university of groningen
faculty of science and engineering
mathematics and applied mathematics

Extreme value statistics of dependent and independent sequences

Bachelor's Project Mathematics

November 2021

Student: X.M. Smid

First supervisor: Dr. A.E. Sterk

Second assessor: Dr.ir. R. Luppes

Abstract

This thesis will be about the branch of mathematics known as extreme value statistics. The partial maximum of sequences of random variables will be considered and the generalized extreme value theorem will be used to study the nature of extreme values of such sequences of random variables. Sequences of independent and dependent random variables will be considered. The sequences of dependent random variables will be generated using certain iterative maps which have some interesting properties. The statistics of the extreme values of dependent sequences will be compared to the statistics of the extreme values of independent sequences. It will be investigated how the statistics of the extreme values of the dependent sequence depend on the way in which the sequence is generated. The block maximum method will be used to study the statistics of extreme values. It will be investigated in which cases the results of the block maximum method converge to the correct value and in some cases this exact value will be calculated analytically.

Contents

1 Introduction

Extreme value statistics is a field of mathematics which revolves around the statistical behaviour of extreme values. Extreme value statistics is a very important branch of mathematics because it leads to a deeper understanding of statistics and probability but also because of the many real life applications where one could be interested in extreme values. These applications could be the weather, geology, finance and many more. It is of course crucial for society to have an understanding of the nature of extreme events in these real life applications. If we consider the weather as an application of extreme value statistics then extreme events could for instance be strong winds or heavy rainfall. Extreme events of the geology of earth could then be earthquakes and tornadoes. The knowledge of how often these things can happen and how extreme they can be is a useful tool for designing safe cities. Having a good understanding of extreme events in finance can be crucial for preventing financial crises from happening. These are only a few of the examples of the real life applications of extreme value statistics but these examples already give an idea of the importance of studying extreme value statistics.

The future can not be predicted which is due to the fact that the data of extreme events in the future for which we are interested in the extremes is random. However by studying available data and making use of the knowledge of extreme value statistics, accurate estimations on the amount at which extreme events happen and the intensity of such extreme events can be made. The data can be of different kinds for different applications. For instance all the randomness of certain aspects in which we are interested in could be independent of previous data or certain connections can exist between the different values. The weather tomorrow is random because it can not be predicted with certainty, however it does depend on the weather that it is today.

The generalized extreme value theorem can be used to investigate the statistical behaviour of extreme values [5]. The generalized extreme value theorem can be applied when the random variables are independent but under certain conditions it can also be applied to dependent random variables [6]. It is however often difficult to check whether these conditions hold for specific cases of dependent random variables. In this thesis multiple different sequences of dependent random variables will be considered and compared to sequences of independent random variables to check if the generalized extreme value theorem applies to these dependent random variables.

2 Theoretical background

2.1 Generalized extreme value theorem

The probability and intensity of extreme values of random events are important to analyse because of the real life applications described earlier. To analyse the probability and intensity of extreme events we need to have data from the field in which we are analysing the extreme values. If for instance one wants to analyse the extreme events of the weather to forecast storms, data is needed of humidity, temperature and many more aspects of the weather. The data can be considered as random variables. The measurements of the data can then be considered as the realisations of those random variables.

Assume that there is no limit on the amount of measurements that can be made and therefore there is also no limit on the amount of random variables. Also assume that each of these infinite random variables $X_1, X_2, \ldots$ are independent of each other and that they are all identically distributed. To analyse the properties of extreme values of this sequence let us consider the partial maximum of this sequence of random variables. The partial maximum is defined to be the following:

Definition 1. *Let $X_1, X_2, \ldots$ be random variables. The partial maximum of order n of this sequence of random variables is defined by:*

$$M_n = \max\{X_1, X_2, \ldots, X_n\}. \tag{1}$$

While this partial maximum is only one maximum out of a finite amount of random variables, it will be useful later to calculate the probability and intensity of extreme values of this sequence. Because the individual X_i are random and not predetermined, the partial maximum will also be a random variable.

For the goal of analysing the probability and intensity of extreme values of the sequence $X_1, X_2, \ldots$ we are in particular interested in the limit of the partial maximum as n approaches infinity. If for example all the random variables X_i are uniformly distributed on the interval $(0, 1)$ and that they are all independent of the other random variables. This means that each random variable will have a realised value somewhere on the interval $(0, 1)$ and the uniform distribution means that each number in that interval will have an equal likelihood of being the realised value of any of the random variables. The realised value of any of the random variables has no influence on the values that the other random variables will obtain since they are all independently distributed. This means that for any random variable X_i the following cumulative density function holds:

$$P\big(X_i < z\big) = z \quad \forall z \in (0, 1).$$

This function can be used to calculate the cumulative density function of the partial maxima M_n:

$$P\big(M_n < z\big) = P\big(X_1 < z, X_2 < z, X_3 < z, \ldots, X_n < z\big).$$

Because each X_i is independent from the other random variables the probability on the right can be written as a product of the individual probabilities

$$P\big(M_n < z\big) = P\big(X_1 < z\big)P\big(X_2 < z\big)\ldots P\big(X_n < z\big).$$

Because each random variable is identically distributed this can easily be transformed to the following.

$$P\big(M_n < z\big) = \big(P\big(X_1 < z\big)\big)^n = z^n.$$

The limit of this function when n goes to infinity is a discontinuous function:

$$P\big(M_n < z\big) = \left\{ \begin{array}{ll} 1, & \text{for } z = 1 \\ 0, & \text{otherwise} \end{array} \right\}.$$

This is called a degenerate probability distribution because it can not be considered as a standard probability distribution.

To avoid the problem of arriving at a degenerate probability distribution for the limit of the partial maximum we can make use of sequences of normalising constants a_n and b_n for which $a_n(M_n - b_n)$ has a proper nondegenerate probability distribution. Such sequences of normalising constants can be found for partial maxima of independent and identically distributed random variables under rather general conditions. Also in this case such normalising constants can be found.

$$\begin{aligned} P\big(a_n(M_n - b_n) < z\big) &= P\left(M_n - b_n < \frac{z}{a_n}\right) \\ &= P\left(M_n < b_n + \frac{z}{a_n}\right) \\ &= \left(b_n + \frac{z}{a_n}\right)^n. \end{aligned} \tag{2}$$

When one chooses $a_n = n$ and $b_n = 1$ the limit of this probability distribution equals the function $\exp(z)$. Nondegenerate probability distributions for the limit of a rescaled partial maximum can be found for sequences of random variables following a different probability distribution. This is described in the generalized extreme value theorem.

Generalized Extreme Value Theorem. *Let $X_1, X_2, \ldots$ be independent and identically distributed random variables and let $M_n = \max\{X_1, X_2, \ldots, X_n\}$ denote the n'th partial maximum. If there exist two sequences a_n and b_n such that the limit of the rescaled partial maximum $a_n(M_n - b_n)$ has some nondegenerate distribution G:*

$$\lim_{n\to\infty} P(a_n(M_n - b_n) \leq z) = G(z)$$

then G is of the form

$$G(z) = \exp\left\{ -\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}.$$

G is called a generalized extreme value (GEV) distribution. GEV distributions have 3 parameters for location, scale and parameter shape (μ, σ, ξ).

For the proof of this theorem the reader is referred to [5].

The generalized extreme value distribution is a family of functions which contains the Gumbel, Fréchet and Weibull distributions. The generalized extreme value distribution suits one of these 3 distributions depending only on the parameter shape: [9]

- $\xi < 0$, corresponds to the Weibull distribution.
- $\xi = 0$, corresponds to the Gumbel distribution.
- $\xi > 0$, corresponds to the Fréchet distribution.

The values of the three parameters depend on the distribution of the individual $X_i's$ and on the values of the sequences a_n and b_n. In some cases multiple different sequences a_n and b_n are possible. In general different values of a_n and b_n correspond to different values of the location and scale parameter, however the parameter shape will remain the same [1]. For our goal of modelling extreme values the parameter shape is much more important than the location or scale parameters since the parameter shape shows the shape of the tail of the distribution which is related to the probability and intensity of extreme values [1].

The generalized extreme value theorem can be extended to the case where the random variables X_i are not independent [6]. When the sequence of dependent and identically distributed random variables satisfies some necessary conditions, the limit of the partial maximum which is rescaled by some sequences a_n and b_n will also follow a generalized extreme value distribution. The conditions for the generalized extreme value theorem to apply to a dependent sequence of random variables are:

- The correlation between the random variables X_i converge to zero sufficiently fast.
- Exceedances of the realisations of the random variables X_i over certain thresholds are not clustering too much.

This gives an idea of the conditions that dependent sequences of random variables need to satisfy for the limit of the rescaled partial maximum to follow a generalized extreme value distribution.

These conditions are of course not very precisely formulated. Describing the exact conditions will take too much space and will therefore not be done here. When one is interested in learning the exact conditions the reader is referred to [6]. The fact that the generalized extreme value theorem can be extended to cases where the random variables are dependent is very useful for real life applications because data in real life is often not completely independent. If the weather today is rainy and cold then the weather of tomorrow will likely be similar.

2.2 Block maximum method

Given a certain probability density function, it is not immediately obvious how one should calculate the values of the three parameters. For some probability density functions it can be done analytically, however this can be very difficult and time consuming for more complicated probability density functions. An analytical approach to calculate the tail parameter of the generalized extreme value distribution for random variables distributed according to the uniform distribution on the interval $(0, 1)$ is given in the previous chapter. It was shown that the generalized extreme value distribution in this case equals

$$G(z) = \exp\{z\}.$$

This corresponds to the values $\mu = -1$, $\sigma = 1$ and $\xi = -1$.

When one wants to find out the value of the tail parameter of the generalized extreme value distribution of the limit of the partial maximum of random variables following a different probability distribution one first has to find sequences a_n and b_n such that the limit of the partial maximum has a nondegenerate distribution, and then one can calculate the tail parameter. However the way to find this tail parameter will be different for every probability density function that the individual random variables follow. In the case where the random variables follow the uniform distribution on the interval $(0, 1)$ this was not too difficult, but for more complicated probability density functions it will be much more difficult.

In real life applications the probability density functions will be much more complicated than the uniform distribution and the probability density functions will also be different for any real life application. It is not efficient and sometimes even not possible to have to find a way to calculate the tail parameter for every real life application where one could be interested in the probability and intensity of extreme values. This is why general methods have been developed to estimate the value of the tail parameter of the generalized extreme value distribution of the limit of the partial maximum of independent and identically distributed random variables.

The method that we will focus our attention on is called the block maximum method. The block maximum method works as follows: given the probability density function which the independent and identically distributed random variables follow, generate a large amount of realisations of these random variables. Given the sequence of realisations of the random variables the following maxima will be calculated:

$$\begin{aligned} B_1 &= \max\{x_1, x_2, \ldots, x_m\} \\ B_2 &= \max\{x_{m+1}, x_{m+2}, \ldots, x_{2m}\} \\ &\vdots \\ B_k &= \max\{x_{(k-1)m+1}, x_{(k-1)m+2}, \ldots, x_{km}\} \end{aligned} \tag{3}$$

The sequence of the realisations is divided into k blocks of length m. Then the block maximum method will only consider the maximum value in each of the blocks. Using these maxima of each block the block maximum method will calculate the maximum likelihood estimator of the tail parameter of the corresponding generalized extreme value distribution.

The maximum likelihood estimator of a random variable is the value of this variable for which the likelihood of the observed realisation is maximized. When a certain outcome of a random process which depends on the value of a random variable is known and fixed, The likelihood of this outcome given a certain value of this random variable can be considered as a function of the random variable. The maximum likelihood estimator of this random variable is the value for which this function is maximized. So when the true value of the random variable equals this maximum likelihood estimator, then the likelihood of the observed outcome having occurred is greater than it would be for any other value of the random variable.

For a sample of n observations $y = y_1, \ldots, y_n$, the maximum likelihood estimators of the three parameters of the generalized extreme value distribution are determined by maximizing the likelihood function [4]:

$$L_n(y|\mu,\sigma,\xi) = \prod_{i=1}^{n} \left[\frac{1}{\sigma} \left(1+\xi \frac{y_i - \mu}{\sigma} \right)^{-(1/\xi)-1} \right] \exp \left\{ -\sum_{i=1}^{n} \left[\left(1+\xi \frac{y_i - \mu}{\sigma} \right)^{-1/\xi} \right] \right\}. \tag{4}$$

Because the likelihood function has positive range the logarithm of this function can be taken. The logarithm is an increasing function which means that the values which maximize the likelihood function also maximize the logarithm of the likelihood function. This logarithm is called the loglikelihood. The loglikelihood is given by:

$$l_n(y|\mu,\sigma,\xi) = -n\log(\sigma) + \sum_{i=1}^{n} \left[\left(-\frac{1}{\xi} - 1 \right) \log(z_i) - {z_i}^{-1/\xi} \right], \tag{5}$$

where $z_i = 1 + \xi(\frac{y_i - \mu}{\sigma})$. The maximum likelihood estimators of the three parameters can then be found by solving the system of equations which set the first derivatives of this loglikelihood with respect to each parameter equal to zero:

$$\begin{aligned} &\frac{1}{\sigma} \sum_{i=1}^{n} \left[\frac{1 + \xi - {z_i}^{-1/\xi}}{z_i} \right] = 0, \\ &-\frac{n}{\sigma} + \frac{1}{\sigma} \sum_{i=1}^{n} \left[\left(\frac{1 + \xi - {z_i}^{-1/\xi}}{z_i} \right) \left(\frac{y_i - \mu}{\sigma} \right) \right] = 0, \\ &-\frac{1}{\xi^2} \sum_{i=1}^{n} \left[\log(z_i) \big(1 + \xi - {z_i}^{-1/\xi} \big) - \left(\frac{1 + \xi - {z_i}^{-1/\xi}}{z_i} \right) \xi \left(\frac{y_i - \mu}{\sigma} \right) \right] = 0. \end{aligned} \tag{6}$$

The calculations of these maximum likelihood estimators were done using a script called gevfit.m which was written by P. Embrechts, C. Klüppelberg, and T. Mikosch [8]. Matlab was used to run this script. Matlab is the program used to calculate the tail parameters and to plot the graphs which are shown in this thesis. The accuracy of the maximum likelihood estimators will of course depend on both the amount of blocks and on the size of the blocks, this will be further discussed in a later chapter.

When the block maximum method is used on sequences for which the generalized extreme value theorem does not apply (for instance if the random variables are dependent but do not satisfy the necessary conditions for the generalized extreme value theorem to apply to dependent random variables or if the random variables are not identically distributed) then the maximum likelihood estimator of the tail parameter will not correspond to the real tail parameter because that does not exist for those sequences. It is therefore still necessary to check whether the generalized extreme value theorem applies to the random variables one is analysing and whether the estimated value of the tail parameter seems correct.

2.3 Peaks over threshold method

There are other methods than the block maximum method to obtain the values of the three parameters such as the peaks over threshold method. The peaks over threshold method focuses on the realisations of the random variables which are exceeding some given threshold. Although there are some advantages to the peaks over threshold method compared to the block maximum method such as the fact that the peaks over threshold method often allows for more flexibility since it is in some practical situations difficult to change the block size [2], It has the problem of determining a proper threshold. There have been many studies done on finding useful thresholds but effective and stable methods have not been found at this moment [3]. The peaks over threshold method also has the disadvantage that in the usual practical situations the values of the three parameters are estimated with less accuracy than is the case when using the block maximum method for the same amount of realisations of the random variables [2]. For these reasons we will focus on the block maximum method instead of the peaks over threshold method.

2.4 Deterministic dynamical systems

In many real life applications such as the weather the current value gives us the ability to calculate the values in the future. For instance the weather today such as temperature, humidity and whether it is raining or not can be used to predict the weather tomorrow. This same principle of course also holds for other applications such as in finance (risks in for instance insurance) and in geology (earthquakes or the behaviour of ocean waves). For this reason let us consider sequences generated by the iterative map

$$X_{i+1} = f(X_i)$$

for some map f. In real life applications the X_i values will often be vectors with many parameters and the map f will be a very difficult function of those multiple parameters. However we will focus on simple maps f which take only one number as its input. This can seem like too much of a restriction but we will see that even very simple maps can have very unpredictable and chaotic outcomes.

Of course when we know the value of X_0 all the other values of the X_i will also be determined, but when X_0 is treated as a random variable then all the other $X_i's$ will be random. X_0 is treated as a random variable because in real life applications it is impossible to know the current value with infinite precision and given the nature of the maps we will consider even very slight differences in initial value can lead to very big changes after only a few steps.

With the applications in mind it is not difficult to see why one would be interested in the amount and severity of extreme values of the sequence generated by the iterative map. The extreme values could for instance mean dangerous storms, financial crises or giant ocean waves capable of destroying a city. Now that we know why one would want to model the likelihood and severity of extremes of sequences generated by the iterative map described above we arrive at the obvious problem that such sequences are indeed random but not necessarily identically distributed and certainly not independent. To deal with this problem it is important to note that for certain maps there exist invariant distributions. This means that if the input for the map is randomly distributed according to its invariant distribution then the output of the map will be another random variable with the same probability distribution.

Let us consider the beta map: $f(x) = \beta x \bmod 1$ which maps numbers from the interval $(0,1)$ to the same interval for some value of β. It can be shown that the uniform distribution is invariant under the beta map when β is a positive integer. This means that when X_0 is distributed according to the uniform distribution on the interval $(0,1)$, then X_1 will also be uniformly distributed on the interval $(0,1)$. Because the output of the map at any point will be the input of the map at the next moment it can be easily seen that every random variable X_i will be distributed with the uniform distribution on $(0,1)$ for all positive integers i. The fact that the uniform distribution on $(0,1)$ is the invariant distribution of the beta map will be proven in later in the subsection about the beta map.

Different maps will have different invariant distributions. When we let X_0 be distributed according to the invariant distribution of the map which we are using to generate the sequence, then the random variables will of course still not be independent since that is the point of the iterative map but at least they will all be identically distributed. The sequence being dependent does not have to be a problem considering the fact that the Generalized Extreme Value Theorem also holds for dependent random variables as long as certain conditions hold. These conditions hold for most of the maps that we will look at but it will not be proven in this thesis that the conditions hold since that will take too much space in this thesis.

This table shows some maps with their invariant distributions:

Interval I	map $f : I \to I$	Density ρ of the invariant distributions
$[0,1)$	Rotation map $f(x) = \omega + x \bmod 1,\ \omega \in \mathbb{R}$	1
$[0,1)$	Doubling map $f(x) = 2x \bmod 1$	1
$[0,1]$	Tent map $f(x) = 1 - \lvert 2x-1 \rvert$	1
$[0,1)$	Beta map $f(x) = \beta x \bmod 1,\ \beta \in \mathbb{Z}_{>1}$	1
$[-1,1]$	Cusp map $f(x) = 1 - 2\sqrt{\lvert x \rvert}$	$(1-x)/2$
$[0,1]$	Logistic map $f(x) = 4x(1-x)$	$1/(\pi\sqrt{x(1-x)})$
$\mathbb{R}$	Newton map $f(x) = \frac{1}{2}(x - \frac{1}{x})$	$1/(\pi(1+x^2))$

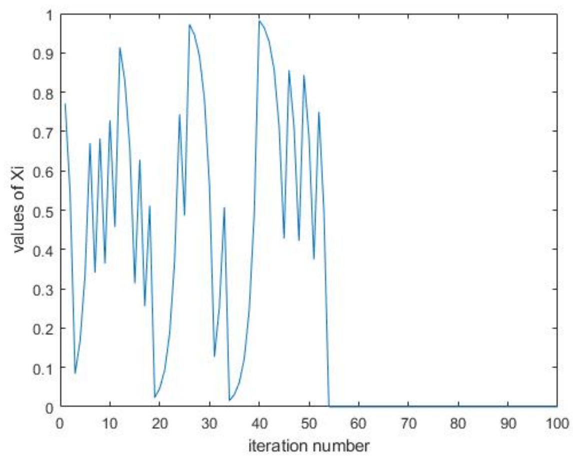

Figure 1: Sequence generated by the doubling map.

Sequences generated by the different maps shown in this table all have different properties. Sequences generated by the doubling map (figure 1) or the tent map (figure 2) oscillate between small and large values. The same is true for the beta map when β is relatively small but for larger β the values of the sequence seem more independent (figure 3). Sequences generated by the cusp map (figure 4) slowly grow until it hits a point at which it will fall down again. The smaller any value is the slower the growth at that point as can be seen in the graph below. The logistic map (figure 5) also has its sequences oscillate just like the doubling map or the tent map, however it oscillates in a different way, namely the farther away any point is from 0, 1, or 0.75 the more it jumps at the next point. The Newton map (figure 6) generates a sequence of numbers in all of $\mathbb{R}$. Most of the realisations lie somewhere close to 0 but sometimes it makes a big jump after which it will again return to 0. In the sequences shown below the starting random variable X_0 is distributed according to the invariant distribution of the map which generates the sequence.

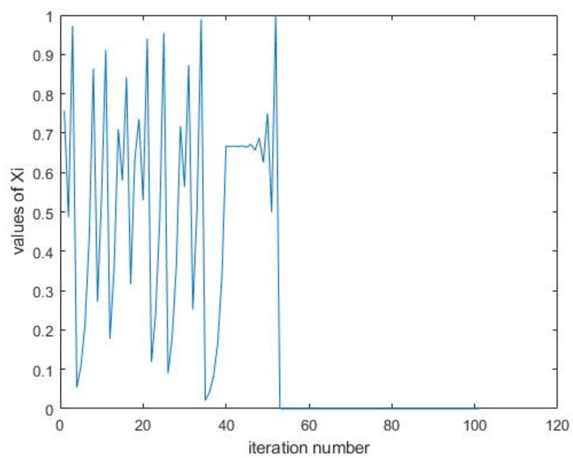

Figure 2: Sequence generated by the tent map.

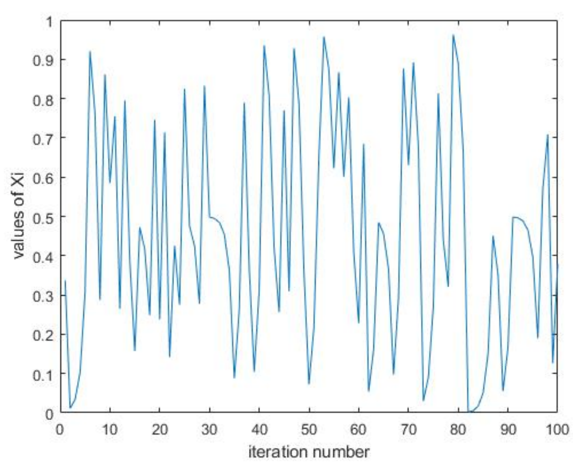

Figure 3: Sequence generated by the beta map with $\beta = 3$.

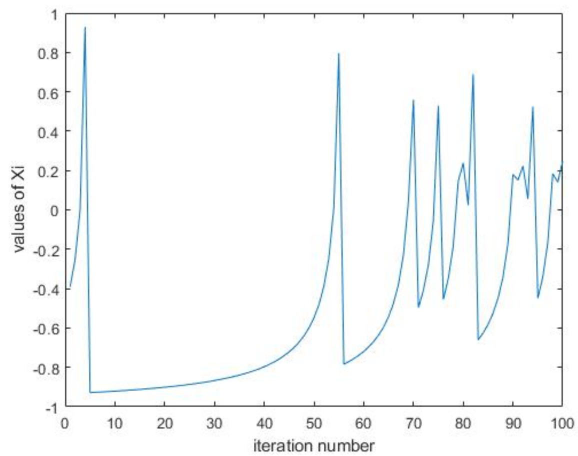

Figure 4: Sequence generated by the cusp map.

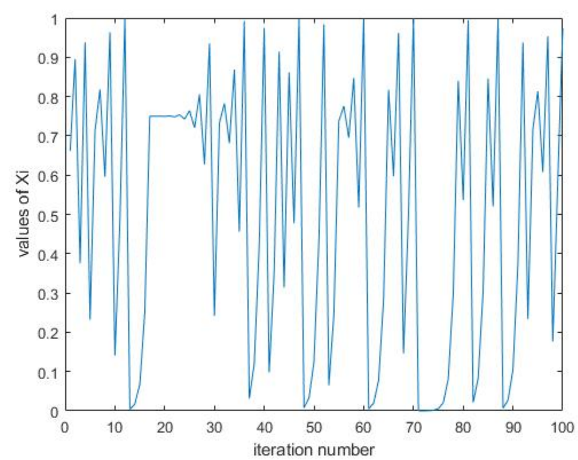

Figure 5: Sequence generated by the logistic map.

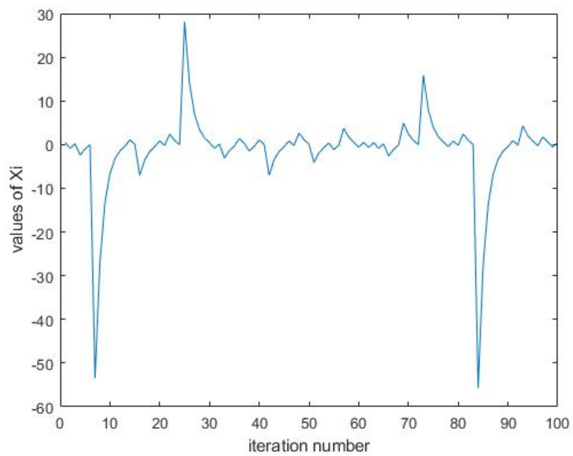

Figure 6: Sequence generated by the Newton map.

It can be seen that sequences generated by the doubling map (figure 1) or tent map (figure 2) will at some point reach 0 after which it will never reach another number. This is not because of the nature of the map but because of a limitation in the calculation software. Matlab is used for plotting these graphs and when a random number is generated in Matlab only a finite number of decimals is generated in base 2. This means that after it has been multiplied by an even number often enough the number will be treated as an integer which is mapped to 0. This same problem applies to the beta map when β is an even number. Therefore we will restrict to the case when β is an odd number.

It is not easy to check for each of these maps if the conditions hold for the generalized extreme value theorem to be applied to dependent sequences. This is why we will apply the block maximum method to dependent sequences generated by each of these maps and compare the results with that of the independent sequences following the invariant distributions of these iterative maps. This way the dependent sequences follow the same distribution as their respective independent sequences and if the dependent sequences satisfy the conditions for the generalized extreme value theorem to be applied to dependent sequences then the results are expected to be similar for the dependent and independent case. Checking if the results for the dependent and independent case are similar is a useful test to help find out which maps do and which do not satisfy the necessary conditions of the generalized extreme value theorem.

3 Case studies

The block maximum method can be used on sequences of random variables to calculate the maximum likelihood estimator of the tail parameter of the generalized extreme value distribution which the limit of the rescaled partial maximum will follow. The value that the block maximum method provides is therefore just an estimator of the value in which we are interested. For many real life applications it is of course important to obtain an estimator of the tail parameter which is as close to the true value as possible, since more accurate estimations give more insight into the likelihood and intensity of extreme values. The more is known about the amount of extreme storms and its intensity the better society can prepare for it. More data means that the block maximum method will likely make a better approximation of the tail parameter.

The block maximum method distributes all the data into different blocks which all have the same amount of data. Therefore more data could mean that each block still has the same amount of data but that there will be more blocks, or it could mean that the amount of blocks remain the same but each block will receive more data. It could also be somewhere in between where there will be more blocks which also all have more data. We will focus on the case where the number of blocks remain the same but the length of the blocks is flexible. This way we can find out if the estimated values converge for larger block lengths and if it converges to the expected value.

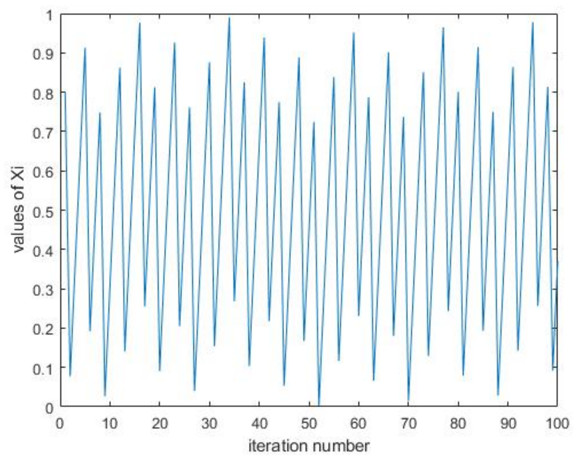

Figure 7: Sequence generated by the rotation map with $\omega = 0.278498218867048$.

3.1 The rotation map

The rotation map is an iterative map as is described in the previous chapter which is as follows

$$f(x) = x + \omega \text{ mod } 1$$

The uniform distribution on the interval $(0, 1)$ is the invariant distribution for this map.

Lemma 1. *The uniform distribution on the interval* $(0, 1)$ *is the invariant distribution of the rotation map described by* $f(x) = x + \omega \ mod\ 1$

Proof. Let X_i be distributed according to the uniform distribution on the interval $(0, 1)$ and let $X_{i+1} = f(X_i)$ where $f(x)$ is the rotation map. then

$$\begin{aligned} P\big(X_{i+1} < z\big) &= P\big(X_i + \omega \text{ mod } 1 < z\big) \\ &= P\big(X_i \text{ mod } 1 < z - \omega\big) + P\big(X_i \text{ mod } 1 > 1 - \omega\big) \\ &= (z - \omega) + \omega \\ &= z. \end{aligned}$$

Therefore X_{i+1} is also distributed according to the uniform distribution on the interval $(0, 1)$ which proves that the uniform distribution on $(0, 1)$ is the invariant distribution of the rotation map. □

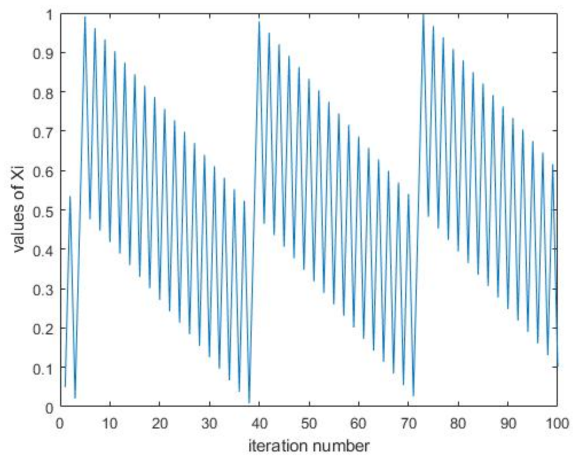

Figure 8: Sequence generated by the rotation map with $\omega = 0.485375648722841$.

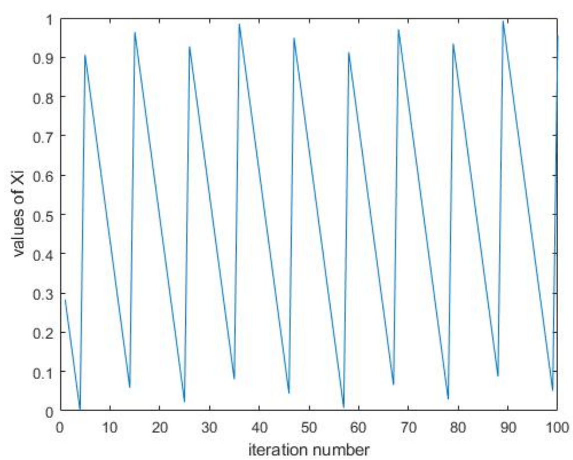

Figure 9: Sequence generated by the rotation map with $\omega = 0.905791937075619$.

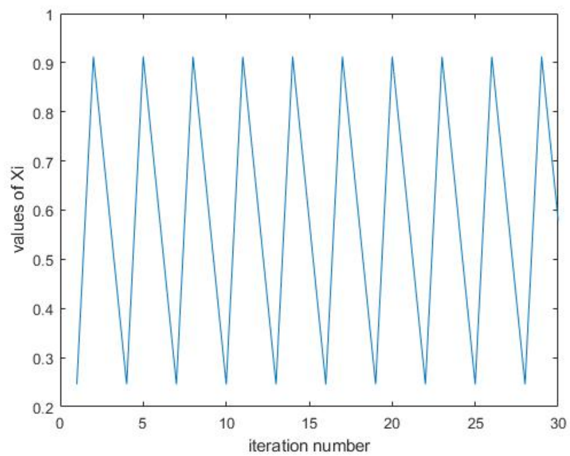

Figure 10: Sequence generated by the rotation map with $\omega = 2/3$.

The rotation map seems similar to the beta map but it behaves rather differently. The only difference between these two maps is that instead of multiplying the rotation map adds the values together. The maps even have the same invariant distribution. However whereas the beta map generates sequences rather similarly for different values of β, the rotation map is very dependent on the value of ω. When ω is an integer every element in the sequence generated by the rotation map will be the same since we are only looking at the decimal numbers. In this case it obviously does not make much sense to try to model the partial maxima since $M_n = X_0$ for every $n \in N$. For other values of ω there is more randomness but there can still be some weird sequences generated by it.

When ω is some rational number $\frac{p}{q}$ for some integers p and q, after q steps the sequence begins again: $X_q = \frac{p}{q} + X_{q-1} \mod 1 = \frac{p}{q} + \frac{p}{q} + X_{q-2} \mod 1$. Repeating this argument one obtains $X_q = q\frac{p}{q} + X_0 \mod 1 = p + X_0 \mod 1 = X_0$. When n is sufficiently large the partial maximum M_n will be equal to the maximum of the loop and afterwards the partial maximum will never increase again. There is still some randomness to the limit of the partial maximum since the maximum of the loop depends on the value of X_0 but the sequences that are generated in this case are clearly not behaving similarly to sequences of independent random variables which are distributed uniformly on the interval $(0, 1)$.

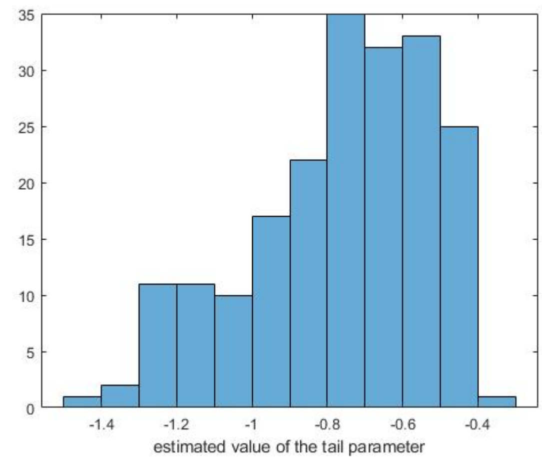

Figure 11: Histogram of the estimated tail parameter values for a sample of random uniform values of omega.

The conditions of the Generalized Extreme Value Theorem are not satisfied for sequences generated by the rotation map. Namely the correlations between the random variables do not converge to zero. This means that the block maximum method will not give good results for the estimation of the parameters because the estimated values of the parameters will depend heavily on the value of ω that is used. This map can therefore not be used to model the tail parameter of the generalized extreme value distribution for sequences of independent random variables distributed uniformly on the interval $(0, 1)$. This can be seen in 12.

The histogram of figure 11 shows the estimated values of the tail parameters for 200 sequences generated by the rotation map, all with a different random value of ω. The values were estimated using the block maximum method with a block length of 200 and an amount of 1000 blocks. Because the random values of ω were generated with a finite number of decimals they are technically all rational numbers however for practical purposes, because the amount of decimals that were generated was so large, they can be regarded as irrational numbers.

The histogram of figure 11 shows that the block maximum method estimates the tail parameters for all of the sequences to be smaller than -0.3. However when one generates sequences with the rotation map for rational values of ω with a small denominator, the values of the tail parameters are estimated to be very close to 0 and almost always bigger than -0.3. All values of ω will cause the uniform distribution on $(0, 1)$ to be the invariant distribution of the rotation map, but all values of ω will also cause the block maximum method to estimate very different tail parameters based on the value of ω.

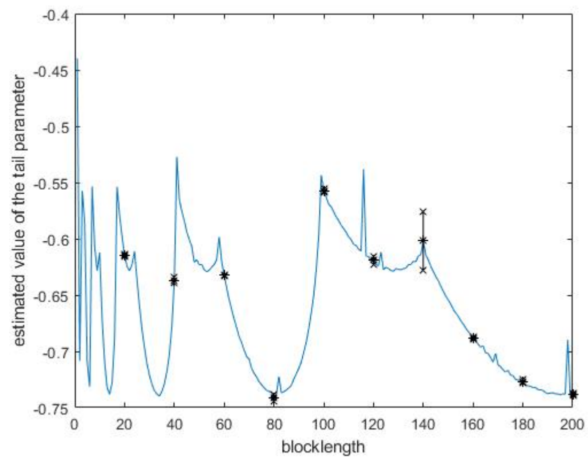

Figure 12: Estimated value of the tail parameter for different blocklengths.

Figure 12 plots the estimated value of the tail parameter against the length of the blocks where the amount of blocks is fixed to equal 1000. 20 different graphs of the estimated value of the tail parameter for different blocklengths were created. The blue line in figure 12 represents the average value of those 20 different graphs for that given blocklength and the vertical black lines show the 95% confidence interval of the maximum likelihood estimator of the tail parameter. Each sequence of random variables is generated by the rotation map with a value of $\omega = \frac{1}{2}\sqrt{2}$. It can be seen that the estimated value of the tail parameter depends heavily on the length of the blocks and it is obvious that the estimation does not converge to any single value. The fact that this is true for the average of 20 different graphs shows that the estimated value of the tail parameter never converges when using sequences generated by the rotation map and that this is not simply a bad result happening once. This shows that the generalized extreme value theorem does not apply to sequences generated by the rotation map.

The actual tail parameter for independent sequences generated by the uniform distribution on $(0, 1)$ equals -1 which was calculated in 2. However only few of the dependent sequences generated by the rotation map will cause the block maximum method to estimate the tail parameter to equal -1. This all shows that the rescaled limit of the partial maximum of dependent sequences generated by the rotation map does not follow a generalized extreme value distribution. The block maximum method therefore does not work for this specific case.

3.2 The beta map

The beta map is an iterative map which was described in an earlier section. The beta map is defined as follows:

$$f(x) = \beta x \text{ mod } 1.$$

It was discussed in an earlier section that the uniform distribution on the interval $(0, 1)$ is the invariant distribution of the beta map, this will now be proven:

Lemma 2. *The uniform distribution on the interval* $(0, 1)$ *is the invariant distribution of the beta map described by* $f(x) = \beta x$ *mod* 1

Proof. Let X_0 be a random variable which is uniformly distributed on the interval $(0, 1)$ and let $X_1 = f(X_0)$ where f is the beta map.

$$\begin{aligned} P\big(X_1 < z\big) &= P\big(\beta X_0 \text{ mod } 1 < z\big) \\ &= \sum_{k=0}^{\beta-1} P\big(\frac{k}{\beta} < X_0 < \frac{k+z}{\beta}\big) \\ &= \sum_{k=0}^{\beta-1} \frac{z}{\beta} \\ &= z. \end{aligned} \tag{7}$$

Therefore X_1 is also distributed according to the uniform distribution on the interval $(0, 1)$. □

The program Matlab was used to generate sequences of dependent random variables generated by the beta map. Matlab generates random numbers with only a finite amount of decimals in base 2. This has the unfortunate effect that after a sufficient amount of multiplication of this randomly generated number with an even number, No decimals in base 2 will remain and the resulting number will be an integer. Any integer mod 1 will equal 0 and 0 will be mapped to itself under the beta map. This means that after sufficient iterations with the beta map when β is an even number any randomly generated number will reach the value zero after which it will never reach another value.

Such a sequence which only has a finite amount of nonzero elements is not useful for our purpose of analysing the properties of extreme values of random sequences. For this reason we will focus on the case when $\beta = 3$. This removes the problem of arriving at a never ending sequence of zeros. The beta map does not have similar problems to the ones in the case of the rotation map, it will not generate periodic sequences. For this reason the estimated value of the tail parameter using the block maximum method will behave more appropriately as it can be seen that for this value of β the estimated value of the tail parameter converges when the blocklength increases.

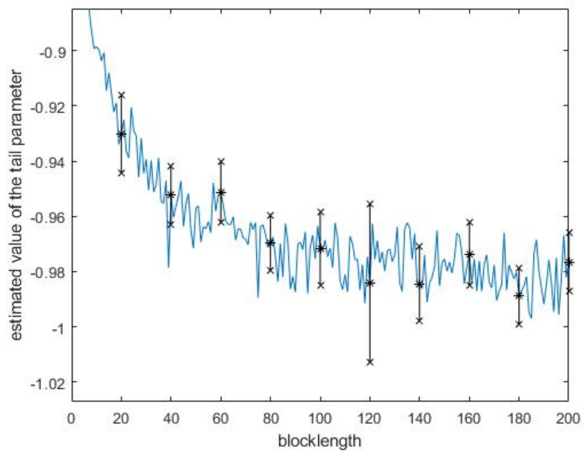

Figure 13: Estimated value of the tail parameter using the beta map for different blocklengths.

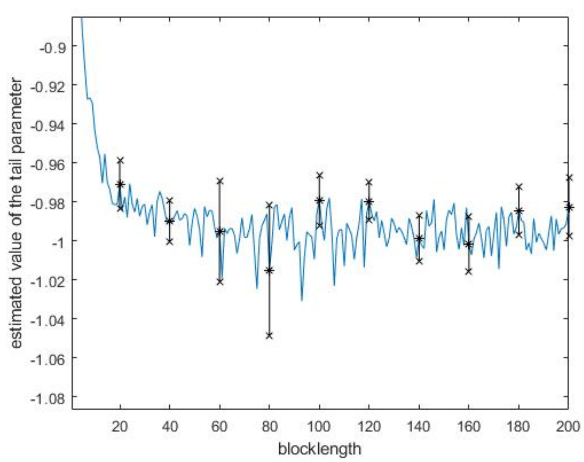

Figure 14: Estimated value of the tail parameter using independent uniformly generated random variables for different blocklengths.

In figure 13 the estimated value of the tail parameter can be seen for sequences generated by the beta map and in figure 14 the tail parameter is estimated for sequences of independent uniformly distributed random variables. In both figures 1000 blocks were used. The two figures look very similar which is a good indication that sequences generated by the beta map satisfy the conditions for the generalized extreme value theorem to apply to dependent sequences. These sequences can therefore be used to estimate the tail parameter of the generalized extreme value distribution of the rescaled limitting partial maximum of independent uniformly distributed random variables. In the case of the dependent sequence only one random number has to be generated which is then iterated many times in a simple map. In the case of the independent sequence many random numbers need to be generated. Because it is quicker to calculate the output of a simple function than it is to generate a random number, using sequences generated by the beta map is a more efficient way of estimating the tail parameter of a sequence of uniformly distributed random variables than using independent random variables.

3.3 The cusp map

The cusp map is an iterative map which is defined as follows:

$$f(x) = 1 - 2\sqrt{|x|}.$$

The invariant distribution for this map is the following function on the interval $[-1, 1]$:

$$\rho(x) = \frac{1-x}{2}.$$

The proof that this function is the invariant distribution of the cusp map is a bit more complicated than the similar proofs for the other maps which we have looked at so far but it is still a fairly simple proof.

Lemma 3. *The distribution given by* $\rho(x) = \frac{1-x}{2}$ *is the invariant distribution of the cusp map given by* $f(x) = 1 - 2\sqrt{|x|}$

Proof. Let X_0 be a random variable distributed according to the distribution $\rho(x) = \frac{1-x}{2}$. This means that

$$\begin{aligned} P\big(X_0 < z\big) &= \int_{-1}^{z} \frac{1-x}{2}\, dx \\ &= -\frac{1}{4}z^2 + \frac{1}{2}z + \frac{3}{4}. \end{aligned}$$

Now let X_1 be defined to be $f(X_0)$ where f is the cusp map. Then

$$\begin{aligned} P\big(X_1 < z\big) &= P\big(1 - 2\sqrt{|X_0|} < z\big) \\ &= P\left(|X_0| > \left[\frac{1-z}{2}\right]^2\right) \\ &= P\left(X_0 < -\left[\frac{1-z}{2}\right]^2\right) + P\left(X_0 > \left[\frac{1-z}{2}\right]^2\right) \\ &= -\frac{1}{4}z^2 + \frac{1}{2}z + \frac{3}{4} \\ &= P\big(X_0 < z\big). \end{aligned}$$

X_1 has therefore the same probability distribution as X_0 which shows that $\rho(x)$ is the invariant distribution of the cusp map. ☐

When X_0 is distributed according to this invariant distribution then all the random variables will be identically distributed. To generate a random variable according to this distribution, one can make use of generalized inverse functions [7].

Definition 1. *For an increasing function* $T : \mathbb{R} \to \mathbb{R}$ *with* $T(\infty) = \lim_{x\to-\infty} T(x)$ *and* $T(\infty) = \lim_{x\to\infty} T(x)$*, the generalized inverse* $T^- : \mathbb{R} \to [-\infty, \infty]$ *of* T *is defined by*

$$T^-(y) = \inf\{x \in \mathbb{R} : T(x) \geq y\}, y \in \mathbb{R}, \tag{8}$$

with the convention that $\inf\{\} = \infty$. *If* $T : \mathbb{R} \to [0, 1]$ *is a distribution function,* $T^- : [0, 1] \to [-\infty, \infty]$ *is also called the quantile function of* T.

With this definition we can make use of an important proposition.

Proposition 1. *Let* F *be a distribution function and* $X \sim F$.

1. *if* F *is continuous,* $F(X) \sim U[0, 1]$.
2. *if* $U \sim U[0, 1], F^-(U) \sim F$

When the generalized inverse function of the distribution function of F is known then we can simply generate a uniformly distributed random variable and use it as the input for the generalized inverse function of F to generate a random variable which is distributed according to the distribution function F.

Note that since F is an increasing function on the interval $[-1, 1]$, the generalized inverse F^{-} is simply the inverse function of F. Let u be the realisation of a random variable following the uniform distribution on the interval $(0, 1)$ and let $x = F^{-}(u)$, then $u = F(x)$ and x is the realisation of a random variable following the distribution function F.

$$P\big(X < x\big) = F(x) = -\frac{1}{4}x^2 + \frac{1}{2}x + \frac{3}{4} = u,$$
$$-\frac{1}{4}x^2 + \frac{1}{2}x + \big(\frac{3}{4} - u\big) = 0.$$

We can use the quadratic formula to calculate x as a function of u.

$$x = 1 \pm 2\sqrt{1-u},$$

where x must lie somewhere on the interval $[-1, 1]$. This means that the inverse function of F equals:

$$F^{-}(u) = 1 - 2\sqrt{1-u}.$$

Now that we can generate random variables which follow the invariant distribution of the cusp map we can compare sequences of dependent random variables generated by the cusp map with sequences of independent random variables which follow the same distribution. Figures 15 and 16 show the estimated value of the tail parameter for different blocklengths for the dependent case and the independent case. In both cases 1000 blocks were used. It can be seen that the results of the two graphs are very different. In the case where all the random variables are independent the value of the tail parameter is estimated to be close to -0.5 while in the case where the sequences are generated with the cusp map the tail parameter is estimated to be around -1.0. Another clear difference is the fact that in the dependent case the estimated value of the tail parameter is not stable which seems to imply that the block maximum method does not produce a good estimation of the tail parameter.

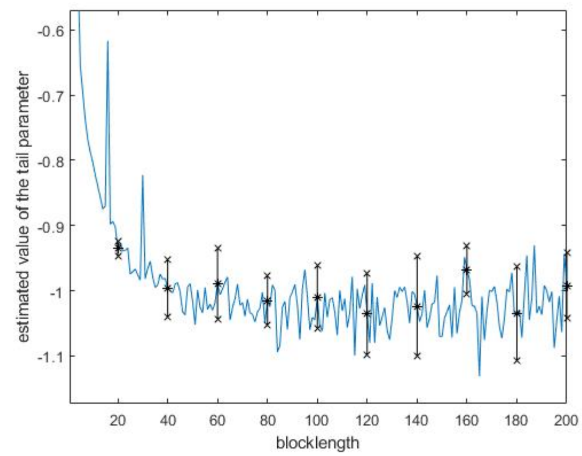

Figure 15: Estimated value of the tail parameter using sequences generated by the cusp map for different blocklengths.

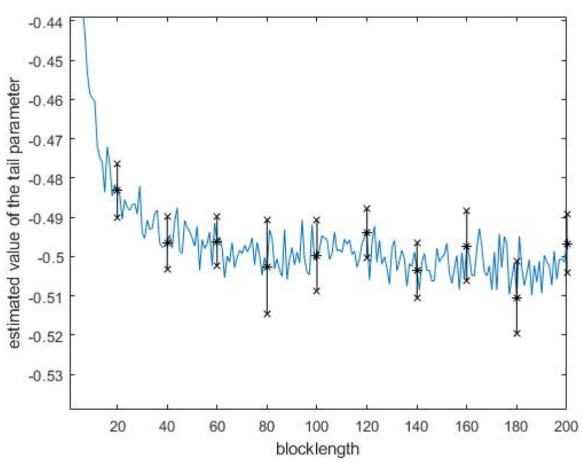

Figure 16: Estimated value of the tail parameter using independent random variables following the invariant distribution of the cusp map for different blocklengths.

It can be shown analytically that the tail parameter of independent sequences of random variables distributed according to the invariant distribution of the cusp map equals -0.5. This can be done in the same way as in equation 2:

$$\begin{aligned}P\big(a_n(M_n - b_n) < z\big) &= P\left(M_n < \frac{z}{a_n} + b_n\right)\\ &= \left[P\left(X_1 < \frac{z}{a_n} + b_n\right)\right]^n\\ &= \left[-\frac{1}{4}\left(\frac{z}{a_n} + b_n\right)^2 + \frac{1}{2}\left(\frac{z}{a_n} + b_n\right) + \frac{3}{4}\right]^n\\ &= \left[\left(-\frac{1}{4a_n^2}\right)z^2 + \left(\frac{1}{2a_n} - \frac{b_n}{2a_n}\right)z + \left(-\frac{1}{4}b_n^2 + \frac{1}{2}b_n - \frac{1}{4}\right)\right]^n.\end{aligned} \tag{9}$$

choosing $a_n = \frac{1}{2}\sqrt{n}$ and $b_n = 1$ gives:

$$P\left(\frac{1}{2}\sqrt{n}(M_n - 1) < z\right) = \left[1 + \frac{-z^2}{n}\right]^n.$$

The limit of this expression when n tends to infinity equals $\exp\{-z^2\}$. This corresponds to the values of the parameters $\mu = 1$, $\sigma = -\frac{1}{2}$ and $\xi = -\frac{1}{2}$. This value of the tail parameter is exactly what would be expected given figure 16. This also shows that the estimated values of the tail parameter using sequences generated by the cusp map do not converge to the true value of the tail parameter for sequences of independent random variables distributed according to the invariant distribution of the cusp map.

These results can be explained when one takes a closer look at the behaviour of the cusp map and sequences generated by it. In the figure 4 it can be seen that sequences generated by the cusp map do not behave as if they were all independent. When the sequence reaches a value close to -1 it tends to stick to such small values for a long time. The periods where the sequence sticks to small values can be arbitrarily long when the values at the start of such a period are sufficiently small. When the block maximum method places the values into the different blocks, there can be blocks where all the values are close to -1 while this is near impossible in the case when all the random variables are independent. This explains why the dependent case and the independent case give such different results.

The generalized extreme value theorem can be extended to the case where the random variables are dependent but then certain conditions must hold. The dependent sequence of random variables must behave similarly to a sequence of independent random variables. This is clearly not the case for sequences generated by the cusp map since small values stick so much together. Therefore using the cusp map is not a good way to estimate the tail parameter of the generalized extreme value distribution for sequences of independent random variables which are all distributed according to the invariant distribution of the cusp map.

3.4 The logistic map

The logistic map is another iterative map with a different invariant distribution. The logistic map is defined by:

$$f(x) = 4x(1 - x).$$

The invariant distribution of the logistic map is given by:

$$\rho(x) = \frac{1}{\pi\sqrt{x(1-x)}}.$$

Lemma 4. *The distribution given by* $\rho(x) = \frac{1}{\pi\sqrt{x(1-x)}}$ *is the invariant distribution of the logistic map given by* $f(x) = 4x(1-x)$.

Proof. Let X_0 be a random variable distributed according to the density function $\rho(x) = \frac{1}{\pi\sqrt{x(1-x)}}$ and let $X_1 = f(X_0)$ where $f(x)$ is the logistic map. Then

$$\begin{aligned} P\big(X_0 < z\big) &= \int_0^z \rho(x)\,dx \\ &= \int_0^z \frac{1}{\pi\sqrt{x(1-x)}}\,dx \\ &= 1 - \frac{2}{\pi}\arcsin\big(\sqrt{1-z}\big), \end{aligned}$$

and

$$\begin{aligned} P\big(X_1 < z\big) &= P\big(f(X_0) < z\big) \\ &= P\big(4X_0(1-X_0) < z\big) \\ &= P\big(X_0(1-X_0) < \frac{1}{4}z\big) \\ &= P\big(-{X_0}^2 + X_0 - \frac{1}{4}z < 0\big). \end{aligned}$$

The quadratic formula can be used to compute the intervals in which X_0 must lie for this condition to hold:

$$
\begin{aligned}
&P\big(-{X_0}^2 + X_0 - \frac{1}{4}z < 0\big) \\
&= P\big(X_0 < \frac{1}{2} - \frac{1}{2}\sqrt{1-z}\big) + P\big(X_0 > \frac{1}{2} + \frac{1}{2}\sqrt{1-z}\big) \\
&= 1 - \frac{2}{\pi}\arcsin\left(\sqrt{1 - (\frac{1}{2} - \frac{1}{2}\sqrt{1-z})}\right) + 1 - \left[1 - \frac{2}{\pi}\arcsin\left(\sqrt{1 - (\frac{1}{2} + \frac{1}{2}\sqrt{1-z})}\right)\right] \\
&= 1 - \frac{2}{\pi}\left[\arcsin\left(\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{1-z}}\right) - \arcsin\left(\sqrt{\frac{1}{2} - \frac{1}{2}\sqrt{1-z}}\right)\right].
\end{aligned}
$$

We can make use of the trigonometric identity
$\arcsin(x) - \arcsin(y) = \arcsin\left(x\sqrt{1-y^2} - y\sqrt{1-x^2}\right)$
with $x = \sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{1-z}}$ and $y = \sqrt{\frac{1}{2} - \frac{1}{2}\sqrt{1-z}}$.
This leads to:

$$
\begin{aligned}
P\big(-{X_0}^2 + X_0 - \frac{1}{4}z < 0\big) &= 1 - \frac{2}{\pi}\arcsin\left(\sqrt{1-z}\right) \\
&= P\big(X_0 < z\big)
\end{aligned}
$$

X_1 has the same distribution as X_0 which proves that $\rho(x) = \frac{1}{\pi\sqrt{x(1-x)}}$ is the invariant distribution of the logistic map. □

To generate random variables following the invariant distribution of the logistic map we need to make use of the generalised inverse again. Since F is a cumulative density function its generalised inverse is again just its inverse.

$$
\begin{aligned}
&P\big(X < x\big) = F(x), \\
&1 - \frac{2}{\pi}\arcsin\left(\sqrt{1-x}\right) = u, \\
&x = 1 - \sin^2\left(\frac{1}{2}\pi[1-u]\right) = F^{-}(u).
\end{aligned}
$$

With the inverse function F^{-} random variables following the invariant distribution of the logistic map can be generated. This allows us to compare the estimated value of the tail parameter for sequences of independent random variables with the estimated value for dependent sequences generated by the logistic map. Figure 17 shows the estimated value of the tail parameter for dependent sequences of random variables generated by the logistic map and figure 18 shows the estimated value of the tail parameter for sequences of independent random variables following the invariant distribution of the logistic map. It can be seen that the two figures look very similar. This suggests that the logistic map satisfies the conditions of the generalized extreme value theorem such that rescaled limitting partial maxima of sequences generated by the logistic map follow a generalized extreme value distribution.

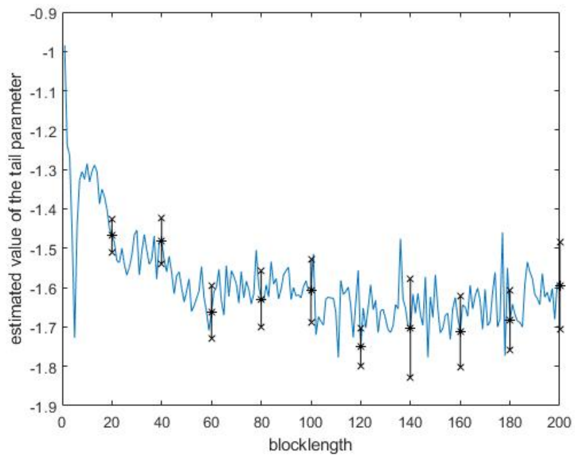

Figure 17: Estimated value of the tail parameter using sequences generated by the logistic map for different blocklengths.

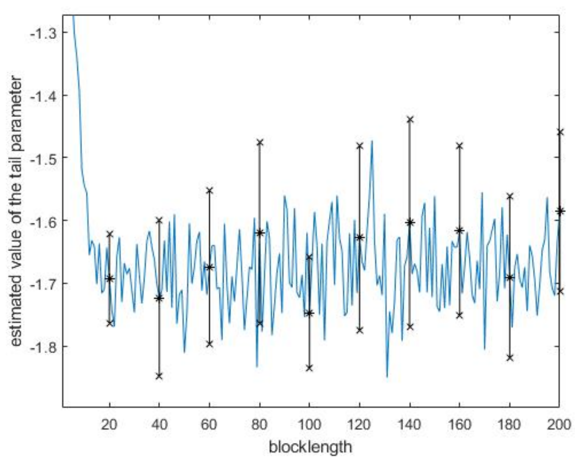

Figure 18: Estimated value of the tail parameter using independent random variables following the invariant distribution of the logistic map for different blocklengths.

3.5 The Newton map

The Newton map is an iterative map which is defined by:

$$F(x) = \frac{1}{2}\big(x - \frac{1}{x}\big)$$

and its invariant distribution is given by:

$$\rho(x) = \frac{1}{\pi(1+x^2)}.$$

The Newton map is what happens when one tries to find the roots of then function $\phi(x) = x^2 + 1$ using the Newton method. The Newton method uses the following iterative map to estimate roots of functions: $X_{n+1} = X_n - \frac{f(X_n)}{f(X_n)}$. When the function of which one wants to find roots is the one one described above then the Newton method consists of the map

$$\begin{aligned} X_{n+1} &= X_n - \frac{\phi(X_n)}{\phi'(X_n)} \\ &= \frac{1}{2}\big(X_n - \frac{1}{X_n}\big) \end{aligned} \tag{10}$$

which is the Newton map.

Lemma 5. *The distribution given by* $\rho(x) = \frac{1}{\pi(1+x^2)}$ *is the invariant distribution of the Newton map given by* $f(x) = \frac{1}{2}(x - \frac{1}{x})$

Proof. Let X_0 be distributed by the distribution function $\rho(x) = \frac{1}{\pi(1+x^2)}$. This means that

$$\begin{aligned} P\big(X_0 < z\big) &= \int_{-\infty}^{z} \frac{1}{\pi(1+x^2)}\,dx \\ &= \frac{1}{\pi}\int_{-\infty}^{z} \frac{1}{1+x^2}\,dx \\ &= \frac{1}{\pi}\big(\arctan(z) + \frac{\pi}{2}\big) \\ &= \frac{1}{\pi}\arctan(z) + \frac{1}{2}. \end{aligned}$$

Now let $X_1 = f(X_0)$ where $f(x)$ is the Newton map.

$$\begin{aligned} P\big(X_1 < z\big) &= P\big(\frac{1}{2}(X_0 - \frac{1}{X_0}) < z\big) \\ &= P\big(X_0 - \frac{1}{X_0} < 2z\big) \end{aligned}$$

In the case that $X_0 < 0$ This means:

$$\begin{aligned} P\big(X_0 - \frac{1}{X_0} < 2z\big) &= P\big({X_0}^2 - 1 > 2zX_0\big) \\ &= P\big({X_0}^2 - 2zX_0 - 1 > 0\big) \end{aligned}$$

Using the quadratic formula and our assumption that $X_0 < 0$:

$$P\big({X_0}^2 - 2zX_0 - 1 > 0\big) = P\big(X_0 < z - \sqrt{z^2+1}\big)$$

In the case that $X_0 > 0$:

$$\begin{aligned} P\big(X_0 - \frac{1}{X_0} < 2z\big) &= P\big({X_0}^2 - 1 < 2zX_0\big) \\ &= P\big({X_0}^2 - 2zX_0 - 1 < 0\big) \end{aligned}$$

Using the quadratic formula and our assumption that $X_0 > 0$:

$$P\big({X_0}^2 - 2zX_0 - 1 < 0\big) = P\big(0 < X_0 < z + \sqrt{z^2+1}\big)$$

combining these two distinct cases:

$$\begin{aligned} P\big(X_0 - \frac{1}{X_0} < 2z\big) &= P\big(X_0 < z - \sqrt{z^2+1}\big) + P\big(0 < X_0 < z + \sqrt{z^2+1}\big) \\ &= \frac{1}{\pi}\arctan\left(z - \sqrt{z^2+1}\right) + \frac{1}{2} + \frac{1}{\pi}\arctan\left(z + \sqrt{z^2+1}\right) + \frac{1}{2} - \frac{1}{\pi}\arctan(0) - \frac{1}{2} \\ &= \frac{1}{\pi}\Big[\arctan\left(z - \sqrt{z^2+1}\right) + \arctan\left(z + \sqrt{z^2+1}\right)\Big] + \frac{1}{2}. \end{aligned}$$

We can make use of the trigonometric identity
$\arctan(x) + \arctan(y) = \arctan\left(\frac{x+y}{1-xy}\right)$ with $x = z - \sqrt{z^2+1}$ and $y = z + \sqrt{z^2+1}$.
This leads to:

$$\begin{aligned} P\big(X_0 - \frac{1}{X_0} < 2z\big) &= \frac{1}{\pi}\arctan(z) + \frac{1}{2} \\ &= P(X_0 < z) \end{aligned}$$

This shows that X_1 is also distributed by $\rho(x) = \frac{1}{\pi(1+x^2)}$ which proves that this distribution is the invariant distribution of the Newton map. □

To generate random variables which are distributed according to the invariant distribution of the Newton map one again has to make use of generalised inverses. The generalised inverse of a cumulative density distribution is simply its inverse. Let X be distributed according to the invariant distribution of the Newton map and let $u = F(x)$. Then u is a uniformly distributed random variable on the interval $(0, 1)$.

$$\begin{aligned}
&F(x) = u,\\
&\frac{1}{\pi}\arctan(x) + \frac{1}{2} = u,\\
&\frac{1}{\pi}\arctan(x) = u - \frac{1}{2},\\
&\arctan(x) = \pi\big(u - \frac{1}{2}\big),\\
&x = \tan\big(\pi\big[u - \frac{1}{2}\big]\big) = F^{-}(u).
\end{aligned}$$

With this inverse function we can generate sequences of dependent random variables using the Newton map and sequences of independent random variables following the same distribution. We can then compare the estimated value of the tail parameter for the two different ways of generating sequences. Figure 19 shows the estimated value of the tail parameter of sequences generated by the Newton map for different blocklengths and figure 20 shows the estimated value of the tail parameter for independent sequences of random variables distributed according to the invariant distribution of the Newton map. The two figures look very similar and in both cases the estimated value of the tail parameter seems to be close to 1. This suggests that sequences generated by the Newton map satisfy the conditions for the generalized extreme value theorem to be applicable to dependent sequences.

It can be seen that in both figures 19 and 20 the block maximum method sometimes estimates a very different value for the tail parameter than expected. This is because there is a small probability that random variables following the invariant distribution of the Newton map will attain very big values in their realisations. If by chance this happens often inside a block or it does not happen at all then this can affect the outcome of the maximum likelihood estimators.

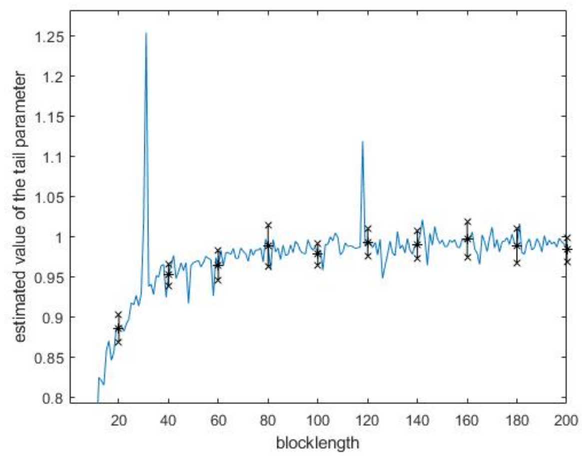

Figure 19: Estimated value of the tail parameter using sequences generated by the Newton map for different blocklengths.

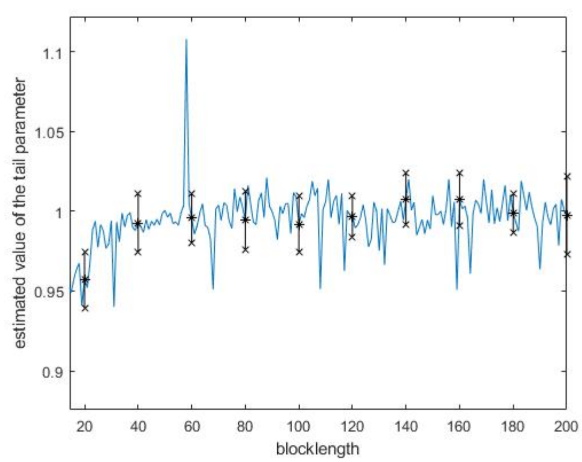

Figure 20: Estimated value of the tail parameter using independent random variables following the invariant distribution of the Newton map for different blocklengths.

4 Conclusion

The tail parameter of the generalized extreme value distribution of the limit of the rescaled partial maximum of a sequence of random variables gives a lot of insight into the statistical behaviour of the extreme values of the random variables. This is what extreme value statistics is about. The problem is that the limit of the rescaled partial maximum only follows a generalized extreme value distribution when the random variables are independent or when additional conditions are satisfied. It is useful to estimate the tail parameter using dependent sequences of random variables which satisfy the conditions of the generalized extreme value theorem because dependent sequences are easier to generate than independent sequences of random variables. This is because it is easier to compute the output of a simple function of a given realisation than it is to generate a random variable. Therefore it is useful to know which dependent sequences satisfy the conditions of the generalized extreme value theorem and which do not.

Multiple sequences of random variables have been analysed by comparing the dependent sequences with independent sequences following the same distribution. The dependent sequences were generated using iterative maps and the independent sequences followed the invariant distributions of these iterative maps. The generalized extreme value theorem was applicable to the independent random variables so comparing their results of the block maximum method with those of the dependent random variables following the same distribution can give a good indication whether the dependent random variables satisfy the conditions for the generalized extreme value theorem to be applicable to dependent random variables.

Dependent sequences generated by the logistic map, Newton map and beta map when β is odd do seem to satisfy the necessary conditions of the generalized extreme value theorem because the results of the block maximum method for sequences generated by those maps seem very similar to the results of independent sequences following the same distribution. This is not the case for dependent sequences generated by the rotation map or the cusp map. In the case of the cusp map the block maximum method estimates completely different values of the tail parameter when the sequence is dependent compared to when the sequence is independent. In the case of the rotation map the estimated values of the tail parameter do not even converge to any value for increasing blocklengths.

These results are not unexpected given the way in which dependent sequences generated by these iterative maps behave. Sequences generated by the cusp map have consecutive random variables attain only small differences in realisations and sequences generated by the rotation map seem very periodic. sequences generated by the logistic map, Newton map or beta map when β is odd seem independent even though each realisation is predetermined by the realisation of the first random variable. This suggests that sequences generated by the logistic map, Newton map or the beta map when β is odd do satisfy the conditions for the generalized extreme value theorem to be applicable to dependent sequences while sequences generated by the rotation map or cusp map do not.

References

[1] H.W.Broer A.E.Sterk, R.Vitolo. Het onvoorspelbare venijn van de staart. *Nieuw Archief Voor Wiskunde*, 3:164–167, 2013.

[2] L.de Haan A.Ferreira. On the block maxima method in extreme value theory: Pwm estimators. *The Annals of statistics*, 43:276–298, 2015.

[3] Bawa J. Trenner L. Dorazio P. Coles, S. *An Introduction to Statistical Modeling of Extreme Values.* London: Springer, 2001.

[4] J.R.Stedinger E.S.Martins. Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Research*, 36:737–744, 3 2000.

[5] F.Takens H.W.Broer. *Dynamical Systems and Chaos.* Springer, 2011.

[6] H. Rootzen M.R. Leadbetter, G. Lindgren. *Extremes and Related Properties of Random Sequences and Processes.* Springer-Verlag, 1980.

[7] M.Hofert P.Embrechts. A note on generalized inverses. *Mathematical Methods of Operations Research*, 77:423–432, 2013.

[8] T.Mikosch P.Embrechts, C.Klüppelberg. *Modelling Extremal Events for Insurance and Finance.* Springer, 1997.

[9] Çiğdem Özarı, Özge Eren, and Hasan Saygin. A new methodology for the block maxima approach in selecting the optimal block size. *Tehnicki Vjesnik*, 26:1292–1296, 2019.