Flood volume estimation in Switzerland using synthetic design hydrographs - a multivariate statistical approach

Manuela Irene Brunner[1,2]**;** Olivier Vannier, Dr.[1]; Anne-Catherine Favre, Prof.[1]; Daniel Viviroli, Dr.[2,3]; Paul Meylan[4];
Anna Sikorska[2], Dr.; Jan Seibert, Prof.[2,5]

ABSTRACT

Accurate estimations of flood peaks, volumes and hydrographs are needed to design safe and cost-effective hydraulic structures. In this study, we propose a statistical approach for the estimation of the design variables peak and volume by constructing a synthetic design hydrograph. Our approach is based on fitting probability density functions to observed flood hydrographs and takes the dependence between the two variables peak and volume into account. The method consists of the following six steps: sampling of flood events, baseflow separation, normalization of the hydrographs, fitting of the hydrographs with statistical density functions, modeling of peak and volume considering their dependence, and construction of the synthetic design hydrograph. The method was developed and tested based on data from nine meso-scale catchments in Switzerland, and has been shown to provide reliable synthetic design hydrographs for all of these catchments. While the method has so far been applied to gauged catchments, it is foreseen to make it applicable to ungauged catchments using regionalization approaches.

KEYWORDS

Synthetic design hydrographs; flood volume estimation; bivariate analysis, copulas

INTRODUCTION

An accurate flood estimation is needed for resilient flood risk management, to design hydraulic structures such as dam spillways, bridges, road culverts, and levees, and to manage residential area zoning, floodplains, and urban design. The major quantity of interest in flood estimation is the magnitude of the flood peak for a specific return period (Rosbjerg et al., 2013). However, flood peaks provide only a limited description of a flood event. For the prevention of flood damage and for designing hydraulic structures, it is also important to know the flood volume and the shape of the entire flood hydrograph. The aim of this study was to develop a flood volume estimation method for meso-scale catchments (i.e. 20-1000 km^2) which can easily be applied by practitioners. Although different methods to derive

1 Université Grenoble-Alpes, Grenoble INP, LTHE, Grenoble, FRANCE, manuela.brunner@geo.uzh.ch
2 Department of Geography, University of Zurich, SWITZERLAND
3 belop gmbh, Sarnen, SWITZERLAND
4 AIC Ingénieurs Conseil SA, Lausanne, SWITZERLAND
5 Department of Earth Sciences, Uppsala University, Uppsala, SWEDEN

design flood hydrographs are described in the literature, most of them are of limited use for practitioners because of their complexity (Yue et al., 2002).

Therefore, we propose a statistical approach for the estimation of not only flood peaks but also flood volumes and the entire flood hydrographs using synthetic design hydrographs. The method has been developed as a simple tool that can support practitioners in consulting or engineering companies with reasonable efforts. The focus lies on meso-scale catchments with natural runoff conditions, i.e. without significant human alterations, and with an insignificant degree of glacierized areas. The basic idea relies on fitting probability density functions to observed flood hydrographs, while considering the dependence between flood peak and flood volume. This is essential because flood events are multivariate and a univariate frequency analysis does not allow for a complete assessment of the probability of their occurrence (Yue et al., 2002). The approach does not rely on a rainfall-runoff model, which makes it less demanding with respect to input data and methods. The transfer to ungauged catchments is foreseen in a next step.

DATA

The proposed method has been developed and tested using data from a representative set of nine meso-scale study catchments in Switzerland. The selected catchments cover different sizes, elevations, and regime types. To avoid hydrograph shapes modified by direct human impacts, we selected only catchments with flow conditions neither altered through hydropower plants nor lake regulation. For the development of the method, long term observations were required. We also exclude highly glacierized catchments because unimpaired flow records are scarce. Moreover, they are usually only sparsely populated and therefore exhibit a low damage potential. The characteristics of the nine study catchments selected are listed in Table 1.

Table 1: List of study catchments and their characteristics.

Catchment	*Station name*	*Area [km^2]*	*Station elevation [m]*	*Mean elevation [m]*	*Glacierized area*	*Regime type (according to (Weingartner and Aschwanden 1992)*	*Observation period*
Birse	Moutier, La Charrue	183	519	930	0 %	*nivo-pluvial jurassien*	1974-2013
Emme	Eggiwil, Heidbüel	124	745	1189	0 %	*nivo-pluvial préalpin*	1975-2013
Emme	Emmenmatt	443	638	1070	0 %	*nivo-pluvial préalpin*	1997-2013
Emme	Wiler, Limpachmündung	939	458	860	0 %	*pluvial supérieur*	1974-2013
Ilfis	Langnau	188	685	1051	0 %	*nivo-pluvial préalpin*	1989-2013
Minster	Euthal, Rüti	59	894	1351	0 %	*nival de transition*	1974-2013
Plessur	Chur	263	573	1850	0 %	*nival alpin*	1974-2013
Somvixer-Rhein	Somvix, Encardens	22	1490	2450	6.7 %	*b-glacio-nival*	1977-2013
Suze	Sonceboz	150	642	1050	0 %	*nivo-pluvial jurassien*	1961-2013

METHODS

The method for the construction of synthetic design hydrographs (SDHs) relies on the fitting of statistical density functions to observed flood hydrographs considering the dependence between the design variables $\boldsymbol{Q}_{max}$ and $\boldsymbol{V}$. This is crucial, because a univariate frequency analysis of $\boldsymbol{Q}_{max}$ or $\boldsymbol{V}$ alone cannot provide an accurate evaluation of the corresponding probabilities. The bivariate analysis was implemented using a copula model (Genest and Favre, 2007). The entire method of the SDH construction is divided into six main steps (Figure 1).

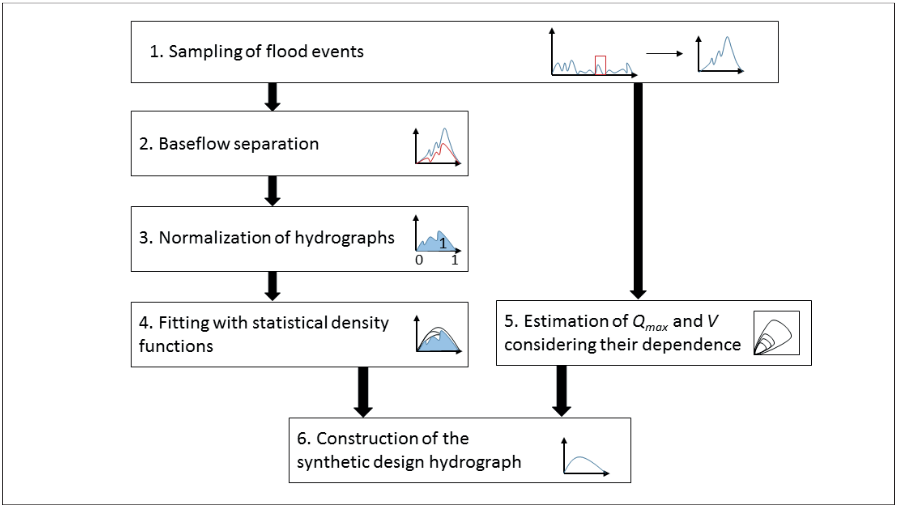

Figure 1: Overview of the method proposed for the construction of synthetic design hydrographs.

These steps include:

1. Flood sampling: The SDH construction method is based on observed runoff data. Therefore, historical flood event hydrographs were selected in nine Swiss catchments (Table 1) independently from precipitation information. We used a peak-over-threshold (POT) approach to sample flood events. The threshold for the peak discharge was chosen iteratively to fulfill a target condition of a defined number of events per year (here four). The independence between successive events was ensured by prescribing a minimum time interval between them (here 72 hours). According to the extreme value theory, POT values follow a generalized Pareto distribution (GPD) (Coles, 2001). Therefore, we used a GPD to fit the peak discharge values. The volume values, on the contrary, follow a generalized extreme value distribution (GEV) because the threshold was only applied to the peak discharge and not to the volume. The goodness-of-fit of the GPD to the peak discharges and the GEV to the volumes was found to be good using the Akaike and the Bayesian information criteria (Meylan et al., 2012).

2. Baseflow separation: The SDH approach describes only the quick flow component of the event hydrograph. Thus, it is necessary to distinguish between the slow and the fast runoff components to analyze the statistical properties of flood hydrographs (Yue et al., 2002). In this study, we applied a recursive digital filter (Eckhardt, 2005), whose two parameters need to be estimated for each catchment. This method allows for the separation of the baseflow from the quick flow and is easily applicable to a wide variety of catchments and provides reliable results.
3. Normalization of the hydrographs: The quick flow component of the hydrographs was normalized so that both the base width and the volume of the modified hydrographs were equal to one. This was done by dividing the base width of each flood hydrograph by its duration D and then dividing the ordinate of each hydrograph by the mean runoff (V/D).
4. Fitting of the normalized hydrographs with statistical density functions: The shape of a normalized hydrograph can be fitted by a probability density function (PDF) because both the area under the normalized hydrograph and the area under the PDF are equal to one. To select the best density, eight different PDFs were fitted to all of the normalized hydrographs in the nine study catchments: Normal, Lognormal, Fréchet, Weibull, Logistic, Gamma, inverse Gamma, and Beta (Nadarajah, 2007; Serinaldi and Grimaldi, 2011). The parameters of the distributions were estimated so that the PDFs approximate the shape of the normalized hydrographs as well as possible. The goodness-of-fit of each distribution function was ranked according to the three performance criteria: Nash-Sutcliffe efficiency (NSE), volumetric efficiency (VE), and correlation coefficient (R2). In each catchment, a representative normalized hydrograph was determined as the PDF that best fitted the median normalized hydrograph. We chose the median normalized hydrograph instead of the mean of the normalized hydrographs because it refers to a real event, which is not the case for the mean of the normalized hydrographs.
5. Modelling the statistical dependence between Q_{max} and V using copulas: The pair of design variables, Q_{max} and V associated with a defined return period T was estimated using the marginal distributions of the variables and a copula to model their dependence. For the marginal distributions, we assumed a GEV distribution for the flood volumes and a GPD for the peak discharges. Copulas are multivariate distribution functions whose marginal distributions are uniform. The main advantage of this approach is that the selection of an appropriate model for the dependence between variables, represented by the copula, can then proceed independently from the choice of the marginal distributions (Genest and Favre, 2007). In contrast to standard multivariate distributions, copula models thus allow the variables to be characterized by different marginal distributions. The dependence between the two variables Q_{max} and V was tested graphically by plotting all pairs of Q_{max} and V and numerically by computing two rank correlation coefficients, Kendall's Tau and Spearman's Rho. Six copula models of the Archimedean copula family, namely the Gumbel, Clayton, Joe, Frank, Ali-Mikhail-Haq (AMH), and the independence copula plus the normal copula, were fitted and tested using both graphical approaches and a goodness-of-fit test based on the Cramér-von Mises statistic (Genest and Favre, 2007). A p-value for

the Cramér-von Mises statistic of each copula was estimated using a statistical bootstrap procedure (Genest et al.,2009). The copula model with the best performance was used to estimate the two design variables for a given return period.

6. Construction of the SDH: The SDH was obtained from scaling the representative normalized hydrograph by the two estimated design variables using:

$$Q_T(t) = f(t)V_T/D_T$$

Formula 1

where $\boldsymbol{f}(\boldsymbol{t})$ is the representative normalized hydrograph, and $\boldsymbol{Q}_{max}$ and $\boldsymbol{V}$ are the design variables for a given return period $\boldsymbol{T}$. In bivariate frequency analysis, in contrast to the univariate case, the definition of an event with a given return period is not clear (Yue and Rasmussen, 2002). The return period used to describe bivariate events can be defined in two different ways. The first of these approaches uses the conditional probability to determine a conditional return period, while the second approach uses joint probability distributions to calculate joint return periods (Gräler et al., 2013). Here, we relied on a conditional probability approach to estimate the design variable pairs $\boldsymbol{Q}_{max}$ and $\boldsymbol{V}$ because the potential end-users of our approach are more familiar with conditional probabilities than joint probabilities. Thereby, the estimation of the variable $\boldsymbol{V}$ for the given return period was calculated according to its marginal distribution. Once the volume was estimated, the estimation of the second variable peak discharge could take place using the conditional cumulative distribution of the copula (Salvadori et al., 2011), as described in Step 5. The SDH was then calculated using the representative normalized hydrograph, scaled by the computed estimates of the design variables $\boldsymbol{Q}_{max}$ and $\boldsymbol{V}$. At the end of the procedure, the baseflow, removed from the hydrograph in Step 2 of the procedure, needs to be added to the SDH.

RESULTS

The proposed method was applied to a sample of nine meso-scale study catchments in Switzerland with an average record length of 40 years (Table 1). The results are illustrated by one of these study catchments, the Birse catchment at Moutier-la-Charrue.

Eight PDFs were fitted to the normalized hydrographs that were derived from the raw data. In the Birse catchment, the Fréchet and Logistic distributions were most often ranked as the best PDFs for all considered performance criteria (Figure 2). Interestingly, the Fréchet distribution together with the normal distribution was also most often ranked as the worst PDF. The fact that the same statistical distribution was ranked both as the best and as the worst PDF could be due to differences in the causative mechanism of the flood events. The tendency of the Fréchet and Logistic distribution to provide good results was also observed in the other eight study catchments.

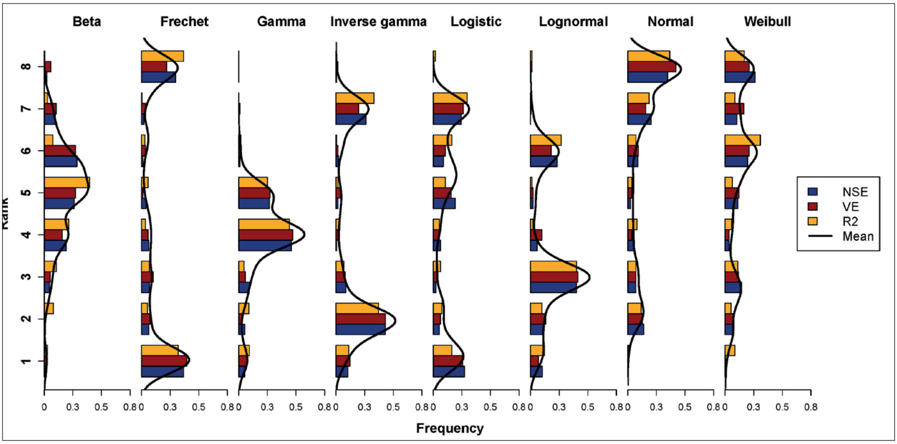

Figure 2: Distributions of the ranks obtained by the eight statistical densities fitted to the normalized hydrographs in the Birse catchment. For each event, the eight densities were ranked according to the three performance criteria NSE, VE, and R2.

Q_{max} and V were clearly dependent in the Birse catchment with a Kendall's Tau of 0.323 and a Spearman's Rho of 0.477. Figure 3 shows the observed pairs of Q_{max} and V (red crosses) and 10 000 pairs simulated using the seven fitted copula models (black crosses).

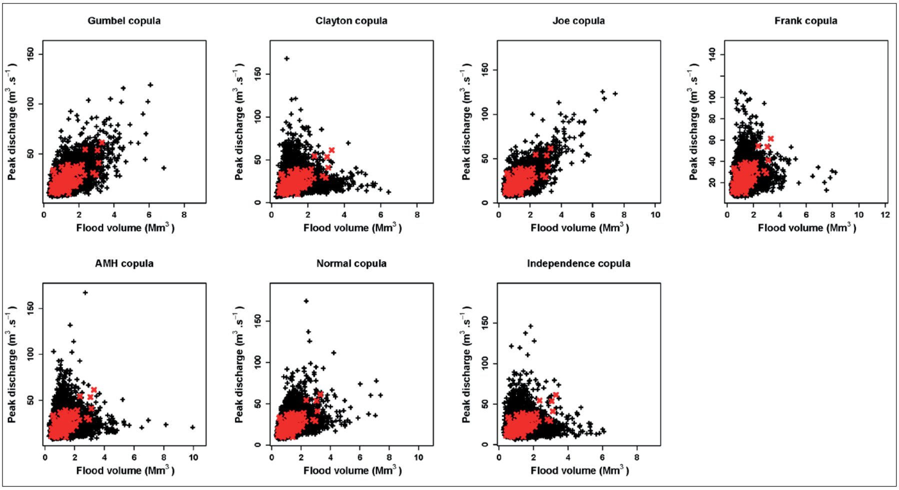

Figure 3: Representation of the bivariate distribution of the variables Q_{max} and V for the Birse catchment. The observations are plotted as red crosses. The black crosses represent 10 000 pairs of Q_{max} and V values simulated according to one of the seven copulas which were fitted to the data and tested subsequently.

Different behaviours among the modelled bivariate distributions can be observed. The Gumbel and Joe copulas tend to model a clear dependence between Q_{max} and V with extreme values expanding towards the upper-right corner of the plot while other copulas do not indicate a dependence and rather behave in a way similar to the independence copula. As a quantitative assessment of a copula's ability to represent the dependence between Q_{max} and V, a goodness-of-fit test based on the Cramér-von Mises statistic was applied. This confirmed that the Gumbel and Joe copulas are best suited to model the dependence between Q_{max} and V. The other tested copulas were rejected at a significance level of 0.05. This tendency is also visible in most of the other study catchments.

The SDHs were computed using seven pairs of design variables, Q_{max} and V, for the 100-year return period, computed with the seven different copulas mentioned above (Figure 4). The SDHs are associated with uncertainty envelopes which represent all the traces resulting of the fits made with different statistical densities. There was a clear difference between all the SDHs derived using different copula models (Figure 4). All calculated values of Q_{max} for a 100-year return period are larger than the value obtained by a univariate analysis (SDH independence). A univariate analysis would neglect the dependence between the two variables Q_{max} and V. This tendency was observed in all nine study catchments. In addition, most of the computed values of the pair Q_{max} and V exceed the highest recorded value.

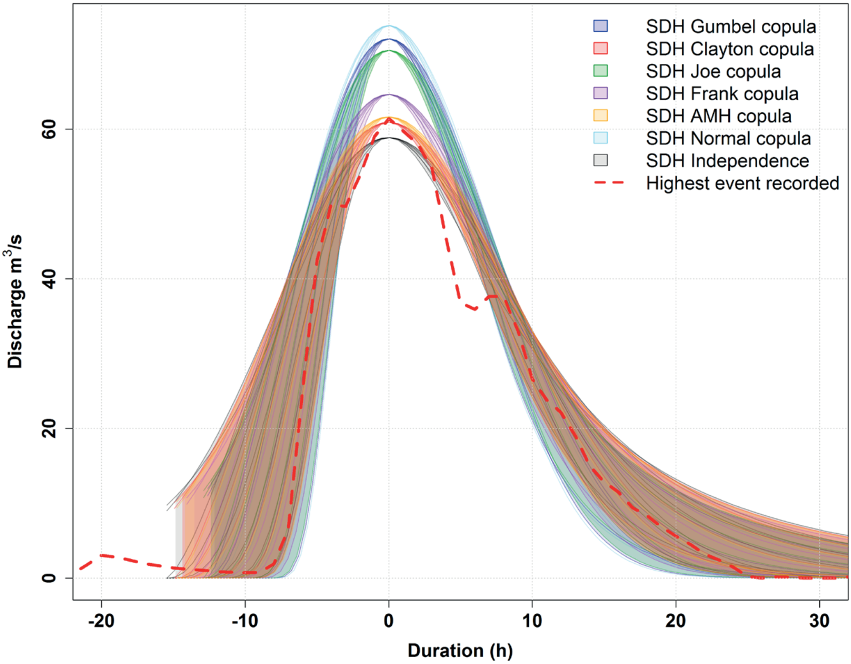

Figure 4: SDHs for an event of a 100-year return period in the Birse catchment. The SDHs were calculated with seven copulas. The envelopes correspond to the eight statistical densities used for the fit. The hydrograph of the highest event (in terms of peak discharge) observed in the Birse catchment is shown as a reference.

CONCLUSIONS & OUTLOOK

The method presented proved to provide reliable SDHs for all nine study catchments. Although the method was developed using a small number of catchments, it can potentially be applied in any meso-scale catchment with no or small glacierized areas and unaltered flow conditions for which observational data are available. The major advantages of the method are its ease of application, its independence from a rainfall-runoff model, and the possibility to account for the dependence between peak discharge and volume. The importance of considering the dependence between peak discharge and volume was clearly shown in all study catchments, where the use of a classical univariate approach would have resulted in a clear underestimation of the magnitude of the design flood hydrograph compared to the approaches where the dependence between the two design variables peak discharge and flood volume is considered.

In the next step, this method will be regionalized to ungauged catchments to also allow for the estimation of flood volume and hydrograph without any runoff measurements. Thus, several parameters used in the method need to be regionalized. To aid this regionalization, relationships between the modelled shape of hydrographs and a typology of causative flood mechanisms will be explored because different flood types are usually characterized by typical hydrograph shapes. Furthermore, relationships between catchment characteristics and parameters of the SDH will be quantified.

ACKNOWLEDGEMENTS

We thank the Swiss Federal Office for the Environment (FOEN) for funding the project under contract 13.0028.KP / M285-0623 and for providing observed runoff data. We also thank the reviewers for their constructive comments.

REFERENCES

- Coles S. (2001). An introduction to statistical modeling of extreme values. Springer, London, 208.
- Eckhardt K. (2005). How to construct recursive digital filters for baseflow separation. Hydrological Processes 19: 507-515.
- Genest C. and Favre A.-C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. Journal of Hydrologic Engineering 12: 347-367.
- Genest C., Rémillard B. and Beaudoin D. (2009). Goodness-of-fit tests for copulas: A review and a power study. Insurance: Mathematics and Economics 44: 199-213.
- Gräler B., van den Berg M. J., Vandenberghe S., Petroselli A., Grimaldi S., Baets B. D. and - Verhoest N. E. C. (2013). Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation. Hydrological Earth System Sciences 17: 1281-1296.
- Meylan P., Favre A.-C. and Musy A. (2012). Predictive hydrology. A frequency analysis approach. Science Publishers. St. Helier, Jersey, British Channel Islands, 212.

- Nadarajah S. (2007). Probability models for unit hydrograph derivation. Journal of Hydrology 344: 185-189.
- Rosbjerg D., Blöschl G., Burn D. H., Castellarin A., Croke B., Baldassarre G. D., 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. and Viglione A. (2013). Prediction of floods in ungauged basins. In Runoff prediction in ungauged basins. A synthesis across processes, places and scales (G. Blöschl, M. Sivapalan, T. Wagener, A. Viglione, and H. Savenije, eds.), pp. 189-226. Cambridge University Press, Cambridge.
- Salvadori G., DeMichele C. and Durante F. (2011). On the return period and design in a multivariate framework. Hydrological Earth System Sciences 15: 3293-3305.
- Serinaldi F. and Grimaldi S. (2011). Synthetic design hydrographs based on distribution functions with finite support. Journal of Hydrologic Engineering 16: 434-446.
- Yue S., Ouarda T. B. M. J., Bobée B., Legendre P. and Bruneau P. (2002). Approach for describing statistical properties of flood hydrograph. Journal of Hydrologic Engineering 7: 147-153.
- Yue S. and Rasmussen P. (2002). Bivariate frequency analysis: discussion of some useful concepts in hydrological application. Hydrological Processes 16: 2881-2898.