Responses to Referee #3

Dear Referee #3,

We are very grateful to you for the professional and constructive comments on our manuscript entitled “Comparative study of flood projections under the climate scenarios: links with sampling schemes, probability distribution models, and return level concepts” (Number: hess-2016-619). These comments are all valuable and very helpful not only for improving this paper but also beneficial for our research. We have carefully followed these comments and accordingly made the relevant revisions. The point-by-point responses to the comments and the corresponding correction to the manuscript are explained as follows:

General comment 1:

In the submitted paper authors compare different methods that can be selected to perform the flood frequency analysis (FFA). Both stationary annual maximum (AM) method and peaks over threshold (POT) method are tested using the daily discharge data from the Huaxian station (Weihe basin) in China. Further, a nonstationary methodology is also applied and compared with the stationary approach. Several interesting and important aspects of the flood frequency analysis approach are analyzed and discussed: application of the expected-number-of-events (ENE) method, selection of different distribution functions in the AM method, evaluation of the suitability of the Poisson and negative binomial (NB) distributions to model the annual number of exceedances above the threshold, comparison between the stationary and nonstationary approaches using both AM and POT methods, sensitivity analysis regarding the influence of the precipitation and temperature on the nonstationary flood frequency analysis results. The paper is relatively well written and the presented topic is interesting for the hydrological society due to the importance of the flood frequency approach for the design of different hydro-technical structures. The paper is in the scope of the journal. However, I would suggest that authors try to put more focus on next points related to the practical applications of the FFA, because the presented paper does not develop a new theory but compares different aspects of the FFA approach:

Response:

We thank the referee very much for the positive evaluation concerning our manuscript. The following comments have been addressed point by point as below.

General comment 2:

2.1 The authors have performed detailed analysis and comparison of different methods that can be used to carry out the flood frequency analysis. Is it possible to point out which method should be used by practitioners to determine the design flood (taking into account the larger sensitivity of the nonstationary AM method compared to the POT, more complicated POT analysis compared to the AM method and other conclusions stressed in this paper)

Response:

Great thanks to the referee for this constructive comment and suggestion. We would like to address this comment from three aspects in accordance with the title of this paper:

(1) *Concerning the sampling schemes*. As we stated in the original manuscript, we suggest that the POT sampling scheme should receive more attention (or at least as much as the AM) in flood return level (RL) analysis, especially when research is conducted on nonstationary condition. First, the POT can describe not only the magnitude of extreme events but also the frequency of the events. Changes in the frequency of extreme events have increasingly attracted public attention (Villarini et al., 2012), and the POT is an effective tool to study this. Second, based on the results of this study, RL derived with POT shows a lower sensitivity to climate change than that with AM. This implies that the established nonstationary POT model, when extrapolated to future application, may have less uncertainty introduced by the climate scenarios than the AM. The plausible reason has been given in Specific comment (22). Third, the sampling method of AM, which stipulates one event per year, is simple to perform, but may lose the significance of real flood and thereby leads to uncertainty of RL. Although the POT sampling method has the uncertainty problem brought by the complicated sampling criteria (e.g., threshold selection), this problem may be relieved to some extent by selecting a range of acceptable thresholds and then comparing the RL results of them. Finally, there is often the case of short data records that renders the inapplicability of AM for both statistical analysis and operational application (due to the insufficient representativeness of population). However, the POT sampling scheme can offer longer sample size and more than one time series. It is found from Figure 5 in the original manuscript that on both stationary and nonstationary conditions, RL derived with POT series yield narrower confidence limits than the AM series.

(2) *Concerning the probability distribution models*. According to the specific study on description of POT arrival rate, the negative binomial (NB) model has been demonstrated to be superior to the Poisson model when POT arrival rate shows a behavior of

over-dispersion. This result coincides with the theoretical basis of the NB distribution that is characterized by variance-to-mean ratio greater than unity. Besides, the NB distribution includes the Poisson distribution as a special case (Anscombe, 1950). When it is doubtful whether the observed POT arrival rate really comes from Poisson distribution, the NB distribution can be preferable for practical application. For both modeling the nonstationarity and making prediction, covariates with physical meaning have been proven to be more reasonable than those without a cause-effect relationship with the nonstationarity of extreme events (e.g., time) (López and Francés, 2013; Du et al., 2015; Prosdocimi et al. 2015). In the real-world application, practitioners have to carefully select the appropriate physical covariates that should be closely related to the flood variability in the actual study region.

(3) *Concerning the return level concept.* In hydrologic studies, flood return level analysis has been conducted with either the assumption of traditional stationarity strategy or the methods adapted to nonstationarity. From the perspective of the practitioners, it is more advisable to employ the method developed with nonstationarity if flood variability really happens over a certain long period, together with a sensitivity/uncertainty analysis as the current work did to enhance the understanding of the possible impact of nonstationarity on return level estimation. But above all, whenever nonstationarity has been considered for flood extrapolation, traditional stationarity strategy must be always retained as the baseline inference. Please see more explanations in the response to General comment 2.2. The ENE method studied in this paper is one of the pertinent ways to extrapolate design floods for both AM and POT series on either stationary or nonstationary condition. Further studies can focus on more other (e.g., expected waiting time) or new methods, especially for POT series, to carry out broader studies on future flood projection.

According to the referee’s suggestion, in the revised manuscript, we have added a brief discussion on the practical application of nonstationary flood return level analysis.

<u>Reference</u>

Anscombe, F.J.: Sampling theory of the negative binomial and logarithmic series distributions, Biometrika, 37, 358-382, doi:10.2307/2332388, 1950.

Du, T., Xiong, L., Xu, C.-Y., Gippel, C.J., Guo, S., and Liu, P.: Return period and risk analysis of nonstationary low-flow series under climate change, J. Hydrol., 527, 234-250, doi:10.1016/j.jhydrol.2015.04.041, 2015.

López, J., and Francés, F.: Non-stationary flood frequency analysis in continental Spanish rivers, using climate and reservoir indices as external covariates, Hydrol. Earth Syst. Sci., 17, 3189-3203, doi:10.5194/hess-17-3189-2013, 2013.

Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J., Dettinger, M.D., and Krysanova, V.: On Critiques of "Stationarity is Dead: Whither Water Management?", Water Resour. Res., 51, 7785-7789, doi:10.1002/2015WR017408, 2015.

Prosdocimi, I., Kjeldsen, T.R., and Miller, J.D.: Detection and attribution of urbanization effect on flood extremes using nonstationary flood frequency models, Water Resour. Res., 51, 4244-4262, doi:10.1002/2015WR017065, 2015.

Villarini, G., Smith, J.A., Serinaldi, F., Ntelekos, A.A., and Schwarz, U.: Analyses of extreme flooding in Austria over the period 1951-2006, Int. J. Climatol., 32, 1178-1192, doi:10.1002/joc.2331, 2012.

2.2 Should the standard procedures to perform the FFA in China be modified after the results of this study?

Response:

Thank you very much for this constructive and valuable comment. For practical application, the more realistic suggestion is to take into account nonstationarity when flood observations really exhibit significant changes at a certain long time period, carefully analyze the physical mechanism behind the detected changes in floods to select reasonable explanatory (physical) covariates, and extrapolate design floods to the near future together with a sensitivity/uncertainty analysis to enhance the understanding of possible impact of physical covariates on nonstationary flood return level analysis. In addition, as we stated in the introduction and actually did in this study, the standard procedures of stationarity strategy to perform the flood return level analysis should always be retained as the baseline for references. This solution can anyway be more informative than the insistence of single use of the traditional standard to perform flood frequency analysis while variability in flood really happens.

It is worth mentioning that the above suggestion should be practicably acceptable and very useful only over the decades-long design horizon of engineering (e.g., some projects are always related to a specific multi-year design plan) but not the very distant future. On the one hand, the presence of flood variability caused by the changing climate cannot readily be neglected as the variability in flood series (that leads to not identically distributed sample) may cause big uncertainty for both modeling and estimating floods when flood return level analysis is still performed by traditional stationarity strategy (Milly et al. 2015). On the other hand, due to incorporation of nonstationarity, model complexity certainly will increase, and thereby could also induce other sources of uncertainty. Therefore, the extrapolation of design floods should be confined in a specific scope with the underlying assumption that the pre-determined nonstationary distribution model for flood extrapolation will be applicable for a near future.

Reference:

Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J., Dettinger, M.D., and Krysanova, V.: On Critiques of "Stationarity is Dead: Whither Water Management?", Water Resour. Res., 51, 7785-7789, doi:10.1002/2015WR017408, 2015.

2.3 What is the trade-off (if any) between the model complexity and uncertainty in the flood frequency analysis results?

Response:

Thanks for this constructive comment. In this study, there are two main aspects related to the trade-off between the model complexity and uncertainty. The first aspect is to perform flood return level analysis under both stationarity and nonstationarity as we have explained in the response to General comments 2.1 and 2.2. The second aspect is to assume a linear dependence on physical covariates for modeling distribution parameter. This assumption should be regarded as a desirable tradeoff between the high-order modeling of covariates (to reduce bias of fit) and the suspicion of over-fitting (to reduce uncertainty) given the previous referential experience (e.g., Villarini et al., 2009; Salas and Obeysekera, 2014; Xiong et al., 2015a, b). With the progress in understanding and modeling of nonstationarity, newly well-defined mathematical treatment that can better describe the physical relationship with the flood variability should be aspired in future study.

Reference

Salas, J.D., and Obeysekera, J.: Revisiting the concepts of return period and risk for nonstationary hydrologic extreme events, J. Hydrol. Eng., 19, 554-568, doi:10.1061/(ASCE)HE.1943-5584.0000820, 2014.

Villarini, G., Smith, J.A., Serinaldi, F., Bales, J., Bates, P.D., and Krajewski, W.F.: Flood frequency analysis for nonstationary annual peak records in an urban drainage basin, Adv. Water Resour, 32, 1255-1266, doi:10.1016/j.advwatres.2009.05.003, 2009.

Xiong, L., Du, T., Xu, C.-Y., Guo, S., Jiang, C., and Gippel, C.J.: Non-stationary annual maximum flood frequency analysis using the norming constants method to consider non-stationarity in the annual daily flow series, Water Resour. Manag., 29, 3615-3633, doi:10.1007/s11269-015-1019-6, 2015a.

Xiong, L., Jiang, C., Xu, C.-Y., Yu, K.-X., and Guo, S.: A framework of change-point detection for multivariate hydrological series, Water Resour. Res., 51, 8198-8217, doi:10.1002/2015WR017677, 2015b.

General comment 3:

Related to the previous point, different methods yielded diverse FFA results. For example, the 50-year flood was estimated to be between approximately 4000 and 8000 m^3/s with the consideration of the confidence intervals (Fig. 5). Can the authors suggest some guidance for selection of the most appropriate method to carry of the FFA?

Response:

We thank the referee very much for this comment. The confidence intervals in the estimations of return levels reflect a certain level of uncertainty involved in the flood inference. Overall, in this study, POT series have much narrower confidence intervals than the AM series, implying that it may provide more practically acceptable design flood values than the AM series when nonstationarity has been taken into account in flood return level analysis. Please refer to the response to General comment 2.1 for more explanations, which is very similar to this comment.

General comment 4:

Looking at the results of the nonstationary approach (AM method) shown in Fig. 5 it seems that the return level increases to about 30-year return period and then it is almost constant for larger return periods? Does this means that the 50-year flood is the same as the e.g. 200-year flood? Please explain.

Response:

We thank the referee's valuable comment. As questioned by the referee, the result of T-year ($30 < T < 90$) return level (RL) derived with AM series on nonstationary condition in Fig. 5, indeed does not increase much as T becomes larger. However, T-year RL, e.g., 200-year RL ($x_{T=200}^{non\_s}$), should not be taken for granted as a natural extension of this result. We give a brief interpretation with an example of the nonstationary GEV model (with parameters as functions of climatic covariates) and apply this model to derive 50- and 80-year RLs (denote as the values $x_{T=50}^{non\_s}$ and $x_{T=80}^{non\_s}$ hereinafter) by the ENE method.

Denote $F(\cdot|\theta_t)$ being the GEV distribution function with time-varying parameters θ_t. According to the ENE method, $x_{T=50}^{non\_s}$ and $x_{T=80}^{non\_s}$ should satisfy the inference formula as

$\sum_{t=t_0+1}^{t_0+50}[1-F(x_{T=50}^{non\_s}|\theta_t)]=1$ and $\sum_{t=t_0+1}^{t_0+80}[1-F(x_{T=80}^{non\_s}|\theta_t)]=1$, respectively. The difference in cumulative probability between $\sum_{t=t_0+1}^{t_0+50}[1-F(x_{T=50}^{non\_s}|\theta_t)]=1$ and $\sum_{t=t_0+1}^{t_0+80}[1-F(x_{T=50}^{non\_s}|\theta_t)]$ over the level $x_{T=50}^{non\_s}$ denotes as $\sum_{t=t_0+51}^{t_0+80}[1-F(x_{T=50}^{non\_s}|\theta_t)]$. If $\sum_{t=t_0+51}^{t_0+80}[1-F(x_{T=50}^{non\_s}|\theta_t)]$ is approximate to zero, we have $\sum_{t=t_0+1}^{t_0+80}[1-F(x_{T=50}^{non\_s}|\theta_t)]\approx\sum_{t=t_0+1}^{t_0+50}[1-F(x_{T=50}^{non\_s}|\theta_t)]=\sum_{t=t_0+1}^{t_0+80}[1-F(x_{T=80}^{non\_s}|\theta_t)]=1$ and thus $x_{T=50}^{non\_s}\approx x_{T=80}^{non\_s}$, otherwise, we have $\sum_{t=t_0+1}^{t_0+80}[1-F(x_{T=50}^{non\_s}|\theta_t)]>\sum_{t=t_0+1}^{t_0+50}[1-F(x_{T=50}^{non\_s}|\theta_t)]=\sum_{t=t_0+1}^{t_0+80}[1-F(x_{T=80}^{non\_s}|\theta_t)]=1$ and thus $x_{T=80}^{non\_s}>x_{T=50}^{non\_s}$. In accordance with these illustrations, the result in this study has indeed found that $\sum_{t=t_0+51}^{t_0+80}[1-F(x_{T=50}^{non\_s}|\theta_t)]$ is approximately zero, and $x_{T=50}^{non\_s}\approx x_{T=80}^{non\_s}$ (the differences are within $80m^3/s$).

The aforementioned result is caused by the negative trend found in the AM series and the sensitivity to climate change. Therefore, in this study, under the impact of future changing climate, T-year RL (T > 90 years) derived on nonstationary condition should be determined strictly by the ENE inference at any time. Whether $x_{T=200}^{non\_s}$ can be same as $x_{T=50}^{non\_s}$ remains to be tested. However, the priori information of climatic scenario required to solve the equation $\sum_{t=t_0+1}^{t_0+200}[1-F(x_{T=200}^{non\_s}|\theta_t)]=1$ is insufficient, $x_{T=200}^{non\_s}$ could be thus inconclusive. Additionally, according to the response to General comment 2.2, future projection of $x_{T=200}^{non\_s}$ might have too big uncertainties, and should be an excessive extrapolation for future.

General comment 5:

It would be interesting to make a comparison of the nonstationary approach where the model parameters change with time (e.g., Obeysekera and Salas, 2014; Salas and Obeysekera, 2014; Sraj et al., 2016; Vogel et al., 2011) and not only with P and T. The English is understandable, but it could benefit from some improvements, therefore I recommend editing for English language.

Response:

Thanks a lot for this good comment and suggestion. We agree with the referee that it can be more informative to compare the results of nonstationary flood return level analysis when taking time and climatic factors as covariates, respectively. Admittedly, using time rather than climatic covariates has the advantage of simplicity in model structure and ease of extrapolation to future (i.e., the priori information of time as covariate is already known). It is in fact a more popular method and has long been used by hydrologists worldwide. However, as we illustrated in the response to General comment 2.1, many recent studies concerning the nonstationary analysis of extreme events have concluded that physical covariates-dependent models should be practically more reasonable than those purely employing time as covariate in both description and extrapolation of nonstationarity for application. Actually, the superiority of physical covariate has also been found in this study, for which we would like to show an example here (only the AIC and BIC values for brief) in Table 1 as below.

Table 1 The optimal GEV models using time and physical covariates as covariates, respectively.

GEV model	Estimated parametric functions	AIC	BIC
Using time *t*	$\mu_t = 1814.942 - 9.857t$ $\ln(\sigma) = 6.644$ $\xi = 0.313$	854.8	862.5
Using physical covariates T_{mean}, P_{total}	$\mu_t = 1789.594 + 3.818P_{total} - 215.657T_{mean}$ $\ln(\sigma_t) = 9.736 - 0.336T_{mean}$ $\xi = 0.108$	832.3	843.8

Inclusion of the results with time as covariate would be lengthy for this paper and seems not very consistent with our study objectives. For example, we analyzed how climate change could influence flood projections by the sensitivity analysis and assuming an increment of climatic covariates. These analyses would be meaningless when using time as covariate which is undoubtedly a monotonically increasing variable.

Therefore, in the revised manuscript, we have directed the interested readers to the studies as mentioned by the referee, in which, time has been used as covariate, instead of a detailed illustration in this study. Following the referee's suggestion, the English writing of the revised manuscript will be improved by a language editing service.

Specific comments

(1) Page 13, line 257: I would suggest adding a reference for the GAMLSS package.

Response:

Thanks for this suggestion. We have added the reference (Rigby and Stasinopoulos, 2005) in the revised manuscript.

<u>Reference</u>

Rigby, R.A., and Stasinopoulos, D.M.: Generalized Additive Models for Location, Scale and Shape, J. Roy. Stat. Soc., 54, 507-554, doi:10.1111/j.1467-9876.2005.00510.x, 2005.

(2) Page 16, lines 313-314: I would suggest rephrasing this sentence.

Response:

Thanks, we have rephrased the sentence in the revised manuscript.

(3) Page 16, line 321: What is "dramatic" or "pointless" for the authors? This can be very subjective, thus I would suggest avoiding such statements.

Response:

Thanks a lot for pointing out the improper and subjective statement. In the original paper, the word ("dramatic" or "pointless") is used to imply that *T*-year return level (RL) would have a large range of estimation values if stationarity strategy continues to be employed on nonstationary condition. For example, the 50-year RL estimated by the nonstationary LNO3 model for AM series (corresponding to the exceedance probability of 0.02) varies from year to year with large discrepancy (ranging roughly from 3000 to 9500) as shown by the red line in Figure 1 below (the calculation is based on the data and model used in the original manuscript). This result highlights the improper application of stationarity strategy under nonstationary condition. Following the referee's suggestion, we have rephrased the sentences in the revision of manuscript to correct the improper and subjective statement.

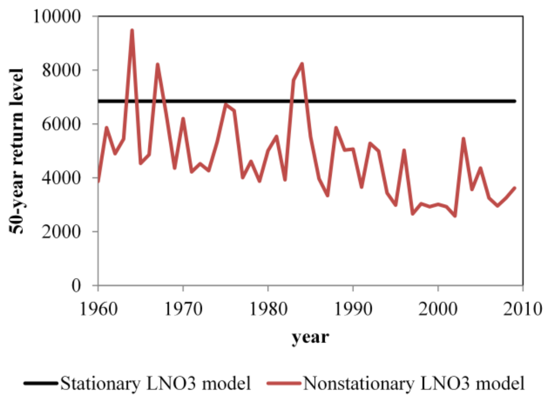

Figure 1 50-year RLs (corresponding to the exceedance probability of 0.02) estimated based on the stationary LNO3 (black line) and nonstationary LNO3 models (red line), respectively.

(4) Page 19, line 387: Which Sensitivity package (a reference should be added)?

Response:

Thanks for this comment and suggestion. The reference (Pujol et al., 2017) has been added in the revision.

Reference

Pujol, G. et al.: Package "sensitivity", Version 1.14.0, 2017, (retrieved from https://cran.r-project.org).

(5) Page 21, line 408: Replace “134,800” with “134 800” (and also in some other parts of the manuscript).

Response:

Thanks, it has been replaced by “134 800” together with the alterations for other parts in the revised version of manuscript (e.g., 106 498).

(6) Page 21, line 414: Upstream and not downstream?

Response:

Thanks for pointing out this mistake. It has been corrected in the revision.

(7) Page 22, line 422: Replace “Thiessen polygon” with “Thiessen polygons”.

Response:

Thanks, the words “Thiessen polygon” has been replaced by “Thiessen polygons”.

(8) Page 24, lines 487-488: Any particular physical reason for this negative trend? It would be interesting to see the discharge data used in study.

Response:

Thanks for this comment. The negative trend detected in AM flood series is mainly dominated by the changes in climate as has been specifically studied in the previous literatures (Xiong et al., 2014, 2015; Jiang et al., 2015). In the revised manuscript, we have modified the relevant text to explicitly show the physical reason for this negative trend. Following the reviewer’s suggestion, we have added the figures to visually display the variations of the discharge series used in this study.

<u>Reference</u>

Jiang, C., Xiong, L., Wang, D., Liu, P., Guo, S., and Xu, C.-Y.: Separating the impacts of climate change and human activities on runoff using the Budyko-type equations with time-varying parameters, J. Hydrol., 522, 326-338, doi:10.1016/j.jhydrol.2014.12.060, 2015.

Xiong, L., Jiang, C., and Du, T.: Statistical attribution analysis of the nonstationarity of the annual runoff series of the Weihe River, Water Sci. Technol., 70, 939-946, doi:10.2166/wst.2014.322, 2014.

Xiong, L., Du, T., Xu, C.-Y., Guo, S., Jiang, C., and Gippel, C.J.: Non-stationary annual maximum flood frequency analysis using the norming constants method to consider non-stationarity in the annual daily flow series, Water Resour. Manag., 29, 3615-3633, doi:10.1007/s11269-015-1019-6, 2015.

(9) Page 24, line 492: Again, what does "dramatically" means?

Response:

Thanks for pointing out the ambiguous expression. In the revised manuscript, "dramatically" has been replaced by "significantly".

(10) Page 25, lines 499-503: Is this the case for all 22 analyzed stations?

Response:

Thanks for this comments. The trend test is in fact conducted on the five series of climatic covariates, i.e., annual total precipitation, annual maximum precipitation on consecutive one, three, and seven days, and annual mean air temperature. The at-site daily total precipitation and daily mean temperature series of the 22 stations have been processed into the areal average series by the method of Thiessen polygons. Then the areal average series provide the series for the selected five climatic covariates. The relevant details have been shown in Lines 409-423 in the original manuscript.

(11) Page 26, lines 526-529: I would suggest rephrasing this sentence.

Response:

Thanks for this suggestion. We have reworded this sentence in the revised version.

(12) Page 26, line 537: "much lower" this is subjective; I would suggest using the % to show the difference.

Response:

Thanks for this comment. In the revised version, we have revised this vague expression as suggested.

(13) Page 29, lines 569-570: What is reason for this large difference and what does this mean from the perspective of the practitioners?

Response:

Thanks a lot for this valuable comment. In the question as raised by the referee, the return level is derived with the GEV model under stationarity (x_T^s) and nonstationarity ($x_T^{non\_s}$) for the future period. The difference between x_T^s and $x_T^{non\_s}$ should be attributed to the significantly decreasing trend in the AM series, the significant increase in future temperature and the higher sensitivity of $x_T^{non\_s}$ to temperature. Assuming that the return level derived under nonstationarity remains invariant as x_T^s and substituting x_T^s into the nonstationary GEV distribution function $F(\cdot|\boldsymbol{\theta}_t)$, the exceedance probability of each year $1-F(x_T^s|\boldsymbol{\theta}_t)$ will become lower and lower as t increases, the obtained cumulative exceedance probability of T years satisfies $\sum_{t=t_0+1}^{t_0+T}[1-F(x_T^s|\boldsymbol{\theta}_t)]<1$, which implies that the return period is no longer T years under nonstationarity. To satisfy the ENE inference formula $\sum_{t=t_0+1}^{t_0+T}[1-F(x_T|\boldsymbol{\theta}_t)]=1$, we can thus obtain $x_T^{non-s}<x_T^s$. The result indicates the uncertainty in x_T^s as introduced by the climate change when flood return level analysis is performed without consideration of the impact of nonstationarity while it has existed in flood series.

(14) Page 29, lines 583-585: These are relatively large differences. Which POT threshold is suggested by the authors and why?

Response:

Thanks for this valuable comment. The large differences among the results of POT2, POT3, and POT4 mainly originate from the different threshold values *u*. In the newly revised manuscript, we have added the test of mean excess plot for the selected threshold of POT2, POT3, and POT4. The results indicate that the POT4 threshold almost approaches (but is within) the lower bound of domain where the mean excess should be an approximately linear function of *u* for a valid choice of the generalized Pareto (GP) distribution. Therefore, POT4 series might not be the best appropriate choice in this study. There is no evidence of invalidation of threshold test for POT2 and POT3, both of which should be thus taken as more reasonable choices for practical application. From the perspective of engineering security, POT2 with higher threshold value might be preferred as it ensures a high level of flood magnitudes. From the perspective of statistics, POT3, as has a larger amount of magnitude values, could be considered to ensure a relatively lower uncertainty caused by sample size. The conservative and also often adopted solution is to give a fair comparison of both acceptable choices of POT2 and POT3 and then make a decision related to the actual requirements of a specific engineering project.

(15) Page 30, line 618: Dot is missing at the end of the sentence.

Response:

Thanks for pointing out this omit. It has been added.

(16) Page 31, line 642: Replace “shows” with “show”.

Response:

Thanks. It has been corrected.

(17) Page 32, line 652: "there is not much difference" looking at Fig. 5 I would say that differences are relatively large for some cases?

Response:

Thanks for pointing out this imprecise expression. It has been rephrased in the revised manuscript. Here the statement "there is not much difference" should have been specified to the results of some return periods *T* (e.g., 30 years) when using the LNO3 and LP3 models for flood estimation.

(18) Page 33, line 672: Replace "if we allowing" with "if we are allowing".

Response:

Thanks, the syntax error has been corrected.

(19) Page 33, line 679: Replace "requires" with "require".

Response:

Thanks, it has been revised.

(20) Page 36, line 748: Reason for this difference?

Response:

Thanks for this good question. Please refer to the response to Specific comment (13) that is very similar to this comment.

(21) Page 36, lines 748-751: What does this conclusion means for the practical application of the FFA?

Response:

Thanks for this valuable comment. This conclusion is meant to remind practitioners to pay attention to the impact of not only significant changes in flood observation series but also the physical causes (e.g., climate change) behind the changes when conducting a nonstationary flood return level analysis for future projection. In this study, the climate change, as has been demonstrated previously to cause the significantly decreasing trend in AM floods (Xiong et al., 2014), has exerted an important influence on determining the distribution model and accordingly on return level estimations. Due to the climatic impact, the significantly negative trend in AM series (though has been proven) does not mean a sure result of $x_T^{non\_s} < x_T^s$. x_T^s

and x_T^{non-s} are T-year return levels derived with stationary and nonstationary models of the same type of distribution, respectively.

To further illustrate the possible different results of x_T^s and x_T^{non-s} derived with AM series under the changing climate, we here show some sketch diagrams in Figure 2 below to give the details in the ENE formula. $\Pr$ represents the exceedance probability of a single type of distribution. Figure 2a shows the stationary case x_T^s that satisfies $\Pr(X > x_T^s) = 1/T$. Figure 2b shows the most common result of $x_T^{non-s} < x_T^s$, in which the downtrend in AM series only leads to a slightly left-ward shift in the probability density curve (pdc). This result is easily taken for granted before conducting an actual calculation. However, overlooking the possible changes in the shape of pdc (especially the behavior of upper tail) may lead to the wrong conjecture as exemplified in Figure 2c. Figure 2c gives the special case for $x_T^{non-s} = x_T^s$ where the shape of pdc has changed from year to year, the final accumulation of exceedance probability $\sum_{t=t_0+1}^{t_0+T} \Pr(X > x_T^s) = \sum_{t=t_0+1}^{t_0+T} \Pr(X > x_T^{non-s}) = 1$ results in $x_T^{non-s} = x_T^s$.

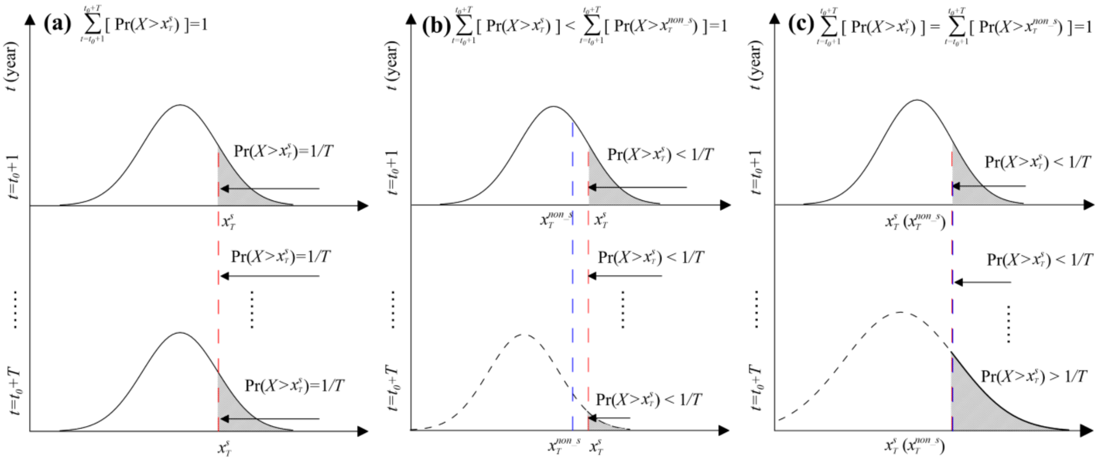

Figure2 Sketch diagrams for return level inference with the ENE method in the case of stationarity (a), nonstationarity with a left-ward shift in pdc (b), nonstationarity with both a left-ward shift and a varying shape in pdc.

Reference

Du, T., L, Xiong, C.-Y. Xu, C.J. Gippel, S. Guo, and P. Liu (2015), Return period and risk analysis of nonstationary low-flow series under climate change, J. Hydrol., 527, 234–250, doi:10.1016/j.jhydrol.2015.04.041.

Jiang, C., Xiong, L., Wang, D., Liu, P., Guo, S., and Xu, C.-Y.: Separating the impacts of climate change and human activities on runoff using the Budyko-type equations with time-varying parameters, J. Hydrol., 522, 326-338, doi:10.1016/j.jhydrol.2014.12.060, 2015.

Xiong, L., Jiang, C., and Du, T.: Statistical attribution analysis of the nonstationarity of the annual runoff series of the Weihe River, Water Sci. Technol., 70, 939-946, doi:10.2166/wst.2014.322, 2014.

(22) Page 37, lines 760-763: This is very important conclusion but is it true only for this case study or there is a theoretical background for it?

Response:

Thanks a lot for this valuable comment. In this paper, the conclusion "that the return level (RL) derived with AM is more sensitivity to climate change than that with the POT" is obtained in the nonstationary context for the study basin of the Weihe. Due to the diversity of the behaviors of nonstationarity and climate change found at different basins, the conclusion obtained in this basin may not be valid for other basins. However, we would like to give two plausible reasons based on this study, which may be helpful for other practitioners. One reason is that the POT series has been demonstrated to be best described by stationary magnitude and nonstationary frequency, while the AM series (only magnitude) has been optimally fitted by the nonstationary model. This indicates that the AM series, which perhaps include small magnitude samples, may overestimate the impact of climate change on variabilities in floods in the real flood. The other reason may be that the AM sample size is too limited in comparison to that of the POT series, thereby leading to the relatively larger uncertainty.

(23) Page 39, lines 807-810: But this "relatively complicated sampling criteria" still exists and if we compare the POT sampling methodology with the nonstationary approach used in this study I would say that it is even more complicated (than the stationary approach) and it requires additional knowledge?

Response:

Thank the referee very much for this valuable comment and good suggestion. In the revised manuscript, we have revised the sentences to clarify why "we suggest that POT series should be warranted more attention in nonstationary flood frequency analysis". It is known that many available studies have been still limited to annual maximum extremes such as AM floods, while POT series have so far not received as much emphasis as the AM series, whether on stationary or nonstationary condition. The most likely reason is the complicated sampling criteria of POT series. We should admit that the flood return level analysis with POT series on nonstationary condition does become more complicated than that under the stationary condition. However, the topic on investigation of changes in POT floods is anyway unavoidable as the evidence of nonstationarity has appeared not only in the magnitude of extreme events but also in the frequency of the events. So we suggest that POT series should be given more focus in nonstationary flood return level analysis. Please see the response to General comment 2.1 for more explanations.

(24) Page 40, 820-823: What does this means from the practical perspective?

Response:

Thanks a lot for this good question. We have shown detailed explanation in the response to Specific comment (21) which is very similar to this comment.

(25) Page 40, lines 830-832: Does this hold for this case study or in general?

Response:

Thanks for pointing out the vague statement. This finding was obtained based on the specific hydro-climatic circumstances in the Weihe basin where flood magnitude observations show a significant decreasing trend under the impact of climate change, but did not refer to the general conclusion for other basins. For application in different regions, how climate change would influence design floods should be studied according to the actual situations in those regions. In the revised manuscript, we have rephrased the vague statement.

Thanks again for your professional and valuable comments which greatly improve our research and paper.

With best wishes

Yours sincerely

Lihua Xiong, PhD, Professor

State Key Laboratory of Water Resources and Hydropower Engineering Science

Wuhan University

Wuhan 430072, PR China

E-mail: xionglh@whu.edu.cn

Telephone: +86-13871078660

Fax: +86-27-68773568