Responses to Anonymous Referee #1:

All the authors are extremely grateful to you for providing your excellent comments and valuable advice on this paper. Thank you for your comments, which are very helpful to improve our manuscript. We agree with your suggestions and we have made a major revision to our paper. We have given more explanations about used data and results.

In the following sections, your comments are in black text while corresponding actions taken by us following in blue.

General comments

The authors had proposes an improved logistic probability prediction model for water shortage risk and applied to Beijing (China) to predict water-shortage risk probability in 2020. Water supply is important in metropolitan-area development and mitigation of water-shortage is an essential measure to reduce shortage induced economic losses. This manuscript presents an important idea in such issues. However, parts of this manuscript including used data and results are not well explained in this manuscript. Major revisions are needed before this manuscript can be considered for acceptance. Specific comments are given below.

Responses: Good suggestions. We agree with your suggestions and we have made a major revision to our paper. We have given more explanations about used data and results. More details are shown in the following point-by-point response. We sincerely hope for your satisfaction with our revision. Thank you again for your kind suggestion.

Specific comments

1. Line 8. It is not clear what is the meaning of the “information flow” when it is firstly appeared in the Abstract. The authors should briefly explain the “information flow” in the Abstract.

Responses: Good suggestions. Information flow was proposed by Liang (2014), is a method for unraveling the cause-effect relation between time series. We have briefly define the “information flow” in the abstract. We sincerely hope for your satisfaction with our revision. Thank you again for your kind suggestion.

2. Line 139, equation (1). What is the definition of the total amount water resources (W)?

Responses：Good suggestions. The total amount water resources (W) refers to the water yield of surface water and groundwater formed by precipitation. We have given its definition in the revised manuscript. We sincerely hope for your satisfaction with our revision. Thank you again for your kind suggestion.

3. Line 164. What is the unit of the information flow? What is the “nats” per unit time?

Responses：Good suggestions. Information flow is expressed by the rate of information flowing (units: nats per unit time) from the latter to the former. The unit of the information flow is nats per unit time. Nat is the unit of entropy and one nat equals to 1.44 bits. We have given some explanations in the revised manuscript. We sincerely

hope for your satisfaction with our revision. Thank you again for your kind suggestion.

4. Lines 189, 190, 193, 194, 196, 198, 201, … What is the meaning of "L"?

Responses：Good suggestions. "L" in Eq. (12) is the log-likelihood function of maximum likelihood estimation, and its expression is as follows.

$$L = \sum_{i=1}^{n}\left[y_i \left(\alpha + \sum_{j=1}^{m} \beta_j x_{ij} \right) - \ln\left(1 + \exp(\alpha + \sum_{j=1}^{m} \beta_j x_{ij}) \right) \right]$$

The second referee believed that Eq. (12) is unnecessary because we didn't use the maximum likelihood estimation. So we have deleted Eq. (12). We sincerely hope for your satisfaction with our revision. Thank you again for your kind suggestion.

5. Line 193. Does pi denote the conditional probability of "water shortage"?

Responses：Good suggestions. pi denotes the conditional probability of "water shortage". We sincerely hope for your satisfaction with our revision. Thank you again for your kind suggestion.

6. Lien 195, equation (11). What is the definition of "K" in equation (1)?

Responses：Good suggestions. There is no "K" in equation (11). "K" is in equation (10). In equation (10), "K" is the number of unit time interval. We have given some explanations in the revised manuscript. We sincerely hope for your satisfaction with our revision. Thank you again for your kind suggestion.

7. Lines 306-307. What is the criterion to remove the Sr, IWp, and DWp should

be clearly defined in the manuscript.

Responses： Good suggestions. According to Table 3, the value of the normalized information flow of AW_p is only 0.0031, and it is very little. It was concluded that the AW_p does not result in a water shortage risk. Therefore, AW_p was removed as risk factors. Moreover, the values of the normalized information flow of S_r, IW_p, and DW_p are only 0.1247, 0.1320 and 0.1164 (Table 3). Therefore, S_r, IW_p, and DW_p are alternative factors to be removed. According to Table 4, the correlation coefficients between S_r and many other factors (e.g. W_p, U_r and P) are very large, so S_r was removed as risk factors. Similarly, IW_p, and DW_p are also removed as risk factors. Therefore, the selected factors for logistic regression model were W_c, W_p, U_r, P and DS_r. We have given more explanations about the criterion to remove AW_p, S_r, IW_p, and DW_p in the revised manuscript.

We sincerely hope for your satisfaction with our revision. Thank you again for your kind suggestion.

8. Figure 3 and Table 5. The authors should discuss why prediction of 1982 and two other years is that no risk occurs, but risk actually occurs.

Responses: Good suggestions. Because the logistic probability prediction model is a kind of statistical model, its accuracy rate can't reach 100%. As long as the accuracy rate can be greater than 80%, the model is applicable. The accuracy rate of our model is 91.18%, so the ability of the logistic regression model to predict water shortage is very strong.

We sincerely hope for your satisfaction with our revision. Thank you again for your kind suggestion.

9. Lines 358-360. Does the total amount of water resources from 1956 to “2020” mean 1956 to “2012”?

Responses：Good suggestions. We made a mistake. The total amount of water resources should be from 1956 to 2012. We sincerely hope for your satisfaction with our revision. Thank you again for your kind suggestion.

10. Figures 4, 5, and 6. How to obtain the prediction of water shortage in 2020 should be more clearly described in manuscript, especially the input data. For example, (1) Lines 354-356. How to estimate the GDP in 2020?

Responses：Good suggestions. We have given more details about the prediction of water shortage in 2020. The GDP of 2020 was the sum of the gross agricultural product, gross industrial product, and gross product of the third industry (In China, the third industry is also known as the service industry, and includes the traffic and transportation industry, communication industry, and commercial industry). The gross product of agriculture, industry and the third industry were predicted by a dynamic forecasting model with recursive compensation by grey numbers of identical dimensions. The theory of the dynamic forecasting model with recursive compensation by grey numbers of identical dimensions are shown in Qian et al. (2016).

We sincerely hope for your satisfaction with our revision. Thank you again for

your kind suggestion.

(2) Lines 358-360. The total amount of water resources from 1956 to 2020 (should be 2012) were considered in 2020. However, the total amount of water resources from 1979 to 2012 were provided Beijing Hydrological Station (lines 111-112), which is inconsistent with the statement in lines 358-360. What total amount of water resources used in 2020 should be precisely described in the manuscript.

Responses：Good suggestions. We are so sorry that we made a mistake. The total water resources in 2020 is 2.88 billion m^3, predicated by the autoregression model (confidence level: 95%; accuracy: 97%) using the total amount of water resources from 1979 to 2012. The total amount of water resources used in 2020 has been precisely described in the revised manuscript.

We sincerely hope for your satisfaction with our revision. Thank you again for your kind suggestion.

(3) Are the same year of precipitation and total amount of water resources used to predict water-shortage risk probability in 2020 shown in Figures 4, 5, and 6 ?

Responses：Good suggestions. Because the precipitation of 2020 is unknown, the precipitation in 2020 is assumed to be any annual precipitation from 1956 to 2012. The total water resources in 2020 was predicted by the autoregression model using the total amount of water resources from 1979 to 2012. Then we computed the sequences of W_c, W_p, U_r, P and DS_r in 2020 and substituted them into Eq. (20), the probability that a

water shortage risk will occur in 2020 under the inflow scenarios of 1956–2012 was predicted. The prediction of water-shortage risk probability in 2020 shown in Figures 4, 5, and 6.

We sincerely hope for your satisfaction with our revision. Thank you again for your kind suggestion.

11. Other minor editorial suggestions include: (1) Lines 35-36. “Giannikopoulou et al., 2015” is inconsistent with “Giannikopoulou et al., 2017” (lines 491-493) in the section of References. (2) Line 61. “Balakrishnan, 1992” is inconsistent with “Balakrishnan, 1991” (lines 476-477) in the section of References. (3) References listed in the section of References should be alphabetically listed.

Responses：Good suggestions. “Giannikopoulou et al., 2015” has been revised as “Giannikopoulou et al., 2017”. “Balakrishnan, 1992” has been revised as “Balakrishnan, 1991”. References listed in the section of References have been alphabetically listed. We sincerely hope for your satisfaction with our revision. Thank you again for your kind suggestion.