

## A regional comparative analysis of empirical and theoretical flood peak-volume relationships

Ján Szolgay<sup>1</sup>, Ladislav Gaál<sup>1,2</sup>, Tomáš Bacigál<sup>3</sup>, Silvia Kohnová<sup>1\*</sup>, Kamila Hlavčová<sup>1</sup>, Roman Výleta<sup>1</sup>, Juraj Parajka<sup>4</sup>, Günter Blösch<sup>4</sup>

<sup>1</sup> Slovak University of Technology in Bratislava, Faculty of Civil Engineering, Department of Land and Water Resources Management, Radlinského 11, 810 05 Bratislava, Slovakia.

<sup>2</sup> MicroStep-MIS, s.r.o., Čavojského 1, SK 841 04 Bratislava 4, Slovakia.

<sup>3</sup> Slovak University of Technology in Bratislava, Faculty of Civil Engineering, Department of Mathematics and Constructive Geometry, Radlinského 11, 810 05 Bratislava, Slovakia.

<sup>4</sup> Institute for Hydraulic and Water Resources Engineering, Vienna University of Technology, Karlsplatz 13/223, A-1040 Vienna, Austria.

\* Corresponding author. E-mail: silvia.kohnova@stuba.sk

**Abstract:** This paper analyses the bivariate relationship between flood peaks and corresponding flood event volumes modelled by empirical and theoretical copulas in a regional context, with a focus on flood generation processes in general, the regional differentiation of these and the effect of the sample size on reliable discrimination among models. A total of 72 catchments in North-West of Austria are analysed for the period 1976–2007. From the hourly runoff data set, 25 697 flood events were isolated and assigned to one of three flood process types: synoptic floods (including long- and short-rain floods), flash floods or snowmelt floods (both rain-on-snow and snowmelt floods). The first step of the analysis examines whether the empirical peak-volume copulas of different flood process types are regionally statistically distinguishable, separately for each catchment and the role of the sample size on the strength of the statements. The results indicate that the empirical copulas of flash floods tend to be different from those of the synoptic and snowmelt floods. The second step examines how similar are the empirical flood peak-volume copulas between catchments for a given flood type across the region. Empirical copulas of synoptic floods are the least similar between the catchments, however with the decrease of the sample size the difference between the performances of the process types becomes small. The third step examines the goodness-of-fit of different commonly used copula types to the data samples that represent the annual maxima of flood peaks and the respective volumes both regardless of flood generating processes (the traditional engineering approach) and also considering the three process-based classes. Extreme value copulas (Galambos, Gumbel and Hüsler-Reiss) show the best performance both for synoptic and flash floods, while the Frank copula shows the best performance for snowmelt floods. It is concluded that there is merit in treating flood types separately when analysing and estimating flood peak-volume dependence copulas; however, even the enlarged dataset gained by the process-based analysis in this study does not give sufficient information for a reliable model choice for multivariate statistical analysis of flood peaks and volumes.

**Keywords:** Flood types; Regionalisation; Flood peaks; Flood volumes; Copulas; Goodness-of-fit; Comparative hydrology.

### INTRODUCTION

The relationship between flood peaks and flood event volumes, and the choice of appropriate statistical models for these are interesting topics both from a theoretical perspective (as it is a statistical fingerprint of catchment response) and from a practical point of view (the design of retention basins and identifying the extent and duration of flooding in flood hazard zones). Most of the recent research on statistical models of flood peak-volume relationships revolved around copulas for individual catchments (Bačová Mitková, 2012; Bačová Mitková and Halmová, 2014; Ben-Aissia et al., 2012; Chowdhary et al., 2011; Favre et al., 2004; Ganguli and Reddy, 2013; Genest and Favre, 2007; Karmakar and Simonovic, 2009; Poulin et al., 2007; Reddy and Ganguli, 2012; Shiau et al., 2006; Sraj et al., 2014; Zhang and Singh, 2006). These studies examined parameter estimation methods and the suitability of particular copula types on the basis of a single or a few catchments. Sets of catchments were examined by Zhang and Sing (2006) or Grimaldi and Serinaldi (2006); though, none of them adopted a spatial perspective.

Estimating bivariate distributions from observed peak-volume pairs, however, is associated with substantial uncertain-

ty due to the larger number of degrees of freedom relative to univariate distributions. Observed flood records of the required length are rarely available, and this poses a problem for reliable estimations of flood risk in bivariate design cases. This difficulty, however, may be overcome by adopting the principles of univariate regional flood frequency analysis to the multivariate case. The usual procedure of estimating flood peak distributions from short records is to pool a number of similar catchments in a region and estimate the parameters from the combined sample (Salinas et al., 2014). There is a vast amount of well-known methods of regional flood frequency analysis ranging from the index flood method to more refined techniques, which will not be referenced here (e.g., Dalrymple, 1960; Gaál et al., 2013; IH, 1999; Kohnová and Szolgay, 1999). They all build on the notion of catchment similarity (e.g., Rosbjerg et al., 2013). A similar pooling procedure as is usual for univariate distributions could, in principle, also be applied to bivariate distributions. Such an attempt for the bivariate case was reported by combining copulas with bivariate quantile curves in Chebana and Ouarda (2009), who also extended the well-known Hosking and Wallis tests (Hosking and Wallis, 1997) for the multivariate case (Chebana and Ouarda, 2007, 2009). These mostly theoretical studies were complemented by Requena et al. (2016) in

order to arrive at a comprehensive statistical multivariate regional procedure using the multivariate index-flood models as a basis including guidelines for application.

Moreover, it is being increasingly recognized that the problem of estimating bivariate distributions cannot be approached from only a purely statistical perspective. In this respect, Serinaldi and Kilsby (2013) and Szolgay et al. (2016) highlighted the importance of studying the relationships between processes that generate the design variables and also the statistical techniques used to model them.

In this respect, however, less is known about the similarity of both empirical and theoretical peak-volume relationships between catchments attributable to different flood generation processes in a region from the perspective of comparative hydrology, which could aid to process-based statistical model selection. Gaál et al. (2012) therefore analysed the ratio of flood volumes and peaks (termed as flood time scale) based on the concept of comparative hydrology and compared catchments with contrasting characteristics in a regional context in Austria in order to understand the hydrologic and climatic controls on this ratio in a holistic way. The results described the roles of climate (through the type of precipitation generated), together with the attributes of the environment and flow generation processes (through antecedent soil moisture and soil characteristics) on the magnitude of flood time scales. Gaál et al. (2014) attempted to understand the causal hydrological factors controlling the strength of the relationship between flood peaks and volumes for the same data quantified by Spearman's rank correlation coefficient. The results suggested that the factors controlling the strength of the dependence are more related to climate instead of catchment characteristics. Szolgay et al. (2015, 2016) aimed at analysing the formal suitability of various copula models of flood peaks and flood volumes, with a particular focus flood generating seasons (summer and winter floods) and processes (synoptic floods, flash floods or snowmelt floods) with the goal of going a little beyond the statistics alone in the choice of the copula functions from the perspective of engineering applications. Szolgay et al. (2015) also illustrated the importance of considering the influence of the length of data series through two simple simulation experiments. The results indicated that the acceptance of copula models can be conditioned on the flood types but the length of the series and possibly also the homogeneity of generation processes within a particular flood type within a region may play an important role and requires a more detailed analysis, which is attempted in this study.

The purpose of this paper is also to shed light on the dependence of flood peaks – flood event volumes from a regional/subregional perspective. In order to go beyond purely statistical analyses and understand the process controls, the paper adopts the concept of comparative analysis of flood process types in a climatically rather homogeneous region and its different, geologically homogeneous subregions. A more differentiated regional look at the selection of a hydrological process-oriented copula model for flood peak/volume relationships is attempted as it was done in Szolgay et al. (2015, 2016), and the role of the length of the available data series is also analysed in more detail.

Specifically, the paper addresses four science questions: (i) How similar are the empirical peak-volume dependence structures of different flood types for a given catchment in a larger region and its subregions? (ii) How similar are the empirical peak-volume dependence structures between catchments for a given flood type in a larger region and its subregions? (iii) What copula types are most suitable for a given flood process type in a larger region and its subregions? (iv) How limiting are

the usually available length of data series in aiding a process-oriented regional bivariate copula-based model selection?

The paper is structured as follows. The next section introduces the pilot region with its 3 subregions, the runoff data used in the analysis and the rules which were used to identify and classify the rainfall–runoff events. Next, an individual section offers a brief overview of the copula methods. After this, the results of the study are presented. These are based on a statistical testing of equality of empirical copulas when, first, different flood processes are compared locally, and second, for a given flood type, data samples from different sites are compared regionally. In the last parts of the Results section, the goodness-of-fit of selected copulas to different data sets are shown. Finally, the last section discusses the outcomes of the study and indicates some possible paths to go along with such analyses for the future.

## STUDY REGION AND DATA

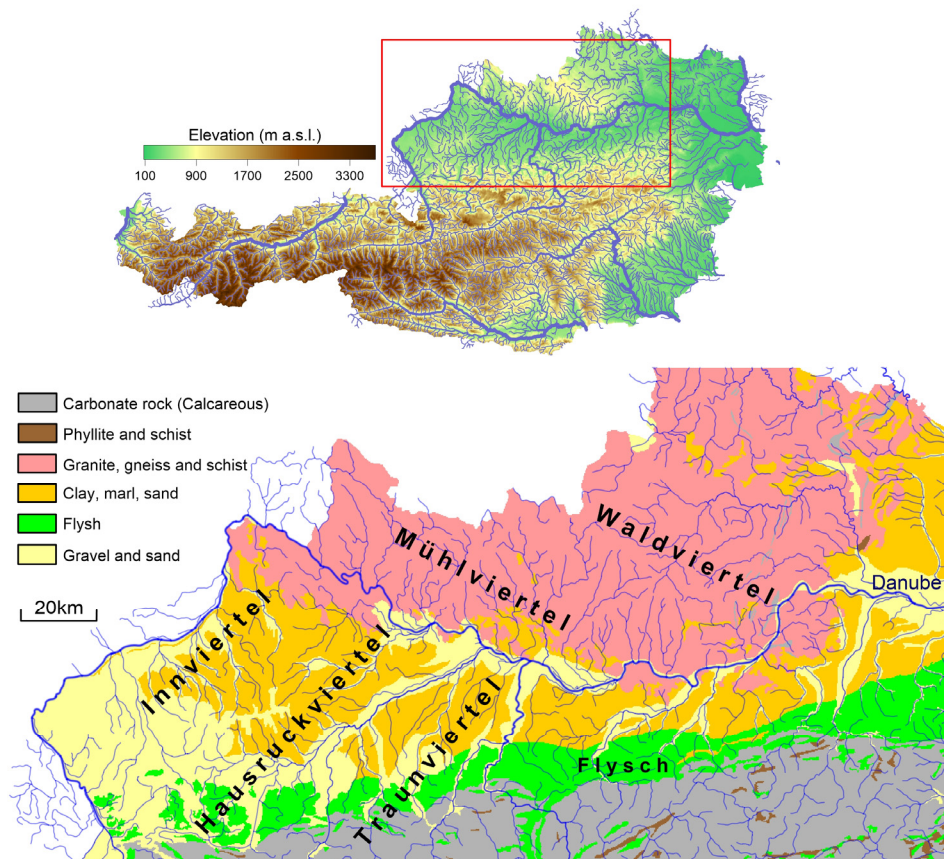
### Study region

There is a wide variety of flood generation mechanisms across Austria due to different hydrological, climatological and geological settings (e.g., Gaál et al., 2012; Parajka et al., 2010) which also challenge the analysis of flood peak-volume relationships. In order to reduce this complexity, here, as in Szolgay et al. (2015, 2016), we restricted our analysis to a geographically limited area, namely to the Northern Lowlands region of Austria (Fig. 1). The region covers approximately 20 000 km<sup>2</sup>, with elevations ranging from about 400 to 1500 m a.s.l. From the climatological point of view, the western parts of the region are mainly under the influence of Atlantic air flows from the West or North-West. Annual rainfall ranges from about 500 to 1500 mm/year with a decreasing trend from the West to the East. Floods occur both in the summer and in the winter seasons. Summer floods are either produced by synoptic weather systems or localised convective events (flash floods). Winter floods are usually induced by rain-on-snow processes when antecedent snowmelt saturates the soils and relatively low rainfall intensities may then cause significant floods. Snowmelt floods without rain contributions also occur but are less important due to the relatively low elevations.

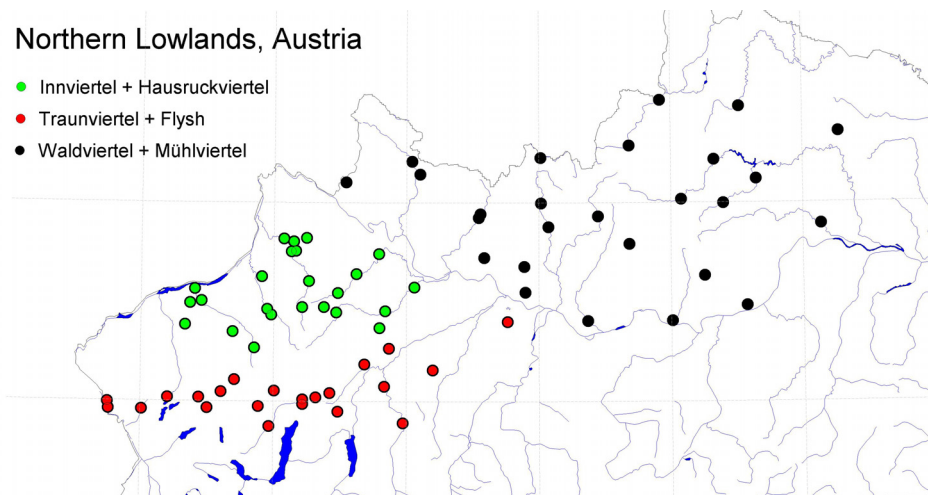
As in Szolgay et al. (2016), three subjectively delineated, spatially contiguous subregions were considered within the target region, based on geographical/geological units (Figs. 1 and 2):

1. Subregion #1: Innviertel and Hausruckviertel,
2. Subregion #2: Traunviertel and Flysch,
3. Subregion #3: Weinviertel and Mühlviertel.

Subregion #3 is located in the North, with a prevailing geology made of granite. In this area, floods appear throughout the year with prevailing synoptic flood types. The spring floods are apparently associated with snow as indicated by the daily cycle of some of the events. In this subregion, the meteorological forcing appears to be due to a range of processes, including convective storms in summer as indicated by the slim shape of some of the summer hydrographs. Subregion #2 is located mostly in the Southern-South-Western parts of the pilot region where the geology is made of marl and sandstone. The catchments have more summer floods and flash floods. Finally, in Subregion #1 in the North-Western parts of the pilot region, catchments have even more winter floods than the other subregions. The geology in this subregion is prevalently formed of gravel and sand.



**Fig 1.** Location of the pilot region in Austria and its prevailing geology.



**Fig. 2.** Location of subregions and the catchment outlets in the pilot region. Note that the colour coding for the subregions will remain the same throughout the whole paper.

### Runoff data

Runoff data from 72 catchments were used in this paper. These small and mid-sized catchments have areas in the range of 10.6 to 444.3 km<sup>2</sup> (median: 78.6 km<sup>2</sup>), while the range of their mean elevations is from 342 to 888 m a.s.l. (median: 571 m a.s.l.). The time resolution of the runoff data was one hour. The basic climatic and physiographic characteristics of the catchments are listed in Table 1.

### Classification of flood generation processes

The classification of flood generation processes used here builds on the flood database of Austria developed and described in detail in Merz and Blöschl (2003), Merz et al. (2006) and Merz and Blöschl (2009). For this database, individual flood events were isolated from the runoff records. As a first step of this procedure, base flow and direct runoff were separated by the recursive digital filter of Chapman and Maxwell (1996).

**Table 1.** The basic catchments characteristics used in this study. All sites = all sites included (72 sites with no regional delineation), Subregion #1 = Innviertel + Hausruckviertel (24 sites), Subregion #2 = Traunviertel + Flysh (22 sites) and Subregion #3 = Mühlviertel + Waldviertel (26 sites). *Area* – catchment area [km<sup>2</sup>], *MCE* – mean catchment elevation [m a.s.l.], *MCS* – mean catchment slope [–], *PFA* – percentage of forest area [–], *MAP* – mean annual precipitation [mm], *MRC* – mean runoff coefficient of maximum annual flood events [–], *MTSS* (*MTSF*, *MTSW*) – the mean time scale for the synoptic (flash, snowmelt) floods [h].

		<i>Area</i> [km <sup>2</sup> ]	<i>MCE</i> [m a.s.l.]	<i>MCS</i> [–]	<i>PFA</i> [–]	<i>MAP</i> [mm]	<i>MRC</i> [–]	<i>MTSS</i> [h]	<i>MTSF</i> [h]	<i>MTSW</i> [h]
All sites	maximum	444.3	888.0	0.33	1.00	1698.8	0.85	47.63	16.38	90.92
	median	78.6	571.0	0.09	0.30	952.6	0.49	26.40	9.62	36.06
	minimum	10.6	342.0	0.02	0.00	577.9	0.19	16.65	5.27	25.82
Subregion #1	maximum	361.8	598.0	0.14	0.46	1251.0	0.77	33.58	16.34	42.61
	median	67.4	450.5	0.10	0.13	979.8	0.60	22.64	8.69	32.93
	minimum	14.2	385.0	0.07	0.01	887.7	0.36	16.65	5.72	25.82
Subregion #2	maximum	444.3	839.0	0.33	1.00	1698.8	0.74	34.87	14.52	47.89
	median	55.8	571.0	0.10	0.34	1292.2	0.49	22.39	9.08	31.73
	minimum	12.0	342.0	0.02	0.00	862.9	0.20	17.34	5.27	26.22
Subregion #3	maximum	305.9	888.0	0.16	0.82	1042.3	0.85	47.63	16.38	90.92
	median	120.6	707.5	0.09	0.48	821.8	0.42	36.79	11.77	52.12
	minimum	10.6	480.0	0.03	0.26	577.9	0.19	22.48	6.77	34.50

Second, the start and the end of the flood events were identified based on criteria related to the base flow and direct flow at the beginning of the event, at the time of peak and at the end of the event. Finally, in some cases for snowmelt and rain-on-snow events a simple rainfall–runoff model (using hourly rainfall and snowmelt inputs) was fitted to the direct hydrograph to obtain more accurate estimates of event durations and volumes than when directly analysing the observed (and oscillating) direct flow hydrograph (see Fig. 4). In 72 catchments of the pilot region this procedure resulted in a total of 25 697 flood events, or an average of ~357 events per catchment.

The flood type classification of Merz and Blöschl (2003) was the starting point for the classification used in this paper. They defined five flood categories (long-rain floods, short-rain floods, flash floods, snowmelt floods and rain-on-snow floods) based on the meteorological situation (the spatial extent of the event, the intensity and the type of precipitation, solar radiation etc.) and the state of the catchment (antecedent rainfall, saturation of soils, snow cover, snowmelt etc.) before and during the individual flood events. Gaál et al. (2014) reduced the number of the flood process types from five to three by merging similar categories. Their three classes consist of synoptic floods (originally long-rain and short-rain floods), flash floods (no change in the classification) and snowmelt floods (originally rain-on-snow floods and snowmelt floods). Note that while for previous studies focused on Austria not all events were classified, in this paper we classified all flood events in the database for the pilot region. From this relatively huge dataset, meteorologically and hydrologically independent events were selected for this study. A flood event was considered independent from the previous event if at least one of the following conditions applied:

1. it began at least 7 days after the end of the previous event,
2. it followed a period of at least 7 days without significant rainfall ( $\leq 0.1$  mm/h).

The 7-day threshold was motivated by the observation that cyclonic situations in Central Europe do not usually persist longer than 7 days (e.g., Werner and Gerstengarbe, 2010), so events separated by longer time periods will likely have different atmospheric origins.

The flood event types were identified by first classifying events as snowmelt events, if a snow depth of at least 5 cm was observed in the highest elevation zone of the catchment. The snow depths were based on interpolated station readings from

Parajka et al. (2007). The altitudinal range of the zones was 200 m. From the remaining events, flash floods were selected if all of the following criteria applied:

1. event occurrence between May and September,
2. event duration of no more than 5 hours, and
3. air temperatures during event of at least 15 degrees of Celsius, as an indication of thermal convectivity.

All the remaining flood events were classified as synoptic floods. Note that the occurrence of synoptic flood types is not restricted to the summer period only; they can occur throughout the whole year (and these floods include short- and long-rain floods, cold and warm fronts and orography-induced events).

Before multivariate analysis, the flood peak and the corresponding volume data samples (for annual maxima of floods and associated volumes; for each flood type peaks and associated volumes, respectively) were statistically tested for stationarity and serial independence. Namely, the Mann–Kendall test (Kendall, 1955) was used to detect presence of a trend, while the Ljung–Box test (Ljung and Box, 1978) was used to check whether any of the first *m* autocorrelations of a time series is different from zero (therefore, the test is sometimes referred to as portmanteau test). As a result, no truly significant violation of both assumptions was found across the region and the few exceptions were further not investigated in detail here, given the regional character of the study, the known properties and the strength of the tests.

A comprehensive analysis of trends in annual maximum floods during last decades in Austria was carried out in Blöschl et al. (2011). From all analysed stations, only 17% of the gauged catchments showed increasing trends during 1976–2007 (with the majority of the stations showing no significant change) with a general tendency for increasing trends in the North, decreasing trends in the South, increasing trends in winter floods in the West and decreasing trends in the South-East of Austria.

Concerning our pilot region, no significant trends of floods were detected. At the Steyr and Traun Rivers, the magnitude of the events in years 1897 and 1899 exceeded all those observed during the 20<sup>th</sup> century. In Innviertel and the Mühlviertel regions, no trends in floods were detected. In Waldviertel, the 2002 events were the largest ones from the whole observation period. On the upper Danube River, upward trends in the small floods since 1850 can be observed, but no trends in the major floods were detected.

## Copula methods and analyses

Empirical copulas and nine frequently used one-parametric copulas from several classes were used in this study. The mathematical background is just briefly summarized since it is treated in detail in a number of papers and, therefore, not repeated here (e.g., Genest and Favre, 2007).

Copulas split the problem of constructing bivariate probability distributions into two parts: (1) the marginal one-dimensional distribution functions and (2) the dependence structure. These parts can be studied and estimated separately and then rejoined to form a joint distribution function. Formally, a bivariate joint distribution function can be written as

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

where  $F_X$  and  $F_Y$  are the respective marginal distribution functions of the random variables  $X$  and  $Y$  (which, in this paper, stand for flood peak and flood event volume, respectively). The term  $F_{XY}$  represents the joint distribution function of random vector  $(X, Y)$ , and  $C$  is a copula, i.e., a function  $C: [0, 1]^2 \rightarrow [0, 1]$  satisfying boundary conditions  $C(t, 0) = C(0, t) = 0$ ,  $C(t, 1) = C(1, t) = t$  (uniform margins) for any  $t \in [0, 1]$ , and the so-called 2-increasing property, which is analogous to the non-decreasing property of a cumulative distribution function in a univariate case. As such, a copula can be viewed as a standardized joint distribution function (Nelsen, 2006).

Usually, the marginal distribution functions  $F_X$  and  $F_Y$  are not known. Here they are considered solely on the basis of the observations of the random variables  $(X_i, Y_i), i = 1, \dots, n$ , using a corresponding empirical distribution function (sometimes referred to as plotting position formula):

$$F_X(x) = \sum_i \mathbf{1}(X_i \leq x) / (n+1) \quad (2)$$

An analogous relationship holds for  $F_Y$ . The individual values of  $F_X$  and  $F_Y$  are called pseudo-observations and are denoted as  $U_{j,i} (j = 1, 2, i = 1, \dots, n)$ . The empirical copula  $C_n$  is practically a two-dimensional extension of the univariate plotting position formula:

$$C_n(u_1, u_2) = \sum_i \mathbf{1}(U_{1,i} \leq u_1) \mathbf{1}(U_{2,i} \leq u_2) / (n+1) \quad (3)$$

In this paper, we are interested in the similarity (equality, in statistical sense) of empirical copulas of pairs of flood peaks and corresponding event volumes, which is tested by the approach of Remillard and Scaillet (2009). It is based on a Cramér-von Mises type of distance measure:

$$S = \left( \frac{1}{n_1} + \frac{1}{n_2} \right) \int_{[0,1]^2} [C_{n_1}(u_1, u_2) - C_{n_2}(u_1, u_2)]^2 \quad (4)$$

where  $C_{n_1}$  and  $C_{n_2}$  denote two empirical copulas constructed from samples of length  $n_1$  and  $n_2$ , respectively. The probability distribution of  $S$ , given that the null hypothesis ( $H_0$ : empirical copulas  $C_{n_1}$  and  $C_{n_2}$  come from the same – unknown – bivariate distribution) holds, is unknown and needs to be bootstrapped. The bootstrap simulations are performed here using the R package *TwoCop* (Remillard and Plante, 2012). The result of the similarity testing is a  $p$ -value, which is the percentage of how many simulations of the test statistic (under  $H_0$ ) exceeded the estimator from observations.

As in Szolgay et al. (2015, 2016) we used here the same nine symmetric (in the sense of non-exchangeability) copula types to represent the population of sample flood peak-volume relationships: Clayton, Frank, Galambos, Gumbel-Hougaard, Hüsler-Reiss, Joe, Normal, Plackett and Student-t. These one-parametric copula families are often used in flood studies (e.g., Bezak et al., 2014; Chowdhary et al., 2011; Karmakar and Simonovic, 2009; etc.). Archimedean copulas are usually favoured for their simplicity. Extreme-value (EV) copulas are particularly appropriate for modelling dependence between extremes of random variables. Elliptical copulas are simply the copulas of elliptically contoured distributions. All Archimedean (except for Frank) and EV copulas are non-symmetric with respect to the secondary diagonal, i.e., low and high extremes have different behaviour. Some of these copulas can model non-zero lower tail or upper tail dependence (Table 2, Fig. 3).

All the copulas are single parameter copulas except for the Student t-copula but, for consistency with the other copulas and for the reason of simplicity, we fixed the second parameter (the degrees of freedom) to 4 to make it effectively a single parameter copula and to distinguish it from the normal copula.

The parameter  $\theta$  controls the degree of association between peaks and volumes and can be directly related to the rank correlation coefficient (e.g., Nelsen, 2006). The parameter  $\theta$  was estimated from the flood peak-volume data by maximizing the so-called pseudo-likelihood function

$$L(\theta) = \sum_i \log [c_\theta(U_{1,i}, U_{2,i})] \quad (5)$$

where  $c_\theta$  denotes the copula density and  $U_{j,i} (i = 1, \dots, n, j = 1, 2)$  are the pseudo-observations of flood peak and volume.

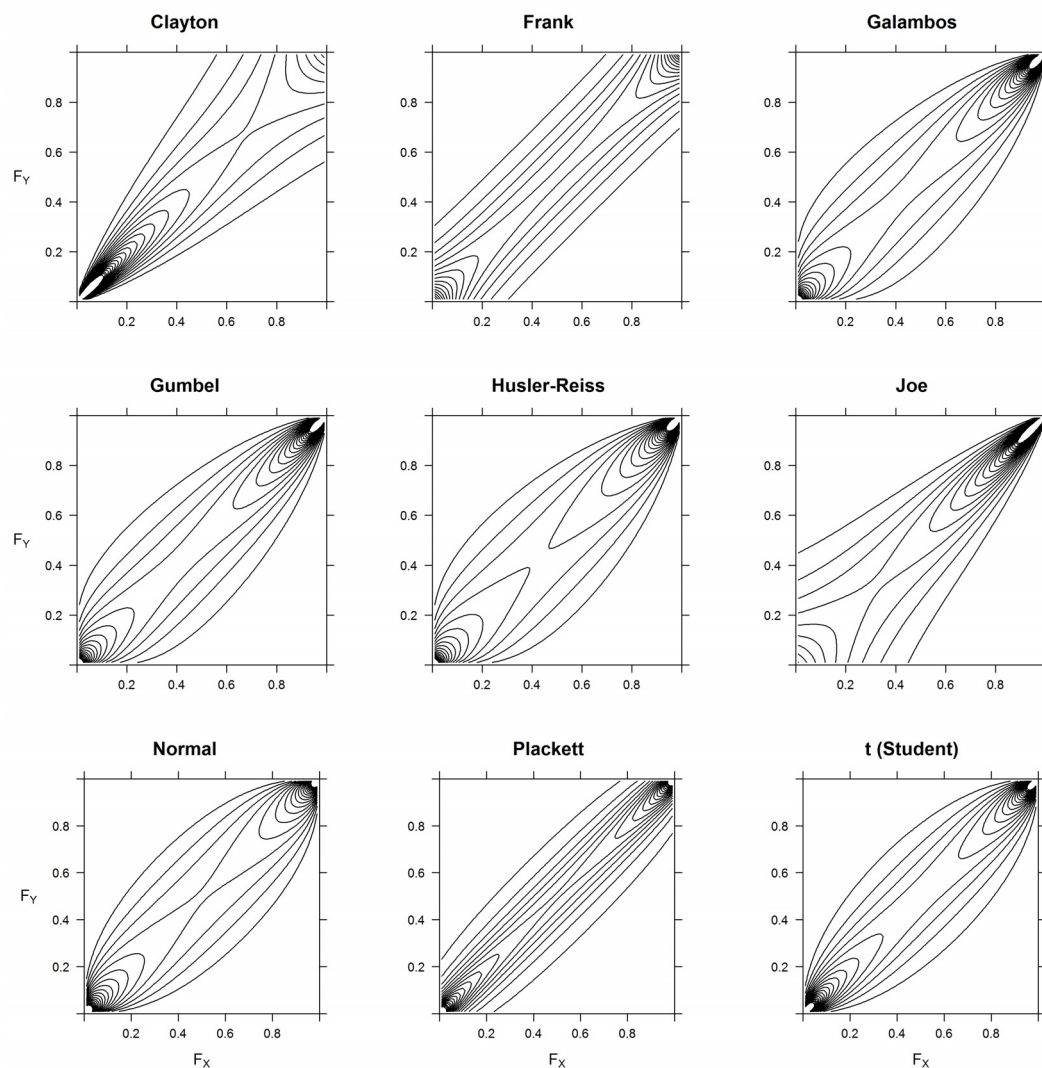
The goodness-of-fit of the copulas to the data was tested by one of the 'blanket' tests (as denoted and designed in Genest et al., 2009, to voice their general applicability with no necessity for strategic choice of any parameters, kernels or weights) with the Cramér-von Mises measure of distance

$$S_n = \sum_i [C_\theta(U_{1,i}, U_{2,i}) - C_n(U_{1,i}, U_{2,i})]^2 \quad (6)$$

between parametric copula  $C_\theta$  and empirical copula  $C_n$ . The probability distribution of  $S_n$ , given that the null hypothesis ( $H_0$ :  $C_\theta$  fits well) holds, is, again, unknown and has been bootstrapped.

**Table 2.** Characteristics of the 9 copula types used in this paper.

Copula type	Archimedean copula	Extreme value copula	Lower tail dependence	Upper tail dependence
<i>Clayton</i>	yes	no	yes	no
<i>Frank</i>	yes	no	no	no
<i>Galambos</i>	no	yes	no	yes
<i>Gumbel-Hougaard</i>	yes	yes	no	yes
<i>Hüsler-Reiss</i>	no	yes	no	yes
<i>Joe</i>	yes	no	no	yes
<i>Normal (Gaussian)</i>	no (elliptical)	no	no	no
<i>Plackett</i>	no	no	no	no
<i>t (Student)</i>	no (elliptical)	no	yes	yes



**Fig. 3.** The 9 copula types used in this paper. Contours indicate the density of the bivariate distribution.  $F_x$  and  $F_y$  relate to flood peaks and event volumes. Copula parameters in each case correspond to Kendall’s rank correlation coefficient of  $\tau = 0.67$ . Note that different values of  $\tau$  would produce different density of contours.

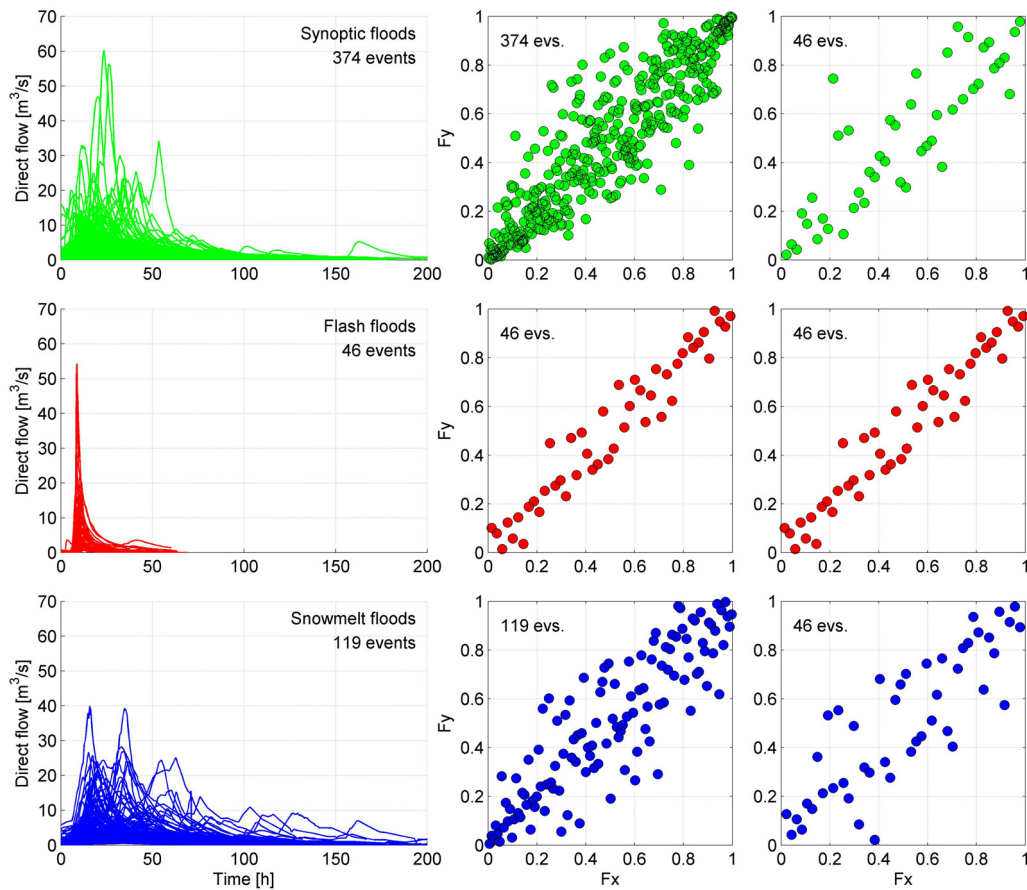
**RESULTS**

**Flood processes in the pilot region**

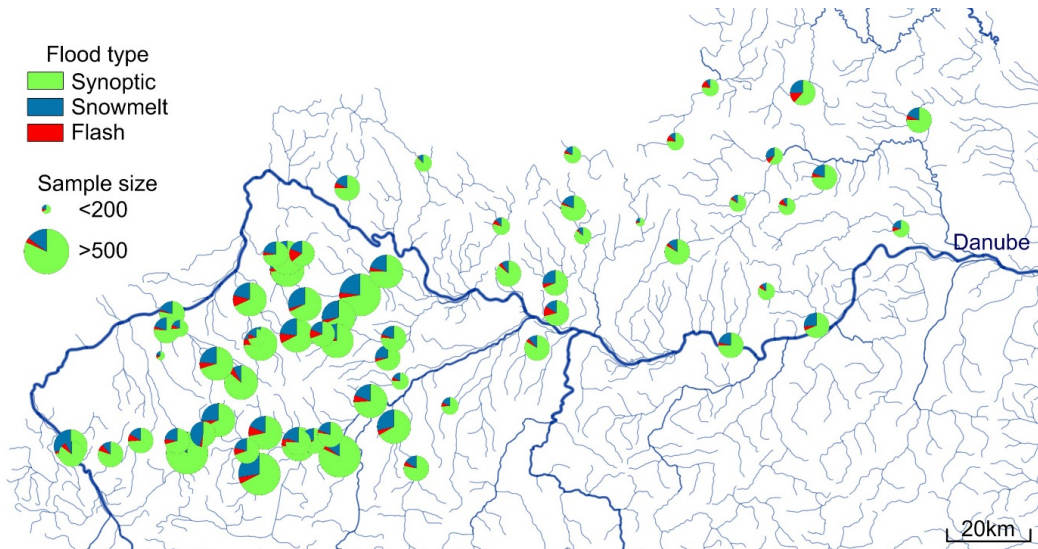
To illustrate the regional typology of hydrograph shapes and the properties of flood peak–volume relationships, which are at the base of the comparative analysis of the similarity of empirical copulas and suitability of theoretical copulas, Fig. 4 presents all flood hydrographs from the extended flood database for station St. Georgen in Attergau at the stream Dürre Agger, located in Subregion #2 (more precisely, the catchment has one

of the southernmost position in central parts of the Subregion ‘Traunviertel + Flysh’). The hydrographs of the flash floods show consistent peaky pattern and short durations. The hydrographs of the synoptic and snowmelt floods are more diverse. Some of the snowmelt floods show daily oscillations due to snowmelt. The right column of the panels of Fig. 4 shows a closer association of peak and volumes for the flash floods than for the other flood types, which is a reflection of the more consistent hydrograph patterns.





**Fig. 4.** Hydrographs of observed synoptic, flash and snowmelt floods (left) and peak-volume relationships (middle and right) for St. Georgen in Attergau / Dürre Agger (57.6 km<sup>2</sup> of catchment area).  $F_x$  and  $F_y$  are the pseudo-observations of the peaks and event volumes, respectively. The middle column shows peak-volume relationships in a unit square for all events, while in the right column, relationships for resampled subsamples are shown where each sample size equals to that of flash floods.



**Fig. 5.** Map of the study region and its subregions in Northern Austria with pie charts of per cent flood types of individual catchments as indicated by the colours. Size of the pie charts corresponds to total number of events per catchment.

The regional distribution of the per cent flood types for each catchment is shown in Fig. 5 and Table 3. There is a clear prevalence of synoptic floods (74% of all events). Snowmelt and flash floods constitute 19% and 7%, respectively. Locally, the percentage of the flood types ranges between 51 and 87% for the case of synoptic floods, and 8 and 46%, and 2 and 17% for

the snowmelt and flash floods, respectively. The relatively small number of snowmelt floods is due to the modest elevations in the region. The small number of flash floods is partly related to the sizes of the gauged catchments (median of 78.6 km<sup>2</sup>). If data for smaller catchments were available, the percentage of flash floods would likely increase.

**Table 3.** Statistics of the number of identified flood events, stratified by regions and process types. The last three columns of the table indicates statistics of local data that is expressed in %.

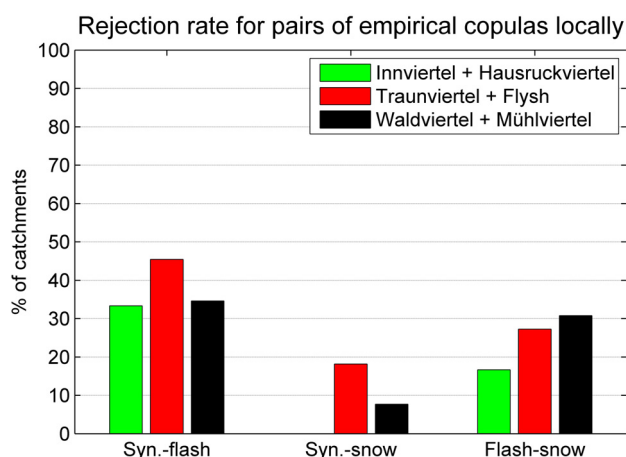
		In absolute values				In % (locally)		
		Total	Synoptic	Flash	Snowmelt	Synoptic	Flash	Snowmelt
All sites	minimum	181	123	6	18	51.4	1.6	7.8
	median	333	249	20	63	74.3	5.8	19.4
	maximum	549	396	70	150	87.2	16.7	45.9
	total	25 697	19 093	1 733	4 871	---	---	---
Subregion #1	minimum	181	148	6	27	71.9	2.1	9.1
	median	405	319	23	61	76.2	5.4	17.5
	maximum	508	396	70	85	86.3	14.9	22.1
	total	9 396	7 325	596	1 475	---	---	---
Subregion #2	minimum	212	168	7	18	51.4	1.6	7.8
	median	392	286	25	74	74.6	6.3	17.6
	maximum	549	389	63	150	87.2	16.7	45.9
	total	8 788	6 448	652	1 688	---	---	---
Subregion #3	minimum	182	123	6	29	61.4	2.0	9.5
	median	293	203	15	64	70.3	5.1	23.3
	maximum	341	272	45	91	79.8	13.7	30.6
	total	7 513	5 320	485	1 708	---	---	---

The total number of events identified does show a modest spatial pattern. It is largest in the North-West (the Innviertel + Hausruckviertel region), smaller in the South-West (Traunviertel + Flysh) and the smallest in the North-East (Mühlviertel + Waldviertel). The average number of flood events is at the similar level in the two regions on the West (~392 and ~399), while the lowest average can be found in the Mühlviertel + Waldviertel region (~289). These differences are related to the spatial distribution of the typical time scales of floods that have been defined by Gaál et al. (2012) as the ratio of flood volume and peak flow and are presented in the Table 1. The medians of mean flood time scales for the whole region are about 26, 10 and 36 hrs for synoptic, flash and snowmelt floods, respectively. The shorter the flood time scales, the larger the number of independent events that can be identified. The differences in the flood time scales are related to differences in both climate and catchment response characteristics (Gaál et al., 2012).

**Comparison of the similarity of empirical copulas for different flood types locally**

In this part of the analysis, we were interested in whether different flood types, for the same catchment, were distinguishable in terms of their empirical flood peak–volume copulas. The analysis was carried out for each catchment separately and the flood samples of process types were compared pairwise, i.e., synoptic floods vs. flash floods, synoptic floods vs. snowmelt floods and flash floods vs. snowmelt floods in each subregion. A similar analysis was first carried out for the pilot region as whole in Szolgay et al. (2016).

The results in the spatially contiguous subregions with different geology, which are presented in Fig. 6, show the same general pattern as found in Szolgay et al. (2016) for the region as a whole, i.e., that synoptic and snowmelt floods could belong more often to the same unknown copula than is the case for the other combinations in each subregion. This feature may be illustrated in the case of the catchment shown in Fig. 4. When the full data samples are used (middle column in Fig. 4), the comparison of synoptic and flash floods gives a *p*-value of 0.004, indicating rejection of the null hypothesis (i.e., the processes are different in terms of their empirical copulas). The same holds true for the comparison of snowmelt and flash floods with *p* = 0.021. On the other hand, the comparison of synoptic and snowmelt processes results in a *p*-value of 0.93;

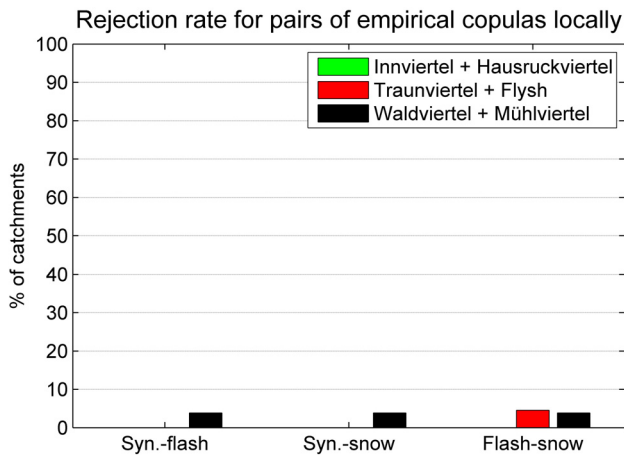


**Fig. 6.** Results of the comparison of the empirical copulas locally for the original data: per cent ratio of the catchments in each subregion, where the equivalence of the pairs of empirical copulas was rejected at the significance level of  $\alpha = 0.1$ .

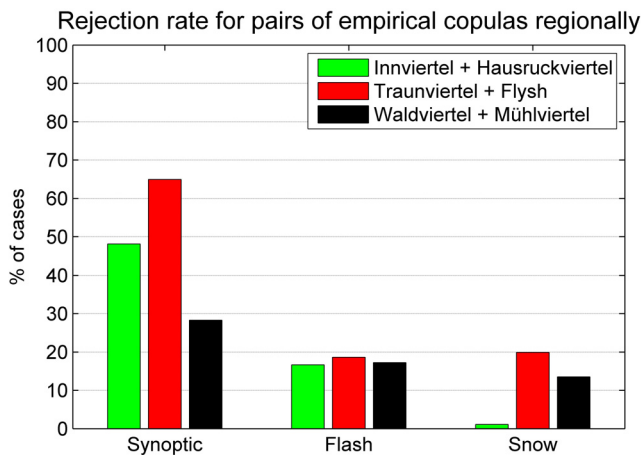
i.e., one cannot reject the null hypothesis about the equality of empirical copulas. This is also suggested by a pure visual comparison of the scatterplots that indicate that the synoptic and snowmelt floods are more similar to each other than the remaining process pairs; or, in other words, flash floods tend to be more dissimilar from both the synoptic and snowmelt floods (in terms of their empirical copulas). This could be intuitively partly related to a much stronger upper tail dependence of flash floods and their specific (similar) hydrograph shapes (Gaál et al., 2014). However, the overall rejection rates are rather small and the analysis has not brought really conclusive results for the subsequent selection of theoretical copula models for engineering design, which would logically follow such an analysis in practice. Moreover, the relatively small number of events for such type of analysis in general and the different number of events in the respective flood types in particular, may also play a role.

To illustrate the potential effect of the sample size, subsamples were drawn from the individual (full) data samples. Since at each site, the flash flood category was the one with the lowest number of the flood events, this number was used to define the size of the subsets both from the synoptic and the snowmelt

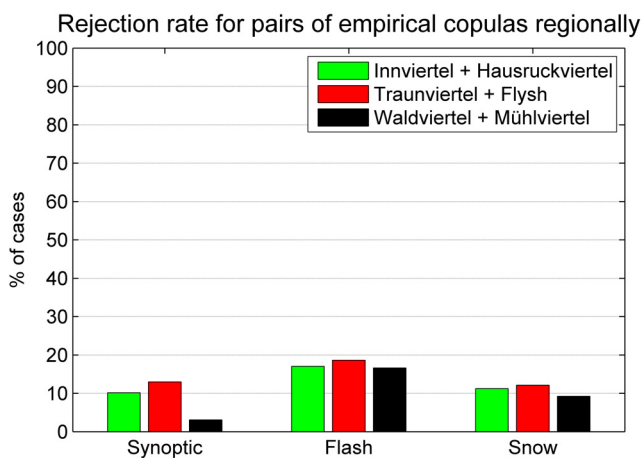




**Fig. 7.** Results of the comparison of the empirical copulas locally for the resampled events: per cent ratio of the catchments in each subregion, where the equivalence of the pairs of empirical copulas was rejected at the significance level of  $\alpha = 0.1$ .



**Fig. 8.** Rejection rates of the comparison of the empirical copulas (based on the original data) subregionally at the significance level of  $\alpha = 0.1$ .



**Fig. 9.** Rejection rates of the comparison of the empirical copulas (based on the resampled data) subregionally at the significance level of  $\alpha = 0.1$ .

floods at each site. Ten of the total number of 100 subsamples were then kept, where the proximity as expressed by means of the difference in Kendall's  $\tau$  between the particular subsample and the one corresponding to all data was small. Preserving the rank correlation between floods and volumes means that at each site, we only eliminate the effect of the different sample sizes in a comparison the empirical copulas. The final  $p$ -values are then estimated as the median of all 10 subsamples kept for the analysis.

Figure 7 presents the results of the test which only compares empirical copulas constructed on the basis of the same number of data pairs at each site. As a result, even the rather low significance of dissimilarity from Fig. 6 is practically lost. There are basically no cases where the empirical copulas were found to be significantly different. This can, again, be intuitively illustrated in the right column of Fig. 4. None of the  $p$ -values from the comparison of flash vs. synoptic (0.327), flash vs. snowmelt (0.253) and of snowmelt vs. synoptic floods (0.915) is significant at the level of  $\alpha = 0.1$ . A visual comparison of the scatter plots of subsamples also suggests that the synoptic and snowmelt floods are not as different from the flash floods as in the case of full data sets; therefore, the statistical test is unable to come to significant conclusions.

Overall, the analysis suggests that flash floods are more often distinguishable from the synoptic and snowmelt floods in terms of their empirical peak-volume copulas than the synoptic and the snowmelt floods from each other. In other words, the copulas of synoptic floods and snowmelt floods are slightly more similar than the other process pairs. The length of the data series is, however, an objective factor, which cannot be overcome in the framework of comparative hydrology, when using only observed data (even in the case of a relatively large sample of data pairs of all independent floods). These conclusions are similar to those in Szolgay et al. (2016) formulated for a geologically heterogeneous region; herein, they were shown to be true in three different geological conditions.

### A comparison of similarity of empirical copulas for each flood type regionally

Next, a comparison of the empirical copulas for each flood type was performed in each subregion. We were interested in the question, whether different catchments with the same flood type are distinguishable in terms of their empirical peak-volume copulas (Fig. 8). Within the set of sites in each subregion, and for each flood process separately, pairwise comparisons were carried out. The colour bars of Fig. 8 indicate the per cent ratio of the cases where the null hypothesis about the equality of the empirical copulas for the given flood process was rejected in the given subregion.

It can be seen that the empirical copulas of synoptic floods are the least similar between the catchments within all three subregions. This is a similar behaviour as observed in Szolgay et al. (2016) for the region as a whole. Here it is shown to be the same in the three geologically homogeneous subregions. Spatially, the phenomenon is more pronounced in the South-West than in the North-Western and North-Eastern parts of the pilot region.

The highest rejection rates for the synoptic flood processes may seem to be surprising at the first glance. Nevertheless, the higher ratio of rejections of the synoptic events across different pairs of sites in all three subregions could be explained by the more complex temporal rainfall structure (that is the mix of long- and short-rain flood processes in the dataset), resulting in a lower degree of similarity between the different events. In the

case of flash and snowmelt floods, the difference between the process types is smaller; the analysis suggests that most catchment pairs, for a given flood type are practically not distinguishable in terms of their empirical peak-volume copulas regardless of the geological conditions. Moreover, dissimilarity may become more pronounced, when the sample size is larger, as it is in the case of synoptic floods when compared to the other two flood types. Since the similarity/dissimilarity patterns in the results seem to more or less similar for each geological setting, they could be attributed to the length of the available data series. Therefore, similarly as in the previous sub-section, an experiment with the resampled data sets was conducted. For a given catchment, the numbers of synoptic, flash and snowmelt floods were equal. Within the set of such sites in each subregion, and for each flood process separately, pairwise comparisons were carried out. The colour bars in Fig. 9 indicate the per cent ratio of the cases where the null hypothesis about the equality of the empirical copulas for the given flood process was rejected in the given subregion.

Overall, the analysis suggests that most catchment pairs, for a given flood type, are not distinguishable in terms of their empirical peak-volume copulas in the resampled data space. The results also show that the differences in the rejection rates between the process types are small for all three geological settings, although the flash floods do show slightly less spatial homogeneity.

The experiment with the subsampled data sets, again, suggests that the length of the data series is a factor, which cannot be neglected in the framework of comparative hydrology and is restricting the usefulness of such an approach in practice. A more detailed analysis of the phenomenon is beyond the goals of the present study and could possibly be investigated only using stochastically simulated data (e.g., Grimaldi et al., 2016).

### The effect of the sample size on the goodness-of-fit of copula types

Szolgay et al. (2016) examined the question what copula types fit the flood peak-volume data of individual catchments best, separately for seasonal floods and the three individual flood types, in the whole region and the subregions, using the full data set of independent floods. It was concluded that modeling dependence structure by treating flood processes separately in a regional context may prove beneficial with respect to narrowing the choice of acceptable models. However, results indicated that even the relatively large sample size of the process-based dataset of independent flood events (when compared to sample sizes of the commonly used pairs of annual maxima of flood peaks and respective event volumes) may not help to overcome some difficulties in restricting the choice of acceptable models in engineering studies. This problem is further examined here.

As a basis for comparison and for the illustration of the usefulness of using a larger dataset, in Fig. 10 the goodness-of-fit tests of the nine copula types at the 72 catchments for annual maximum floods are shown. These annual maxima of floods and their respective volumes are derived using different approaches: (i) in a traditional way, i.e., with no information about the flood process types, and (ii) for the annual maxima of each flood process type separately. The copula types (each bar represents a copula type) are organized alphabetically. Black colour indicates the percentage of  $p \leq 0.05$ , i.e., a rejection of the null hypothesis  $H_0$ .

For annual maxima of flood events, the general pattern of the overall regional acceptance rate of particular models is in general the same as in the case of all independent floods reported in Szolgay et al. (2016); however, the rejection rates are

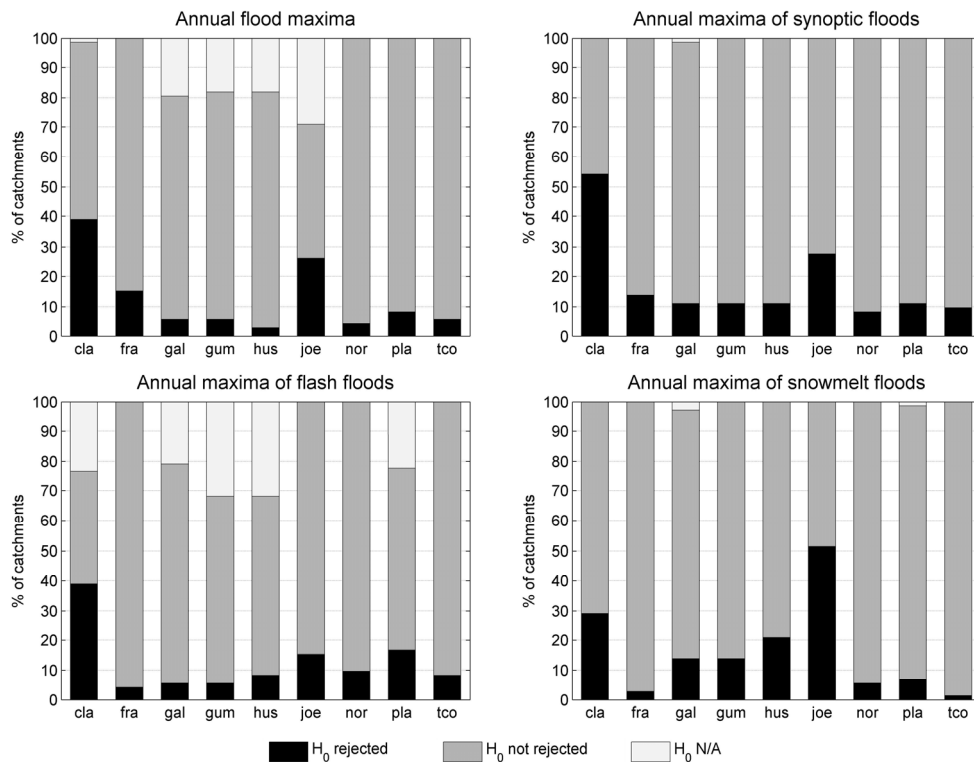
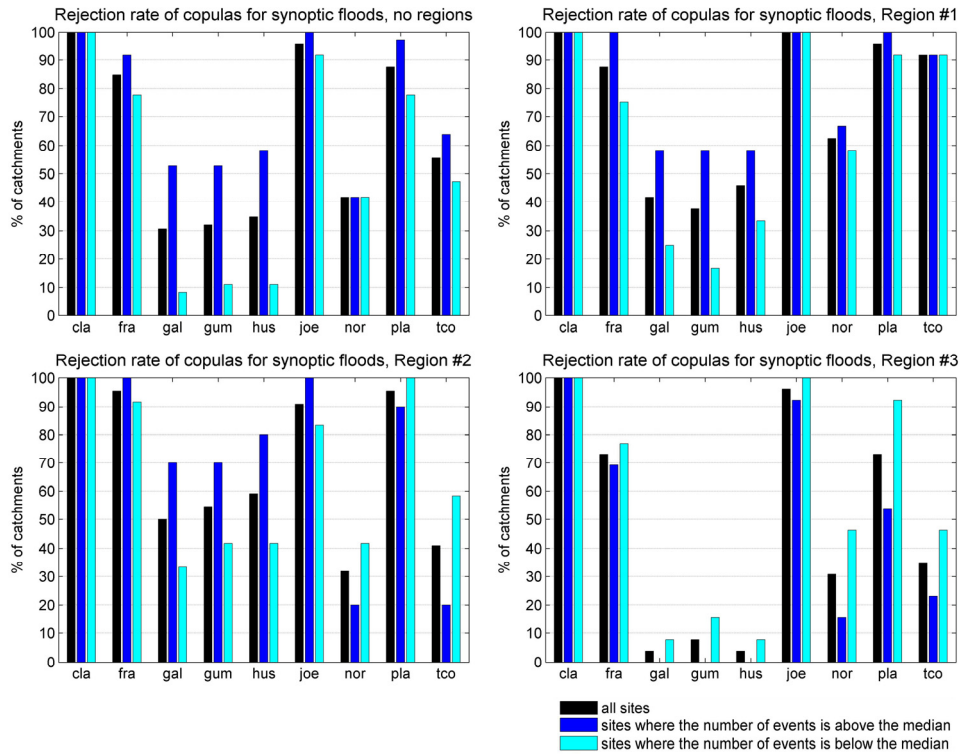
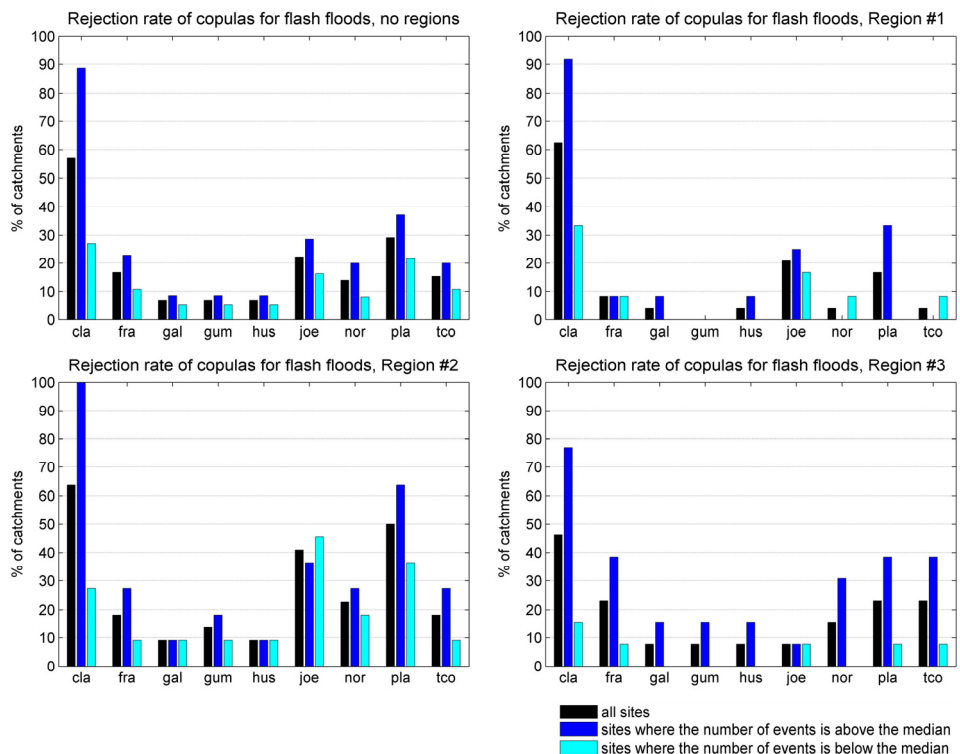


Fig. 10. Results of the goodness-of-fit tests of the nine copula types for annual maxima of floods. The black colour indicates the percentage of  $p \leq 0.05$ , i.e., a rejection of the null hypothesis  $H_0$ .



**Fig. 11.** Rejection rates of the goodness-of-fit tests (at the significance level of  $\alpha = 0.05$ ) of floods from the database of all independent floods for the **synoptic floods** in the whole pilot region and the three subregions separately.



**Fig. 12.** The same as in Fig. 11 but for the **flash floods**.

lower in the whole region in almost each case. (Note, that in some cases the available data was not sufficient to reliably fit a particular model.) Except for the Clayton and Joe copulas, all models could succeed in a particular engineering case study. This is to be attributed to the low number of events and less to

the mixture of events types in the data samples of annual flood maxima, as it will be shown below. This can have severe consequences, when the copulas are used for prediction, and the prediction uncertainty due to the model choice would not be taken into account in the design.

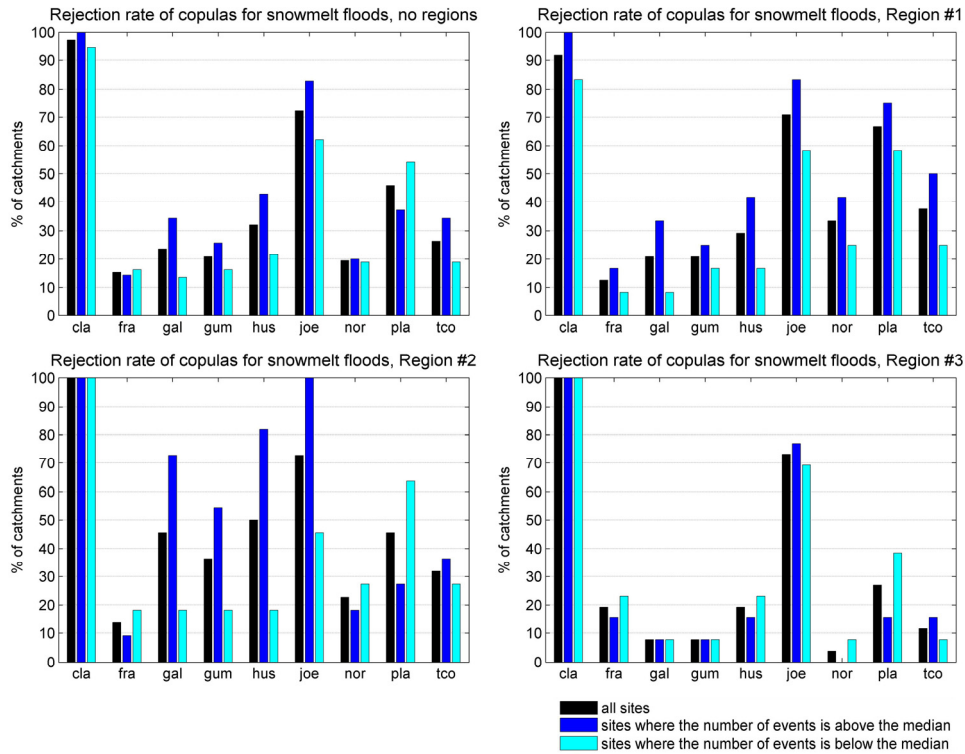


Fig. 13. The same as in Fig. 11 but for the snowmelt floods.

In Figs. 11–13 we show the results of the goodness-of-fit tests of floods from the database of all independent floods for the three flood types in the region and also for the three subregions separately, respectively. The unique feature of Figs. 11–13 is that beyond the sets including all catchments (for the particular flood type and region/subregion), also halves of these sets constructed on the basis of the median value of the sample sizes are analysed. For instance, in Subregion #1 that consists of 24 catchments, the median size of the synoptic floods is 319. The rejection rates for a particular copula type in Fig. 11 are therefore evaluated for (i) the entire set of 24 catchments (black bars), (ii) for the set of 12 catchments where the number of synoptic flood events exceed 319 (dark blue bars), and (iii) similarly, for the set of 12 catchments with the number of synoptic flood events below 319 (light blue bars). The colour bars indicate the percentage of  $p \leq 0.05$ , i.e., a rejection of the null hypothesis  $H_0$ . The copula types (each column represents a copula type) are organized alphabetically.

The rationale behind such an analysis is, again, to examine the effect of the sample size on the rejection rates; however, only using the existing ‘real’ data, without any resampling experiment (see the previous two sub-sections). Note that the black bars only serve as a basis for comparison since they repeat the results from Szolgay et al. (2016). The analyses suggest that, for synoptic and flash floods, the extreme value copulas (Galambos, Gumbel and Hüsler-Reiss) tend to perform better than the others. For snowmelt floods, the Frank copula shows the best fit. On the other hand, the Clayton and the Joe copula often show poor performance.

When comparing the rejection rates for cases with sample sizes above/below the median we can conclude, that for larger samples (i.e., for the case ‘above the median’), in general, the selectivity of the test increased (with a few exception in the case of snowmelt floods in Subregions #2 and #3). This also means, that with the increase of information on the variability

of flood typology we cannot give a straightforward recommendation for a regionally/subregionally acceptable model (within the region and set of models considered here). This fact also indicates that less data really broadens the model choice both in the region as a whole and in its subregions. On the contrary, for the practitioners, this also means that the set of models considered here (and preferred due their relative simplicity in many copula studies) may not be adequate for the variety of flood shape types and other copula models may be needed to be considered. Alternative copula models include, e.g., *vine copulas*, which allow for constructing a multivariate copula based on the mixing of bivariate copulas (e.g., Gräler et al., 2013; Pham et al., 2015; Vernieuwe et al., 2015) or *entropy copulas*, which integrate the concept of copulas with the principle of maximum entropy (i.e., the entropy variables are mutually independent from each other; AghaKouchak, 2014).

## DISCUSSION AND CONCLUSIONS

The analysis suggests that, when empirical copulas for different flood processes are compared locally at each site, flash floods are more often distinguishable from synoptic and snowmelt floods than are synoptic and snowmelt floods from each other in the given region and its subregions of Austria. It is important to note that the sample size played a key role in deciding whether empirical copulas can be regarded as different or not. In the answer to the first science question when the test of Remillard and Scaillet (2009) was adopted for the full data sets, we observed significantly different empirical copulas in a number of cases. When subsamples were drawn to verify the conclusions, however, the test showed decreased ability to yield significant results. These findings could support the choice of common theoretical models in the region/subregions in engineering applications or support the merging data from similar catchment to aid that. It is therefore concluded that analyses

like this should rather be based on data sets (1) that are much larger, and (2) which have comparable sizes for different flood types. In the current analysis, the samples of flash floods did not meet any of the suggested criterions. For the future, it is advised to analyse data sets that are more balanced in terms of sample size, e.g., data gained in simulation studies through the combination of stochastic rainfall generators with rainfall-runoff models (Grimaldi et al., 2016).

Therefore, we conclude in general, that the findings (along with the results of the goodness-of-fit test of copula types) suggest that there is less value in the concept of process-based comparative at-site analysis of empirical peak-volume relationships than may be expected. The stratification of the data by processes will invariably reduce the sample size, so some spatial pooling or data extension by simulation will be required in practical applications in order to limit the sampling uncertainty.

The second science question discussed the similarity of empirical peak-volume dependence structures between catchments for a given flood type that were analysed regionally and sub-regionally. The outcomes of the analysis suggest that, for a given flood type, the majority of catchment pairs is not distinguishable in terms of their empirical peak-volume copulas. This finding supports some doubts whether regional process-based pooling of catchments can be attempted only on the basis of flood process types based on real life sizes of data samples.

The third science question revolved around copula types that are most suitable for a given flood process type. Small differences between copula types were found here for the case of annual maxima of floods. The differences in suitable copula types between flood processes do suggest that there is (limited) merit in treating the flood types separately when analysing and estimating flood peak-volume copulas and using all independent floods. Obviously, smaller sample sizes make it difficult to distinguish between copula types. As a general recommendation resulting from this study, it is advisable to select models from the extreme value class of copulas (in the given region).

Answers to the fourth question support our conclusion formulated in Szolgay et al. (2016) with new and spatially differentiated results with respect to the usefulness of a comparative treatment of the model choice problem. Based on the results, it can be concluded that modeling dependence structure by treating flood processes separately in a regional context may, to a limited extent, prove beneficial with respect to narrowing the choice of acceptable models; nevertheless, it was also shown that beside the more detailed differentiation of the flood types and subregions, the sample size introduces into the selection of the model a larger degree of uncertainty than expected in Szolgay et al. (2015, 2016), which does not make the task easier for an analyst in engineering studies in practice.

As shown in a comparative hydrology framework, the choice of the copula model that best fits the observed data and is regionally acceptable in terms of flood typology, is not a trivial issue, even if more than statistical aspects are taken into consideration since the lack of sufficient data makes the analysis difficult. It was also shown that with increasing data availability the commonly used model choice needs to be revisited and models which would better reflect the variability in flood wave shapes have to be tested. This supports Favre et al. (2004) and Serinaldi and Kilsby (2013), who emphasized that further work is needed to choose the best copulas capable of reproducing the dependence structure of multivariate hydrological variables.

Uncertainties inherent in the copula-based bivariate frequency analysis itself (caused, among others, also by the relatively small sample sizes for consistent copula model selection, or upper tail dependence characterization, etc.) may not be over-

come in the scope of such a regional comparative analysis even in the case of using all available independent flood events differentiated by processes. Based on this comparative study and results of other more advanced studies (e.g., Serinaldi, 2013, 2015) it can be concluded that if reliable predictions are required for an important engineering application, the benefits of regional bivariate frequency analysis methods could be further explored (e.g., Ben Aissia et al., 2015) or the potential of the combination of rainfall generators, rainfall-runoff models, analysis of historical floods and advanced statistics considering uncertainty might be utilized as, e.g., in Grimaldi et al. (2016).

*Acknowledgements.* We would like to thank the Austrian Academy of Sciences (International Strategy for Disaster Reduction Programme, IWHRE2008), the ERC (Advanced Grant on Flood Change) and the FWF (project no. P 23723-N21) for financial support. This research was also supported by the Slovak Research and Development Agency under contract No. APVV-15-0497, by the Slovak Grant Agency VEGA under the project no 1/0776/13 and by the project IMPALA (FP7-PEOPLE-2011-IEF-301953) of the Marie Curie Intra European Fellowship within the 7th European Community Framework Programme. Furthermore, the support of the Competence Center for SMART Technologies for Electronics and Informatics Systems and Services, ITMS 26240220072, funded by the Research & Development Operational Programme from the ERDF is acknowledged. The authors would also like to thank to Francesco Serinaldi for a very inspiring open discussion on the topic during a review of another paper.

## REFERENCES

- AghaKouchak, A., 2014. Entropy-copula in hydrology and climatology. *Journal of Hydrometeorology*, 15, 2176–2189. doi:10.1175/JHM-D-13-0207.1.
- Báčová Mitková, V., 2012. Vzájomný vzťah objemu a dĺžky trvania povodňových vln. [The relationship between volume of the flood wave and the time duration of flood events]. *Acta Hydrologica Slovaca*, 13, 1, 165–174. (In Slovak.)
- Báčová Mitková, V., Halmová, D., 2014. Joint modeling of flood peak discharges, volume and duration: a case study of the Danube River in Bratislava. *Journal of Hydrology and Hydromechanics*, 62, 3, 186–196. doi:10.2478/johh-2014-0026.
- Ben-Aissia, M.-A., Chebana, F., Ouarda, T.B.M.J., Roy, L., Desrochers, G., Chartier, I., Robichaud, É., 2012. Multivariate analysis of flood characteristics in a climate change context of the watershed of the Baskatong reservoir, Province of Québec, Canada. *Hydrological Processes*, 26, 130–142. doi:10.1002/hyp.8117.
- Ben Aissia, M.-A., Chebana, F., Ouarda, T.B.M.J., Bruneau, P., Barbet, M., 2015. Bivariate index-flood model: case study in Québec, Canada. *Hydrological Sciences Journal*, 60, 2, 247–268. doi:10.1080/02626667.2013.875177.
- Bezák, N., Mikoš, M., Šraj, M., 2014. Trivariate frequency analyses of peak discharge, hydrograph volume and suspended sediment concentration data using copulas. *Water Resources Management*, 28, 8, 2195–2212. doi:10.1007/s11269-014-0606-2.
- Blöschl, G., Viglione, A., Merz, R., Parajka, J., Salinas, J.L., Schöner, W., 2011. Auswirkungen des Klimawandels auf Hochwasser und Niedrigwasser. [Climate impacts on floods and low flows]. *Österreichische Wasser- und Abfallwirtschaft*, 63, 1, 21–30. doi:10.1007/s00506-010-0269-z. (In German.)



- Chapman, T.G., Maxwell, A.I., 1996. Baseflow separation—comparison of numerical methods with tracer experiments. In: 23<sup>rd</sup> Hydrology and Water Resources Symposium: Water and the Environment, Natl. Conf. Publ., 96/05, Inst. of Eng., Barton, A.C.T., Australia, pp. 539–545.
- Chebana, F., Ouarda, T.B.M.J., 2007. Multivariate L-moment homogeneity test. *Water Resources Research*, 43, W08406. doi:10.1029/2006WR005639.
- Chebana, F., Ouarda, T.B.M.J., 2009. Index flood-based multivariate regional frequency analysis. *Water Resources Research*, 45, W10435. doi 10.1029/2008WR007490.
- Chowdhary, H., Escobar, L.A., Singh, V.P., 2011. Identification of suitable copulas for bivariate frequency analysis of flood peak and flood volume data. *Hydrology Research*, 42, 2–3, 193–216. doi:10.2166/nh.2011.065.
- Dalrymple, T., 1960. Flood frequency analysis. U.S. Geological Survey Water-Supply Paper, 1543-A, 80 p.
- Favre, A.-C., El Adlouni, S., Perreault, L., Thiémondge, N., Bobée, B., 2004. Multivariate hydrological frequency analysis using copulas. *Water Resources Research*, 40, W01101. doi:10.1029/2003WR002456.
- Gaál, L., Szolgay, J., Kohnová, S., Parajka, J., Merz, R., Viglione, A., Blöschl, G., 2012. Flood timescales: Understanding the interplay of climate and catchment processes through comparative hydrology. *Water Resources Research*, 48, 4, W04511. doi:10.1029/2011WR011509.
- Gaál, L., Kohnová, S., Szolgay, J., 2013. Regional flood frequency analysis in Slovakia: Which pooling approach suits better? In: Klijn, F., Schreckendiek, T. (Eds.): *Comprehensive Flood Risk Management: Research for Policy and Practice*. London, CRC Press/Balkema, pp. 27–30.
- Gaál, L., Szolgay, J., Kohnová, S., Hlavčová, K., Parajka, J., Viglione, A., Merz, R., Blöschl, G., 2014. Dependence between flood peaks and volumes – A case study on climate and hydrological controls. *Hydrological Sciences Journal*, 60, 6, 968–984. doi:10.1080/02626667.2014.951361.
- Ganguli, P., Reddy, M.J., 2013. Probabilistic assessment of flood risks using trivariate copulas. *Theoretical and Applied Climatology*, 111, 341–360. doi:10.1007/s00704-012-0664-4.
- Genest, C., Favre, A.-C., 2007. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, 12, 4, 47–368. doi: 10.1061/(ASCE)1084-0699(2007)12:4(347).
- Genest, C., Rémillard, B., Beaudoin, D., 2009. Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, 44, 199–213. doi:10.1016/j.insmatheco.2007.10.005.
- Gräler, B., van den Berg, M.J., Vandenberghe, S., Petroselli, A., Grimaldi, S., De Baets, B., Verhoest, N.E.C., 2013. Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation. *Hydrology and Earth System Sciences*, 17, 1281–1296. doi:10.5194/hess-17-1281-2013.
- Grimaldi, S., Serinaldi, F., 2006. Asymmetric copula in multivariate flood frequency analysis. *Advances in Water Resources*, 29, 8, 1155–1167. doi: 10.1016/j.advwatres.2005.09.005.
- Grimaldi, S., Petroselli, A., Salvadori, G., De Michele, C., 2016. Catchment compatibility via copulas: A non-parametric study of the dependence structures of hydrological responses. *Advances in Water Resources*, 90, 116–133. doi:10.1016/j.advwatres.2016.02.003.
- Hosking, J.R.M., Wallis, J.R., 1997. *Regional Frequency Analysis: An Approach Based on L-moments*, Cambridge University Press, Cambridge, UK.
- IH, 1999. *Flood Estimation Handbook*. Institute of Hydrology: Wallingford, UK.
- Karmakar, S., Simonovic, S.P., 2009. Bivariate flood frequency analysis. Part 2: A copula-based approach with mixed marginal distributions. *Journal of Flood Risk Management*, 2, 32–44. doi:10.1111/j.1753-318X.2009.01020.x.
- Kendall, M.G., 1955. *Rank Correlation Methods*. Hafner Publishing, New York.
- Kohnová, S., Szolgay, J., 1999. Regional estimation of design summer flood discharge in small catchments of northern Slovakia. In: Gottschalk, L., Olivry, C., Reed, D., Rosbjerg, D. (Eds.): *Hydrological Extremes: Understanding, Predicting, Mitigating*. IAHS publ. 255, IAHS Press, Wallingford, pp. 265–268.
- Ljung, G.M., Box, G.E.P., 1978. On a measure of lack of fit in time series models. *Biometrika*, 65, 297–303. doi:10.1093/biomet/65.2.297.
- Merz, R., Blöschl, G., 2003. A process typology of regional floods. *Water Resources Research*, 39, 12, 1340–1347. doi:10.1029/2002WR001952.
- Merz, R., Blöschl, G., 2009. A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria. *Water Resources Research*, 45, 1, W01415. doi:10.1029/2008WR007163.
- Merz, R., Blöschl, G., Parajka, J., 2006. Spatio-temporal variability of event runoff coefficients. *Journal of Hydrology*, 331, 3–4, 591–604. doi:10.1016/j.jhydrol.2006.06.008.
- Nelsen, R.B., 2006. *An Introduction to Copulas*. 2<sup>nd</sup> edition. Springer-Verlag, New York.
- Parajka, J., Merz, R., Blöschl, G., 2007. Uncertainty and multiple objective calibration in regional water balance modelling – Case study in 320 Austrian catchments. *Hydrological Processes*, 21, 435–446. doi:10.1002/hyp.6253.
- Parajka, J., Kohnová, S., Bálint, G., Barbuc, M., Borga, M., Claps, P., Cheval, S., Dumitrescu, A., Gaume, E., Hlavčová, K., Merz, R., Pfaundler, M., Stancalie, G., Szolgay, J., Blöschl, G., 2010. Seasonal characteristics of flood regimes across the Alpine–Carpathian range. *Journal of Hydrology*, 394, 1–2, 78–89. doi:10.1016/j.jhydrol.2010.05.015.
- Pham, M.T., Vernieuwe, H., Baets, B.D., Willems, B., Verhoest, N.E.C., 2015. Stochastic simulation of precipitation-consistent daily reference evapotranspiration using vine copulas. *Stochastic Environmental Research and Risk Assessment*, 18 p. doi: 10.1007/s00477-015-1181-7.
- Poulin, A., Huard, D., Favre, A.-C., Pugin, S., 2007. Importance of tail dependence in bivariate frequency analysis. *Journal of Hydrologic Engineering*, 12, 4, 394–403. doi:10.1061/(ASCE)1084-0699(2007)12:4(394).
- Reddy, M.J., Ganguli, P., 2012. Bivariate flood frequency analysis of Upper Godavari River flows using Archimedean copulas. *Water Resources Management*, 26, 14, 3995–4018. doi:10.1007/s11269-012-0124-z.
- Remillard, B., Plante, J.-F., 2012. TwoCop: Nonparametric test of equality between two copulas. R package version 1.0 (19-10-2012). <http://cran.r-project.org/web/packages/TwoCop>.
- Remillard, B., Scaillet, O., 2009. Testing for equality between two copulas. *Journal of Multivariate Analysis*, 100, 377–386. doi:10.1016/j.jmva.2008.05.004.
- Requena, A.I., Chebana, F., Mediero, L., 2016. A complete procedure for multivariate index-flood model application. *Journal of Hydrology*, 535, 559–580. doi:10.1016/j.jhydrol.2016.02.004.
- Rosbjerg, D., Blöschl, G., Burn, D.H., Castellarin, A., Croke, B., DiBaldassarre, G., Iacobellis, V., Kjeldsen, T.R., Kuczera, G., Merz, R., Montanari, A., Morris, D., Ouarda,

- T.B.M.J., Ren, L., Rogger, M., Salinas, J.L., Toth, E., Viglione, A., 2013. Prediction of floods in ungauged basins. Chapter 9. In: Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., Savenije, H. (Eds.): *Runoff Prediction in Ungauged Basins - Synthesis across Processes, Places and Scales*. Cambridge University Press, Cambridge, UK, pp. 135–162.
- Salinas, J.L., Castellarin, A., Viglione, A., Kohnová, S., Kjeldsen, T.R., 2014. Regional parent flood frequency distributions in Europe – Part 1: Is the GEV model suitable as a pan-European parent? *Hydrology and Earth System Sciences*, 18, 4381–4389. doi:10.5194/hess-18-4381-2014.
- Serinaldi, F., 2013. An uncertain journey around the tails of multivariate hydrological distributions. *Water Resources Research*, 49, 10, 6527–6547. doi:10.1002/wrcr.20531.
- Serinaldi, F., 2015. Can we tell more than we can know? The limits of bivariate drought analysis in the United States. *Stochastic Environmental Research and Risk Assessment*, 14 p. doi:10.1007/s00477-015-1124-3.
- Serinaldi, F., Kilsby, C.G., 2013. The intrinsic dependence structure of peak, volume, duration, and average intensity of hyetographs and hydrographs. *Water Resources Research*, 49, 3423–3442. doi:10.1002/wrcr.20221.
- Shiau, J.-T., Wang, H.-Y., Tsai, C.-T., 2006. Bivariate flood frequency analysis of floods using copulas. *Journal of the American Water Resources Association*, 42, 6, 1549–1564. doi:10.1111/j.1752-1688.2006.tb06020.x.
- Sraj, M., Bezak, N., Brilly, M., 2014. Bivariate flood frequency analysis using the copula function: a case study of the Litija station on the Sava River. *Hydrological Processes*, 29, 2, 225–238. doi:10.1002/hyp.10145.
- Szolgay, J., Gaál, L., Bacigál, T., Kohnová, S., Hlavčová, K., Výleta, R., Blöschl, G., 2016. A regional look at the selection of a process-oriented model for flood peak/volume relationships. *IAHS publ. 373*. IAHS Press, Wallingford, pp. 61–69. doi: 10.5194/piahs-373-1-2016.
- Szolgay, J., Gaál, L., Kohnová, S., Hlavčová, K., Výleta, R., Bacigál, T., Blöschl, G., 2015. A process-based analysis of the suitability of copula types for peak-volume flood relationships. *Proc. IAHS*, 370, 183–188. doi: 10.5194/piahs-370-183-2015.
- Vernieuwe, H., Vandenberghe, S., De Baets, B., Verhoest, N.E.C., 2015. A continuous rainfall model based on vine copulas. *Hydrology and Earth System Sciences*, 19, 6, 2685–2699. doi:10.5194/hess-19-2685-2015.
- Werner, P.C., Gerstengarbe, F.-W., 2010. *Katalog der Grosswetterlagen Europas (1881–2009) nach Paul Hess und Helmut Brezowsky, 7., verbesserte und ergänzte Auflage*. [Catalog of Large Weather Conditions of Europe (1881-2009), after Paul Hess and Helmut Brezowsky, 7<sup>th</sup>, Improved and Completed Edition]. PIK-Report No. 119, Potsdam Institute for Climate Impact Research, Potsdam, 146 p. (In German.)
- Zhang, L., Singh, V.P., 2006. Bivariate flood frequency analysis using the copula method. *Journal of Hydrologic Engineering* 11, 150–164. doi:10.1061/(asce)1084-0699(2006)11:2(150).

Received 17 June 2016

Accepted 10 August 2016