Hydrology and
Earth System
Sciences
Discussions

Interactive comment on "Defining flood risk in a multivariate framework: Application on the Panaro watershed" *by* Eleni Maria Michailidi and Baldassare Bacchi

F. Serinaldi (Referee)

francesco.serinaldi@ncl.ac.uk

General comments

I apologize in advance if the tone of this report will be a little bit harsh in some parts, but unfortunately reading this paper was very unpleasant. Even though I am not English native, I have to say that the language is very poor (especially for a document submitted in an open review journal), the text confused, confusing, and badly organized, and finally the content is almost uniformative. I will be even more explicit. I am reviewing more and more papers written by first- or second-year PhD students or even master students. Based on a quick web research, I think this is one of these cases. Since students have little experience by definition (nessuno nasce 'imparato' = no one was

born 'learnt'), if a submitted paper shows poor language, bad organization, confuse discussions, and even a confused abstract (which is the business card of a research document), the responsible is the senior author, who did not exerted the necessary control. In this respect, I am rather disappointed because a bad paper can be a serious damage in terms of reputation for young researchers at the early stage of their career, while they have limited impact for tenured co-authors. The latter should feel this responsibility when they revise documents (if they do this. . . and they should) before submission.

As mentioned above, the abstract itself is quite confused and likely reflects a lack of focus on the problem under study. In fact, the Authors talk about return periods, multivariate settings, inaccuracy of univariate analysis, making also examples in the abstract, while the content can easily be summarized in a few lines: "In this study, the Authors derive the 'univariate' empirical cumulative distribution function of reservoir level MWL by routing simulated flood hydrographs whose peak (P) and volume (V) preserve their correlation. Moreover, they explore the relationship between the pairs (P, V) corresponding to MWL values with specific 'univariate' return period, investigating the impact of the shape of the dimensionless typical hydrograph on such a relationship". That's all.

Now, it seems that the Authors are aware of some of my papers, but some aspects seem to be missing, thus generating substantial confusion in the discussion. As many other papers, this work devotes much space to describe standard copula inference procedures (some of which are ineffective or incorrectly applied), and overlook the main and most interesting aspects, i.e., the actual nature of the problem at hand, and uncertainty.

As far as the first aspect is of concern, the Authors should recognize that the multivariate return periods do not enter at all into the studied problem, and multivariate frequency analysis is rather secondary owing to the high uncertainty. In more depth, in every design procedure we have usually a unique target (design) variable and the associated probability of exceedance (the word 'risk' in risk analysis has a different meaning

and, as the Authors know, it involves exposure and vulnerability, at least). When there are multiple variables, multi-criteria optimization should be used; in this case, the target variable is usually the cost of the infrastructure (or whatever else), and the probability of exceedance is calibrated on the damage. The proliferation of more or less exotic conditional/multivariate RPs should be revised in light of design purposes that hydrologists and engineers should know very well. In this context, AND, OR, Kendall RPs do not matter very much because the underlying probabilistic models do not describe your problem. What actually matters is the structural approach proposed (without this name) by De Michele et al (2005) in their seminal work and revitalized more recently by Volpi and Fiori (2014) (see also Salvadori et al. (2015) for another recent application). This approach is the only one that allows operational choices (this is the Authors' choice indeed) and it is basically univariate, as P and V are only nuisance input variables.

The discussion above can seem purely conceptual or excessively critical; however, it explains why the Authors do not successfully show what they want in this paper. In fact, their aim is simply to define the pairs (P, V) corresponding to MWL values with a specified univariate return period (... in my opinion 'annual probability of exceedance' is less misleading, considering the widespread confusion on return periods, testified by this paper as well) which are linked by a generic (structural) function, namely $MWL_p = g(X_{MWL_p}, Y_{MWL_p})$. The function g can be simple, analytical and invertible, or very complex. In the first case analytical solutions are possible (Volpi and Fiori, 2014), while in the latter, Monte Carlo simulations can solve the problem (Salvadori et al. (2015), Serinaldi (2016)). In this contest, the only part of possible interest in the present paper is Section 4.3. However, the Authors devoted almost all the text to copula inference (which is better explained and applied in hundreds of papers), instead of thinking a bit more about the actual problem, and setting up a suitable simulation strategy. In fact, the lack of three clear patterns corresponding to the three hydrograph shapes, in Figure 9, is due to the ineffectiveness of the simulation procedure. With a suitable simulation, three almost parallel curves would likely emerge.

C3

Fortunately, a suitable algorithm is provided by Serinaldi (2016) where I suggested how to effectively simulate pairs X, Y corresponding to specified Z quantiles, provided that $Z = g(X, Y)$. Even though the example in that paper is quite simple, the procedure is general and holds true for whatever situation (under the usual minimal conditions), by replacing e.g. g with the reservoir routing rule. Moreover, that algorithm takes into account the parameter uncertainty due to sampling variability (see also Serinaldi (2013), and Dung et al. (2015)); thus, it allows for the definition of both level curves (i.e. what the Authors are looking for), and uncertainty areas around them.

As mentioned above, sampling uncertainty is another aspect neglected in this paper. In my opinion, when dealing with 52 pairs of data points, instead of discussing Chi-plots, K-plots, or supposed asymptotic properties (... frankly, do you really believe that 52 data points can even vaguely give information on whatever asymptotic property?), you should quantify at least the sampling uncertainty. In this respect, all the inference procedure is performed without reporting neither confidence intervals of point estimates nor the performance of other models. My guess is that more than one marginal and copula passed the tests and probably selection relied on the smallest value of AIC and/or BIC (P8L28) and largest p-values of GoF tests. If it is so, consider that (i) information criteria are not absolute indices but relative indices (i.e. selection should be based on Akaike weights, evidence ratios, etc.; see e.g. Burnham and Anderson, (2002;2004)), and the applied GoF tests are not distribution-free (i.e. the distribution of the test statistics changes for each family and each set of parameter values, thus making a comparison impossible).

To summarize, my suggestion is to focus on engineering problems instead of contributing to what Vit Klemes defined 'dilettantism in hydrology', referring to the abuse and misuse of statistics and 'mathematistry' i.e. 'the unscientific concept of mindless fitting that dominates contemporary hydrologic modelling'. Please, help to make 'dilettantism in hydrology' a transition rather than destiny. From a more practical point of view, the paper does not introduce any methodology and does not shed any light on any hydrological problem. As mentioned above, the 90% of the manuscript reports (superficial)

C4

inferential analysis, neglecting the intrinsic lack of information coming with 52 observations, while the actual problem is relegated in half a page before conclusions, without recognizing that it was already addressed in the literature (accounting for the uncertainty mentioned above). In this respect, I would like to stress that the bibliography in the ICSH website is not very much up-to-date, unless the Authors really think that only five (one) papers on the topic were published in 2015(2016). I think it may be better using Scopus and ISIwebofknowledge.

Specific comments

Abstract: abstracts should be clear and concise. Unless the paper is an essay, an opinion paper or whatever else justifying an original approach (and this is not the case), abstract should show: (i) the problem addressed (here, identification of input pairs (X, Y) corresponding to a given quantile of a target variable Z linked by $Z = g(X, Y)$, in the context of flooding evens triggering reservoir level... please, avoid discussions on return periods, etc; they are out of context here, as you focus on the target Z); (ii) the method proposed to solve the problem in (i) (here, the simulation procedure mentioned above); (iii) a brief summary of results. Please, avoid digressions and sentences such as 'a univariate approach in complex problems ignores the effect of other significant variables leading to different risk levels for each quantity of interest and resulting in an inaccurate estimate of the risk- often wrongly set equal to the risk of the hydrological event.' It does not make sense as your event is Z, is univariate, and your inference is based on univariate distributions (Fig. 8); X and Y (and the hydrograph shape) are only nuisance input variables, similar to covariates in GLMs; their return period does not matter at all in the present context.

Section 1.1: I have to say that my first two papers (ten years ago) start with such a type of introduction: 'Author 1 et al. (yyyy) used Arch1 copula for... Author 2 et al. (yyyy) used Arch2 copula for...'. My suggestion for young authors is to avoid such a kind of acritical and uninformative lists. Instead, use the introduction to better ex-

C5

plain the problem you address, why you do that, who already did something similar (cite only references dealing with the same problem, instead of every 'copula' paper on flood/drought/rainfall frequency analysis: the topic is not 'copula inference' but the relationship $Z_p = g(X_{Z_p}, Y_{Z_p})$), and why the previous efforts where not sufficient. Finally, briefly introduce what you intend to do in the remainder. This allows the reduction of section 1.1, which should be merged with section 1.2.

Section 1.2: the first part should be completely reworded and expanded. The last but one paragraph (P3L33-P4L5) should be moved in Section 2 as it describes the methodology. Once expanded, actually this the only methodological description required in the paper along with the details of a suitable simulation strategy (see Serinaldi, 2016).

P4L11: In a multivariate context, the hydrograph selection is not so straightforward. Selecting the hydrographs corresponding to annual maximum peaks means that we already focus on a dominant (driving) variable and the shape is conditioned on that. On the other hand, focusing on e.g. maximum volumes, hydrograph shapes are conditioned on volumes. If the shapes are different for e.g. fast floods with high peaks driven by convective storms and high-volume floods driven by frontal systems, it follows that we are already restricting the range of possible shapes S. So, we are in an incoherent situation: we want to develop a joint distribution of (X, Y), but the output of the joint model feeds dimensionless hydrographs whose shape is conditioned on e.g. X, i.e. $S|X$. This aspect, as many others, is usually overlooked in such a type of papers (...I faced this problem a few years ago in Serinaldi and Grimaldi (2011)). Instead of wasting time to describe copula inference and randomly applying conditional and joint distributions, hydrologists should focus on building probabilistic descriptions coherent with the physical problem at hand.

Section 2: the 90% of this section reports material already published in hundreds of papers with less and less accuracy as the time passes and people attempt to reword it in order to avoid possible plagiarism. Nevertheless, these procedures are well known... and misused.

C6

Section 2.1: reporting the formulas of AIC and BIC is useless if they are then used to rank models according to their absolute values; what matters to allow the assessment of results is to say how these indices are used.

Section 2.2: we do not need chi-plots and K-plots to know that data are correlated because this correlation might be a numerical effect (see Serinaldi and Kilsby (2013))

Section 2.3: Why introducing upper tail dependence coefficient when dealing with 52 pairs of data points? Furthermore, why introducing the most biased estimator among the available ones? I recognize that λ_U^{CFG} is the only estimator with closed form formula, and so the easiest to implement. However, after the performance assessment shown by Serinaldi et al. (2015), it should be clear to everyone that such an estimator yields UTD coefficients close to EV copulas independently of the actual model (samples from Frank copulas show tail dependence different from zero!!!) and sample size (look a bit better at Figures 1 and 2 in that paper). Anyway, the very illogical aspect is the application of such asymptotic indices to a few tens of observations.

Section 2.4: Why introducing such return periods if they are not related to the problem at hand in any way, and they are not used at all throughout the paper? I recognize that this a fashionable topic, but adding this list does not make this paper a good reference.

Sections 4.1 and 4.2: Without showing results for other models, it is not possible to assess the model selection. The superficial description highlights misuses and misinterpretation of results (e.g. questionable use of BIC scoring, and meaningless discussion on UTD); however, without suitable tables and diagrams it is not possible to say much more. In this respect, Figure 7 is uninformative, as these diagrams do not allow any quantitative assessment of marginal and copula fitting. By the way, what means 'The very heavy tail of the Generalised Extreme Value (GEV) distribution appears to overestimate the peak in the upper quantile (5%) and consequently, was not preferred'? We apply information criteria and three statistical tests, but a distribution is discarded because it does not look like we want! About the occurrence of apparently wild val-

ues, please have look at the discussion provided by Serinaldi (2013; Section 2, Fig. 4) Furthermore, 'It should be noted that the assumption that the data set is characterized by upper tail dependence is based on past research and is somewhat intuitive': No! it is not intuitive at all, because, in your selection procedure, only the peak is a block maximum, while the corresponding volume could not (and generally is not). UTD and EV copulas are (theoretically) justified for n-wise maxima (e.g. annual maxima or POT from multiple stations), and hold asymptotically. For finite sizes and finite blocks, convergence is usually very slow. So, nothing is intuitive; it is only a matter of theory and its understanding.

Section 4.3: The title is misleading as this section concerns the definition of the law $Z_p = g(X_{Z_p}, Y_{Z_p})$, where $p = 1/RP$ is simply the probability of univariate quantiles of interest. Moreover, as mentioned above, the present simulation strategy is not devised for the problem under study, which was, in turn, already addressed elsewhere. Comparisons of point estimates resulting from models fitted on 52 data points are also quite uninformative, if the main sources of uncertainty are not accounted for, at least.

Conclusions: 'Understanding the mechanisms of failure of our problem is crucial and can guide our decision regarding the return period definition which differentiates greatly the results.' I agree, but the same understanding is required for the statistical methods providing the correct probabilistic description of such mechanisms of failure. In 2006, fitting a joint distribution to a data set could be enough to justify publication. Nowadays, one can perform this exercise (a bit better than in this paper) in 10 minutes by using R or Matlab packages. Hydrologists should therefore make the effort to think a bit more deeply about the actual problem they face and then applying such tools where and when they provide effective improvements, considering the quality, quantity and nature of the available data.

References

Burnham K. P. and Anderson D. R. (2002) Model Selection and Multimodel Inference:

A Practical Information-Theoretic Approach, Springer

Burnham K. P. and Anderson D. R. (2004) Multimodel Inference: Understanding AIC and BIC in Model Selection, Sociological Methods Research 2004; 33; 261

Dung NV, Merz B, Bárdossy A, Apel H (2015) Handling uncertainty in bivariate quantile estimation – an application to flood hazard analysis in the Mekong Delta. J Hydrol 527:704–717

Salvadori, G., F. Durante, G.R. Tomasicchio, F. D'Alessandro (2015) Practical guidelines for the multivariate assessment of the structural risk in coastal and off-shore engineering, Coastal Engineering, 95, 2015, 77-83

Serinaldi F. (2013) An uncertain journey around the tails of multivariate hydrological distributions. Water Resources Research, DOI: 10.1002/wrcr.20531.

Serinaldi F. (2016) Can we tell more than we can know? The limits of bivariate drought analysis in the United States. Stochastic Environmental Research and Risk Assessment, 30(6), 1691-1704.

Serinaldi F, Grimaldi S. (2011) Synthetic design hydrographs based on distribution functions with finite support. Journal of Hydrologic Engineering, 16(5), 434-446.

Serinaldi F, Kilsby CG. (2013) The intrinsic dependence structure of peak, volume, duration and average intensity of hyetographs and hydrographs. Water Resources Research, 49(6), 3423-3442.

C9