*European Water* 57: 19-25, 2017.


# Flood risk and flood processes in a changing environment

A.H. Schumann* and S. Fischer
*Ruhr- University Bochum, Institute of Hydrology and Water Management, D-44801 Bochum, Germany*
* *e-mail: andreas.schumann@rub.de*

**Abstract:** Driven by the EU- Flood Directive, flood risk management became one of the focusing points of water policy. Flood risk means the combination of the probability of a flood event with its adverse consequences. Unfortunately, the consequences of a changing environment on floods are hidden by high uncertainties about the stochastic behavior of flood inducing processes and their possible changes. Tools, which are able to combine statistical assessments with deterministic considerations, are rare. In this contribution, the problem of instationarity is discussed to demonstrate the need to close the gap between stochastic and deterministic flood hydrology. Some methodological tools will be presented which are useful for this purpose.

**Key words:** Floods, multivariate statistics, timescales, flood types, peak over thresholds

## 1. INTRODUCTION

In the last 25 years, floods moved into the focus of European water management. This resulted from a large number of catastrophic events that happened in the last 25 years in several parts of Europe. The political goal of flood management was propagated in the European Flood Directive (EFD), which aims to reduce the likelihood of flooding as well as its harmful consequences. By definition flood risk management has to consider the eventuality of flood damages, even if the causing floods would have "a low probability" or belong in the category of "extreme event scenarios" (European Commission 2007). According to the Merriam- Webster Dictionary the word "management" means not only the act of managing: "the conducting or supervising of something" but also the "judicious use of means to accomplish an end". In this sense flood risk management is aimed at managing whole flooding systems by their catchments in an integrated way that accounts for all of the potential interventions that may alter flood risk (Sayers et al. 2002). At this extended spatial scale, their effects on local, regional and supra-regional flood risks may differ and often depend on event-specific interactions between tributaries and the alterations of floods by other interventions in the flood regime, which may differ in their impacts between events also (Schumann, 2017). However, if the specification of effective measures for flood management is essential, assessments and comparisons of economic efficiencies are also indispensable. With regard to the stochasticity of floods, this efficiency of flood protection measures depends strongly on the exceedance probabilities of design floods. The economic aspects of flooding can be assessed using the Expected Annual Damage (EAD) concept (e.g. Chow et al. 1988). The reduction of damage costs by flood management depends on the choice of the design, which is selected to adapt the system, but also on the expected changes of the damage curve after adaptation (Rosbjerg 2017). Both problems mentioned above, the improved characterisation of hydrological loads to specify the performance of flood management measures and the estimation of flood probabilities under consideration of the multivariate flood characteristics, demand new methodological approaches in flood statistics. It is insufficient to specify floods by the exceedance probabilities of their peak flow only. The variability between flood events with similar peaks, but different hydrograph and volumes or the spatial differences of flood formation between events becomes more and more relevant if we try to characterize options and limitations of flood management in a more realistic way than by single design floods only.

## 2. MULTIVARIATE OR PROCESS-BASED STATISTICS?

For flood retention, the assessment of flood volumes for design floods, which are statistically specified by the exceedance probability of their peak flows, is essential. One main application of copulas is to support the sizing of a design flood to consider the dependencies between peak and volume (De Michele et al. 2005; Shiau et al. 2006; Klein et al. 2010). With copulas, a joint distribution of correlated random variables can be expressed as a function of the univariate marginal distributions. Copulas allow modeling of the dependence structure of two (or more) random variables. This structure often shows a variability over the range of flood peaks, which is an indicator for different flood processes. To analyze these differences, the "flood timescale" that was proposed by (Gaál et al. 2012) can be applied. It is the quotient of the flood volume and flood peak. In another study (Gaál et al. 2015) analyzed the consistency of the peak–volume relationship quantified by the Spearman rank correlation coefficient for a large sample of watersheds in Austria. The main conclusion of this study is that a mix of different flood types reduces the consistency between flood peaks and volumes. The authors explain a weak dependence in the high alpine catchments by the mix of flood types, including long-duration snowmelt, synoptic floods and flash floods, which are a characteristic of these regions. An important conclusion of this publication is that climate-related factors are more important than catchment-related factors in controlling the consistency between peaks and volumes. We use these results to go a step further. By an analysis of the correlation between peaks and volumes, we differentiated the monthly flood peaks by groups, which characterize certain types. In Fig. 1 an example is given.

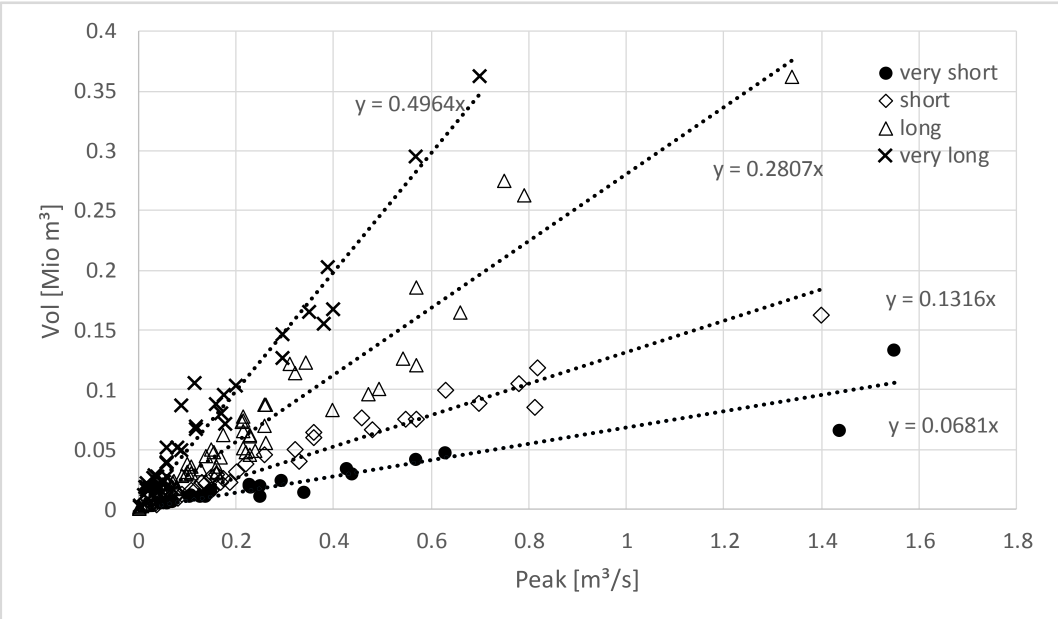


*Figure 1. Relationships between peak and volume (direct runoff only) for floods at the gauge Derenburg/ Hellbach in Germany.*

According to the timescale we differentiated into four groups from "very short" to "very long". There are some accumulations of short events in the months June to August (convective storms) and of very long events in March to April (snowmelt). However, the most of the floods happen in different seasons of the year. This result is similar to the experience of other authors (Brunner et al. 2016). As a result, we can separate samples of flood peaks according to their event-type. With a POT- approach, where different thresholds are used for the different flood types, probability distribution functions can be estimated for all of these samples. These distributions can be combined by a methodology, which was proposed by (Zhou et al. 2016). This mixing POT-model is an extension to the classical POT-models using the Generalized Pareto distribution (see e.g. Fischer and Schumann, 2016). The common distribution of the $n_t$ different event types is given by:

$$F(x) = 1 - \sum_{i=1}^{n_t} \left(1 - G_i(x;\theta_i, u_i)\right)\left(1 - F_i(u_i)\right)\omega_i \qquad (1)$$

where $G_i$ is the Generalized Pareto distribution with parameter $\theta_i$ fitted to event type i exceeding a threshold $u_i$ and $F_i$ is the marginal distribution of this event type. The weighting $\omega_i$ corresponds to the probability of occurrence of the event type. It estimates the joint distribution of flood peaks under consideration of the frequencies of events of certain types. In total, we get a statistics of yearly flood peaks but also the probabilities that a flood peak is caused by a certain flood type and the expected values of the flood volumes. Compared with the multivariate copula-based approach, this differentiation of the entity of observed floods into flood types improves our opportunities to specify the upper reach of the copula under consideration of physical boundaries and a shift of flood type frequencies within the range of observations. Examples are given in Fig. 2 and 3.

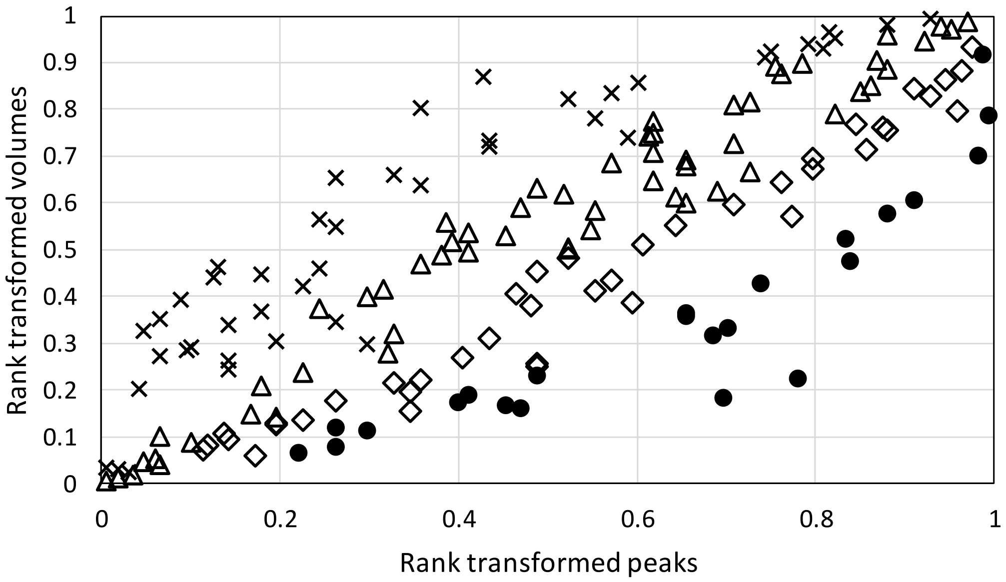


*Figure 2. Scatter plot of rank transformed volumes above rank transformed peaks for flood events at the gauge Derenburg/ Hellbach in Germany. The marking of flood types is the same as in Fig. 1*

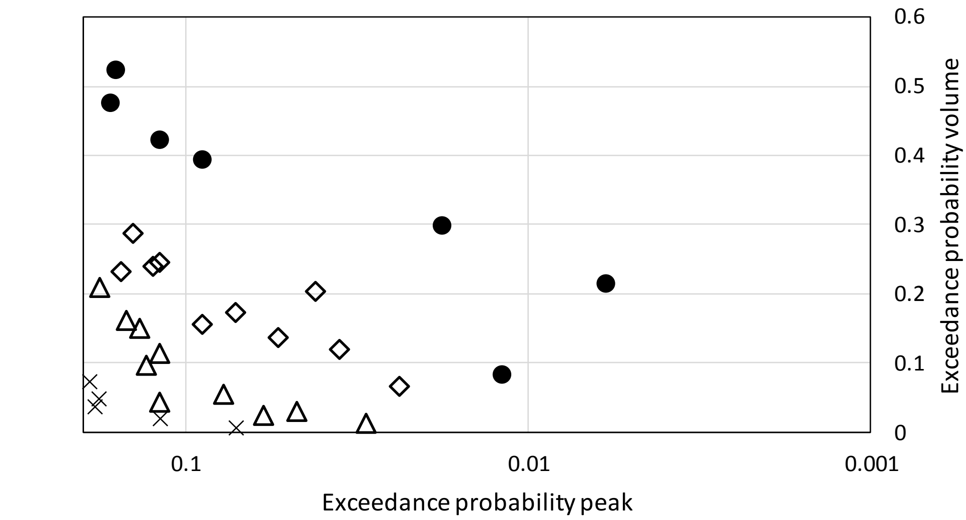


*Figure 3. The variety of flood volumes for the highest flood events of all four types, gauge Derenburg/ Hellbach*

It can be seen that in the range of peaks with exceedance probabilities above 0.1, estimated from all events above the thresholds, the categories "very short" and "short" floods are much more represented than events belonging to the long categories. The volumes of these events have higher exceedance probabilities. It has to be emphasized that the mixed POT- approach results in annual flood probabilities, which are nearly the same as derived from the classical Annual Maxima Series (AMS) by application of the Generalized Extreme Value distribution and a parameter estimation with L- moments. By the proposed methodology, it becomes possible to specify the dominating flood type for this catchment and to consider this result in flood regionalization.

## 3. CHANGING FREQUENCIES OF FLOOD TYPES

The characteristics of flood types become highly relevant for impact assessments of climatic change. The relevance of global warming on snowmelt and rain-on-snow floods is obvious and was discussed in several studies (e.g. Köplin et al. 2014; Støren et al. 2012; Köplin et al. 2014; Vormoor et al. 2015). But also changes of other climatic characteristics as the frequencies of large- scale circulation pattern (Jacobeit 2003) or of the soil moisture state (Nied et al. 2014) could have impact on flood statistics. Such changes could affect the frequencies of floods with specific types only. Especially for tropical extreme events observations reveal a distinct link between rainfall extremes and temperature, with heavy rain events increasing during warm periods and decreasing during cold periods (Allan and Soden 2008). Flash floods could happen more often caused by short and intensive rain events in the summer. How an increase of frequencies of short but intensive rainfall events would affect the flood statistics in summer will be demonstrated in the following. For this study, we use the "maximum Winter-Summer-Types approach" (WST) (Fischer et al., 2016). It is based on a subdivision of the year in a summer and a winter period. By multiplication of non-exceedance probabilities in winter and summer, the probability of a flood peak can be estimated. The annual maximum is, in any case, the maximum of the summer or winter annual maximum. Vice versa, if the annual maximum $X_A$ does not exceed some value x, then the summer and winter annual maximum must not exceed this value. Thus, the annual non-exceedance probability can be estimated multiplicatively from the probabilities of winter $X_W$ and summer $X_S$ peaks (Todorovic and Rouselle, 1971):

$$P(X_A \le x) = P(X_S \le x) \cdot P(X_W \le x) \tag{2}$$

As described above, several flood types could be identified for both seasons by their timescales. With regard to the question raised above, we differentiate here between long and short summer events only. The statistics of winter events will not be modified. For this purpose, the seasonal mixing model was extended to a model that considers the different distributions of winter and summer floods as well as the differences between short and long summer events. Considering several flood types (long and short floods) in summer demands a statistical mixing model to specify the resulting flood probabilities in summer. The maximum mixing model benefits from the fact that the non-exceedance of a summer flood peak x demands that it is not exceeded by any short or long flood event in this half-year. Thus, the multiplicative approach of Todorovic and Rouselle (1971) can be applied as follows:

$$P(X_A \le x) = \left(P(X_{S,l} \le x) \cdot P(X_{S,s} \le x)\right) \cdot P(X_W \le x) \tag{3}$$

where $X_{S,s}$ and $X_{S,l}$ are the flood peaks of short and long summer runoffs.

For flood peaks in winter we know the complete sample. This is not the case if we differentiate summer maxima into short and long summer floods. As only a member from one of both groups is the maximum of the summer half-year, we know just the highest summer flood per year which can be characterized by a long or by a short timescale. Smaller values are overlaid. To estimate the distributions of long and short flood events, we have to correct the statistical parameters of both subsamples. For this purpose we applied a method of “filling” both data series.

It can be shown by statistical simulations that maxima series resulting from two different distributions are always dominated by the distribution that has the larger location parameter. If all other parameters (scale and shape) are equal, shares of this dominating series of approximately 97% can be obtained. If the other series has a larger scale, this relation is diminished, but the series with the larger location parameter still dominates the maxima series (about 80%). The shape parameter has only a small influence on the shares of the two distribution functions in this series. This leads to the following procedure to adjust statistical parameters. The shares are estimated by counting the number of events belonging to one or the other distribution that occur in the maxima series. As only a member from one of both groups is the maximum of the summer half-year, we only know the highest summer flood per year which can be characterized by a long or by a short timescale. We estimated the location parameter $\mu_S$ of the whole summer annual maxima series. We know that the event type with a higher share in this sample determines this parameter. For this part of the sample we used the location parameter $\mu_S$, for the other part which has a relative share RS (with RS<0.5) we replaced the location parameter with $(RS+0.5)\cdot\mu_S$. The more equal the shares of both event types within the summer series, the more equal both location parameters are to the location parameter of the summer series. To replace the missing values, we can use the fact that for the subseries that dominates the summer annual maxima, the missing values scatter around the parameter $\mu_S$, because none of the overlaid events belongs to the sample of large events, which were observed. No events from the right tail occur in the series of missing values (these are already included in the non-overlaid series), so we can assume that these are distributed symmetrically. There is no option to estimate the variance of this subsample directly. Relatively large events are possibly overlaid by large events belonging to the other subsample. That is why we considered a spectrum around the location parameter to introduce an artificial variance. We used the robust estimator MAD (median absolute deviation) to estimate this variance from the known part of this series which is based on the median of each subsample. Within the interval $[RS\cdot\mu_S\text{-MAD}; RS\cdot\mu_S\text{+MAD}]$ resp. $[\mu_S\text{-MAD}; \mu_S\text{+MAD}]$ the missing values in the two subsamples were chosen from a uniform distribution. Further details are given in (Fischer et al., 2016).

In our case study the increased frequency of short and intensive rain events in summer results in a higher frequency of floods with high peaks and short timescales. This leads to a reduced skew but a higher location and scale of the distribution of such events since the mid of the distribution increases. That is, the right tail of the distribution becomes “more probable”. Simultaneously the location of the long summer is reduced since the long events have less magnitude. The proposed statistical model can consider such changes. To demonstrate this we use an example of the flood statistics at the gauge Berthelsdorf, which is located at the Freiberger Mulde River in Germany. The relationship between short and long flood events at gauge Berthelsdorf is 0.4 to 0.6. We invert this relationship now. As a result the parameters of the distributions of flood peaks in summer are changed. For short events we’ve got a reduced skew but a higher location and scale of the distribution, for long events the location was reduced since the long events have less magnitude. The modification of the mean $\mu$, standard deviation $\sigma$ and skewness $\xi$ are shown in Table 1.

*Table 1. Parameters of the GEV distribution for short and long summer events at the gauge Berthelsdorf/ Freiberger Mulde under assumed climatic change*

| | **Current conditions** | | **Climate change scenario** | |
|---|---|---|---|---|
| | summer (short) | summer (long) | summer (short) | summer (long) |
| $\mu$ | 11.718 | 14.307 | 12.7 | 10.307 |
| $\sigma$ | 3.988 | 7.416 | 6 | 7.416 |
| $\xi$ | 0.752 | 0.32 | 0.7 | 0.32 |

With these parameters, short and long summer series were simulated and combined via the maxima to obtain a series of half- years maxima in summer. The simulated series of winter and summer floods were combined to a simulated annual maximum series, which is needed for the statistics of annual flood peaks. To compare the changes of return periods under the climate change scenario, the annualities of the five largest observed floods (note that these were summer events) are listed in Table 2.

*Table 2. Estimated annualities of the five largest floods at the gauge Berthelsdorf/Freiberger Mulde before and after an assumed climatic change*

| | **Current conditions** | | **Climate change scenario** | |
|---|---|---|---|---|
| **HQ [m³/s]** | **AMS** | **Maximum WST** | **AMS** | **Maximum WST** |
| 68.2 | 11 | 17 | 10 | 11 |
| 120 | 35 | 62 | 35 | 34 |
| 122 | 41 | 64 | 36 | 35 |
| 140 | 55 | 83 | 50 | 44 |
| 360 | 476 | 375 | 461 | 198 |

It becomes obvious how the increased frequency of short summer floods, which are critical for the flood regime in this watershed, results in decreased return periods. For the highest observed flood (360 $m^3/s$) peak, which was caused by a short summer event in the year 2002, the return periods estimated with the AMS model would not change. Using the seasonal approach and considering more frequent events of this type in summer, the WST approach estimated higher exceedance probabilities. For this flood event, the return period would be halved.

## 4. CONCLUSIONS

The basic assumption of flood statistics is that the observations are homogeneous, i.e., subject to a common set of forces. This is often not the case. Snow-, rain-on-snow, convective or synoptic rain flood events are mixed in annual series of maxima. If the different types are separated, we gain a better understanding about the main impacts on hydrological aspects of flood risk assessments. Floods can be specified by multiple event characteristics (peak, volume, hydrograph). This can be useful for planning of flood protection measures, e.g. for retention facilities. But also aspects of climate variabilities, which affect only some flood types, could be considered in a better way. In this sense, process knowledge can improve and adapt flood statistics to new challenges.

## REFERENCES


Yapo, P. O., Gupta, H. V., Sorooshian, S. (1998). Multi – objective global optimisation for hydrologic models. Journal of Hydrol., 204: 83-97.

Allan, R.P.. Soden, B.J. (2008). Atmospheric warming and the amplification of precipitation extremes. Science, 321(5895): 1481–1484. DOI: 10.1126/science.1160787.

Brunner, M.I., Viviroli, D., Sikorska, A.E., Vannier, O., Favre, A.-C., Seibert, J. (2016). Flood type-specific construction of synthetic design hydrographs. Water Resour. Res., DOI: 10.1002/2016WR019535.

Chow, V.T., Maidment, D.R., Mays, L.W. (1988). Applied hydrology. Series in water resources and environmental engineering, McGraw-Hill, New York

De Michele, C., Salvadori, G., Canossi, M., Petaccia, A., Rosso, R. (2005). Bivariate Statistical Approach to Check Adequacy of Dam Spillway. In: J. Hydrol. Eng., 10(1): 50–57. DOI: 10.1061/(ASCE)1084-0699(2005)10:1(50).

European Commission (2007). Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the assessment and management of flood risks. Online available: http://ec.europa.eu/environment/water/flood_risk/index.htm

Fischer, S., Schumann, A.H. (2016). Robust flood statistics - comparison of classical annual maxima and Peak over Threshold approaches. Hydrological Science Journal, 61(3): 457-470, DOI:10.1080/02626667.2015.1054391

Fischer, S., Schumann, A., Schulte, M. (2016). Characterisation of seasonal flood types according to timescales in mixed probability distributions. Journal of Hydrology 539: 38–56. DOI: 10.1016/j.jhydrol.2016.05.005.

Gaál, L., Szolgay, J., Kohnová, S., Hlavčová, K., Parajka, J., Viglione, A. et al. (2015). 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.

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 Resour. Res., 48(4), DOI: 10.1029/2011WR011509.

Jacobeit, J. (2003). Links between flood events in central Europe since AD 1500 and large-scale atmospheric circulation modes. Geophys. Res. Lett. 30(4). DOI: 10.1029/2002GL016433.

Klein, B., Pahlow, M., Hundecha, Y., Schumann, A. (2010). Probability Analysis of Hydrological Loads for the Design of Flood Control Systems Using Copulas. J. Hydrol. Eng., 15 (5): 360–369.

Köplin, N., Schädler, B., Viviroli, D., Weingartner, R. (2014). Seasonality and magnitude of floods in Switzerland under future climate change. Hydrol. Process., 28(4): 2567–2578. DOI: 10.1002/hyp.9757.

Nied, M., Pardowitz, T., Nissen, K., Ulbrich, U., Hundecha, Y., Merz, B. (2014). On the relationship between hydro-meteorological patterns and flood types. Journal of Hydrology, 519: 3249–3262. DOI: 10.1016/j.jhydrol.2014.09.089.

Rosbjerg, D. (2017). Optimal adaptation to extreme rainfalls in current and future climate. Water Resour. Res. DOI: 10.1002/2016WR019718.

Sayers, P.B., Hall, J.W., Meadowcroft, I.C. (2002). Towards risk based flood hazard management in the UK. In: Proceedings of the ICE, Civil Engineering 150(5): 36–42. DOI: 10.1680/cien.150.5.36.38631.

Shiau, J.-T., Wang, H.-Y., Tsai, C.-T. (2006). Bivariate Frequency Analysis of Floods Using Copulas. Journal of the Americal Water Resources Association 42(6): 1549–1564.

Støren, E.N., Kolstad, E.W., Paasche, Ø. (2012). Linking past flood frequencies in Norway to regional atmospheric circulation anomalies. J. Quaternary Sci., 27(1): 71–80. DOI: 10.1002/jqs.1520.

Vormoor, K., Lawrence, D., Heistermann, M., Bronstert, A. (2015). Climate change impacts on the seasonality and generation processes of floods - projections and uncertainties for catchments with mixed snowmelt/rainfall regimes. Hydrol. Earth Syst. Sci. 19(2): 913–931. DOI: 10.5194/hess-19-913-2015.

Zhou, X.-Y., Schmidt, F., Toutlemonde, F., Jacob, B. (2016). A mixture peaks over threshold approach for predicting extreme bridge traffic load effects. Probabilistic Engineering Mechanics 43: 121–131. DOI: 10.1016/j.probengmech.2015.12.004.