7th International Water Resources Management Conference of ICWRS,
18–20 May 2016, Bochum, Germany, IWRM2016-68-2

A regional look at the selection of a process-oriented model for flood peak/volume relationships

J. Szolgay[1], L. Gaál[1,2], T. Bacigál[3], S. Kohnová[1], K. Hlavčová[1], R. Výleta[1], and G. Blöschl[4]

[1] Department of Land and Water Resources Management, Faculty of Civil Engineering, Slovak University of Technology, Bratislava, Slovakia

[2] MicroStep-MIS, spol. s r.o., Bratislava, Slovakia

[3] Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Slovak University of Technology, Bratislava, Slovakia

[4] Institute for Hydraulic and Water Resources Engineering, Vienna University of Technology, Vienna, Austria

Correspondence to: Jan Szolgay (jan.szolgay@stuba.sk)

Abstract

Research on the bivariate flood peak/volume frequency analysis has mainly focused on the statistical aspects of the use of various copula models. The interplay of climatic and catchment processes in discriminating among these models has attracted less interest. In the paper we analyse the influence of climatic and hydrological controls on flood peak and volume relationships and their models, which are based on the concept of comparative hydrology in the catchments of a selected region in Austria. Independent flood events have been isolated and assigned to one of the three types of flood processes: synoptic floods, flash floods and snowmelt floods. First, empirical copulas are regionally compared in order to verify whether any flood processes are discernible in terms of the corresponding bivariate flood-peak relationships. Next the types of copulas, which are frequently used in hydrology are fitted, and their goodness-of-fit is examined in a regional scope. The spatial similarity of copulas and their rejection rate, depending on the flood type, region, and sample size are examined, too. In particular, the most remarkable difference is observed between flash floods and the other two types of flood. It is concluded that treating flood processes separately in such an analysis is beneficial, both hydrologically and statistically, since flood processes and

the relationships associated with them are discernible both locally and regionally in the pilot region.

However, uncertainties inherent in the copula-based bivariate frequency analysis itself (caused, among others, also by the relatively small samples sizes for consistent copula model selection, upper tail dependence characterization and reliable predictions) may not be overcome in the scope of such a regional comparative analysis.

1 Introduction

Bivariate distributions of flood peaks and flood event volumes are needed for a range of practical purposes, including, e.g., the design of retention basins and identifying the extent and duration of flooding in flood hazard zones. For the statistical analysis of flood peaks and volumes, identical marginal distributions for both random variables were used in the past (e.g., Goel et al., 1998; Yue et al., 2002). Recently, the use of copula-based multivariate models has become widespread. These allow for separate studies of the marginal distributions of the component variables and the correlation/dependence structure between them. Numerous studies have been published on this topic (e.g., Shiau, 2003; De Michele et al., 2005; Chowdhary et al., 2011; Requena et al., 2013), including recommendations and how to select appropriate copula models (e.g., Favre et al., 2004; Genest and Favre, 2007). Despite numerous studies, as mentioned by Chowdhary et al. (2011), the use of copula-based multivariate distributions for hydrological designs still cannot be regarded as having been satisfactorily resolved. The selection of the types of bivariate distributions and the estimation of their parameters from observed peak-volume pairs are associated with far greater uncertainties when compared to univariate distributions, since observed flood records of the required length are rarely available. This poses a problem for reliable estimations of flood risks in bivariate design cases. It is being increasingly recognized that the problem cannot be approached from only a purely statistical perspective. The crucial step for predictions in the multivariate modelling of flood characteristics by copulas is the choice of the copula which best fits the data (Favre et al., 2004). In this respect Serinaldi and Kilsby (2013) highlighted the importance of studying the relationships between processes that generate the design variables and also the statistical techniques used to model them. In our previous studies we have attempted to better understand the hydrological factors controlling hydrographic shapes that have implications for the dependence between peaks and volumes. In Gaál et al. (2012), we analyzed the ratio of both quantities based on the concept of comparative hydrology in a regional context in Austria and compared catchments with contrasting characteristics in order to understand the controls 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 (e.g., through antecedent soil moisture and soil characteristics). In Gaál et al. (2015), our aim was 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. These coefficients ranged from about 0.2 in high alpine catchments to about 0.8 in lowlands. The weak dependence in high alpine catchments was attributed to a mix of flood types. The results also suggested that the factors controlling the strength of the dependence are more related to climate instead of catchment characteristics. In Szolgay et al. (2015), we aimed at analyzing the formal suitability of various copula-based bivariate relationships between flood peaks and flood volumes, with a particular focus on two basic flood generating seasons (summer and winter floods) with the goal of going a little beyond the statistics alone in the choice of the copula functions for the engineering applications. It was concluded that, for rainfall-fed floods, three extreme value copulas performed best in the pilot region (the Galambos, Gumbel-Hougaard and Hüsler-Reiss copulas) followed by a normal copula. The other copulas were not regarded as regionally acceptable. For winter floods the best performer was the Frank copula, followed by the normal and Plackett copulas and the three extreme value models. The Clayton and Joe copulas indicated an unacceptable performance for both seasons. In Szolgay et al. (2015), we also illustrated the importance of considering the influence of the length of data series through two simple simulation experiments. The preferences for the choice of copulas were still visible, but less evident. 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 the flood types within a region may play an important role.

Here, these approaches are followed in a more differentiated regional look at the selection of a hydrological process-oriented copula model for flood peak/volume relationships. Specifically, we are again interested in the following questions:

- how similar are peak-volume relationships of different flood types in a climatically rather homogeneous region but in geologically different subregions,

- which flood generation processes may play a role in forming the similarity of copula models in the subregions and
- if recommendations could be formulated for engineering studies with respect to the suitability of some copula types for the given flood processes ?

2 The pilot region

There are a wide variety of flood-generation mechanisms across Austria (e.g., Merz and Blöschl, 2003), which result in complex flood peak-volume relationships (Gaál et al., 2012, 2015; Szolgay et al., 2015). In order to decrease the complexity of runoff generation schemes in this analysis, we decided to reduce the climate variability through the selection of a geographically limited area, i.e., the Northern Lowlands region of Austria (Fig. 1), as a pilot region. The area is dominated by lowlands and hilly sites, with elevations ranging from about 400 to 1500 m a.s.l. The region is generally under the influence of air masses from the Atlantic. The mean annual precipitation in the target region shows a decreasing western-to-eastern gradient. Orographic enhancement is not significant; the annual rainfall amounts (from about 500 to 1500 mm) are significantly lower than in the Austrian Alps. The basic climatic and physiographic characteristics of the catchments are listed in Table 1.

Floods in the Northern Lowlands region may occur during the whole year. The winter floods are usually induced by snowmelt and rain-on-snow processes, when antecedent snowmelt saturates the soils, and temperature increases and/or relatively low rainfall intensities may then cause significant floods. Flash floods are mostly caused by convective events or cold fronts; synoptic floods depend on the particular circulation pattern, but westerly circulation prevails.

The data set used in this paper builds on the Austrian flood data described in detail in Szolgay et al. (2015) and the papers referenced therein. The 72 small and mid-sized catchments analyzed have areas in the range of 10.6 to 444.3 km^2 (median: 78.6 km^2), 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. Three subregions were selected that were formed on traditional geographical units, using results of Gaál et al. (2012), which were based on characteristic geology, geomorphology, and pedology (Fig. 1). The first subregion consists of the Weinviertel and Mühlviertel areas in the north, with a prevailing geology made of granite. In this region 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. The second subregion consists of the Innviertel and Hausruckviertel areas, where catchments have even more winter floods than the other subregions. The geology in this subregion is prevailingly formed of gravel and sand. In the third subregion, which consists of the Traunviertel and parts of the Flysh, the catchments have more summer floods and flash floods. The geology in this subregion is made of marl and sandstone.

3 Methodology

Instead of the traditional engineering approach which often deals with flood volumes associated with the annual maxima of flood events, the current analysis intends to include all the flood events in the region which can be hydrologically regarded as independent. According to our understanding, two subsequent flood events are independent, when they do not originate from the same synoptic situation. We assumed that after a seven-day period, on the average, a completely different atmospheric circulation situation occurs in Central Europe (Gaál et al., 2015; Szolgay et al., 2015). The flood type classification according to the genesis of events in the region was introduced by Merz and Blöschl (2003) and modified in Gaál et al. (2015). As a further modification, 25 697 flood events from the 1976–2007 period were selected and classified as synoptic floods (originally long or short rain-induced floods in Merz and Blöschl, 2003), flash floods, and snowmelt floods (originally rain-on-snow floods or snowmelt floods).

As in Szolgay et al. (2015), nine frequently-used copulas were chosen to be fitted locally, specifically from the Archimedean class (Clayton, Frank, Gumbel-Hougaard and Joe copulas), the extreme-value class (Gumbel-Hougaard, Galambos, Hüsler-Reiss), the elliptical class (normal, Student-t), and finally the Plackett copula. (The abbreviations of the copulas used throughout the paper are: *cla* = Clayton, *fra* = Frank, *gal* = Galambos, *gum* = Gumbel-Hougaard, *hus* = Hüsler-Reiss, *joe* = Joe, *nor* = Normal, *pla* = Plackett, *tco* = Student t copula). All the copulas are single parameter copulas with the exception of the Student-t copula. In order to be consistent with the other copulas, we fixed the second parameter to effectively make it a single parameter copula. We chose a value of four for the second parameter to distinguish it from the normal copula.

In this paper we are also interested in the similarity of empirical copulas of flood peak-volume pairs, which we tested accordingly to the approach of Remillard and Scaillet (2009). It is based on the Cramér-von Mises type of distance measure:

$$S = \sum_i \left[C_{n1}\left(U_{1,i}, U_{2,i}\right) - C_{n2}\left(U_{1,i}, U_{2,i}\right)\right]^2 \tag{1}$$

where C_{n1} and C_{n2} denote two empirical copulas. The probability distribution of S, given that the null hypothesis (*H0*: the empirical copulas come from the same – unknown – bivariate distribution) holds, is unknown and needs to be bootstrapped. The result of the similarity testing is a p-value, which is the percentage of how many simulations of the test statistic (under *H0*) exceeded the estimator from the observations. Since the test statistic represents a distance, the lower values (a smaller departure from the null hypothesis statement) lead to larger p-values, which in turn can be used as a measure of dissimilarity.

As in Szolgay et al. (2015), the parameter of the copulas was estimated by maximizing the so-called pseudo-likelihood function:

$$L(\theta) = \sum_i log\left[c_\theta\left(U_{1,i}, U_{2,i}\right)\right], \tag{2}$$

which, in addition to the copula density c_θ, contains pseudo-observations $U_{j,i}$ $(i = 1, \dots n, j = 1, 2)$, i.e., a transformation of n real observations of random variable X_j, by means of a corresponding empirical distribution function (sometimes referred to as the “plotting position”). The goodness-of-fit was examined locally as well as analyzed in a regional scope. The goodness-of-fit of the parametric copulas under consideration was tested by a ‘blanket’ test (Genest et al., 2009) with the Cramér-von Mises measure of distance:

$$S_n = \sum_{i=1}^{n} \left[C_\theta\left(U_{1,i}, U_{2,i}\right) - C_n\left(U_{1,i}, U_{2,i}\right)\right]^2 \tag{3}$$

between the parametric copula C_θ and the empirical copula defined by:

$$C_n(u_1, u_2) = \frac{1}{n} \sum_{i=1}^{n} 1\left(U_{1,i} \leq u_1\right) 1\left(U_{2,i} \leq u_2\right). \tag{4}$$

The probability distribution of the test statistic S_n, given that the null hypothesis (*H0*: C_θ fits well) holds, is unknown and needs to be bootstrapped. Consequently, the p-value is a percentage of how many simulations of the test statistic (under *H0*) exceeded the estimator from the actual observations.

4 Results and discussion

The regional distribution of the flood types for each subregion is shown in Table 2. Locally, the percentage of the flood types ranges between 51 and 74% in the case of synoptic floods and 8 and 46% and 2 and 19% for the snowmelt and flash floods, respectively. In the selection there are no catchments with a missing flood type. The relatively smaller number of snowmelt floods is due to the modest elevations in the region. The fact that flash floods represent 6.7% of all the events is not negligible, since in Gaál et al. (2015), the flash floods only represented 1.8% based on the annual maxima flood events. The total number of 25 697 independent flood events identified in the target region does show a spatial pattern. It is largest in the southwest (the Innviertel and Hausruckviertel regions), smaller in the central parts of the target region (the Traunviertel and Flysh), and smallest in the north (the Mühlviertel and the Waldviertel) (Table 2).

First, a comparison of the empirical copulas for the different flood types locally was performed. In this 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 the process types were compared pairwise, i.e., synoptic floods vs. flash floods, synoptic floods vs. snowmelt floods, and flash floods vs. snowmelt floods. The results, which are shown in Figure 2, suggest that synoptic and snowmelt floods seem to belong more often to the same unknown copula than is the case for the other combinations. This suggests that the synoptic and snowmelt floods are more similar to each other (in terms of their empirical copulas) than the other process pairs; or, in other words, flash floods tend to be more dissimilar from both the synoptic and snowmelt floods. This could be partly related to much stronger upper tail dependence of flash floods and their specific (similar) hydrograph shapes (Gaál et al., 2015). However, the relatively small number of events for such type of analysis in general and the differing number of events in the respective flood types in particular, may also play a role in the fact, that the analysis has not brought conclusive results. These are objective factors, which cannot be overcome in the framework of comparative hydrology, when based on real data.

Next, a comparison of the empirical copulas for each flood type regionally was performed where we were interested in whether different catchments with the same flood type are distinguishable in terms of their empirical peak-volume copulas (Figure 3). It can be seen that the empirical copulas of synoptic floods are the least similar between the catchments. This seems to be surprising; a closer look into the subregions showed that this phenomenon is more pronounced in the southwest than in the north. The high ratio of rejections of the synoptic events across different pairs of sites is therefore likely to be related to the more complex temporal rainfall structure, the mix of long and short rain processes in the dataset, and the more complex geology resulting in a lower degree of similarity between the different events and sites which may become more evident, when the sample size increases (as it is in the case of synoptic floods when compared to the other two types). A more detailed analysis of the phenomenon is beyond the goals of the present study. In the case of flash floods and snowmelt, the difference between the process types is smaller; the analysis suggests that most catchment pairs, for a given flood type are not distinguishable in terms of their empirical peak-volume copulas.

Next, we attempted a more detailed flood typology and regional differentiation in fitting the copulas to the data than that in Szolgay et al. (2015). The goodness-of-fit test of the nine copula types at the 72 catchments for all the floods merged into one set; and the three flood types separately are shown in Figure 4, stratified by subregions. The copula types (each column represents a copula type) are organized alphabetically, while the subregions are visualized by different color bars, which indicate $p \leq 0.05$, i.e., a rejection of the null hypothesis H_0. The number of catchments in the subregions is comparable, but the number of events analyzed was naturally different for each flood type.

In the case of analyzing all the floods together, we can see, that the three extreme value copulas (the Galambos, Gumbel-Hougaard and Hüsler-Reiss copulas) and the normal copula performed best in all the subregions (except in the southwest, where the extreme value models clearly outperform the normal). In 15% of the catchments all nine models were rejected, the acceptance rate oscillates around 50 per cent in general. This relatively low rate (when compared to the process-wise analysis below) could be attributed to the mix of flood types in the data sample, but the effect of a relatively large sample size when compared to the other types cannot be excluded (larger sample decreases the uncertainty even to an extent, where the models considered in engineering applications are not suitable at all). The acceptance rate of these four copulas improved for the synoptic floods (still in 7% of the catchment no one model was found suitable), but was highly variable across the subregions, which was not expected. This could be the result of the mix of short and long rain processes

(Merz and Blöschl, 2003) and a smaller number of events in the subsamples; however, a detailed analysis is beyond the goals of this study. Interestingly, the extremal copulas exhibited a larger rejection rate in the Traunviertel and Flysh subregion, and the Student copula can also be considered as applicable in practice. The rejection/acceptance rate pattern in the case of snowmelt floods is clearly different and adds the Frank, Plackett and Student models into the basket of acceptable models (again in two catchments no suitable model was found). Nevertheless, the behavior of snowmelt floods within the subregion Traunviertel and Flysh is again different. An explanation could be found in the fact, that in this region, the underlying geology is more variable; the variability of the mean catchment elevation covers the other two regions (for two clearly disjunctive sets) and the time scales and their variability partially overlaps with the Waldviertel–Mühlviertel region (Gaál et al., 2012). However such an analysis was beyond the scope of the research in this case. Further, one must also suspect, that the different regional patterns of the copula suitability in winter can also be related to the fact that the number of winter events is lower compared to the summer floods (though larger than in many copula studies), Szolgay et al. (2015). The behavior of the flash floods (the smallest sample size) in the subregions is very similar to that observed in Szolgay et al. (2015) for the whole region and allows for a choice of almost all the models (despite the rather small variability of the flood wave shapes in the region; see, e.g., Gaál et al., 2015). In all the catchments at least one suitable model was found.

These results indicate that the acceptance of a particular copula model can be conditioned on the processes but that the size of the data samples and possibly also the homogeneity of the region with respect to the flood formation factors and flood types within the data set plays a role. However, uncertainties inherent in the bivariate frequency analysis methodology itself in real world applications (e.g., small samples for reliable model identification and upper tail dependence description) may not be overcome in the scope of such a regional comparative analysis as conducted here.

5 Conclusions

Not much attention has been paid thus so far to directing a multivariate analysis of floods toward the selection of models for specific runoff generation processes. Here, this issue was addressed in a regional context by the differentiation of the flood types into three categories. Based on the results, it can be 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, since the suitability patterns of acceptable copula types are distinguishably different for the subregions/flood-types considered. This could help analysts to overcome some difficulties in the choice of the model caused by the inadequate length of a data series. On the other hand, it was also shown that a more detailed differentiation of the flood types and subregions opens in the selection of the model a greater degree of uncertainty than expected in Szolgay et al. (2015), which does not make the task easier for an analyst in practice. Despite that shortcoming, given that usually more than one statistically suitable dependence model exists, a regional analysis and an uncertainty analysis of the design values in the engineering studies resulting from the choice of a model can be recommended, especially for important water resources projects.

Our results support 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. But, as shown in a comparative hydrology framework above, the choice of the copula model that best fits the observed data and is regionally acceptable in term 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 (if not even impossible). 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).

Note that even the adoption of generally accepted and widely used copula models may not lead to a successful bivariate fitting. Uncertainties inherent in the copula-based bivariate frequency analysis itself (caused, among others, also by the relatively small samples sizes for consistent copula model selection, upper tail dependence characterization and reliable predictions) may not be overcome in the scope of such a regional comparative analysis.

Based on this comparative study and results of other more advanced studies (e.g., Serinaldi, 2013, 2015) it can be concluded, if reliable predictions will be required for an important engineering application, the benefits of regional bivariate frequency analysis methods could be further explored (e.g., Ben Aissisa 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.

Acknowledgements

We would like to thank the Austrian Academy of Sciences (International Strategy for Disaster Reduction Programme, IWHRE2008, 2008–2013), for its financial support. This research was also supported by the Slovak Research and Development Agency under Contract No. APVV 0496-10, by the Slovak Grant Agency VEGA under Project No. 1/0776/13 and by the IMPALA project (FP7-PEOPLE-2011-IEF-301953) of the Marie Curie Intra European Fellowship. This publication was further supported by Competence Center for SMART Technologies for Electronics and Informatics Systems and Services, ITMS 26240220072, funded by the Research & Development Operational Programme from the ERDF.

The authors would also like to thank Francesco Serinaldi for a very inspiring open discussion on the topic during a review of their other paper.

References

Ben Aissia, M. A., Chebana, F., Ouarda, T. B. M. J., Bruneau, P., and Barbet, M.: Bivariate index-flood model: case study in Québec, Canada. Hydrol. Sci. J., 60, 247–268, doi:10.1080/02626667.2013.875177, 2015.

Chowdhary, H., Escobar, L. A., and Singh, V. P.: Identification of suitable copulas for bivariate frequency analysis of flood peak and flood volume data, Hydrol. Res., 42, 193–216, doi:10.2166/nh.2011.065, 2011.

De Michele, C., Salvadori, G., Canossi, M., Petaccia, A., and Rosso, R.: Bivariate statistical approach to check adequacy of dam spillway, J. Hydrol. Eng., 10, 50–57, 2005.

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

Gaál, L., Szolgay, J., Kohnová, S., Parajka, J., Merz, R., Viglione, A., and Blöschl, G.: Flood timescales: Understanding the interplay of climate and catchment processes through comparative hydrology, Water Resour. Res., 48, W04511, doi:10.1029/2011WR011509, 2012.

Gaál, L., Szolgay, J., Kohnová, S., Hlavčová, K., Parajka, J., Viglione, A., Merz, R., and Blöschl, G.: Dependence between flood peaks and volumes – A case study on climate and hydrological controls, Hydrol. Sci. J., 60, 968–984, doi:10.1080/02626667.2014.951361, 2015.

Genest, C. and Favre, A.-C.: Everything you always wanted to know about copula modeling but were afraid to ask, J. Hydrol. Eng., 12, 347–368, 2007.

Genest, C., Rémillard, B., and Beaudoin, D.: 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, 2009.

Goel, N. K., Seth, S. M., and Chandra, S.: Multivariate modeling of flood flows, ASCE, J. Hydraul. Eng., 124, 146–155, 1998.

Merz R. and Blöschl, G.: A process typology of regional floods, Water Resour. Res., 39, 1340, doi:10.1029/2002WR001952, 2003.

Requena, A. I., Mediero, L., and Garrote, L.: Bivariate return period based on copulas for hydrologic dam design: comparison of theoretical and empirical approach, Hydrol. Earth Syst. Sci., 17, 3023–3038, doi:10.5194/hess-17-3023-2013, 2013.

Remillard, B. and Scaillet, O.: Testing for equality between two copulas, Journal of Multivariate Analysis, 100, 377–386, doi:10.1016/j.jmva.2008.05.004, 2009.

Shiau, J. T.: Return period of bivariate distributed extreme hydrological events, Stoch. Environ. Res. Risk Assess., 17, 42–57, doi:10.1007/s00477-003-0125-9, 2003.

Serinaldi, F. and Kilsby, C. G.: The intrinsic dependence structure of peak, volume, duration, and average intensity of hyetographs and hydrographs, Water Resour. Res., 49, 3423–3442, doi:10.1002/wrcr.20221, 2013.

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

Serinaldi, F.: Can we tell more than we can know? The limits of bivariate drought analysis in the United States, Stoch. Environ. Res. Risk Assess. doi:10.1007/s00477-015-1124-32015, 2015.

Szolgay, J., Gaál, L., Kohnová, S., Hlavčová, K., Výleta, R., Bacigál, T, and Blöschl, G.: 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, 2015.

Yue, S, Ouarda, T. B. M. J., Bobeé, B., Legendre, P., and Bruneau, P.: Approach for describing statistical properties of flood hydrograph, J. Hydrol. Eng., 7, 147–53, 2002.

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^2], *MCE* – mean catchment elevation [m a.s.l.], *MRC* – mean runoff coefficient of maximum annual flood events [-], *MCS* - mean catchment slope [-], *PFA* – percentage of forest area [-].

		AREA [km^2]	***MCE*** [m a.s.l.]	***MRC*** [-]	***MCS*** [-]	***PFA*** [-]
All sites	maximum	444.3	888.0	0.85	0.33	1.00
	median	78.6	571.0	0.49	0.09	0.30
	minimum	10.6	342.0	0.19	0.02	0.00
Subregion #1	maximum	361.8	598.0	0.77	0.14	0.46
	median	67.4	450.5	0.60	0.10	0.13
	minimum	14.2	385.0	0.36	0.07	0.01
Subregion #2	maximum	444.3	839.0	0.74	0.33	1.00
	median	55.8	571.0	0.49	0.10	0.34
	minimum	12.0	342.0	0.20	0.02	0.00
Subregion #3	maximum	305.9	888.0	0.85	0.16	0.82
	median	120.6	707.5	0.42	0.09	0.48
	minimum	10.6	480.0	0.19	0.03	0.26

Table 2. Number of identified independent flood events, stratified by regions and flood types.

	Number of catchments	All floods	Synoptic floods	Flash floods	Snowmelt floods
All sites	72	25697	19093	1733	4871
Subregion #1	24	9396	7325	596	1475
Subregion #2	22	8788	6448	652	1688
Subregion #3	26	7513	5320	485	1708

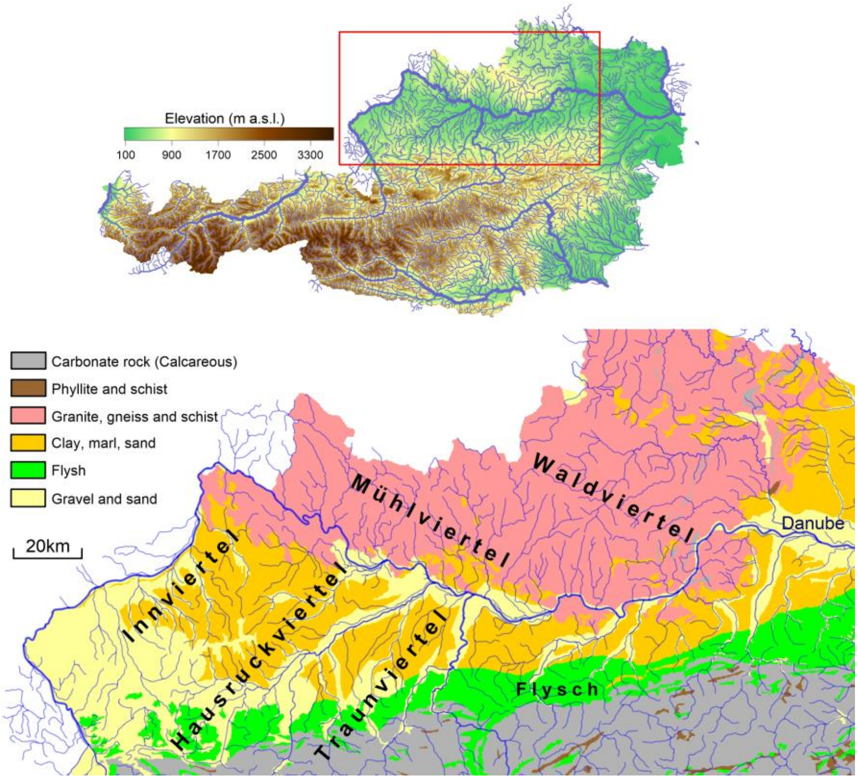

Figure 1. Location and geology of the pilot region in Austria and the subregions considered.

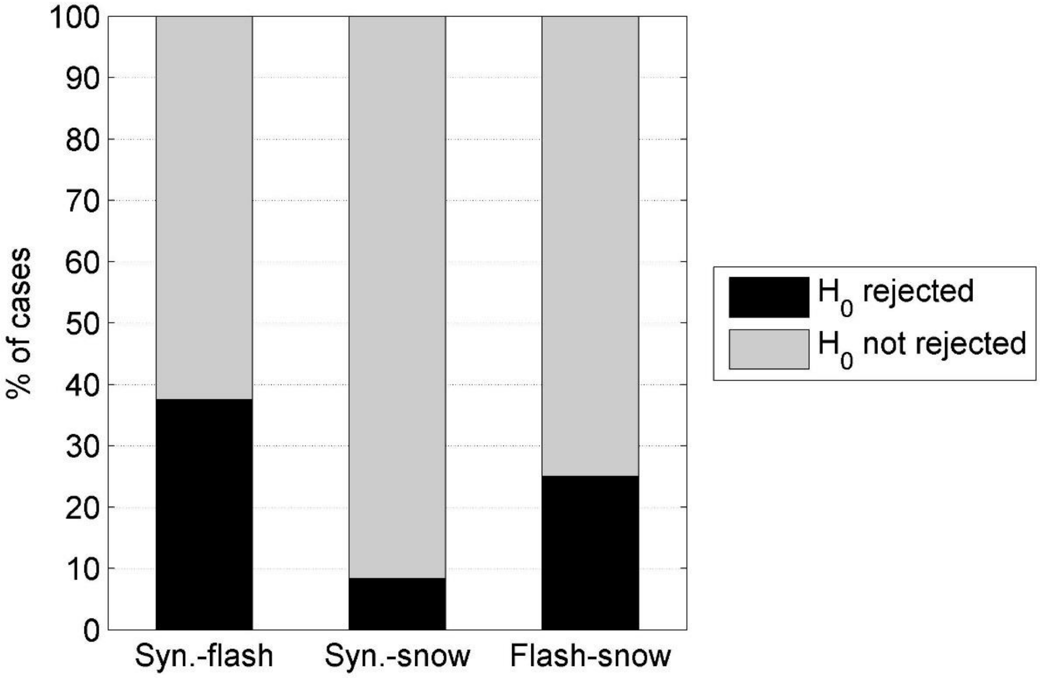

Figure 2. Results of the comparison of the empirical copulas locally: per cent ratio of the catchments where the equivalence of the pairs of empirical copulas was rejected or could not be rejected at the given level of significance (here α = 10%). The black color indicates the per cent ratio of the catchments where the null hypothesis about the equality of the empirical copulas for the given two flood processes was rejected.

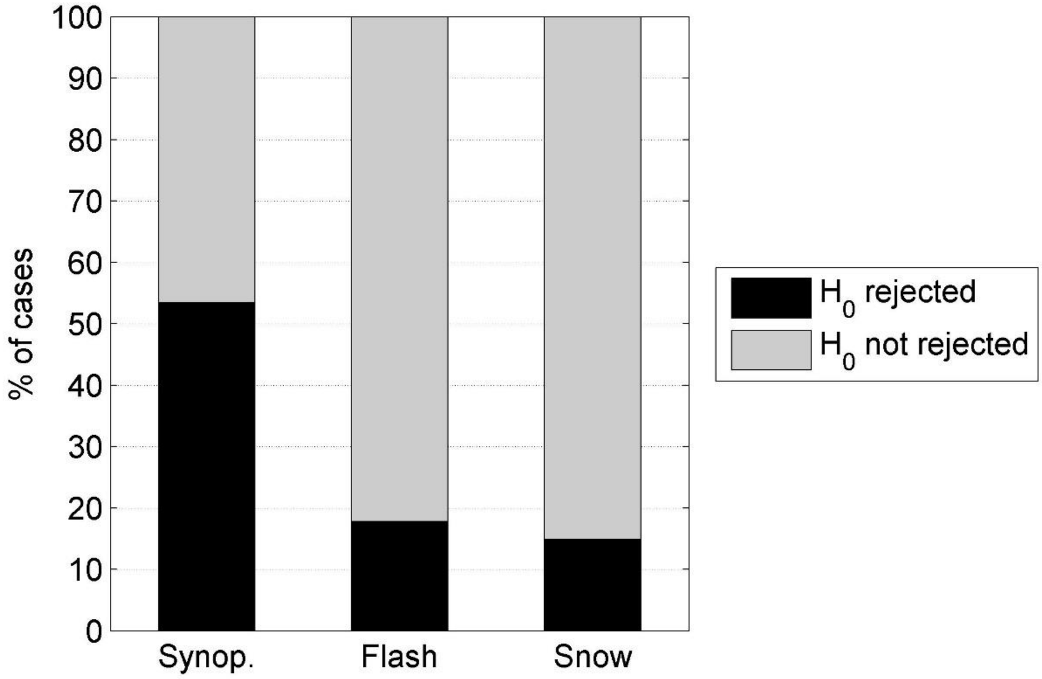

Figure 3. Results of the comparison of the empirical copulas regionally: per cent ratio of the cases where the equivalence of the empirical copulas was rejected or could not be rejected at the given level of significance (here $\alpha = 10\%$). Within the set of 72 sites, and for each flood process separately, altogether 2556 (72*71/2) pairwise comparisons were carried out. The black color indicates the percent ratio of the cases where the null hypothesis about the equality of the empirical copulas for the given flood process was rejected.

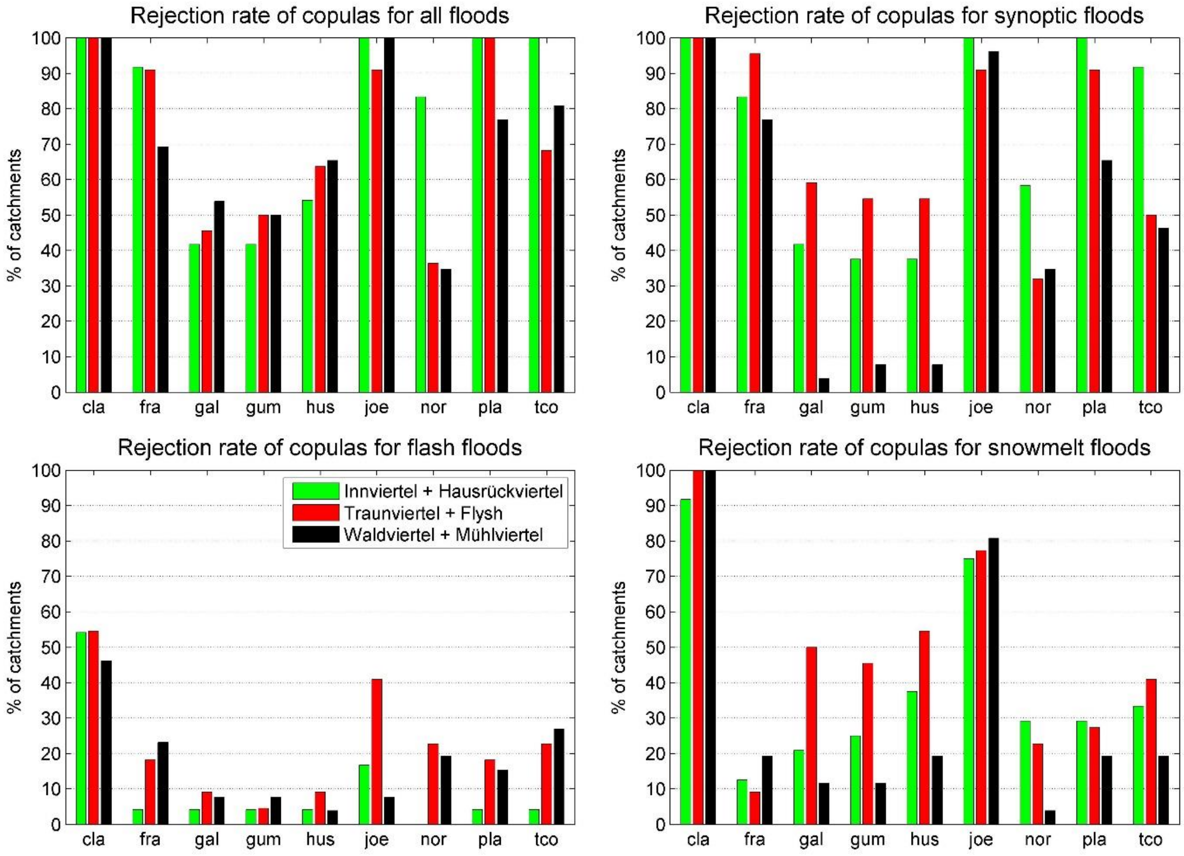

Figure 4. Results of the goodness-of-fit test of the selected nine copula types within the three subregions (indicated by colors), for the all the floods merged in a single data set (top left) and for flood types treated separately. The bars indicate per cent ratio of the catchments where the given copula type was rejected.