td>0.75</td>
<td>7.06</td>
<td>11.00</td>
<td>18</td>
</tr>
<tr>
<td>1.00</td>
<td>5.71</td>
<td>11.00</td>
<td>70</td>
</tr>
<tr>
<td>1.25</td>
<td>4.54</td>
<td>10.00</td>
<td>141</td>
</tr>
<tr>
<td>1.50</td>
<td>3.54</td>
<td>10.00</td>
<td>888</td>
</tr>
<tr>
<td>1.75</td>
<td>2.71</td>
<td>10.00</td>
<td>8161</td>
</tr>
<tr>
<td>2.00</td>
<td>2.01</td>
<td>6.00</td>
<td>213</td>
</tr>
<tr>
<td>2.25</td>
<td>1.45</td>
<td>6.00</td>
<td>1314</td>
</tr>
<tr>
<td>2.50</td>
<td>1.00</td>
<td>5.00</td>
<td>1656</td>
</tr>
<tr>
<td>2.75</td>
<td>0.66</td>
<td>4.00</td>
<td>1657</td>
</tr>
<tr>
<td>3.00</td>
<td>0.41</td>
<td>4.00</td>
<td>1.53E+04</td>
</tr>
<tr>
<td>3.25</td>
<td>0.23</td>
<td>3.00</td>
<td>1.08E+04</td>
</tr>
<tr>
<td>3.50</td>
<td>0.11</td>
<td>2.00</td>
<td>4800</td>
</tr>
<tr>
<td>3.75</td>
<td>0.04</td>
<td>2.00</td>
<td>7.71E+04</td>
</tr>
<tr>
<td>4.00</td>
<td>0.01</td>
<td>2.00</td>
<td>4.66E+06</td>
</tr>
<tr>
<td>4.25</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
</tr>
</tbody>
</table>

In the above tables, very high return periods are found (in general many thousands of years), also in distributions that account for a likely climate change since 1985. Therefore, it is tried, which additional shift in the location parameter, again using σ=0.638 and ξ=-0.311, would lower return periods for all of the threshold exceedance frequencies to less than 1000 years. This would happen for μ=3.185. For this μ, return periods for most thresholds are reduced to several decades, the highest return period (at threshold 4) being 1000 years.

On the basis of the yearly maximum of the years 2006 and 2007, also an estimate can be made of the return period of the 2006-2007 sequence of extremes. The year maximum of the year 2006 is 4.06, for july 2006. The year maximum for 2007 is at least 4.21, for april 2007, which is the last month of 2007 in the dataset. On the basis of the 1910-1985 fit, the return period of these values would both be not defined, as the distribution is bounded, and gives p=0 for all values greater than 3. On the basis of the fit shifted to the warm side, with μ=2.502, the return periods would be 42000 years for the 2006 year maximum, and would still not be defined for the 2007 year maximum. When also the scale parameter is raised to σ=0.638, return periods reduce to 144 respectively 675 years.

The outcome of the procedure followed to obtain tables 4, 5 and 6 is remarkable. The hypothesis of a '2006 climate jump' can not be rejected by this approach, as it can not explain the May 2006-April 2006 high temperature extremes from a linear warming trend, even when an increased climatic variance is considered. The same holds for the outcome of the procedure followed to assess the 2006 and 2007 year maxima. The method followed, does however not allow for hard conclusions about this.
5. Conclusions and recommendations

A time series of estimates of the distribution of high temperature extremes in month temperatures from 1951-2007 has been obtained (see attachment III). From this, it can be concluded that recent climate warming has induced a strong increase in the location parameter of high temperature extreme distributions, corresponding to a shift in the mean. Also, some evidence has been found for an increase in the scale parameter of these distributions, which would correspond to an increase in variability. No evidence for any systematic trend has been found in the shape parameter, which is the parameter that determines the characteristics of shape and boundedness of the distribution. Furthermore, it has been found that data transformed as in this study can be described much better by extreme value distributions than data that has not been transformed (still incorporating the seasonality in variance and distribution skewness). The use of a variable threshold does not improve the model fit significantly. This might be caused by some inability of variable thresholds to account for a seasonal variance and skewness, and also by the problem of non-identical distribution the non-transformed month temperatures, and thus the more fundamental constraint of the invalidity of a GEV distribution.

From the trends found above, estimations have been made of extreme value distributions before, and after recent climate warming. These have been used to attempt to explain the high month temperatures that occurred between May 2006 and April 2007. From this, it can be concluded that these high temperature extremes might be extremely rare in the estimated extremal distributions, even incorporating recent climate warming. This raises some evidence on the idea, that extreme value distributions as obtained from measurements in the 20th century, might have changed quite fundamentally in the past years. This could be caused by non-linear effects in the climate system, for example a change or shift in general circulation. Because the methods used, and the significance of the time period assessed, do not allow for hard, more than indicative conclusions, the above implications remain uncertain.

It could be investigated, if the transformation approach used in this study could be applied in other studies on temperature extremes, because it has several advantages. It makes research on extremes more convenient, as no longer 12 different, but one homogeneous time series can be considered. Furthermore, conclusions on the transformed time series can be transformed back into conclusions per month, which can be considered as obtained by using information from all months. This way, it might be possible to obtain a model for month temperatures for every month, having used much more data than the observed month temperatures from that month only. This also broadens the field of possibilities for investigating the response of extremes to future climate changes.

A final recommendation would be, to put some effort in a further analysis of the high temperature extremes in the period after May 2006. Using more sophisticated methods (and, possibly, a more interdisciplinary view) than were used in this study, it might be interesting to further assess what implications this succession of extremely high month temperatures could have for the present state of the climate in The Netherlands or Europe.
References

KNMI, 2006(a): Uitzonderlijk warm minder zeldzaam geworden, KNMI press release KS 06/24

KNMI, 2007: Warme winter voorbode toekomst: nieuwe berekeningen, KNMI press release KS 07/15

Van Oldenborgh, 2006: Juli en september beide warmste in zeker drie eeuwen, KNMI

KNMI, date unknown: Gehomogeniseerde reeks maandtemperaturen De Bilt


Van de Putte, 2006: The influence of the creation of lake Kariba on local precipitation (MSc Thesis, Wageningen UR)

KNMI, 2006(b): Kleine Dstijl (online ‘Nader verklaard’ article)

KNMI, 2003: De toestand van het klimaat in Nederland 2003 (public report)

Peltier, Wilcox and Sharp, 1998: Technical Note: Application of the Box-Cox Data Transformation to Animal Science Experiments

Reiss and Thomas, 2001: Statistical analysis of extreme values : from insurance, finance, hydrology, and other fields, Birkhäuser


Gilleland and Katz, 2006: Analyzing seasonal to interannual extreme weather and climate variability with the extRemes toolkit, Research Applications Laboratory, National Center for Atmospheric Research

Katz, Parlange and Naveau, 2002: Statistics of extremes in hydrology

URL 1: Box-Cox Transformations on Linear Regression, KU Leuven Java applet showing Box-Cox transformations: http://ucs.kuleuven.be/java/version2.0/Applet015.html (viewed 26-6-2007)


Attachment I

trans.txt, script for finding Box-Cox transformations and applying them to the data, by using the functions boc.cox.powers and bc from the car toolkit.

data<-read.table("afwijken positief19101985.txt", header=TRUE)
source("scripts/maanden.txt")
library(car)

pjan<-box.cox.powers(jan)
pfeb<-box.cox.powers(feb)
pmar<-box.cox.powers(mar)
papr<-box.cox.powers(apr)
pmay<-box.cox.powers(may)
pjun<-box.cox.powers(jun)
pjul<-box.cox.powers(jul)
paug<-box.cox.powers(aug)
psep<-box.cox.powers(sep)
poct<-box.cox.powers(oct)
pnov<-box.cox.powers(nov)
pdec<-box.cox.powers(dec)

jan2<-bc(jan,pjan$lambda)
feb2<-bc(feb,pfeb$lambda)
mar2<-bc(mar,pmar$lambda)
apr2<-bc(apr,papr$lambda)
may2<-bc(may,pmay$lambda)
jun2<-bc(jun,pjun$lambda)
jul2<-bc(jul,pjul$lambda)
aug2<-bc(aug,paug$lambda)
sep2<-bc(sep,psep$lambda)
 oct2<-bc(oct,poct$lambda)
nov2<-bc(nov,pnov$lambda)
dec2<-bc(dec,pdec$lambda)

logmatrix<-matrix(nrow=12,ncol=3)

logmatrix[1,1]<-mean(jan2)
logmatrix[1,2]<-sd(jan2)
logmatrix[1,3]<-pjan$lambda
logmatrix[2,1]<-mean(feb2)
logmatrix[2,2]<-sd(feb2)
logmatrix[2,3]<-pfeb$lambda
logmatrix[3,1]<-mean(mar2)
logmatrix[3,2]<-sd(mar2)
logmatrix[3,3]<-pmar$lambda
logmatrix[4,1]<-mean(apr2)
logmatrix[4,2]<-sd(apr2)
logmatrix[4,3]<-papr$lambda
logmatrix[5,1]<-mean(may2)
logmatrix[5,2]<-sd(may2)
logmatrix[5,3]<-pmay$lambda
logmatrix[6,1]<-mean(jun2)
logmatrix[6,2]<-sd(jun2)
logmatrix[6,3]<-pjun$lambda
logmatrix[7,1]<-mean(jul2)
logmatrix[7,2]<-sd(jul2)
logmatrix[7,3]<-pjul$lambda
logmatrix[8,1]<-mean(aug2)
logmatrix[8,2]<-sd(aug2)
logmatrix[8,3]<-paug$lambda
logmatrix[9,1]<-mean(sep2)
logmatrix[9,2]<-sd(sep2)
logmatrix[9,3]<-psep$lambda
logmatrix[10,1]<-mean(oct2)
logmatrix[10,2]<-sd(oct2)
logmatrix[10,3]<-poct$lambda
logmatrix[11,1]<-mean(nov2)
logmatrix[11,2]<-sd(nov2)
logmatrix[11,3]<-pnov$lambda
logmatrix[12,1]<-mean(dec2)
logmatrix[12,2]<-sd(dec2)
logmatrix[12,3]<-pdec$lambda

write.table (logmatrix,"translog.txt")

write.table(jantrans,"jan.txt")
write.table(febtrans,"feb.txt")
write.table(martrans,"mar.txt")
write.table(aprtrans,"apr.txt")
write.table(maytrans,"may.txt")
write.table(juntrans,"jun.txt")
write.table(jultrans,"jul.txt")
write.table(augtrans,"aug.txt")
write.table(septrans,"sep.txt")
write.table(octtrans,"oct.txt")
write.table(novtrans,"nov.txt")
write.table(dectrans,"dec.txt")

l<-length(aprtrans)*12
k<-1
n<-1
j<-1
onderelkaar<-vector(length=l)
while (k<l)
{
    n<-1
    onderelkaar[k]<-jantrans[j]
    onderelkaar[k+1]<-febtrans[j]
    onderelkaar[k+2]<-martrans[j]
    onderelkaar[k+3]<-aprtrans[j]
    onderelkaar[k+4]<-maytrans[j]
    onderelkaar[k+5]<-juntrans[j]
    onderelkaar[k+6]<-jultrans[j]
    onderelkaar[k+7]<-augtrans[j]
    onderelkaar[k+8]<-septrans[j]
    onderelkaar[k+9]<-octtrans[j]
    onderelkaar[k+10]<-novtrans[j]
    onderelkaar[k+11]<-dectrans[j]
    j<-j+1
}
k<-k+12
}
write.table(onderelkaar,"onderelkaar_trans.txt")

**ppfit.txt**, script for performing a moving 15-year point process fit, using the function fpp from the extRemes toolkit

library(extRemes)
data1<-read.table("onderelkaar_trans.txt")
data<-data1$x[1:1876]
nmax<-length(data)
n<-nmax
parameter<-matrix(nrow=nmax,ncol=3)
se<-matrix(nrow=nmax,ncol=3)
nllh<-vector(length=nmax)
code<-vector(length=nmax)
while (n>0)
{
 test<-data[(n-180):n]
v<-fpp(test,0.5,npy=12)
parameter[n,1:3]<-v$mle
se[n,1:3]<-v$se
nllh[n]<-v$nllh
code[n]<-v$conv
}
write.table(parameter[n:nmax,1:3],"mle.txt")
write.table(se[n:nmax,1:3],"se.txt")
write.table(nllh[n:nmax],"nll.txt")
write.table(code[n:nmax],"code.txt")

n<-n-1
}

**maanden.txt**, function for converting the matrix of yearly data by month into monthly vectors

jan<-data$jan
feb<-data$feb
mar<-data$mrt
apr<-data$apr
may<-data$mei
jun<-data$jun
jul<-data$jul
aug<-data$aug
sep<-data$sep
oct<-data$okt
nov<-data$nov
dec<-data$dec
year<-data$yearav
Attachment II

Temperature deviations from 1936-1985 monthly means (DATA1) with centered moving averages, length 24 and 120 months

Attachment II: DATA1 values, plotted with their 24 and 120 months centered moving averages
Attachment III: MLEs of point process GEV parameters moving in time, length 15 years (missing values: cases of failed parameter convergence). Black lines: MLE plus and minus standard error.