ERASMUS UNIVERSITY ROTTERDAM

Quantifying Climate Change Impact on Extreme Weather Events, a Space-Time Extreme Value Analysis

Master Thesis Quantitative Finance

December 23, 2020

Abstract

The specific local impact of global warming on extreme weather events remain uncertain. This thesis applies recently developed extreme value techniques for heteroscedastic extremes (Einmahl et al., 2016, 2020), to investigate how climate change impacts future extreme weather probabilities. The choice of weather variables is motivated by an insurance industry perspective. We use historical daily data for precipitation and windstorm, starting from January 1971 to June 2020, from 13 automated weather stations distributed across the Netherlands. For extreme precipitation in the summer season and for extreme wind speed in the winter season we reject constant frequency of extremes over time. Finally, we forecast extreme weather probabilities until 2040. We conclude that the probability of precipitation extremes in the summer season has increased over the considered time period and most likely continues to increase in the years to come, whereas the probability of extreme windstorms in the winter season has decreased after a peak around 1990 and potentially continues to decrease in the years to come.

Keywords— Climate Change Risk, Extreme Weather Events, Non-Stationary Extreme Value Theory, Heteroscedastic Extremes

Author P.S.A. de Graaf

Supervisor Prof. dr. C. Zhou
Second Assessor dr. H.J.W.G. Kole

Contents

1 Introduction **3**

2 Literature Review **4**

2.1 Trends in extreme weather events 5

2.2 Extreme Value Theory 6

3 Data **8**

3.1 Seasonality 10

3.2 Data quality control 11

3.3 Descriptive statistics 11

3.4 Declustering of extreme events 13

4 Methodology **13**

4.1 Peak Over Threshold 14

4.1.1 Exceedance probabilities 16

4.2 Estimating temporal trends in extremes 16

4.3 Testing for temporal trends in extremes at each station 18

4.4 Estimating space-time trends in extremes 19

4.5 Testing for space-time trends in extremes 21

4.6 Pseudo Maximum Likelihood estimation 22

4.7 Heteroscedastic exceedance probabilities 23

5 Results **26**

5.1 24 hour precipitation accumulation 26

5.1.1 Univariate analysis per station 26

5.1.2 Multivariate analysis 29

5.2 Windstorm peak wind speed 34

5.2.1 Univariate analysis per station 34

5.2.2 Multivariate analysis 38

6 Forward looking scedasis **43**

6.1 24 hour precipitation accumulation 44

6.2 Windstorm peak wind speed 46

7 Conclusion **47**

8 Discussion **48**

References **51**

Appendices **54**

1 Introduction

Biological and chemical studies report that extreme weather events *will* become more frequent due to global warming (Easterling et al., 2000; Coumou & Rahmstorf, 2012). An increasing frequency of extreme weather events would pose a risk to companies that provide insurance in a high environmental risk market (Botzen & Van Den Bergh, 2008; Mills, 2009). Nevertheless, the specific local impact (if any) of current global warming on extreme weather events remain uncertain.

Detecting and quantifying temporal changes in the likelihood of extreme events is not straightforward, since by definition extreme observations are scarce. This also holds for meteorological variables, which in addition exhibit large natural variability in their extremes. Fortunately, the area of Extreme Value Theory is developed to accurately model extremes for which classical statistical inference is unreliable. Recently Einmahl et al. (2016) introduced a theoretical framework for independent but non-identically distributed extremes over time, referred to as *heteroscedastic extremes*. Their methodology is developed to model the variation in tail distribution continuously over time via a so-called scedasis function. Ferreira et al. (2017) and Einmahl et al. (2020) extend the framework to a space-time dimension and propose a test statistic to formally test for constant frequency of extremes in time and space.

This thesis applies these recently developed extreme value techniques for heteroscedastic extremes (Einmahl et al., 2016; Ferreira et al., 2017; Einmahl et al., 2020), to investigate the impacts of global warming on the frequency of local weather extremes in the Netherlands. Our choice of weather variables is motivated by an insurance industry perspective. The Dutch *Centrum voor Verzekeringsstatistiek* recently concluded that extreme precipitation and windstorm are among the most important loss generating events. Therefore we investigate the existence and magnitude of temporal trends in the frequency of extreme events for precipitation and windstorm in a space-time framework.

We use historical daily data, starting from January 1971 to June 2020, from 13 automated weather stations distributed across the Netherlands. We assume a constant extreme value index (i.e., the shape parameter governing heaviness of the tail distribution function) throughout time. This allows us to model a non-parametric trend in the scale parameter of the Generalized Pareto distribution using a scedasis function (Einmahl et al., 2016). Thereafter we test for constant frequency of weather extremes in both space and time (Einmahl et al., 2020). Then we incorporate

the (potential) time varying behavior into (joint) exceedance probabilities (Ferreira et al., 2017). Finally, we parameterize the non-parametric scedasis function using a (log) linear trend (Mefleh et al., 2020) to extrapolate (potential) time varying exceedance probabilities into the future.

We contribute to the existing literature as a first empirical application of the recently published results for non-identically distributed extremes (Einmahl et al., 2016, 2020). The findings of this research provide insights in the local climate change impact faced by society in general. Furthermore, this research provides a framework that allows insurance companies to quantify climate change impact on the hazard probability of their underwriting portfolio. Finally, a better understanding of the existence and magnitude of temporal trends in weather extremes allow for adaption, planning and mitigation efforts by governments and regulators. However, as this research will focus solely on the Netherlands, and climate impacts depends heavily on geographic region (Bouwer, 2013), the results of this research cannot be directly incorporated outside the Netherlands. This methodology, however, could be applied to many other (weather) variables for which trends in extremes are of interest.

We conclude that the probability of precipitation extremes in the summer season has increased over the considered time period and most likely continues to increase in the years to come. On the contrary, we conclude that the probability of extreme windstorms in the winter season has decreased after a peak around 1990 and potentially continues to decrease in the years to come. over the considered time period and potentially continues to decrease in the years to come. In addition, we reject the constant frequency of precipitation and windstorm extremes over space for the summer and winter season.

The remainder of this paper is structured as follows, Section 2 reviews the existing literature, Section 3 describes the data, Section 4 discusses the applied methodology, Section 5 presents the results, Section 6 presents the forward looking analysis and results, Section 7 provides a conclusion and finally Section 8 provides a discussion.

2 Literature Review

Within the framework of climate change, it is an academically well established fact that greenhouse gas emissions cause the mean temperature on earth to rise, see for example Stocker et al. (2013). It remains, however, uncertain how and if local extreme weather patterns will change as the mean

temperature rises. Several qualitative studies e.g. Botzen & Van Den Bergh (2008) and Mills (2009), suggest that climate change could have profound implications for property and casualty insurance companies. Despite the potential risk, currently the majority of the insurance companies do not integrate climate risk in their risk assessments (Thistlethwaite & Wood, 2018).

However, recently European regulators started the discussion of climate risk related capital requirements for insurance companies. Therefore our choice of weather variables is motivated by an insurance industry perspective. Based on Dutch non-life insurance claim data from 2000 to 2013 the Dutch *Centrum voor Verzekeringsstatistiek* concluded that extreme precipitation and windstorm are among the most important loss generating events. Hence, we review existing literature with regards to trends in extreme precipitation and extreme windstorms.

2.1 Trends in extreme weather events

Many researchers attempted to investigate the existence and magnitude of temporal trends in extreme precipitation. For instance, Klein Tank & Können (2003) estimates trends in indices of climate extremes based on daily series of temperature and precipitation observations from weather station distributed across Europe. They use a simple least squares regression where the regression coefficient estimates the change per decade. For precipitation, all Europe-average indices of extremes increase between 1946 and 1999, although the spatial coherence of the trends is low.

Groisman et al. (2005) analyzes changes in precipitation extremes, selected by event frequency thresholds. In contrast to Klein Tank & Können (2003), Groisman et al. (2005) did not focus solely on linear changes and used an additional a non-parametric test to check for a monotonic change of the time series. It was found that both the empirical evidence from the period of instrumental observations and model projections of a greenhouse-enriched atmosphere indicate an increasing probability of intense precipitation events for many extra-tropical regions including the United States.

Similar to Klein Tank & Können (2003), Alexander et al. (2006) analyzes trends in indices of climate extremes. They conduct two different analyses of temporal changes in the indices. The first examines trends in station and grid point data using a non-parametric estimator for the slope based on Kendall's tau, and the second compares empirical probability distribution functions for various time intervals. Precipitation changes showed a widespread and significant increase, but again the

spatial coherence is low, compared with temperature change.

Zolina et al. (2009) makes the assumption that the probability density function of daily precipitation follows a gamma distribution which is used to define extreme precipitation indices. Their results are based on long-term daily precipitation rain gauge observations over Europe. Analysis of these indices show an increase in extreme precipitation of up to 3% per decade in central western Europe.

Despite its practical relevance, existing literature that investigates the existence and magnitude of temporal trends in extreme windstorms are scarce. Easterling et al. (2000) raises concerns for increasing extreme weather events in their review of a number of observational studies. They conclude that limited research has been published on trends in extreme windstorms, and existing results are inconclusive. In addition,Meehl et al. (2000) indicate that there have been few observational studies of local changes in the frequency and intensity of windstorms. However, they notice a growing number of climate projection models that aim to address possible changes in windstorm activity using a scenario analysis. For instance, an analysis of four climate change experiments with hypothetical increased levels of CO_2 and sulfate aerosols on the northern hemisphere showed an increase in the number of extreme windstorms. In general, however, there is little agreement on the behaviour of storms in a warming climate.

2.2 Extreme Value Theory

Since we aim at analyzing the frequency of extreme events, we take an Extreme Value Theory perspective. The area of Extreme Value Theory is developed to accurately model the probabilities of highly extreme outcomes. It is the study of the tail of a probability distribution near the endpoint, where inference is challenging due to the scarcity of data. It is introduced by Fisher & Tippett (1928) and Gnedenko (1943) and extensively applied to model financial, hydrological and environmental variables, of which the risk of extreme outcomes is of interest. For instance, Rootzén & Tajvidi (1997) applies extreme value statistics to model extremes in insured wind storm losses.

We refer to Coles et al. (2001) for an extensive overview of the classical methods in the field of Extreme Value Theory and solely introduce the underlying theory to connect to the existing literature. Consider an independent and identically distributed sample $X_1, X_2, \cdots, X_n$ with distribution function $F(x) = P(X \leq x)$. Define the sample maxima as $M_n = \max(X_1, X_2, \ldots, X_n)$. Suppose

that there are proper normalizing series $b_n \in \mathbb{R}$ and $a_n > 0$, such that,

$$\lim_{n\to\infty} P\left(\frac{M_n - b_n}{a_n} \leq x\right) = G(x), \tag{1}$$

For such distribution functions F and G, we write $F \in MDA(G)$, that is F is in the maximum domain of attraction of G. Gnedenko (1943) proves that, if $F \in MDA(G)$ then G is a Generalized Extreme Value distribution. The Generalized Extreme Value distribution is a parameterization of three distributions into one defined as,

$$G_\gamma(x) = \exp\left(-(1+\gamma x)^{-1/\gamma}\right), \tag{2}$$

where $-\infty < \gamma < \infty$ is known as the shape parameter or the extreme value index. The extreme value index distinguishes across the three types of distributions: the Fréchet distribution if $\gamma > 0$, the Gumbel distribution if $\gamma = 0$ and the Weibull distribution if $\gamma < 0$.

Since classical Extreme Value Theory is based on independent and identically distributed random variables, it can not be readily applied to detect trends in extremes. We can distinguish two deviations, dependence and non-identical distributions. To detect trends in extremes we need to consider the latter deviation and discuss existing literature that copes with extremes that are drawn from gradually shifting tail distributions.

A number of studies have explored a trend in the parameters of a limit distributions in extreme value theory. Smith (1989) proposed a time dependent generalization of the location parameter b_n from Equation (1) defined as $b_n(t)$. They fitted a linear trend $b_n(t) = b_n(0) + tc$ with an analysis based on the point process strategy. However, no asymptotic properties are derived and the trend for $b_n(t)$ could get out of range if $\gamma < 0$. Differently, Davison & Smith (1990) proposes a linear trend on the shape and scale parameters of the Generalized Pareto (GP) distribution and Coles et al. (2001) considered a linear trend in the log-scale parameter of the GP distribution. However, no asymptotic properties of the estimators are derived and a trend in the (log) scale parameter of the GP distribution is difficult to interpret. The methods proposed were parametric. Hall & Tajvidi (2000) suggest an adaptive, non-parametric approach. They argue that in most cases the trend is more complex than a simple parametric analysis would allow. They propose non-parametric methods to fit time-varying parameters. However, again, no asymptotic properties are derived. Hence these existing methods are not completely satisfactory, especially not for testing the existence of trends in the frequency of extremes.

Differently, De Haan et al. (2015) proposes a method that models the exceedance probability of a pre-defined high threshold over time. They introduce a simple model for relative risk, using (possibly) non-identical distribution functions of the random variables defined for each point in time. A quotient of probabilities is modelled such that the exceedance probability at a given time point, is a factor times the exceedance probability of a baseline distribution. In contrast to the work discussed earlier, this methodology has the desirable property that the results are easy to interpret. However the analysis requires multiple observations at each time point and focuses solely on one specific parametric trend.

More general, Einmahl et al. (2016) introduces a flexible theoretical framework referred to as *heteroscedastic extremes* i.e. time varying frequency of extremes. They propose to model (potential) changes in the tail of a distribution function continuously over time via the positive *scedasis* function. In addition, they propose a test statistic to formally test for constant frequency of extremes in time. This approach applies an extreme value analysis method based on the domain of attraction rather than the limit situation. It does not impose any (local) parametric model on the scedasis function and the analysis requires only one single observation at each time point. Finally, the scedasis function can be estimated non-parametrically, with asymptotic theories.

Most recently, the heteroscedastic extremes framework of Einmahl et al. (2016) is extended in several research fields. For example Ferreira et al. (2017) extends the heteroscedastic extremes framework to space-time dimension via a space-time scedasis function that allows for trends both in time and space. In addition, Ferreira et al. (2017) proposes a homogenization procedure to obtain stationary pseudo observations that can be used to incorporate the time-varying behavior of extremes in the estimation of (joint) exceedance probabilities using the non-parametric scedasis function. Furthermore, Einmahl et al. (2020) extends Ferreira et al. (2017) and propose a test statistic to test for constant frequency of extremes in time and space, which they apply to estimate space–time trends of precipitation extremes in North-Western Germany.

3 Data

The Royal Netherlands Meteorological Institute (KNMI) manages several automated weather stations distributed across the Netherlands. The weather stations continuously measure the wind direction and speed, temperature, humidity, precipitation intensity, radiation from the sun, visibility and air pressure, as well as type of precipitation. Following the Dutch CVS we collect

historical daily weather data regarding precipitation and windstorm from multiple weather stations distributed across the Netherlands. Selecting the specific variables of interest regarding windstorm and precipitation is crucial for the relevance of this research. From the insurance perspective we aim to select variables that are able to capture losses of extreme windstorms and excessive precipitation.

Several studies conclude that the intensity of the precipitation significantly explains the claim sizes. Furthermore it can be concluded that additional or persistent precipitation does not increase damage on existing claims but it does increase the total number of claims. More specifically the Dutch *Verbond Van Verzekeraars* reports a threshold of 38 mm per 24 hour beyond which the number of claims increases rapidly. To conclude, both the intensity as well as the daily sum are of interest. Hence, to analyse precipitation extremes over space and time we collect daily data for the 24 hour precipitation accumulation in mm.

Claims generated by windstorms are mainly caused by extreme wind speeds rather than an extreme duration of the storm (Dorland et al., 1999). On average windstorms generate in the Netherlands approximately 50 million euros in claims per year. To analyse windstorm extremes over space and time we collect daily data for the peak wind speed measured in m/s.

Figure 1: Weather station locations geographically plotted across the Netherlands

From the set of available KNMI weather stations in the Netherlands, we select 13 stations based on their data availability. For the peak wind speed we collect data from January 1971 until June 2020, which yields approximately 18000 observations per station. For 24 hour precipitation accumulation we collect data from June 1974 until June 2020, which yields approximately 16800 observations per station. the weather stations *Deelen*, *Gilze-Rijen* and *Eindhoven* are not considered due to limited data availability. Figure 1 plots the locations of the weather stations over space. It can be observed that the stations spatially cover all the regions of interest in the Netherlands. That is, we consider both coastal and inland weather stations as well as stations in all of the four cardinal directions.

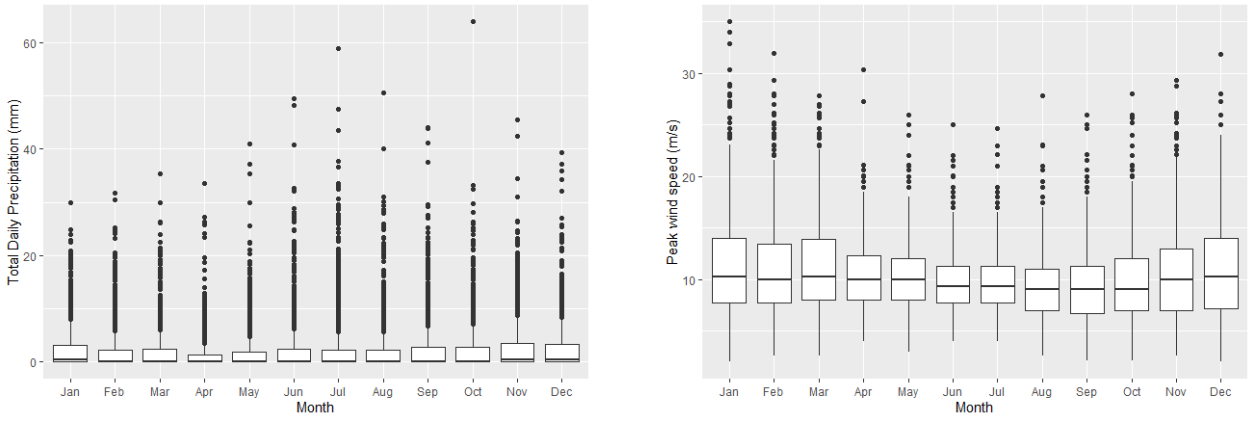

Figure 2: Seasonality box-plot for 24 hour precipitation accumulation and peak wind speed in De Bilt, the black dots represent the extremes.

3.1 Seasonality

Hydrological and environmental variables tend to exhibit seasonality. Figure 2 provide seasonal box-plots of the daily data from the oldest weather station De Bilt for peak wind speed and 24 hour precipitation accumulation. We can clearly identify a seasonal pattern for the peak wind speed. Peak wind speed in the winter months is on average higher than in summer months. However, 24 hour precipitation accumulation tend to be higher in summer months compared to winter months.

Since our mean data exhibits seasonal behavior we assume their extremes to do so as well. Inspecting the extremes (black dots) in the box-plots in Figure 2 confirms this. We separate the analyses for summer and winter periods to overcome the problem of seasonality. To control for the transition periods we define May up to and including September as the summer season and November up to and including March as the winter season, following Einmahl et al. (2020). Hence we allow for seasonal differences in the impact of climate change on extreme weather patterns.

3.2 Data quality control

Klein Tank et al. (2009) provide an extensive discussion on the importance of data quality in climate change related research, in which they refer to the *World Meteorological Organisation* (WMO). The WMO provides guidance in developing long-term, high-quality and reliable time series records to asses climate change impact. The basic quality control proposed by WMO consists of multiple steps. Daily meteorological time series data tend to have missing values. Note that the presence of a highly extreme event might be the cause of the weather station to fail, leading to potentially missing extremes. Furthermore the WMO raises concerns for obvious wrong values, for example nonexistent dates and negative precipitation. Finally, it is worth noting that in general long-term climatic time series are influenced by inhomogeneities, i.e., non-climatic factors, like changes in observational instruments and practices, station reallocation's and changes in station environment.

Following Klein Tank et al. (2009), we allow for at most four missing observations per station per year, which we define as the completeness criteria. Missing observations are then compared with neighbouring weather stations to ensure that there were no extremes present locally. According to the KNMI, precipitation observations $< 0.1mm$ are not accurately recorded therefore we manually set these to zero for Total Daily Precipitation. The KNMI provides homogenized data that is manually corrected for non-climatic factors, and does not consist of obvious wrong values.

3.3 Descriptive statistics

Table 1 provide the descriptive statistics for 24 hour precipitation accumulation for the 10 weather stations in the summer and winter season, respectively. The descriptive statistics for 24 hour precipitation accumulation are reasonably spatially coherent, for both summer and winter. Comparing the descriptive statistics for summer and winter season reveals again the presence of seasonality. We observe that summer season 24 hour precipitation accumulation is higher when compared to winter season 24 hour precipitation accumulation, in both their standard deviation, 95% empirical quantile and maximum.

Table 2 provide the descriptive statistics for daily peak wind speed observations for the 13 weather stations in the summer and winter season, respectively. We recognize the variability across space for both summer and winter. The stations differ both in their mean, standard deviation, 95% empirical quantile and their maximum. When we compare *De Kooy* and *Vlissingen* with *Twenthe* it follows that this variability across space is mainly driven by the difference between coastal and

Table 1: Descriptive statistics for daily observed 24 hour precipitation accumulation for 10 weather stations distributed across the Netherlands from June 1974 until June 2020.

Descriptive Statistics	De Kooy	Schiphol	De Bilt	Leeuwarden	Eelde	Twenthe	Vlissingen	Rotterdam	Volkel	Maastricht
Summer Season										
Mean	2.08	2.32	2.31	2.32	2.26	2.17	2.08	2.38	2.07	2.19
Std. Deviation	4.85	5.15	4.94	5.03	4.63	4.85	4.77	5.39	4.71	5.03
Quantile 95%	11.00	12.41	12.30	12.00	11.31	11.31	11.20	12.51	11.20	11.60
Maximum	66.30	67.20	58.90	69.70	51.30	106.40	80.90	101.40	68.00	68.60
Observations	7039	7039	7039	7039	7039	7039	7039	7039	7039	7039
Winter Season										
Mean	2.14	2.23	2.33	2.17	2.21	2.14	2.07	2.36	2.04	2.09
Std. Deviation	3.73	3.99	4.22	3.76	3.85	3.91	3.68	4.21	3.86	3.83
Quantile 95%	9.91	10.10	11.00	10.10	10.30	10.40	9.50	11.21	9.80	9.81
Maximum	37.30	44.50	45.50	34.30	36.70	45.00	39.20	38.50	52.70	69.60
Observations	6958	6958	6958	6958	6958	6958	6958	6958	6958	6958

Note: The table consists of descriptive statistics for the 24 hour precipitation accumulation in mm, separated in northern hemispheric summer and winter periods. The descriptives are based on daily data collected from June 1974 until June 2020. Observations $< 0.1mm$ *are not accurately recorded and manually set to zero. Since 24 hour precipitation accumulation has a natural lower bound, the minimum is not reported.*

Table 2: Descriptive statistics for daily observed peak wind speed for 13 weather stations distributed across the Netherlands from January 1971 until June 2020.

Descriptive Statistics	De Kooy	Schiphol	De Bilt	Leeuwarden	Deelen	Eelde	Twenthe	Vlissingen	Rotterdam	Gilze-Rijen	Eindhoven	Volkel	Maastricht
Summer Season													
Mean	11.71	11.20	9.62	10.88	10.52	10.52	9.62	12.02	10.72	10.18	10.23	9.66	9.93
Std. Deviation	3.75	3.41	3.02	3.34	3.14	3.22	3.04	4.10	3.45	3.02	3.21	3.18	3.41
Quantile 95%	18.50	17.50	15.00	17.00	16.00	16.00	15.00	19.50	17.00	15.40	16.00	15.00	16.00
Maximum	30.00	36.00	27.80	28.80	30.00	31.40	26.80	36.00	35.00	30.40	32.00	34.00	32.00
Observations	7498	7528	7498	7465	7351	7495	7482	7514	7506	7458	7488	7456	7498
Winter Season													
Mean	14.02	12.95	10.91	12.67	11.34	11.93	10.64	14.11	12.55	11.08	10.99	10.72	11.31
Std. Deviation	5.42	5.32	4.58	5.05	4.63	4.96	4.42	5.74	5.11	4.63	4.61	4.59	4.82
Quantile 95%	23.70	23.00	19.00	22.00	20.00	21.00	19.00	25.00	22.00	19.50	19.50	19.00	20.10
Maximum	40.60	41.20	35.00	39.10	34.00	38.60	34.50	41.20	41.70	36.00	34.50	38.60	34.00
Observations	7456	7500	7456	7436	7386	7456	7421	7381	7489	7397	7436	7384	7456

Note: The table consists of descriptive statistics for peak wind speed in m/s*, separated in northern hemispheric summer and winter periods. The descriptives are based on daily data collected from January 1971 until June 2020. Since peak wind speed has a natural lower bound, the minimum is not reported.*

inland stations. Furthermore, comparing the descriptive statistics for summer and winter season reveals again the presence of seasonality. We observe that winter season peak wind speed is higher when compared to summer season peak wind speed, in both their mean, 95% empirical quantile and maximum.

3.4 Declustering of extreme events

Hydrological variables tend to exhibit substantial serial dependence. Furthermore there exist evidence that exceedances over high thresholds occur in clusters as well. We therefore employ a declustering procedure to remove potential temporal dependence following Ferreira et al. (2017); Einmahl et al. (2020). In general, the idea is to exclude some dates in the data set in order to create gaps between consecutive observations. After declustering, extremes in our data can be regarded as serially independent.

For both variables of interest we perform a separate declustering procedure described as follows. We calculate for each point in time $i = 1, \dots n$ the station wise maxima $M_1, \dots, M_n$ across the 13 (10) weather stations. The number of time points n differs per variable and per season. That is, for 24 hour precipitation accumulation $n = 7039$ in the summer season and $n = 6958$ in the winter season, whereas for peak wind speed $n = 7529$ in the summer season and $n = 7502$ in the winter season. We order the station wise maxima $M_1, \dots, M_n$ from high to low. Select the pair of dates with the largest and second largest station-wise maxima. If the date of the second station-wise maxima is within two days around the first station-wise maxima, then this corresponding date is excluded from our original data set. Hence all 13 observations are removed. Otherwise both dates are retained.

We roll this approach forward over all station-wise maxima from high to low. Eventually we remove dates if their station-wise maxima is recorded within two consecutive days of any of the previously retained dates. This procedure results in the serially independent data set used for modelling trends in extremes.

4 Methodology

We aim to investigate the impact of climate change on local weather extremes in the Netherlands. Following Klein Tank et al. (2009) we define climate change as any change occurring to the climate either permanently or lasting for long periods of time, regardless of cause. The applied methodology

is structured in a multiple step analysis which we perform for the daily observations of 24 hour precipitation accumulation and peak wind speed in both the summer and winter season.

First, we perform a Peak Over Threshold analysis with assumed i.i.d observations for each station using Maximum Likelihood estimation. We use the estimated extreme value parameters to estimate stationary exceedance probabilities of thresholds for which exceedances are rarely observed in the data. Second, we relax the i.i.d. assumption and test for constant frequency of extremes in time at each station with the heteroscedastic extremes framework (Einmahl et al., 2016). Third, we allow for spatial dependence to test for constant frequency of extremes in time and space (Ferreira et al., 2017; Einmahl et al., 2020). Fourth, we estimate the extreme value parameters for non-stationary, spatially dependent extremes with pseudo Maximum Likelihood estimation by pooling observations from all stations (Einmahl et al., 2020). Finally, depending on the presence of heteroscedastic extremes, we apply a homogenization procedure to obtain stationary pseudo observations that are used together with the non-parametric space-time scedasis function to estimate (joint) exceedance probabilities over time (Ferreira et al., 2017).

What follows is a detailed explanation of the methodology. The data referred to below, is the cleaned and serially independent data as discussed in Section 3.

4.1 Peak Over Threshold

Let $X_1, X_2, \ldots, X_n$ denote the set of (assumed) independent and identically distributed observations, with distribution function $F(x) = P\{X \leq x\}$. Since we are interested in modelling the probabilities of extreme events, we perform an Extreme Value Analysis using the Peak Over Threshold approach. In general, the idea is to consider solely *extreme* observations in the data. We let $X_{n:1} \leq X_{n:2} \leq \cdots \leq X_{n:n}$ be the order statistics of $X_1, X_2, \cdots, X_n$. Then for an intermediate sequence $k = k(n)$, i.e. a sequence satisfying

$$\lim_{n\to\infty} k = \infty \text{ and } \lim_{n\to\infty} \frac{k}{n} = 0, \tag{3}$$

we select the extreme value threshold as upper order statistic $X_{n:n-k}$. Correspondingly, k denotes the number of threshold exceedances in the sample. Then $\{Y_1, \ldots, Y_k\} = \{X_i - X_{n:n-k} | X_i \geq X_{n:n-k}\}$ denotes the set of independent excesses above the threshold $X_{n:n-k}$ with conditional excess distribution defined as $F_k(x) = P(X - X_{n:n-k} \leq x \mid X > X_{n:n-k})$. Pickands et al. (1975) proves that if $F \in MDA(G_\gamma)$ then the conditional excess distribution F_k is approximately a Generalized Pareto

(GP) distribution. For $\sigma_k > 0$, the GP distribution is defined as,

$$H_\gamma(x) = \begin{cases} 1 - \left(1 + \frac{\gamma x}{\sigma_k}\right)^{-1/\gamma} & \text{if } \gamma \neq 0 \\ 1 - \exp\left(-\frac{x}{\sigma_k}\right) & \text{if } \gamma = 0 \end{cases} \tag{4}$$

where $x \geq 0$, if $\gamma \geq 0$ and $0 \leq x \leq -\frac{\sigma_k}{\gamma}$ if $\gamma < 0$. The parameters of the GP distribution are uniquely determined by those of the corresponding Generalized Extreme Value distribution G_γ.

For further details we refer to Coles et al. (2001) and especially the references therein. We assume the required conditions are satisfied. Then we define the extreme value threshold $X_{n:n-k}$, number of exceedances k, independent excesses $Y_1, \ldots, Y_k$ and conditional excess distribution F_k as above. We model the independent excesses as a GP distribution, as given in Equation (4).

To estimate the γ and σ_k we use the Maximum Likelihood estimator. Let the log-likelihood function derived from Equation (4) be given by,

$$L\left(\gamma, \sigma_k \mid Y_1, \ldots, Y_k\right) = -k \log \sigma_k - \left(1 + \frac{1}{\gamma}\right) \sum_{j=1}^{k} \log\left(1 + \frac{\gamma}{\sigma_k} Y_j\right). \tag{5}$$

In general, $\hat{\sigma_k}$ and $\hat{\gamma}$ are Maximum Likelihood estimates if they maximize $L\left(\gamma, \sigma_k \mid Y_1, \ldots, Y_k\right)$, solving the score equations,

$$\begin{cases} \dfrac{\partial}{\partial \gamma} L\left(\gamma, \sigma_k, \mid Y_1, \ldots, Y_k\right) = 0 \\ \dfrac{\partial}{\partial \sigma} L\left(\gamma, \sigma_k \mid Y_1, \ldots, Y_k\right) \ = 0. \end{cases} \tag{6}$$

We solve the score equations for γ and σ numerically using the *Newton-Raphson* algorithm for $Y_1, \ldots, Y_k$. Note that this estimation procedure is valid for $\gamma \in \mathbb{R}$ as required for meteorological data.

It is evident that choosing a proper threshold $X_{n:n-k}$ is essential to this method, this comprises a trade-off between variance and bias. If the threshold is too low, estimates are biased as the estimator takes into account non-extreme values, whereas choosing a threshold too high results in parameter estimates that are uncertain. Several heuristics for choosing the optimal number of upper order statistics k exists. We estimate parameter stability plots as a function of k. Since the scale estimator is a function of k, we need to select one k for all stations to be able to compare the scale parameter across space.

4.1.1 Exceedance probabilities

The strength of extreme value theory is the accurate estimation of tail event probabilities, especially for events that are scarce or even never observed in the data. Using the estimated GP distribution parameters we can estimate exceedance probabilities of high level extreme thresholds. We have,

$$\Pr(X - X_{n:n-k} > x \mid X > X_{n:n-k}) = \frac{\Pr(X - X_{n:n-k} > x)}{\Pr(X > X_{n:n-k})}. \tag{7}$$

Let p_z for all $z > u$ denote $\Pr(X > z)$. We substitute the estimated conditional excess distribution F_k and let $\lambda_k = \Pr(X > X_{n:n-k})$ denote the exceedance rate, re-writing then gives, for all $x > X_{n:n-k}$

$$\hat{p}_z = \lambda_k \left(1 + \hat{\gamma}\left(\frac{z - X_{n:n-k}}{\hat{\sigma}}\right)\right)^{-1/\hat{\gamma}}. \tag{8}$$

Finally, we estimate λ_k as the empirical threshold exceedance rate, such that

$$\hat{p}_z = \frac{k}{n}\left(1 + \hat{\gamma}\left(\frac{x - X_{n:n-k}}{\hat{\sigma}}\right)\right)^{-1/\hat{\gamma}}. \tag{9}$$

We perform this Peak Over Threshold analysis for each station independently. Note that, the scale and shape parameters of the GP distribution are estimated as being constant over time. However we allow for varying scale and shape parameters across space.

4.2 Estimating temporal trends in extremes

In order to model trends in extremes we relax the assumption of i.i.d observations. Hence, observations $X_1, X_2, \ldots, X_n$ are now possibly non-identically but independent distributed observations, collected over time.

We employ a framework developed for modelling non-identically distributed observations over time introduced by Einmahl et al. (2016). We define (possibly) mildly variating continuous distribution functions $F_{n,i}(x) = P\{X_i \leq x\}$ for all $i = 1, \ldots, n$. Within the heteroscedastic extremes framework the distribution of extreme events evolves continuously over time via the positive scedasis function c defined on $[0, 1]$. Mathematically we have,

$$\lim_{x \to x*} \frac{1 - F_{n,i}(x)}{1 - F_0(x)} = c\left(\frac{i}{n}\right), \tag{10}$$

where x^* denotes a common right end point and F_0 some baseline distribution. Here we impose a condition that

$$\int_0^1 c(s)ds = 1. \tag{11}$$

Within this framework, c can be interpreted as the relative frequency of extreme events over the considered time period. The case $c(s) \equiv 1$ is called homoscedastic extremes i.e. constant frequency of extremes.

Similar to classical Extreme Value Theory, we assume that the baseline distribution F_0 belongs to the domain of attraction of a Generalized Extreme Value distribution with extreme value index γ. That is, there exists a $\gamma \in \mathbb{R}$ and a positive scale function a such that, for all $x > 0$ we have,

$$\lim_{t\to\infty} \frac{U(tx) - U(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}, \tag{12}$$

where

$$U := \left(\frac{1}{1-F}\right)^{\leftarrow} \tag{13}$$

and $\leftarrow$ denotes the left continuous inverse function.

Einmahl et al. (2016) consider only the case where $\gamma > 0$. In general, meteorological data such as precipitation accumulation and peak wind speed have small or even slightly negative extreme value index γ. However Einmahl et al. (2020) generalize this methodology to the case where $\gamma \leq 0$, as desired when modelling meteorological data.

Combining the proportional tail condition in Equation (10) with the domain of attraction condition in Equation (12) yields a flexible non-parametric model. This model implies that all mildly varying distribution functions $F_{n,i}$ belong to the domain of attraction of the same extreme value distribution, i.e. they have the same extreme value index γ. The trend in the frequency of extremes is modelled by the non-parametric scedasis function c which impacts the scale parameter of extremes.

We can estimate the scedasis function c directly. Recall that Equation (11) ensures that the scedasis function can be interpreted as the relative frequency of extreme events over the considered time period. If we interpret the time interval $[0, 1]$ as the domain of a random variable, then the scedasis function c can be interpreted as a probability density function. Hence, it is intuitive to estimate c with probability density estimation techniques.

Since we do not want to assume a monotonic or parametric scedasis we estimate c using a Kernel Density Estimation method. We define $K(x)$ as the continuous, symmetric biweight kernel function defined on $[-1, 1]$ as $K(x) = 15\left(1 - x^2\right)^2/16$. We have that $\int_{-1}^{1} K(x)dx = 1$ and we set $K(x) = 0$ for $|x| > 1$. Let $h := h_n > 0$ be a bandwidth such that $h \to 0$ and $kh \to \infty$, when $n \to \infty$. Now we define the non-parametric scedasis estimator as,

$$\hat{c}(s) = \frac{1}{kh} \sum_{i=1}^{n} \mathbb{I}_{\{X_i > X_{n:n-k}\}} K\left(\frac{s - i/n}{h}\right). \tag{14}$$

Note that for Kernel Density Estimation in general and for scedasis estimation in particular, we require a boundary correction for unbiased estimation within the bandwidth h of both boundaries 0 and 1. We follow Jones (1993) and we estimate the non-parametric scedasis for $s \in [0, 0+h)$ and $s \in (1-h, 1]$ as,

$$\hat{c}(s) = \frac{1}{kh} \sum_{i=1}^{n} \mathbb{I}_{\{X_i > X_{n:n-k}\}} K_b\left(\frac{s - i/n}{h}\right), \tag{15}$$

with boundary kernel K_b defined as,

$$K_b(x) = \frac{\int_0^1 u^2 K(u)\mathrm{d}u - x\int_0^1 uK(u)\mathrm{d}u}{\frac{1}{2}\int_0^1 u^2 K(u)\mathrm{d}u - \left\{\int_0^1 uK(u)\mathrm{d}u\right\}^2} K(x). \tag{16}$$

4.3 Testing for temporal trends in extremes at each station

The estimated scedasis function is easy to interpret and eminently suited for visual exploratory analysis. However, for formal testing we use the integrated scedasis function defined as,

$$C(s) = \int_0^s c(t)\mathrm{d}t, \text{ for } s \in [0,1]. \tag{17}$$

The function C should be proportional to the number of exceedances above a high threshold in the first $[ns]$ observations. Then for an intermediate sequence $k = k(n)$, i.e. a sequence satisfying

$$\lim_{n\to\infty} k = \infty \text{ and } \lim_{n\to\infty} \frac{k}{n} = 0, \tag{18}$$

we define the non-parametric integrated scedasis estimator as,

$$\hat{C}(s) = \frac{1}{k} \sum_{i=1}^{[ns]} \mathbb{I}_{[X_i > X_{n:n-k}]}. \tag{19}$$

Hence, integrated scedasis estimator $\hat{C}(s)$ can be estimated using solely the pseudo order statistics of $X_1, X_2, \ldots, X_n$. To ensure the asymptotic normality of these estimators we require second-order conditions on the rate of convergence for which we refer to Einmahl et al. (2016).

We test the null-hypothesis that $c(s) \equiv 1$, which corresponds to testing that the integrated scedasis $C(s)$ is the identity function on the considered time period. That is, for $0 \le s \le 1$ we have,

$$\begin{aligned} H_0 &: C(s) = s, \\ H_a &: C(s) \neq s. \end{aligned} \tag{20}$$

If we can reject this null hypothesis, we reject the presence of homoscedastic extremes, or more specifically, we reject the hypothesis that the frequency of extreme events is constant over the considered time period. Einmahl et al. (2016) propose a Kolmogorov-Smirnov-type test based on the process $\sqrt{k}\left(\hat{C}_j(s) - s\right)$ for $0 \leq s \leq 1$. Under the null hypothesis we have,

$$\left\{\sqrt{k}\left(\hat{C}_j(s) - s\right)\right\}_{s\in[0,1]} \xrightarrow{d} \{B(s)\}_{s\in[0,1]}. \tag{21}$$

We consider Kolmogorov-Smirnov-type test statistic T_1 defined as,

$$T_1 := \sup_{0\leqslant s\leqslant 1} \left|\hat{C}(s) - s\right|. \tag{22}$$

Einmahl et al. (2016) provide the asymptotic distribution of the test statistic under the null hypothesis. We assume the required conditions are satisfied, then as $T \to \infty$,

$$\sqrt{k}T_1 \xrightarrow{\mathrm{d}} \sup_{0\leqslant s\leqslant 1} |B(s)|, \tag{23}$$

where B denotes a standard Brownian bridge. We perform this analysis for each station independently.

4.4 Estimating space-time trends in extremes

We extend our framework to a situation with *spatially dependent* observations following Ferreira et al. (2017) and Einmahl et al. (2020), to model trends in space and to investigate spatial coherence for trends in time. Instead of considering the weather station separately, we now consider the weather stations in a multivariate setting. Let m denote the number of weather stations in the data, We assume that the distribution of the random vectors $(X_{i,1}, \ldots, X_{i,m})$ for $i = 1, \ldots, n$ is in the domain of attraction of a multivariate extreme value distribution.

Similar to the temporal dimension we model proportional tails over time. Let $F^j_{n,i}$ denote the marginal distribution function of $X_{i,j}$ for all $i = 1, \ldots, n$ and $j = 1, \cdots, m$. For some baseline continuous distribution function F_0 in the maximum domain of attraction of an extreme value distribution G_γ we have

$$\lim_{x\uparrow x^*} \frac{1 - F^j_{n,i}(x)}{1 - F_0(x)} = c\left(\frac{i}{n}, j\right) \in (0, \infty), \tag{24}$$

for all $i = 1, \ldots, n$ and $j = 1, \cdots, m$, where the space-time scedasis $c(\frac{i}{n}, j)$ is a positive continuous function for each station j. In addition, to ensure that the space-time scedasis function is uniquely defined we impose the condition,

$$\sum_{j=1}^{m} \frac{1}{m} \int_0^1 c(s, j)ds = 1. \tag{25}$$

The space-time scedasis $c(\frac{i}{n}, j)$ can be interpreted as the relative frequency of extremes at time i for station j. Since we assume that baseline distribution $F_0 \in D(G_\gamma)$ i.e. is in the maximum domain of attraction of G_γ, we have $F^j_{n,i} \in D(G_\gamma)$ for all time points $i = 1, \cdots, n$ and all stations $j = 1, \cdots, m$. Hence the extreme value index γ within this framework is assumed to be constant over time and space.

To model the spatial dependence we define F_i as the joint distribution function of the vector $(X_{i,1}, \ldots, X_{i,m})$. We assume that $F^j_{n,i}$ is continuous for all time points $i = 1, \cdots, n$ and all stations $j = 1, \cdots, m$, similar to Equation (13) we let $U^j_{n,i} := (\frac{1}{1-F^j_{n,i}})^{\leftarrow}$. We further assume that

$$\tilde{F}(x_1, \ldots, x_m) := F_i\left(U^1_{n,i}(x_1), \ldots, U^m_{n,i}(x_m)\right), \tag{26}$$

holds for all $i = 1, \ldots, n$ and is in the domain of attraction of a multivariate extreme value distribution. It follows that the multivariate spatial tail dependence structure is time invariant within this framework. Now we define R_{j_1,j_2} as the tail copula of the weather stations j_1 and j_2, that is,

$$R_{j_1,j_2}(x, y) = \lim_{s \downarrow 0} \frac{1}{s} P\left(1 - F^{j_1}_{n,i}(X_{i,j_1}) \leq sx, 1 - F^{j_2}_{n,i}(X_{i,j_2}) \leq sy\right). \tag{27}$$

with $j_1 \neq j_2$ and for $(x, y) \in [0, \infty]^2 \backslash \{(\infty, \infty)\}$.

As previously, we use the integrated scedasis function for formal testing. Let C_j denote the integrated scedasis function for station j defined as,

$$C_j(s) = \frac{1}{m} \int_0^s c(t, j) \text{ for } s \in [0, 1]. \tag{28}$$

Note that the condition in Equation (25) ensures that,

$$\sum_{j=1}^{m} C_j(1) = 1. \tag{29}$$

Similar to Equation (19), $C_j(s)$ could be estimated as the number of exceedances over a high quantile at station j. However, since we want to test for a trend in the frequency of extreme events across space we use the same threshold for all weather stations. Following Ferreira et al. (2017) and Einmahl et al. (2020), we define the common threshold as a high empirical quantile of all $N = n \cdot m$ observations, let $X_{N:N-k}$ denote the $(N-k)$-th order statistic of all observations i.e. the set $\{X_{i,j}\}$ for all time points $i = 1, \ldots, n$ and all stations $j = 1, \ldots, m$. Then for an intermediate sequence k, we define the non-parametric integrated scedasis estimator as,

$$\hat{C}_j(s) := \frac{1}{k} \sum_{i=1}^{[ns]} \mathbb{I}_{\{X_{i,j} > X_{N:N-k}\}}. \tag{30}$$

4.5 Testing for space-time trends in extremes

We test if there are significant differences in the cumulative frequency of extreme events across the Netherlands i.e. whether the frequency of extreme events is constant over space. More specifically, we test weather $C_j(1) = \frac{1}{m}$ for all $j = 1, \ldots, m$, that is,

$$\begin{aligned} H_0 &: C_j(1) = \tfrac{1}{m} \quad \text{for all } j = 1, \ldots, m \\ H_a &: C_j(1) \neq \tfrac{1}{m} \quad \text{for some } j = 1, \ldots, m. \end{aligned} \tag{31}$$

Let $M = I_m - \frac{1}{m}\mathbb{I}_m\mathbb{I}_m'$ where I_m denotes the m-dimensional identity matrix and $\mathbb{I}_m$ the m-dimensional unit vector. Einmahl et al. (2020) proves that under the null hypothesis $D = \sqrt{k}\left(\hat{C}_1(1) - \frac{1}{m}, \ldots, \hat{C}_m(1) - \frac{1}{m}\right)'$ is asymptotically an m-dimensional multivariate normally distributed with zero mean vector and covariance given by $M\Sigma M'$. Here Σ denotes a symmetric matrix with entries given by,

$$\sigma_{j_1,j_2} = \frac{1}{m}\int_0^1 R_{j_1,j_2}\left(c\left(t, j_1\right), c\left(t, j_2\right)\right) dt, \tag{32}$$

for all j_1 and j_2 with $j_1 \neq j_2$. We assume that the inverse of Σ exists, then $M\Sigma M'$ has rank $m-1$. Therefore we restrict attention to the first $m-1$ elements. We estimate Σ using the empirical counterpart of Equation (32), that is for all combinations j_1 and j_2 with $j_1 \neq j_2$

$$\hat{\sigma}_{j_1,j_2} = \frac{1}{k}\sum_{i=1}^{n} \mathbb{I}\left\{X_{i,j_1} > X_{N:N-k}, X_{i,j_2} > X_{N:N-k}\right\}. \tag{33}$$

Finally we define the test statistic as,

$$T_2 = D_{m-1}'\left(\left(M\hat{\Sigma}M'\right)_{m-1}\right)^{-1} D_{m-1}, \tag{34}$$

where the subscript $m-1$ indicates our restriction to the first $m-1$ elements in D and first $m-1$ rows and columns in $M\hat{\Sigma}M'$. Einmahl et al. (2020) provide the joint asymptotic properties under the null hypothesis. We assume the required conditions are satisfied, then as $n \to \infty$,

$$T_2 \xrightarrow{\mathrm{d}} \chi^2_{m-1}. \tag{35}$$

Furthermore we test whether the scedasis $c(\frac{i}{n}, j)$ exhibits a temporal trend, that is, for all stations $j = 1, \ldots, m$,

$$\begin{aligned} H_0 &: C_j(s) = sC_j(1), \\ H_a &: C_j(s) \neq sC_j(1). \end{aligned} \tag{36}$$

If we can reject this null hypothesis, we prove the presence of heteroscedastic extremes, i.e. we prove that the frequency of extreme events is not constant over the considered time period. Note that this test is similar to the test proposed in Section 4.3. However recall that, differently from Section 4.3, we defined a common threshold $X_{N:N-k}$, an high empirical quantile of all $N = n \cdot m$ observations to investigate spatial coherence, therefore the results of these tests may differ.

Similar to Einmahl et al. (2016) in the time dimension, Einmahl et al. (2020) propose a Kolmogorov-Smirnov-type test based on the process $\sqrt{k}\left(\hat{C}_j(s) - s\hat{C}_j(1)\right)/\sqrt{\hat{C}_j(1)}$ for $0 \leq s \leq 1$. For all stations $j = 1, \ldots, m$, under the null hypothesis we have,

$$\left\{\sqrt{k\hat{C}_j(1)}\left(\frac{\hat{C}_j(s)}{\hat{C}_j(1)} - s\right)\right\}_{s\in[0,1]} \xrightarrow{d} \{B(s)\}_{s\in[0,1]}. \tag{37}$$

We consider a Kolmogorov-Smirnov-type test statistic T_3 defined as follows,

$$T_3 := \sup_{0\leqslant s\leqslant 1}\left|\hat{C}(s) - s\hat{C}_j(1)\right|\frac{1}{\sqrt{\hat{C}_j(1)}}. \tag{38}$$

Einmahl et al. (2020) provide the asymptotic distribution of the test statistic under the null hypothesis. We assume the required conditions are satisfied, then as $T \to \infty$,

$$\sqrt{k}T_3 \xrightarrow{\mathrm{d}} \sup_{0\leqslant s\leqslant 1}|B(s)|, \tag{39}$$

where B denotes a standard Brownian bridge.

4.6 Pseudo Maximum Likelihood estimation

We estimate the extreme value parameters in the case of spatially dependent and non-identically distributed observations via a pseudo Maximum Likelihood procedure based on pooling all $N = n \cdot m$ observations (Einmahl et al., 2020).

Recall that $X_{N:N-k}$ denotes the $(N-k)-th$ order statistic of the pooled observations. From Section 4.1 we recall that *independent* exceedances over a high threshold are approximately distributed as a GP distribution. Hence, we let the misspecified log-likelihood function with parameter space $(\gamma, \sigma) \in \mathbb{R} \times (0, \infty)$ be given by:

$$L\left(\gamma, \sigma_{N/k} \mid N, k\right) = \sum_{i=1}^{k} \ell\left(\gamma, \sigma_{N/k}, X_{N:N-i+1} - X_{N:N-k}\right), \tag{40}$$

where $\ell(\gamma, \sigma, x)$ denotes the well known GP distribution, with score equations given by,

$$\begin{cases} \dfrac{\partial}{\partial \gamma} L\left(\gamma, \sigma_{N/k} \mid N, k\right) = 0 \\ \dfrac{\partial}{\partial \sigma} L\left(\gamma, \sigma_{N/k} \mid N, k\right) = 0. \end{cases} \tag{41}$$

For $\gamma > -\frac{1}{2}$, under the assumption that the required conditions are met, there exists a unique sequence of estimators $(\hat{\gamma}, \hat{\sigma}\left(\frac{N}{k}\right))$ that solve the score equations and tend to the true unknown extreme value parameters (Einmahl et al., 2020).

In the space-time framework, this $\hat{\gamma}$ represents estimated constant extreme value index over time and space. Estimating the extreme value index using the pooled observations greatly improves estimation accuracy. Furthermore, the estimated parameters can be interpreted as respectively the shape and scale parameter of the baseline distribution F_0 in Equation (10).

It is evident that choosing a proper common threshold is essential to this method, recall that this comprises a trade-off between variance and bias. Several heuristics for choosing the optimal extreme value threshold $X_{N:N-k}$ exists, most of which are based on eye-balling estimated parameter graphs as a function of upper pseudo order statistics k. We use the sequence of pseudo Maximum Likelihood estimators $(\hat{\gamma}, \hat{\sigma}\left(\frac{N}{k}\right))$ as a function of k to identify a plateau of stability.

4.7 Heteroscedastic exceedance probabilities

Finally, we use the estimated space-time scedasis and extreme value parameters to estimate time varying risk measures. More specifically, *if* we identify heteroscedastic extremes, we extend the estimation of exceedance probabilities in the i.i.d. case by incorporating heteroscedasticity via the estimated scedasis function in univariate and bivariate setting.

In the univariate setting (i.e. for a given point in time $(\frac{i}{n})$ at location j) this heteroscedastic exceedance probability is defined as,

$$\hat{p}_z\left(\frac{i}{n}, j\right) = \operatorname{pr}\left\{X_{i,j} > z\right\}. \tag{42}$$

Let Z denote observations from baseline distribution F_0. Equation (10) provides a relation between the exceedance probability of $X_{i,j}$, scedasis function c and the exceedance probability of Z. Obviously we do not have any observations of baseline distribution F_0. Ferreira et al. (2017) propose a homogenization procedure to obtain (high order) pseudo observations of Z from real observations $X_{i,j}$ for $i = 1, \ldots, n$ and $j = 1, \ldots, m$.

More specifically, the tail quantile function of the original process is related to the tail quantile function of a stationary process Z via the scedasis function such that we can obtain pseudo observations $\hat{Z}_{i,j}$ for $i = 1, \cdots, n$ and $j = 1, \cdots, m$, as

$$\hat{Z}_{i,j} = \left\{ \hat{c}\left(\frac{i}{n}, j\right) \right\}^{-\hat{\gamma}} X_{i,j} - \hat{\sigma}\left(\frac{N}{k}\right) \frac{1 - \left\{ \hat{c}\left(\frac{i}{n}, j\right) \right\}^{-\hat{\gamma}}}{\gamma} \left\{ 1 - \frac{\hat{\gamma} X_{N:N-k}}{\hat{\sigma}\left(\frac{N}{k}\right)} \right\}, \tag{43}$$

with $\hat{\gamma}, \hat{\sigma}\left(\frac{N}{k}\right)$ and $\hat{c}$ estimated as discussed in the previous sections. Using the pseudo observations of Z, we can estimate the exceedance probability for a given point in time $(\frac{i}{n})$ at location j as

$$\hat{p}_z\left(\frac{i}{n}, j\right) = \hat{c}\left(\frac{i}{n}, j\right) \frac{k}{N} \left\{ 1 + \hat{\gamma} \frac{z - Z_{N:N-k}}{\hat{\sigma}\left(\frac{N}{k}\right)} \right\}^{-1/\tilde{\gamma}}. \tag{44}$$

In the multivariate setting (i.e. for a given point in time $(\frac{i}{n})$ the heteroscedastic exceedance probability is defined as,

$$p_z\left(\frac{i}{n}, j_1, j_2\right) = \text{pr}\left\{ X_{i,j_1} > z \text{ and } X_{i,j_2} > z \right\}. \tag{45}$$

Estimating joint exceedance probabilities requires an extremal dependence structure. Due to the spatial context of our framework it is reasonable to consider a parametric approach based on longitudinal and latitudinal distances. Since we have obtained homogenized pseudo observations $\hat{Z}_{i,j}$ for $i = 1, \cdots, n$ and $j = 1, \cdots, m$, we can apply spatial extreme value statistics for stationary extremes. We characterize variation in space via longitudinal ($h_1 \in \mathbb{R}$) and latitudinal distances ($h_2 \in \mathbb{R}$), respectively. Following Ferreira et al. (2017), we use the *Brown-Resnick* process associated with the power variogram that allows for potential geometric anisotropy to model the spatial tail dependence structure. For further details we refer to Ferreira et al. (2017) and especially the references therein.

Let the L-dependence function be given by,

$$L_{j_1,j_2}(x, y) = \lim_{t \to \infty} t \, \text{pr}\left\{ Z_{j_1} > U_Z(tx) \text{ or } Z_{j_2} > U_Z(ty) \right\}. \tag{46}$$

and let the power variogram be given by,

$$v_\vartheta(h_1, h_2) = \left\| \begin{pmatrix} b_1 \cos\theta & b_1 \sin\theta \\ -b_2 \sin\theta & b_2 \cos\theta \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \right\|^\alpha, \tag{47}$$

with $b_1, b_2 > 0, \theta \in (-\pi/2, \pi/2]$ and $\alpha \in (0, 2]$. In the case of *Brown-Resnick* models the bivariate marginals are known. Following De Haan et al. (2006) and Kabluchko et al. (2009) we have that,

$$L_{j_1,j_2}(x,y) = \frac{1}{x}\Phi\left[\frac{v_\vartheta(h_1,h_2)^{1/2}}{2} + \frac{\log(\frac{y}{x})}{v_\vartheta(h_1,h_2)^{1/2}}\right] + \frac{1}{y}\Phi\left[\frac{v_\vartheta(h_1,h_2)^{1/2}}{2} + \frac{\log(\frac{x}{y})}{v_\vartheta(h_1,h_2)^{1/2}}\right]. \quad (48)$$

We estimate the parameters as follows. First, we estimate the well-known tail dependence coefficient with asymptotically normal non-parametric estimator given by,

$$\hat{L}_{j_1,j_2}(1,1) = \frac{1}{k'}\sum_{l=1}^{n} 1_{\left\{\hat{Z}_{i,j_1} > \hat{Z}_{n:n-k',j_1} \text{ or } \hat{Z}_{i,j_2} > \hat{Z}_{n:n-k',j_2}\right\}}, \quad (49)$$

where we use $k' = \min\left\{\hat{C}_{j_1}(1), \hat{C}_{j_2}(1)\right\} \times k$. Rewriting Equation (48) for $x = y = 1$ with estimated tail dependence coefficient $\hat{L}_{j_1,j_2}(1,1)$ yields the estimator for the *empirical* variogram, that is,

$$\hat{v}\left(h_1,h_2\right) = 4\left[\Phi^{\leftarrow}\left\{\frac{\hat{L}_{j_1,j_2}(1,1)}{2}\right\}\right]^2. \quad (50)$$

Thereafter, we estimate the variogram parameter estimates by numerically minimizing the squared difference between the *empirical* and parametric variogram for $\vartheta = (b_1, b_2, \theta, \alpha)$, that is,

$$\min_{b_1,b_2,\theta,\alpha} \sum_{1 \le j_1, j_2 \le m} \left\{\hat{v}\left(h_1,h_2\right) - v_\vartheta\left(h_1,h_2\right)\right\}^2. \quad (51)$$

To estimate the joint exceedance probability for a given point in time $(\frac{i}{n})$ at locations j_1 and j_2, we use that the joint probability (intersection) is the sum of the marginal probabilities minus the union. Hence, combining the univariate exceedance probabilities and the L-dependence function with the variogram parameter estimates $\hat{\vartheta} = (\hat{b}_1, \hat{b}_2, \hat{\theta}, \hat{\alpha})$ we obtain,

$$\hat{p}_z\left(\frac{i}{n}, j_1, j_2\right) = \hat{p}_z\left(\frac{i}{n}, j_1\right) + \hat{p}_z\left(\frac{i}{n}, j_2\right) - \frac{k}{N}\hat{L}_{j_1,j_2}\left(\left(\frac{N}{k}\hat{p}_z\left(\frac{i}{n}, j_1\right)\right)^{-1}, \left(\frac{N}{k}\hat{p}_z\left(\frac{i}{n}, j_2\right)\right)^{-1}\right). \quad (52)$$

5 Results

5.1 24 hour precipitation accumulation

5.1.1 Univariate analysis per station

We assume the 24 hour precipitation accumulation extremes are i.i.d. and we model their exceedances above threshold $X_{n:n-k}$ as a GP distribution for all stations individually. Choosing a proper threshold $X_{n:n-k}$, i.e. the number of upper order statistics k, is essential to this method. Since the scale estimator is a function of k, we preferably select one k for all stations to be able to compare the scale parameter across space. Stability plots for the Maximum Likelihood estimator of the extreme value index for the summer and winter season are displayed in the appendix. For the summer season we can identify a reasonably stable plateau around $k = 150$ for all stations. For the winter season we can identify a reasonably stable plateau around $k = 200$ for all stations.

Maximum Likelihood estimates for the shape and scale parameters of the GP distribution are presented in Table 9. We observe that the extreme value index γ, which drives the severity of extremes, seems to vary over space. However, we also recognize the substantial standard error of the estimate. As expected, precipitation accumulation has a small positive or slightly negative extreme value index.

Table 3: Parameter estimates for GP distribution for 24 hour precipitation accumulation for the 10 weather stations separately in the summer and winter season.

		De Kooy	Schiphol	De Bilt	Leeuwarden	Eelde	Twenthe	Vlissingen	Rotterdam	Volkel	Maastricht
Summer Season	Shape $\hat{\gamma}$	-0.01	0.1	0	0.02	-0.05	0.13	0.06	0.13	0.17	0.07
		(0.09)	(0.1)	(0.09)	(0.09)	(0.08)	(0.08)	(0.08)	(0.09)	(0.1)	(0.09)
	Scale $\hat{\sigma}$	100	86.2	75.28	89.53	86.02	80.56	90.43	88.71	64.57	88.66
		(11.99)	(11.08)	(9.36)	(10.87)	(9.96)	(9.18)	(10.38)	(11.03)	(8.26)	(10.85)
Winter Season	Shape $\hat{\gamma}$	-0.07	-0.04	-0.01	-0.09	-0.08	-0.02	-0.07	-0.07	0.01	0.09
		(0.06)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
	Scale $\hat{\sigma}$	54.01	59.7	59.8	54.18	56.84	56.17	56.94	62.29	57.82	47.57
		(5.13)	(6.1)	(6.04)	(5.55)	(5.62)	(5.44)	(5.74)	(6.36)	(5.75)	(4.65)

Note: We determine the number of upper pseudo order statistics using parameter stability plots. We choose $k = 150$ for the summer season and $k = 200$ for the winter season.

For the summer season, γ is comprised between -0.05 and 0.17. Recall that the right endpoint x^*, of the parent distribution F is determined by the extreme value index. Since we find for

the summer season that the majority of the weather stations $\gamma > 0$, we conclude that 24 hour precipitation extremes in the summer season have tend to have an infinite right endpoint. For the winter season, γ is comprised between -0.09 and 0.09 with less spatial variation. Since we find for the winter season that the majority of the weather stations $\gamma < 0$, we tend to conclude that 24 hour precipitation extremes in the winter season have a finite right endpoint, however, due to the large standard error we might not be able to reject $\gamma > 0$.

Additionally we notice that the scale parameter σ, which drives the frequency of extremes, seems to varies over space. As expected, increasing the number of pseudo order statistics k used to estimate the parameters, e.g. comparing the summer and winter season estimates, decreases the standard error of the parameter estimates.

We use the estimated extreme value parameters to estimate exceedance probabilities of high level thresholds, for which exceedances are scare in the data. We aim to select a threshold that is known to generate claims if passed. Recall that the Dutch *Verbond Van Verzekeraars* reports a threshold of *38mm* per 24 hour, after which the number of incurred claims rapidly increases. We select two thresholds, *40mm* and *50mm* per 24 hour. Table 4 presents the estimated exceedance probabilities.

Table 4: Estimated exceedance probabilities of high level thresholds for 24 hour precipitation accumulation for the 10 weather stations separately in the summer and winter season.

		De Kooy	Schiphol	De Bilt	Leeuwarden	Eelde	Twenthe	Vlissingen	Rotterdam	Volkel	Maastricht
Summer Season	*40mm*	0.165	0.170	0.127	0.146	0.079	0.147	0.146	0.216	0.106	0.163
	50mm	0.059	0.071	0.041	0.051	0.019	0.065	0.058	0.098	0.045	0.066
Winter Season	*40mm*	0.004	0.014	0.026	0.002	0.004	0.013	0.005	0.014	0.022	0.021
	50mm	0.000	0.002	0.004	0.000	0.000	0.002	0.000	0.001	0.004	0.006

Note: Exceedance probabilities are reported in % and should be interpreted as the probability of exceeding the respective threshold on a given day in the respective season.

Despite the variation in extreme value parameters, the exceedance probabilities are to a certain extent spatially coherent. The estimated exceedance probabilities of *40mm* per 24 hour are comprised between 0.08% and 0.22% for the summer season and between $< 0.01\%$ and 0.03% for the winter season. The estimated exceedance probabilities of *50mm* per 24 hour are comprised between 0.2% and 0.1% for the summer season and between $< 0.01\%$ and 0.01% for the winter season. The exceedance probabilities should be interpreted as the probability of exceeding the respective

threshold on a single given day. Hence, as expected the probability of extreme 24 hour precipitation is higher in the summer season. In the estimated exceedance probabilities of *50mm* per 24 hour we recognize the potential finite right endpoint of winter season extremes.

Threshold exceedance probabilities on a daily level can easily be transformed to more tangible risk measures if we assume independence. Dividing 1 by the estimated daily exceedance probability, yields the expected number of days in which precisely one extreme risk materializes. For instance, it is expected that the risk of an extreme precipitation event (*40mm*) materializes in *Rotterdam* on average once every 455 days in the summer season. Whereas for *Eelde* it is expected that the risk of an extreme precipitation event (*40mm*) materializes on average once every 1250 days in the summer season.

Figure 3: Estimated scedasis functions from June 1974 until June 2020 for 24 hour precipitation accumulation at *Schiphol* weather station in summer and winter season, respectively.

We relax the assumption of i.i.d. observations and model the (potentially) continuously evolving distribution functions over time using continuous and positive scedasis function c. In this framework, the scedasis function can be interpreted as the relative frequency of extremes over time. We estimate the non-parametric scedasis function using boundary corrected Kernel Density Estimation for $0 \leq s \leq 1$. Figure 3 depicts the estimated non-parametric scedasis function at weather station *Schiphol* for summer and winter season, respectively.

We recognize the inter annual variability in the frequency of extremes, for both summer and winter season. In addition, the summer season scedasis function seems to indicate an upward sloping trend, indicating increasing frequency of extremes over time. On the contrary, we do not recognize a trend that indicates long term changes in frequency of extremes for the winter season. The estimated scedasis functions of the remaining weather stations are deferred to the appendix.

Since the scedasis function is eminently suited for visual exploratory analysis we estimate the integrated scedasis function $C(s)$ on $0 \leq s \leq 1$ to formally test for heteroscedastic extremes in time. We calculate the test statistic T_1 as the supremum on $0 \leq s \leq 1$ of the absolute difference between estimated integrated scedasis function $\hat{C}(s)$ and the identity function. Table 5 reports the thresholds used based on the selected number of upper pseudo order statistics k. In addition, Table 5 presents the estimated test statistic T_1 and corresponding P-value derived from the Monte Carlo simulated probability density function of $\sup_{0 \leqslant s \leqslant 1} |B(s)|$, for all stations in the winter and summer season.

Table 5: Results of Kolmogorov-Smirnov type test statistic for 24 hour precipitation accumulation for the 10 weather stations separately in the summer and winter season.

		De Kooy	Schiphol	De Bilt	Leeuwarden	Eelde	Twenthe	Vlissingen	Rotterdam	Volkel	Maastricht
Summer Season	Threshold	146.40	153.20	161.06	153.73	139.18	141.30	138.78	162.97	147.69	149.79
	T_1	0.09	0.13	0.11	0.11	0.11	0.05	0.09	0.08	0.06	0.08
	P-value	0.19	0.01	0.04	0.04	0.06	0.76	0.16	0.27	0.69	0.27
Winter Season	Threshold	112.75	114.65	127.71	108.81	111.84	113.40	104.91	125.62	111.15	106.50
	T_1	0.04	0.05	0.04	0.06	0.04	0.06	0.04	0.06	0.06	0.05
	P-value	0.88	0.55	0.78	0.55	0.88	0.34	0.88	0.34	0.44	0.67

Note: P-values for T_1 depend on the number of upper pseudo order statistics which we determine using parameter stability plots. We choose $k = 150$ for the summer season and $k = 200$ for the winter season.

We cannot compare the thresholds across the seasons due to difference in upper pseudo order statistics k used for estimation. For the summer season, homoscedastic extremes in time are rejected at 5% significance level for *Schiphol, De Bilt* and *Leeuwarden*. The rejection of constant frequency of extremes coincides with our initial exploratory observation of the trend in the estimated scedasis function in Figure 3 for *Schiphol*. For all other stations we cannot reject homoscedastic extremes in time at any meaningful significance level. We conclude that the spatial coherence is low. For the winter season, homoscedastic extremes in time cannot be rejected at any meaningful significance level for all stations. Hence, again, the outcome of the formal test coincides with our initial exploratory observation of the estimated scedasis function in Figure 3.

5.1.2 Multivariate analysis

We further relax the assumption of spatial independence, and invoke a space-time framework. This allows us to estimate the extreme value parameters using by pooling all $N = n \cdot 10$ observations,

which greatly improves estimation accuracy. We estimate $(\gamma, \sigma(\frac{N}{k}))$ using the pseudo Maximum Likelihood Estimator. These parameters can be interpreted as respectively the extreme value index and the scale parameter of the baseline distribution F_0 in Equation (10). In addition, γ, represents the common extreme value index over space and time.

This choose a proper threshold $X_{N:N-k}$, we estimate stability plots for the Maximum Likelihood estimator of the extreme value index as a function of k. In this framework, k represents the number of upper order statistics from the set of pooled observations. Figure 4 presents the estimated stability plots for the Maximum Likelihood estimates of the extreme value index γ, as a function of the number of pseudo order statistics k, for summer and winter season.

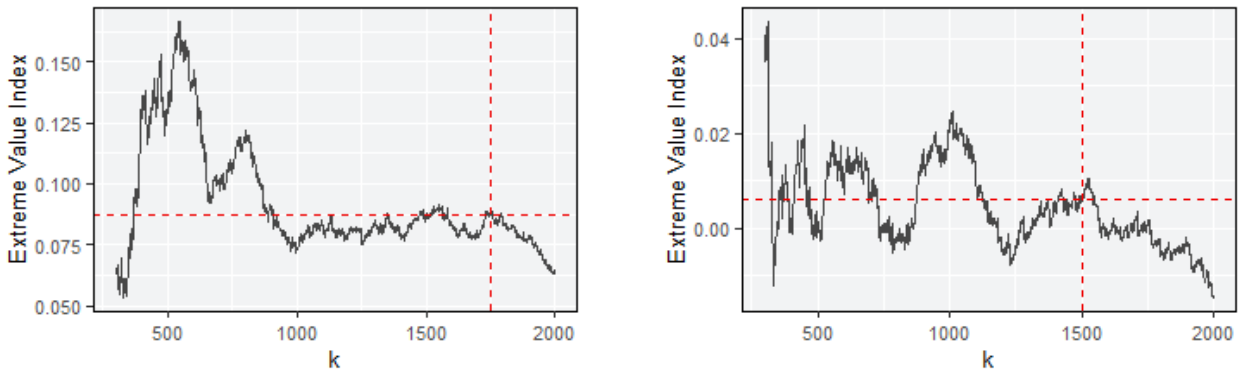

Figure 4: Parameter stability plots of the Maximum Likelihood estimates of the extreme value index γ, as a function of the number of upper pseudo order statistics k, for 24 hour precipitation accumulation for summer and winter season.

We aim to select k as large as possible, without including non-tail observations that yield bias. For the summer season we identify a plateau of stability around $k = 1750$ after which we recognize bias via a steeply decreasing trends. Fixing $k = 1750$, we find parameter point estimates $\hat{\gamma} = 0.087$ and $\hat{\sigma}(\frac{N}{k}) = 81.132$. For the winter season we identify a plateau of stability around $k = 1500$ after which we recognize bias via a steeply decreasing trends. Fixing $k = 1500$, we find parameter point estimates $\hat{\gamma} = 0.006$ and $\hat{\sigma}(\frac{N}{k}) = 53.37$.

To test whether there are significant differences in cumulative number of extreme events across the Netherlands i.e. whether the frequency of extreme events is constant over space, we estimate the space-time integrated scedasis function at the end of the considered time period. Hence we obtain $\hat{C}_j(1)$ for $j = 1, \ldots, 10$. We formally test for constant frequency of extremes over space using test statistic T_2. Under the null hypothesis (constant frequency of extremes) T_2 follows a χ^2-distribution, with 9 degrees of freedom.

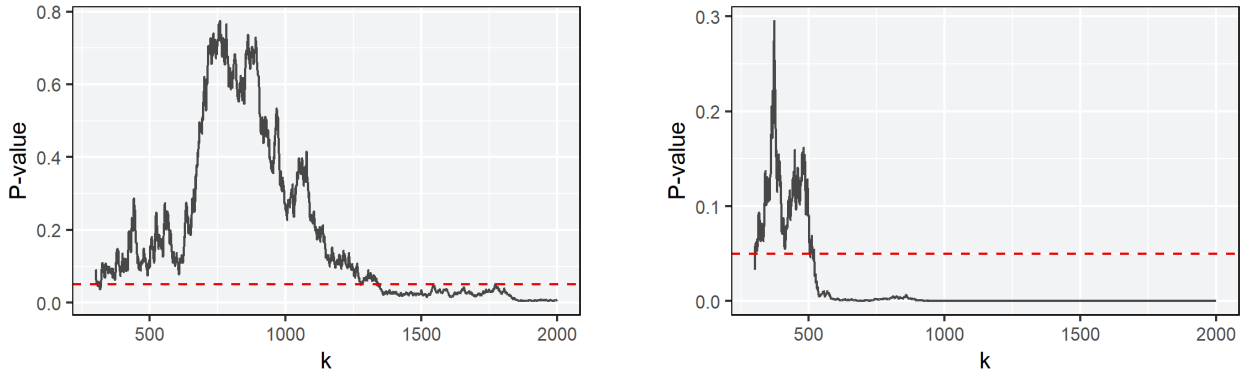

Figure 5: P-values obtained for test of constant frequency of extremes over space, as a function of the number of upper pseudo order statistics k for 24 hour precipitation accumulation for summer and winter season.

Figure 5 presents the P-values for the test statistic T_2 as a function of the number of pseudo order statistics k, for summer and winter season. In the summer season, for $250 < k < 1300$ the obtained P-values stay above 5%, in other words, there is no statistical evidence that the frequency of extremes varies across space for $250 < k < 1300$. On the contrary, for $k > 1300$ the obtained P-values stay below 5%. In the winter season the test results are less ambiguous, for $k > 500$ the P-values are virtually equal to zero. Based on parameter stability, we selected $k = 1750$ as a proper threshold in the summer season. Therefore we conclude that we have evidence that the frequency of having 24 hour precipitation accumulation extremes is not constant across space in both the summer and winter season.

Furthermore, similar to in the univariate framework, we test whether the space-time scedasis exhibits a temporal trend. However, since we aim to compare the results over space we use our common extreme value threshold $X_{N:N-k}$. As a consequence, the number of observations per station may differ and are given by $k \cdot \hat{C}_j(1)$ for all $j = 1, \ldots, 10$. Hence, the results of the multivariate and univariate test for heteroscedastic extremes over time could be different. We estimate the space-time integrated scedasis function $C_j(s)$ on $0 \leq s \leq 1$, such that we can calculate the test statistic T_3.

Table 6 reports the calculated test statistics and corresponding P-values derived from the Monte Carlo simulated probability density function of $\sup_{0 \leqslant s \leqslant 1} |B(s)|$. The results confirm the findings in the univariate framework, as expected given the relatively stable thresholds across space as reported in Table 5.

Table 6: Results Kolmogorov-Smirnov type test statistic for 24 hour precipitation sum in the summer and winter season

		De Kooy	Schiphol	De Bilt	Leeuwarden	Eelde	Twenthe	Vlissingen	Rotterdam	Volkel	Maastricht
Summer Season	T_3	0.02	0.05	0.04	0.03	0.03	0.02	0.03	0.03	0.02	0.02
	P-value	0.21	0.00	0.02	0.04	0.06	0.54	0.20	0.15	0.22	0.35
Winter Season	T_3	0.01	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.02	0.03
	P-value	0.92	0.56	0.55	0.77	0.74	0.49	0.25	0.26	0.48	0.18

Note: Thresholds corresponding to the summer and winter season are 13.8mm and 13.1mm, respectively. P-values for T_3 depend on the number of upper pseudo order statistics which vary across the stations as $k \cdot \hat{C}_j(1)$ for all $j = 1, \ldots, 10$.

Since we have identified heteroscedastic extremes for a number of weather stations in the summer season, we estimate time varying exceedance probabilities. We incorporate heteroscedasticity of extremes via the estimated space-time scedasis function. We perform the homogenization procedure to obtain pseudo observations $Z_{i,j}$ by using estimates γ, $\sigma(\frac{N}{k})$ and $\hat{c}\left(\frac{i}{n}, j\right)$ for $i = 1, \ldots, n$ and $j = 1, \ldots, 10$. Thereafter we estimate the time varying exceedance probability using homogenized threshold $Z_{N:N-k}$.

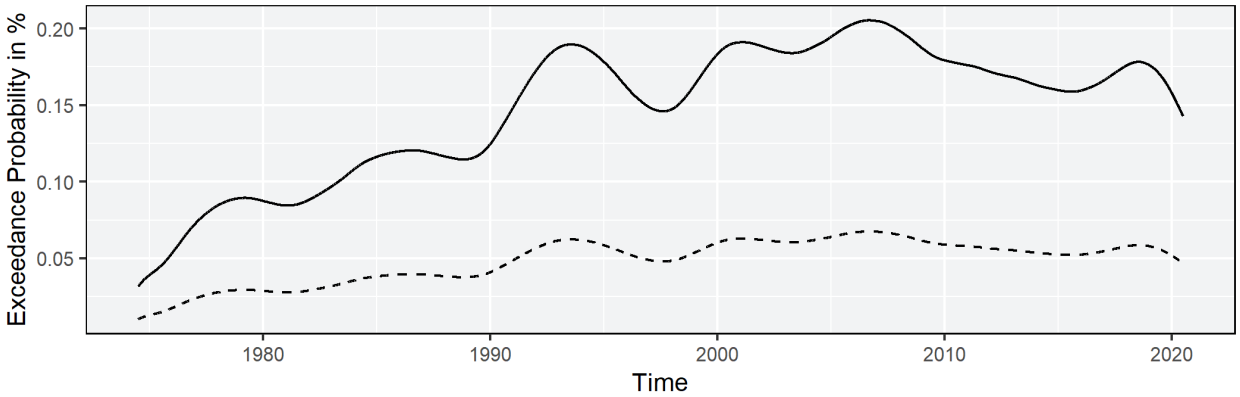

Figure 6: Estimated time varying daily exceedance probabilities in % from June 1974 until June 2020 of *40mm* and *50mm* at *Schiphol* for the 24 hour precipitation accumulation in the summer season.

Figure 6 depicts the estimated time varying exceedance probabilities of *40mm* and *50mm* at *Schiphol* for the 24 hour precipitation sum in the summer season. We clearly identify an increase in extreme event probability. In the summer of 1971 the exceedance probability of *40mm* per 24 hour is estimated to be less than 0.05%, that is, one extreme precipitation event (*40mm*) is expected approximately every 2000 days. Whereas in the summer of 2005, the exceedance probability is

slightly under 0.2%, that is, one extreme precipitation event (*40mm*) is expected approximately every 500 days. Hence, we identify that the probability of observing at least *40mm* precipitation on a given day, has quadrupled between 1971 and 2005. The estimated time varying exceedance probabilities of the remaining weather stations are deferred to the appendix.

Table 7: Variogram parameter estimates for 24 hour precipitation accumulation in the summer and winter season, respectively.

	Summer Season				**Winter Season**			
Parameters	$\hat{b}_1$	$\hat{b}_2$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{b}_1$	$\hat{b}_2$	$\hat{\theta}$	$\hat{\alpha}$
Estimate	0.078	0.118	0.559	0.663	0.028	0.018	1.571	0.882
Std. Error	(0.011)	(0.020)	(0.088)	(0.035)	(0.004)	(0.003)	(0.227)	(0.067)

Note: Parameters are estimated using longitudinal and latitudinal distances between the 10 weather stations.

In general, insurance companies underwrite a spatially diversified portfolio. Therefore, the time varying joint exceedance probabilities are of interest. We use obtained homogenized pseudo observations to characterize variation in space via longitudinal and latitudinal distances. We estimate the well-known tail dependence coefficient, after which we estimate the *empirical* variogram. Thereafter, we estimate the variogram parameter estimates by numerically minimizing the squared difference between the *empirical* and parametric variogram. Table 7 provides the estimated variogram parameters and standard errors for both the summer and winter season, respectively.

By definition, the tail dependence coefficient $L(1,1)$ has lower bound 1, referred to as completely tail dependent, and upper bound 2 referred to as completely tail independent. Figure 7 graphs the estimated empirical tail dependence coefficients $\hat{L}(1,1)$ for all bivariate combinations $(j_1, j_2) \in j = 1, \ldots, 10$ as a function of the Euclidean distance in km, for summer (red) and winter (blue) season. In addition, we graph the parametric tail dependence as a function of Euclidean distance in km for summer (red) and winter (blue) season. We observe higher tail dependence for the weather stations in the winter season compared to the summer season. This is not surprising given the convection thunderstorms in the summer season. Convection thunderstorms arise over a highly heated earth's surface and are therefore a rather local phenomenon.

We use the estimated variogram and tail dependence structure to estimate time varying joint exceedance probabilities that incorporate heteroscedasticity of extremes via the estimated space-time scedasis function. Figure 8 presents the estimated time varying joint exceedance probabilities of *40mm* and *50mm* at *Schiphol* and *De Bilt* (left) and at *Schiphol* and *Eelde* (right). We clearly

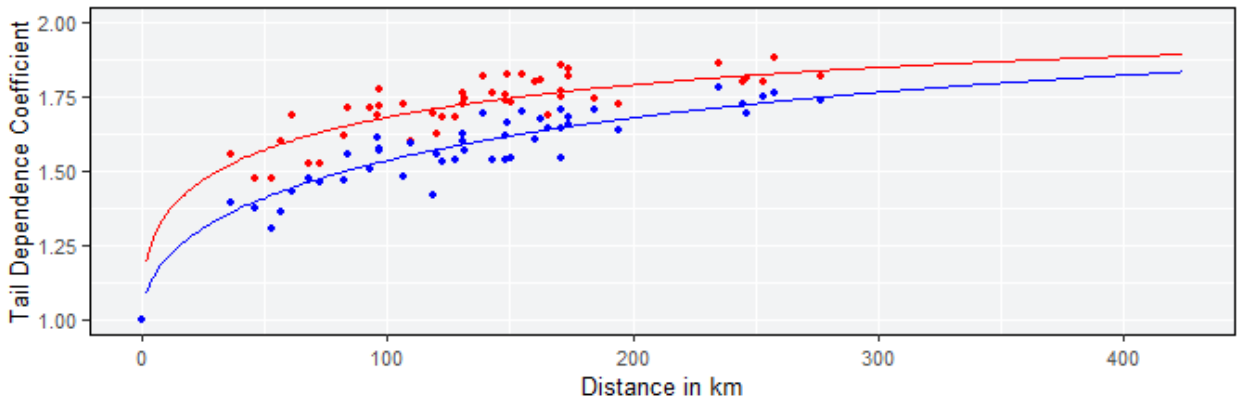

Figure 7: Estimated empirical tail dependence coefficients for all bivariate combinations $(j_1, j_2) \in j = 1, \ldots, 10$ and estimated parametric tail dependence as a function of the Euclidean distance in km for summer (red) and winter (blue) season.

identify the time varying behaviour stemming from the scedasis functions, more specifically, we recognize the trend from Figure 6. Indeed, an increase in marginal probability leads to an increase in joint probability (ceteris paribus). Furthermore, we identify the decreasing tail dependence over space when comparing the joint exceedance probability at *Schiphol* and *De Bilt* (36km distance) and at *Schiphol* and *Eelde* (150.7km distance).

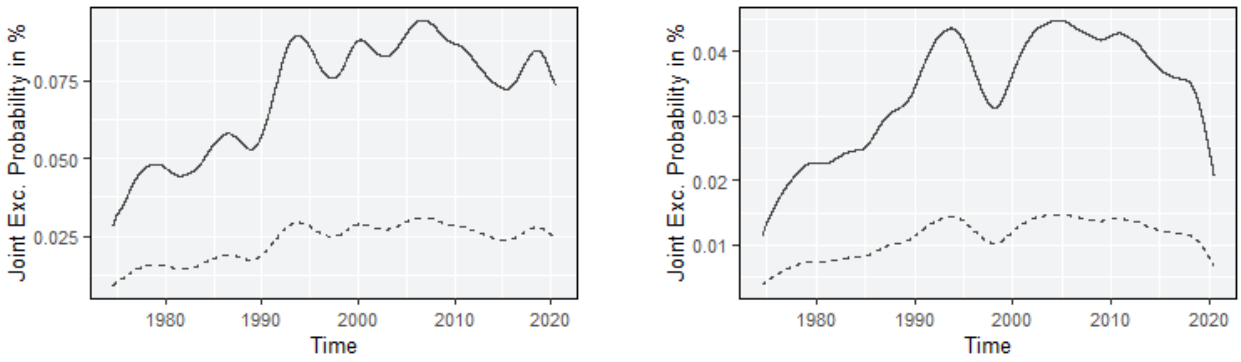

Figure 8: Estimated time varying daily joint exceedance probabilities in % from June 1974 until June 2020 of *40mm* and *50mm* at *Schiphol* and *De Bilt* (left) and at *Schiphol* and *Eelde* (right) for 24 hour precipitation extremes in the winter season.

5.2 Windstorm peak wind speed

5.2.1 Univariate analysis per station

We assume that peak wind speed extremes are i.i.d. and we model their exceedances above threshold $X_{n:n-k}$ as a GP distribution for all stations individually. We select one k for all stations to be able to compare the scale parameter across space. Stability plots for the Maximum Likelihood estimator

of the extreme value index for the summer and winter season are displayed in the appendix. For both the summer ans winter season we can identify a reasonably stable plateau around $k = 200$ for all stations

Maximum Likelihood estimates for the shape and scale parameters of the GP distribution are presented in Table 8. We observe that the extreme value index γ varies substantially over space. However, we also recognize the substantial standard error of the estimate. As expected, peak wind speed has a negative or slightly positive extreme value index.

Table 8: Parameter estimates for GP distribution for peak wind speed for the 13 weather stations separately in the summer and winter season.

		De Kooy	Schiphol	De Bilt	Leeuwarden	Deelen	Eelde	Twenthe	Vlissingen	Rotterdam	Gilze-Rijen	Eindhoven	Volkel	Maastricht
Summer Season	Shape $\hat{\gamma}$	-0.18	-0.13	-0.09	-0.19	0.12	-0.1	-0.16	-0.11	-0.13	0.05	-0.05	0	-0.09
		(0.07)	(0.04)	(0.07)	(0.05)	(0.1)	(0.07)	(0.05)	(0.05)	(0.04)	(0.1)	(0.06)	(0.06)	(0.06)
	Scale $\hat{\sigma}$	31.14	34.79	26.25	31.95	20.8	23.27	25.64	31.5	32.66	19.64	26.92	27.3	31
		(3.12)	(2.85)	(2.59)	(2.81)	(2.62)	(2.36)	(2.25)	(2.75)	(2.59)	(2.39)	(2.51)	(2.5)	(2.87)
Winter Season	Shape $\hat{\gamma}$	-0.14	-0.15	-0.06	-0.05	-0.14	-0.13	-0.15	-0.16	-0.1	-0.09	-0.02	-0.08	-0.29
		(0.05)	(0.05)	(0.08)	(0.07)	(0.07)	(0.06)	(0.07)	(0.06)	(0.05)	(0.07)	(0.09)	(0.06)	(0.05)
	Scale $\hat{\sigma}$	37.78	40.64	31.92	31.3	36.58	38.51	40.82	44.1	34.96	35.94	32.99	37.78	45.9
		(3.36)	(3.6)	(3.41)	(3.04)	(3.63)	(3.5)	(3.91)	(4.2)	(3.02)	(3.57)	(3.82)	(3.59)	(3.83)

Note: We determine the number of upper pseudo order statistics using parameter stability plots. We choose $k = 200$ for both the summer and the winter season.

For the summer season, γ is comprised between -0.19 and 0.12, where especially *Deelen* stands out. For the winter season, γ is comprised between -0.29 and -0.01 with more spatial coherence. Since we find for the summer and winter season the majority of the weather stations have $\gamma < 0$, we tend to conclude that peak wind speed in the summer and winter season have a finite right endpoint, however, due to the large standard error we might not be able to reject $\gamma > 0$. Additionally we notice that the scale parameter σ, which drives the frequency of extremes, seems to varies over space.

We use the estimated extreme value parameters to estimate exceedance probabilities of high level thresholds, for which exceedances are scare in the data. We select two extreme thresholds that are not often exceeded in the data, *25m/s* and *30m/s*. Table 4 presents the estimated exceedance probabilities.

Table 9: Estimated exceedance probabilities of high level thresholds for peak wind speed for the 13 weather stations separately in the summer and winter season.

		De Kooy	Schiphol	De Bilt	Leeuwarden	Deelen	Eelde	Twenthe	Vlissingen	Rotterdam	Gilze-Rijen	Eindhoven	Volkel	Maastricht
Summer Season	*25 m/s*	0.27	0.23	0.03	0.10	0.10	0.03	0.01	0.51	0.15	0.04	0.08	0.16	0.10
	30 m/s	0.01	0.02	0.00	0.00	0.02	0.00	0.00	0.06	0.01	0.01	0.01	0.03	0.01
Winter Season	*25 m/s*	2.09	1.78	0.56	1.12	0.65	0.97	0.49	3.09	1.21	0.58	0.63	0.62	0.83
	30 m/s	0.44	0.41	0.09	0.19	0.10	0.18	0.08	0.89	0.22	0.10	0.13	0.13	0.10

Note: Exceedance probabilities are reported in % and should be interpreted as the probability of exceeding the respective threshold on a given day in the respective season.

The estimated exceedance probabilities of *25m/s* are comprised between 0.01% and 0.27% for the summer season and between 0.49% and 3.09% for the winter season. The estimated exceedance probabilities of *30m/s* are comprised between $< 0.01\%$ and 0.6% for the summer season and between 0.1% and 0.89% for the winter season. We identify the immediate difference between summer and winter season. Furthermore, we identify the coastal stations via their extreme peak wind speed probability, *De Kooy* and *Vlissingen*.

In the winter season the risk of extreme peak wind speed (*25m/s*) in *vlissingen* is expected to materialize approximately once every 32 days, and can therefore hardly be defined as extreme. On the contrary, in *Twenthe* extreme peak wind speed (*25m/s*) is expected once every 10000 days in the winter season, and can therefore be defined as highly extreme. In the estimated exceedance probabilities of *30m/s* in the summer season we recognize the (potential) finite right endpoint of the extremes.

Figure 9: Estimated scedasis functions from January 1971 until June 2020 for 24 hour precipitation accumulation at *Schiphol* weather station in summer and winter season, respectively.

Next, we relax the assumption of i.i.d. observations and model the (potentially) continuously evolving distribution functions over time using continuous and positive scedasis function c. In this framework, the scedasis function can be interpreted as the relative frequency of extremes over time. We estimate the non-parametric scedasis function using boundary corrected Kernel Density Estimation for $0 \leq s \leq 1$. Figure 9 depicts the estimated non-parametric scedasis function at weather station *Schiphol* for summer and winter season, respectively.

We recognize the inter annual variability in the frequency of extremes, for both summer and winter season. In addition, the summer season scedasis function does not seem to indicate a trend. On the contrary, for the winter season we do recognize a steep decrease after 1990 and correspondingly, a lower frequency of peak wind speed extremes in the second half of the considered time period (1995-2020). The estimated scedasis functions of the remaining weather stations are deferred to the appendix.

Since the scedasis function is eminently suited for visual exploratory analysis we estimate the integrated scedasis function $C(s)$ on $0 \leq s \leq 1$ to formally test for heteroscedastic extremes in time. We calculate the test statistic T_1 as the supremum on $0 \leq s \leq 1$ of the absolute difference between estimated integrated scedasis function $\hat{C}(s)$ and the identity function. Table 10 reports the thresholds used based on the selected number of upper pseudo order statistics k. In addition, Table 10 presents the estimated test statistic T_1 and corresponding P-value derived from the Monte Carlo simulated probability density function of $\sup_{0 \leqslant s \leqslant 1} |B(s)|$, for all stations in the winter and summer season.

Table 10: Results of Kolmogorov-Smirnov type test statistic for peak wind speed for the 13 weather stations separately in the summer and winter season.

		De Kooy	Schiphol	De Bilt	Leeuwarden	Deelen	Eelde	Twenthe	Vlissingen	Rotterdam	Gilze-Rijen	Eindhoven	Volkel	Maastricht
Summer Season	Threshold	189	175	150	170	164	164	150	200	170	159	160	154	160
	T_1	0.12	0.04	0.06	0.07	0.14	0.10	0.09	0.06	0.05	0.10	0.09	0.06	0.06
	P-value	0.01	0.88	0.55	0.26	0.00	0.05	0.07	0.44	0.67	0.02	0.05	0.55	0.44
Winter Season	Threshold	237	230	200	220	200	210	185	252	220	195	200	195	201
	T_1	0.12	0.14	0.14	0.16	0.17	0.12	0.08	0.14	0.10	0.15	0.12	0.09	0.12
	P-value	0.01	0.00	0.00	0.00	0.00	0.01	0.20	0.00	0.05	0.00	0.01	0.07	0.01

Note: P-values for T_1 depend on the number of upper pseudo order statistics which we determine using parameter stability plots. We choose $k = 200$ for the summer and winter season.

For the summer season, homoscedastic extremes in time are rejected at 5% significance level for *De Kooy, Deelen* and *Gilze-Rijen*. For all other stations we cannot reject homoscedastic extremes in time at any meaningful significance level. Hence the formal test confirms our initial exploratory observation of the trend in the estimated scedasis function in Figure 9 for *Schiphol*. We conclude that the spatial coherence is low. For the winter season, homoscedastic extremes in time are rejected at 5% significance level for all stations except *Twenthe* and *Volkel*. Hence the formal test confirms our initial exploratory observation of the estimated scedasis function in Figure 3 for *Schiphol*. We conclude that for the winter season, the spatial coherence is high.

5.2.2 Multivariate analysis

We further relax the assumption of spatial independence, and invoke a space-time framework. This allows us to estimate the extreme value parameters using by pooling all $N = n \cdot 13$ observations, which greatly improves estimation accuracy. We estimate $(\gamma, \sigma(\frac{N}{k}))$ using the pseudo Maximum Likelihood Estimator. These parameters can be interpreted as respectively the extreme value index and the scale parameter of the baseline distribution F_0 in Equation (10). In addition, γ, represents the common extreme value index over space and time.

This choose a proper threshold $X_{N:N-k}$, we estimate stability plots for the Maximum Likelihood estimator of the extreme value index as a function of k. In this framework, k represents the number of upper order statistics from the set of pooled observations. Figure 10 presents the estimated stability plots for the Maximum Likelihood estimates of the extreme value index γ, as a function of the number of pseudo order statistics k, for summer and winter season.

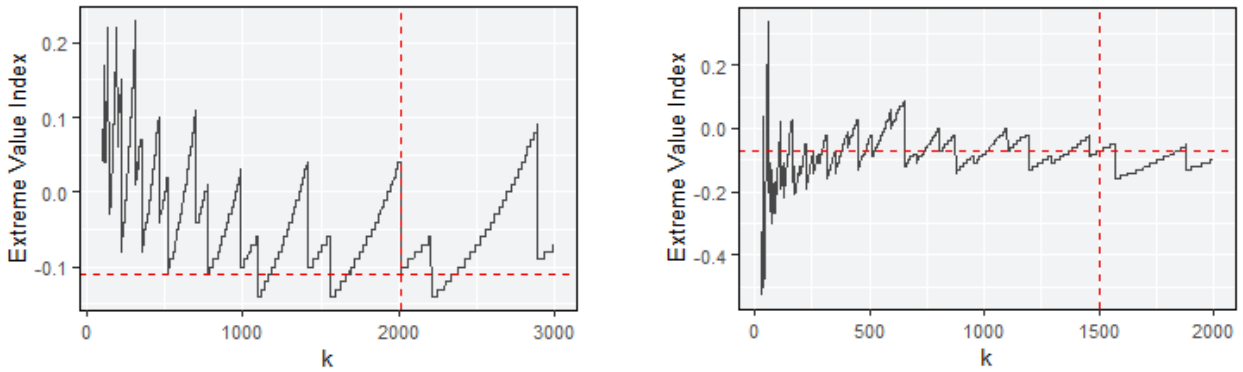

Figure 10: Parameter stability plots of the Maximum Likelihood estimates of the extreme value index γ, as a function of the number of upper pseudo order statistics k, for peak wind speed for summer and winter season.

Since the peak wind speed data presents many ties, we observe stair-like graphs in Figure 10. For the summer season it is hard to identify a plateau of stability. Note that for $1000 \leq k \leq 2000$, $\hat{\gamma}$ is relatively stable around -0.1 after each staircase drop. Since we aim to select k as large as possible, without including non-tail observations that yield bias, we fix $k = 2000$. This yields parameter point estimates $\hat{\gamma} = -0.105$ and $\hat{\sigma}(\frac{N}{k}) = 30.79$. For the winter season we identify a plateau of stability around $k = 1500$. Fixing $k = 1500$, we find parameter point estimates $\hat{\gamma} = -0.071$ and $\hat{\sigma}(\frac{N}{k}) = 33.30$.

To test whether there are significant differences in cumulative number of extreme events across the Netherlands i.e. whether the frequency of extreme events is constant over space, we estimate the space-time integrated scedasis function at the end of the considered time period. Hence we obtain $\hat{C}_j(1)$ for $j = 1, \ldots, 13$. We formally test for constant frequency of extremes over space using test statistic T_2. Under the null hypothesis (constant frequency of extremes) T_2 follows a χ^2-distribution, with 12 degrees of freedom. For all k, this yields P-values that are virtually zero, both in the summer and winter season.

Table 11: Integrated scedasis at $s = 1$ for peak wind speed for summer and winter season for all stations separately.

		De Kooy	Schiphol	De Bilt	Leeuwarden	Deelen	Eelde	Twenthe	Vlissingen	Rotterdam	Gilze-Rijen	Eindhoven	Volkel	Maastricht
Summer Season	$\hat{C}_j(1)$	0.14	0.10	0.03	0.08	0.06	0.06	0.03	0.20	0.09	0.04	0.05	0.05	0.06
Winter Season	$\hat{C}_j(1)$	0.13	0.12	0.05	0.08	0.05	0.07	0.03	0.20	0.09	0.04	0.05	0.05	0.06

Note: Integrated scedasis at $s = 1$ can be interpreted as the relative fraction of extremes from January 1971 until June 2020 over space.

Table 11 presents the integrated scedasis at at the end of the considered time period for all stations. These represent the relative cumulative frequency of extremes over the considered time period. We clearly identify again the coastal weather stations *De Kooy* and *Vlissingen* that represent a fraction of 14% and 20% of all threshold exceedances in the summer season, respectively. Hence, we conclude that we have evidence that the frequency of having peak wind speed extremes is not constant across space in both the summer and winter season.

Furthermore, similar to in the univariate framework, we test whether the space-time scedasis exhibits a temporal trend. However, since we aim to compare the results over space we use our common extreme value threshold $X_{N:N-k}$. As a consequence, the number of observations per station

may differ and are given by $k \cdot \hat{C}_j(1)$ for all $j = 1, \ldots, 13$. Hence, the results of the multivariate and univariate test for heteroscedastic extremes over time could be different. We estimate the space-time integrated scedasis function $C_j(s)$ on $0 \leq s \leq 1$, such that we can calculate the test statistic T_3.

Table 12: Results Kolmogorov-Smirnov type test statistic for peak wind speed in the summer and winter season.

		De Kooy	Schiphol	De Bilt	Leeuwarden	Deelen	Eelde	Twenthe	Vlissingen	Rotterdam	Gilze-Rijen	Eindhoven	Volkel	Maastricht
Summer Season	T_3	0.02	0.01	0.01	0.02	0.02	0.01	0.04	0.01	0.01	0.02	0.02	0.02	0.01
	P-value	0.51	0.85	0.79	0.41	0.52	0.96	0.00	0.84	0.82	0.20	0.24	0.41	0.73
Winter Season	T_3	0.04	0.04	0.04	0.05	0.04	0.04	0.03	0.05	0.02	0.05	0.04	0.02	0.03
	P-value	0.01	0.01	0.01	0.00	0.04	0.01	0.07	0.00	0.65	0.00	0.00	0.33	0.27

Note: Thresholds corresponding to the summer and winter season are 17.5m/s and 23.7m/s, respectively. P-values for T_3 depend on the number of upper pseudo order statistics which vary across the stations as $k \cdot \hat{C}_j(1)$ for all $j = 1, \ldots, 13$.

Table 12 reports the calculated test statistics and corresponding P-values derived from the Monte Carlo simulated probability density function of $\sup_{0 \leqslant s \leqslant 1} |B(s)|$. Recall that the P-values for T_3 depend on the number of upper pseudo order statistics which vary across the stations as $k \cdot \hat{C}_j(1)$ for all $j = 1, \ldots, 13$. We conclude that using a common threshold based on pooling all observations yields slightly different results compared to the univariate framework. This is due the high relative fraction of extremes stemming from the coastal stations as presented in Table 11.

For the summer season, homoscedastic extremes in time are rejected at 5% significance level only for *Twenthe*. For all other stations we cannot reject homoscedastic extremes in time at any meaningful significance level. We conclude that the spatial coherence is high. For the winter season, homoscedastic extremes in time are rejected at 5% significance level for 9 out 13 stations. For *Twenthe, Rotterdam, Volkel* and *Maastricht* we cannot reject homoscedastic extremes in time at any meaningful significance level. Examining Table 11 and 12 simultaneously, reveals that this is not necessarily due to the varying fraction of extremes across space, i.e. the number observations per station.

Since we have identified heteroscedastic extremes for 9 out of 13 weather stations in the winter season, we estimate time varying exceedance probabilities. We incorporate heteroscedasticity of extremes via the estimated space-time scedasis function. We perform the homogenization proce-

dure to obtain pseudo observations $Z_{i,j}$ sing estimates γ, $\sigma(\frac{N}{k})$ and $\hat{c}\left(\frac{i}{n}, j\right)$ for $i = 1, \ldots, n$ and $j = 1, \ldots, 13$. Thereafter we estimate the time varying exceedance probability using homogenized threshold $Z_{N:N-k}$.

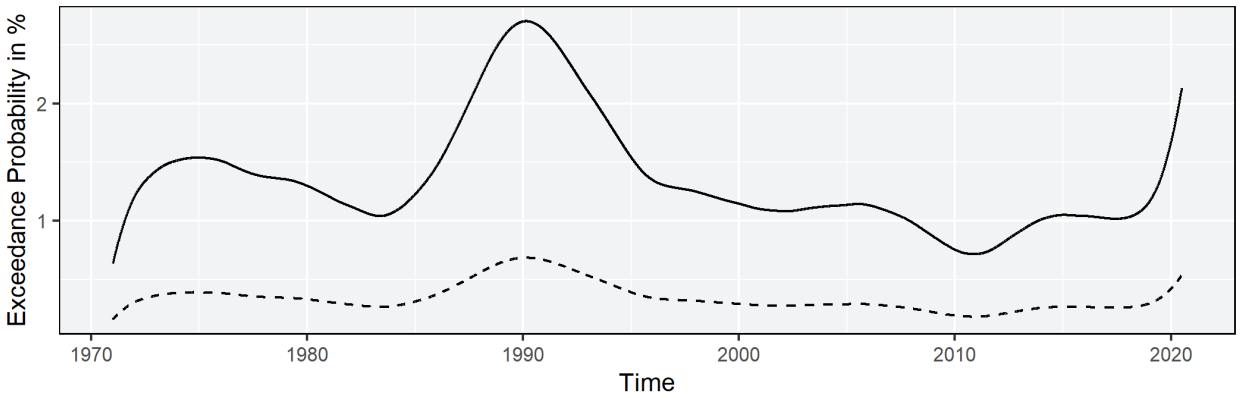

Figure 11: Estimated time varying daily exceedance probabilities in % from January 1971 until June 2020 of *25m/s* and *30m/s* at *Schiphol* for peak wind speed in the winter season.

Figure 11 depicts the estimated time varying exceedance probabilities of *25m/s* and *30m/s* at *Schiphol* for peak wind speed in the winter season. We clearly identify a large peak in extreme event probability around 1990, and a slow decay for the years thereafter. The exceedance probability of *25m/s* varies substantially between 0.75% and 2.75% over the considered time period at *Schiphol*. However, we cannot detect a monotone trend. The estimated time varying exceedance probabilities of the remaining weather stations are deferred to the appendix.

Next, we use the obtained homogenized pseudo observations to characterize variation in space via longitudinal and latitudinal distances. We estimate the well-known tail dependence coefficient, after which we estimate the *empirical* variogram. Thereafter, we estimate the variogram parameter estimates by numerically minimizing the squared difference between the *empirical* and parametric variogram. Table 13 provides the estimated variogram parameters and standard errors for both the summer and winter season, respectively.

Table 13: Variogram parameter estimates for peak wind speed in the summer and winter season, respectively.

	Summer Season				**Winter Season**			
Parameters	$\hat{b}_1$	$\hat{b}_2$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{b}_1$	$\hat{b}_2$	$\hat{\theta}$	$\hat{\alpha}$
Estimate	0.073	0.216	0.743	0.344	0.017	0.040	0.850	0.579
Std. Error	(0.026)	(0.107)	(0.066)	(0.053)	(0.003)	(0.009)	(0.067)	(0.075)

Note: Parameters are estimated using longitudinal and latitudinal distances between the 13 weather stations.

Figure 12 graphs the estimated empirical tail dependence coefficients $\hat{L}(1,1)$ for all bivariate combinations $(j_1, j_2) \in j = 1, \ldots, m$ as a function of the Euclidean distance between stations in km, for summer (red) and winter (blue) season. In addition, we graph the estimated parametric tail dependence as a function of the Euclidean distance in km for summer (red) and winter (blue) season. For distances below 200km we observe higher tail dependence for the weather stations in the winter season compared to the summer season, however, this difference diminishes as the distance between stations increases.

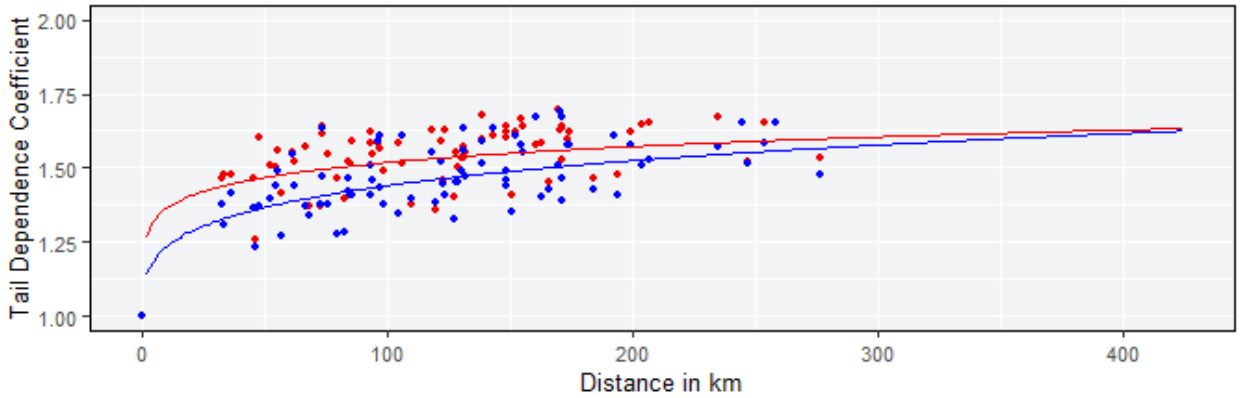

Figure 12: Estimated empirical tail dependence coefficients for all bivariate combinations $(j_1, j_2) \in j = 1, \ldots, 13$ and estimated parametric tail dependence as a function of the Euclidean distance in *km* for summer (red) and winter (blue) season.

We use the estimated variogram and tail dependence structure to estimate time varying joint exceedance probabilities that incorporate heteroscedasticity of extremes via the estimated space-time scedasis function. Figure 13 presents the Estimated time varying joint exceedance probabilities in % of *25m/s* and *30m/s* at *Schiphol* and *De Bilt* (left) and at *Schiphol* and *Eelde* (right). We clearly identify the time varying behaviour stemming from the scedasis functions, more specifically, we recognize the decreasing trend between 1990 and 2010 from Figure 11. Indeed, a decrease in marginal probability leads to a decrease in joint probability (ceteris paribus).

However, we cannot identify the decreasing tail dependence over space when comparing the joint exceedance probability at *Schiphol* and *De Bilt* (36*km* distance) and at *Schiphol* and *Eelde* (150.7*km* distance). Comparing Figure 7 and 12 we recognize that peak wind speed has a higher and less spatial dependent tail dependence coefficient. In addition we identified that marginal exceedance probabilities in *De Bilt* are lower compared to *Eelde*. Hence, therefore we find comparable joint exceeding probabilities despite the spatial difference.

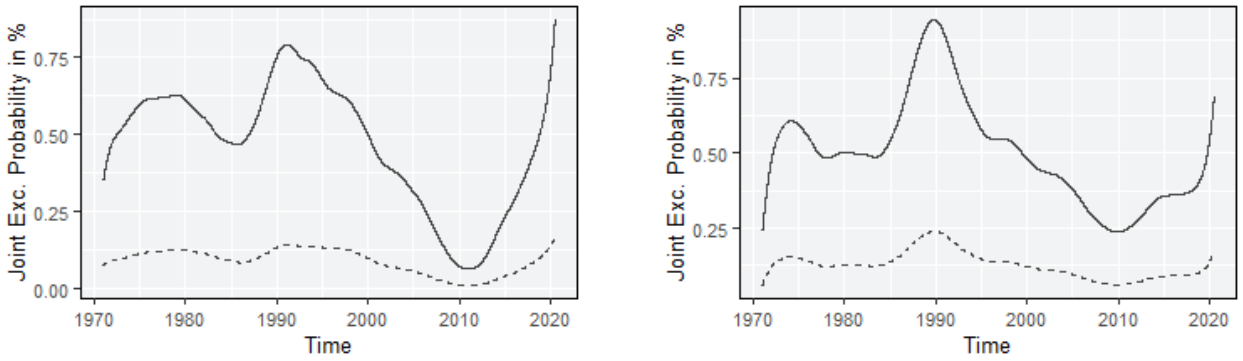

Figure 13: Estimated time varying daily joint exceedance probabilities in % from January 1971 until June 2020 of *25m/s* and *30m/s* at *Schiphol* and *De Bilt* (left) and at *Schiphol* and *Eelde* (right) for peak wind speed in the winter season.

6 Forward looking scedasis

Since we rejected constant frequency of extremes for 24 hour precipitation accumulation in the summer season and peak wind speed in the winter season, it is of interest to estimate the potential trend going forward. However, to forecast time varying (joint) exceedance probabilities we require a parametric scedasis function that can be extrapolated into the future, rather than the non-parametric scedasis function discussed so far.

Mefleh et al. (2020) builds upon the heteroscedastic extremes framework by Einmahl et al. (2016). Instead of non-parametric estimation of the scedasis function, it considers a monotone parametric scedasis function and proves the consistency and asymptotic normality of the corresponding parameter estimators. Parametric models are more rigid compared to non-parametric ones. However, the parametric model forces the scedasis function to be monotone which is interesting from the climate change perspective. In addition, this monotone parametric scedasis function can be extrapolated into the future.

Using a point process approach Mefleh et al. (2020) shows that the standardized threshold exceedance times are asymptotically independent and identically distributed with density $c = c_\theta$, for some parametric density function c_θ. The limiting distribution of the exceedance values depends on the extreme value index, which we assume to be constant over time. Therefore the parameter θ can be estimated using solely the standardized threshold exceedance times, denoted by $T_1, \ldots, T_k$.

We estimate log-linear and linear scedasis functions. The log-linear scedasis function defined on is of the form,

$$c_{\theta_1}(s) = e^{\theta_1 s - h(\theta_1)} \quad \text{where} \quad \theta_1 \in \mathbb{R}, \tag{53}$$

The function $h(\theta_1)$ ensures that the scedasis function integrates to 1 for $0 \leq s \leq 1$, as required by Equation (11). The linear scedasis function is of the form,

$$c_{\theta_2}(s) = \theta_2(2s-1) + 1 \text{ where } -1 < \theta_2 < 1. \tag{54}$$

which again integrates to 1 for $0 \leq s \leq 1$. We estimate θ_1 and θ_2 using numerical Maximum Likelihood estimation on the standardizes exceedance times $T_1, \ldots, T_k$. We refer to Mefleh et al. (2020) for a detailed explanation of the parameter estimation procedure and the asymptotic theory.

6.1 24 hour precipitation accumulation

Given the presence of heteroscedastic extremes for 24 hour precipitation accumulation in the summer season we estimate the log-linear and linear scedasis functions, to model a monotone trend and perform forecasting. We estimate the parameters θ_1 and θ_2 using numerical Maximum Likelihood estimation on the standardizes exceedance times $T_1, \ldots, T_k$. Using the estimated parameter $\hat{\theta_1}$ and $\hat{\theta_2}$ for log-linear and linear scedasis, respectively, we extrapolate the scedasis for the next 25 years. Then we use the extrapolated (forward looking) scedasis function to forecast the time varying exceedance probability over the next 25 years.

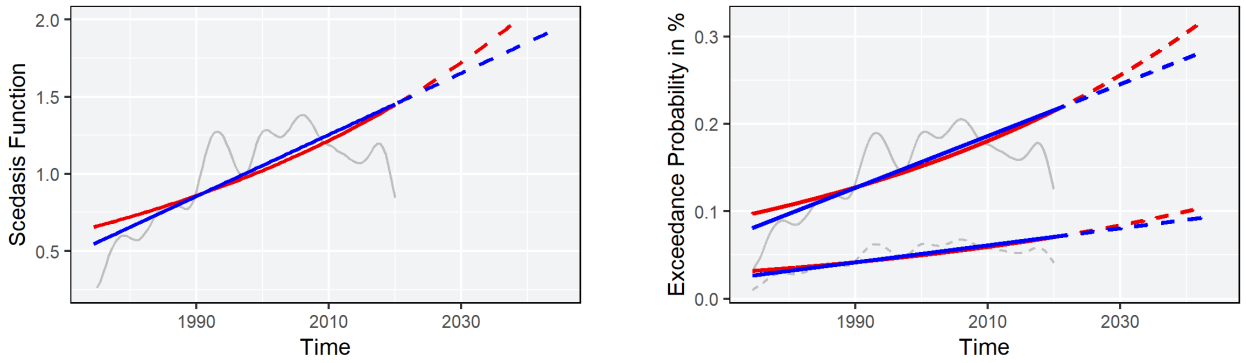

Figure 14: Forward looking scedasis function and forecasted exceedance probabilities of *40mm* and *50mm* from June 1974 until June 2045 at weather station *Schiphol*, parameterized using linear trend (blue) and log-linear trend (red), for summer season.

Figure 14 depicts the estimated forward looking scedasis function and forecasted exceedance probability at weather station *Schiphol*, parameterized using linear trend and log-linear trend, for the summer season. We identify a clear upward sloping trend. Table 14 presents threshold exceedance probability forecasts for 2030 and 2040 of *40mm* and *50mm* precipitation accumulation in 24 hours, for all stations, for both log-linear and linear scedasis in the summer season.

Table 14: Forecasted threshold exceedance probabilities for the years 2030 and 2040 of *40mm* and *50mm* precipitation accumulation in 24 hours, in the summer season.

			De Kooy	Schiphol	De Bilt	Leeuwarden	Eelde	Twenthe	Vlissingen	Rotterdam	Volkel	Maastricht
Summer 2030	Log-linear	*40mm*	0.17	0.24	0.21	0.20	0.16	0.14	0.16	0.20	0.16	0.17
		50mm	0.06	0.08	0.07	0.07	0.05	0.05	0.05	0.07	0.05	0.06
	Linear	*40mm*	0.17	0.24	0.20	0.20	0.16	0.14	0.16	0.19	0.16	0.17
		50mm	0.05	0.08	0.07	0.06	0.05	0.05	0.05	0.06	0.05	0.06
Summer 2040	Log-linear	*40mm*	0.18	0.28	0.23	0.22	0.17	0.15	0.17	0.22	0.17	0.19
		50mm	0.06	0.09	0.07	0.07	0.06	0.05	0.06	0.07	0.06	0.06
	Linear	*40mm*	0.18	0.26	0.22	0.21	0.17	0.15	0.17	0.21	0.17	0.18
		50mm	0.06	0.09	0.07	0.07	0.06	0.05	0.05	0.07	0.05	0.06

Note: Exceedance probabilities are reported in % and should be interpreted as the probability of exceeding the respective threshold on a single day.

We observe that, although *Eelde* has the lowest exceedance probability in 2020, given the estimated linear scedasis in 2040, the probability of extremes in *Eelde* is higher compared to *Twenthe, Vlissingen* and *Volkel*. In general the daily probability of extremes slowly increases over time with around 0.01% per decade.

We compare the estimated exceedance probabilities for 2020 in Figure 6 to the forecasted exceedance probabilities in Table 14 to further specify the magnitude of this trend. It follows that, in the summer season of 2020, extreme precipitation accumulation (*40mm*) at *Schiphol* is expected approximately once every 665 days. Whereas, assuming a log-linear trend, extreme precipitation accumulation (*40mm*) in the summer season of 2040 at *Schiphol* is expected approximately once every 355 days. Hence, we can identify a substantial increase in hazard probability for extreme precipitation accumulation going forward.

6.2 Windstorm peak wind speed

Given the presence of heteroscedastic extremes for peak wind speed in the winter season we estimate the log-linear and linear scedasis functions. Similar to the analysis above, we use the extrapolated (forward looking) scedasis function to forecast the time varying exceedance probability over the next 25 years.

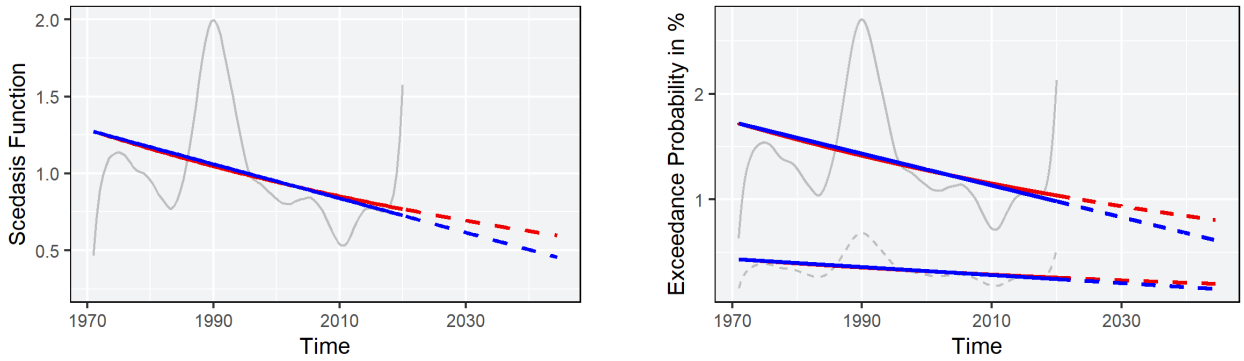

Figure 15: Forward looking scedasis function and forecasted exceedance probabilities of *25m/s* and *30m/s* from January 1971 until June 2045 at weather station *Schiphol*, parameterized using linear trend (blue) and log-linear trend (red), for winter season.

Figure 15 depicts the estimated forward looking scedasis function and forecasted exceedance probability at weather station *Schiphol*, parameterized using linear trend and log-linear trend, for the winter season. We identify a clear downward sloping trend. Table 15 presents threshold exceedance probability forecasts for 2030 and 2040 of *25m/s* and *30m/s* peak wind speed, for all stations, for both log-linear and linear scedasis in the winter season.

Table 15: Forecasted threshold exceedance probabilities for the years 2030 and 2040 of *25m/s* and *30m/s* peak wind speed, in the winter season.

			De Kooy	Schiphol	De Bilt	Leeuwarden	Deelen	Eelde	Twenthe	Vlissingen	Rotterdam	Gilze-Rijen	Eindhoven	Volkel	Maastricht
Winter 2030	Log-linear	*25m/s*	1.05	0.94	0.26	0.51	0.35	0.46	0.27	1.78	1.04	0.22	0.27	0.74	0.50
		30m/s	0.27	0.24	0.07	0.13	0.09	0.12	0.07	0.45	0.26	0.06	0.07	0.19	0.13
	Linear	*25m/s*	0.97	0.83	0.18	0.39	0.32	0.35	0.21	1.73	1.03	0.18	0.21	0.71	0.44
		30m/s	0.24	0.21	0.04	0.10	0.08	0.09	0.05	0.44	0.26	0.05	0.05	0.18	0.11
Winter 2040	Log-Linear	*25m/s*	0.94	0.85	0.21	0.43	0.31	0.39	0.24	1.66	1.03	0.18	0.22	0.81	0.46
		30m/s	0.24	0.21	0.05	0.11	0.08	0.10	0.06	0.42	0.26	0.05	0.06	0.20	0.12
	Linear	*25m/s*	0.79	0.68	0.07	0.21	0.26	0.21	0.16	1.57	1.03	0.10	0.11	0.76	0.38
		30m/s	0.20	0.17	0.02	0.05	0.06	0.05	0.04	0.40	0.26	0.03	0.03	0.19	0.10

Note: Exceedance probabilities are reported in % and should be interpreted as the probability of exceeding the respective threshold on a single day.

In general the daily probability of extremes decreases over time with 0.05%-0.1% per decade, strongly depend on the assumption of a log-linear or linear scedasis function. This corresponds with Figure 15 where the log-linear scedasis tend to flatten after 2030, whereas the linear scedasis (by definition) keeps decreasing.

Again we can compare the estimated exceedance probabilities for 2020 in Figure 11 to the forecasted exceedance probabilities in Table 15. It follows that, in the winter season of 2020, extreme peak wind speed (*25m/s*) at *Schiphol* is expected approximately once every 66 days. Whereas, assuming a log-linear trend, extreme peak wind speed (*25m/s*) at *Schiphol* in the winter season of 2040 is only expected approximately once every 120 days. Hence, we can identify a substantial decrease in hazard probability for extreme peak wind going forward.

7 Conclusion

This thesis applies recently developed techniques for heteroscedastic extremes to investigate the impact of global warming on local weather extremes in the Netherlands. The choice of variables is motivated by an insurance industry perspective. For extreme precipitation we use 24 hour precipitation accumulation as leading variable, for which we collect daily data from June 1974 until June 2020. For extreme windstorm we use peak wind speed as leading variable, for which we collect daily data from January 1971 until June 2020. We separate the analysis for summer and winter periods to remove potential seasonality in extremes. Furthermore, we employ a declustering procedure to remove potential temporal dependence.

Next, we assume extremes are i.i.d. and we model exceedances above threshold a GP distribution for all stations individually. Thereafter we relax the assumption of i.i.d. observations and model the continuously evolving distribution functions over time using a continuous and positive scedasis function (Einmahl et al., 2016). Further, we relax the assumption of spatial independence, and invoke a space-time framework (Einmahl et al., 2020). For both precipitation and windstorm extremes we pool all observations to estimate common extreme value parameters over space and time using a common threshold.

Furthermore, we illustrate how to incorporate time varying frequency of extremes (heteroscedasticity) in the estimation of well-known risk measures. As a leading example we estimate time varying exceedance probabilities. We use the estimated scedasis function and extreme value parameters to

perform a homogenization procedure (Ferreira et al., 2017). Combining the estimated scedasis, extreme value parameters and homogenized pseudo-observations we estimate (joint) exceedance probabilities of high level extremes over time, using an extremal dependence structure over space. Finally, we parameterize the non-parametric scedasis function using a (log) linear trend (Mefleh et al., 2020) to forecast time varying (joint) exceedance probabilities.

On the one hand we conclude that the probability of 24 hour precipitation accumulation extremes in the summer season has increased over the considered time period and might continue to increase in the years to come. On the other hand, we conclude that the probability of extreme peak wind speeds in the winter season has decreased after a steep peak around 1990 and might continue to decrease in the years to come. In addition, we conclude that the frequency of precipitation and windstorm extremes is not constant over space for summer and winter season. In general the test outcomes for constant frequency of extremes in the multivariate framework (using a common threshold) confirm the conclusion obtained in the univariate framework.

Since we estimated trends over time regardless of cause (Klein Tank et al., 2009), we cannot attribute these changes directly to anthropogenic climate change. Hence the time varying behaviour can be due to both anthropogenic climate change and natural internal processes within the climate system. However, still these trends provide valuable information on the first-order changes over time. From the insurance industry perspective one could argue that insurance companies should take time varying claim behaviour into account when pricing insurance in a high environmental peril risk market. Incorporating time varying claim probabilities in their underwriting risk modelling presumably impacts both their premium pricing and obligated loss reserving.

8 Discussion

We aim to investigate the impact of global warming on local weather extremes in the Netherlands. However, our methodology contains a number of limitations which could be considered as suggestions for further research.

The main underlying assumption of the applied methodology is a constant extreme value index γ over time and space, which should preferably be validated. The extreme value index controls the magnitude of extremes. One might argue that, apart from the frequency of weather extremes, also the magnitude of weather extremes could be impacted by climate change. Einmahl et al. (2016)

propose to estimate γ for several, interval based, subsamples and compare these estimates over time. However the estimator and test statistic considered are only justifiable if extreme value index $\gamma > 0$. For $\gamma \in \mathbb{R}$ no formal test exists (yet) and therefore we simply assume constant extreme value index over time and space and subsequently refer to Buishand et al. (2008) and Klein Tank et al. (2009) to strengthen this assumption.

In addition, we explicitly assume that the multivariate tail dependence structure is time invariant. However, one might argue that, apart from frequency and magnitude of weather extremes, also the spatial tail dependence structure of weather extremes could be impacted by climate change. For instance, Castro-Camilo et al. (2018) and Mhalla et al. (2019) propose a non-parametric kernel estimator and a vector generalized additive model, respectively. However it is not straightforward (yet) to implement these methods within the heteroscedastic extremes framework.

Following Klein Tank et al. (2009) we investigated trends in extreme weather events regardless of cause. Detecting a trend or change in extremal behaviour is different from attributing that change to a human induced climate changing factor. More specifically, statistical analysis can detect trends, however this does not determine whether this trend is anthropogenic or caused by natural inter-annual climate variability. For the attribution problem one requires climate models to quantify the climate system's response to external driving forces. Within the context of extreme weather events, climate models should adequately capture extreme event statistics. However, according to Min et al. (2011) existing climate system models tend to underestimate extreme behaviour for windstorms and precipitation in particular.

From the insurance perspective it is worth noting that, even if extreme weather events are not getting more frequent or more intense over the considered time period, extreme events may cause more damage than they did years ago. This can be explained by increased exposure in many regions, for instance, population growth and expanding urban zones. In other words, if we model trends in extremes for incurred insurance claims, we might detect a spatially coherent increasing monotone trend. The question then remains if this should be attributed to climate change, and if this should be taken into account when investigating impact of global warming. This is outside the scope of this thesis.

Finally, we apply the methodology of Einmahl et al. (2016), Ferreira et al. (2017) and Einmahl et al. (2020) to model trends in the frequency of extreme precipitation and peak wind speed. This this contributes to the existing literature as a first empirical application of the recently published

methodology for non-identically distributed extremes. Hence, the multi-step analysis in this thesis could be applied to many other (slightly) heavy-tailed (weather) variables for which trends in extremes are of interest.

References

Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A., ... others (2006). Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research: Atmospheres*, *111*(D5).

Botzen, W. J., & Van Den Bergh, J. C. (2008). Insurance against climate change and flooding in the netherlands: present, future, and comparison with other countries. *Risk Analysis: An International Journal*, *28*(2), 413–426.

Bouwer, L. M. (2013). Projections of future extreme weather losses under changes in climate and exposure. *Risk Analysis*, *33*(5), 915–930.

Buishand, T., De Haan, L., Zhou, C., et al. (2008). On spatial extremes: with application to a rainfall problem. *The Annals of Applied Statistics*, *2*(2), 624–642.

Castro-Camilo, D., de Carvalho, M., Wadsworth, J., et al. (2018). Time-varying extreme value dependence with application to leading european stock markets. *The Annals of Applied Statistics*, *12*(1), 283–309.

Coles, S., Bawa, J., Trenner, L., & Dorazio, P. (2001). *An introduction to statistical modeling of extreme values* (Vol. 208). Springer.

Coumou, D., & Rahmstorf, S. (2012). A decade of weather extremes. *Nature climate change*, *2*(7), 491–496.

Davison, A. C., & Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society: Series B (Methodological)*, *52*(3), 393–425.

De Haan, L., Pereira, T. T., et al. (2006). Spatial extremes: Models for the stationary case. *The annals of statistics*, *34*(1), 146–168.

De Haan, L., Tank, A. K., & Neves, C. (2015). On tail trend detection: modeling relative risk. *Extremes*, *18*(2), 141–178.

Dorland, C., Tol, R. S., & Palutikof, J. P. (1999). Vulnerability of the netherlands and northwest europe to storm damage under climate change. *Climatic change*, *43*(3), 513–535.

Easterling, D. R., Meehl, G. A., Parmesan, C., Changnon, S. A., Karl, T. R., & Mearns, L. O. (2000). Climate extremes: observations, modeling, and impacts. *science*, *289*(5487), 2068–2074.

Einmahl, J. H., de Haan, L., & Zhou, C. (2016). Statistics of heteroscedastic extremes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *78*(1), 31–51.

Einmahl, J. H., Ferreira, A., de Haan, L., Neves, C., & Zhou, C. (2020). Spatial dependence and space-time trends in extreme events.

Ferreira, A., Friederichs, P., de Haan, L., Neves, C., & Schlather, M. (2017). Estimating space-time trend and dependence of heavy rainfall.

Fisher, R. A., & Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical proceedings of the cambridge philosophical society* (Vol. 24, pp. 180–190).

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of mathematics*, 423–453.

Groisman, P. Y., Knight, R. W., Easterling, D. R., Karl, T. R., Hegerl, G. C., & Razuvaev, V. N. (2005). Trends in intense precipitation in the climate record. *Journal of climate*, *18*(9), 1326–1350.

Hall, P., & Tajvidi, N. (2000). Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statistical Science*, 153–167.

Jones, M. C. (1993). Simple boundary correction for kernel density estimation. *Statistics and computing*, *3*(3), 135–146.

Kabluchko, Z., Schlather, M., De Haan, L., et al. (2009). Stationary max-stable fields associated to negative definite functions. *The Annals of Probability*, *37*(5), 2042–2065.

Klein Tank, A., & Können, G. (2003). Trends in indices of daily temperature and precipitation extremes in europe, 1946–99. *Journal of climate*, *16*(22), 3665–3680.

Klein Tank, A., Zwiers, F., & Zhang, X. (2009). Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation. *World Meteorological Organization*.

Meehl, G. A., Zwiers, F., Evans, J., Knutson, T., Mearns, L., & Whetton, P. (2000). Trends in extreme weather and climate events: issues related to modeling extremes in projections of future climate change. *Bulletin of the American Meteorological Society*, *81*(3), 427–436.

Mefleh, A., Biard, R., Dombry, C., & Khraibani, Z. (2020). Trend detection for heteroscedastic extremes. *Extremes*, *23*(1), 85–115.

Mhalla, L., de Carvalho, M., & Chavez-Demoulin, V. (2019). Regression-type models for extremal dependence. *Scandinavian Journal of Statistics*, *46*(4), 1141–1167.

Mills, E. (2009). A global review of insurance industry responses to climate change. *The Geneva Papers on Risk and Insurance-Issues and Practice*, *34*(3), 323–359.

Min, S.-K., Zhang, X., Zwiers, F. W., & Hegerl, G. C. (2011). Human contribution to more-intense precipitation extremes. *Nature*, *470*(7334), 378–381.

Pickands, J., et al. (1975). Statistical inference using extreme order statistics. *the Annals of Statistics*, *3*(1), 119–131.

Rootzén, H., & Tajvidi, N. (1997). Extreme value statistics and wind storm losses: a case study. *Scandinavian Actuarial Journal*, *1997*(1), 70–94.

Smith, R. L. (1989). Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. *Statistical Science*, 367–377.

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., ... others (2013). Climate change 2013: The physical science basis. *Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*, *1535*.

Thistlethwaite, J., & Wood, M. O. (2018). Insurance and climate change risk management: rescaling to look beyond the horizon. *British Journal of Management*, *29*(2), 279–298.

Zolina, O., Simmer, C., Belyaev, K., Kapala, A., & Gulev, S. (2009). Improving estimates of heavy and extreme precipitation using daily records from european rain gauges. *Journal of Hydrometeorology*, *10*(3), 701–716.

Appendices

Standard Brownian Bridge

Provided that the required conditions are satisfied, the univariate ($\sqrt{k}T_1$) and multivariate ($\sqrt{k}T_3$) test statistics for heteroscedastic extremes over time are asymptotically distributed as $\sup_{0\leqslant s\leqslant 1}|B(s)|$, where B denotes a standard Brownian bridge.

Let $W(s)$ be a Wiener process, that is, a continuous stochastic process $W(s), s \geq 0$ such that $W(0) = 0$, for $s_1 < s_2, W(s_2) - W(s_1) \sim N(0, s_2 - s_1)$ and for any $s_1 \text{and} s_2$ increments $W(s_1) - W(s_1)$ are independent. Now we construct a Brownian bridge B for $0 \leq s \leq 1$ as $B(s) = W(s) - s \cdot W(1)$. Hence the Brownian bridge is a continuous stochastic process defined on the interval $[0, 1]$, for which $B(0) = B(1) = 0$.

To derive the approximate distribution of $\sup_{0\leqslant s\leqslant 1}|B(s)|$, we perform a Monte Carlo simulation study as follows. We simulate a discretization of the Wiener process with step size 0.001 for $0 \leq s \leq 1$, we construct the corresponding Brownian Bridge as $B(t) = W(t) - t \cdot W(1)$, we calculate and store $\sup_{0\leqslant s\leqslant 1}|B(s)|$ and repeat this simulation scheme 100000 times. Figure 16 depicts the Monte Carlo estimate of the distribution of $\sup_{0\leqslant s\leqslant 1}|B(s)|$. This Monte Carlo simulation allows us to find the P-value corresponding to a given test statistic. The 95% and 99% quantiles are respectively given by, 1.34 and 1.61.

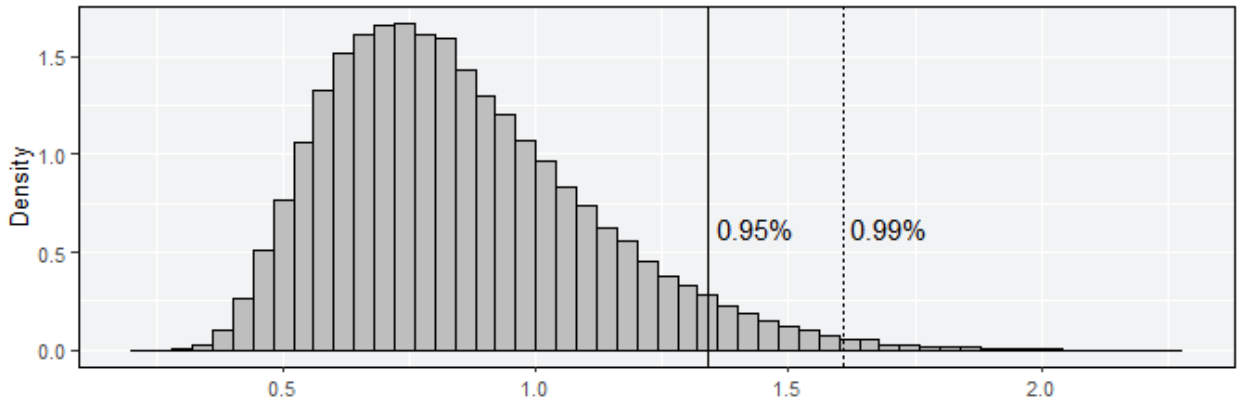

Figure 16: Monte Carlo simulated probability density function for $\sup_{0\leqslant s\leqslant 1}|B(s)|$

Distances

Table 16: Longitude, latitude and measuring altitude for the weather stations

Weather Station	Longitude (East)	Latitude (North)	Altitude
De Kooy	4.781	52.928	1.2
Schiphol	4.790	52.318	-3.3
De Bilt	5.180	52.100	1.9
Leeuwarden	5.752	53.224	1.2
Deelen	5.873	52.056	48.2
Eelde	6.585	53.125	5.2
Twenthe	6.891	52.274	34.8
Vlissingen	3.596	51.442	8.0
Rotterdam	4.447	51.962	-4.3
Gilze-Rijen	4.936	51.566	14.9
Eindhoven	5.377	51.451	22.6
Volkel	5.707	51.659	22.0
Maastricht	5.762	50.906	114.3

Table 17: Great-Circle distances between weather stations in km and measuring altitudes per station in m

Weather Station	De Kooy	Schiphol	De Bilt	Leeuwarden	Deelen	Eelde	Twenthe	Vlissingen	Rotterdam	Gilze-Rijen	Eindhoven	Volkel	Maastricht
De Kooy		67.9	96.1	72.8	122.1	122.8	160.2	184.1	109.9	152.0	169.4	154.7	234.9
Schiphol	67.9		36.0	119.9	79.5	150.7	143.1	127.4	46.0	84.3	104.6	96.6	170.9
De Bilt	96.1	36.0		130.9	47.7	148.4	118.4	131.4	52.5	61.8	73.5	61.0	138.9
Leeuwarden	72.8	119.9	130.9		130.3	56.7	130.7	246.7	165.9	192.7	199.0	174.2	258.0
Deelen	122.1	79.5	47.7	130.3		128.4	73.6	171.2	98.3	84.5	75.5	45.6	128.2
Eelde	122.8	150.7	148.4	56.7	128.4		97.0	276.6	194.2	206.6	203.7	173.7	253.4
Twenthe	160.2	143.1	118.4	130.7	73.6	97.0		244.7	170.6	155.6	138.7	106.2	171.1
Vlissingen	184.1	127.4	131.4	246.7	171.2	276.6	244.7		82.4	93.9	123.6	148.1	162.5
Rotterdam	109.9	46.0	52.5	165.9	98.3	194.2	170.6	82.4		55.5	85.7	93.0	148.8
Gilze-Rijen	152.0	84.3	61.8	192.7	84.5	206.6	155.6	93.9	55.5		33.1	54.3	93.3
Eindhoven	169.4	104.6	73.5	199.0	75.5	203.7	138.7	123.6	85.7	33.1		32.5	66.4
Volkel	154.7	96.6	61.0	174.2	45.6	173.7	106.2	148.1	93.0	54.3	32.5		83.9
Maastricht	234.9	170.9	138.9	258.0	128.2	253.4	171.1	162.5	148.8	93.3	66.4	83.9	

Further Results Precipitation

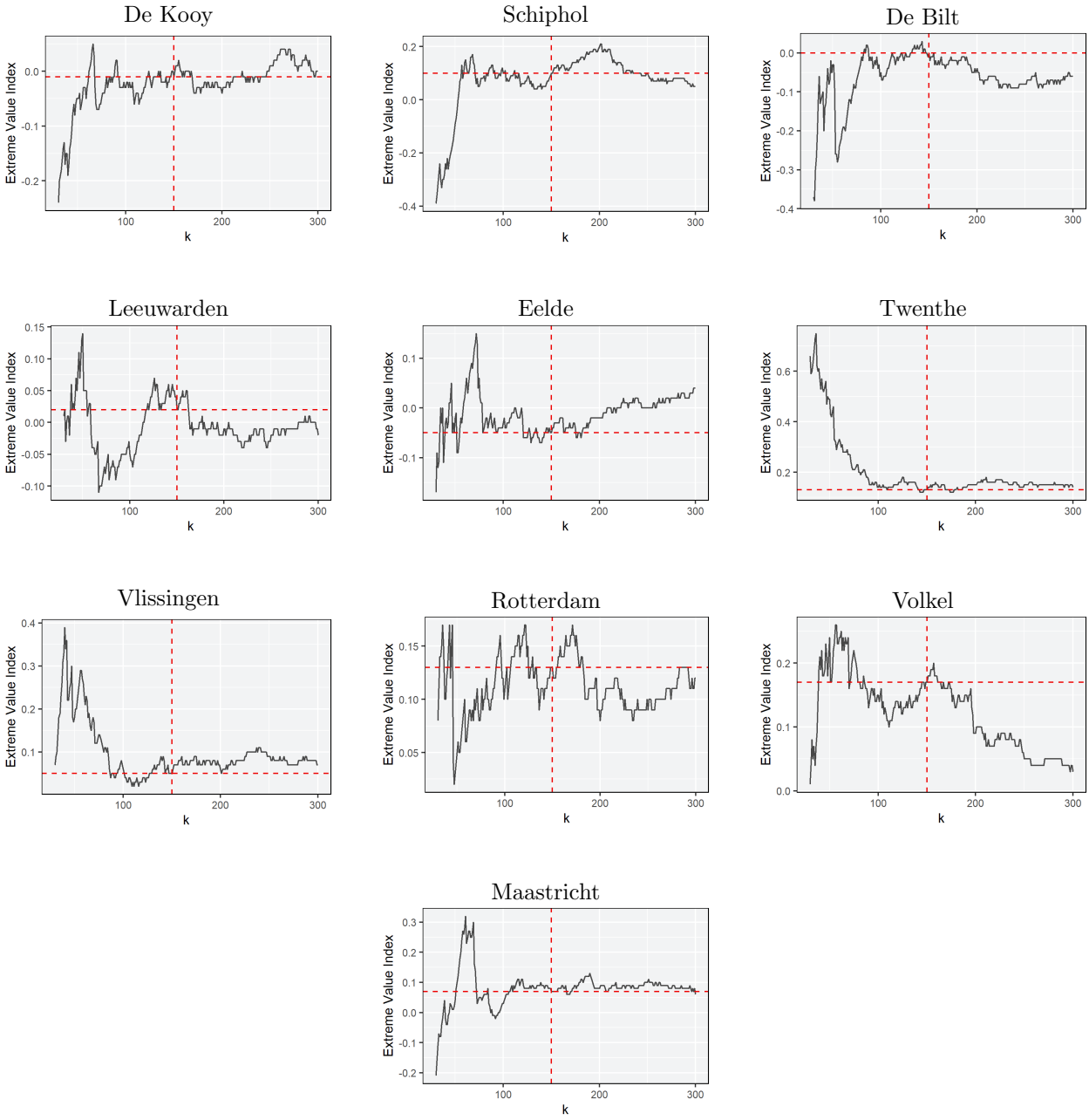

Figure 17: Parameter stability plots for the extreme value index as a function of upper pseudo order statistics k.

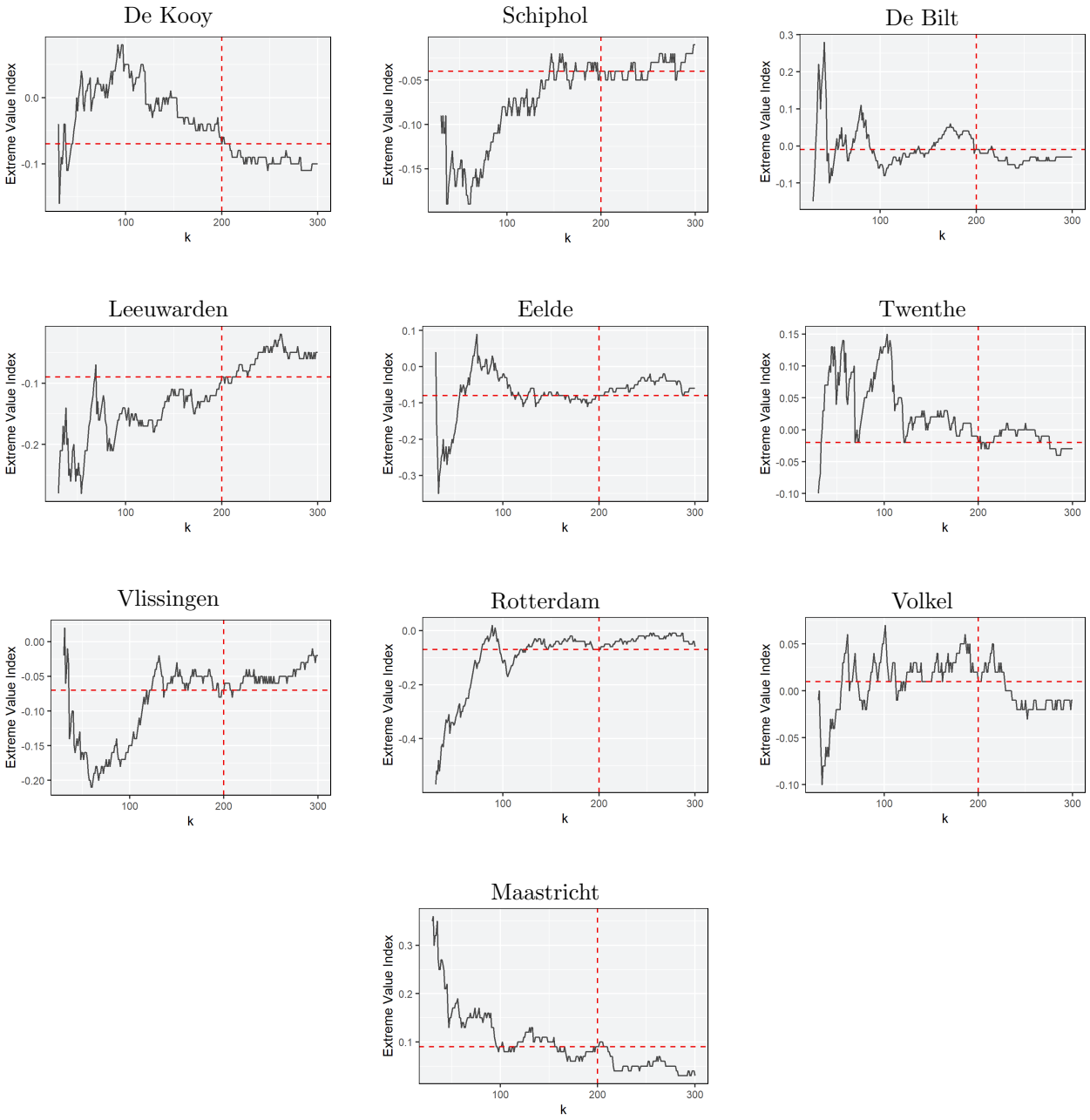

Figure 18: Parameter stability plots for the extreme value index as a function of upper pseudo order statistics k.

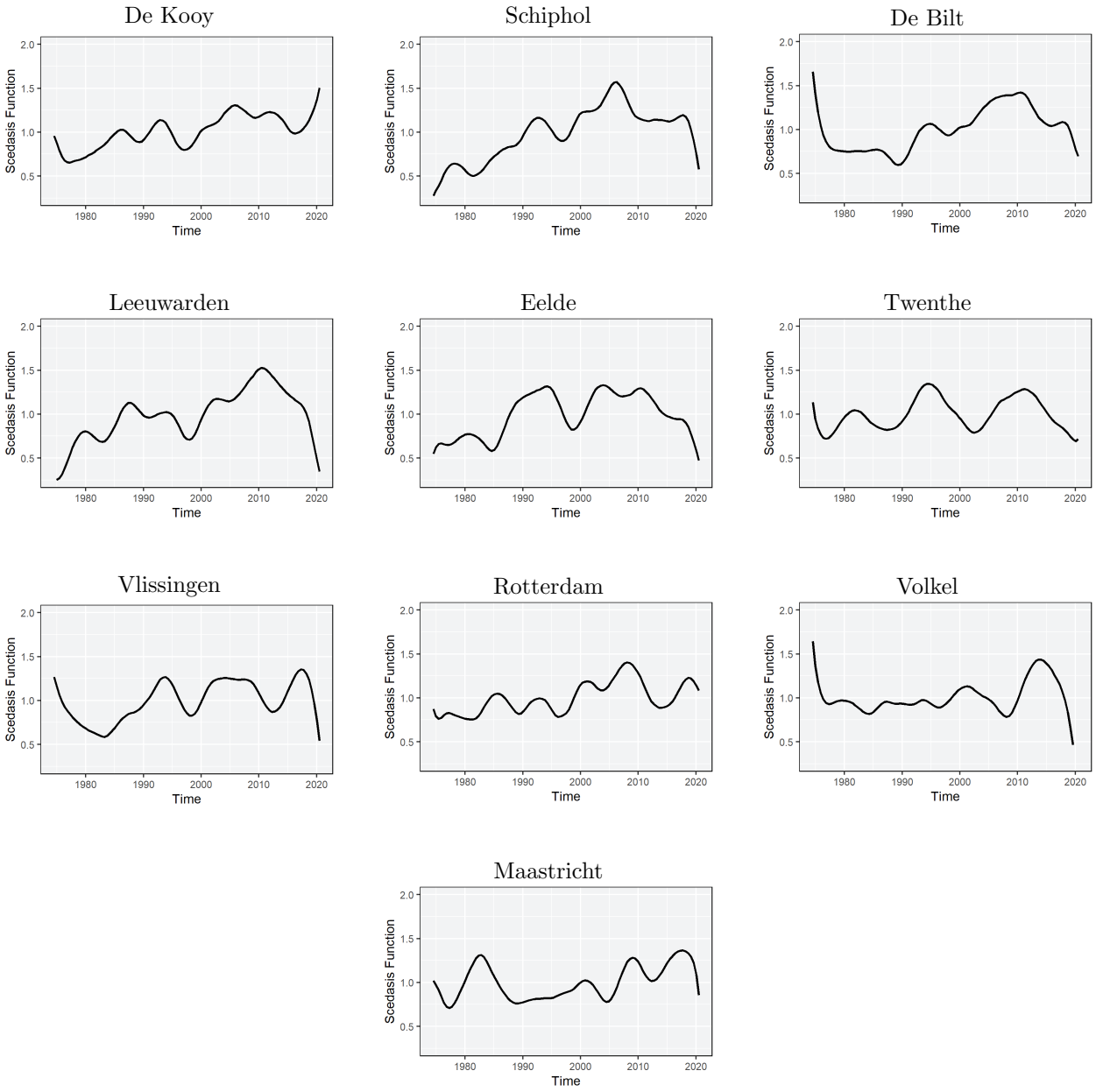

Figure 19: Plotted estimated scedasis functions for 24 hour Precipitation accumulation in summer season.

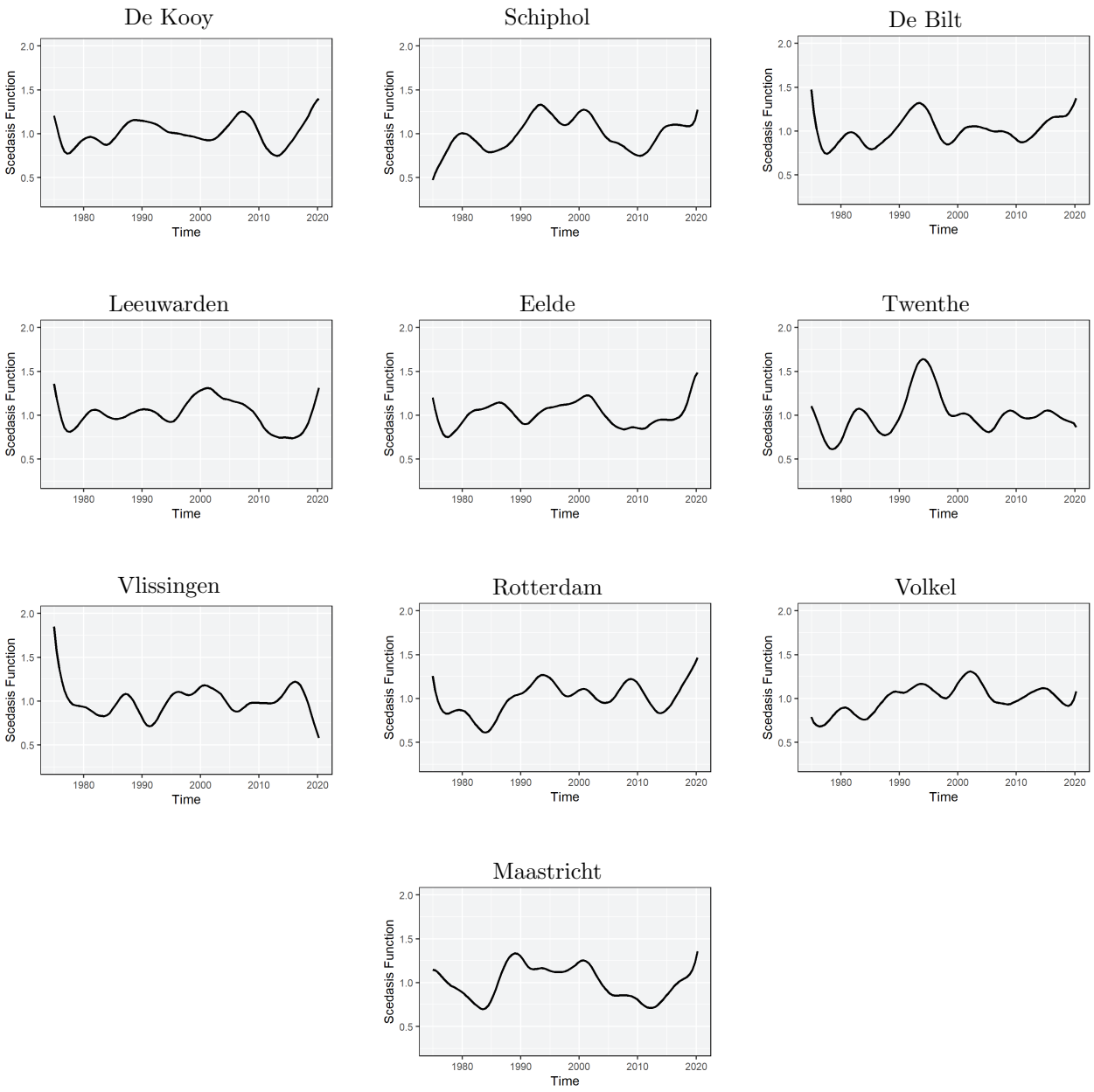

Figure 20: Plotted estimated scedasis functions for 24 hour Precipitation accumulation in winter season.

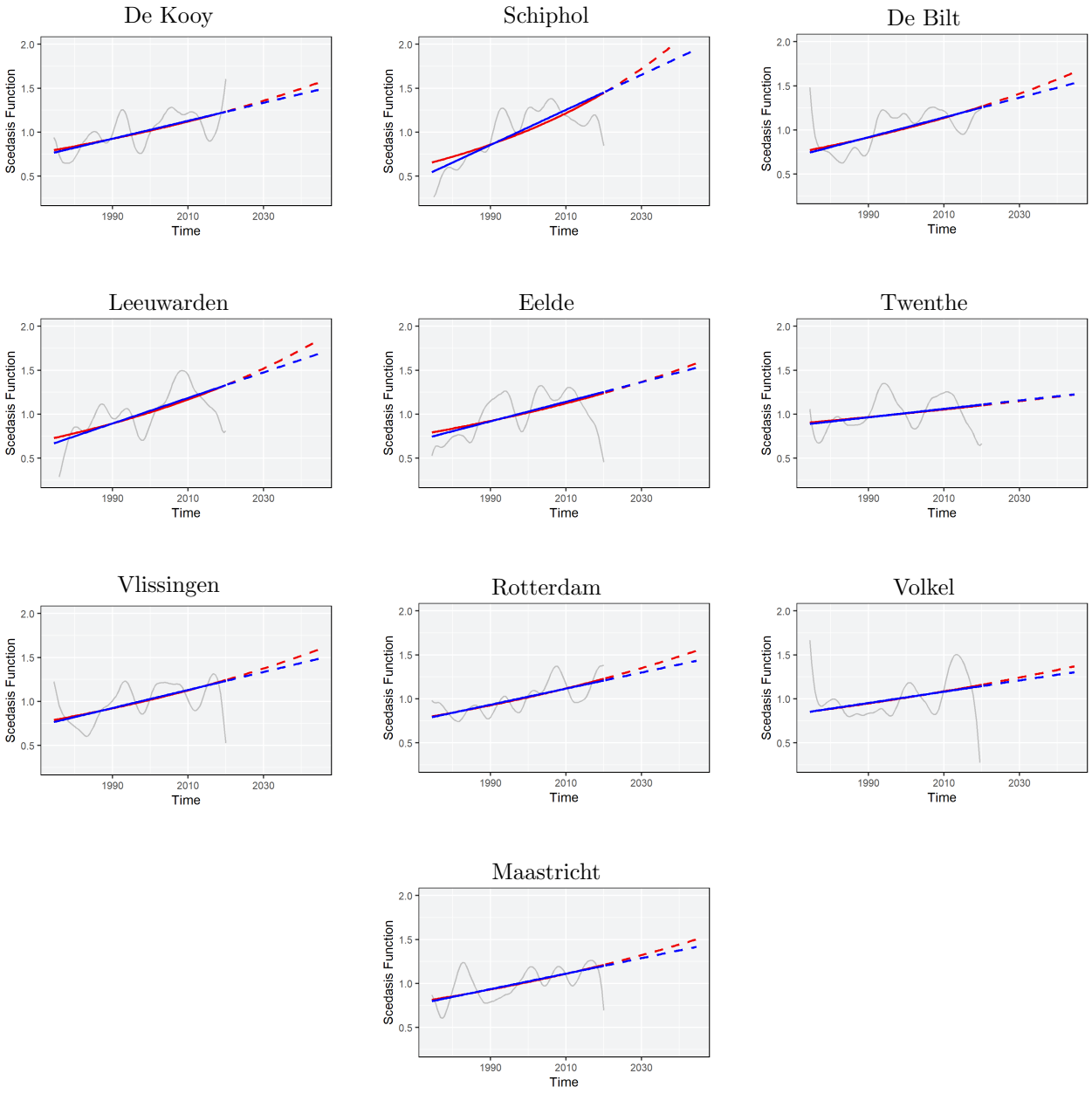

Figure 21: Forward looking scedasis function, parameterized using linear trend (blue) and log-linear trend (red) for the summer season.

Further Results Windstorm

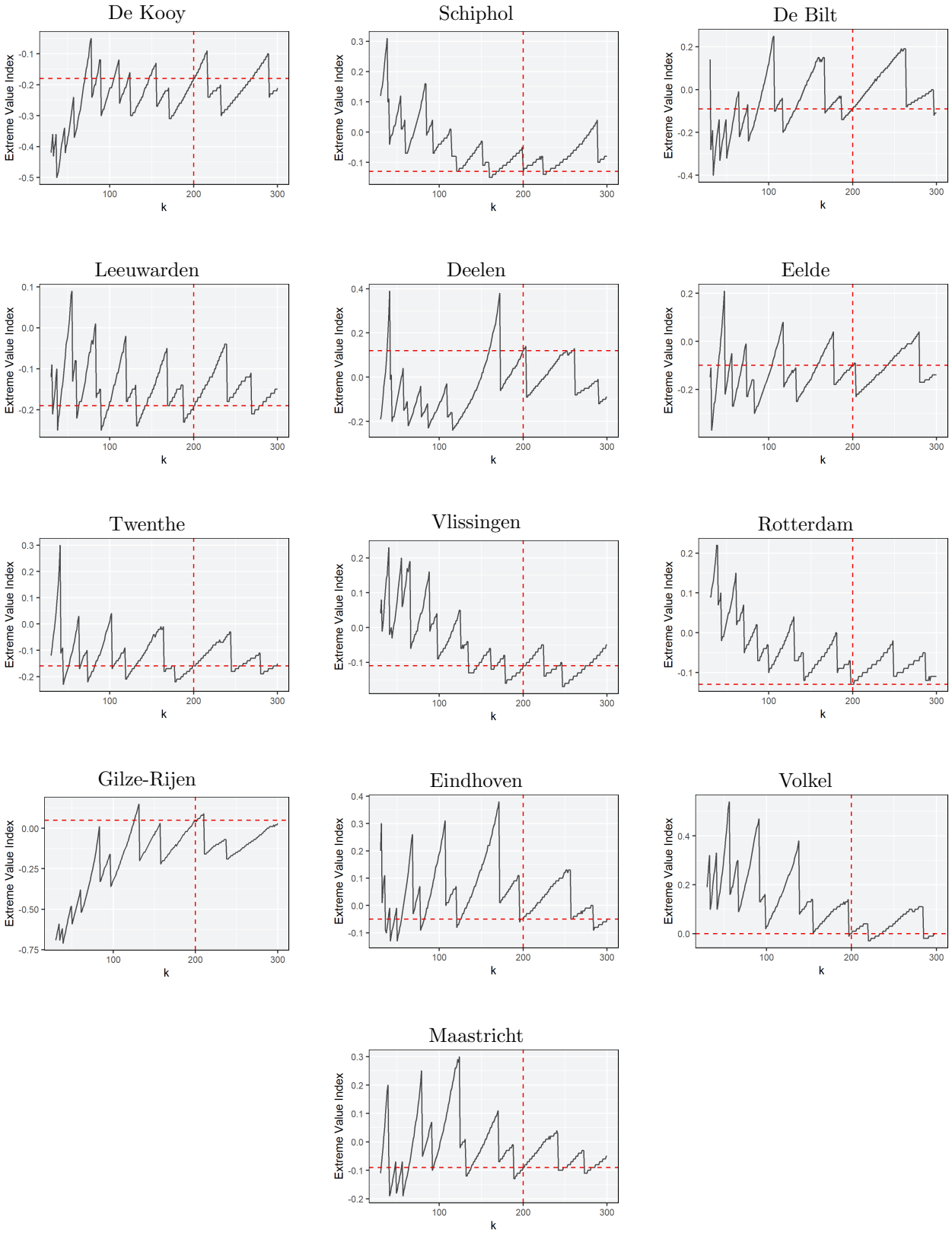

Figure 22: Parameter stability plots for the extreme value index as a function of the number of upper pseudo order statistics k.

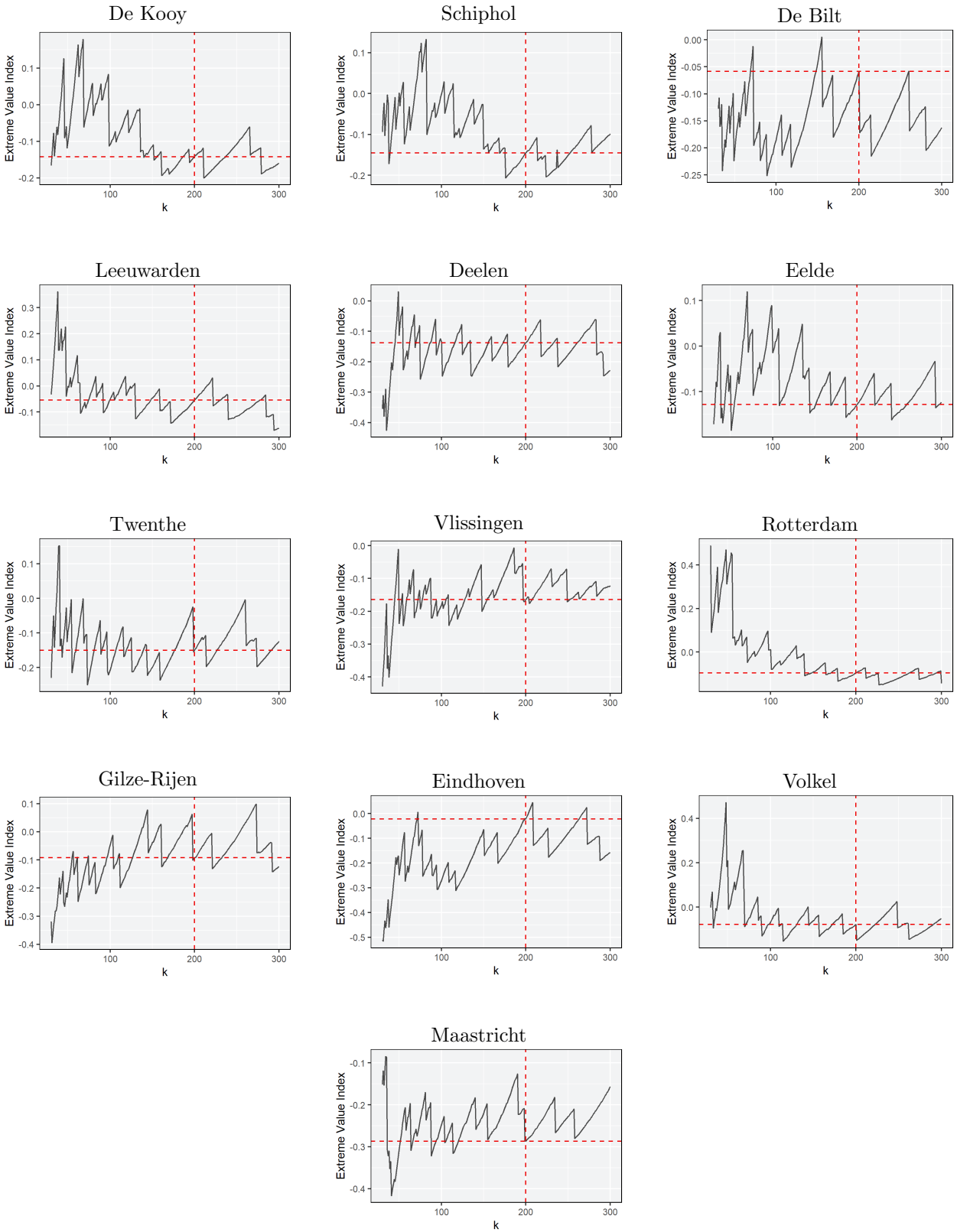

Figure 23: Parameter stability plots for the extreme value index as a function of the number of upper pseudo order statistics k.

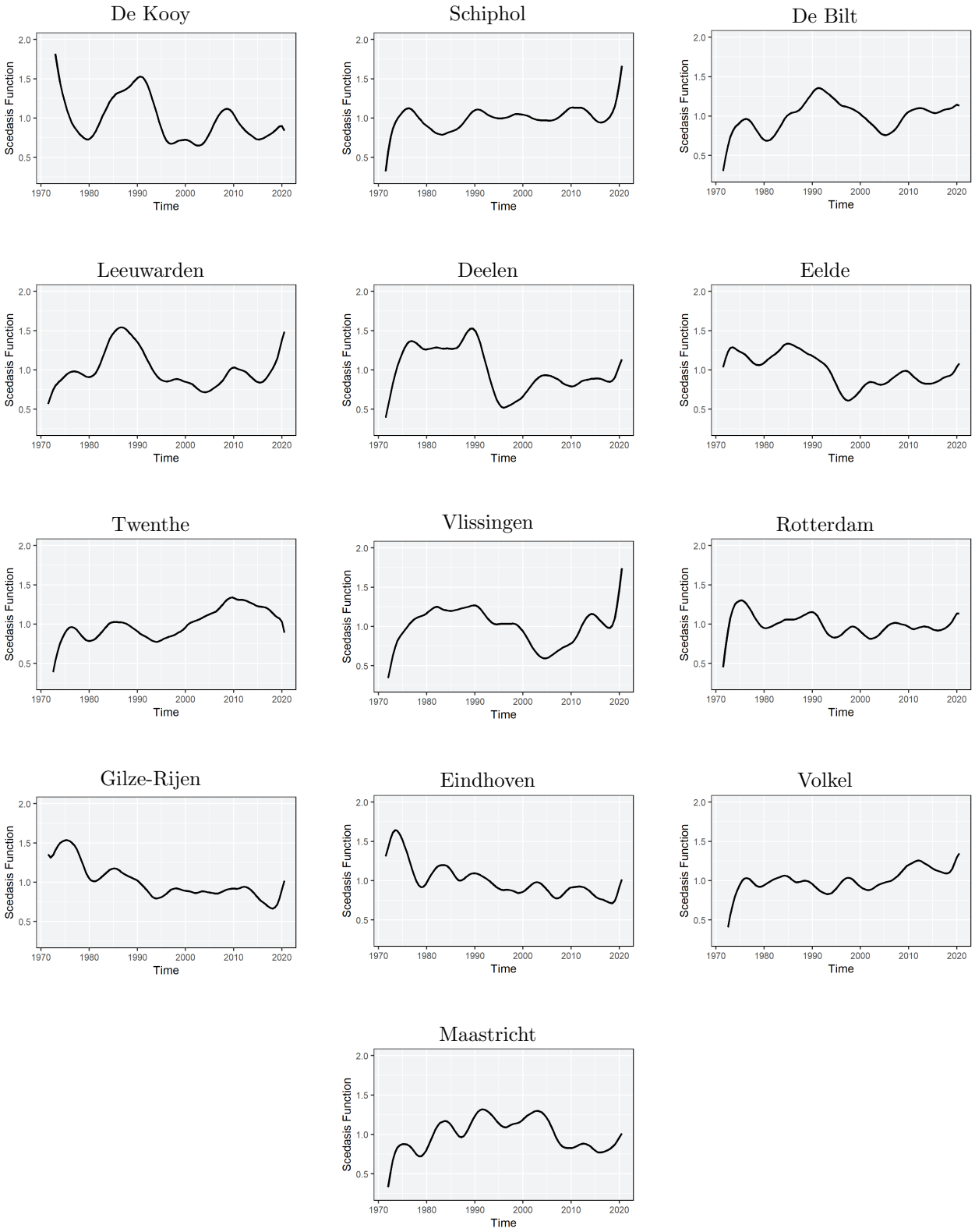

Figure 24: Plotted estimated scedasis functions for peak wind speed in summer season.

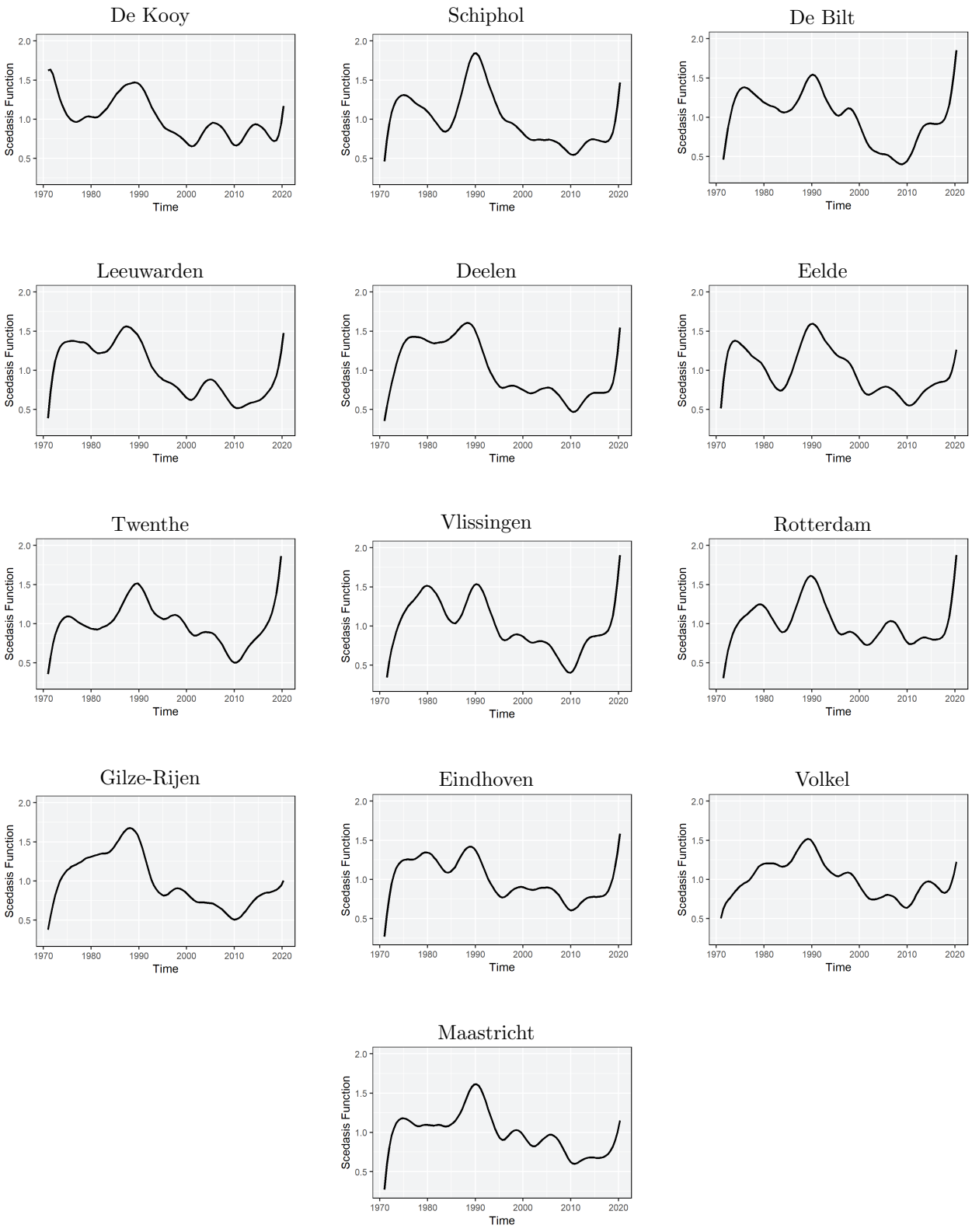

Figure 25: Plotted estimated scedasis functions for peak wind speed in winter season.

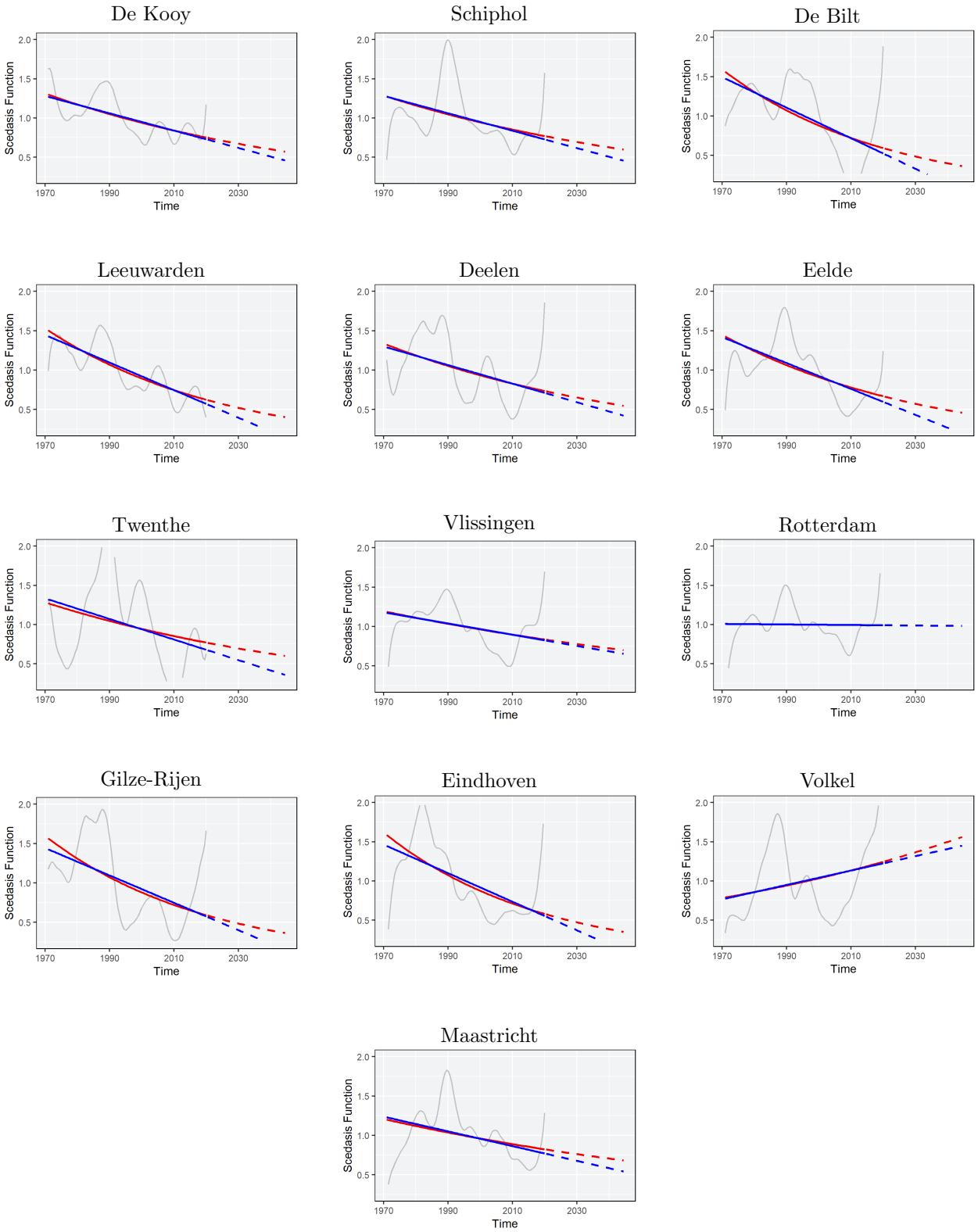

Figure 26: Forward looking scedasis function, parameterized using linear trend (blue) and log-linear trend (red) for winter season.