Reducing uncertainty in the selection of bi-variate distributions of flood peaks and volumes using copulas and hydrological process-based model selection

Jan Szolgyay (1), Ladislav Gaál (1,2), Tomáš Bacigál (3), Silvia Kohnová (1), and Günter Blöschl (4)

(1) Department of Land and Water Resources Management, Faculty of Civil Engineering, Slovak University of Technology, Radlinského 11, SK 81005 Bratislava, Slovakia, (2) MicroStep-MIS, spol. s r. o. Cavojského 1, 84104 Bratislava, Slovakia, (3) Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Slovak University of Technology, Radlinského 11, SK 81005 Bratislava, Slovakia, (4) Institute for Hydraulic and Water Resources Engineering, Vienna University of Technology, Karlsplatz 13/223, A-1040 Vienna, Austria

Bi-variate distributions of flood peaks and flood event volumes are needed for a range of practical purposes including e.g. retention basin design and identifying extent and duration of flooding in flood hazard zones. However, the selection of the types of bi-variate distributions and estimating their parameters from observed peak-volume pairs are associated with far larger uncertainties compared to uni-variate distributions, since observed flood records of required length are rarely available. This poses a serious problem to reliable flood risk estimation in bi-variate design cases.

The aim of this contribution was to shed light on the possibility of reducing uncertainties in the estimation of the dependence models/parameters from a regional perspective. The peak-volume relationships were modeled in terms of copulas. Flood events were classified according to their origin. In order to reduce the uncertainty in estimating flood risk, pooling and analyzing catchments of similar behavior according to flood process types was attempted. Most of the work reported in the literature so far did not direct the multivariate analysis toward discriminating certain types of models regionally according to specific runoff generation processes.

Specifically, the contribution addresses these problems:
- Are the peak-volume relationships of different flood types for a given catchment similar?
- Are the peak-volume dependence structures between catchments in a larger region for given flood types similar?
- Are some copula types more suitable for given flood process types and does this have consequences for reliable risk estimation?

The target region is located in the northern parts of Austria, and consists of 72 small and mid-sized catchments. Instead of the traditional approach that deals with annual maximum floods, the current analysis includes all independent flood events in the region. 24 872 flood events from the period 1976-2007 were identified, and classified as synoptic, flash and snow-melt floods. First, empirical copulas for the individual processes were compared at each site separately in order to assess whether peak-volume relationships are different for different flood processes. Next, the similarity of empirical distributions was tested in a regional perspective process-wise. In the last step, the goodness-of-fit of frequently used copula types was examined both for process based data samples (the current approach, based on a wider database of flood events) and annual maximum floods (the traditional approach that makes use of a limited number of events). It was concluded, that in order to reduce the uncertainty in model selection and parameter estimation, it is necessary to treat flood processes separately and analyze all available independent floods.

Given that usually more than one statistically suitable copula model exists in practice, an uncertainty analysis of the design values in engineering studies resulting from the model selection is necessary. It was shown, that reducing uncertainty in the choice of model can be attempted by a deeper hydrological analysis of the dependence structure/model’s suitability in specific hydrological environments or by a more specific distinction of the typical flood generation mechanisms.