Response to reviews

Reviewer #1:

The Authors present a methodology to estimate a methodology for a design life level-based risk analysis that takes into account univariate and bi-variate non-stationarity. The methodology presented is organized in 3 steps. The research presented is of potential interest for the readers of HESS but there some points need to be revised by the Authors. I suggest to accept the manuscript with major review.

Response to Reviewer 1:

Great appreciation for this comment!

In order to make the description of the methodology more readable and easier to comprehend, considerable revision has been made to the original manuscript. We have taken the review's suggestions into consideration. A detailed point-by-point reply has been made as follows:

General comments

There are some points that I would like the Authors to address in the manuscript:

1) For the design of which structure are the variables *Im* and *Ps* significant?

Response: Copulas have been widely used as for multivariate frequency analysis of extreme rainfall events (Zhang and Singh, 2007; Kao and Govindaraju, 2008; Rauf and Zeephongsekul, 2014; Vandenberghe et al., 2010). In this study, the proposed nonstationary model can not only used to make hazard assessment for extreme rainfall events but also for flood and drought events. We focus more on the proposed method and show how to implement it in the multivariate hazard assessment. *Ps* denotes annual extreme rainfall volume while *Im* (annual maximum daily precipitation) is just the annual rainfall intensity. In my opinion, variables *Im* and *Ps* would be significant for the urban flood control and drainage facilities. Not like the flood event, the volume and peak attributes has direct relation with the dam spillway. Here, variables *Im* and *Ps* would be indirectly related to the urban drainage facilities. If the *Ps* and *Im* intensify,

```
```

the pressure of the drainage facilities would be larger.

2) In the Introduction climate change is indicated as one of the motivation to propose a nonstationary design life level-based risk analysis, but there is any attempt to project the results of the manuscript in the future. Shall the Authors provide some indication of what shall we expect in the future? Do the Authors compare their projection of hydrological extremes with the trends that can be derived from climate models?

Response: In the revised manuscript, the average annual reliability (AAR) method to quantify the probability that a hydraulic system would be safe over its planning period in univariate case (Salas and Obeysekera, 2014; Read and Vogel, 2015) is adopted because of its agreement with nonstationarity. We have accepted this good suggestions of predicting the future extremes in the future. Since the nonstationary models with time as covariates, it can be used to predict the extreme value in the future 5 years (2018-2022). Considering new dataset being added to the original data, the parameters of the time-varying or stationary models would change. We limited the years of future to 5 years. In the same way, the ending year of the design life period is set as 2022 (**Figure 5**).

In this study, the time-varying marginal and copula models are established on the observed dataset (1958-2017). It would show different trend analysis results with the synthetized time series from the seven GCM simulations (BCC-CSM 1.1, CanESM2, CNRM-CM5, CSIRO-Mk3.6.0, IPSL-CM5A-LR, IPSL-CM5A-MR, NorESM1-M) involved in CMIP5 under RCP4.5 scenarios as shown in Table R1. Based on the analysis of manuscript, the extreme values or their dependence structure showed a significant trend. So we compared the series from these 4 stations. The difference between the observed extreme series and the simulated series are obvious. And if we make the trend analysis based on the predicted data set from the GCM models for the period of 1961-2100, the trend is significant (Z for Ps of station 1 is -5.82 while Im of station 1 is -4.32) at 5% level. And the results of Mann-Kendall trend analysis is bound to be the length of extreme series. And the above of trend comparison is beyond our scope of this paper, we focus more on the guidelines of how to establish the nonstationary models and the importance of considering the nonstationarity in hydrologic design by quantile estimation based a certain AAR level.

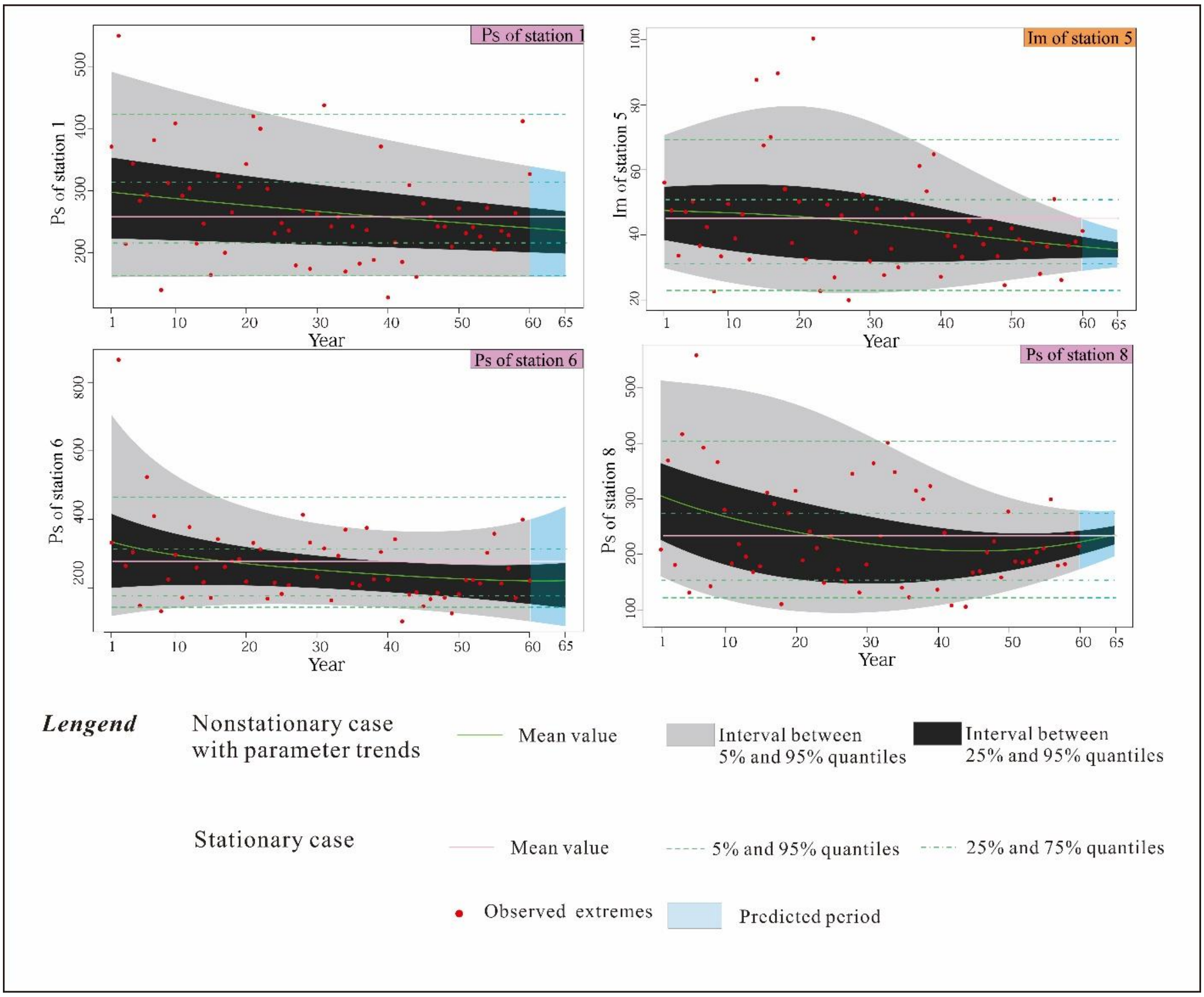

Figure 5. Nonstationary marginal distributions with trends in parameter during observed and predicted periods

Table R1. The trend analysis results of synthesized time series (1961-2017) from Seven GCM models

Station No.	Attribute	AC Test	Univariate MK	Multivariate MK
		p.value	Z	Z
1	***Ps***	0.69(0.93)	-0.34(-1.72)	-0.197(-0.76)
	Im	0.75(0.79)	0.07(0.31)	
5	***Ps***	0.67(0.89)	-1.32(-0.53)	-0.22(-1.25)
	Im	0.65(0.57)	-0.81(-1.79)	
6	***Ps***	0.57(0.99)	-0.98(-1.33)	-0.25(-1.32)
	Im	0.79(0.97)	-0.68(-1.16)	
8	***Ps***	0.67(0.53)	-0.69(-1.35)	-1.12(-1.38)
	Im	0.58(0.52)	-0.58(-1.22)	

Note: data in the bracket is corresponding value calculated on the observed dataset from 1961-2017.

```
```

3) Why do the Authors limits their analysis to six rain-gauges when there are several more in the area (e.g. https://doi.org/10.1002/hyp.9607)?

Response: We have taken the suggestions of the comment suggested by Reviewer 2. Also, in the following comment, the selected extreme series did not show a significant trend at 5% level. So we considered more stations in the Haihe River Basin. In addition, the original 95-th percentile threshold for *Ps* has changed to 0.90. Because of the limitation data length and missing data, the daily precipitation with period 1958-2017 was available at 16 stations. Fortunately, the absolute value of Z statistics of the Ps of station 2 and dependence structure of station 6 is larger than 1.96 which the threshold value of 5% significance level. And also, the change-point (CP) tests recommend change point existed in the *Ps* and dependence structure of *Ps* and *Im* from station 7. Based on the trend and change-point analysis, the single extremes or their dependence structure of station 2, 6, 7 showed a monotonic trend or change point. And based on LR statistic tests which are used to test whether linear or nonlinear trend existed in parameter, only 1,5, 6 and 8 showed trend in parameters. For the space limitations, we only show eight stations which contains all the nonstationary series in these stations. And this study mainly focus on how to establish the nonstationary copula models to make the hazard assessment from the hydrologic design. Based on the LR and nonparametric tests (Mann-Kendall and change point), there are 11 stations out of these 16 stations did not have stationarity.

```
```

Table R2 Trend analysis for the additional stations which are not shown in manuscript

No.	Attribute	AC Test	Univariate MK	Univariate CP Test	Multivariate MK	Multivariate CP Test
		p.value	Z	p.value	Z tatistics	p.value
9	Ps	0.48	-0.13	1.29	-0.58	0.57
	Im	0.76	-0.94	0.56		
10	Ps	0.70	-0.93	0.61	-0.86	0.40
	Im	0.69	-0.68	1.12		
11	Ps	0.54	-0.38	1.20	-0.61	0.69
	Im	0.17	-0.76	0.88		
12	Ps	0.14	-1.71	0.09	-1.58	0.11
	Im	0.46	-1.51	0.18		
13	Ps	0.68	-0.24	1.38	-0.21	0.83
	Im	0.85	-0.13	1.41		
14	Ps	0.54	-0.41	0.75	-0.31	0.43
	Im	0.49	-0.18	0.87		
15	Ps	0.58	-0.82	0.14	-1.08	0.25
	Im	0.94	-1.25	0.93		
16	Ps	0.53	-1.47	0.32	-0.91	0.35
	Im	0.74	-0.25	0.51		

4) Which are the limits/problems relate to use an upper bounded distribution (i.e. GEV when κ <0) to describe variables that potentially range between [0 +∞).

Response: This is a formula expression error. In order to make the fit of marginal distribution fit more objectively, we take five kinds of extreme value distribution into consideration: two kinds of 3-parameter distribution (Generalized Extreme Value, GEV; Pearson type III, PIII) and three kinds of 2-parameter distributions (Gamma, Weibull and Lognormal).

```
```

5). There is a significant difference between quantiles reported in Figure 4 and those computed using equation 2 and parameters reported in bold in Tables 3(a) and 3(b) with t=1960, 1970, 1980, 1990, 2000, 2010, 2020.

Response: As stated in response to comment 4, the selected extreme series did not show a significant trend at 5% level. On the one hand, we considered more stations in the Haihe River Basin. On the other hand, the original 95-th percentile threshold for *Ps* has changed to 0.90. And potential marginal models are shown in **Table S1 (a)-(d)**. Finally, based on the LR tests, GOF tests and AICc criterion. For marginal distribution, the extreme attributes (*Ps* or *Im*) extracted from station 1, 5, 6, 7 and 8 showed nonstationarity. The nonstationarity for station 1 , 5, 6 and 8 was because of existed trend in parameter while the existed change point (year 1979) was the reason why incorporate the nonstationarity in station 7.

```
```

S1 (a) Patterns of the time-varying models for 3-parameter marginal distribution (GEV in this study)

Model	μ	σ	κ
GEV0	constant	constant	constant
GEV1	$\mu = \mu_0 + \mu_1 t$	constant	constant
GEV2	$\mu = \mu_0 + \mu_1 t + \mu_2 t^2$	constant	constant
GEV3	constant	$ln\sigma = \sigma_0 + \sigma_1 t$	constant
GEV4	constant	$ln\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	constant
GEV5	constant	constant	$\kappa = \kappa_0 + \kappa_1 t$
GEV6	constant	constant	$\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$
GEV7	$\mu = \mu_0 + \mu_1 t$	$ln\sigma = \sigma_0 + \sigma_1 t$	constant
GEV8	$\mu = \mu_0 + \mu_1 t$	constant	$\kappa = \kappa_0 + \kappa_1 t$
GEV9	constant	$ln\sigma = \sigma_0 + \sigma_1 t$	$\kappa = \kappa_0 + \kappa_1 t$
GEV10	$\mu = \mu_0 + \mu_1 t + \mu_2 t^2$	$ln\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	constant
GEV11	constant	$ln\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	$\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$
GEV12	$\mu = \mu_0 + \mu_1 t + \mu_2 t^2$	constant	$\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$
GEV13	$\mu = \mu_0 + \mu_1 t$	$ln\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	constant
GEV14	constant	$ln\sigma = \sigma_0 + \sigma_1 t$	$\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$
GEV15	$\mu = \mu_0 + \mu_1 t$	constant	$\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$
GEV16	$\mu = \mu_0 + \mu_1 t + \mu_2 t^2$	$ln\sigma = \sigma_0 + \sigma_1 t$	constant
GEV17	constant	$ln\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	$\kappa = \kappa_0 + \kappa_1 t$
GEV18	$\mu = \mu_0 + \mu_1 t + \mu_2 t^2$	constant	$\kappa = \kappa_0 + \kappa_1 t$
GEV19	$\mu = \mu_0 + \mu_1 t + \mu_2 t^2$	$ln\sigma = \sigma_0 + \sigma_1 t$	$\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$
GEV20	$\mu = \mu_0 + \mu_1 t + \mu_2 t^2$	$ln\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	$\kappa = \kappa_0 + \kappa_1 t$
GEV21	$\mu = \mu_0 + \mu_1 t$	$ln\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	$\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$
GEV22	$\mu = \mu_0 + \mu_1 t$	$ln\sigma = \sigma_0 + \sigma_1 t$	$\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$
GEV23	$\mu = \mu_0 + \mu_1 t + \mu_2 t^2$	$ln\sigma = \sigma_0 + \sigma_1 t$	$\kappa = \kappa_0 + \kappa_1 t$
GEV24	$\mu = \mu_0 + \mu_1 t$	$ln\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	$\kappa = \kappa_0 + \kappa_1 t$
GEV25	$\mu = \mu_0 + \mu_1 t$	$ln\sigma = \sigma_0 + \sigma_1 t$	$\kappa = \kappa_0 + \kappa_1 t$
GEV26	$\mu = \mu_0 + \mu_1 t + \mu_2 t^2$	$ln\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	$\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$

```
```

S1 (b) Patterns of the time-varying models for 3-parameter marginal distribution (PIII in this study)

Model	μ	σ	κ
PIII0	constant	constant	constant
PIII1	$\mu = M * \sin(\mu_0 + \mu_1 t)$	constant	constant
PIII2	$\mu = M * \sin(\mu_0 + \mu_1 t + \mu_2 t^2)$	constant	constant
PIII3	constant	$\sigma = \sigma_0 + \sigma_1 t$	constant
PIII4	constant	$\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	constant
PIII5	constant	constant	$ln\kappa = \kappa_0 + \kappa_1 t$
PIII6	constant	constant	$ln\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$
PIII7	$\mu = M * \sin(\mu_0 + \mu_1 t)$	$\sigma = \sigma_0 + \sigma_1 t$	constant
PIII8	$\mu = M * \sin(\mu_0 + \mu_1 t)$	constant	$ln\kappa = \kappa_0 + \kappa_1 t$
PIII9	constant	$\sigma = \sigma_0 + \sigma_1 t$	$ln\kappa = \kappa_0 + \kappa_1 t$
PIII10	$\mu = M * \sin(\mu_0 + \mu_1 t + \mu_2 t^2)$	$\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	constant
PIII11	constant	$\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	$ln\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$
PIII12	$\mu = M * sin(\mu_0 + \mu_1 t + \mu_2 t^2)$	constant	$ln\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$
PIII13	$\mu = M * \sin(\mu_0 + \mu_1 t)$	$\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	constant
PIII14	constant	$\sigma = \sigma_0 + \sigma_1 t$	$ln\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$
PIII15	$\mu = M * \sin(\mu_0 + \mu_1 t)$	constant	$ln\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$
PIII16	$\mu = M * \sin(\mu_0 + \mu_1 t + \mu_2 t^2)$	$\sigma = \sigma_0 + \sigma_1 t$	constant
PIII17	constant	$\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	$ln\kappa = \kappa_0 + \kappa_1 t$
PIII18	$\mu = M * sin(\mu_0 + \mu_1 t + \mu_2 t^2)$	constant	$ln\kappa = \kappa_0 + \kappa_1 t$
PIII19	$\mu = M * \sin(\mu_0 + \mu_1 t + \mu_2 t^2)$	$\sigma = \sigma_0 + \sigma_1 t$	$ln\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$
PIII20	$\mu = M * sin(\mu_0 + \mu_1 t + \mu_2 t^2)$	$\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	$ln\kappa = \kappa_0 + \kappa_1 t$
PIII21	$\mu = M * sin(\mu_0 + \mu_1 t)$	$\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	$ln\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$
PIII22	$\mu = M * sin(\mu_0 + \mu_1 t)$	$\sigma = \sigma_0 + \sigma_1 t$	$ln\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$
PIII23	$\mu = M * \sin(\mu_0 + \mu_1 t + \mu_2 t^2)$	$\sigma = \sigma_0 + \sigma_1 t$	$ln\kappa = \kappa_0 + \kappa_1 t$
PIII24	$\mu = M * \sin(\mu_0 + \mu_1 t)$	$\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	$ln\kappa = \kappa_0 + \kappa_1 t$
PIII25	$\mu = M * \sin(\mu_0 + \mu_1 t)$	$\sigma = \sigma_0 + \sigma_1 t$	$ln\kappa = \kappa_0 + \kappa_1 t$
PIII26	$\mu = M * \sin(\mu_0 + \mu_1 t + \mu_2 t^2)$	$\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	$ln\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$

Note: M represents the minimum value of the observed time series.

```
```

S1(c) Patterns of the time-varying models for 2-parameter marginal distribution containing scale and shape parameter (Weibull, Gamma in this study).

Model	σ	κ
WE0/GA0	constant	constant
WE1/GA1	$ln\sigma = \sigma_0 + \sigma_1 t$	constant
WE2/GA2	$ln\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	constant
WE3/GA3	constant	$ln\kappa = \kappa_0 + \kappa_1 t$
WE4/GA4	constant	$ln\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$
WE5/GA5	$ln\sigma = \sigma_0 + \sigma_1 t$	$ln\kappa = \kappa_0 + \kappa_1 t$
WE6/GA6	$ln\sigma = \sigma_0 + \sigma_1 t$	$ln\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$
WE7/GA7	$ln\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	$ln\kappa = \kappa_0 + \kappa_1 t$
WE8/GA8	$ln\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$	$ln\kappa = \kappa_0 + \kappa_1 t + \kappa_2 t^2$

S1(d) Patterns of the time-varying models for 2-parameter marginal distribution containing location and scale parameter (Lognormal function in this study).

Model	μ	σ
LOGN0	constant	constant
LOGN1	$\mu = \mu_0 + \mu_1 t$	constant
LOGN2	$\mu = \mu_0 + \mu_1 t + \mu_2 t^2$	constant
LOGN3	constant	$ln\sigma = \sigma_0 + \sigma_1 t$
LOGN4	constant	$ln\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$
LOGN5	$\mu = \mu_0 + \mu_1 t$	$ln\sigma = \sigma_0 + \sigma_1 t$
LOGN6	$\mu = \mu_0 + \mu_1 t$	$ln\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$
LOGN7	$\mu = \mu_0 + \mu_1 t + \mu_2 t^2$	$ln\sigma = \sigma_0 + \sigma_1 t$
LOGN8	$\mu = \mu_0 + \mu_1 t + \mu_2 t^2$	$n\sigma = \sigma_0 + \sigma_1 t + \sigma_2 t^2$

```
```

Table 4(a). Results of marginal models with no trend for the distribution parameters corresponding to both extreme attributes

Station	Attribute	Model	μ	σ	κ	*AICc*	*KS*
2	***Ps***	LOGN0	5.49 [5.42,5.57][a]	-0.95 [-1.10,-0.81]	—	720.40	0.88
	Im	GEV0	61.11 [55.58,69.15]	3.30 [3.11,3.44]	0.24 [0.065,0.43]	607.31	0.99
3	***Ps***	GA0	—	2.17 [1.85,2.41]	2.79 [2.55,3.13]	599.85	0.95
	Im	GEV0	35.08 [32.85,38.23]	2.43 [2.61,2.23]	0.25 [0.055,0.47]	504.21	0.79
4	***Ps***	GA0	—	2.64 [2.31,2.87]	2.51 [2.26,2.82]	638.59	0.84
	Im	GA0	—	1.80 [1.44,2.04]	2.06 [1.80,2.39]	508.01	0.85
7	***Ps***^{b}	GA0	—	3.46 [2.67,3.93]	2.12 [1.71,2.85]	252.40	0.72
	Ps^{c}	LOGNO	5.28 [5.20,5.39]	-1.03 [-1.24,-0.87]	—	435.43	0.81
	Im	GA0	—	2.44 [2.09,2.66]	1.94 [1.68,2.34]	577.74	0.99

[a]parameter uncertainties with 90% confidence bands; *Ps*^{b} represents the *Ps* time series before change point 1979. *Ps*^{c} represents the *Ps* time series after change point 1979.

```
```

Table 4(b). Results of nonstationary marginal models with a certain trend for the parameters

Station	Attribute	Model	μ			σ			κ	AIC_c	LR^2	KS
			μ_0	μ_1		σ_0	σ_1					
1	***Ps***	LOGN5	5.64 [5.51,5.78]	-0.0031 [-0.0065, 0.0022]		-1.07 [-1.46,-0.84]	-0.0069 [-0.018,0.0022]		—	691	0.021	0.75
	Im	LOGN0	4.03 [3.93, 4.09]			-1.04 [-1.17,-0.93]			—	533	—	0.66
5	***Ps***	GAO	—			2.45 [2.12,2.72]			2.57 [2.34,2.89]	620	—	0.45
			μ_0	μ_1		σ_0	σ_1	σ_2(*1e-4)				
	Im	LOGN6	3.83 [3.69,3.96]	-0.0041 [-0.0076,-0.0012]		-1.37 [-2.08,-1.02]	0.036 [0.0048,0.091]	-7.8 [-15.6,-3.5]	—	479	2.2e-4	0.73
			μ_0	μ_1		σ_0	σ_1	σ_2(*1e-4)				
6	***Ps***	LOGN6	5.67 [5.50,5.82]	-0.006 [-0.0098,-0.0014]		-0.57 [-1.07,-0.34]	-0.04 [-0.067,-0.031]	5.8[2] [-0.8,10.5]	—	712	1.2e-5	0.68
	Im	GEV0	58.84 [53.92,65.44]			3.10 [2.91,3.29]			0.283 [0.078,0.43]	587	—	0.57
			μ_0	μ_1(*1e-2)	μ_2(*1e-4)	σ_0	σ_1	σ_2(*1e-4)				
8	***Ps***	LOGN8	5.680 [5.40,5.93]	-2.1 [-3.9,-0.5]	2.7[2] [0.3,5.3]	-1.07 [-1.72,-0.80]	0.032 [-0.0016,0.071]	-7.7[2] [-14.4,-3.2]	—	702	8.5e-4	0.95
	Im	LOGN0	4.18 [4.10,4.27]			-0.86 [-1.02,-0.74]			—	572	—	0.68

[1]parameter uncertainties with 90% confidence bands;[2] this value is value in the*1e-4 because of the table space limitations;[2]p.value of LR tests.

```
```

6) In the presence of a statistically significant change point, is it correct to use the same parameter formulation to describe data “before” and “after” the abrupt change point? Moreover, which would be the results of trend analysis if the time series are split as before” and “after” the change point. Do the Mann-Kendall test’ results change considering the before” and “after” segment of the time series separately?

Response: Firstly, thanks a lot for this suggestion by reviewer 1. Your idea is just in accordance with the paper by Salvadori et al. (2018), which is just the correct and objective procedure when the change point did exist in the extreme series. Following this suggestions, we made univariate and multivariate change point analysis for these 8 stations with results in **Table 3**. The *Ps* and dependence structure of *Ps* and *Im* of station 7 showed change point at year 1979. Also the Mann-Kendall (MK) trend analysis was also implemented for these two spilt series. For station 7, ***Before change point (1979),*** Z=0.26 for univariate MK test for *Ps*, Z=0.99 for Multivariate MK tests; ***After change point (1979)*** Z=0.45 for univariate MK for *Ps* and Z=0.62 for Multivariate MK tests. Compared to the MK tests for the whole data, the statistics (Z) changed a lot from 1.54 to 0.26 and 0.45 for *Ps*. The same situation existed in Multivariate MK tests. For the sake of showing more visually, the change point at year 1979 is plotted in **Figure 3**. In the following analysis of ARR quantiles, the quantile of each AAR quantile was calculated for these two spilt series no matter in univariate or bivariate case.

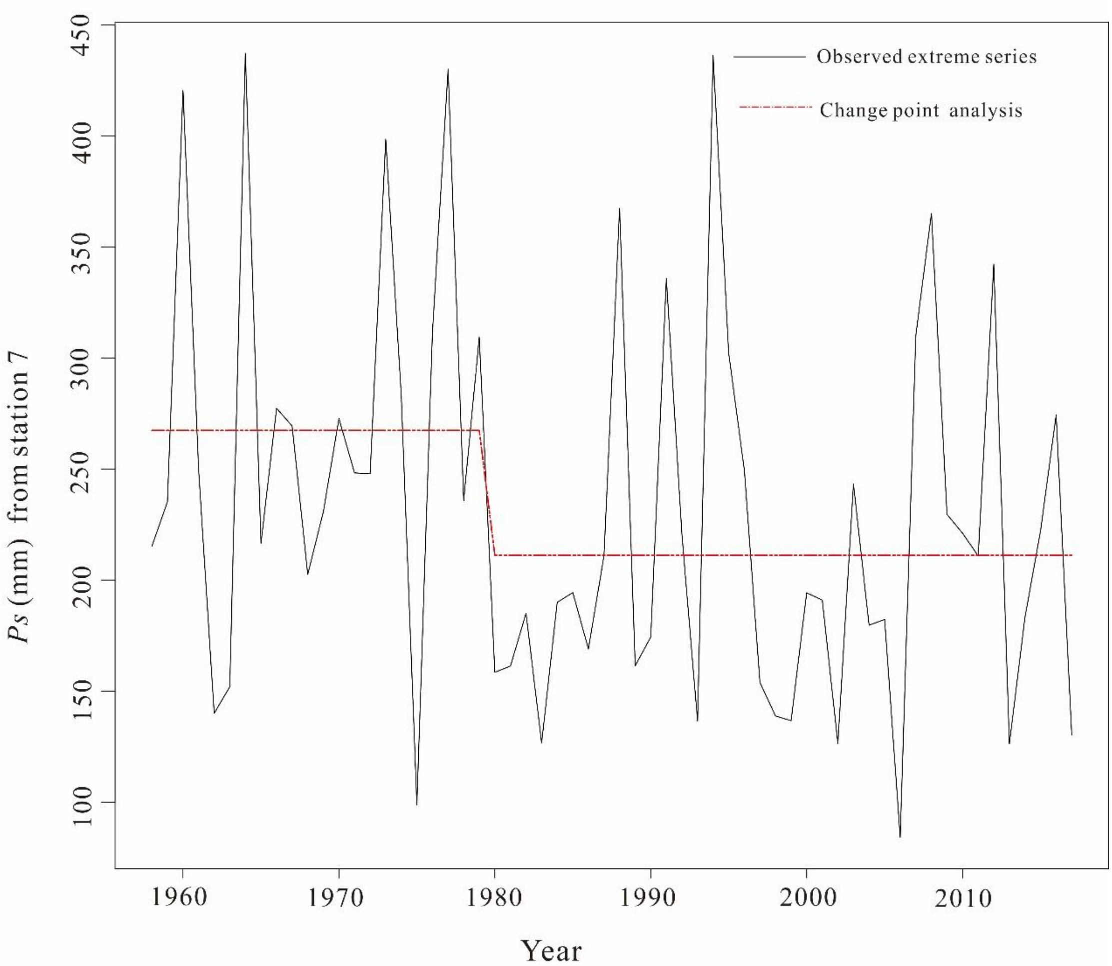

Figure 3. Plots of the change point according to the CP tests**.**

```
```

Table 3. Detection of trends and change points in extreme rainfall attributes collected from six stations

Station No.	Attribute	AC Test	Univariate MK	Multivariate MK	Univariate CP Test		Multivariate CP Test	
		p.value	Z	Z	p.value	CP	p.value	CP
1	***Ps***	0.93	**-2.03***	-0.86	0.069	—	0.062	—
	Im	0.79	0.42		0.48	—		
2	***Ps***	0.79	-0.63	-0.43	0.67	—	0.70	—
	Im	0.62	-0.16		0.53	—		
3	***Ps***	0.19	-0.20	-0.72	0.96	—	0..66	—
	Im	0. 32	-1.13		0.33	—		
4	***Ps***	0.05	0.71	0.61	0.64	—	0.46	—
	Im	0.17	-0.76		0.43	—		
5	***Ps***	0.89	-0.82	-1.48	0.14	—	0.15	—
	Im	0.57	-1.93		0.071	—		
6	***Ps***	0.99	-1.91	**-2.02***	0.16	—	0.055	—
	Im	0.97	-1.93		0.12	—		
7	***Ps***	0.38	-1.54	-1.70	**0.033**	1979	**0.025**	1979
	Im	0.77	-1.66		0.11	—		
8	***Ps***	0.53	-1.45	-1.45	0.17	—	0.31	—
	Im	0.52	-1.26		0.33	—		

MK: Mann-Kendall tests; CP: change point tests; AC: Autocorrelation tests; For station 7, ***Before change point (1979),*** Z=0.26 for univariate MK test for *Ps*, Z=0.99 for Multivariate MK tests; ***After change point (1979)*** Z=0.45 for univariate MK for *Ps* and Z=0.62 for Multivariate MK tests.

```
```

7) As last I would like to suggest the Author to add:

a. A section on climate change projection and analysis that can be of interest for future infrastructure design

b. one table reporting the basic statistics (min/max/mean/standard deviation) of the Ps and Im variables and the values of the 95-th percentile threshold to help understanding the variability of datasets;

c. one figure showing the time series with the indication of the change point year of occurrence according to Pettitt test.

Response: for suggestion (a), we have add this suggestion in section 3.3 as follows: "Since the nonstationary models with time as covariates, it can be used to predict the extreme value in the future. Considering new dataset being added to the original data, the parameters of the time-varying models would change. In this study, we assume the parameters of the selected time-varying models did not change a lot in the future 5 years. So the estimated time varying models from the original extreme series were used to predict year 2018-2022 (t=61-65). Based on the same assumption, the ending year of design life period in the following AAR-based quantiles calculation is set as 2022. As shown in **Figure 5**, mean value of Ps from station 1 and station 6 exhibited a downward trend while mean value of Im from station 5 and Ps from station 8 exhibited a NMT trend. For the predicted period, the predicted nonstationary marginal distributions for the extremes extracted from these 4 stations presented smaller mean values than those of the stationary distributions. Furthermore, the divergence of mean values between them are becoming larger as time goes on for

```
```

station 1, 5 and 6. But for Ps from station 8, the divergence of mean values between them are becoming smaller as time goes on.”

And in section 3.5-“Nonstationary hazard assessment based on AAR metrics for univariate and bivariate cases”，the design life period is set as 1983-2022. The future 5 years from 2018 to 2022. Considering new dataset being added to the original data, the parameters of the time-varying or stationary models would change. We limited the years of future to 5 years for cautiousness. These two adding parts are of great interest for the future hydrologic design.

For suggestion (b), we add **Table 2.** to show the basic statistics (min/max/mean/standard deviation) of the *Ps* and *Im* variables and the values of the 90-th percentile threshold to help understanding the variability of datasets.

Table 2. Basic statistics of the extreme variables observed at the original 8 stations

No.	ID	*Ps*(mm)					*Im*(mm)			
		Min	Max	Mean	Sd	$\overline{Th}$	Min	Max	Mean	Sd
1	53588	127.9	549.3	269.0	81.2	22.0	28.0	113.4	59.6	21.2
2	53986	80.1	636.2	263.2	107.7	26.6	24.3	414.0	85.0	56.5
3	54208	70.1	229.6	143.0	35.5	16.4	20.3	154.7	45.2	22.3
4	54311	65.8	299.2	172.5	49.1	18.7	18.1	100.9	47.3	17.3
5	54401	65.9	273.4	152.7	42.1	18.1	19.9	100.4	43.3	15.8
6	54511	102.2	865.9	263.6	115.1	28.7	32.8	253.5	79.7	44.7
7	54518	84.1	437.3	231.8	87.6	26.5	29.8	181.4	79.8	30.6
8	54602	87.2	559.0	230.1	93.1	27.3	23.7	185.6	71.8	32.9

$\overline{Th}$: represents the mean value of the 90% percentile threshold value for the attribute *Ps*.

For suggestion (c), we plot the change point for the Ps extreme series from station 7 (**Figure 3**).

```
```

Specific comments

Line 116 The definition given of Ps variable recall me the index R95pTOT used in climate change studies (http://etccdi.pacificclimate.org/list_27_indices.shtml). Is it the same index? In addition, could the Authors specificy the period of observation they used to set the 95-th percentile threshold? According to R95pTOT index the reference period to set the 95-th percentile threshold is 1961-1990.

Response: In this study, the Ps is different from the R95pTOT index. It is just annual total precipitation of the daily precipitation more than the 90th percentile threshold for each year. That means, each year has a unique threshold value. We have made its definition more readable and clear.

Line 290 -291 The Authors write where R_i^{ns} and R_i^{s} are nonstationary risk and stationary risk of a certain hydraulic structure for a design life of i years", but 'i' goes from 1 to n. I would expect that 'n' indicates the design life and 'i' indicates the i-th year from now (i.e. the year the project "starts") to n-th year (end of the project's life).

Response: This risk quantified metric is deleted from the manuscript. We accept the average annual reliability (AAR) method to quantify the probability that a hydraulic system would be safe over its planning period in univariate or bivariate case (Salas and Obeysekera, 2014; Read and Vogel, 2015).

```
```

Lines 367-368 The Authors write “Except for stations 4 and 5, the best distributions for the other stations were parallel for nonstationarity tests shown in Section 4.1”. Is it possible that the mismatch between the nonstationarity test results and the best fitting distribution for Im (station 4 and 5) and Ps (station 5) was to the choice of the Author to ignore the test’s results?

Response: as stated by the reviewer, it is of great possibility the mismatch between the nonstationarity test results and the best fitting distribution for Im (station 4 and 5) and Ps (station 5) was to the choice of the Author to ignore the test’s results. And we use 10% significance level which entailing a large probability of rejecting Null Hypothesis of non-stationarity. So in revised forms, we take a careful statistical tests with 5% significance level for Mann-Kendall (MK), change point tests. We also proposed the Log Likelihood ratio (LR) tests (Coles (2001)), which is more rigorous trend detection methods than nonparametric methods (MK). For most cases, the results of trend analysis from LR tests are consistent with that by MK tests. However, as shown in Table 3(b), the trend in the parameter existed in Im of station 5, Ps of station 6 and 8 based on the LR tests at the 5% significance level which recommends another situation different from the previous MK tests. Through analysis, it would be caused by the he opposing trend in the location and scale parameters.

Lines 387-390 The Authors write “Contrary to station 5, the nonstationary St copula fitted better than did the stationary model for stations 1 and 6 which was not in accordance with the nonstationarity tests for these two stations (Table 2).” It is true that

```
```

according to bivariate MK test results station 1 and 6 should stationary, but at station 1 bivariate Pettitt test shows the presence of a change point; the presence of a change point could have influenced the results of LL and AICc ? What will happen if Im and Ps time series are "broken" before and after the change point to LL and AICc estimates?

Response: we have taken this suggestion into considerations. The extreme series from station 7 in the revised manuscript showed change point. The *Ps* and dependence structure of *Ps* and *Im* of station 7 showed change point at year 1979. Also the Mann-Kendall (MK) trend analysis was also implemented for these two spilt series. For subseries, no trend can be detected based on Mann-Kendall and LR tests. The best marginal distribution and copula is shown in **Table 4(a)** and **Table 5(a)**.

Line 411 (and Conclusions) The Authors report a value of 355 mm for the 100-year Ps quantile in station 1 under stationary circumstances, but using the parameters reported in Table 3(a) the 100-year Ps quantile in station 1 under stationary circumstances is about 383 mm. It is probably a matter of approximation in the parameters values (355 mm corresponds to a report period of about 62 yr) but I will suggest the Authors to check these values.

Response: Since average annual reliability (AAR) method to quantify the probability of the hydraulic structure is of great potentiality in communicating hazard of failure under both stationary and nonstationary conditions (Read and Vogel, 2015), we adopt ARR method to estimate the quantiles for a design life period including future 5 years.

```
```

Table 5(a). The stationary copula for the dependence structure of the *Ps* and *Im* for station 1-5 and 7

Station	Model	Kendall's tau	θ	β	$AICc$	AD_{RT}
1	FR0	0.55	6.65 [4.53,8.36][1]	—	-43.69	0.46[3]
2	JC0	0.65	-0.22 [-0.55,0.18]	-1.22 [-1.29,-1.15]	-81.59	0.77
3	SJC0	0.56	-0.78 [-1.15,-0.52]	-1.23 [-1.29,-1.16]	-57.27	0.64
4	SJC0	0.58	-0.35 [-0.98,0.11]	-1.28 [-1.44,-1.18]	-63.46	0.76
5	SJC0	0.55	-0.26 [-0.61,0.01]	-1.34 [-1.46,-1.25]	-57.81	0.88
7[1]	FR0	0.4	4.43 [2.21,8.69]	—	-5.71	0.76
7[2]	SGU0	0.61	0.43 [-0.58,0.70]	—	-34.73	0.63
8	SGU0	0.58	0.33 [0.021,0.61]	—	-56.63	0.59

```
```

Minor corrections

Around the manuscript there are some typos like “Pettist” instead of “Pettitt”; missing spaces and so on (e.g Lines 226, 228), please check the text.

Response: Following the study of Salvadori et al. (2018), the non-parametric change-point statistical tests were implemented to check that whether the marginal or joint distributions are sensitive to changes. These tests can be manipulated in the R package npcp (Kojadinovic, 2017). We replaced Pettitt tests with the above statistical change point tests as suggested by Review #2.

Lines 172 and 173 Is the limit “(μ-σ)/κ” for lower (upper) boundary of x value correct? According to parameter’s estimates in Tables 3(a)-3(b), when κ<0, x can assume only negative values, that is non coherent with the variables Ps and Im that are positively defined.

Response: It is an error of formula definition.

Lines 249-250 The Authors write “Let JRP_{s-and} and JRP_{s-ken} represent the three types of return period in the stationary case”, but the return periods presented are only 2.

Response: it is an error of station. It should be stated as follows: Let we calculate the joint return period from the AND and Kendall Scenario.

Line 260 “JPRs” probably was “JRPs”

Response: we have made deep self-checking of the notations.

```
```

Line 343 and Line 426 check the correct location of Figure 3

Response: we have checked this figure's location.

Lines 359-365 The Authors write "The best fitted model was selected by performing the minimum *DIC* criterion combined with the Bayes factor (*BF*) test", but looking at bold rows in tables 3(a) and 3(b) the criterion of minimum DIC seems not be respected for Im at station 2 where EVns-1 is in bold instead of GEVns-2 (minimum DIC value).

Response: The reviewers' suggestions and views are right. The selected extreme series from the manuscript did not show a significant trend at 5% level. We have checked the process of extreme value extraction. We have changed the original 95-th percentile threshold for *Ps* has changed to 0.90-th percentile for *Ps*. And we take data from more stations into consideration. And based on the analyzed results, the nonparametric tests were consistent with the LR tests in most cases.

Line 360-365 Comparing these lines with Table 3(a), for station 1, the variable described as GEVns-2 appears to be Ps and not Im. BF for Im variable in station 1 is >1. Please clarify this point.

Response: in the revised manuscript, we use the LR tests to select the nonstationary models with trend in parameter. The process can be defined as follws:

"Let ST represent the stationary model for the extreme attribute (Ps and Im in this study) and let NST be the time-varying model with trend existed in the parameter. In the same way, let LL_1 and LL_0 denote the log-likelihoods under model ST and NST.

The log likelihood ratio statistics (LR) would converge to the χ^2_Δ distribution with Δ being the divergence of the parameter number between the ST and NST models. And LR can be formulated as follows (Coles, 2001):

$$LR = 2(LL_1 - LL_0) \tag{3}$$

The LR statistics can be regarded as a good criterion to check whether the Null trend hypothesis ($\mathcal{H}_0$: there is no trend for the distribution parameter) can be rejected or not. If the value of LR is more than the upper-α point of the χ^2_Δ distribution, the above Null assumption can be rejected at a significance level of α (in this study α is equal to 5%)."

Line 387 Please define "MK"

Response: we have taken this suggestion to show the full definition of the Abbreviation.

Lines 409-411 Figure 3 illustrates the results of nonstationary tests. Figure 4 reports the extreme rainfall quantiles. Please check the text.

Response: We have checked the text.

##

References

```
```

Line 62 Does “Assia et al., 2014” refer to “Aissia, M.A.B., Chebana, F., Ouarda, T.B.M.J., Roy, L., Bruneau, P., and Barbet, M.: Dependence evolution of hydrological characteristics, applied to floods in a climatechange context in Quebec, J. Hydrol., 519, 148–163, https://doi.org/10.1016/j.jhydrol.2014.06.042, 2014” ?

Response: We have made revision for the reference quotation.

Line 98 Does “(Jakob, 2013)” refer to “Jakob, D., AghaKouchak, A. Easterling, D., Hsu, K., Schubert, S., and Sorooshian, S. (Eds.): Nonstationarity in extremes and engineering design, Springer, New York, 2013” ?

Response: We have made revision for the reference quotation.

Line 100 “Read and Vogel (2015)” there is no correspondence in the references

Response: we have added the “Read, L.K., and Vogel, R.M.: Reliability, return periods, and risk under nonstationarity, Water Resour. Res., 51, 6381-6398, https://doi.org/10.1002/2015WR017089, 2015.” to reference list.

Line 126 Does “Nelson (2007)” refer to “Nelsen, R.B.: An introduction to copulas, Springer, New York, 2007.”?

Response: We have made revision for the reference quotation.

Line 212 “Genest et al., 1995” there is no correspondence in the references

Response: In revised manuscript, this reference is deleted

```
```

Line 213 “Hurvich and Tsai, 1989” there is no correspondence in the references

Response: we have added the “Hurvich, C. M. and Tsai, C. L.: Regression and time series model selection in small samples, Biometrika, 76, 297–307, https://doi.org/10.2307/1271469, 1989.” to the reference list

Line 235 “Fernandez and Salas, 1999” there is no correspondence in the references.

Response: we have used AAR metrics which did not contain the reference.

Ghanbari, M., M. Arabi, J. Obeysekera, and Sweet, W.: A coherent statistical model for coastal flood frequency analysis under nonstationary sea level conditions, Earth's Future, 7, 162-177, https://doi.org/10.1029/2018EF001089, 2017.” The publication year is 2019.

Response: we have changed the publication year.

“Zhang, Q. , Gu, X. , Singh, V. P. , and Chen, X.: Evaluation of ecological instream flow using multiple ecological indicators with consideration of hydrological alterations, J. Hydrol., 529, 711-722, https://doi.org/10.1016/j.jhydrol.2015.08.066, 2015 . ” should be moved at the end of the reference list.

Response: we have deleted the reference which is not included in the manuscript.

```
```

###

Table (1) I suggest the Authors to change Longitude and Latitude with Longitude E and Latitude N, respectively coherently with the choice of indicating geographical coordinates in degree/minutes format.

Response: we have accepted the suggestion and made revision to Table 1.

Table (2) "Ps" and "Im" should be in italic. For station 3 and multivariate MK test the "*" should be close to the Z-statistic value not to the p-value. I suggest the Authors to add the indication of the year at which the change point is detected for both univariate and bivariate Pettitt test.

Response: we have accepted the suggestion and showed the change point if it exists by change point tests. "Ps" and "Im" have be made in italic form.

Table 3(a) e 3(b) please specify the meaning of bold row, I guess that bold indicates the "best" fitting model, but in this case why for Im variable at station 2 the best model is GEVns-1 if GEVns-2 shows the minimum DIC?

Response: we have taken this suggestion in the revised manuscript.

Table 3(b) refers to (Station 4-6) not to (Station 2-6) and the '-' symbol is missing for BF values of stationary GEV in station 5.

Response: we have taken this suggestion in the revised manuscript.

```
```

Table 4(a) and 4(b) reports the meaning of bold and underlined text.

Infinity symbol cited in caption does not appear in the table, probably substituted by "NaN".

Response: we have taken this suggestion in the revised manuscript and improved the definition of table symbols.

###

Figure (1) Step S2 is omitted.

Response: we have taken this suggestion in the revised manuscript and improved the logical order of the flowchart.

Figure (2) I would like to suggest the Authors to add the Haihe river to the map.

Response: we have add the main stream of Haihe River to the map.

```
```

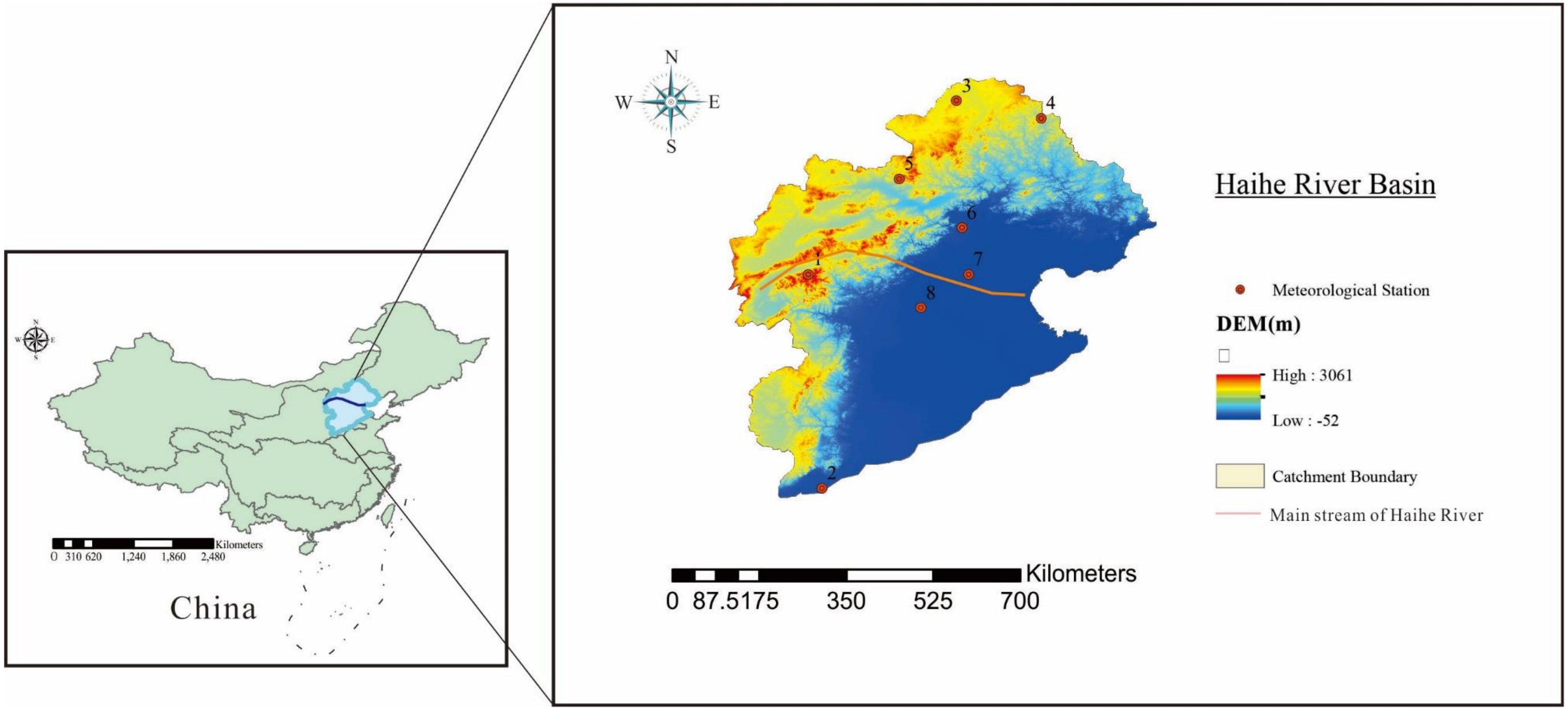

Figure 2. Selected meteorological stations in Haihe River basin

```
```

Figure (3) In the caption there is a typo “Mann-Kendalld” instead of “Mann-Kendall”. Please check the legend, the description of the last item (purple backward arrow) is equal to the one of the third one (green upward arrow). The “+” symbol is redundant with the test that already specify if the trend/change point is statistically significant.

Response: we have taken this suggestion in the revised manuscript

Figure (4) the “star” symbol is not defined.

Response: we have checked clearly in revised manuscript.

Figures (4), (6), (7) I would like to suggest the Authors to improve the quality of these figures.They seems to be a collection of screenshots with different size and background colour. Figure 4, in particular, seems to lack of organization in the sub-figures arrangement.

Response: we have taken this suggestion to improve the quality of the figures.

```
```