Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-619-RC2, 2017

Hydrology and Earth System Sciences Discussions
Open Access

Interactive comment on "Comparative study of flood projections under the climate scenarios: links with sampling schemes, probability distribution models, and return level concepts" *by* Lingqi Li et al.

F. Serinaldi (Referee)

francesco.serinaldi@ncl.ac.uk

Received and published: 3 January 2017

General comments

In this study, the Authors perform quite a standard nonstationary frequency analysis. It can seem surprising to talk about "standard analysis" when dealing with relatively new "nonstationary" fashion, but the main point is that this paper, as a large part of those on this topic, is a simple application of a set of routines already implemented in R packages. Thus, most of the results are quite speculative, as they overlook the theoretical

Printer-friendly version

Discussion paper

concepts behind statistical tools and some significant papers explaining these issues.

Moreover, I regret to say that a very similar version of this paper was already declined by Advances in Water Resources in 2015. In that case, I suggested a major revision, but I see that the Authors did not account for most of my suggestions. Some of them are reported below once again with updates.

The main contribution of this study should be the use of the expected-number-of-events (ENE) method and the negative binomial distribution to replace the Poisson distribution under overdispersion conditions. Concerning the ENE method, it is only one of the possible approaches to define return periods and corresponding return levels. It yields the general equation $T = \frac{T}{m\sum_{t=1}^{T}(1-F_t(x_T))} = \frac{1}{mE[1-F_t(x_T)]}$, which reduces to the classical $T = \frac{1}{m(1-F(x_T))}$ if $F_t(x_T) \equiv F(x_T)$ for each t. In other words, the return period corresponds the expected value of the reciprocal of the exceedance probability of a fixed quantile x_T, which is constant if $F_t(x_T)$ is constant. Therefore, sentences such as "This advantage makes the method able to provide unique design value for reference even though the flood behaviors observe nonstationarity, which is beyond the capacity of traditional stationarity strategy" make little sense because return periods and return levels, being expected values taken over T (for ENE) or ∞ (for expected waiting time), are always unique values in both stationary and nonstationarity framewoks (a discussion is provided by Serinaldi (2015)). As far as the negative binomial distribution is concerned, it was already discussed in stationary flood frequency analysis, and compared with Poisson by Bhunya et al. (2013), while its introduction and theoretical justification in stationary and nonstationarity frameworks was presented by Eastoe and Tawn (2010). In particular, the latter highlighted how the overdispersion is not necessarily a consequence of nonstationarity. In fact, overdispersion can easily results from (hidden) persistence (see e.g. Serinaldi F, Kilsby CG. (2016a) and references therein) and/or mixed effects (random fluctuations of the rate of occurrence). Moreover, saying that "over-dispersion of observations" is a possible source of nonstionarity (P9L161-168) seems to me logically flawed because overdispersion is not a cause but an effect of non-Poissonian behaviour, which can have many different causes. Moreover, other models such as generalized Poisson can be used (Raschke, 2015). Again, nonstationarity is not a necessary condition. However, what really matters is that the distribution of the number of event under nonstationarity is neither Poisson nor negative binomial, but Poisson binomial (Tejada and den Dekker, 2011; Obeysekera and Salas, 2016). Thus, a more careful literature review should be performed before running (a bit blindly) computer codes/packages.

Most of the conclusions in the case study are quite speculative because the behaviour of the return periods under nonstationarity depends on many factors, such as the link functions, the relationships between distributions' parameters and covariates (linear or polynomial regression are surely convenient but also quite arbitrary and surely not physically based), as well as the nature of the distributions (fat tailed, heavy tailed, etc.). In this respect, conclusions are quite fair as they reflect the overall uncertainty of the empirical results, which is however exacerbated by lack of theoretical reasoning on the rationale and true nature of the methods used. By the way, it is worth noting that the models with parameters depending on covariates that exhibit stochastic fluctuations (such as rainfall and temperature) are not nonstationary but simply doubly stochastic. Nonstationary models require that the distribution (marginal and joint) change $with time$ according to some well-defined function holding true for whatever instant along the time axis. In this respect, trend analysis can only detect local changes in a very small interval (i.e., the period of record), and this explains why nonstationarity cannot be inferred from trend analysis but requires exogenous information, i.e. attribution based on physical reasoning rather than statistical correlation analysis.

Some additional specific remarks are provided below. I also refer the Authors to my report for the previous AWR version.

Specific comments

L141-143: see Serinaldi (2015) for a wider discussion.

L146: “other sampling methods. . .seem”

L189: TFPW does not perform any prewhitening and does not preserve the nominal significance level. This explains why the results reported in the literature for MK and TFPW MK are often close to each other (see Serinaldi and Kilsby (2016b) for an analytical and numerical proof)

L202-205: The interpretation of the partial MK test is not correct. Moreover, the interpretation reflects a widespread merging of Neyman-Pearson testing procedure and Fisher’s p-values, whose values cannot be interpreted as proxies of the strength of the relationship between target variables and covariates.

L268-275: AIC and BIC have different meaning and are relative measures. Thus, model selection should be based on Akaike weights and/or evidence ratios (see Burnham and Anderson (2002,2004)).

L289: If the process is not stationary, the empirical frequencies cannot be computed by the Gringorten formula. Classical qq plot does not make sense in a (true) nonstationary framework. Moreover, in GAMLSS, the residuals are not the normal quantile transform of the observed values (this holds only for stationary models) but the difference between the predictions (given the covariates) on the observations (for the same covariates). Furthermore, qq plots and coefficient of determination are not formal tests but diagnostic plots and measures of performance, respectively. In particular, no tests can be performed at the 5% significance level for coefficients of determination (unless ad hoc MC experiments are set up).

L494-496: This result is not so surprising because the number of exceedances decreases as the threshold increases, and therefore the clustering of extreme events is more evident given the short time series.

Section 4.3: GAMLSS are nothing but an advanced form of regression. Using the fitted model with covariates taking values beyond the fitting range is never a good idea because we do not know if the fitted relationship still holds true in that range.

Editing remarks

English should be revised and some typos fixed.

Sincerely,
Francesco Serinaldi

References

Bhunya P.K., R. Berndtsson, S. K. Jain, R. Kumar (2013) Flood analysis using negative binomial and Generalized Pareto models in partial duration series (PDS), Journal of Hydrology, 497, 121-132.

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

Eastoe, E. F., and J. A. Tawn (2010), Statistical models for overdispersion in the frequency of peaks over threshold data for a flow series, Water Resour. Res., 46, W02510, doi:10.1029/2009WR007757.

Obeysekera, J. and J. D. Salas (2016) Frequency of Recurrent Extremes under Non-stationarity, Journal of Hydrologic Engineering, 21(5), doi: 10.1061/(ASCE)HE.1943-5584.0001339

Raschke, M. (2015) Statistical detection and modeling of the over-dispersion of winter storm occurrence, Nat. Hazards Earth Syst. Sci., 15, 1757-1761.

Serinaldi F. Dismissing return periods! (2015). Stochastic Environmental Research and Risk Assessment, 29(4), 1179-1189

Serinaldi F, Kilsby CG. (2016a) Understanding persistence to avoid underestimation of collective flood risk. Water, 8(4), 152.

Serinaldi F, Kilsby CG. (2016b)The importance of prewhitening in change point analysis under persistence. Stochastic Environmental Research and Risk Assessment, 30(2), 763-777

Tejada, A., and den Dekker, A. J. (2011). The role of Poisson's binomial distribution in the analysis of TEM images. Ultramicroscopy, 111(11), 1553–1556.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016.