Reply to Referee #1

General comments

The paper deals with an issue of interest for the readers of HESS. In my opinion, a few critical issues must be fixed. Below, please find some indications: the objections should be read in a constructive way, since they may help the Authors improve the paper. Some useful bibliography is given at the end of this review.

Response:

We are very grateful for your kind evaluation as well as insightful comments on this paper. All your comments and suggestions have been addressed in revising the manuscript.

Specific comments

(1) Page(s) 3, Line(s) 48–49.

Author(s). . . . while the exceedance probability of a multivariate flood event could have multiple definitions (Salvadori et al., 2011; Vandenberghe et al., 2011).

Referee. Here, an additional reference is Salvadori and De Michele (2004), where the non-uniqueness of return periods in a multivariate setting was first pointed out.

Response:

Thanks for this comment. The corresponding reference has been added in this sentence.

Newly cited reference:

Salvadori, G., and De Michele, C.: Frequency analysis via copulas: theoretical aspects and applications to hydrological events, Water Resour. Res., 40, W12511, https://doi.org/10.1029/2004WR003133, 2004.

(2) Page(s) 3, Line(s) 54.

Referee. The citation to Salvadori and De Michele (2004) should be added here. Actually, there are (at least) four types of approaches, as recently outlined and discussed in Salvadori et al. (2016). The Authors are missing the Survival Kendall approach, first introduced in Salvadori et al. (2013): this latter avoids possible problems of divergence of the marginals on the boundaries of the domain.

Response:

Thank you for this kind suggestion. The Survival Kendall approach and structural approach have been added in the review on the current multivariate design methods.

Newly cited references:

Salvadori, G., Durante, F., and De Michele, C.: Multivariate return period calculation via survival functions, Water Resour. Res., 49, 2308–2311, https://doi.org/10.1002/wrcr.20204, 2013.

Salvadori, G., Durante, F., De Michele, C., Bernardi, M., and Petrella, L.: A multivariate Copula-based framework for dealing with Hazard Scenarios and Failure Probabilities, Water Resour. Res., 52(5), 3701–3721, https://doi.org/10.1002/2015WR017225, 2016.

Requena, A. I., Mediero, L., and Garrote, L.: A bivariate return period based on copulas for hydrologic dam design: Accounting for reservoir routing in risk estimation, Hydrol. Earth Syst. Sci., 17(8), 3023–3038, https://doi.org/10.5194/hess-17-3023-2013, 2013.

Salvadori, G., Durante, F., Tomasicchio, G. R., and D'Alessandro, F. : Practical guidelines for the multivariate assessment of the structural risk in coastal and off-shore engineering, Coastal Eng., 95, 77–83, https://doi.org/10.1016/j.coastaleng.2014.09.007, 2015.

(3) Page(s) 3, Line(s) 55.

Author(s). Due to changing climatic conditions, as well as some anthropogenic driving force. . .

Referee. Here, the Authors should cite the paper by Milly et al. (2008).

Response:

Thank you for this suggestion. The paper by Milly et al. (2008) has been listed as the reference.

Newly cited reference:

Milly, P., Betancourt, J., Falkenmark, M., Hirsch, R., Kundzewicz, Z., Lettenmaier, D., and Stouffer, R.: Climate change - Stationarity is dead: Whither water management? Science, 319(5863), 573–574, https://doi.org/10.1126/science.1151915, 2008.

(4) Page(s) 3, Line(s) 59–60.

Author(s). The multivariate flood distribution could exhibit more complex nonstationarity behaviors than the univariate distribution. . .

Referee. Here, the Authors should cite the multivariate distributional approach outlined in Vezzoli et al. (2017)—where Sklar's Theorem representation is used to test nonstationarity—as well as the guidelines for multivariate change-point detection illustrated in Salvadori et al. (2018).

Response:

Thanks for your recommendation of these two papers, which are very helpful to support the view of this paper.

Newly cited references:

Vezzoli, R., Salvadori, G., and De Michele, C.: A distributional multivariate approach for assessing performance of climate-hydrology models, Scientific Reports, 7(12071), https://doi.org/10.1038/s41598-017-12343-1, 2017.

Salvadori, G., Durante, F., Michele, C. D., and Bernardi, M.: Hazard assessment under multivariate distributional change-points: Guidelines and a flood case study. Water 10 (6), 751–765, https://doi.org/10.3390/w10060751, 2018.

(5) Page(s) 4, Line(s) 87–88.

Author(s). These design criteria assess the risk or reliability of hydraulic structures. . .

Referee. As a further approach involving Failure Probabilities, and design under nonstationarity and extreme value marginals, the Authors should cite the seminal paper by Salvadori et al. (2018).

Response:

We have added this citation in the revised manuscript as follows:

Salvadori et al. (2018) associated hydrological designs with both given life times and failure probabilities to calculate the design values under nonstationarity.

(6) Page(s) 5, Line(s) 95–96.

Author(s). . . . these design events are generally not equivalent because their joint probability density values (i.e. likelihood) are usually different. . .

Referee. This fact was first pointed out in Salvadori et al. (2011).

Response:

We have added this citation in the revised manuscript.

(7) Page(s) 6, Line(s) 106–108.

Author(s). . . . the cases for higher-dimensional hydrologic designs as well as under nonstationary conditions have not yet been covered in current studies.

Referee. In Salvadori et al. (2018) appropriate confidence intervals are computed in a nonstationary case via suitable Monte Carlo techniques (the case study is a bivariate one, but the procedures can be generalized to any dimension).

Response:

In the revision, this sentence have been rephrased as follows:

Very few studies covered the cases for higher-dimensional hydrologic designs as well as under nonstationary conditions (Salvadori et al., 2018).

(8) Page(s) 6–ff., Section(s) 2.1.

In my opinion, the Authors should check for possible nonstationarity by using the functions provided by the R package "npcp" (Kojadinovic, 2017), which works both for the marginals and the copula, and provides approximate p-values concerning distributional changes.

Response:

Thank you for this constructive suggestion. In the revision work, we have employed the methods provide by R package "npcp" to detect the change points of the multivariate flood series. The results indicate that neither marginal distributions nor dependence (copula) display significant change points. But current finding shows that both the margins and dependence (copula parameter) are nonstationary due to the human activities of both urbanization and reservoir regulation. This difference in nonstationarity judgment should be attributed to the opposite roles of these two driving factors, that urbanization generally enlarges the mean value of the flood series and weakens the dependence, while reservoirs usually decrease the mean value and strengthen the dependence. In other words, the nonstationarities induced by these two factors could be offset by each other. As the result, the nonstationarities of the multivariate flood series might fail to be captured by the methods in "npcp". In the revised manuscript we have discussed this finding in the final of section 4.2 as follows:

Additionally, the change-point detection method based on Cramér-von Mises statistic (Bücher et al., 2014) is employed to detect the possible nonstationarities in both marginal distributions and dependence of the multivariate flood series (Q_1, V_3, V_7, V_{15}). For specific implementation steps of change-point detection, readers are referred to Bücher et al. (2014) and Kojadinovic (2017). The results indicate that neither marginal distributions nor dependence display change points at the 0.05 significance level, while the previous analysis suggests nonstationary margins and dependence due to the joint effects of urbanization and reservoir regulation. This difference in nonstationarity judgment should be attributed to the opposite roles of urbanization and reservoir regulation on shifting the multivariate flood distribution, that the former generally enlarges the mean values of the flood series and weakens their dependence, while the latter decreases the mean values and strengthen the dependence. In other words, the nonstationarities induced by these two factors could been offset by each other. As the result, the nonstationarities of (Q_1, V_3, V_7, V_{15}) might fail to be captured by the statistical method based on Cramér-von Mises statistic. This finding highlights the significance of cause-effect analysis in judging the nonstationarity of hydrological series (Serinaldi and Kilsby, 2015; Xiong et al., 2015).

Newly cited references:

Kojadinovic, I.: npcp: Some Nonparametric CUSUM Tests for Change-Point Detection in Possibly Multivariate Observations, R Package Version 0.1-9, Vienna, Austria, 2017.

Bücher, A., Kojadinovic, I., Rohmer, T., Segers, J.: Detecting changes in cross-sectional dependence in multivariate time series, J. Multivar. Anal., 132, 111–128, http://dx.doi.org/10.1016/j.jmva.2014.07.012, 2014.

(9) Page(s) 7, Eq(s) 1.

The Authors should explain here the meaning of the undefined variable/parameter t in Eq. (1). In addition, the mathematical notation used is wrong. Usually, in Probability, upper-case letters (e.g., $Q1;t$) denote random variables, but here the arguments of the distribution functions F 's are real variables (i.e., lower-case letters). Please fix the notation throughout the manuscript.

Response:

Thanks for comment. The symbol 't' denotes time measured by years, and this point has been explained in the revised manuscript. The mistakes in mathematical notation are also fixed throughout the manuscript.

(10) Page(s) 8, Line(s) 169.

The covariates $(x_1; : : : ; x_k)$ should be specified here.

Response:

The covariates $(x_1; : : : ; x_k)$ indicate the factors leading to marginal nonstationarities, i.e. urban population and reservoir index. In the revision, the covariates have been specified.

(11) Page(s) 8, Line(s) 172–174.

Author(s). The higher-order distribution parameters such as scale and shape parameters are assumed to be kept constant to avoid possible larger uncertainty in parameter estimation, although they could also be nonstationary.

Referee. This assumption can make the work rather weak. In fact, here maxima are investigated, and the shape parameter plays a fundamental role in distributions like GEV or GPD, in order to rule the strength of the extremes: keeping these parameters constant could be unrealistic, and may yield strongly biased estimates. As shown in Salvadori et al. (2018), the uncertainties could really be large, but "constraining" the model by fixing the most relevant parameters is not a proper way to solve the problem. Please duly comment this point.

Response:

Great thanks for this constructive comment. It is really true that the tradeoff between reducing estimation bias and increasing model uncertainty is debatable in the application of nonstationary models. In the revision, this sentence has been rephrased as follows:

In reality, all parameters of the flood distribution could be nonstationary, but in this paper only the nonstationarity of the location parameter μ (referring to the first moment or mean of flood series) is considered. Given the limited length of the flood series used in this study, the higher-order distribution parameters such as scale and shape parameters are fixed to avoid the possible large uncertainty.

In addition, a comment for this point is also added in the final paragraph of discussion section:

In this study, considering the limited length of the flood series, the nonstationarities in both higher-order marginal distribution parameters and copula parameters in roots T2 and T3 are ignored to avoid possible larger uncertainty in parameter estimation. It is also important to note that keeping these parameters constant could be unrealistic and yield biased estimates.

(12) Page(s) 8, Line(s) 178–180.

Author(s). The goodness of fit (GOF) of the probability distributions is examined by using the Kolmogorov-Smirnov (KS) test (Frank and Massey, 1951).

Referee. This is a critical statistical point: how was the *p*-value computed? Just comparing the values of the KS test statistics with the ones reported in some table? This would be wrong. In fact, as is well known (e.g., simply read the help of Matlab), the KS test requires that the theoretical distribution be known a priori, it cannot be the fitted one. In the latter case, suitable (but simple) Monte Carlo techniques can be used to estimate an approximate p-value. Please fix the issue.

Response:

Great thanks for this comment. In the revision, we have employed the Monte Carlo technique to estimate the *p*-value of the KS test for probability distributions. The results indicate the chosen distributions perform well in fitting the flood series.

(13) Page(s) 9, Line(s) 191–192.

Author(s). In this study the dependence structure of $(Q_1; V_3; V_7; V_{15})$ is constructed by the pair copula method.

Referee. To the best of my knowledge, the current software for vine-copulas does not provide reliable p-values of GoF tests: this may represent a statistical weakness of this approach. A comment is required here.

Response:

Thanks for this comment. It is really true that the GoF test methods for vine-copulas are very limited. In the study, the Probability Integral Transform (PIT) method is employed to perform the GoF test for vine copulas, and this method should be a reliable way to examine the GoF of vine copulas (Aas et al., 2009). In the revised manuscript, we have added a comment for this point and supplemented more information about the PIT test as follows:

The goodness of fit (GoF) tests for vine copulas are very limited, and the Probability Integral Transform (PIT) test (Rosenblatt, 1952) seems to be one of the few reliable methods (Aas et al., 2009). Under the null hypothesis that the multivariate flood variables (Q_1, V_3, V_7, V_{15}) follow a given C-vine copula, the PIT converts the dependent flood variables into a new set of variables that are independent and uniformly distributed on $[0,1]^4$. Then the next step is to verify whether the resulting variables really are independent and uniform in [0,1]. This work can be finished by using chi-square test, and the significance level of the test is 0.05. For more details of the PIT test, reader are referred to Aas et al. (2009).

(14) Page(s) 11, Line(s) 227.

The covariates $(x_1; : : : ; x_l)$ should be specified here.

Response:

The covariates have been specified in the revised manuscript.

(15) Page(s) 12, Line(s) 252–253.

Author(s). . . . the exceedance probability p_t of $(q_1;\ v_3;\ v_7;\ v_{15})$. . .

Referee. What is the exceedance probability of a multivariate event? It should be properly defined here, or, at least, the Authors should put a reference to the next Section 2.2.2.

Response:

The exceedance probability p_t of $(q_1;\ v_3;\ v_7;\ v_{15})$ indicates the danger of flood event, and is defined as the occurrence probability of the event more dangerous than $(q_1;\ v_3;\ v_7;\ v_{15})$ in a specific hazard scenario. In the revision, we have added this definition.

(16) Page(s) 12, Line(s) 262.

Author(s). . . . OR case, AND case and Kendall case. . .

Referee. Here, the best reference is Salvadori et al. (2016), where these cases are, for the first time, properly and rigorously defined in terms of suitable Hazard Scenarios based on the notions of Copulas and Lower/Upper Sets.

Response:

In the revision, we have added this reference.

(17) Page(s) 13, Eq(s) 11.

The U's notation in Eq. (11) should be upper-case: the U's used as arguments of the copula must be random variables, otherwise it make no sense to calculate a probability. Please fix this point.

Response:

Great thanks. We have made the correction.

(18) Page(s) 17, Line(s) 373–374.

Author(s). The four candidate models of the time-varying margins are formulated as follows. . .

Referee. Please provide due comments/justifications about these choices.

Response:

Thank you for this constructive suggestion. In the revised manuscript, we have added some comments for these four candidate models.

As above, the first equation defines the most complex nonstationary model, where it is assumed that both urban population and reservoir index are the driving factors of marginal distributions; the second and third equations indicate that the marginal nonstationarity is only linked to urban population and reservoir index, respectively; and the fourth and final one is the simplest and stationary model.

(19) Page(s) 18, Line(s) 376–ff.

Author(s). In terms of the fitting quality assessed by AIC. . .

Referee. The fitting procedure should work in the reverse sense. Viz., first the admissible distributions (among the ones of interest) are identified (if any) via a GoF test—e.g., KS, typically via a Monte Carlo algorithm, not by using values from statistical tables, as already mentioned above. Then, a "best" distribution is chosen (e.g., via AIC) only among the admissible ones. It makes no sense to compute the AIC of a non-admissible distribution. Please fix this point.

Response:

We have adjusted the fitting procedure for marginal distributions in the revision according to this comment.

(20) Page(s) 18, Line(s) 380–382.

Author(s). According to the regression functions of the location parameters μ, the means of the flood series are generally positively related to the urban population *Pop* , while negatively related to the reservoir index *RI*.

Referee. This point is not clear. If a regression has been performed, the corresponding p-value should be shown, in order to decide whether the regression is statistically significant. In general, this work lacks of a solid statistical base. Please fix this point.

Response:

In the revision, the p-values of regression parameters have been calculated and displayed in Table 2. The results indicate that all parameters are statistically significant at the 0.05 level.

(21) Page(s) 19, Eq(s) 25.

The formulas given in Eq.s (25) look different from the one shown in Eq. (7): is this correct?

Response:

Thanks. It is a typing mistake. In the revision, we have fixed it.

(22) Page(s) 21, Line(s) 457–458.

Author(s). These differences among the OR, AND and Kendall exceedance probabilities induce the different design strategies.

Referee. As discussed in Serinaldi (2015); Salvadori et al. (2016), comparing results induced by the usage of different Hazard Scenarios could be misleading, if not meaningless. The Hazard Scenario should be chosen a priori, not as a result of the consequences that the choice of a given scenario might entail. This looks like a methodological flaw. Please comment this point.

Response:

Great thanks for this insightful comment. It is really true that the choice of design strategy is generally priori, and depends on the specific situations of design requirements and mechanisms of failure for hydraulic structures. In the revision, we have added the following comment:

These differences among the OR, AND and Kendall exceedance probabilities indicate the different design strategies. In engineering practice, the choice of design strategy is generally priori, and depends on the specific situations of design requirements and mechanisms of failure for hydraulic structures (Serinaldi, 2015; Salvadori et al., 2016).

Newly cited reference:

Serinaldi, F.: Dismissing return periods!, Stoch. Eev. Res. Risk. A., 29 (4), 1179–1189, http://dx.doi.org/10.1007/s00477-014-0916-1, 2015.

(23) Page(s) 23, Line(s) 511.

Author(s). In this paper, we present the methods addressing the multivariate hydrologic design. . .

Referee. The claim “the methods” is incorrect, and too strong: you present “some possible methods”. Please fix the sentence.

Response:

In the revision, we have rephrased this sentence according your suggestion.