

Modeling the spatial dependence of floods using the Fisher copula

Manuela I. Brunner, Reinhard Furrer, Anne-Catherine Favre

▶ **To cite this version:**

Manuela I. Brunner, Reinhard Furrer, Anne-Catherine Favre. Modeling the spatial dependence of floods using the Fisher copula. Hydrology and Earth System Sciences, 2019, 23, pp.107-124. 10.5194/hess-23-107-2019 . insu-03706499

HAL Id: insu-03706499

https://insu.hal.science/insu-03706499

Submitted on 28 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution| 4.0 International License

Hydrol. Earth Syst. Sci., 23, 107–124, 2019
https://doi.org/10.5194/hess-23-107-2019
© Author(s) 2019. This work is distributed under
the Creative Commons Attribution 4.0 License.

Hydrology and Earth System Sciences
Open Access
EGU

Modeling the spatial dependence of floods using the Fisher copula

Manuela I. Brunner[1,2,3], **Reinhard Furrer**[4], **and Anne-Catherine Favre**[2]

[1]Department of Geography, University of Zurich, Zurich, Switzerland
[2]Univ. Grenoble Alpes, CNRS, IRD, Grenoble INP, IGE, Grenoble, France
[3]Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf ZH, Switzerland
[4]Department of Mathematics and Department of Computational Science, University of Zurich, Zurich, Switzerland

Correspondence: Manuela I. Brunner (manuela.brunner@wsl.ch)

Received: 28 March 2018 – Discussion started: 10 April 2018
Revised: 4 December 2018 – Accepted: 14 December 2018 – Published: 8 January 2019

Abstract. Floods often affect not only a single location, but also a whole region. Flood frequency analysis should therefore be undertaken at a regional scale which requires the considerations of the dependence of events at different locations. This dependence is often neglected even though its consideration is essential to derive reliable flood estimates. A model used in regional multivariate frequency analysis should ideally consider the dependence of events at multiple sites which might show dependence in the lower and/or upper tail of the distribution. We here seek to propose a simple model that on the one hand considers this dependence with respect to the network structure of the region and on the other hand allows for the simulation of stochastic event sets at both gauged and ungauged locations. The new Fisher copula model is used for representing the spatial dependence of flood events in the nested Thur catchment in Switzerland. Flood event samples generated for the gauged stations using the Fisher copula are compared to samples generated by other dependence models allowing for modeling of multivariate data including elliptical copulas, R-vine copulas, and max-stable models. The comparison of the dependence structures of the generated samples shows that the Fisher copula is a suitable model for capturing the spatial dependence in the data. We therefore use the copula in a way such that it can be used in an interpolation context to simulate event sets comprising gauged and ungauged locations. The spatial event sets generated using the Fisher copula well capture the general dependence structure in the data and the upper tail dependence, which is of particular interest when looking at extreme flood events and when extrapolating to higher return periods. The Fisher copula was for a medium-sized catchment found to be a suitable model for the stochastic simulation of flood event sets at multiple gauged and ungauged locations.

1 Introduction

Reliable flood estimates are needed to protect settlements and infrastructure against future floods. Such estimates have traditionally been derived by looking at runoff observations at a single measurement station (Neal et al., 2013). Severe precipitation events, however, often affect not only one sub-catchment, but several ones. This results in widespread floods whose probability is of interest for flood mitigation and for assessing financial risk in the reinsurance industry (Keef et al., 2013). To estimate the probability of widespread floods, a regional multivariate frequency analysis is required which jointly considers discharge values at multiple locations (Serinaldi and Kilsby, 2016). Such an analysis must take into account the dependence of events in different sub-catchments (Thibaud et al., 2013). Spatial dependence was found to be important because no single event caused the maximum flood extent at all locations and assuming perfect correlation between tributaries overestimated flood hazard (Neal et al., 2013). A multivariate approach is therefore needed which represents the spatial dependence between floods from different watersheds (Schulte and Schumann, 2015) which might be radially asymmetric and show non-null tail dependence since more extreme events might be more strongly related than less extreme events. A multivariate approach should allow for the generation of stochastic event sets at multiple stations and for the consideration of the network structure of the catchment. Such event sets can be used for the estimation of

rare events which are missing in the observed records and which affect a larger area. In addition to the generation of event sets at gauged stations, the approach should also allow for the simulation of stochastic event sets at ungauged locations since flood estimates might be required at locations where no runoff observations are available. Both the fitting of the models and the interpolation process should be simple such that they can be easily interpreted and be applied by practitioners.

The spatial dependence of flood events at several locations has been assessed using different types of approaches comprising conditional exceedance models (Keef et al., 2013; Neal et al., 2013), hierarchical Bayesian models (Yan and Moradkhani, 2015), the multivariate skew-t distribution (Ghizzoni et al., 2010, 2012), max-stable models (Wang et al., 2014), and copula models such as pair-copula constructions (Bevacqua et al., 2017; Gräler, 2014; Schulte and Schumann, 2015), hierarchical Kendall copulas (Fischer et al., 2017), and factor copula models (Lee and Joe, 2017). Max-stable and copula models are two classes of approaches which are well established in multivariate extreme value modeling and have been used in other contexts than those of floods. Max-stable distributions arise out of the study of the limiting behavior of vectors of component-wise maxima. There exist a number of parametric max-stable models (Segers, 2012). The consideration of concomitance in annual maxima, however, only informs whether maxima tend to cooccur in the same year without accounting for their timing within a year (Blanchet et al., 2018). To overcome this, the max-stable framework has been extended to the peak-over-threshold setting (Thibaud et al., 2013). Max-stable models have often been applied in the context of extreme precipitation (see, e.g., Blanchet et al., 2018; Davison et al., 2012; Stephenson et al., 2016) and less in the context of floods since most models presupposed that extremal dependence depends only on Euclidean distance (Asadi et al., 2015). While max-stable models describe the marginal distribution and the dependence structure at the same time, copula models allow for the modeling of the dependence structure of multivariate distributions separately from their univariate marginals (Genest and Favre, 2007).

Copula-based models have also been found to be well suited for the spatial interpolation of extremes (Gräler, 2014) since they overcome the limitations of the classical variogram which is sensitive to outliers and influenced by the marginal distribution of the observations (Kazianka and Pilz, 2010). Copula models are therefore a flexible alternative to max-stable models. Bárdossy and Li (2008) have introduced an interpolation method based on copulas. They have used both the Gaussian and a v-transformed normal copula, which was used by Durocher et al. (2016) for predicting flood quantiles at ungauged basins based on physiographical locations. Quessy et al. (2016) applied the general class of chi-square copulas in a spatial context. The family of Fisher copulas generalizes the class of centered chi-square dependence models (Favre et al., 2018). Gräler (2014) introduced spatial vine copulas which are a combination of bivariate copulas not limited to a single copula family. Because of that, they allow for both varying strength of dependence and a changing dependence structure with distance. However, the use of spatial vine copulas requires fitting of several models and makes interpretation more difficult since subsequent pair copulas are conditioned on previous ones.

The models mentioned above usually do not fulfill all of the requirements to satisfactorily generate stochastic flood event sets in a nested catchment with gauged and ungauged locations. Max-stable models usually do not allow for the consideration of the network structure. An exception is the model of Asadi et al. (2015) which distinguishes between flow-connected and flow-unconnected stations. Conditional exceedance models and spatial vine copulas require the fitting of several models. Classical copula models usually do not allow for spatial interpolation and those which allow for interpolation do not allow for both radial asymmetry and non-null tail dependence. The normal copula has well-established theoretical properties, but it can only represent dependence structures that are radially symmetric, and does not incorporate tail dependence. The Student copula and other elliptical copulas allow for tail dependence, but are also radially symmetric (Favre et al., 2018). The chi-square copulas introduced by Bárdossy (2006) are radially asymmetric, but their tail dependence coefficients are null. Favre et al. (2018) have shown that the limitations of these copulas in terms of tail dependence and asymmetry can be overcome by the recently introduced Fisher copula. It allows for high-dimensional modeling, upper tail dependence, and radial asymmetry, and was therefore found to be well suited for modeling multi-site precipitation data in Switzerland. So far, the Fisher copula has only been used in a simulation context for generating stochastic precipitation event sets at gauged locations. However, it has not yet been used in a spatial context to interpolate to ungauged locations.

In this study, we aim at using the Fisher copula such that it can be applied not only to generate stochastic flood event sets at multiple locations, but also in an interpolation context for the generation of event sets including ungauged locations. More specifically, we address the following research questions.

- What is a suitable distance measure for explaining the correlation of events at different locations within a nested catchment?

- Can the Fisher copula capture the dependence structure of the floods observed at several locations within a nested catchment and what is its performance compared to other dependence models in use?

- How can the Fisher copula be used to generate event sets for multiple locations, including ungauged locations?

These questions are answered by applying several dependence models in the nested Thur catchment in Switzerland which is described in the next section. Section 3 then describes how a synchronous flood event set is sampled from observed runoff time series, how the marginal behavior and the dependence structure of the flood sample are modeled, and how the Fisher copula is extended to simulate floods at ungauged locations. In Sect. 4, we compare the dependence structure and the marginal behavior of event sets simulated by the Fisher copula to those of event sets generated by other types of dependence models including two max-stable models and several copula models including elliptical copulas, an extreme value copula, and the R-vine copula. Furthermore, the dependence structure of event sets including ungauged locations is analyzed. Section 5 discusses limitations and perspectives of the proposed approach.

2 Study area

This study was conducted using flood samples from 10 gauging stations situated in the nested Thur catchment in northeastern Switzerland (Fig. 1). The Thur catchment is a tributary of the Rhine and neither has considerable retention areas (Guldener and Wieland, 1980) nor is it strongly regulated (Girons Lopez and Seibert, 2016). The catchment has an area of 1696 km^2 and its elevation ranges from 356 to 2503 m a.s.l. with a mean elevation of 770 m a.s.l. (Federal Office for the Environment FOEN, 2009). The flow regime is snowmelt dominated and the average rainfall is 1350 mm. This rainfall is distributed over all seasons, with the highest amounts in summer. Large precipitation events in the headwaters might cause a rapid discharge build-up in the basin due to the steep terrain and short concentration times (Girons Lopez and Seibert, 2016). The 10 gauging stations are distributed over the whole catchment area and subdivide the catchment into 10 sub-catchments. These are listed in Table 1 together with their main catchment characteristics and an abbreviation that will be used in the sequel of this paper to refer to the individual stations. For each of the measurement stations, an hourly runoff time series spanning 30 to 40 years was available from the Federal Office for the Environment (FOEN) and used as the basis for flood sampling.

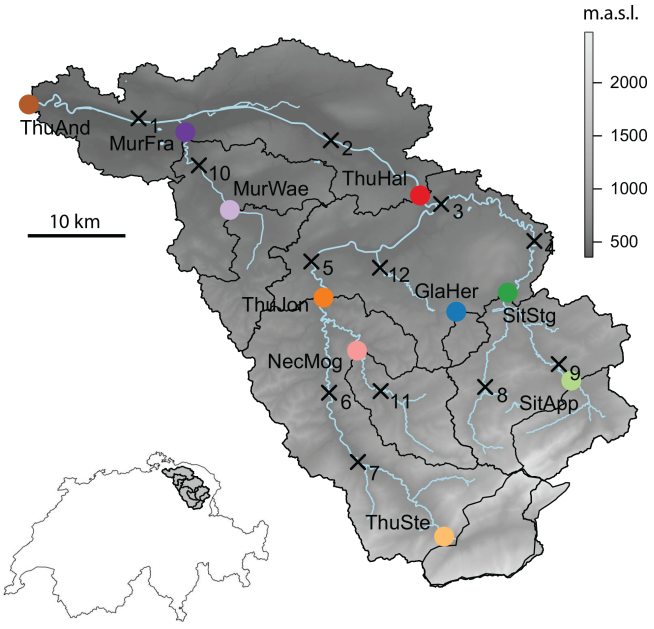


Figure 1. Gauging stations within the Thur catchment in northeastern Switzerland (colored dots) and stations used for spatial interpolation (crosses).

3 Methods

3.1 Event definition

The multivariate analysis was based on a set of flood events which were relevant at the regional scale since a univariate analysis would be sufficient for events relevant only at a local scale. These regional flood events were identified via the three-step procedure illustrated in Fig. 2.

First, flood events were identified at a local scale for each individual station using a peak-over-threshold approach with the 0.9975 quantile of the individual station as a threshold. The dates of occurrence of these local events were listed. Within all the dates of occurrence, independent events were identified by allowing for only one event per week. Second, a synchronous event set was composed by identifying the maximum event magnitudes corresponding to these independent events for each of the individual stations. This procedure allowed for the composition of an event set with events during which at least one station exceeded its 0.9975 quantile. This event set represents an upper set of events that are considered to be dangerous and therefore represent a hazard scenario as defined by Salvadori et al. (2016). However, this set consisted of both events that were only relevant at a local scale where univariate frequency analysis is sufficient and events that were relevant at a regional scale where the consideration of the dependence between events is relevant. Third, these regional events were therefore separated from the local events. To qualify for a regional event, where several stations are jointly affected by a flood event, two criteria had to be fulfilled: (1) the event had to be of similar importance at the individual stations and, (2) over all stations, the event had to belong to the most important ones. The event highlighted in step (3) of Fig. 2 would for example be chosen because it fulfills both criteria. In contrast, the third event might not be chosen because it was not of similar importance at the individual stations. These criteria were quantified in terms of the ranks of the events within the individual station series in order to make observations in catchments with a different size and therefore different event magnitudes comparable. An event could have a low rank in one series but

Table 1. Catchments used for fitting the spatial dependence models tested in this study. The station abbreviation is given together with the full name of the river and the gauging station, the catchment area, the station elevation, the mean elevation of the catchment, and the record period.

Abbreviation	River	Station name	Catchment area (km^2)	Station elevation (m a.s.l.)	Mean elevation (m a.s.l.)	Record (from–to)
ThuSte	Thur	Stein	84	850	1448	1974–2004
NecMog	Necker	Mogelsberg	88	606	959	1974–2013
ThuJon	Thur	Jonschwil	493	534	1030	1974–2013
GlaHer	Glatt	Herisau	16	679	840	1974–2013
SitApp	Sitter	Appenzell	73	790	1186	1974–2013
SitStg	Sitter	St.Gallen	256	636	975	1980–2008
ThuHal	Thur	Halden	1085	456	910	1974–2013
MurWae	Murg	Wängi	79	466	650	1974–2013
MurFra	Murg	Frauenfeld	212	390	580	1974–2013
ThuAnd	Thur	Andelfingen	1696	356	770	1974–2013

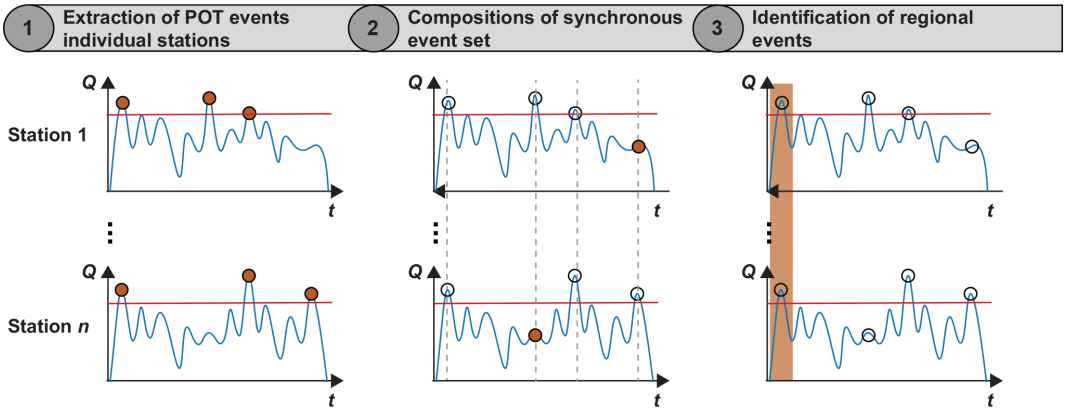


Figure 2. Illustration of the flood event identification procedure consisting of three main steps: (1) extraction of peak-over-threshold (POT) events at individual stations, (2) composition of a synchronous event set, and (3) identification of regional events within the synchronous event set.

a high rank in another series, which would lead to a high variability in ranks across stations. By contrast, an individual event could be assigned similar ranks at different stations, which would lead to a low variability in ranks across stations. Criterion one was a low variability of the ranks of a specific event across different stations (standard deviation of ranks < 50) and criterion two a high rank sum over all stations (rank sum > 1500). The marginal distributions of local and regional events were compared to check whether they were actually distinct. The location parameter was higher for the regional than for the local events and the shape parameters of the two event sets were also clearly distinct. The shape parameter of local events showed a low mean and a high variability across stations, while the regional events showed a higher mean and low variability. The spatial dependence of events assessed via Kendall's tau was higher for the regional than for the local events. The local events were found to be not necessarily regionally important and were therefore excluded from subsequent analyses. The 63 regional events identified occurred in all seasons with a concentration in summer. The regional event set is subsequently referred to as the event set.

3.2 Spatial dependence

The spatial dependence of events was assessed via Kendall's tau for all pairs of stations (Fig. 3). It shows that there is a generally positive dependence between events at most stations, while the dependence is strongest at stations that are closely linked along the stream (e.g., MurWae and MurFra or ThuHal and ThuAnd). Both upper and lower tail dependence were present in the data according to the estimator of Schmidt and Stadtmüller (2006), which needs to be used with care since it provides unreliable estimates for small sample sizes (Serinaldi, 2015). However, upper tail dependence can also be assumed to be present since extreme precipitation events, which might cause extreme flood events, have been shown to exhibit upper tail dependence (see, e.g., Evin et al., 2018; Naveau et al., 2016).

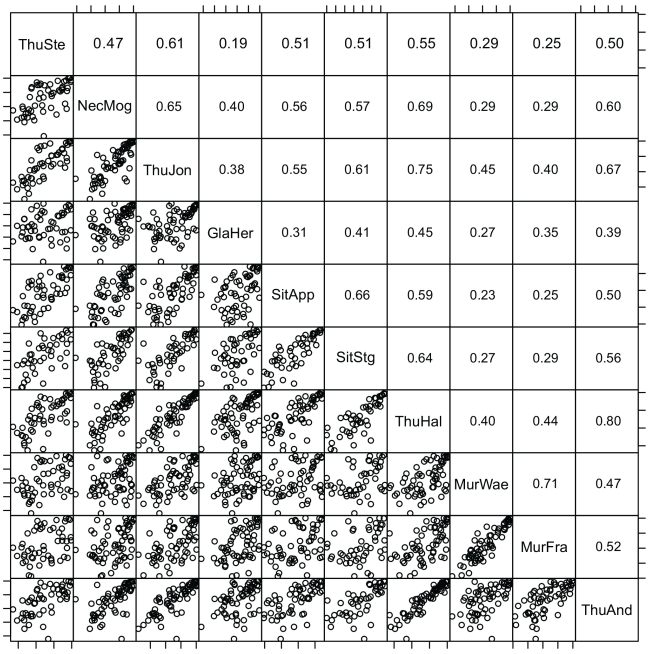


Figure 3. Pairs plot (lower triangular matrix) and Kendall's correlation coefficients (upper triangular matrix) of regional peak-over-threshold events. The stations are labeled with their abbreviations listed in Table 1. Patterns of positive association are visible in all cases.

Spatial methods often relate the dependence between events at two stations to the distance between these stations. Traditionally, the Euclidean distance has been used to do so, which might not be very relevant in the case of floods since they evolve along a river network. A comparison of different distance measures (Euclidean, river distance, physiographical distance) showed that river distance and distance regarding mean catchment elevation explained Kendall's tau best. We used river distance as a distance measure since it has a hydrological meaning. R-package *riverdist* (Tyers, 2017) was used to compute river distance based on a line shapefile of the river network.

A model depicting the behavior of joint events at multiple stations needs to account for both the marginal behavior of the variables and their dependence structure. These two elements are addressed in turn.

3.3 Marginal model

The marginal distribution of extreme values is usually modeled using a GEV distribution in an annual maxima framework or a generalized Pareto distribution (GPD) in a peak-over-threshold framework (Coles, 2001). Strictly speaking, we were in neither of the two frameworks since we were working with a synchronous dataset that has been composed of peak-over-threshold events at individual stations and complemented with the corresponding events at the remaining stations. The event sets at the individual stations were well fit by the GEV distribution. The generalized extreme value distribution (GEV) (Coles, 2001) was not rejected for both types of events in most catchments according to the Anderson–Darling test statistic computed using the procedure proposed by Chen and Balakrishnan (1995) (level $\alpha = 0.05$). This statistic was introduced to test the validity of skewed distributions when unknown parameters must be estimated from the sample data. The marginal GEV distributions were fit for each station separately using maximum likelihood estimation. A regional estimation of the shape parameter ξ would have been desirable (Evin et al., 2016). Pooling the events of different stations, however, was not sensible since even the specific discharges were dependent on the catchment area. A spatial model of the marginal distributions was derived using trend surfaces (Davison et al., 2012). In a first step, the location and scale parameters were normalized by catchment area. In a second step, a regression model was fit for each parameter of the GEV distribution, which allowed for predictions at a given location using catchment characteristics only. Suitable models were identified using stepwise backward regression (Harrell, 2015) on a set of 53 catchment characteristics (for a list and a short description, see Table A1 in Viviroli et al., 2009). The final linear models used the following explanatory variables.

- Location parameter: X-coordinate, percentage area of hard rock, variability of the Julian date of the annual maximum hourly precipitation.
- Scale parameter: X-coordinate, Y-coordinate, catchment area, catchment elevation.
- Shape parameter: catchment shape, soil topographic index, percentage area of hardrock, pasture and arable land in contributing areas, urban area in contributing areas.

They resulted in absolute prediction errors over the 10 stations of 0.11, 0.21, and 0.15, respectively; i.e., the use of the spatial marginal model already leads to prediction errors, which is, however, not the focus of this study.

3.4 Dependence model

The dependence structure was treated independently of the marginal distributions. We modeled the spatial dependence of the flood events using two types of approaches: copula-based approaches and max-stable approaches. While the copula approach assumes uniform marginals, max-stable process models assume unit Fréchet marginals (Cooley et al., 2012). Although the copula and the extreme value frameworks share some connections, only a few authors from the extreme value community adopt the copula framework for multivariate extremes (Ribatet and Sedki, 2013). The different approaches are summarized in Table 2 and described in the paragraphs below.

3.4.1 Copula models

The copula approach has its origin in the representation theorem of Sklar (1959) which states that the joint cumulative distribution function $F_{XY}(x, y)$ of any pair of continuous random variables (X, Y) can be written as

$$F_{XY}(x, y) = C\{F_X(x), F_Y(y)\}, x, y \in \mathbb{R}, \quad (1)$$

where $F_X(x)$ and $F_Y(y)$ are the marginal distributions and $C : [0, 1]^2$ is the copula. C is unique if the marginals are continuous. One of the main advantages of the copula approach is that the selection of an appropriate copula for modeling the dependence between X and Y can proceed independently of the choice of their marginal distributions (Genest and Favre, 2007). An important property of a copula is whether the dependence is the same for high and low values (Bárdossy and Li, 2008), which implies tail symmetry. For a theoretical introduction to copulas, the reader is referred to Durante and Sempi (2015), Joe (2014), or Nelsen (2005), and for an introduction with an engineering focus to Genest and Favre (2007), Salvadori et al. (2007), and Salvadori and De Michele (2007). Guidelines for using copulas in a hydrological and natural hazard context are provided in Favre et al. (2004), Salvadori and De Michele (2004), Salvadori et al. (2014), Salvadori et al. (2015) and Salvadori et al. (2016).

We tested five different types of copulas, which all allow for modeling in high dimensions and therefore at multiple sites: (1) the Gaussian copula, (2) the Student t copula, (3) the Gumbel copula, (4) the R-vine copula, and (5) the recently developed Fisher copula.

1. Gaussian copula. The Gaussian copula is completely determined by the knowledge of the correlation matrix Σ and the parameters of the Gaussian copula are simple to estimate. However, the Gaussian copula does not have tail dependence (Malevergne and Sornette, 2003) and the dependence is symmetrical (Bárdossy and Li, 2008).

2. Student t copula. The description of a Student t copula relies on two parameters: the correlation matrix Σ and the number of degrees of freedom ν. In contrast to the Gaussian copula, it allows for tail dependence (Malevergne and Sornette, 2003), but the dependence is symmetrical (Bárdossy and Li, 2008). Both the Gaussian and Student t copulas allow for the description of spatial variability and therefore for the interpolation to ungauged locations.

3. Gumbel copula. The Gumbel copula belongs both to the class of Archimedean and extreme value copulas. It allows for upper tail dependence (Poulin et al., 2007).

4. R-vine copula. Multivariate data can be modeled using a cascade of simple building blocks, called pair copulas, in a flexible way (Aas et al., 2009). There is an enormous number of possible R-vine tree sequences to choose from. Dißmann et al. (2013) therefore proposed an automated model selection and estimation technique which is implemented in R-package *VineCopula* (Schepsmeier et al., 2017).

5. Fisher copula. The Fisher copula arises when components of a Student random vector are squared. This construction procedure follows the construction of the chi-square copula family which is obtained when squaring the components of a normal random vector (Bárdossy, 2007). The Fisher copula is characterized by the two parameters ν and Σ, similar to the Student t copula. The Fisher copula generalizes the class of centered chi-square dependence models since it tends to the centered chi-square copula as $\nu \to \infty$. For $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_d)^T$, the d-dimensional Fisher copula can be expressed by

$$C^{\mathrm{F}}_{\Sigma,\nu}(u_1, \ldots, u_d) = \sum_{\boldsymbol{\epsilon}\in\{-1,1\}^d} \left(\prod_{j=1}^{d} \epsilon_j\right) C^{\mathrm{t}}_{\Sigma,\nu}\left(\frac{1+\epsilon_1 u_1}{2}, \ldots, \frac{1+\epsilon_d u_d}{2}\right), \quad (2)$$

where $C^{\mathrm{t}}_{\Sigma,\nu}$ is the Student copula and u are the uniform marginals. The elements of the parameter ϵ take either the value -1 or 1. The parameter does not have a particular role with respect to the dependence structure (Quessy et al., 2016). The Fisher copula allows for modeling at multiple sites, non-vanishing upper tail dependence, radial asymmetry, and a pairwise structure that helps to interpret results (Favre et al., 2018).

The Gaussian, Student t, and Gumbel copulas were fitted using maximum pseudo-likelihood estimation on pseudo-observations. The R-vine copula was estimated using the automated model selection and estimation technique by (Dißmann et al., 2013). The Fisher copula was estimated using the two-step pseudo-maximum likelihood estimator proposed by Favre et al. (2018), which allows for estimation of the parameters of the Fisher copula (Σ and ν) based on the relationship between each entry of Σ and the pairwise values of Kendall's tau (Favre et al., 2018).

3.4.2 Max-stable models

Max-stable processes extend extremal models to the spatial context. Various parametric models of max-stable processes have been proposed in the literature (Blanchet et al., 2018), including the Gaussian extreme value (Smith et al., 1990) and the Brown–Resnick process (Kabluchko et al., 2009), which were found to fit precipitation data better than other max-stable processes (Davison et al., 2012). Max-stable processes assume asymptotic dependence (i.e., dependence will not disappear at very large distances between stations) and usually use variables converted to Fréchet margins. For a

Table 2. Modeling capabilities of copula and max-stable models tested to model spatial dependence. Asymmetry and tail dependence are not specified in the context of max-stable models.

Model	Asymmetry	Upper tail dependence	Lower tail dependence	Asymptotic dependence	Simulation at multiple stations	Spatial interpolation	Simple model setup
Gaussian copula	No	No	No	No	Yes	Yes	Yes
Student t copula	No	Yes	Yes	No	Yes	Yes	Yes
Gumbel copula	Yes	Yes	No	No	Yes	No	Yes
R-vine copula	Yes	Yes	Yes	No	Yes	Yes	No
Fisher copula	Yes	Yes	No	No	Yes	Yes	Yes
Gaussian max-stable	–	–	–	Yes	Yes	Yes	Yes
Brown–Resnick max-stable	–	–	–	Yes	Yes	Yes	Yes

detailed overview of max-stable processes, the reader is referred to Padoan (2013). Standard procedures for fitting max-stable processes were established for coordinates and Euclidean distances. We therefore used classical multidimensional scaling (Borg and Groenen, 2010) to search for coordinates which could correspond to the river distance matrix. The max-stable models were then fitted using the Euclidean distances of the coordinates obtained by multidimensional scaling.

3.5 Simulation for gauged locations

Each of the dependence models outlined above can be used to generate a random flood event set for the gauged stations in the region under study. We used each of these models to simulate $n = 1000$ flood event sets. These event sets consisted of variables with uniform margins in the case of copula models and of variables with Fréchet margins in the case of max-stable models. To get values on the original scale, these values had to be back-transformed using a probability integral transform (Genest and Rivest, 2001). The back-transformation was achieved by using the predicted parameters of the GEV distribution obtained by the marginal model (see Sect. 3.3).

3.6 Validation

The multivariate distribution per se could not directly be validated since no quantitative goodness of fit test is available to help in rejecting unsuitable models (Ghizzoni et al., 2010). There are a few validation techniques which are called multivariate, but they are essentially aggregation metrics (see, e.g., Li and Lu, 2018) and therefore not truly multivariate. Each of the samples obtained by simulation was therefore compared to the observations with respect to their marginal behavior and with respect to the dependence structure separately. For the assessment of the dependence structure, we computed different dependence measures such as the F-madogram which is closely related to the extremal coefficient and usually used in the context of max-stable processes, tail dependence coefficients which are usually computed in the context of copulas, and Kendall's tau which tells us something about the general dependence structure of the data. For the assessment of the marginal behavior, we displayed QQ plots.

The F-madogram summarizes the spatial dependence structure of the data (Cooley et al., 2006) and is expressed as

$$v^F(h) = \frac{1}{2}\mathrm{E}|F[Z(x+h)] - F[Z(x)]|, \quad (3)$$

where the margin of $Z(x)$ equals Fréchet margins, $F(x) = \exp(-1/x)$, and h is the distance between a pair of stations.

The F-madograms computed for the different simulated samples were compared to the F-madogram of the observations. This allowed for the evaluation of whether the dependence in relation to the distance between stations was captured by the models. Tail dependence describes the dependence in the upper or lower tail of a distribution and the upper tail dependence coefficient describes the probability that one margin exceeds a high threshold given that the other margin also exceeds a high threshold (Poulin et al., 2007). We used the tail dependence estimator of Schmidt and Stadtmüller (2006) with a parameter $\kappa = 5$ to produce upper tail dependence plots for the simulated samples and the observations. Since tail dependence estimators react quite sensitively to small sample sizes (Serinaldi, 2015), we also computed Kendall's tau, which is a non-parametric measure of dependence and does not only focus on the tails of the distributions (Poulin et al., 2007).

3.7 Interpolation to ungauged location(s)

The model found to be most suitable for modeling the spatial dependence in the regional flood data was used for the simulation of event sets including ungauged locations. The results shown in Sect. 4 (Figs. 7 and 8) indicate that the Fisher copula best reproduces the dependence structure in the data. The Fisher copula is non-parametric and uses an empirical correlation matrix (Favre et al., 2018). Each entry of the correlation matrix Σ can be related to the value of Kendall's tau of

a corresponding data pair. The Fisher copula can be adapted to a spatial context since pairwise distances between stations can be incorporated into the model.

Using only the gauged stations in the study catchment, the correlation matrix Σ had a dimension of 10×10. If a new, ungauged location was added, it was converted to a 11×11 matrix, if s ungauged locations were added to a $(10+s) \times (10+s)$ matrix. The new matrix still consisted of the former 10×10 matrix in the upper left corner completed with s vectors of correlation values for the new stations with respect to the other stations. The vectors of correlations could be derived via an empirical correlogram. The empirical correlogram consisted of pairwise correlations plotted against the river distance between stations. To derive the values for an ungauged location, its distance to all gauged stations was determined. The correlogram could then be used to derive the correlations corresponding to these distances. These were used to extend the matrix Σ. The new extended matrix could finally be used as a parameter of the Fisher copula to generate a flood event set for the gauged and ungauged locations in the dataset. The individual steps of this procedure are outlined in more detail below.

1. An empirical correlogram was set up which relates the entries of the Σ matrix corresponding to pairs of stations to the distance between these locations (Fig. 4, black points).

2. A non-parametric or parametric model for the correlogram was chosen. The empirical correlogram was on the one hand smoothed using smoothing splines and on the other hand fit by an exponential model. Both models indicated a gradual decrease in correlation with increasing distance if the smoothing parameter of the spline was chosen high enough. Subsequently, we used the exponential model; however, the procedure could also be applied using the smoothing spline.

3. One or several ungauged locations of interest had to be identified. We here used 12 ungauged locations distributed along the river network which could be prone to flooding.

4. For each of these locations, the river distance with respect to all other locations was determined. These values were used as the new entries of the extended Σ matrix.

5. The extended matrix was then used to simulate $n = 1000$ random flood event sets for a set of stations comprising gauged and ungauged locations using the new Σ as a parameter in the Fisher copula. The second parameter of the Fisher copula ν, which corresponds to the number of degrees of freedom of the Student distribution and is by the Fisher copula required to be an integer, was found to be only weakly sensitive to an extension of Σ. This was assessed by reestimating ν using the new simulated data. The estimated ν remained at $\nu = 7$ when adding only one station and slightly increased for $s > 10$ to $\nu = 8$.

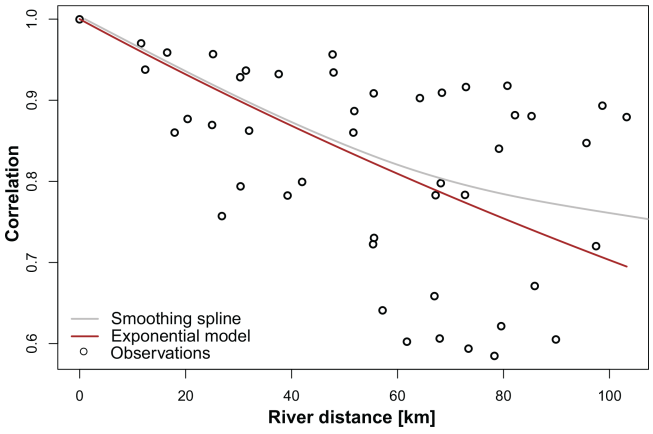


Figure 4. Correlation (entries of the Σ matrix) plotted against river distance. The observations (dots) are given together with a smoothing spline (grey line) and an exponential model (brown line).

The adaptation of the Fisher copula to the spatial context was validated by using leave-one-out cross-validation. To do so, 1 of the 10 gauged stations was considered ungauged, and a Fisher copula fit to this reduced dataset. The Sigma matrix was then completed by applying the procedure described above to the station that was considered to be ungauged. The completed matrix was used to simulate $n = 1000$ event sets which were visually summarized by Kendall's correlation matrix. A comparison of the correlation matrices resulting from excluding one station at a time with the correlation matrix of the observations showed that these were rather similar.

4 Results

4.1 Spatial dependence

The exploratory data analysis of the observed flood event sets showed that spatial dependence was present in the data. This is illustrated in Fig. 5 by the weekly occurrence of the annual maxima events and the ranks of their magnitudes. A correspondence in the week of occurrence indicates temporal dependence and a similarity in the ranks of the magnitudes across stations indicates spatial dependence. Very severe events were recorded at all stations within the nested catchment and were ranked high in most catchments (e.g., events 1977, 1978, 1999, and 2013). This implies dependence in the upper tail of the multivariate distribution. The most severe events were summer events (see week of occurrence). The ranks of rather weak events were also related (see event 2003, which was a very dry year), which indicates lower tail dependence in the multivariate distribution.

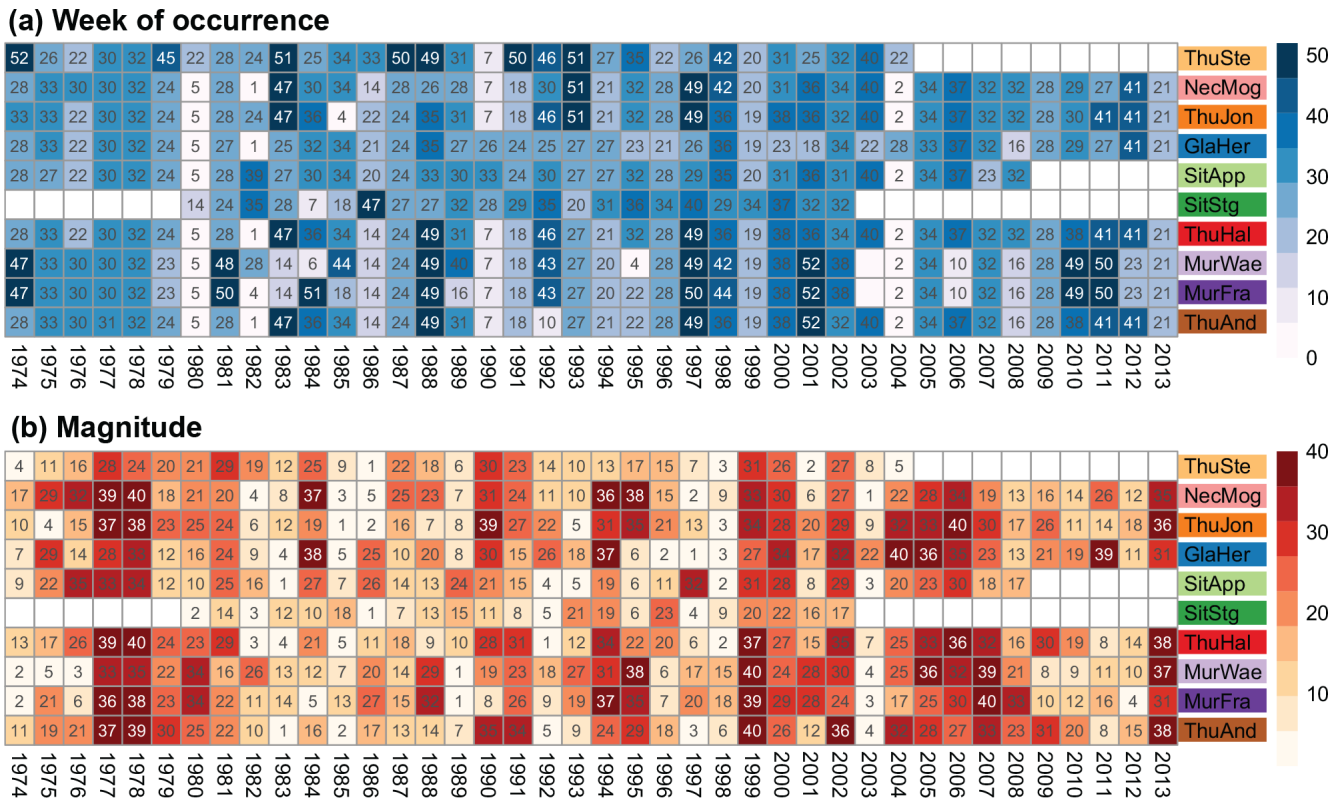


Figure 5. Heatplots of the week of occurrence **(a)** and the magnitude **(b)** of annual maxima of the 10 stations within the study area. The darker the color, the higher the ranks. Missing records lead to empty, white cells.

4.2 Marginal model

Figure 6 shows how well the marginal distributions of the flood events at the individual stations were modeled by a fitted GEV distribution (see column "Fitted marginals"), by a GEV model with regionalized parameters (see column "Regionalized marginals"), and by samples generated using several dependence models and back-transformed using the regionalized GEV model (see columns three to nine). The p-values computed by the Kolmogorov–Smirnov test using the individual samples and the observations are provided in Table 3. The marginal distributions of the observations were accurately modeled by the fitted GEV model. As expected, the GEV model using the regionalized GEV parameters introduced some bias, mainly in the upper part of the distribution. The samples generated using the dependence models and back-transformed using the regionalized GEV parameters showed a very similar picture to these regionalized marginal distributions.

4.3 Dependence model

Figure 7 shows the F-madogram of the observations and the different dependence models. It shows that the Gaussian, Student t, and Gumbel copula introduced dependence structures that did not reflect the decreasing dependence among pairs of stations with increasing distance (virtually constant estimates). The Gaussian max-stable model did introduce dependence decreasing with the distance among pairs of stations, but this decrease was too strong. The Brown–Resnick max-stable model and the R-vine and Fisher copulas introduced a dependence pattern that closely resembled the one of the observations. The results of the F-madogram were confirmed by Kendall's tau matrices (Fig. 8) which showed that the general dependence structure in the data was best reproduced by the R-vine and Fisher copulas, while the Gaussian, Student t and Gumbel copula and the Brown–Resnick max-stable models introduced overly uniform dependence structures.

Figure 9 shows how well the upper tail dependence in the data was reproduced by the different dependence models. The tail dependence was rather underestimated by the Gaussian and R-vine copulas, while it was overestimated by the Student t and Gumbel copulas and the Gaussian max-stable model. The Fisher copula model generated samples with variable tail dependence across pairs of stations and most closely reproduced the tail dependence of the observations even though it also led to an overestimation of tail dependence.

The lower tail dependence in the data, which was rather weak for most pairs of stations, was not reproduced by any of

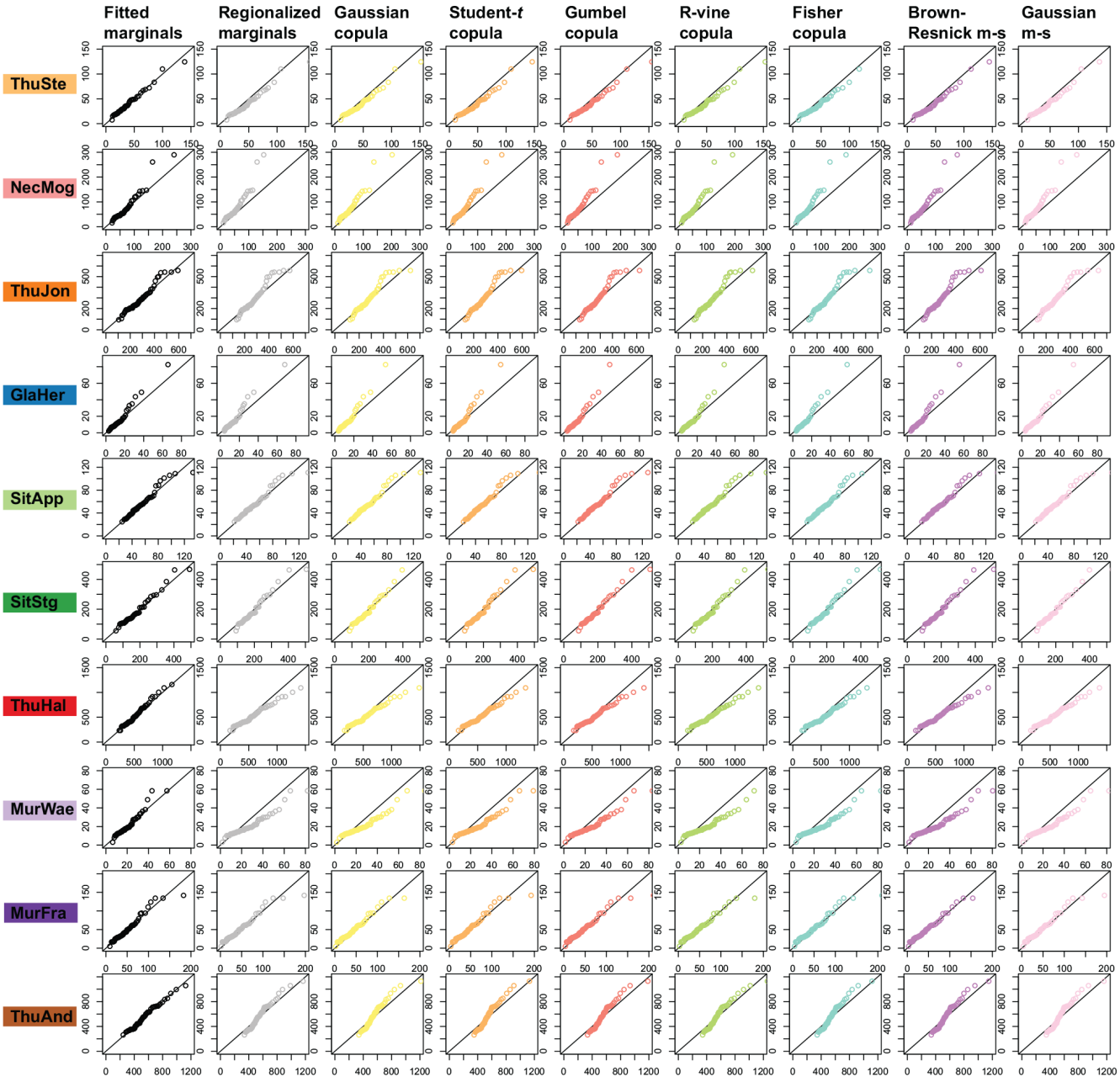


Figure 6. The first two columns show QQ plots obtained by the fitted GEV parameters (fitted marginals) and by the regionalized GEV parameters (regionalized marginals) for the 10 study stations. The following columns show QQ plots of samples simulated using different models and back-transformed using the regionalized GEV distribution.

the models tested. It was overestimated by the Student *t* copula and the Gaussian max-stable model and underestimated by all other models.

4.4 Simulation for gauged locations

Figure 10 displays the return periods of peak discharges simulated at all stations corresponding to the 100-year event at a reference station. The boxplots indicate the range of return periods obtained by repeating the simulation 2000 times. The first subplot of Fig. 10 for example chose subcatchment Thur-Stein as the reference station. After having generated a stochastic event set using the Fisher copula, the 100-year event at this station was determined using univariate frequency analysis. Then, the return period of the values corresponding to the same generated event set at the other stations was determined by looking at their univariate, empirical distributions. The simulation procedure was repeated 2000

Table 3. p-values for all stations obtained when comparing two samples, observations, and simulations, using the Kolmogorov–Smirnov test. The name of the simulated dataset is indicated in the column name. Note that the p-values are not adjusted for multiple testing.

Station	Fitted marginals	Regionalized marginals	Gaussian copula	Student t copula	Gumbel copula	R-vine copula	Fisher copula	Brown–Resnick max-stable	Gaussian max-stable
ThuSte	0.97	0.13	0.20	0.08	0.20	0.13	0.13	0.08	0.13
NecMog	0.69	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ThuJon	0.94	0.94	0.99	0.94	0.94	0.94	0.94	0.94	0.94
GlaHer	0.94	0.54	0.83	0.69	0.69	0.54	0.69	0.54	0.69
SitApp	1.00	0.79	0.49	0.64	0.49	0.64	0.64	0.79	0.49
SitStg	0.96	0.37	0.37	0.25	0.52	0.37	0.37	0.37	0.52
ThuHal	0.54	0.83	0.83	0.69	0.69	0.83	0.69	0.69	0.83
MurWae	0.83	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.02
MurFra	0.83	0.54	0.41	0.29	0.29	0.41	0.41	0.29	0.29
ThuAnd	0.99	0.06	0.06	0.06	0.06	0.09	0.06	0.06	0.06

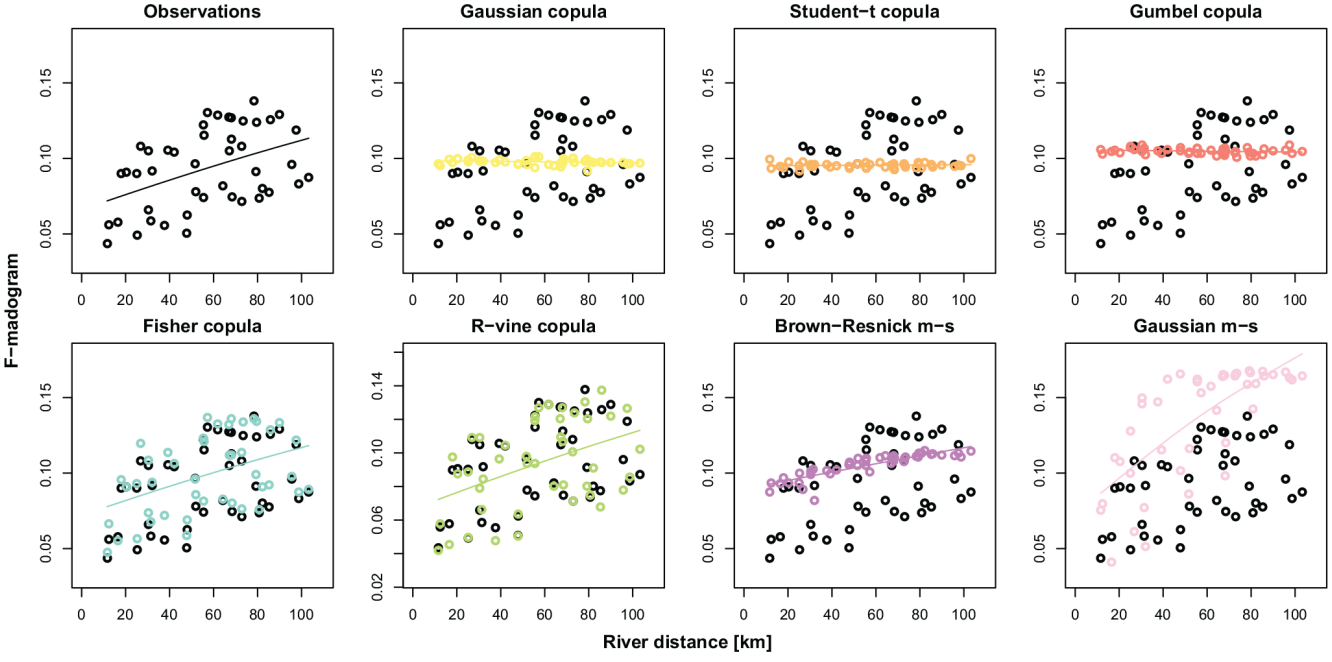


Figure 7. F-madograms of observations and samples simulated by different types of dependence models. A smoothed line was added to the data. Each simulated dataset is displayed in a separate panel to increase readability.

times, which resulted in a range of return periods for each station corresponding to the 100-year event at the reference station. A 100-year event at one station was usually associated with events with return periods higher than 100 years at some stations and with events with return periods lower than 100 years at other stations. Some stations were rather independent of others since their return periods differed strongly from the return periods of the same events at other stations. The values generated for the tributaries of the Murg and the Glatt could have rather low return periods, while the Thur at Andelfingen could show high return periods, and vice versa (see subplots GlaHer, MurFra, and ThuAnd).

4.5 Interpolation to ungauged location(s)

The Fisher copula was applied in a spatial context to generate event samples for a set of locations consisting of both gauged and ungauged locations. The dependence structure of a sample generated using the Fisher copula is shown in Fig. 11. The dependence structure among the gauged catchments was similar to the one obtained when using the model consisting of the gauged stations only. The dependence structure obtained for the ungauged catchments could not be directly compared to observations, but it could be verified by comparing the dependence values to the locations of the stations on the river network (Fig. 1). Stations close to one another in terms of river distance should have a higher de-

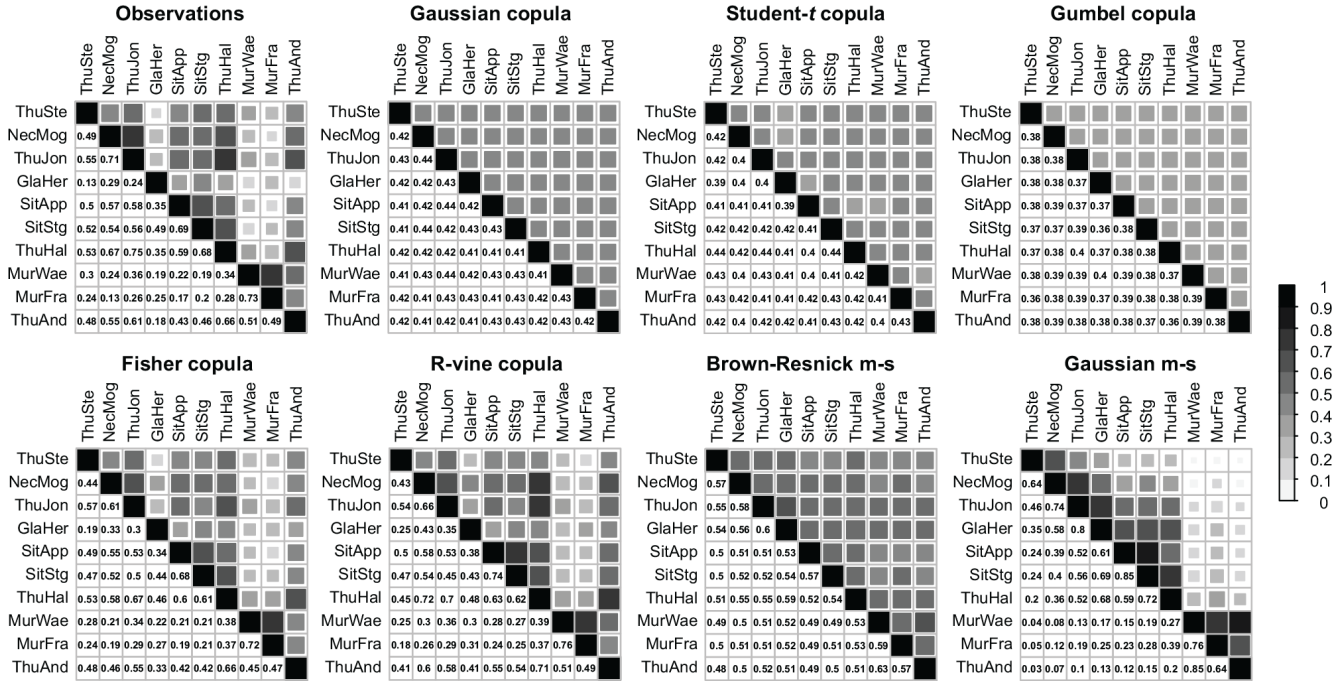


Figure 8. Kendall's tau matrices for the observations and the event samples generated by different types of dependence models. The darker the color, the stronger the dependence.

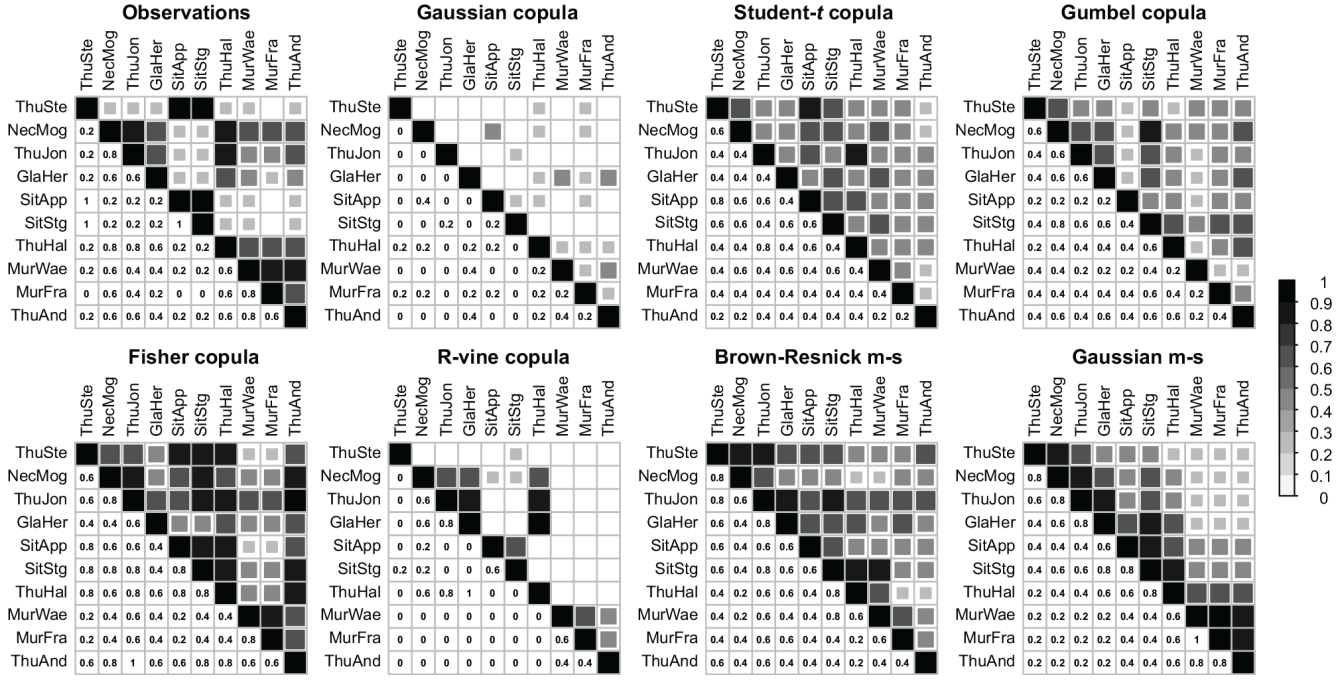


Figure 9. Upper tail dependence matrices for the observations and the event samples generated by different types of dependence models. The darker the color, the stronger the dependence.

pendence than stations which are not part of the same subcatchment and separated by longer river distances. Gauged station Thur-Halden and ungauged location 3 for example showed high dependence in terms of Kendall's tau since they are very close to each other. The same could be observed for ungauged locations 1 and 10. In contrast, gauged station Thur-Stein and ungauged location 8 showed a low dependence because they are separated by a long river distance

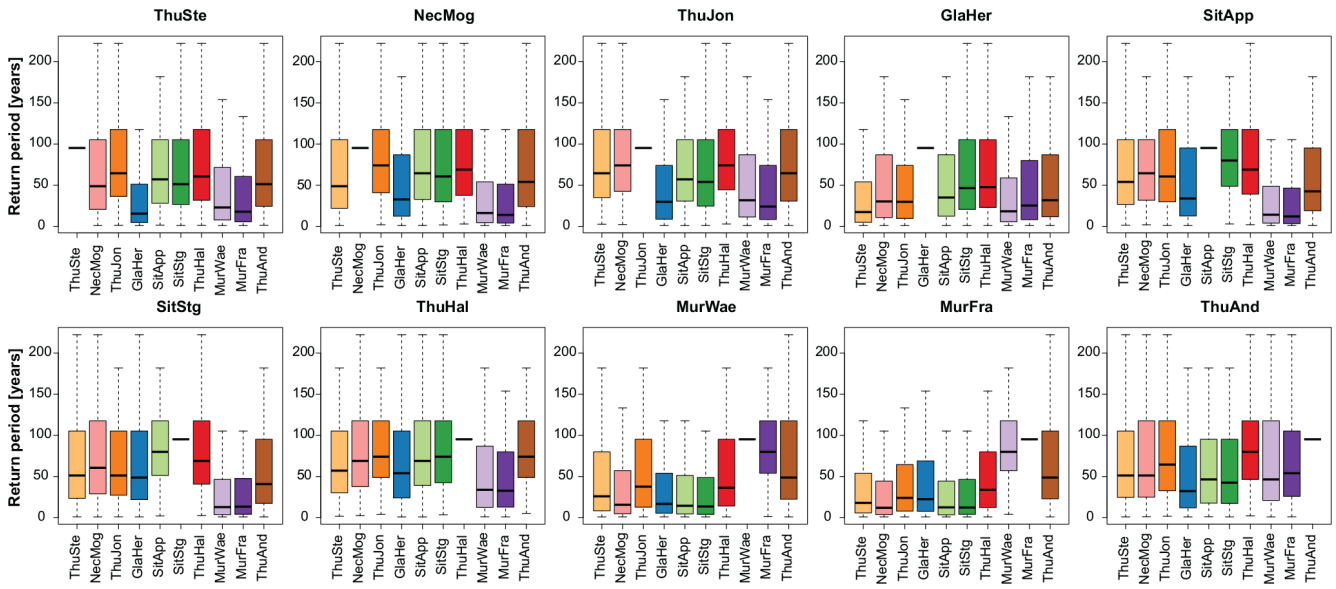


Figure 10. Return period of the events of all stations corresponding to the 100-year flood event determined at one single reference station. The reference station is indicated by the horizontal tick at a return period of 100 years. The interquartile range and the whiskers (last value within 1.5 times the interquartile range) indicate the range of the return periods of the values corresponding to the same event set measured at the other stations.

even though they are quite close in terms of Euclidean distance. The same was also true for, e.g., ungauged locations 8 and 11.

5 Discussion

The Euclidean distance is used as a standard distance measure when looking at spatial extremes such as precipitation, temperature, and wind (Asadi et al., 2015). However, our study as well as the one by Asadi et al. (2015) for the Danube River have shown that the Euclidean distance between stations has low explanatory power in the case of regional floods. We found that river distance is a suitable dependence measure for describing the change in dependence along the river network. Other distances such as distance with respect to mean altitude were also found to be suitable since it is strongly related to distance along the river network. Our results have shown that correlations and the F-madogram between stations at the same river distance can still vary quite a bit since two pairs of stations could be located in the upper or lower part of the catchment, which seems to have an impact on the correlation coefficient. This clearly shows the need for a multivariate assessment of dependence.

Most models tested were not able to reproduce the dependence structure in the flood event data of the Thur catchment with respect to the river distance between the individual locations. This was indicated by the non-variable F-madogram estimates for the Gaussian, Student t, and Gumbel copulas. The Gaussian and Student t copulas were not able to reproduce the dependence structure in the data because they have symmetric tail dependence and were not able to model the decrease in dependence with increasing distance between the stations. This is in line with results by Davison et al. (2012), who found that the use of Gaussian copulas was not satisfactory when looking at extreme precipitation. Similarly, the Gumbel copula was not able to model the dependence structure in the data despite its asymmetry in tail dependence. The max-stable models also captured neither the general dependence structure in the data nor the upper or lower tail dependence. The F-madogram produced by the Gaussian max-stable process was more variable than the one of the Brown–Resnick max-stable process, which better fit the observed data. The difference between the two max-stable models might be explained by the higher flexibility of the Brown–Resnick process compared to the Gaussian process (Cooley et al., 2012). The R-vine copula was thanks to its flexible structure able to reproduce the general dependence structure. This is in line with findings by Schulte and Schumann (2015), who worked with pair-copula constructions and found that the use of flexible Archimedean copulas allowed representation of spatial dependence in flood event data. This is essentially due to the fact that Archimedean copulas can model asymmetric lower and upper tail dependence, but they are only available for lower dimensions. In our case, the tail dependence was not well reproduced by the R-vine copulas but, as mentioned previously, currently available tail dependence estimators are not very reliable (Serinaldi et al., 2015) and all the results regarding tail dependence need to be interpreted with care.

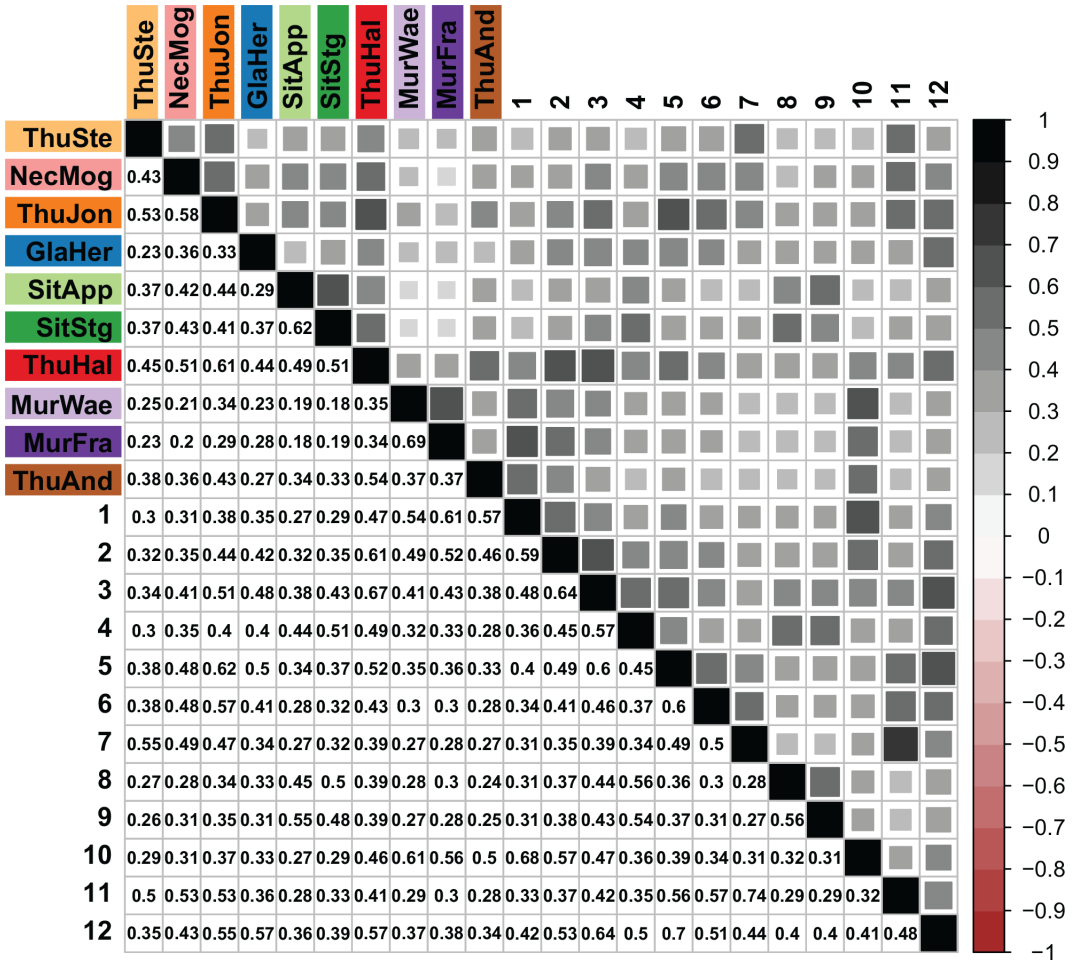


Figure 11. Kendall's tau matrix of samples generated by the Fisher copula for the 10 gauged (colored abbreviations) and 12 ungauged (numbers 1–12) locations (see Fig. 1). Kendall's tau matrix of the samples generated does not equal the matrix displayed in Fig. 8 for the Fisher copula since the samples are randomly generated and therefore not perfectly reproducible.

The Fisher copula in contrast to most other models tested reproduced well the general dependence structure in the data as described by Kendall's tau. This was possible since the empirical correlation matrix was used for computations and there exist many degrees of freedom. The Fisher copula was also able to model the upper tail of the distribution, while it was unable to reproduce the lower tail dependence structure in the data. This is not surprising since the Fisher copula has a non-null upper tail dependence coefficient but no non-null lower tail dependence coefficient (Favre et al., 2018). If one wanted to give even more weight to the upper tail of the distribution, the inference of the Fisher copula parameter ν could be based on the upper tail dependence matrix instead of Kendall's tau matrix. Currently, to our knowledge, no copula model in more than three dimensions is available which models asymmetric lower and upper tail dependence. Such a class of copula could be developed by squaring the components of elliptical random vectors other than the Student distribution (e.g., Laplace or logistic). One copula potentially able to model asymmetric lower and upper tail dependence might be the copula developed by Khoudraji (1995) and generalized by Durante and Salvadori (2010) to a d-dimensional space. However, the tail dependence of this type of copula has not been studied yet. Ghizzoni et al. (2010) and Ghizzoni et al. (2012) have shown that the multivariate version of the skew-t distribution can be an alternative to model spatial dependence in flood events since it allows for asymmetric upper and lower tail dependence and can be used in high dimensions. However, this approach has not yet been extended to an interpolation context where event sets can be simulated for both gauged and ungauged locations. Another alternative for flexibly modeling the spatial dependence in the data is the conditional exceedance model by Heffernan and Tawn (2004) which has successfully been applied in the context of flood events by Keef et al. (2013) and Neal et al. (2013). This model accounts for varying strengths of extremal dependence; however, a number of models must be fit conditional

on each gauging station in turn, and no interpolation to ungauged locations is possible so far.

Contrary to the conditional exceedance and the multivariate skew-t distribution approaches, the Fisher copula can be applied for the simulation of event sets including ungauged locations. The Fisher copula was adjusted by extending the Σ correlation matrix to the desired dimension using the values of a fitted correlogram corresponding to the river distances among stations. An alternative method to this statistical approach would be to use a rainfall–runoff model on stochastically generated precipitation series for the generation of event sets including ungauged locations (Haberlandt et al., 2011). Rainfall–runoff modeling implicitly creates spatial dependence between flows at different locations (Neal et al., 2013). However, the use of a rainfall–runoff model compared to a statistical analysis of flow records introduces additional sources of uncertainties such as model and parameter uncertainty of the hydrological model (Sikorska et al., 2015). Focusing on flood event data avoids complexity and reduces the number of uncertainty sources (Keef et al., 2013). In the approach presented here, uncertainty arises from both the marginal model and the spatial dependence model. The uncertainty coming from these two uncertainty sources would need to be assessed in a subsequent study. Such a study could be set up in a similar way to studies which have been performed in the context of uncertainty of flood estimates characterizing flood events in terms of the dependent variables peak discharge, flood volume, and duration. Brunner et al. (2018b), Dung et al. (2015), and Serinaldi (2013) have shown how the uncertainty due to the marginal distributions and the dependence between variables can jointly be assessed.

In addition to the spatial dependence, the marginal distributions had to be regionalized to ungauged catchments. They were regionalized using a simple linear model which led to an acceptable performance. However, more sophisticated regionalization techniques such as the use of nonlinear regression techniques (Brunner et al., 2018a), the use of the region of influence approach (Hosking and Wallis, 1997), or the copula-based clustering algorithm proposed by Pappadà et al. (2018) could be applied if the marginal distributions were of particular interest and if a larger dataset was available. Alternatively, new regionalization techniques based on catchment similarity in terms of empirical copulas (Grimaldi et al., 2016) could be developed. A larger dataset would also be desirable for reliable estimation of the dependence structure. One possibility of increasing the dataset would be to exploit the whole content of a continuous runoff series instead of only using a flood event sample as suggested by Serinaldi and Kilsby (2017).

So far, the spatial Fisher copula model has only been applied to the Thur study region. However, its application is not limited to this particular catchment. The Fisher copula can potentially be fitted to the flood event data of other regions since it is quite flexible. When transferring the method to other regions one might need to answer the question of whether Alpine and lowland catchments or humid and dry catchments show similar dependence structures or whether the spatial dependence structure varies with the topography and the hydro-climatology of the catchment. Keef et al. (2009) have shown that spatial dependence is strong in areas where catchment characteristics are similar. A point which has not been addressed here is whether the spatial dependence structure is constant over time. Blanchet et al. (2018) have found that there was a trend in the co-occurrence of extreme precipitation events. Such a trend analysis should ideally also be performed on a flood dataset covering a longer observation period.

6 Conclusions

River distance was found to be a suitable distance measure for explaining the spatial dependence in the flood event data of the Thur catchment in Switzerland. In contrast to other copula and max-stable models, the Fisher copula was able to model both the general dependence structure of the data and its upper tail dependence. However, it failed at capturing the lower tail dependence in the data. Thanks to the main interest in the upper tail of the distribution when modeling flood events, it was still found to be a suitable model for representing the spatial dependence of flood events. Due to its flexibility it can be adjusted to a context where both gauged and ungauged catchments can be modeled. Flood event sets generated by the Fisher copula can be used in spatial flood risk analyses by combining them with hydraulic models and flood loss models.

Data availability. The data used are available upon order from the FOEN using the form at https://www.bafu.admin.ch/bafu/en/home/topics/water/state/data/obtaining-monitoring-data-on-the-topic-of-water/hydrological-data-service-for-watercourses-and-lakes.html (last access: 1 March 2018).

Author contributions. The idea and setup for the paper were jointly developed by the three co-authors. The analyses were performed by MIB and discussed with the co-authors. MIB wrote the first draft of the manuscript which was revised and edited by RF and ACF.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank the Federal Office for the Environment (FOEN) for funding the project (contract 13.0028.KP/M285-0623) and for providing runoff measurement data. We also thank Daniel Viviroli for providing the catchment characteristic data for the 10 gauged catchments.

Edited by: Alberto Guadagnini
Reviewed by: two anonymous referees

References

Aas, K., Czado, C., Frigessi, A., and Bakken, H.: Pair-copula constructions of multiple dependence, Insur. Math. Econ., 44, 182–198, https://doi.org/10.1016/j.insmatheco.2007.02.001, 2009.

Asadi, P., Davison, A. C., and Engelke, S.: Extremes on river networks, Ann. Appl. Stat., 9, 2023–2050, https://doi.org/10.1214/15-AOAS863, 2015.

Bárdossy, A.: Copula-based geostatistical models for groundwater quality parameters, Water Resour. Res., 42, 1–12, https://doi.org/10.1029/2005WR004754, 2006.

Bárdossy, A.: Calibration of hydrological model parameters for ungauged catchments, Hydrol. Earth Syst. Sci., 11, 703–710, https://doi.org/10.5194/hess-11-703-2007, 2007.

Bárdossy, A. and Li, J.: Geostatistical interpolation using copulas, Water Resour. Res., 44, 1–15, https://doi.org/10.1029/2007WR006115, 2008.

Bevacqua, E., Maraun, D., Hobæk Haff, I., Widmann, M., and Vrac, M.: Multivariate statistical modelling of compound events via pair-copula constructions: analysis of floods in Ravenna (Italy), Hydrol. Earth Syst. Sci., 21, 2701–2723, https://doi.org/10.5194/hess-21-2701-2017, 2017.

Blanchet, J., Aly, C., Vischel, T., Panthou, G., Sané, Y., and Kane, M. D.: Trend in the co-occurrence of extreme daily rainfall in West Africa since 1950, J. Geophys. Res.-Atmos., 123, 1–16, https://doi.org/10.1002/2017JD027219, 2018.

Borg, I. and Groenen, P.: Modern Multidimensional Scaling: Theory and Applications, Springer, New York, 2010.

Brunner, M. I., Furrer, R., Sikorska, A. E., Viviroli, D., Seibert, J., and Favre, A.-C.: Synthetic design hydrographs for ungauged catchments: A comparison of regionalization methods, Stoch. Env. Res. Risk A., 32, 1993–2023, https://doi.org/10.1007/s00477-018-1523-3, 2018a.

Brunner, M. I., Sikorska, A. E., Furrer, R., and Favre, A.-C.: Uncertainty assessment of synthetic design hydrographs for gauged and ungauged catchments, Water Resour. Res., 54, 1–20, https://doi.org/10.1002/2017WR021129, 2018b.

Chen, G. and Balakrishnan, N.: A general purpose approximate goodness-of-fit test, J. Qual. Technol., 2, 154–161, 1995.

Coles, S.: An introduction to statistical modeling of extreme values, Springer, London, 2001.

Cooley, D., Naveau, P., and Poncet, P.: Variograms for spatial max-stable random fields, in: Lecture Notes in Statistics. Dependence in Probability and Statistics, Springer, New York, 373–390, 2006.

Cooley, D., Cisewski, J., Erhardt, R. J., Jeon, S., Mannshardt, E., Omolo, B. O., and Sun, Y.: A survey of spatial extremes: Measuring spatial dependence and modeling spatial effects, Revstat-Stat. J., 10, 135–165, 2012.

Davison, A. C., Padoan, S. A., and Ribatet, M.: Statistical Modeling of Spatial Extremes, Stat. Sci., 27, 161–186, https://doi.org/10.1214/11-STS376, 2012.

Dißmann, J., Brechmann, E. C., Czado, C., and Kurowicka, D.: Selecting and estimating regular vine copulae and application to financial returns, Comput. Stat. Data An., 59, 52–69, https://doi.org/10.1016/j.csda.2012.08.010, 2013.

Dung, N. V., Merz, B., Bárdossy, A., and Apel, H.: Handling uncertainty in bivariate quantile estimation – An application to flood hazard analysis in the Mekong Delta, J. Hydrol., 527, 704–717, https://doi.org/10.1016/j.jhydrol.2015.05.033, 2015.

Durante, F. and Salvadori, G.: On the construction of multivariate extreme value models via copulas, Environmetrics, 21, 143–161, 2010.

Durante, F. and Sempi, C.: Principles of copula theory, CRC Press, Taylor & Francis Group, Boca Raton, 2015.

Durocher, M., Chebana, F., and Ouarda, T. B. M. J.: Delineation of homogenous regions using hydrological variables predicted by projection pursuit regression, Hydrol. Earth Syst. Sci., 20, 4717–4729, https://doi.org/10.5194/hess-20-4717-2016, 2016.

Evin, G., Blanchet, J., Paquet, E., Garavaglia, F., and Penot, D.: A regional model for extreme rainfall based on weather patterns subsampling, J. Hydrol., 541, 1185–1198, https://doi.org/10.1016/j.jhydrol.2016.08.024, 2016.

Evin, G., Favre, A.-C., and Hingray, B.: Stochastic generation of multi-site daily precipitation focusing on extreme events, Hydrol. Earth Syst. Sci., 22, 655–672, https://doi.org/10.5194/hess-22-655-2018, 2018.

Favre, A.-C., Adlouni, S. E., Perreault, L., Thiémonge, N., and Bobée, B.: Multivariate hydrological frequency analysis using copulas, Water Resour. Res., 40, W01101, https://doi.org/10.1029/2003WR002456, 2004.

Favre, A.-C., Quessy, J.-F., and Toupin, M.-H.: The new family of Fisher copulas to model upper tail dependence and radial asymmetry: properties and application to high-dimensional rainfall data, Environmetrics, 29, 1–17, https://doi.org/10.1002/env.2494, 2018.

Federal Office for the Environment FOEN: Verzeichnis der eidgenössischen hydrometrischen Stationen auf Ende 2009, Hydrologisches Jahrbuch Schweiz, 59–78, 2009.

Fischer, S., Gräler, B., and Schumann, A.: Spatio-temporal assessment of flood events by hierarchical Kendall-copulas, in: EGU General Assembly 2017, Geophysical Research Abstracts, vol. 19, Vienna, 2017.

Genest, C. and Favre, A.-C.: Everything you always wanted to know about copula modeling but were afraid to ask, J. Hydrol. Eng., 12, 347–367, https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(347), 2007.

Genest, C. and Rivest, L. P.: On the multivariate probability integral transformation, Stat. Probabil. Lett., 53, 391–399, https://doi.org/10.1016/S0167-7152(01)00047-5, 2001.

Ghizzoni, T., Roth, G., and Rudari, R.: Multivariate skew-t approach to the design of accumulation risk scenarios for the flooding hazard, Adv. Water Resour., 33, 1243–1255, https://doi.org/10.1016/j.advwatres.2010.08.003, 2010.

Ghizzoni, T., Roth, G., and Rudari, R.: Multisite flooding hazard assessment in the Upper Mississippi River, J. Hydrol., 412–413, 101–113, https://doi.org/10.1016/j.jhydrol.2011.06.004, 2012.

Girons Lopez, M. and Seibert, J.: Influence of hydro-meteorological data spatial aggregation on streamflow modelling, J. Hydrol., 541, 1212–1220, https://doi.org/10.1016/j.jhydrol.2016.08.026, 2016.

Gräler, B.: Modelling skewed spatial random fields through the spatial vine copula, Spat. Stat., 10, 87–102, https://doi.org/10.1016/j.spasta.2014.01.001, 2014.

Grimaldi, S., Petroselli, A., Salvadori, G., and De Michele, C.: Catchment compatibility via copulas: A non-parametric study of the dependence structures of hydrological responses, Adv. Water Resour., 90, 116–133, https://doi.org/10.1016/j.advwatres.2016.02.003, 2016.

Guldener, H. and Wieland, C.: The Thur river and its highfloods, in: INTERPRAEVENT 1980, Bad Ischl, 225–235, 1980.

Haberlandt, U., Hundecha, Y., Pahlow, M., and Schumann, A. H.: Rainfall Generators for Application in Flood Studies, in: Flood Risk Assessment and Management, edited by: Schumann, A., chap. 7, Springer Science+Business Media, Dordrecht, 117–147, 2011.

Harrell, F. E.: Regression Modeling Strategies. With applications to linear models, logistic and ordinal regression, and survival analysis, Springer, Cham, 2015.

Heffernan, J. E. and Tawn, J.: A conditional approach to modelling multivariate extreme values, J. R. Stat. Soc. B, 66, 497–546, https://doi.org/10.1111/j.1467-9868.2004.02050.x, 2004.

Hosking, J. R. M. and Wallis, J. R.: Regional frequency analysis, Cambridge University Press, Cambridge, https://doi.org/10.1017/CBO9780511529443, 1997.

Joe, H.: Dependence modeling with copulas, Chapman & Hall/CRC, Boca Raton, 2014.

Kabluchko, Z., Schlather, M., and de Haan, L.: Stationary max-stable fields associated to negative definite functions, Ann. Probab., 37, 2042–2065, https://doi.org/10.1214/09-AOP455, 2009.

Kazianka, H. and Pilz, J.: Spatial interpolation using copula-based geostatistical models, in: geoENV VII – Geostatistics for Environmental Applications, edited by: Atkinson, P. and Lloyd, C., Springer, Southampton, 307–319, 2010.

Keef, C., Svensson, C., and Tawn, J. A.: Spatial dependence in extreme river flows and precipitation for Great Britain, J. Hydrol., 378, 240–252, https://doi.org/10.1016/j.jhydrol.2009.09.026, 2009.

Keef, C., Tawn, J. A., and Lamb, R.: Estimating the probability of widespread flood events, Environmetrics, 24, 13–21, https://doi.org/10.1002/env.2190, 2013.

Khoudraji, A.: Contribution à l'étude des copules et à la modélisation de valeurs extrêmes bivariées, PhD thesis, Université Laval, Québec, Canada, 1995.

Lee, D. and Joe, H.: Multivariate extreme value copulas with factor and tree dependence structures, Extremes, 1–30, https://doi.org/10.1007/s10687-017-0298-0, 2017.

Li, L. and Lu, Z.: A new method for model validation with multivariate output, Reliab. Eng. Syst. Saf., 169, 579–592, https://doi.org/10.1016/j.ress.2017.10.005, 2018.

Malevergne, Y. and Sornette, D.: Testing the Gaussian copula hypothesis for financial assets dependences, Quant. Financ., 3, 231–250, 2003.

Naveau, P., Huser, R., Ribereau, P., and Hannart, A.: Modeling jointly low, moderate, and heavy rainfall intensiies without a threshold selection, Water Resour. Res., 52, 2753–2769, https://doi.org/10.1002/2015WR018552, 2016.

Neal, J., Keef, C., Bates, P., Beven, K., and Leedal, D.: Probabilistic flood risk mapping including spatial dependence, Hydrol. Process., 27, 1349–1363, https://doi.org/10.1002/hyp.9572, 2013.

Nelsen, R. B.: An introduction to copulas, Springer Science & Business Media, New York, 2nd edn., 2005.

Padoan, S. A.: Max-Stable Processes, in: Encyclopedia of Environmetrics, John Wiley & Sons Ltd., https://doi.org/10.1002/9780470057339.vnn022, 2013.

Pappadà, R., Durante, F., Salvadori, G., and De Michele, C.: Clustering of concurrent flood risks via Hazard Scenarios, Spat. Stat., 23, 124–142, https://doi.org/10.1016/j.spasta.2017.12.002, 2018.

Poulin, A., Huard, D., Favre, A.-C., and Pugin, S.: Importance of tail dependence in bivariate frequency analysis, J. Hydrol. Eng., 12, 394–403, https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(394), 2007.

Quessy, J. F., Rivest, L. P., and Toupin, M. H.: On the family of multivariate chi-square copulas, J. Multivariate Anal., 152, 40–60, https://doi.org/10.1016/j.jmva.2016.07.007, 2016.

Ribatet, M. and Sedki, M.: Extreme value copulas and max-stable processes, Journal de la Société Française de Statistique, 154, 138–150, 2013.

Salvadori, G. and De Michele, C.: Frequency analysis via copulas: Theoretical aspects and applications to hydrological events, Water Resour. Res., 40, W12511, https://doi.org/10.1029/2004WR003133, 2004.

Salvadori, G. and De Michele, C.: On the Use of Copulas in Hydrology: Theory and Practice, J. Hydrol. Eng., 12, 369–380, https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(369), 2007.

Salvadori, G., Michele, C. D., Kottegoda, N. T., and Rosso, R.: Extremes in nature. An approach using copulas, vol. 56 of *Water Science and Technology Library*, Springer, Dordrecht, 2007.

Salvadori, G., Tomasicchio, G. R., and D'Alessandro, F.: Practical guidelines for multivariate analysis and design in coastal and off-shore engineering, Coast. Eng., 88, 1–14, https://doi.org/10.1016/j.coastaleng.2014.01.011, 2014.

Salvadori, G., Durante, F., Tomasicchio, G. R., and D'Alessandro, F.: Practical guidelines for the multivariate assessment of the structural risk in coastal and off-shore engineering, Coast. Eng., 95, 77–83, 2015.

Salvadori, G., Durante, F., De Michele, C., Bernardi, M., and Petrella, L.: A multivariate copula-based framework for dealing with hazard scenarios and failure probabilities, Water Resour. Res., 52, 3701–3721, https://doi.org/10.1002/2015WR017225, 2016.

Schepsmeier, U., Stoeber, J., Christian, E., Graeler, B., Nagler, T., Erhardt, T., Almeida, C., Min, A., Czado, C., Hofmann, M., Killiches, M., Joe, H., and Vatter, T.: Package "VineCopula", CRAN-R, 1–161, https://doi.org/10.1126/science.1136800, 2017.

Schmidt, R. and Stadtmüller, U.: Non-parametric estimation of tail dependence, Scand. J. Stat., 33, 307–335, 2006.

Schulte, M. and Schumann, A. H.: Extensive spatio-temporal assessment of flood events by application of pair-copulas, Proc. IAHS, 370, 177–181, https://doi.org/10.5194/piahs-370-177-2015, 2015.

Segers, J.: Max-stable models for multivariate extremes, Revstat Statistical Journal, 10, 61–82, 2012.

Serinaldi, F.: An uncertain journey around the tails of multivariate hydrological distributions, Water Resour. Res., 49, 6527–6547, https://doi.org/10.1002/wrcr.20531, 2013.

Serinaldi, F.: Can we tell more than we can know? The limits of bivariate drought analyses in the United States, Stoch. Env.

Res. Risk A., 30, 1691–1704, https://doi.org/10.1007/s00477-015-1124-3, 2015.

Serinaldi, F. and Kilsby, C. G.: Irreversibility and complex network behavior of stream flow fluctuations, Physica A, 450, 585–600, https://doi.org/10.1016/j.physa.2016.01.043, 2016.

Serinaldi, F. and Kilsby, C. G.: A Blueprint for Full Collective Flood Risk Estimation: Demonstration for European River Flooding, Risk Anal., 37, 1958–1976, https://doi.org/10.1111/risa.12747, 2017.

Serinaldi, F., Bardossy, A., and Kilsby, C. G.: Upper tail dependence in rainfall extremes: would we know it if we saw it?, Stoch. Env. Res. Risk A., 29, 1211–1233, 2015.

Sikorska, A. E., Viviroli, D., and Seibert, J.: Flood type classification in mountainous catchments using crisp and fuzzy decision trees, Water Resour. Res., 51, 7959–7976, https://doi.org/10.1002/2015WR017326, 2015.

Sklar, A.: Fonctions de répartition à n dimensions et leurs marges, Publ. Inst Statist Univ. Paris, 8, 229–231, 1959.

Smith, R. L., Tawn, J. A., and Yuen, H. K.: Statistics of multivariate extremes, Int. Stat. Rev., 58, 47–58, 1990.

Stephenson, A. G., Lehmann, E. A., and Phatak, A.: A max-stable process model for rainfall extremes at different accumulation durations, Weather Climate Extremes, 13, 44–53, https://doi.org/10.1016/j.wace.2016.07.002, 2016.

Thibaud, E., Mutzner, R., and Davison, A. C.: Threshold modeling of extreme spatial rainfall, Water Resour. Res., 49, 4633–4644, https://doi.org/10.1002/wrcr.20329, 2013.

Tyers, M.: Package "riverdist", Tech. rep., available at: https://cran.r-project.org/web/packages/riverdist/riverdist.pdf (last access: 1 February 2018), 2017.

Viviroli, D., Mittelbach, H., Gurtz, J., and Weingartner, R.: Continuous simulation for flood estimation in ungauged mesoscale catchments of Switzerland – Part II: Parameter regionalisation and flood estimation results, J. Hydrol., 377, 208–225, https://doi.org/10.1016/j.jhydrol.2009.08.022, 2009.

Wang, Z., Yan, J., and Zhang, Y.: Incorporating spatial dependence in regional frequency analysis, Water Resour. Res., 50, 9570–9585, https://doi.org/10.1002/2013WR014849, 2014.

Yan, H. and Moradkhani, H.: A regional Bayesian hierarchical model for flood frequency analysis, Stoch. Env. Res. Risk A., 29, 1019–1036, https://doi.org/10.1007/s00477-014-0975-3, 2015.