About the return period of a catastrophe

Mathias Raschke[1]

[1]Freelancer/independent researcher, Stolze-Schrey-Str.1, 65195 Wiesbaden, Germany

Correspondence to: Mathias Raschke (mathiasraschke@t-online.de.com)

Abstract. When a natural hazard event like an earthquake affects a region and generates a natural catastrophe (NatCat), the following questions arise: How often does such an event ~~occurs~~occur? What is ~~their~~its return period (RP)? We derive the combined return period (CRP) from a concept of extreme value statistics and theory ~~–~~ the pseudo-polar coordinates. A CRP is the (weighted) average of the local RP of local event intensities. Since CRP's reciprocal is its expected exceedance frequency, ~~which applies to any RP per stochastic definition,~~ the concept is testable. As we show, the CRP is related to the spatial characteristic of the NatCat generating hazard event and ~~their~~the spatial dependence of corresponding local block maxima (e.g., annual wind speed maximum). For this purpose, we extend previous construction for max-stable random fields from extreme value theory and consider ~~a~~the recent concept of area function from NatCat research. Based on the CRP, we also develop a new method to estimate the NatCat risk of a region via stochastic scaling of historical fields of local event intensities (represented by records of measuring stations) and averaging ~~corresponding risk parameters such as~~ the computed event loss ~~with a~~for defined ~~RP.~~

CRP or the computed CRP (or its reciprocal) for defined event loss. Our application example is winter storm (extratropical cyclones) over Germany. We analyze wind station data and estimate local hazard, CRP of historical events, and the risk curve of insured event losses. The most destructive storm of our observation period of 20 years is Kyrill 2002, with ~~weighted~~ CRP 16.97±1.75. The CRPs could be successfully tested statistically. We also state that our risk estimate is higher for the max-stable case than for the non-max-stable. Max-stable means that the dependence measure (e.g., Kendall's τ) for annual wind speed maxima of two wind stations has the same value as for maxima of ~~higher~~larger block size, such as 10 or 100 years since the copula (the dependence structure) remains the same. However, the spatial dependence decreases with increasing block size; a new statistical indicator confirms this. Such control of spatial characteristic and dependence is not realized by the previous risk models in science and industry. We compare our risk estimates to these.

1 Introduction

After a natural hazard event such as a large windstorm or an earthquake has occurred in a defined region (e.g., in a country) and results in a natural catastrophe (NatCat), the question arises, how often does such random event ~~appear~~occurs? What is the corresponding return period (RP, also called recurrence interval)? Before discussing this issue, we underline that the extension of river flood events or windstorms in time and space depend on the scientific and socio-economic event definition. ~~The~~

~~definitions~~This definition may vary by peril and is not our topic even though they influence our research object ~~–~~– the RP of a hazard and NatCat event.

The RP of an event magnitude or index is~~ so far~~ frequently used as a stochastic measure of a catastrophe. For example, there are different magnitudes scales for earthquakes (Bormann and Saul, 2009). ~~But~~However, their RP may not correspond ~~well~~ with the local consequences since the hypocenter position also determines local event intensities and effects. For floods, regional or global magnitude scales are not in use (Guse et al., 2020). For hurricanes, the Saffir–Simpson scale (National Hurricane Centre, 2020) is a magnitude measure; ~~but~~however, the random storm track also influences the extent of destruction. Extratropical cyclones hitting Europe, called winter storms, are measured by a storm severity index (SSI~~;~~, Roberts et al., 2014) or extreme wind index (EWI~~;~~, Della-Marta et al., 2009). Their different definitions result in quite different RP for the same events. In rare scientific publications about risk modelling for the insurance industry, such as by Mitchell-Wallace et al. (2017), better and universal approaches for the RP are not offered. In sum, previous approaches are not ~~very successful~~satisfactory regarding the stochastic quantification of a hazard or NatCat event~~ what~~. This is our motivation ~~is ~~to develop a new approach. Building on results of extreme value theory and statistics, we mathematically derive the concept of combined return period (CRP~~) being~~), which is the average of RPs of local event intensities~~ from an approach of extreme value theory and statistics.~~. As we will show by a combination of ~~previous~~existing and new approaches from stochastic and NatCat research, the concept of CRP is strongly related to the spatial association/dependence between the local event intensities, their RPs, and corresponding block maxima, such as annual maxima.

~~This ~~Spatial dependence is ~~less or inappropriately~~not suitably considered in previous research about NatCat. The issue is only a marginal topic in the book ~~about NatCat modelling for insurance industry ~~by Mitchell-Wallace et al. (2017, Section 5.4.2.5~~). Jongman's~~) about NatCat modelling for insurance industry. Jongman et al.'s (2014) model for European flood risk considers such dependence explicitly. However, their assumptions and estimates are not appropriate according to Raschke (~~2015~~2015b). In statistical journals, max-stable dependence models have been applied to natural hazards without a systematic test of the stability assumption~~, such as~~. Examples are the snow depth model by Blanchet and Davison (2011) for Switzerland and the river flood model by Asadi et al. (2015) for the Upper Danube River system. Max-stable dependence means that the copula (the dependence structure of a bi- or multivariate distribution) and corresponding value of dependence measures are the same for annual maxima as for ten-year maxima or these of a century (Dey et al., 2016). ~~Raschke's~~Also, Raschke et al. (2011) proposed a winter storm risk model for a power transmission grid in Switzerland ~~also implies this~~without validation of stability assumption~~ without a validation~~. The sophisticated model for spatial dependence between local river floods by Keef et al. (2009) is very flexible. However, it needs a high number of parameters, and the spatial dependence cannot be simply interpolated as ~~it is ~~possible with covariance and correlation functions (Schabenberger and Gotway, 2005, Section 2.4). Youngman and Stephenson (2016) suggested a statistical ~~modelling and simulation for~~model to simulate hazard events. As far as we understand, they generate wind fields by a Monte Carlo simulation of a complex random field. However, the random occurrence of a hazard event is more like a point event of a Poisson process than the draw/realization of a random variable. ~~An example illustrates~~The draw of the ~~difference; the ~~annual random variable ~~*local annual loss from catastrophes* is realized~~

~~every~~is sure; the occurrence of a Poisson point event in this year ~~even though~~is not ~~one catastrophe and loss event need to be occurred.~~sure but random. The same concern applies to the idea by Papalexiou et al. (2021) for storm simulation.

In the research of spatial dependence by Bonazzi et al. (2012) and Dawkins and Stephenson (2018), the local extremes of European winter storms are sampled by a pre-defined list of significant events. Such sampling is not foreseen in (multivariate) extreme value statistics; block maxima and (declustered) POT are the established sampling methods (Coles, 2010, Section 3.4 and 4.4; Beirlant et al., 2004, Section 9.3~~,~~ and 9.4). Event-wise spatial sampling is a critical task; the variable time lag between the occurrences at different measuring station, such as river gauging stations, makes it confusing. The corresponding assignment of ~~Asadi~~Jongman et al. (~~2015~~2014) of one local/regional flood peak to peaks at other sites ~~does~~is not ~~convince us completely. The same applies~~convincing, according to ~~Jongman et al. (2014;~~the comments by Raschke~~,~~ (2015b). The sampling of multivariate block maxima is simpler. However, the univariate sampling and analysis ~~is~~are also not trivial~~ as interpretations of~~. An example is the trend over decades in the time series of a wind station in Potsdam (Germany~~) over several decades shows.~~). Wichura (2009) assumes changed local roughness condition over the time as reason~~,~~; Mudelsee (2020) ~~the~~cites climate change as reason.

The research of spatial dependence of natural hazards is not an end in itself~~,~~; the final goal is an answer to the question about the NatCat risk. What is the RP of events with aggregate damage or losses in a region equal to or higher ~~to~~than a defined level? By using CRP, we quantify the risk via stochastic scaling of fields of local intensities of historical events and averaging corresponding risk measures. This new approach significantly extends the methods to calculate a NatCat risk curve. Previous opportunities and approaches for a risk estimate are the conventional statistical models that are fitted to observed or re-analyzed aggregated losses (also called as-if losses~~;~~) of historical events, as used by Donat et al. (2011) and ~~by~~ Pfeifer (2010) for annual sums. The advantages of such simple models are the controlled stochastic assumptions and the small number of parameters; the disadvantages are high uncertainty for widely extrapolated values and limited ~~opportunities~~possibilities to consider further knowledge. The NatCat models in (re)insurance industry combine different components/sub-models for hazard, exposure (building stock or insured portfolio~~)~~), and corresponding vulnerability (Mitchell-Wallace et al. 2017, Section 1.8; Raschke, 2018~~) and offers~~); additionally, they offer better opportunities for knowledge transfer such as the differentiated projection of a market model on a single insurer. However, the corresponding standard error of the risk estimates is frequently not quantified (and cannot be quantified). The numerical burden of such complex models is high. Tens of thousands of NatCat events must be simulated (Mitchell-Wallace eta al., 2017, Chapter 1). Thus, the question arises, what is the stochastic criterion for the simulation of a reasonable event set in NatCat modelling? As far as we know, scientific NatCat models for European winter storms (extratropical cyclones) are based on numerical simulations (Della-Marta et al., 2010; Osinski et al., 2016; Schwierz et al., 2010~~; Osinski et al., 2016~~) and are not intensively validated regarding spatial dependence.

To answer our questions, we start with topics of extreme value statistics in ~~the 2nd~~ Section ~~and illuminate~~ 2, where we recall the concept of max-stability ~~in the univariate sense,~~ for ~~the~~ single random variables, bivariate dependence ~~structure (copula) of the bivariate case and max-stable~~structures (copulas), and random fields. We also extend Schlather's (2002) 1st theorem with focus on spatial dependence. The more recent approaches ~~of hazard event related~~ area functions (Raschke, 2013) and survival functions (Jung and Schindler, 2019) of local event intensities within a region are implemented therein~~ to characterize spatiality.~~. In ~~the 3rd ~~Section 3, we derive the CRP from the concept of pseudo polar coordinates of extreme value statistics and explain its testability~~ and~~, possibility of scaling~~ opportunity,~~ and corresponding risk estimate. Subsequently, in Section 4, we apply the new approaches to winter storms (extratropical cyclones) over Germany to demonstrate their potentials. This application implies several elements of conventional statistics, which are explained in Section 5. Finally, we summarize and discuss our results and ~~give~~provide an outlook in Section 6. Some stochastic and statistical details are presented in the Supplementary and Supplementary data~~ to remain clarity of the main paper and limit its extent.~~. In the entire paper, we must consider several stochastic relations. Therefore, the same mathematical symbol can have different meanings in different subsections. We also expect that the reader is ~~more ~~familiar with ~~statistics~~statistical and stochastic ~~than only with basics about random variables.~~concepts such as statistical significance, goodness-of-fit tests, random fields, ~~or a~~and Poisson (point) ~~process~~processes (Upton and Cook, 2008~~) should be familiar terms.~~).

2 Max-stability in statistics and stochastic

2.1 The univariate case

Before ~~we formulate the ~~introducing CRP and its properties, we discuss ~~their opportunities, we must present, discuss, ~~and extend ~~a corresponding topic – ~~the concept of max-stability in extreme value statistics~~ especial of~~, with focus on random ~~process~~processes and fields. Max-stability has its origin in univariate statistics. The cumulative distribution functions (CDF) $F_n(x)$ of maximum $X_n = Max(X_1, \dots, X_n)$ of n identical and independently distributed (iid) random variables X_i with CDF $F(x)$ (for the non-exceedance probability $Pr(X \le x)$) is

$$F_n(x) = F(x)^n. \tag{1}$$

A CDF $F(x)$ is max-stable if the linear transformed maximum (with parameters a_n and b_n) has the same distribution (Coles, 2001, Def. 3.1)

$$F_n(a_n x + b_n) = F(a_n x + b_n)^n = F(x\cancel{)}). \tag{2}$$

The Fréchet distribution (Beirlant et al., 2004, Tab. 2.1) is such a max-stable distribution, also called extreme value distribution, with CDF

$$G(x) = \exp\left(-\frac{1}{x^\alpha}\right), x \ge 0, \alpha > 0\ . \tag{3}$$

For the unit Fréchet distribution, the parameter is $\alpha = 1$ and the transformation parameters are $b_n = 0$ and $a_n = n$. The most distribution types are not max-stable, but their distribution of maxima (1) converges to an extreme value distribution by increasing sample size n, called the block size in this context (Beirlant et al., 2004, Chapter 3). ~~These are well-known facts, and ~~We can only refer to some of a very high number of corresponding publications (e.g., de Haan and Ferrira, 2007; Falk et al. 2011). Coles (2001) gives a good overview for practitioners.

2.2 Max-stable copulas

It is also well-known that a bivariate CDF $F(x,y)$ can ~~replace~~be replaced by a copula $C(u,v)$ and the marginal CDFs $F_x(x)$ and $F_y(y$~~)~~):

$$F(x,y) = C\left(F_x(x), F_y(y)\right) = Pr(X \le x, Y \le y). \quad (4)$$

The copula approach represents a ~~universal~~basic distinction between the marginal distributions and the dependence structure~~ and~~; it was introduced by Sklar (1959). As there are different univariate distributions (types~~)~~), there are different copulas (types). Mari and Kotz (2001) ~~presents~~present a good overview about copulas, their construction ~~principals~~principles, and different views on dependence. Max-stability is also a property of some copulas, called max-stable copula or extreme copula. A max-stable copula remains the same for pairs of ~~component wise~~componentwise maxima (X_n, Y_n) as it was already for the underlying pairs (X, Y); the copula parameters including dependence measure such as Kendall's (1938) rank correlation are equal. The formal definition is (Dey et al., 2016, (2.3))

$$C_n(u\,,v) = C(u^{\,1/n}, v^{\,1/n})^n. \quad (5)$$

2.3 Max-stability of stochastic processes

The spatial extension of the bivariate situation and corresponding distribution is the random field $Z(x)$ at points x in the space $\mathbb{R}^d$ with d dimensions (e.g., Schlather, 2001). In our application, $\mathbb{R}^2$ is the geographical space and x is the corresponding coordinate vector. At one point/site, x in $\mathbb{R}^d$, $F_x(z)$ is the marginal distribution of the local random variable Z. There are various differentiations and variants such as (non)stationarity or (non)homogeneity. A max-stable random field has max-stable marginal distributions and the copulas between ~~to~~two margins are also max-stable. Schlather (2002) has formulated and proofed a construction of a max-stable random field (we cite his 1st theorem with the same notation)

Theorem 1: *Let Y be a measurable random function and $\mu = \mathbb{E}\int_{\mathbb{R}^d} max\{0, Y(x)\}\, dx \in (0,\infty)$. Let Π be a Poisson process on $\mathbb{R}^d \times (0,\infty)$ with intensity measure $d\Lambda(y,s) = \mu^{-1} dy s^{-2} ds$, and $Y_{y,s}$ i.i.d. copies of Y; then*

$$Z(x) = \sup_{(y,s)\,\in\,\Pi} sY_{y,s}(x-y) = \sup_{(y,s)\,\in\,\Pi} s\, max\{Y_{y,s}(x-y), 0\}, \quad (6)$$

is a stationary max-stable process with unit Fréchet margins.

Extreme value statistics is interested in the max-stable dependence structure (copula) between the margins, the unit Fréchet distributed random variables Z at fixed points x in space $\mathbb{R}^d$. From perspective of NatCat modelling in the geographical space $\mathbb{R}^2$ and with $Y(x) \ge 0$, the entire generating process is interesting. The Poisson (point) process Π represents all hazard events (e.g., storms) of a unit period such as a hazard season or a year~~ and~~; it has two parts, s and y. The point events s on $(0,\infty)$ are a stochastic event magnitude and scale the field of local ~~events~~event intensity $s_x(x$~~), short called intensity~~):

$$s_x(x) = sY_{y,s}(x-y\cancel{)}), \quad (7)$$

which represents all point events $s_x(x)$ at sites x. The random coordinate y is a kind of epicenter in the meaning of NatCat with the (tendentiously) highest local event intensity such as maximum wind speed, maximum hail stone diameter, or peaks of earthquake ground accelerations. The copied random function $Y(x)$ determines the pattern of a single random event in the space $\mathbb{R}^d$. $Y(x)$ or its local expectation converges to 0 or is 0 if the magnitude $\|x\|$ of coordinate vector converges to infinity due to the measurability condition in Theorem 1. This also applies to NatCat events with limited geographical extend.

Schlather (2002) has demonstrated the flexibility of his construction by presenting realizations of maximum fields for different variants of $Y(x)$. Its measurability condition is fulfilled by classical probability density functions (PDF, first derivative of the CDF, Coles, 2001, Section 2.2) of random variables. For instance, Smith (1990, an unpublished and frequently cited paper) used the PDF of the normal distribution. We present ~~in the Supplementary, Section 4,~~ some examples of the random function $Y(x)$ in the Supplementary, Section 4, to illustrate the universality of the approach. $Y(x)$ can also imply random parameters such as ~~verities~~variants of standard deviation of applied PDF~~ or is~~, and it can be combined with a random field.

Both, s and $s_x(x)$ with fixed x, are point events of Poisson processes with intensity $s^{-2}ds$. This is the expected point density and determines the exceedance frequency. The latter is the expected number of point events $s_x(x) > z$ and $s > z$:

$$\cancel{A}\Lambda(z) = \int_z^{\infty} s^{-2}ds = 1/z. \tag{8}$$

The entire construction of Theorem 1 is also a kind of shot noise field according to the definitions of Dombry (2012~~); and~~). Furthermore, Schlather (2002) has also published a construction of max-stable random field without a random function but with a stationary random field. The logarithmic variant of Theorem 1 (logarithm of (6,7)) also results in a max-stable random field~~,~~; however, the marginal maxima are unit Gumbel distributed and (8) would be an exponential function. The Brown-Resnick process~~ –~~ — well-known in ~~stochastic~~extreme value statistics (e.g., Engelke et al., 2011~~) – determines~~) — generates a max-stable random field with such unit Gumbel distributions ~~and use a nonstationary~~ as result of random ~~field.~~walk processes (Pearson, 190). It is implicitly a construction according to Theorem 1~~ since~~, as for exponential transformation (inverse of logarithmic transformation~~)~~), the ~~nonstationary~~ random ~~field~~walk with drift is the random function of Theorem 1. The origin of a Brown-Resnick process in $\mathbb{R}^d$ can be fixed but can also be a random coordinate as y is in Theorem 1.

The construction of Theorem 1 is already used to model natural hazards in the geographical space. Smith (1990) has applied the bivariate normal distribution as $Y(x)$ in a rainstorm modelling. The Brown-Resnick Process has ~~been~~ already been applied to river flood (Asadi et al., 2015). Blanchet and ~~Davison~~Davison's (2011) ~~have applied a max-stable~~ model for snow depth, and Raschke et al.'s (2011) model for winter ~~storm~~storms, both in Switzerland~~. And there~~, are also max-stable. There are also similarities to conventional hazard models. ~~Punge's~~Punge et al.'s (2014) hail simulation includes maximum hail stone diameter that acts like ~~ln (~~$ln(s)$ in (6,7). Raschke (2013) already stated similarity between earthquake ground motion models and Schlather's construction. However, the earthquake magnitude can have a wider influence on the geographical event pattern than ~~a~~ simple scaling. This was one of ~~the~~our motivations to extend and generalize the Schlather's construction (7) with dimension d of $\mathbb{R}^d$:

$$s_x(x) = s^{1+\beta} Y_{y,s}\left(\left((1+\beta)s^{-\beta}\right)^{-\frac{1}{d}}(x-y)\right), \beta > -1, \tag{9}$$

and for the corresponding field of maxima (6~~)~~), we write

$$Z(x) = \sup_{(y,s)\,\in\,\Pi} s^{1+\beta} Y_{y,s}\left(\left((1+\beta)s^{-\beta}\right)^{-\frac{1}{d}}(x-y)\right), \beta > -1. \tag{10}$$

As we show in the Supplementary, Section 2, the marginal Poisson processes $s_x(x)$ in (9) have the same exceedance frequency (8) as (7). Correspondingly, $Z(x)$ in (10) is also unit Fréchet distributed as in (6). Schlather's construction is a special case of (9,10) with $\beta = 0$; (9,10) only implies max stability of spatial dependence in this case, what we discuss in the following section.

2.4 Spatial characteristics and dependence

We now illustrate spatial max-stability and its absence by examples of (9,10) with standard normal PDF as random function $Y(x)$ in a one-dimensional parameter space $\mathbb{R}^{d=1}$. For this purpose, we apply the simulation approach of Schlather (2002) and generate random events within a range (-10,10) for local event intensities within the region/range (-4,4) in $\mathbb{R}^1$ by a Monte Carlo simulation. According to Schlather's procedure~~ that~~, which processes a series of random numbers from a (pseudo) random generator, only the events for the large s are simulated~~ which~~; this implies incompleteness for smaller events. This does not significantly affect the simulated field $Z(x)$ of maxima. However, we can only consider this simulation for $\beta \geq 0$ in (9,10) since the edge effects increase for increasing s if $\beta < 0$. In Figure 1a, we show fields for one realization Π of Schlater's theorem ($n = 1$, equivalent to one year or one season in NatCat modelling) for the max-stable case with $\beta = 0$ in (9,10). With the same series of random numbers, we generate fields of $n = 100$ realizations of Π in Figure 1b. It has the same pattern $n = 1$ and is the same when we linear transform the local intensities s_x, with division by $n = 100$. The entire generating processes are max-stable, just as the resulting marginals~~ and dependence~~ and association/dependence between marginals are. In contrast to this total max-stability, the example with $\beta = 0.1$ results in different ~~pattern~~patterns for $n = 1$ and $n = 100$ in Figure 1 c and d. The shape of the event fields gets sharper for larger s~~,~~; only the marginals are max-stable, not their spatial relations.

To ~~illustrates~~illustrate the effect on spatial dependence quantitatively, we have generated local maxima $Z(x)$ from (10) by Monte Carlo simulation with 100,000 repetitions and computed corresponding dependence measure Kendall's τ (Kendall, 1938; Mari and Kotz, 2001, Section 6.2.6). As depicted in Figure 2a and b, the functions are the same if $\beta = 0$ and differs if $\beta = 0.1$~~,~~; the dependence is decreasing by increasing n if $\beta > 0$. In Figure 2c, the functions are shown for the limit cases full dependence with the same value of $s_x(x)$ at each point x and full independence with $s_x(x) = 0$ everywhere except one point.

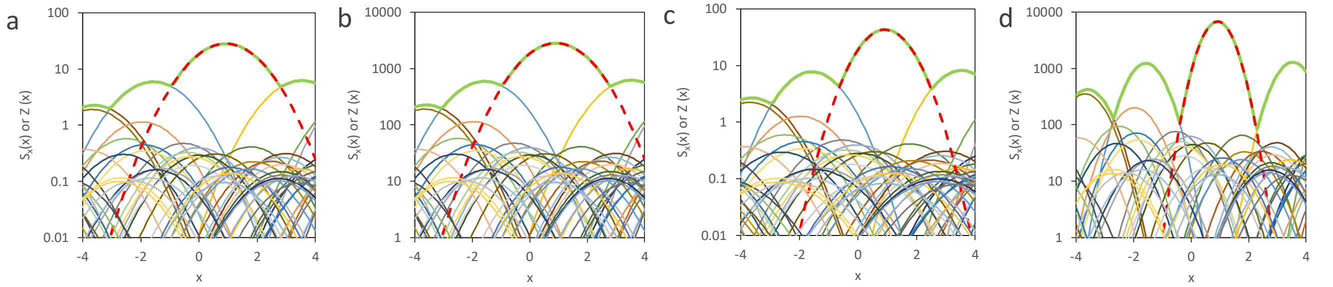


Figure 1: Examples of simulated fields of local event intensities and enveloping field of maxima (bold green line) generated with standard normal PDF as *Y(x)* in (6,7,9,10) and the same series of numbers from pseudo-random generator: a) max-stable and n=1, b) max-stable and n=100, c) non max-stable and n=1, d) non max-stable and n=100. The strongest event has a broken red line.

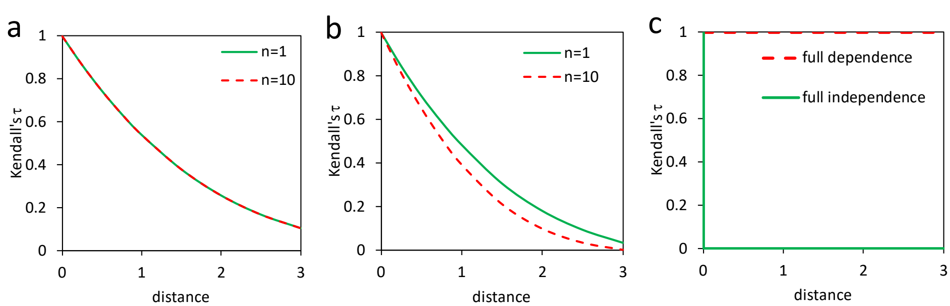


Figure 2: Spatial dependence in relation to the distance measured by Kendall's τ: a) max-stable fields of Figure 1, b) non-max-stable fields of Figure 1, c) limit cases.

Beside our extension of Schlather's theorem, we also consider a more recent approach from NatCat research to understand the spatial characteristic. Raschke (2013) described an earthquake event by its area function for the peak ground accelerations. This is a cumulative function and measures the set of points in the geographical space (the area) with an event intensity higher than the argument of the function. The area function is limited here to a region and is normalized (u and l symbolises the region's bounds, the integral in the denominator is the area of the region in $\mathbb{R}^2$, $\mathbf{1}$ is an indicator function~~)~~):

$$A(z) = \frac{\int_l^u \mathbf{1}(s_x(x) > z)\,dx}{\int_l^u dx}. \qquad (11)$$

~~and~~It is now like a survival function of a random variable (decreasing with the value of functions between 0 and 1~~)~~), which describes the exceedance probability in contrast to a CDF for non-exceedance probability (Upton and Cook, 2008). Jung and Schindler (2019) have already applied such aggregating functions to German winter storm events and call them explicitly survival function. However, not every normalized aggregating decreasing function is based on an actual random variable. ~~And~~Moreover, survival functions are not used in statistics to describe regions of random fields or random function as far as we know. Nonetheless, we use the area function (11) to characterize and research the spatiality of the event field $s_x(x)$ in a defined region. As an example, the area function for the strongest events in Figure 1 is shown in Figure 3a. The differences between the variants $n = 1$ versus $n = 100$ and $\beta = 0$ versus $\beta = 0.1$ ~~corresponds~~correspond with the differences between these events in Figure 1. In Figure 3b, the limit cases~~ of Figure c~~ are depicted to illustrate the underlying link between area function and spatial dependence.

We also use the parameters of a random variable ~~X~~Z with PDF $f(x)$ and CDF $F($~~x~~$z)$, and survival function $\bar{F}$~~(x)~~$(z) = 1 - F($~~x~~$z)$ to characterize our area function. These parameters are ~~exception~~expectation $\mathbb{E}$~~$[X]$~~$[Z]$ (estimated by sample mean/average), variance Var~~$[X]$~~$[Z]$, standard deviation Sd~~$[X]$~~$[Z]$ (the square root of variance), and a coefficient of variation (CV) Cv~~$[X]$ with~~$[Z]$ (Coles, 2001, Section 2.2~~;~~; Upton and Cook, 2008) with

$$\mathbb{E}\cancel{[X] = \int_{-\infty}^{\infty} x f(x) dx}[Z] = \int_{-\infty}^{\infty} z f(z) dz = \cancel{\int_1^0 x d\bar{F}(x)} \int_1^0 z d\bar{F}(z), \mathrm{Var}\cancel{[X]}[Z] = \cancel{\int_1^0 (x - \mathbb{E}[X])^2 d\bar{F}(x)} \int_1^0 (z - \mathbb{E}[Z])^2 d\bar{F}(z), \mathrm{Sd}\cancel{[X]}[Z] = \cancel{\sqrt{\mathrm{Var}[X]}}\sqrt{\mathrm{Var}[Z]}, \mathrm{Cv} = \cancel{\frac{\mathrm{Sd}[X]}{\mathbb{E}[X]}} \frac{\mathrm{Sd}[Z]}{\mathbb{E}[Z]}. \quad (12)$$

According to (12), any scaling of ~~X~~Z by a factor $S > 0$ results in proportional scaling of expectation and standard deviation in (12), and the CV remains constant. Correspondingly, random magnitude s in (9,10) only scales the field $s_x(x)$ in the max-stable case with $\beta = 0$ and influences the expectation of $A(z)$ but not the CV. Thus, the CV is independent on the expectation. This does not apply to the non-max-stable case with $\beta \neq 0$ in (9,10). These different behaviors are detectable for the examples of Figure 1b and d in Figure 3c and d. For the max-stable case, the scale/slope parameter of the linearized regression function does not differ significantly from 0 according to the t-test (Fahrmeir et al., 2013, Section 3.3). For the max-stable case, the regression function is statistically significant with a p-value of 0.00. Linearization is provided by the logarithm of CV and expectation/average. For completeness, the full dependence case of Figure 3b corresponds with ~~an~~a CV of 0.

~~In sum of~~As per Section 2, Schlather's 1st Theorem has parallels to NatCat models, is used already in hazard models, and was extended here to the non-max-stable case regarding spatial dependence and characteristic. Statistical indication for max-stability is the independence of the spatial dependence measure from the block size (e.g., one versus ten years) and independence between CV and expectation of the area function (11). Otherwise, non-max-stability is indicated.

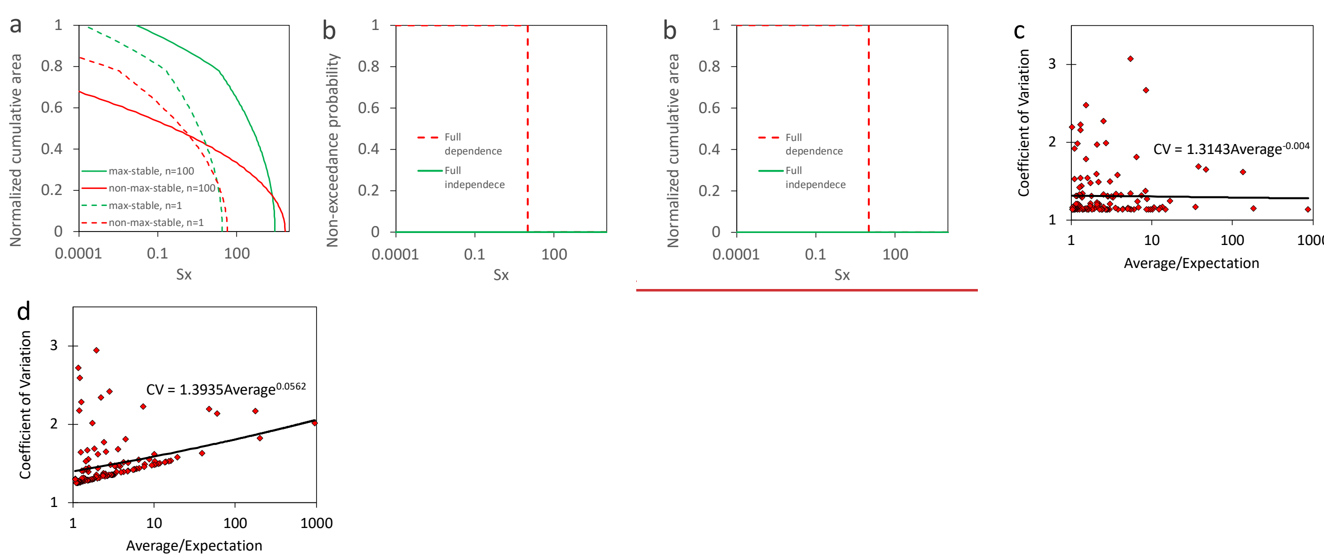


Figure 3: The area function and corresponding characteristics: a) area function of the biggest event of Figure 1, b) area functions for the limit cases (examples), c) relation CV to average for the max-stable case of Figure 1b, d) for the non-max-stable case of Figure 1d (events with average >1, the distance between the support points is 0.1 for the computation of the average in the region ~~(-~~[-4,4~~)).~~]).

3 The combined return period (CRP)

3.1 The stochastic derivation

Let the point event $\cancel{s_{x,i}}Y_{x,i}$ be the local intensity at site x of a hazard event and i as a member of the set of all events of a defined unit period such as a year, hazard season, or half season. This local event intensity might be the maximum river discharge of a flood, the peak ground acceleration of an earthquake, or the maximum wind gust of a windstorm event. The entire number of events with $\cancel{s_{x,i}}Y_{x,i} > \cancel{z}y$ during the unit period is $K = \cancel{\sum_{i=1}^{\infty}\mathbf{1}(s_{x,i} > z)}\sum_{i=1}^{\infty}\mathbf{1}(Y_{x,i} > y)$. K is (at least approximately) a Poisson distributed (Upton and Cook, 2008) discrete random variable with an expectation ~~–~~ — the expected exceedance frequency, that is the local hazard function in a NatCat model (this is not the hazard function/hazard rate of statistical survival analysis, Upton and Cook, 2008~~).~~)

$$\cancel{\Lambda(z)}\Lambda_y(y) = \mathbb{E}[K]. \tag{13}$$

This is the bijective frequency function and the local hazard curve. Its reciprocal determines the hazard curve for the RP

$$\cancel{T(z)}T_y(y) = \cancel{\frac{1}{\Lambda(z)}}\frac{1}{\Lambda_y(y)} = \frac{1}{\mathbb{E}[K]}. \tag{14}$$

As $\cancel{s_x}Y_x$ is a point event ~~it's~~, its RP $T\cancel{(s_x)} = T_y(Y_x)$ is also a point event of a point process with frequency function according to (14) but now with the argument/threshold variable z, since the scale unit is changed:

$$\Lambda_T(z) = 1/z. \tag{15}$$

Since (15) is the same as (8), Schlather's theorem and our extensions directly apply to RP~~.~~ with $T = s_x$ in (7,9). For completeness, the marginal maxima have a CDF for n unit periods (a unit Fréchet distribution for $n = 1$ according to (3))

$$G_n(z) = exp\,\cancel{(-n\Lambda_T(z)) =}\,(-n\Lambda_T(z)) = exp\,(-n/z). \tag{16}$$

This is applicable because the probability of non-exceedance for level z of the block maxima is the same probability ~~that~~as per which no events occur with $\cancel{s_{x,i} >}\,T > z$, which is determined by the Poisson distribution; (6~~,8~~) also ~~imply~~implies this link, and Coles (2002, p. 249 "y_p") has also mentioned this. The same applies to the relation between frequency and maxima of locale event intensity.

Schlather's theorem is also based on and implies the concept of pseudo polar coordinates. According to de Haan (1984) and well explained by Coles (2001, Section 8.3.2), two max-stable linked point processes with expected exceedance frequency (15) and point events T_1 and T_2 are also represented by pseudo polar coordinates with radius R and angle V:

$$\left\{R = T_1 + T_2, V = \frac{T_1}{T_1+T_2}\right\} \Leftrightarrow \left\{T_1 = RV, T_2 = R(1-V) = T_1\frac{1-V}{V}\right\}. \tag{17}$$

As we describe in the Supplementary, Section 1, the expectation of $(1-V)/V$ is 1 and for the conditional expectation of unknown RP T_2 with known T_1 applies (association is provided~~)~~):

$$\cancel{E}\mathbb{E}[T_2|T_1] = T_1. \tag{18}$$

The interest in extreme value theory and statistics (Coles, 2001, Section 3.8; Beirlant et al., 2004, Section 8.2.3; Falk et al. 2011, Section 4.2) is focused on the distribution of pseudo angle V with CDF $H(z)$. As Coles (2001) ~~write~~writes "*the angular spread of points of* N [the entire point processes] *is determined by H, and is independent of radial distance* $[R]$", angle and radius ~~occurs~~occur independently to each other, and H ~~determine~~determines the copula between two marginal maxima $Z(x)$ in Theorem 1.

According to Coles (2001, Section 3.8), the pseudo radius R in (17) is a point event of a Poisson process with frequency $\Lambda(z) = 2/\cancel{x}z$ - the double of (15). This means the average of two ~~RP~~RPs, T_1 and T_2, results in a combined return period (CRP) T_c

$$T_c = \frac{T_1+T_2}{2}\cancel{.} \tag{19}$$

with exceedance frequency function (8,15). We do not have a mathematical proof that (18,19) also applies for non-max-stable associated point processes. However, max-stable and non-max-stable cases have the same limits: full dependence ($T_1 = T_2$) and no dependence/full independence ($T_1 = 0$ if $T_2 > 0$ and vice versa, $T = 0$ represents the lack of a local event). Therefore, (19) should also apply to the non-max-stable case between these limits. This can be ~~validated~~ heuristically validated as we demonstrate by an example in the Supplementary, Section 3.

More than one RP can be averaged since the averaging of two RPs can be done in serial (and the pseudo polar coordinates are also applied to more than two marginal processes). Serial averaging (averaging the last result with a further RP) also implies a weighting; the first considered RPs would be smaller weighted than the last in the final CRP. The general formulation of averaging of RP with weight w is

$$T_c = \frac{\sum_{i=1}^{n} T_i w_i}{\sum_{i=1}^{n} w_i}. \tag{20}$$

The corresponding continuous version within the region's bounds u and l in space $\mathbb{R}^d$ is

$$T_c = \frac{\int_l^u T(x)w(x)dx}{\int_l^u w(x)dx}. \tag{21}$$

If $w(x) = 1$ applies in (21~~)~~), then the denominator is the area of the region ~~and~~. Furthermore, the CRP T_C is the expectation of the area function (11). This also applies for other weightings if we consider it in the area function, here written for RP $T(x$~~)~~),

$$A(z) = \frac{\int_l^u w(x)\mathbf{1}(T(x)\geq z)dx}{\int_l^u w(x)dx}, \tag{22}$$

with empirical version for n measuring station i in the analyzed region:

$$A(z) = \frac{\sum_{i=1}^{n} w_i \mathbf{1}(T_i \geq z)}{\sum_{i=1}^{n} w_i}. \tag{23}$$

The weighting, especially the empirical one, can be used in hazard research to compensate an inhomogeneous geographical distribution of measurement stations or a different focus than the covered geographical area such as the inhomogeneous distribution of exposed values or facilities in NatCat research. It has the same effect on the area function as a distortion of the geographical space as used by Papalexiou et al. (2021). Weighted or not, CRP and CV are parameters of the area function.

3.2 Testability

Before the CRP is applied in stochastic NatCat modelling, it should be tested statistically to validate the appropriateness. A sample of CRPs can be tested by a comparison of its exceedance frequency function (15) and their empirical variant. Therein the empirical ~~EF~~exceedance frequency of the largest CRP in the sample is the reciprocal of the length of the observation period. The 2nd largest CRP is hence associated to twice the exceedance frequency of the largest CRP and so on. It is the same as for empirical exceedance frequency for EQ (e.g., the well-known Gutenberg-Richter relation in Seismology, Gutenberg-Richter, 1956). However, not all small events are recorded; the sample is thinned and incomplete. This completeness issue is well known for earthquakes and is ~~here~~ less important here if only the distribution (16) of maximum CRPs is tested. There are ~~a number~~several goodness-of-fit tests (Stephens, 1986, Section 4.4) for the case of known distribution ~~model~~. The Kolmogorov-Smirnov test is a popular variant.

3.~~2~~3 The scaling ~~opportunity~~property of CRP

The CRP also offers the opportunity of stochastic scaling. The CRP T_c and all n local RPs T_i in (20) (and $T(z)$ in (21)) are scaled by a factor $S > 0$:

$$T_{cs} = \cancel{T_c} S \cancel{= \frac{\sum_{i=1}^{n} ST_i w_i}{\sum_{j=1}^{n} w_i}} T_c = \frac{\sum_{i=1}^{n} ST_i w_i}{\sum_{j=1}^{n} w_i}, T_{s,i} = \cancel{T_i} S \cancel{,}\ T_i. \tag{24}$$

This means for the pseudo polar coordinates in (17), which applies to the max-stable case~~,~~:

$$R_s = ST_1 + ST_2 = SR, V_s = \frac{T_{s,1}}{T_{s,1} + T_{s,2}} = \frac{ST_1}{S(T_1 + T_2)} = V. \tag{25}$$

The pseudo angle V is not changed as expected since pseudo radius and pseudo angle are independent in the pseudo polar coordinate for the max-stable case (Section 3.1). This also means that a scaling must be more complex if there is non-max-stability. We cannot offer a general scaling method for this situation; however, it must consider/reproduce the pattern of the relation CV versus CRP (example in Figure 3d) adequately. Irrespective of this, the corresponding event field of local intensities (e.g., maximum wind gust speed) can be computed for the scaled local RPs via the inverse of the local hazard function: $T(z)$ in (14) or $\Lambda(z)$ in (13).

3.4 Risk estimates by scaling and averaging

The main goal of a NatCat risk analysis is the estimate of a risk curve (Mitchell-Wallace et al., 2017, Section 1), the bijective functional of event loss in a region~~,~~ and corresponding RP, which is called ~~here~~the event loss return period (ELRP) T_E. As aforementioned, there are two approaches for such estimates with corresponding pros and cons.

We introduce an alternative method. Under the assumption of max-stability between ELRP T_E and CRP T_C, according to (18~~),~~) with $T_1 = T_C$ and $T_2 = T_E$, the expectation of an unknown ELRP T_E is the CRP ~~T_c~~T_C of the local event intensities~~;~~. This means that the CRP is an estimate of ~~the ELRP (max-stability between ELRP and CRP provided). To get a good estimate of ELRP, we must~~ ELRP. We can average the ~~T_c~~CRP of many events with the same event loss~~. We cannot observe such, but we~~

~~can stochastically scale historical events respectively their~~ to get a good estimate of ELRP. However, observations of events with the same loss are not available. Nonetheless, we can exploit the stochastic scaling property of CRP to rescale the local intensity observations of historical events to get the required information. The modelled event loss L_E is the sum of the product of local loss ratio L_R, determined by local event intensity ~~$s_{x,i}$~~ $y_{x,i}$ and local exposure value E_i over all sites i (Klawa and Ulbrich, 2003; Della-Marta et al. 2010~~)~~):

$$L_E = \cancel{\sum_{i=1}^{n} L_{R,i}(s_{x,i})E_i} \sum_{i=1}^{n} L_{R,i}(y_{x,i})E_i \qquad (26)$$

with the local vulnerability function $L_{R,i}($~~$s_{x,i}$~~$y_{x,i})$. To get the event loss for the scaled event, the observed ~~$s_{x,i}$~~ $y_{x,i}$ is replaced by

$$\cancel{s_{xs,i} = \Lambda_i^{-1}\left(\frac{\Lambda_i(s_{x,i})}{S}\right) \qquad (27)}$$

$$y_{xs,i} = \Lambda_{y,i}^{-1}\left(\frac{\Lambda_{y,i}(y_{x,i})}{S}\right) = T_{y,i}^{-1}\left(S T_{y,i}(y_{x,i})\right). \qquad (27)$$

with local hazard function ~~$\Lambda(z)$~~ (13~~) for exceedance frequency~~ ,14) and its invers function ~~$\Lambda^{-1}(z)$. The formulation with RP (14) is equivalent.~~. The scaling factor S in (27) is the same for all sites/locations i, respectively, as it is in (24) for the CRP ~~and~~. This factor S must be adjusted ~~in an iteration~~iteratively until the ~~defined event loss is the~~ result of (26~~).~~) converges to the desired event loss. The scheme in Figure 4a includes all elements and relations of the scaling approach. Therein, the numerical determination in the scaling scheme has only one direction, from scaled CRP to the event loss. The idea of CRP averaging is also illustrated by Figure 4b. The standard error of the averaging is the same as for the estimates of an expectation by the sample mean (Upton and Cook, 2008, ~~catchword~~keyword central limit theorem).

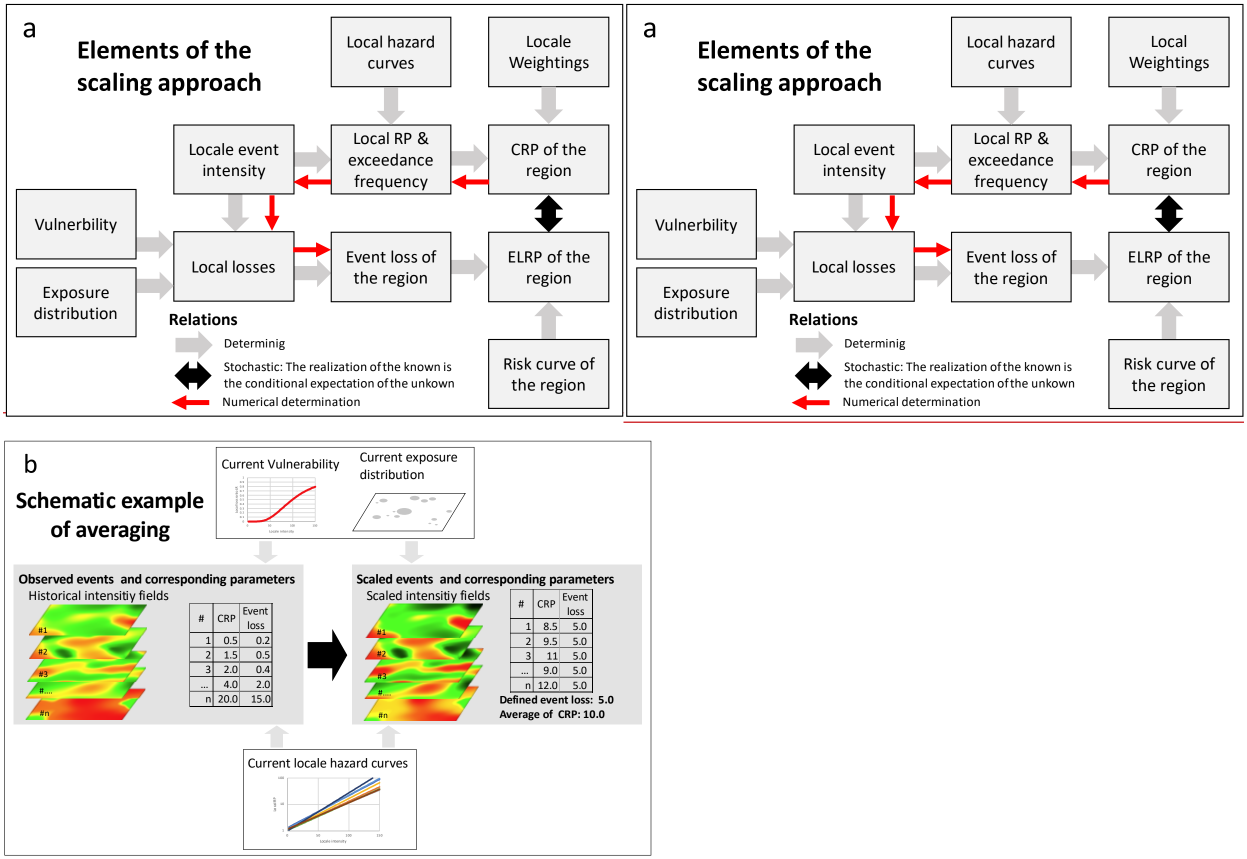


Figure 4: Schemes of the scaling approach: a) elements and relations, b) schematic example for estimation RP of event loss by averaging of CRP.

According to the ~~well-known~~ Delta method~~, well explained by~~ (Coles ~~(~~2001, Section 2.6.4), statistical estimates and their standard error can be approximately transferred in ~~other~~another parameter estimation and corresponding standard error by the determined transfer function and its ~~derivates. In its meaning, we can also average the event loss for a fixed/determined CRP respectively its scaled variant. This also applies to~~ derivatives. Condition of this linear error transfer is a relatively small standard error of the original estimates. The Delta method could be used to compute the reciprocal of ELRP—the exceedance frequency~~, the reciprocal of RP, for a determined/fixed event loss~~ and corresponding standard error, or the exceedance frequency is computed directly by averaging the reciprocal of scaled CRPs, and the transfer proxy acts implicitly. We also apply the idea of linear transfer proxy when we average the modelled event loss for the historical events being scaled to the same defined CRP. The unknown sample of ELRPs, which represents the ELRP's distribution for a fixed CRP, is implicitly transferred to a sample of event loss. If the proxies perform well, the difference to the risk estimates via CRP averaging should be small.

There is a further chain of thoughts as argument for the different variants of averaging~~ the exceedance frequency.~~. The ~~scaled~~scaling implies subsets of intensity fields ~~are like sub-sets in the set~~ of all possible ~~event~~intensity fields. ~~Each of these~~

~~subsets implies a relation exceedance frequency to event loss – a risk curve. Furthermore, we assume an~~ The links between the fields of a subset are determined by the scaling of their CRPs. Correspondingly, every subset generates a risk curve with CRP (now also an ELRP) versus event loss. We also assume a certain unknown probability per subset that ~~this sub-set generates a part of~~ is applied if all these subsets generate the entire risk curve~~. The latter is for a fixed event loss the aggregation of subset exceedance frequency multiplied with their probability. This is basically the same as the definition of~~ via an integral like the expectation of a random variable (12). ~~Therefore, we can apply~~ The corresponding empirical variant (estimator ~~of the expectation in the estimation of a risk curve and~~) is the averaging. However, we can average ~~the~~three values: CRP, exceedance ~~frequencies of risk curves of the subsets – the scaled~~frequency, or event ~~fields~~loss.

All mentioned estimators for risk curve via scaling and averaging over n events are

$$\hat{T}_E(L_E)=\frac{1}{n}\sum_{i=1}^{n}T_{CS,i}(L_E)\,,\hat{\Lambda}_E(L_E)=\frac{1}{n}\sum_{i=1}^{n}1/T_{CS,i}(L_E)\,,\hat{L}_E(T_E)=\frac{1}{n}\sum_{i=1}^{n}L_{E,i}(T_{CS}=T_E). \tag{28}$$

The right side of the equations in (28) implies actual values which can be and are ~~be~~being replaced by estimates. Corresponding uncertainties must be considered in the final error quantification.

We draw attention to the fact that the explained scaling does not change the CV of (23~~) which~~); this implies independence between CRP and CV (Section 2.4). Therefore, the presented scaling only applies to the max-stable case of local hazard. For the non-max-stable case, the scaling factor S in (27) must be replaced event wise by S_i, which ~~reproduce~~reproduces the observed relation between CRP and CV. An example without max-stability was already shown in Figure 3d.

4 Application to German winter storms

4.1 Overview about data and analysis

We have selected the peril winter storms (also called extratropical cyclones or ~~wind~~ winter windstorms) over Germany to demonstrate the opportunities of the CRP because of good data access and since we are familiar with this peril (Raschke et al., 2011; Raschke, 2015). Our analysis follows the scheme in Figure 4a, important results are presented in the subsequent sections, and the technical details are explained in Section 5. At first, we ~~give~~provide an overview.

We analyzed 57 winter storms over 20 years, from autumn 1999 to spring 2019 (Supplementary data, Table 1 and 2) to validate the CRP approach. Different references (Klawa and Ulbrich, 2003; Gesamtverband Deutscher Versicherer [GDV], 2019; Deutsche Rück, 2020) have been considered to select the time window per event. In our definition ~~of a~~, the winter storm season is from September to April of the subsequent year ~~and accept~~. It accepts a certain opportunity of contamination of the sample of block maxima by extremes from convective windstorm events and a certain opportunity of incompleteness ~~since~~from extratropical cyclones ~~can also be observed~~ outside our season definition ~~(Deutsche Rück, 2020).~~. The term winter storm is only based on the high frequency of extratropical cyclones during the winter. The seasonal maximum is also the annual maximum of this peril.

The maxima per half season (bisected by the turn of the year) are analyzed to double the sample size and to increase estimation precision. The appropriateness of this sampling is discussed in Section 5.1. We considered records of wind stations in Germany of DWD (2020; *FX_MN003*, a daily maximum of wind peaks [m/s], usually wind ~~guest~~gust speed) that include minimum record completeness of 90% for analyzed storms, at least 90% completeness for the entire observation period, and minimum 55% completeness per half season. Therefore, we only consider 141 of 338 DWD wind stations (Supplementary data, Table 3). We think this is a good balance between large sample size and high level of record completeness.

The intensity field per event is represented by the maximum wind gust for the corresponding time window of the event at each considered wind station. The local RP per event is computed by a hazard model per wind station. This is an implicit part of the estimated extreme value distribution per station, as explained in Section 5.1. The resulting CRPs per event and corresponding statistical tests are presented in the following Section 4.2. We have considered two weightings per station, capital, and area. Both are computed per wind station by assigning the grid cells with capital data of the Global Assessment Report (GAR data; UNISDR, 2015) via the smallest distance to a wind station. We also use this capital data to spatially distribute our assumed total insured sum 15.23 Trillion € for property exposure (residential building, content, commercial, industrial, agriculture, and business interruption) in Germany in 2018. This is based on Waisman's (2015) assumption for property insurance in Germany and is scaled to exposure year 2018 under consideration of inflation in the building industry (Statistisches Bundsamt, 2020) and increasing building stock according to the German insurance union (GDV, 2020). It is confirmed by the assumptions of the Perils AG (2021~~),~~); however, their data product is not public. We also used loss data of the GDV (2019) for property insurance, when we fitted the vulnerability parameters for the NatCat model. These event loss data of 16 ~~storm~~storms during a period of 17 years are already scaled by GDV to exposure year 2018.

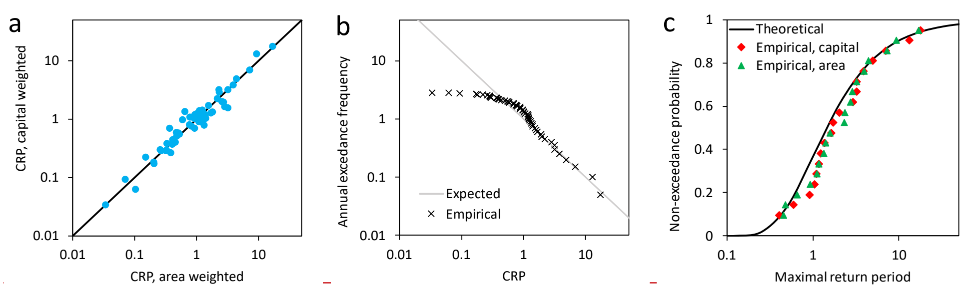
The spatial characteristic is analyzed in Section 4.3 according to ~~the aspects of~~ Section 2.4~~ with focus~~, focusing on the question if there is max-stability or not in spatial dependence and characteristic. Finally, we present the estimated~~ the~~ risk curve for the portfolio of the German insurance market in Section 4.4 including a comparison with previous estimates. Details of the vulnerability model are documented in Section 5.2. The concrete numerical steps, the applied methods to quantify the standard error of estimates, and the consideration of the results from vendor models are explained in ~~Section~~Sections 5.3, 5.4, and 5.5, respectively.

4.2 The CRP of past events and validation

As announced, we have computed the CRP according to (20) with the wind gust peaks listed in the Supplementary data, Table 2, and local hazard models according to (30). Our local hazard models are discussed in Section 5.1 and parameters are presented in the Supplementary data, Table 4. An example for a complete CRP calculation is also provided therein (Table 8). We have considered two weightings for the CRP, a simple area weighting and a capital weighting (Supplementary data, Table 3). In Figure 5a, we compare the estimates which do not differ so much; the approach is robust in the example. The most significant winter storm of the observation period is Kyrill that occurred in 2007. It has CRPs of 16.97±1.75 and 17.64±1.81 years (area and capital). Both are around the middle of the estimated range of 15 to 20 years ~~by~~ (Donat et al~~. (~~., 2011). Further estimates are listed in Supplementary data, Table 1.

In Figure 5b, the results are validated according to Section 3.2. The empirical exceedance frequency matches well with the theoretical one for $T_c \geq 1.65$. Small CRPs are affected by the incompleteness of our record list. In the medium range, the differences between the model and empiricism are not statistically significant. In detail, we observe 27 storms with $T_c \geq 1$ within 20 years; 20 storms were expected ~~were 20~~. According to the Poisson distribution, the probability of 27 exceedances or more is 7.8%. A two-sided test with $\alpha = 5\%$ would reject the model if this exceedance probability ~~would be~~were 2.5% or smaller.

The seasonal/annual maxima of CRP must follow a ~~uniform~~unit Fréchet distribution ($\alpha = 1$ in (3)) according to (16). We plot this and the empirical distribution in Figure 5c. The Kolmogorov-Smirnov (KS) test (Stephens, 1986, Section 4.4) for the fully specified distribution model ~~accepts~~does not reject our model at the very high significance level of 25% for the capital-weighted variant. Usually, only level 5% is considered. This result should not be affected seriously by the absence of one (probably the smallest) maximum due to incompleteness ~~issue~~issues. In summary, we state that the CRP offers a stable, testable, and robust method, to stochastically quantify winter storms over Germany.



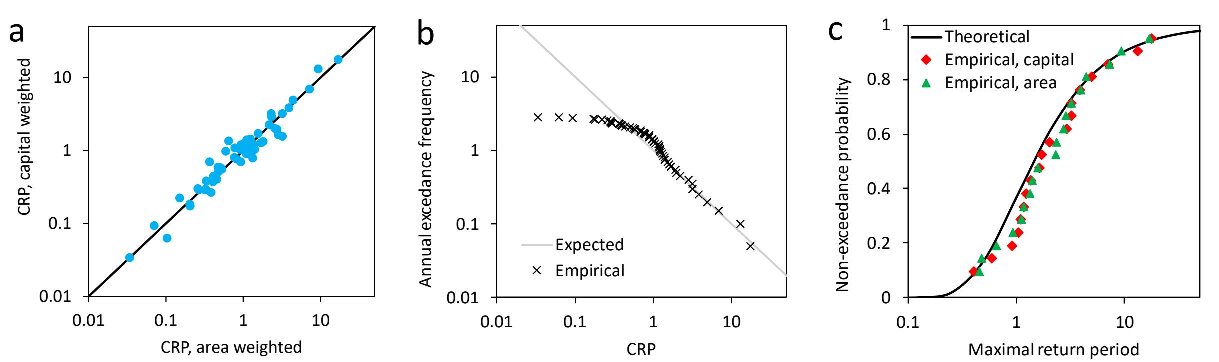


Figure 5: Results of the analysis: a) comparison of area and capital weighted CRPs, b) comparison of theoretical and observed exceedance frequency of capital weighted CRP, c) test of distribution of seasonal maxima of CRP.

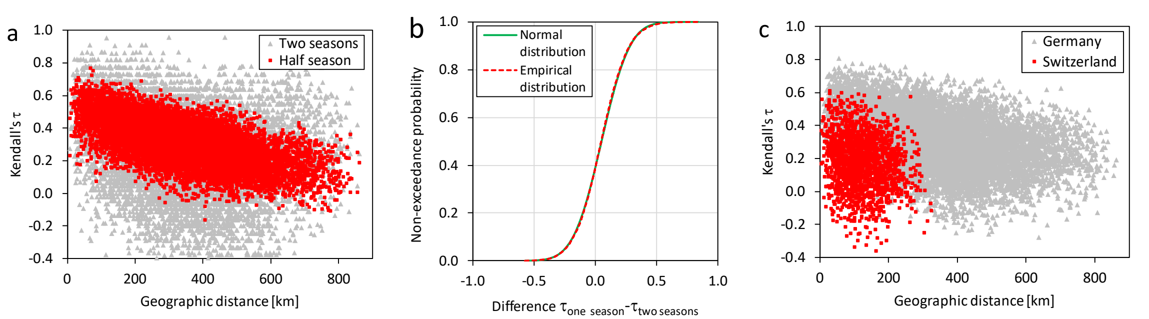
4.3 Spatial characteristic and dependence

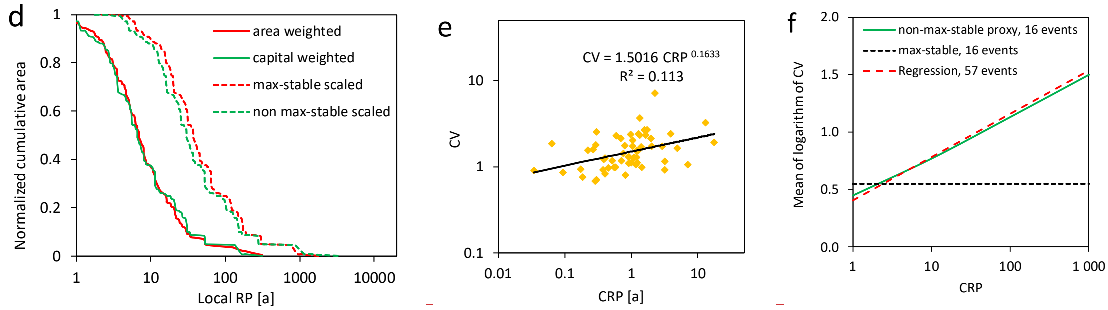
As discussed in Section 2.4, the spatial characteristic is an important aspect from a stochastic perspective. Therefore, we have analyzed the relation between distance and dependence measure. ~~Here~~ We have applied Kendall's τ (Kendall, 1938; Mari and

Kotz, 2001, Section 6.2.6) and show the dependence between the maxima of the half~~-~~ of a season ~~maxima~~ and ~~two season~~ maxima of two seasons for 9,870 pairs of stations in Figure 6a. Since the sample ~~sizes~~size is relatively small, the spreading is strong~~ that~~; it is caused by estimation error. ~~Furthermore, the differences between the estimates for one hazard season maxima and two hazard seasons maxima are almost perfectly normal distributed and should be centered to 0 in case of max-stability (CDF in Figure 6 b). This does not apply with expectation 0.051 and standard deviation 0.182. According to a normally distributed confidence range for the estimated expectation with standard deviation 0.002, the probability, that the actual value is <0, is smaller than 0.00. This is in line with the differences of the non-max-stable example in Figure 2 b.~~ Furthermore, the differences between the estimates of Kendall's τ for maxima of one and two hazard seasons are almost perfectly normally distributed (CDF in Figure 6b) and should be centered to 0 in case of max stability (Figure 2b). This does not apply with sample mean 0.051 and standard deviation 0.182. The corresponding normally distributed confidence range for the expectation has a standard deviation 0.002 and a probability of 0.00 that the actual expectation is 0 or smaller.

For completeness, we compare the current estimates of Kendall's τ with ~~these~~those for Switzerland from Raschke et al. (2011) in Figure 6c. The spatial dependence is higher for Germany. A reason might be differences in the topology.

We have also computed the area functions and show examples in Figure 6d for winter storm Kyrill. The different weightings result in similar area functions. ~~The CRP and CV of all events are plotted in~~ Figure 6e~~.~~ plots the CRP and CV of all events. The regression analysis ~~result in~~reveals the statistical dependence between CRP and CV. For the linearized regression function, the p-value is 0.002 (t-test, Fahrmeir et al., 2013, Section 3.3). ~~Because~~Due to two statistical indications of non-max-stability, we develop ~~for every event with loss information~~ a local scaling that ~~consider~~considers the global scaling factor and the ratio between local RP and CRP~~.~~ for every event with loss information. In this way, we could reproduce the observed pattern (Figure 6f). Details of this workaround are presented in the Supplementary data, Section 7. The differences between the scaling variants (max-stable or not) for storm Kyrill do not seem to be strong (Figure 6d).





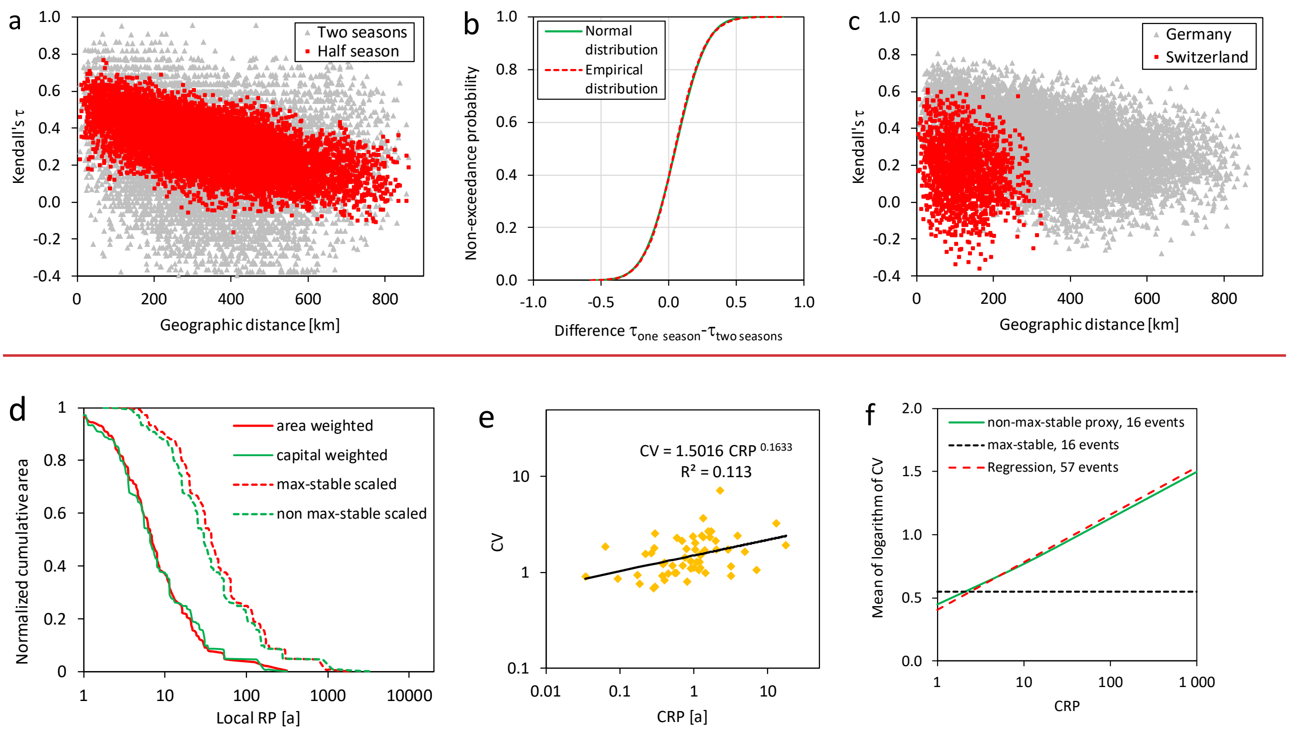


Figure 6: Spatial characteristics of winter storms over Germany: a) estimated Kendall's τ versus distance, b) differences between Kendall's τ for different block sizes, c) estimated Kendall's τ versus distance for seasonal maxima in Germany and Switzerland (Raschke et al., 2011), d) area functions for the storm Kyrill with scaling to CRP 100 years, e) relation CV to CRP of capital weighted area functions, and f) approximation of this relation by special stochastic scaling.

4.4 The risk estimates

Before we estimated risk curves according to the approach of Section 3.4, we must estimate a vulnerability function (31~~)~~), which determines the local loss ratio L_R in the event loss aggregation (26). First, we fit the scaling parameter on the event loss data of the General Association of German ~~Insurer~~Insurers (GDV, 2019) for 16 historical events from 2002 to 2018, as plotted in Figure 7a. The details of the vulnerability function and its parameter fit are explained in Section 5.2. Then, we use the vulnerability function in the three variants of risk curve estimates of Section 3.4~~—~~—averaging ~~of~~ event loss, ~~ELRP~~CRP, or its reciprocal, the exceedance frequency. Details of the numerical procedure are explained in Section 5.3, which corresponds with the scheme in Figure 4a.

In Figure 7b, the three estimated risk curves according to the three estimators in (28) are presented for max-stable scaling and differ less to each other~~ what~~, which indicates the robustness of our approach. The empiricism is presented by the historical event losses and their empirical RP (observation period 17 years of GDV loss data) and capital-weighted CRP. In addition, we present the range of two standard errors of the estimates of loss averaging which implies the simplest numerical procedure. Details of uncertainty quantification are explained in Section 5.4.

The differences between max-stable and non-max-stable scaling in the risk estimates are demonstrated in Figure 7c. For smaller RP, no significant difference can be stated in contrast to higher RP. This corresponds with the differences between the CV in relation to the CRP for max-stable and non-max-stable ~~case~~cases in Figure 6f. These are also higher for higher CRP.

We also compare our results with previous estimates in Figure 7c. For this purpose, we must scale these to provide comparability as ~~good~~well as possible. The relative risk curve of Donat et al. (2011) is scaled simply by our ~~TSI~~ assumption for the total sum insured (TSI) for the exposure year 2018. The vendor models of Waisman (2015) are scaled by the average of ratios between modelled and observed event losses from Storm Kyrill since a scaling via TSI was not possible (uncertain market share and split between residential, commercial, and industry exposure). The result of the standard model of European Union (EU) regulations (European Commission, 2014), also known as Solvency II requirements, is also based on our TSI assumption, split into the Cresta zones by the GAR data. The Cresta zones (www.cresta.org) are an international standard in insurance industry and corresponds to the two-digit postcode zones in Germany.

The risk estimate of Donat et al. (2011) is based on a combination of frequency estimation and event loss distribution by the generalized Pareto distribution, which is fitted on a sample of modelled event losses for historical storms. The corresponding risk curve differs very much from other estimates and ~~obviously overestimate~~overestimates the risk of winter storms over Germany. The standard model of EU only estimates the maximum event loss for RP 200 years~~,~~; the estimated event loss is very high. The vendor models vary but have a similar course as our risk curves. The non-max-stable scaling is in the lower range of the vendor models, whereas the unrealistic max-stable scaling is more in the middle. The concrete names of the vendors can be found in Waisman's (2015) publication. The reader should be aware that the vendors might have updated their winter storm model for Germany in the meantime.

The major result of Section 5 is the ~~successfully~~successful demonstration that the CRP can be applied to estimate reasonable risk curves under controlled stochastic conditions. In addition, we have ~~also~~ discovered the ~~high~~strong influence of the underlying dependence model (max-stable or not) and corresponding spatial ~~characteristic~~characteristics to loss estimates for higher ELRP.

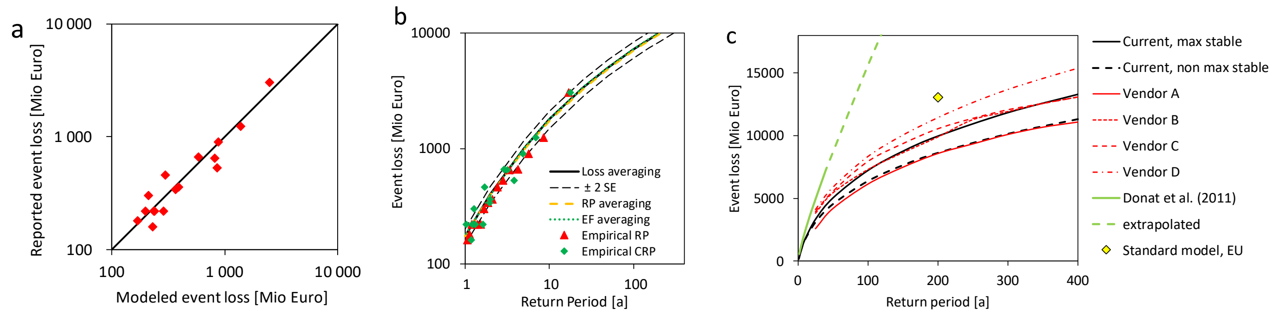


Figure 7: Estimates for insured losses from winter storms in Germany: a) reported versus modelled event losses, b) current risk curves and observations, c) influence max-stable and non-max-stable scaling and comparison to scaled, previous estimates (Donat et al., 2011; Waisman, 2015, European Commission, 2014).

5 Technical details of the application example

5.1 Modelling and estimation of local hazard

As aforementioned, the maximum wind gusts of half seasons of winter storms (extratropical cyclones~~), the~~) block maxima, have been analyzed. Therein, the generalized extreme value distribution (Beirlant~~,~~ et al., 2004, (5.1)) is applied:

$$\cancel{G(x) = \begin{cases} exp\left(-exp\left(\frac{x-\mu}{\sigma}\right)\right), if\ \gamma = 0 \\ exp\left(-\left(1+\gamma\frac{x-\mu}{\sigma}\right)^{-1/\gamma}\right), if\ \gamma \neq 0, with\ x > \mu - \frac{\sigma}{\gamma}\ if\ \gamma > 0\ and\ x < \mu - \frac{\sigma}{\gamma}\ if\ \gamma < 0 \end{cases}} G(y) =$$

$$\begin{cases} exp\left(-exp\left(\frac{y-\mu}{\sigma}\right)\right), if\ \gamma = 0 \\ exp\left(-\left(1+\gamma\frac{y-\mu}{\sigma}\right)^{-1/\gamma}\right), if\ \gamma \neq 0, with\ y > \mu - \frac{\sigma}{\gamma}\ if\ \gamma > 0, and\ y < \mu - \frac{\sigma}{\gamma}\ if\ \gamma < 0 \end{cases}. \quad (29)$$

As discussed below, the Gumbel distribution (Gumbel, 1935, 1941), as a special case in (29) with extreme value $\gamma = 0$, is an appropriate model. The scale parameter is σ, and the location parameter is μ. The local hazard function (13,14) can be derived directly from the estimated variant of (29) according to the link between extreme value distribution and exceedance frequency (16~~);~~) (the accent symbolizes the point estimation~~)~~):

$$\cancel{\hat{T}(x)}\hat{T}_y(y) = 1/\cancel{\hat{\Lambda}(x)}\hat{\Lambda}_y(y) = exp\left(\cancel{\frac{x-\hat{\mu}}{\hat{\sigma}_{cor}}}\right)\left(\frac{y-\hat{\mu}}{\hat{\sigma}_{cor}}\right). \quad (30)$$

We apply the maximum likelihood (ML) method for the parameter estimation (Clarke, 1973, Coles, 2001, Section 2.6.3). The wind records' incompleteness ~~of wind records~~ per half season have been considered in the ML estimates by ~~a modification of~~modifying the procedure as explained in ~~the~~ Supplementary data, Section 5. A Monte Carlo simulation confirms the good performance of our modification. The biased estimate of σ for our sample size $n = 40$ was also detected ~~which~~, where we considered $\hat{\sigma}_{cor} = \hat{\sigma}/0.98$ as corrected estimation. Landwehr et al. (1979) have already stated such bias. ~~A further bias was discovered~~In addition, the ~~EF~~exceedance frequency is well estimated by (30) in contrast to the RP $\hat{T}$~~, this~~. The latter is strongly biased. We also corrected this as documented in ~~the~~ Supplementary data, Section 6. The analyzed half-season maxima, record completeness, and parameter estimates are listed in Supplementary data, ~~Table~~Tables 4, 5 ~~ad~~, and 6.

We have validated the sampling of block maxima per half season. The opportunity of correlation between the first and second half-season maxima has been tested~~,~~ for significant level $\alpha = 5$~~%~~%, and hypothesis H_0: uncorrelated samples. Around 6% ~~fails~~fail the test with ~~Fisher's~~Fisher’s z-transformation (Upton and Cook, 2008). This corresponds to the error of the first kind ~~and is interpreted as~~ (Type I error, e.g., Lindsey, 1996, Section 7.2.3), the falsely rejected correct models. Therefore, we interpret the correlation ~~not being significant.~~as statistical insignificant. Similarly, the Kolmogorov-Smirnov homogeneity test rejects 4% of the sample pairs for period September to December and January to April (first half ~~to second~~ season to subsequent second half season) for a significant level of 5%. ~~This 5%~~There are ~~the expected share of falsely rejected correct models – the first kind of error (type I error, e.g. Lindsey, 1996, Section 7.2.3).~~no significant differences between the samples.

To optimize the intensity measure of the hazard model, we have considered the wind speed with power 1, 1.5, and 2 as the local event intensity in a first fit of the Gumbel distribution by the maximum likelihood method. According to these, power 1.5 offers the best fit of wind gust data to the Gumbel distribution. Such wind measure variants were already suggested by Cook (1986) and Harris (1996).

We do not apply the generalized extreme value distribution in (~~24~~29) with extreme value index $\gamma \neq 0$ but the Gumbel case with $\gamma = 0$ for the following reasons. ~~At first, an extensive physical explanation would be required if some~~ Different stochastic regimes $\gamma<0$ and $\gamma\geq0$ for different wind stations ~~are concerned by a~~imply fundamental physical differences: finite and infinite upper ~~bound for $\gamma < 0$ and other stations not with $\gamma \geq 0$ according to (29). Why should be local~~ bounds of wind ~~hazard short tailed for some~~ speed. Such fundamental differences between different wind stations ~~and heavy tailed for others?~~would need reasonable explanations (especially for very low bounds versus infinite bounds). River discharges at different gauging stations could imply such physical differences since there are variants with laminar and turbulent stream ~~(catchword Reynolds number)~~ or very different retention/storage capacities of catchment areas (e.g., Salazar et al~~.~~., 2012). ~~Such~~Similar significant physical differences do not exist for wind stations ~~which~~that are placed and operated under consideration of rules of meteorology (World Meteorological Organization, 2008, Section 5.8.3) to provide homogeneous roughness ~~condition due to generate comparable data.~~conditions. Besides, we also found several statistical indications for our modelling. Information criteria AIC and BIC (Lindsey, 1996; here over all stations) indicate that the Gumbel distribution is the better model than the variant with a higher degree of freedom. Furthermore, the share of rejected Gumbel distributions of the Goodness-of-fit test (Stephens, 1986, Section 4.10) is with 6% around the defined significance level of 5% (the error of 1st kind~~—~~—falsely rejected correct models). We have also estimated γ for each station and got a sample of point estimates. The sample mean ~~of~~ is ~~with~~ 0.002, very close to $\gamma = 0$ ~~which~~; this confirms our assumption. Moreover, the sample variance is 0.018 which is around the same what as we get for a large sample of estimates $\hat{\gamma}$ for samples of Monte Carlo simulated and Gumbel distributed random variables ($n = 40$). All statistics validate the Gumbel distribution.

To provide reproducible results, we also present a computational example for the CRP in Table 1 with reference to all needed equations and information. The entire calculation for Storm Kyrill is presented in the Supplementary data, Table 8.

Table 1: Example for the computation of a CRP - Station #303 (Baruth) for Storm Kyrill

#	Step	Equation	Supplementary data
1	Event Definition: Kyrill, time window 18-19.01.2007	-	Tab. 1 Considered storm events
2	Wind speed maximum: Station #303 (Baruth) 34.7 m/s	-	Tab. 2 Event maxima
3	Transformation to local intensity measure – the 1.5 power of wind speed: $34.7^{1.5}m^{1.5}/s^{1.5}=204.4m^{1.5}/s^{1.5}$	-	-
4	Computation of local RP for the unit period: 58.07 half season (half of a year)	(30) with $\hat{\sigma}_{cor}$	Tab. 4 Gumbel parameters
5	Bias correction of RP for the unit period: 46.58 half season (half of a year)	(16) in Supplementary	-
6	RP for unit period one year: 46.58/2= 23.29 season (year)	-	-
7	Capital weighting: 23.29*0.0008988= 0.021	(20)	Tab. 1 Considered storm events
8	Aggregation over all stations: 17.65 year	(20)	Tab. 8 CRP example Kyrill
9	Final capital weighted CRP for event Kyrill - Normalization by the sum of weighting: 17.65/1.0= 17.65 year	(20)	Tab. 8 CRP example Kyrill

5.2 Modelling and estimation of vulnerability

To quantify the loss ratio L_R at location (wind station) j and event i in the loss aggregation (26), we use the approach of Klawa and Ulbrich (2003) for Germany~~. The difference to the origin is not significant.~~ with a certain modification. The event intensity ~~x~~ is the maximum wind gust speed v. $v_{98\%}$ is the upper 2% percentile from the empirical distribution of all local wind records. The relation with vulnerability parameter a_L is

$$L_{R,i,j} = a_L\, Max\{0,(v_{i,j}-v_{98\%})\}^3 \cancel{;}. \qquad (31)$$

Donat et al. (2011) have used a similar formulation but with an additional location parameter. This is discarded here since the loss ratio must be $L_R = 0$ for local wind speed $v < v_{98\%}$. This is also a reason~~,~~ why a simple regression analysis (Fahrmeir et al., 2013) is not applied to estimate a_L. We formulate and use the estimator

$$\hat{a}_L = \frac{1}{n}\cancel{\sum_{i=1}^{k}\frac{\sum_{j=1}^{n} E_{j,i}\, Max\{0,(x-x_{98\%})\}^3}{L_{E\,reported,i}}}\sum_{i=1}^{k}\frac{\sum_{j=1}^{n} E_{j,i}\, Max\{0,(v-v_{98\%})\}^3}{L_{E\,reported,i}}, \qquad (32)$$

with k historical storms, corresponding reported event losses L_E, n wind stations, and local exposure value $E_{j,i}$ being assigned to the wind station. $E_{j,i}$ be fixed for every station j if there were wind records for every storm i at each station j. However, the wind records are incomplete and the assumed TSI must be split and assigned to the stations a bit differently for some storms. The exposure share of the remaining stations is simply adjusted so that the sum over all stations remains the TSI.

Our suggested estimator (32) has the advantage that it is less affected by the issue of incomplete data (smaller events with smaller losses are not listed in the data) than the ratio of sums over all events, and the corresponding standard error can be quantified (as for the estimation of an expectation). The current point estimate is $\hat{a}_L$ =9.59E-8±5.97E-9.

An example of our vulnerability function (with the average of $v_{98\%}$ over the wind stations) is depicted in Figure 8 and compared with previous estimates for Germany. It is in the range of previous models. Differences might be caused by different geographical resolutions of corresponding loss and exposure data. A power parameter of 2 in (31) might also be reasonable since the wind load of building design codes (European Union, 2005, Eurocode 1) is proportional to the squared wind speed. The influence of deductibles (Munich Re, 2002) per insured object is not explicitly considered but smoothed in our approach.

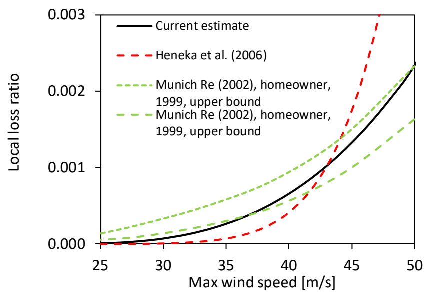


Figure 8: Vulnerability functions, current estimate with the average of local parameters, previous estimates by Heneka and Ruck (2008) and Munich Re (2002) for residential buildings.

5.3 Numerical procedure of scaling

Here, we briefly explain ~~here~~ the numerical procedure to calculate a risk curve via averaging the event loss. For any supporting point of a risk curve during an event loss averaging, the ELRP T_E is defined and ~~determine~~determines the scaled CRP T_{cs} for all historical events. For each historical event, the scaling factor is $S = T_E / T_c$ according to (24) and is applied in (27) together with local hazard function (30) and its invers. The hazard parameters are listed in the Supplementary data, Table 4. For the scaled local intensity, the local loss ratio $L_{R,i}$ is computed with vulnerability function (31). The corresponding parameter $v_{98\%}$ is also listed in the Supplementary data, Table 2. The local loss ratio $L_{R,i}$ and the local exposure value E_i are used in (26) to compute the event loss. The considered values of E_i per event are listed in Supplementary data, Table 7. The incompleteness of wind observation is considered therein. Finally, for the supporting point, the modelled event losses of all scaled historical events are averaged according to (28).

The historical events are also scaled for a defined event loss and the corresponding scaled CRP is averaged. However, the 'Goal ~~Seek~~Seek' function in MS Excel is applied to find the correct scaled CRP T_{sc} and corresponding scaling factor S. For the averaging of the exceedance frequency, the reciprocal of T_{sc} is averaged. All these apply to max-stable scaling. For the non-max-stable scaling, the scaling factor S is adjusted to a local variant according to the description in ~~the~~ Supplementary data, Section 7. Therein, the factor S is adjusted for each station and depends on the relation of local RP to CRP of the historical event. This adjustment is made for each historical storm individually.

5.4 Error propagation and uncertainties

The uncertainty of the local hazard models influences the accuracy of the CRP since the CRP is an average of estimates of local RP. The issue is that there is a certain correlation between the estimated hazard parameters of neighboring wind stations. We consider this by application of the Jack-knife method (Effron and Stein, 1986). According to these, the root of mean squared error (~~MSE~~RMSE, which is the standard error if the estimate is bias free as we assume here) of the original estimated parameter $\hat{\theta}$ is (accents symbolize estimations)

$$\cancel{MSE}RMSE(\hat{\theta}) = \sqrt{\frac{n-1}{n}\sum_{i=1}^{n}(\hat{\theta}_{-i} - \hat{\theta})^2}, \tag{33}$$

with the estimates $\hat{\theta}_{-i}$ for the Jack-knife sample i of observations, being the original sample but without one of the observations/realizations. Therefore, it is ~~also called~~ also called leave-one-out method. The estimator (33) implies a parameter sample of $\hat{\theta}_{-i}$ of size n, with one estimated parameter or parameter vector for each Jack-knife sample i of observations.

To consider any correlation in the error propagation of CRP estimate, the maximum of the same half-season i is left out synchronously when the parameter sample is computed for each wind station. Without changing the order in the parameter sample of each wind station, the CRP $\hat{T}_{c-i}$ of the concrete historical event is computed with the hazard parameters $\hat{\theta}_{-i}$ of each station. Finally, for this storm, the standard error of point estimate $\hat{T}_c$ is computed according to (33).

We use the same approach to consider the error propagation from local hazard models to the risk estimate for the max-stable case in Section 4.4. But the finally estimated parameter $\hat{\theta}$ in (33) is the averaged event loss $\hat{L}_E(T_E)$ for scaled CRP. This only covers a part of uncertainties in risk estimate. We consider two further sources of uncertainty and assume that they influence the risk estimate independently to each other. The uncertainty of loss averaging is the same as during an estimation of an expectation from a sample mean and is determined by sample variance and sample size (number of scaled events). The propagation of the uncertainty of the vulnerability parameter is computed via the Delta method (Coles 2001, Section 2.6.4). The aggregated standard error is the square root of the sum of squared errors. This implies a simple variance aggregation according to the convolution of independent random variables (Upton and Cook, 2008).

The computed standard errors in Figure 7b are in the range of 7.5 to 8.5% of estimated event loss per defined ELRP. The shares of uncertainty components on the error variance (squared SE) of our risk estimates depend on the RP. On average for our supporting point, these are 15% for the limited sample of scaled historical events, 24% for the uncertainty of local hazard parameters, and 61% by the vulnerability ~~model's~~model’s parameter. Unfortunately, we do not know a published error estimation for a vendor model for ~~risk from~~ winter storm ~~over~~risk in Germany ~~and~~. Therefore, we can only compare our estimates with ~~these of Donat’s~~Donat et al~~.~~.`s (2011). Their confidence range ~~indicate~~indicates a smaller precision than ours.

5.5 RP of vendor’s risk estimate

We have compared our results with vendor models in Section 4.4. These have estimated the risk curve for the maximum event loss within a year. This is a random variable, and their ~~pseudo~~ RP is the reciprocal of the exceedance probability and can never be smaller than 1. ~~Under the assumption of a Poisson process,~~ We transform the ~~pseudo~~ RP of annual maxima to ~~an actual and event related one with the relation between EF and CDF in (13,16) With increasing~~the RP of event loss~~,~~ according to the relations in Section 3.1 and the ~~difference between its pseudo RP and the actual ELRP converges to 0.~~explanations by Coles (2002, p. 249, with y_p as exceedance frequency of events). The relative differences between the RPs are around 5% for ELRP 10 years and 0.5% for 100 years.

6 Conclusion, discussion, and outlook

6.1 General

In the beginning, we asked the questions about the RP of a hazard event in a region, the corresponding NatCat risk, and necessary conditions for a reasonable NatCat modelling. To answer our questions, we have mathematically derived the CRP of a NatCat generating hazard event from previous concepts of extreme value theory, the pseudo polar coordinates (17). This implies the important fact that the average of the RPs of random point events remains a RP with exceedance frequency (8,15). Furthermore, we extended Schlatter's 1st theorem for max-stable random fields to non-max-stable spatial dependence and characteristic. We have also considered the normalized variant of the area function of all local RP of the hazard event in a region with parameters CRP and CV. The absence of max-stability in the spatial dependence results in a correlation between CRP and CV, which is a further indicator for non-max-stability beside changes of measures for spatial dependence by changed block size (e.g., annual maxima versus two years maxima).

The derived CRP is a universal, simple, plausible, and testable stochastic measure for a hazard and NatCat event. The weighting of local RP in the computation of the CRP can be used to compensate an inhomogeneous distribution of corresponding measuring stations if the physical-geographical hazard component of a NatCat, the field of local event intensity ~~field~~, is of interest. However, the concentration of human values in the geographical space could also be considered in the weighting to ~~get~~obtain a higher association of the CRP with the ELRP of a risk curve. This link implies the conditional expectation (18) under the assumption of max-stable association between CRP and ELRP, and offers ~~the~~a new opportunity to estimate risk curves, the bijective function event loses to ELRP, via a stochastic scaling of historical intensity fields and averaging of corresponding risk ~~parameters~~parameter. The averaged ~~parameters~~parameter can be the scaled CRP for a defined event loss, corresponding exceedance frequency, or the event loss for a defined/scaled CRP.

The differences between the three estimators are small in our application example, insured losses from winter storm over Germany. In contrast ~~to this~~, the influence of the stochastic assumptions regarding spatial dependence and characteristic (max-stable or not) is significant in the range of higher ELRP. This highlights the importance of realistic consideration of spatial dependence and characteristic of the hazard in a NatCat model. Besides, our risk curves for Germany have a similar course as those derived by vendors (Figure 7d). The risk assumption by EU for Germany with ~~RP~~ELRP 200 years is significantly higher than ours. The estimate by Donat et al. (2011) differs significantly and seems to be implausible for higher ~~RP~~ELRP. A reason might be their statistical modelling by the generalized Pareto distribution as already applied for wind losses by Pfeifer (2001).

The tapered Pareto distribution (Schoenberg and Patel, 2012), also called tempered Pareto distribution (Albrecher et al., 2021), or a similar approach (Raschke, 2020) provide more appropriate proxies for our risk ~~curve's~~curve’s tail.

According to our results, necessary conditions for~~ an~~ appropriate NatCat modelling are the realistic consideration of local hazard and their spatial dependence (max-stable or not?). Correspondingly, the spatial characteristic of NatCat events, described here by relation CRP to CV, must be reproduced. In addition, the CRPs of a simulated set of hazard events in a NatCat model should have an empirical exceedance frequency that follows the theory (15). ~~That~~Finally, the standard error of an estimate should be quantified, the sampling should be appropriate~~, and~~; overfitting ~~should be avoided (catchwords (~~over parametrization and parsimony~~),~~) should be avoided. This principle applies to all scientific models with a statistical component (e.g., Lindsey, 1996).

The advantage of our approach over vendor models is the simplicity and clarity about the stochastic assumptions. The numerical simulations for models in the insurance industry (Mitchell-Wallace et al., 2017, Section 1.8) and science (e.g., Della-Marta et al., 2010) need tens of thousands of simulated storms with unpublished or even unknown (implicit) stochastic assumptions. We have only scaled 16 event fields of historical storms with controlled stochastics and could even quantify the standard error.

6.2 Requirements of the new approaches

Our approach to CRP is based on two assumptions. At first, the local and global events occur as a Poisson process. This is a common assumption or approximation in applied extreme value statistics ~~and its application~~ (Coles, 2001, Chapter 7~~)~~), and the corresponding Poisson distribution of the number of events can be statistically tested (Stephens, 1986; Section 4.17). Moreover, the verified clustering (overdispersion) of winter storms over Germany (Karremann et al., 2014) is statistically not relevant for higher RP (Raschke, 2015). With increasing RP, the number of occurring winter storms converges to a Poisson distribution. Clustering is also influenced by the event definition, which is not the topic here (~~catchword~~keyword declustering; Coles, 2001). We also point out that the assumed Poisson process needs not be homogenous during a defined unit period (year, hazard season, or half-season).

The second prerequisite is robust knowledge about the local RP by a local hazard curve. Unfortunately, there are no appropriate and comprehensive models for the local hazard of every peril and region~~.~~, for example, hail in Europe~~,~~; we only know local hazard curves for Switzerland by Stucki et al. (2007~~)~~), and these were roughly estimated.~~ For flood hazard,~~ There are public hazard maps of flooding areas for defined RP; corresponding local hazard curves are rarer.

Furthermore, existing models for local hazard are partly questionable according to our discussion about local modelling of wind hazard from winter storms in Section 5.1. We have assumed a Gumbel case of the generalized extreme value distribution for local block maxima with extreme value index $\gamma = 0$ for ~~physical reasons and have validated this by~~ several statistical indicators and physical consistency. Youngman’s and Stephenson’s (2016) modelling of winter storms over Europe implies an extreme value index $\gamma < 0$ for the region of Germany, which means a short tail with a finite upper bound. ~~Unfortunately,~~ They have not depicted the spatial distribution of the corresponding finite upper bounds and does not provide a physical explanation for the spatial varying upper physical limit of wind speed maxima. The plausibility of such physical details in a NatCat model should be shown and discussed.

6.3 Opportunities for future research

Since the current model for the local hazard of winter storms over Germany results in considerable uncertainty, it should be improved in the future. This could be realized by a kind of regionalization of the hazard as already known in flood research (Merz and Blöschl, 2003; Hailegeorgis and Alfredsen, 2017) or ~~by~~ a spatial model ~~as suggested by (~~(Youngman and Stephenson ~~(~~, 2016). Besides, more wind stations could be considered in the analysis with better consideration of incompleteness in the records. An extension of the observation period is conceivable if homogeneity of records and sampling is ensured. A more sophisticated approach might be used to discriminate the extremes of winter storms from other windstorm perils at the level of wind station records. The POT methods (Coles, 2001, Section 4.3; Beirlant et ~~a~~al., 2004, Section 5.3) could then be used in the analysis even though the spatial sampling is complicated as stated in the introduction.

Further opportunities for improvements in the winter storm modelling are conceivable. The event field might be ~~more detailed~~ reproduced/interpolated in more detail, as done by Jung and Schindler (2019). They have considered the roughness of land cover at a regional scale besides further attributes. However, they did not consider the local roughness of immediate surroundings, as ~~discussed by~~ Wichura (2009) discussed for a wind station.

Besides, our approach could be used for further hazards such as earthquake, hail, or river flood. The reasonable weighing would not be trivial for river flood. May be, the local expected annual flood loss would be a reasonable weighting if the final goal is a risk estimate for a region. The numerical handling of the case that an event does not occur everywhere in the researched region but ~~has also~~ local RP $T = 0$ must be discussed for some perils, such as hail or river flood.

We also see research opportunities for the community of mathematical statistics, especially extreme value statistics. Does (18) for conditional expected RP also apply to the non-max-stable case? A deeper theoretical understanding of non-max-stable random fields with max stable margins is of great interest from practitioners' perspectives. ~~A~~ Research about the link between normalized area functions (expectation versus CV) and spatial dependence could increase understanding of natural hazard and risk~~.~~, and our construction for the non-max-stable scaling is just a workaround to illustrate the consequences of dependence characteristics; for risk models in practice, a transparent stochastic construction is needed. Furthermore, estimation methods could be extended and examined, such as the bias in estimates of local RP.

7 Code and data availability

A special code was not generated or used. M.S. Excel had carried out our computations ~~had been simply caried out by MS Excel.~~. The wind data were downloaded from the server of the German meteorological service (Deutscher Wetter Dienst, 2020~~).~~), and the exposure data were provided by UNISDR (2015). The loss data are part of the ~~downloaded report of~~ General

Association of German ~~Insurer~~Insurers’ (Gesamtverband Deutscher Versicherer, 2019~~).~~) report. The ~~here~~ considered wind stations and storms ~~considered in the paper~~ are listed in Supplementary data.

8 Author’s contribution

The author has derived the theory, carried out the analysis, and wrote the paper. External help has been used regarding proofreading.

9 Competing interest

The author declares that he has no conflict of interest.

10 Acknowledgement

The author thanks the reviewers for helpful comments.

References

Albrecher, H., Araujo-Acuna, J., & Beirlant, J. Tempered Pareto-type modelling using Weibull distributions. *ASTIN Bulletin*, 51(2), 509-538. doi:10.1017/asb.2020.43, 2021.

Asadi P., Engelke S. and Davison A.C. Extremes on river networks. *Ann. Appl. Stat*. **9**, 2023~~-~~–2050, 2015.

Beirlant, J., Goegebeur, Y., Teugels, J. and Segers, J. *Statistics of Extremes – Theory and Application*. Book Series: Wiley Series in Probability and Statistics, John Wiley & Sons, 2004.

Blanchet, J. & Davison A.C. Spatial Modelling of extreme snow depth. *The Annals of Applied Statistics* **5**, 1699~~-~~–1725, 2011.

Bonazzi, A., Cusack, S., Mitas, C. and Jewson, S. The spatial structure of European wind storms as characterized by bivariate extreme-value Copulas. *Nat. Hazards Earth Syst. Sci*. **12**, 1769~~-~~–1782, 2012.

Bormann, P. and Saul, J. Earthquake Magnitude, in *Encyclopedia of Complexity and Applied Systems Science*, 3, pp. 2473~~-~~–2496, http://gfzpublic.gfz-potsdam.de/pubman/item/escidoc:238827:1/component/escidoc:238826/13221.pdf, 2009.

Clarke, R.T. Mathematical models in hydrology. Irrig. Drain. Pap. 19, Food and Agr. Organ. Of the U.N., Rom. 1973.

Coles, S. *An Introduction to Statistical Modeling of Extreme Values*. Book Series: Springer Series in Statistics, Spinger, 2001.

Cook, N.J. The Designer’s Guide to Wind Loading of Building Structures. Part 1: Background, Damage Survey, Wind Data and Structural Classification. Building Research Establishment, Garston, and Butterworths, London, ~~pp371~~pp. 371, 1986.

Dawkins, L.C. and Stephenson, D.B. Quantification of extremal dependence in spatial natural hazard footprints: independence of windstorm gust speeds and its impact on aggregate losses. *Nat. Hazards Earth Syst. Sci*. **18**, 2933~~-~~–2949, 2018.

De Haan, L. A spectral representation for max-stable processes. *The Annaly of Probability* **12**, 1194~~-~~–1204, 1984.

De Haan, L., and Ferreira, A. Extreme value theory: an introduction. Springer, 2007.

Della-Marta, P., Mathias, H., Frei, C., Liniger, M., Kleinn, J. & Appenzeller, C. The return period of wind storms over Europe. *International Journal of Climatology* **29**, 437~~-~~–459, 2009.

Della-Marta, P.M., Liniger, M. A., Appenzeller, C., Bresch, D. N, Koellner-Heck, P., and Muccione, V. Improved estimates of the European winter windstorm climate and the risk of reinsurance loss using climate model data. *Journal of Applied Meteorology and Climatology* **49**, 2092~~-~~–2120, 2010.

Deutsche Rück, *Sturmdokumentation*, www.deutscherueck.de/downloads/sturmdokumentation/, (download 2020)

Deutscher Wetter Dienst (DWD, German meteorological service), Climate Data Centre (CDC), https://cdc.dwd.de/portal/202007291339/index.html (download Spring 2020).

Dey, D., Jiang, Y:, and Yan, J., Multivariate extreme value analysis. In: *Extreme Value Modeling and Risk Analysis – Methods and Applications*. Ed. D. Dey and J. Yuan, CRC Press, Boca Raton, 2016.

Dombry, C. Extremal shot noises, heavy tails and max-stable random fields. Extremes 15, 129–158, 2012.

Donat, M. G., Pardowitz, T., Leckebusch, G. C., Ulbrich, U. and Burghoff, O. High-resolution refinement of a storm loss model and estimation of return periods of loss-intensive storms over Germany. *Nat. Hazards Earth Syst. Sci*. **11**, 2821~~-~~–2833, 2011.

Engelke, S., Kabluchko, Z, and Schlather, M. An equivalent representation of the Brown–Resnick process, Statistics & Probability Letters 81, 1150~~-~~–1154, 2011.

Efron, B., and Stein, C. The Jackknife Estimate of Variance. *The Annals of Statistics*, **9**(3), S. 586–596, 1981.

European Commission. Valuation and risk-based capital requirements (pillar i), enhanced governance (pillar ii) and increased tranparency (pillar iii), COMMISSION DELEGATED REGULATION (EU) 2015/35 supplementing Directive 2009/138/EC of the European Parliament and of the Council on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II), 2014.

European Union (EU). Eurocode 1: Actions on structures – Part 1-4: General actions – Wind actions. The European Union per Regulation 305/2011, Directive 98/34/EC, Directive 2004/18/EC, 2005.

Fahrmeir, L., Kneib, T., and Lang, S. Regression - Modells, Methods and Applications. Springer, Heidelberg, 2013.

Falk, M., Hüsler, J., and Reiss, R.-D. *Laws of Small Numbers: Extremes and rare Events*. Biskhäuser 3rd ed., Basel, 2011.

Gesamtverband Deutscher Versicherer (GDV, General Association of German Insurer), *Naturgefahrenreport - Serviceteil* (German, www.gdv.de/de/zahlen-und-fakten/publikationen/naturgefahrenreport), 2019.

Gumbel, E.J. Les valeurs extrêmes des distributions statistiques. Annales de ~~l'Institut~~l’Institut Henri Poincaré 5, 115–158, 1935.

Gumbel E.J. The return period of flood flows. The Annals of Mathematical Statistics 12, 163–190, 1941.

Guse, B., Merz, B., Wietzke, L., Ullrich, S., Viglione, A. and Vorogushyn, S. The role of flood wave superposition in the severity of large floods. *Hydrol. Earth Syst. Sci.* **24**, 1633–1648, 2020.

Gutenberg, B., Richter, C. F.: Magnitude and Energy of Earthquakes. Annali di Geofisica, 9: 1–15, 1956.

Harris, R. I. Gumbel re-visted – a new look at extreme value statistics applied to wind speeds. *J. Wind Eng. Ind. Aerodyn.* **59**, 1–22, 1996.

Hailegeorgis, T.T. and Alfredsen, K. Regional flood frequency analysis and prediction in ungauged basins including estimation of major uncertainties for mid-Norway. *Journal of Hydrology: Regional Studies* **9**, 104–126, 2017.

Heneka, P. and Ruck. B. A damage model for assessment of storm damage buildings. *Engineering Structures* **30**, 721–733, 2008.

Jongman, B., Hochrainer-Stigler, S., Feyen, L. et al. Increasing stress on disaster-risk finance due to large floods. *Nature Clim. Change* **5**, 264–268, 2014.

Jung, C. and Schindler, D. Historical Winter Storm Atlas for Germany (GeWiSA). *Atmosphere* **10**, 387, 2019.

Karremann M.K., Pinto J.G., von Bomhard P.J. and Klawa M. On the clustering of winter storm loss events over Germany, *Nat Hazards Earth Sys* **14**, 2041–2052, 2014.

Keef, C., Tawn, J., Svensson, C. Spatial risk assessment for extreme river flows. *J. R. Stat. Soc. C* **58**, 601–61, 2009.

Kendall, M. A. New Measure of Rank Correlation". Biometrika. 30 (1–2): 81–89, 1938.

Klawa, M. & Ulbrich, U. A model for the estimation of storm losses and the identification of severe winter storms in Germany. *Nat. Hazards Earth Syst. Sci*. **3**, 725–732, 2003.

Landwehr, M.J., Matalas, N. C. & Wallis, J. R. Probability weighted moments compared with some traditional techniques in estimating Gumbel Parameters and quantiles. *Water Resources Research* **15**, 1055–1064, 1979.

Lindsey, J. K. *Parametric statistical inference* Clarendon Press, Oxford, 1996.

Mari, D. and Kotz, S. *Correlation and Dependence*. Imperial College Press, 2001.

Merz., R., and Blöschl, G. A process typology of regional floods. *Water Resources Research* **19**, doi.org/10.1029/2002WR001952, 2003.

Mitchell-Wallace, K., Jones, M., Hiller, J., and, Foote, M. *Natural catastrophe Risk Management and Modelling - Practioner's Guid*. Willey Blackwell, Chichester, UK, 2017.

Mudelsee. M. Statistical analysis of climate extremes. Cambridge University Press, Cambridge, pp 124–129, 2020.

Munich RE, GeoRisks Research Department. Winter Storms in Europe (II) Analysis of 1999 losses and loss potentials, 2002.

National Hurricane Centre, Saffir-Simpson Hurricane Wind Scale, web page www.nhc.noaa.gov/aboutsshws.php (last download Spring 2020).

Osinski, R. et al. An approach to build an event set of European windstorms based on ECMWF EPS. *Nat. Hazards Earth Syst. Sci.* **16**, 255–268, 2016.

Papalexiou, S.M, Serinaldi, F., and Porcu, E~~:~~. Advancing Space-Time Simulation of Random Fields: From Storms to Cyclones and Beyond. *Water Resource Research* 57, e2020WR029466, 2021.

Pearson, K. The problem of the random walk. Nature 72(1865), 294, doi:10.1038/0722940b0.

Perils AG. Products – Industry and Loss Database. Web presence, https://www.perils.org/products/industry-exposure-and-loss-database, last visit August 2021.

Pfeifer, D. Study 4: Extreme value theory in actuarial consulting: windstorm losses in Central Europa. In: R.-D. Reiss & M. Thomas: Statistical Analysis of Extreme Values – with Applications to insurance, finance, hydrology and other fields. 2nd Ed., Birkhäuser, Basel, 373~~-~~–378, 2001.

Punge, H.J., Bedka, K.M., Kunz, M. et al. A new physically based stochastic event catalog for hail in Europe. Nat Hazards 73, 1625–1645, 2014.

Raschke, M., Bilis, V. and Kröger, W. Vulnerability of the Swiss electric power grid against natural hazards. In Proceedings of *11th International Conference on Applications of Statistics and Probability in Civil Engineering* (ICASP11), Zurich, Switzerland, 2011.

Raschke, M. Statistical modelling of ground motion relations for seismic hazard analysis. *Journal of Seismology* **17**, 1157~~-~~–1182, 2013.

Raschke, M. Statistical detection and modeling of the over-dispersion of winter storm occurrence. *Nat. Hazards Earth Syst. Sci*. 1**5**, 1757~~-~~–1761, 2015a.

Raschke, M. Statistics of flood risk. *Nature Clim. Change* **4**, 843~~-~~–844, 2015b.

Raschke, M. A Statistical Perspective on Catastrophe Models. 31st International Congress of Actuaries (ICA), Berlin (https://www.researchgate.net/publication/325673290_A_statistical_perspective_on_catastrophe_models/link/5b1ccb60aca272021cf47c03/download), 2018.

Raschke, M. Alternative modelling and inference methods for claim size distributions. *Annals of Actuarial Science* **14**, 1~~-~~–19, 2020.

Roberts, J.F., Champion, A. J., Dawkins, L. C., Hodges, K. I., Shaffrey, L. C., Stephenson, D. B., Stringer, M. A., Thornton, H. E., and Youngman, B. D. The XWS open access catalogue of extreme European windstorms from 1919 to 2012. *Nat Haz Earth Sys Sci* **14**, 2487~~-~~–2501, 2014.

Salazar, S., Francés, F., Komma, J., Blume, T., Francke, T., Bronstert, A., Blöschl, G. A comparative analysis of the effectiveness of flood management measures based on the concept of ~~"~~“retaining water in the landscape~~"~~” in different European hydro-climatic regions. *Nat. Hazards Earth Syst. Sci*. **12**, 1684~~-~~–9981, 2012.

Schabenberger, O., Gotway, C.A: *Statistical Methods for Spatial Data Analysis*. Texts in Statistical Science, ~~Chapmen~~Chapman & Hall, Boca Raton, 2005.

Schlather, M. Models for Stationary Max-Stable Random Fields. *Extremes* **5**, 33–44, 2002.

Schwierz, C., Köllner-Heck, P., Zenklusen Mutter, E. et al. Modelling European winter wind storm losses in current and future climate. Climatic Change 101, 485–514, 2010.

Schoenberg, F.P. and Patel, R.D. Comparison of Pareto and tapered Pareto distributions for environmental phenomena. *Eur. Phys. J. Spec. Top*. **205**, 159–166, 2012.

Simth, R.L. Max-stable processes and spatial extremes. Unpublished manuscript, 1990.

Sklar, A. Fonctions de Répartition à n Dimensions et Leurs Marges. Publications de l'Institut Statistique de l'Université de Paris, 8, 229–231, 1959.

Statistisches Bundesamt (German Office statistics) Preisindizes für die Bauwirtschaft – Mai 2020, 2020.

Stephens M.A. Test based on EDF statistics. in D'Augustino, RB, Stephens, MA (Editors) *Goodness-of-Fit Techniques. statistics*: textbooks and monographs, Vol. 68, Marcel Dekker, New York, 1986.

Stucki, M., & Egli, T. *Synthesebericht - Elementarschutzregister Hagel*. Präventionsstiftung der Kantonale Gebäudeversicherungen, ISBN 978-3-9523300-0-5, 2007.

UNISDR, Global Assessment Report (GAR) Global exposure dataset - population and environmental built, https://data.humdata.org/dataset/exposed-economic-stock, last down load 2020, 2015.

Upton, G. and Cook, I. *A dictionary of statistics*. 2nd rev. Ed., Oxford University Press, 2008.

Waisman, F. European windstorm vendor model comparison (and panel discussion). In *Slides of a presentation at IUA catastrophe risk management conference*, London **30**, (https://www.iua.co.uk/IUA_Member/Events/Catastrophe_Risk_Management_Presentations/European_Windstorm_Vendor_Model_Comparison.aspx) 2015.

Wichura, B. Analyse standortbezogener Windklimatologien als Eingangsgröße für die Bemessung von Bauwerken nach der DIN 1055-4. In book: *Windingenieurwesen in Forschung und Praxis* (pp.157–168) Edition: *WtG-Berichte* **11**, Windtechnologische Gesellschaft e.V., Editor: Udo Peil, 2009.

Youngman, B.D., and Stephenson, D.B. A geostatistical extreme-value framework for fast simulation of natural hazard events. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 472: 2189, 2016.

World Meteorological Organisation, *Guide to Meteorological Instruments and Methods of Observation*, 7th Ed., WMO-No.8, https://www.weather.gov/media/epz/mesonet/CWOP-WMO8.pdf, 2008.