Hydrol. Earth Syst. Sci. Discuss., 8, C3147–C3149, 2011
www.hydrol-earth-syst-sci-discuss.net/8/C3147/2011/


Hydrology and Earth System Sciences Discussions

# *Interactive comment on* "Multivariate design via Copulas" *by* G. Salvadori et al.

**B. Gräler**

ben.graeler@uni-muenster.de

Received and published: 2 August 2011

The authors Salvadori et al. present an approach to introduce a total order in the multidimensional euclidean space induced by the Kendall distribution function of a copula. The theory is well explained and accompanied with an exemplary application to a three dimensional data set. In this example, annual maximum flood volumes, maximum flood peaks and the corresponding initial water levels of the dam are used.

In the univariate case, the notion of extremes (in the applied meaning) coincides with the natural total order relation: larger values denote more extreme and as such more dangerous events. Thus, it is immediate to build a dam that withstands the lower 99.9% of possible values (given annual maxima and a return period of 1000 years). This approach fails in the multivariate case due to the lack of a unique, natural total



ordering in the multidimensional euclidean space, as pointed out by the authors.

A flood is characterized by several parameters, thus, it is desirable to have a multivariate notion of larger floods. Given such an ordering, one can then design a dam to withstand the lower (or less dangerous) 99.9% of all possible events in the multivariate space. In contrast to the univariate case, the area/volume uniting these 99.9% of all possible events can unfortunately be shaped in many ways.

Considering the same trivariate example of the paper, the shape of the sub-critical region induced by $K_C$ for a suitable copula (as introduced by the authors) corresponds to the lower left volume of the unit cube bounded by an iso-surface of the copula (as presented in Figure 4). By construction, this volume will cover approximately 99.9% of all realizations in the long run. The authors suggest to select a single trivariate point $(u_1, u_2, u_3)$ from this critical layer in order to determine a single design event. However, the probability to exceed one or more of the margins is given by $1 - C_{QVL}(u_1, u_2, u_3) = 1 - t^* = 0.053481$, no matter which specific design event is used in the given example. Hence, a simulation of 1000 yearly maximum floods will yield about 53 floods that exceed one or two of the three margins. Depending on the sensitivity of the designed dam to either higher flood peaks, higher volumes or higher initial water levels, this approach may or may not give the right design event.

A similar statement applies to the approach where the total order is derived by the copula itself. The critical layer is then given by the surface of the cuboid with one corner at $(0, 0, 0)$ and another one at the iso-surface $(u_1, u_2, u_3)$ with $C(u_1, u_2, u_3) = t_C = 0.999$. In this approach, any point on the iso-surface can be chosen to derive a design event yielding different critical layers. As before, no rule for a unique choice can be given. However, it is ensured that none of the margins is exceeded with a probability of $0.999$ for any point on the iso-surface. Given a sufficiently low value in one margin, other margins may exceed their threshold without being an extreme (or harmful) event. Hence, the remaining fraction of 0.001 may include harmless events, depending on the real threat to the dam. In this case, marginals tend to be overestimated.

Despite the afore mentioned problems, the presented new approach is a natural choice from a probabilistic perspective. Any point in the sub-critical region has a smaller joint cumulative distribution function value than any point in the super-critical region. However, it is highly questionable that a single point is able to sufficiently represent the properties of all critical points. Given the above commentary, I recommend to use a design ensemble along the critical layer induced by the Kendall distribution (as also suggested by Vandenberghe in an earlier comment). Such an ensemble will improve the representation of the critical layer. An ensemble of design events reduces the risk of underestimating the quantiles of the marginal distributions but adds an additional probabilistic process. This additional process cannot be accounted for completely and an additional confidence level has to be introduced, potentially decreasing the return period of the critical design events under consideration.

The above discussion illustrates the importance of the right choice of a total ordering. In applied sciences, this order should be related to the applied meaning of a larger (more extreme) event in nature. The authors present a sound and valuable alternative to define a rather natural total order in a multidimensional euclidean space. However, the selection of the *right* ordering in every application and i.e. the choice of a particular design event or a design ensemble will remain an active topic of research. An open and wider discussion of this issue in the last part of the presented paper would have been desirable.

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 8, 5523, 2011.