

## Response letter to Reviewer#1

We thank Reviewer#1 for the careful consideration of our work. We agree with his/her constructive and thoughtful comments and suggestions, which led to a much improved and complete manuscript. In this response letter, we have replied (in blue) to all the comments formulated by the Reviewer (in black).

### Comments:

Thank you for sending me the manuscript: "Quantifying the impacts of the Three Gorges Dam on the spatial-temporal water level dynamics in the Yangtze River estuary" by Huayang Cai for review, which I read with great interest.

The authors apply a linear regression model to the tidally averaged water level in the Yangtze estuary to investigate the effects of the Three Gorges dam. The authors find that their regression model predicts the water level in the Yangtze reasonably well. They find that since construction of the dam, low flows have increased while flows during transition from the high to the low flow season have decreased.

The topic is very relevant and the manuscript was interesting to read. The applicability of a regression model to predict water levels in tidal rivers agrees with my own experience in this field. The text and figures are of high quality.

However, the regression model applied here is relatively simple, at least much simpler than previously applied models. This certainly makes it easy to grasp the results, especially for readers who are not experts on the topic. However, this also makes it difficult to identify the physical drivers behind changes in the water levels, and might introduce systematic errors. Below, I provide suggestions on how these issues can be verified and mitigated, if necessary.

**Our reply:** We very much appreciate all the comments and suggestions raised by the reviewer. In the revised manuscript, we shall completely address all the comments.

### Methods

1. The regression model includes both discharge and water level at the upstream station. As they depend on each other, the model is not parsimonious. As a consequence, the columns for  $Q$  and  $Z_{up}$  of the regression matrix will be close to collinear so that small changes (errors) in the data can result in large changes in the coefficients  $\alpha$  and  $\gamma$  even if the fit is good. Possible changes of the coefficients over time might thus be regression

artefacts. This should be ruled out by verifying that  $Q_{up}$  and  $Z_{up}$  are not strongly correlated.

- If the correlation is weak, then the model is robust, but then it would be insightful to elaborate on why the upstream water level and discharge are unrelated. The comment "influenced by the dynamics of [] tributaries" (l. 119) is unclear. The water level is uniquely determined by the backwater curve as long as the daily averaged water level does not change rapidly in time. Therefore, tributaries upstream of the inflow boundary influence downstream levels only through their discharge. Do the authors refer to tributaries downstream of the upstream station?

Our reply: It can be seen from Figure R1 below that the correlation between  $Q_{up}$  and  $Z_{up}$  is indeed strong. However, we observe that the daily averaged water levels are not uniform for identical river discharge (see Figure R1a) due to the external forcing, either the potential influence induced by the tidal forcing or the exerted residual water level slope upstream of the DT hydrological station. Actually, the observed water levels at DT hydrological stations were influenced by both the residual water level slope upstream of the inflow boundary (owing to the relative importance of river discharge between the main stream and the tributaries, especially during the flood season) and that downstream of the inflow boundary (owing to the tidal forcing, especially during the dry season). To account for the influence of residual water level slope, in the previous manuscript we have explicitly introduced the  $z_{up}$  into the regression model.

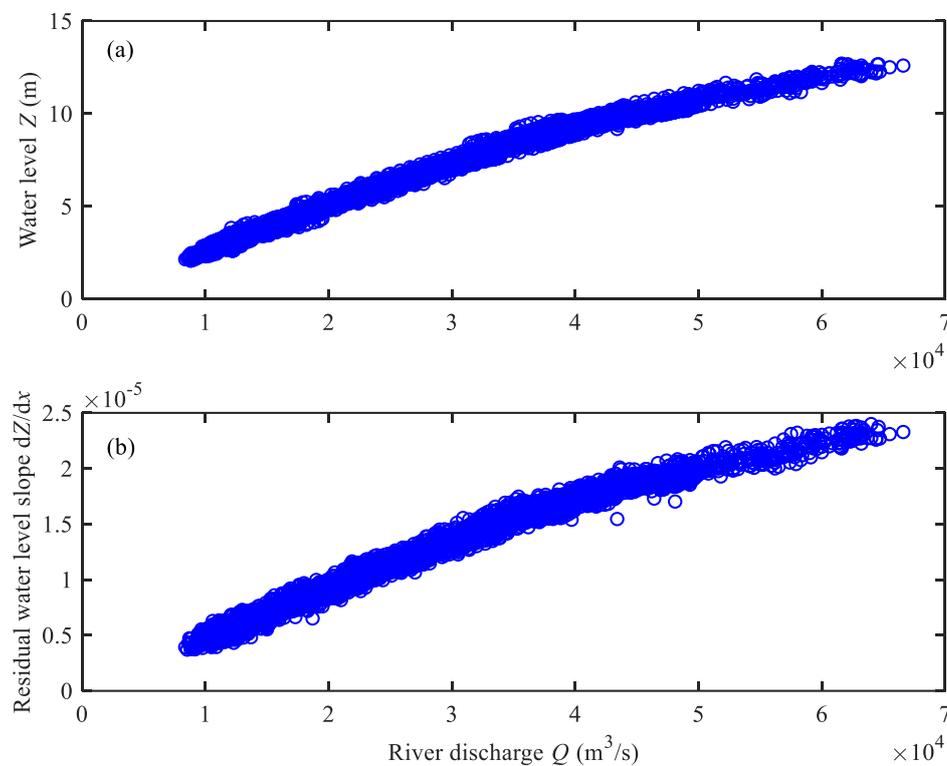


Figure R1. Relationship between water level and river discharge at the DT hydrological station (a) and that between residual water level slope for the whole estuary and river discharge (b).

- If the correlation is strong, then it is better to replace the terms  $\alpha Q + \gamma Z_{\text{up}}$  with the non-linear term  $\alpha Q^b$ . This model is less ambiguous. In my personal experience, the coefficients  $a$  and  $b$  of the non-linear model also give much more insight into the influence of the river discharge on the mean water level along tidal rivers.

Our reply: We very much appreciate the comments raised by the Reviewer. In this case, the regression model can be described by the following equation:

$$Z = Z_0 + \alpha Q^b + \gamma Z_{\text{down}} \quad (\text{R1})$$

where the potential influence of  $Z_{\text{up}}$  on water level dynamics is implicitly accounted by the nonlinear term  $\alpha Q^b$ . It can be seen from Figure R2 and Table R1 that the model performance is more or less the same as the original triple linear regression model, except that the RMSE values are slightly larger at NJ, MAS and WH stations (ranging between 0.17 and 0.21 m) than those using the triple linear regression model (ranging between 0.11 and 0.15 m).

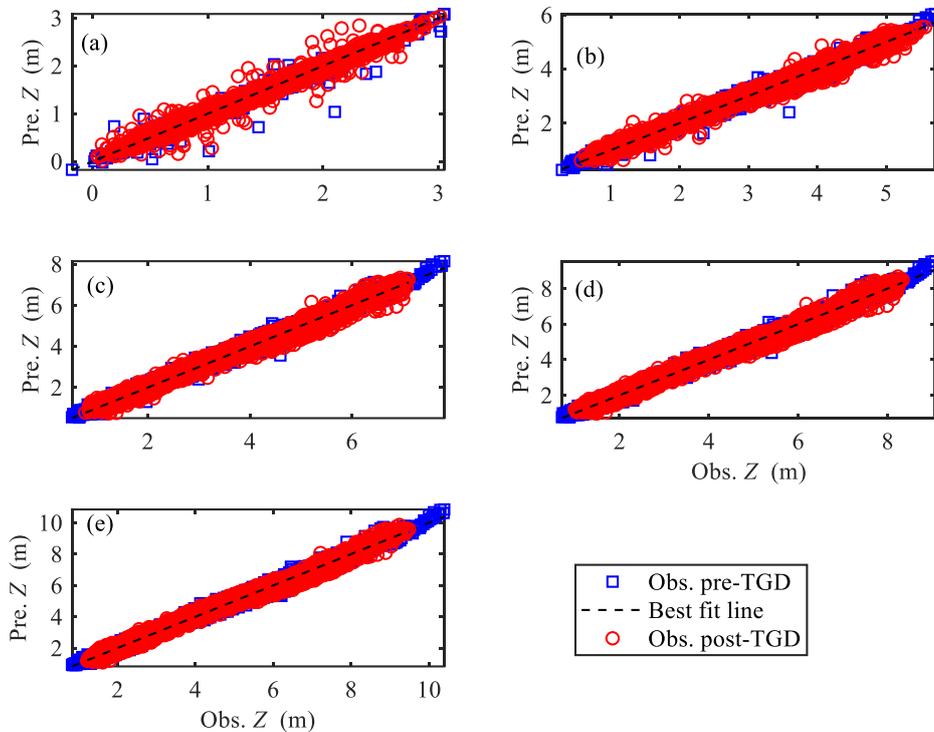


Figure R2. Comparison between predicted and observed daily averaged water levels for both the pre-TGD and post-TGD periods at different gauging stations along the YRE by means of Equation R1: (a) Jiangyin (JY), (b) Zhenjiang (ZJ), (c) Nanjing (NJ), (d) Maanshan (MAS), (e) Wuhu (WH).

Table R1. Calibrated regression coefficients for both the pre-TGD and post-TGD periods along the YRE by means of Equation R1.

Stations		$Z_0$	$\alpha$	$\beta$	$\gamma$	RMSE/m	Standard deviation/m
JY	pre-TGD	-0.10	0.31	0.62	0.49	0.06	0.64
	post-TGD	-0.44	0.55	0.45	0.43	0.08	0.59
ZJ	pre-TGD	0.00	0.89	0.92	0.48	0.13	1.23
	post-TGD	-0.27	1.02	0.82	0.43	0.14	1.12
NJ	pre-TGD	-0.37	1.84	0.81	0.40	0.17	1.72
	post-TGD	-0.55	1.57	0.84	0.40	0.19	1.54
MAS	pre-TGD	-0.57	2.50	0.77	0.38	0.20	2.02
	post-TGD	-0.64	2.03	0.83	0.37	0.21	1.80
WH	pre-TGD	-1.31	3.78	0.66	0.32	0.21	2.35
	post-TGD	-1.36	3.09	0.71	0.32	0.21	2.07

2. The regression model does not include the effect of the tides on the mean water level. However, this effect is not negligible during periods of low river flow (LeBlond, 1978). This introduces a systematic error. As the Three Gorges dam increased river discharge during the low flow season, this can bias the results. It is, therefore, reasonable to include the influence of tides on the mean water level in the regression model. For example, Kukulka and Jay (2003) suggest the regression model linear in  $h^3$ :

$$h^3 \approx aQ_{river}^2 + b|z_{tide}|^2 + c,$$

while (Kastner et al., 2019) suggested linearizing the backwater equation, which can be readily approximated in a regression model linear in  $h$  (or  $z$ ).

Our reply: Actually, the potential effect of the tides on the mean water level is implicitly considered by the  $Z_{down}$  term, which is typically featured by a spring-neap cycle. Figure R3 shows the autoregressive power spectral density estimate of the daily averaged water level observed at TSG gauging station, where significant periodic cycle of 14.8 days was observed.

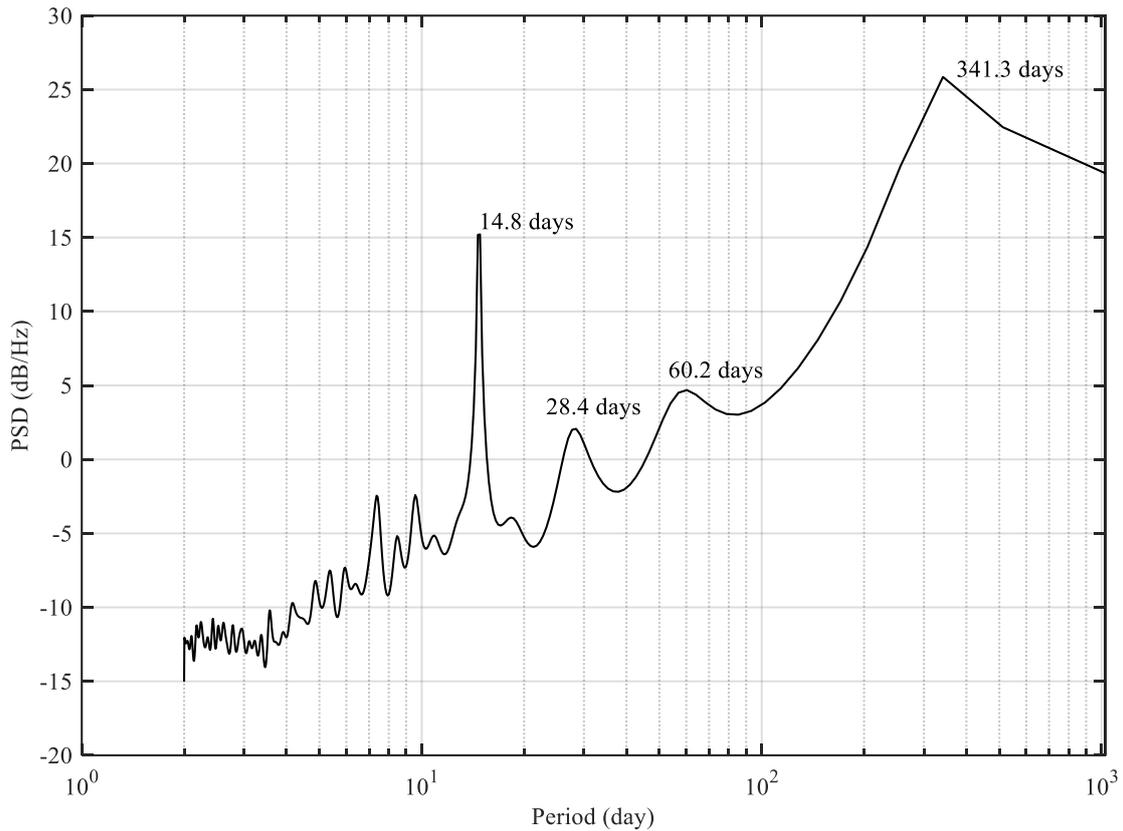


Figure R3. Autoregressive power spectral density estimate of the daily averaged water level observed at TSG gauging station

3. The independent variables are not normalized in the regression model so that the coefficients have very different magnitudes ( $O(\alpha) = 10^{-5}$  while  $O(\beta) = 1$ ). It is thus not obvious which predictor (downstream or upstream level) has the largest influence at a particular location. This can be revealed by normalizing the independent variables by their standard deviation before the regression:

$$Z = Z_0 + \alpha Q / \text{std}(Q) + \beta Z_{\text{down}} / \text{std}(Z_{\text{down}}) + \gamma Z_{\text{up}} / \text{std}(Z_{\text{up}})$$

This is preferable to the order in the study, where variance is normalized after the regression.

**Our reply:** We thank the reviewer to point this out. In the revised manuscript, we shall normalize the input parameters by their standard deviations.

4. Interpolation of slopes and uncertainty estimates (Figure 4 and 5, lines 190\_)

- There is a mistake in the slope calculation. The values should be in the order  $10^{-5}$ , not  $10^{-8}$ . The distance between the stations was probably not converted from km to m.

**Our reply:** You are right! In the revised manuscript, we shall correct this mistake (see Figure R4 below).

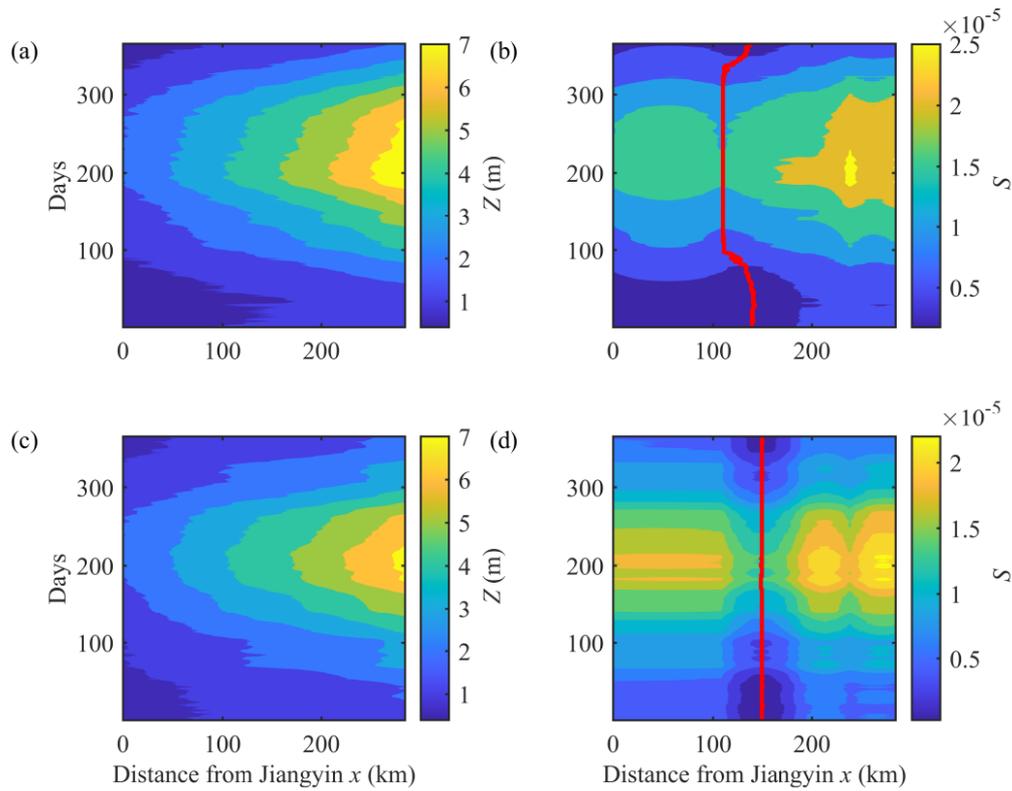


Figure R4. Reconstructed spatial-temporal water levels,  $Z$ , (a, c) and their slopes,  $S$ , (b, d) for the climatological year during both the pre-TGD (a, b) and post-TGD (c, d) periods. The red lines in subplots (b) and (d) indicate the local minimum water level slopes in the central section of the YRE (between Jiangyin and Zhenjiang).

- Determining the slope from by higher-order (Hermite) interpolation is not meaningful here. This is because the error (of the slope) is amplified at the interpolated values between the stations. As a consequence, the interpolated slope has unrealistic local extrema at the midpoints between stations (Figure 5). The error of the cubically (Hermite) interpolated slope at the midpoint between stations is about 1.8 times as large as the errors of the levels at the stations. Since the error of the levels is about 10%, the error of the slope is about 20%. The local maxima of the slope, as well as the difference between the pre- and post-TGD period as indicated in Figure 5 are therefore insignificant. The interpolation error (of the slopes) can be considerably reduced by calculating the slopes at the midpoint between two stations and then linearly interpolating the slopes between the midpoints. In this case the error of the slope is only 0.7 times that of the error in levels. If the authors want to retain cubic interpolation, then the spurious extrema can be suppressed by fitting the 4 coefficients of the cubic polynomial with all 5 stations in a least squares manner.

Our reply: Many thanks for the reviewer's comments on the interpolation of the results. Actually, we only interpolated the daily averaged water level along the estuary, while the slope is computed on the basis of the interpolated water level using the Matlab

“gradient.m” function. However, it is true that the unrealistic local extrema are mainly due to the amplification of the error of the interpolated water level.

- As the model is not parsimonious, it might fit well even if the regression coefficient are uncertain, as multiple parameter combinations can result in similar good model performance. A good way to assess the uncertainty is bootstrapping (Efron and Tibshirani, 1994). Simply split the time series into blocks comprising of one month, this reduces the effect of serial correlation. When there are  $n$  blocks, randomly choose  $\sqrt{n}$  blocks and fit the model. Repeat this a few hundred times. The standard error is simply the standard deviation of the estimated parameters. The standard error of the coefficient, predicted levels and slopes can then be indicated with error bars in Figure 3 and 5. The cubic interpolation results in larger errors at midpoints between sections, so errors bars are best placed there.

Our reply: We very much appreciate the reviewer’s comments on the quantification of the model uncertainty. Actually, since we used the default multiple linear regression function “regress.m” in Matlab, it is possible to output a matrix of 95% confidence intervals for the coefficient estimates, which is similar to the adoption of bootstrapping. In the revised manuscript, we shall update the Figure 3 by including the error bar for each coefficient (see Figure R5 below).

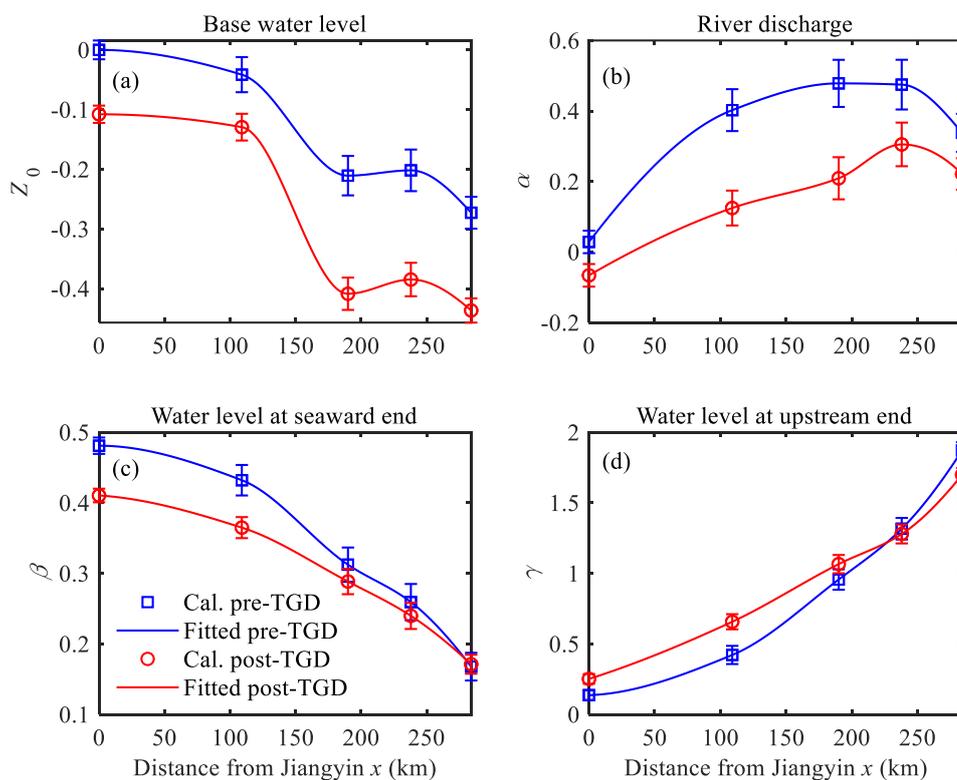


Figure R5. Interpolated linear regression coefficients  $Z_0$  (a),  $\alpha$  (b),  $\beta$  (c),  $\gamma$  (d) with error bar along the YRE (upstream of the Jiangyin gauging station) for both the pre-TGD and post-TGD periods. The vertical error bar was estimated using the Matlab ‘regress.m’ function with 95% confidence intervals.

## Minor

- Title estuary→upper estuary

Our reply: We agree with the reviewer's comment. The title will be revised as: "Quantifying the impacts of the Three Gorges Dam on the spatial-temporal water level dynamics in the upper Yangtze River estuary"

- 35 The term "analytical solution" is misleading, as the water level is still determined by (numerically) integrating an initial value problem (eq. 22 in Cai et al. (2016)).

Our reply: In the revised manuscript, we shall replace "analytical solutions" with "solutions".

- 44 Kukulka and Jay (2003) should be referenced here, as an important regression model for the mean water level of tidal rivers.

Our reply: In the revised manuscript, we shall include the reference of Kukulka and Jay (2003).

- 45 "these methods suggest that water level dynamics in estuaries are highly nonlinear and nonstationary" This sounds as if water levels in tidal are difficult to analyse and predict, and that looking at tidal cycle/average is a novel idea. However, there is a large amount of publications how water levels can be approximated well on a cycle-by-cycle basis, see the works of the groups of Savenije, Godin, Jay, Hoitink, and Friedrichs.

Our reply: In the revised manuscript, we shall revise this sentence as: "these methods suggest that water level dynamics in estuaries are highly nonlinear and nonstationary owing to complex tide-river interactions".

- 49 The reference to Darcy is dubious. Even if the surface level can be predicted by a linear regression model, it is still turbulent flow (quadratic flow resistance), which is very different from groundwater flow (linear flow resistance).

Our reply: In the revised manuscript, we shall remove "similar to Darcy's law for groundwater flow".

- 89 Mark Gaoqiaoju on the map in Figure 1

Our reply: In the revised manuscript, we shall mark Gaoqiaoju gauging station on the Map.

- 93 "we mainly concentrate on the tide-river dynamics" This is not the case, since, as commented by me before, the tidally induced water level offset is not included in the regression model.

Our reply: In the revised manuscript, to be more specific, we shall replace "tide-river dynamics" with "water level dynamics". Actually, since the input parameter  $Z_{down}$  implicitly considered the influence induced by the tidal forcing (especially the spring-neap changes), we actually concentrated on the tide-river dynamics.

- 110 Mention here, which of the stations were chosen as the upstream and the downstream end (Datong and Gaoqiaoju?).

Our reply: We agree with the reviewer's comment. In the revised manuscript, we shall explicitly mention that "In this study, the DT hydrological station was chosen as the upstream end, while the TSG gauging station being the downstream end."

- 169 "linear" is misleading here. The water depth in the upstream estuary most likely scale like  $h \approx Q^{2/3}$ . The non-linearity is just hidden by including  $Z_{up}$  in the regression model.

Our reply: In the revised manuscript, we shall explicitly mention that "which leads support to our hypothesis that the response of water level dynamics to hydrodynamics at both ends of the estuary is largely linear in the YRE owing to the explicit inclusion of  $Z_{up}$  in the regression model."

- 248 The conclusion "[at the downstream stations] tide dominates [the tidally averaged water level]" sounds odd, as the regression model applied in this study does not explicitly account for the tidally induced water level offset. It only includes the tidally averaged water level at the seaward station. However, at the river mouth the tidally induced water level offset is negligible as it integrates along the estuary, c.f. Kastner et al. (2019) and Cai et al. (2016). So, no meaningful conclusion about the tidal influence can be drawn. The model probably indicates that fluctuations of the sea level unrelated to tides, such as wind, ocean-temperature and ocean-salinity, dominate the mean water level dynamics near the sea. It would be insightful to actually determine the tidal influence by including it explicitly in the regression model.

Our reply: Actually, since the input parameter  $Z_{down}$  in the regression model implicitly considered the influence induced by the tidal forcing (especially the spring-neap changes), we actually concentrated on the tide-river dynamics.

- 248 The river discharge influences the salinity gradient, and with it the variation of the water level at the reference station at the sea (Savenije, 2012). The influence on river discharge on the downstream stations might thus be larger than indicated by the model.

Our reply: We agree with the reviewer that the salinity gradient may influence the water level at the reference station at the sea. However, since the study area is out of the maximum salt intrusion length, thus the potential influence due to salinity gradient is negligible.

- 257 This paper has → We have

Our reply: In the revised manuscript, we shall replace "This paper has" with "In this study, we have".

- 263 It was shown → We show

Our reply: We agree with the reviewer's comment.

- 271 How relevant are (seasonal) changes of roughness and bedforms, due to changes in water and sediment supply by the dam?

Our reply: Here we can conclude that the main impact due to changes in water and sediment supply by the dam tends to deepen the riverbed since the alterations caused by geometric changes are negative.

- Figure 2 It would be more meaningful to plot  $(Z_{\text{pred}} - Z_{\text{obs}})$  vs  $Z_{\text{obs}}$  and to use smaller dots which do not overlap that much. This would reveal better any systematic variation.

Our reply: We agree with the reviewer's comments. In the revised manuscript, the Figure 2 will be revised as follows (see Figure R6 below).

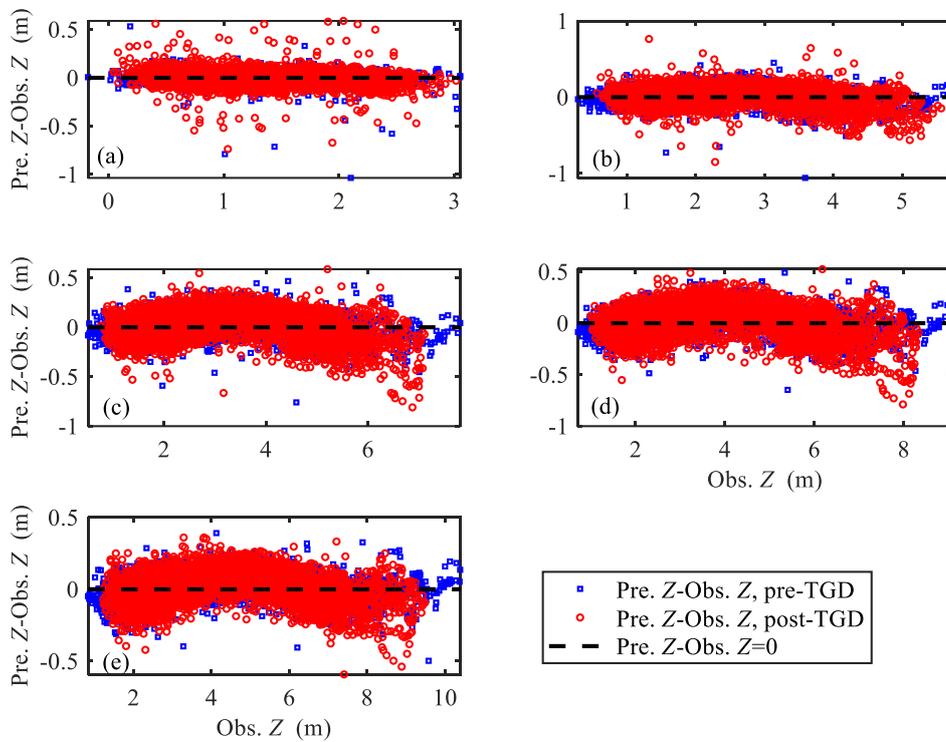


Figure R6. Alterations in difference between predicted and observed daily averaged water levels as a function of observed daily averaged water levels for both the pre-TGD and post-TGD periods at different gauging stations along the YRE: (a) Jiangyin (JY), (b) Zhenjiang (ZJ), (c) Nanjing (NJ), (d) Maanshan (MAS), (e) Wuhu (WH).

- Figure 3 Add subplots titles, like Discharge, Downstream level, Upstream level so that the figure can be interpreted without looking up the meaning of the coefficients  $\alpha$ ,  $\beta$ ,  $\gamma$ .

Our reply: We agree with the reviewer's comment. In the revised manuscript, we shall include the subplots titles (see Figure R5 above).

■ Figure 3 begins from Jiangyin→upstream of Jiangyin

Our reply: We agree with the reviewer's comment.

■ Figure 7 The average annual average hydrograph of the post-TGD period is corrupted by high-frequent fluctuations of the hydrograph. The graph would be clearer if the fluctuation is removed it through by smoothing with a sliding window. A triangular window with a width of 30 days seems appropriate. Smooth the data for the pre-TGD period as well, for better comparison.

Our reply: We agree with the reviewer's comments. In the revised manuscript, the Figure 7 will be revised as follows (see Figure R7 below).

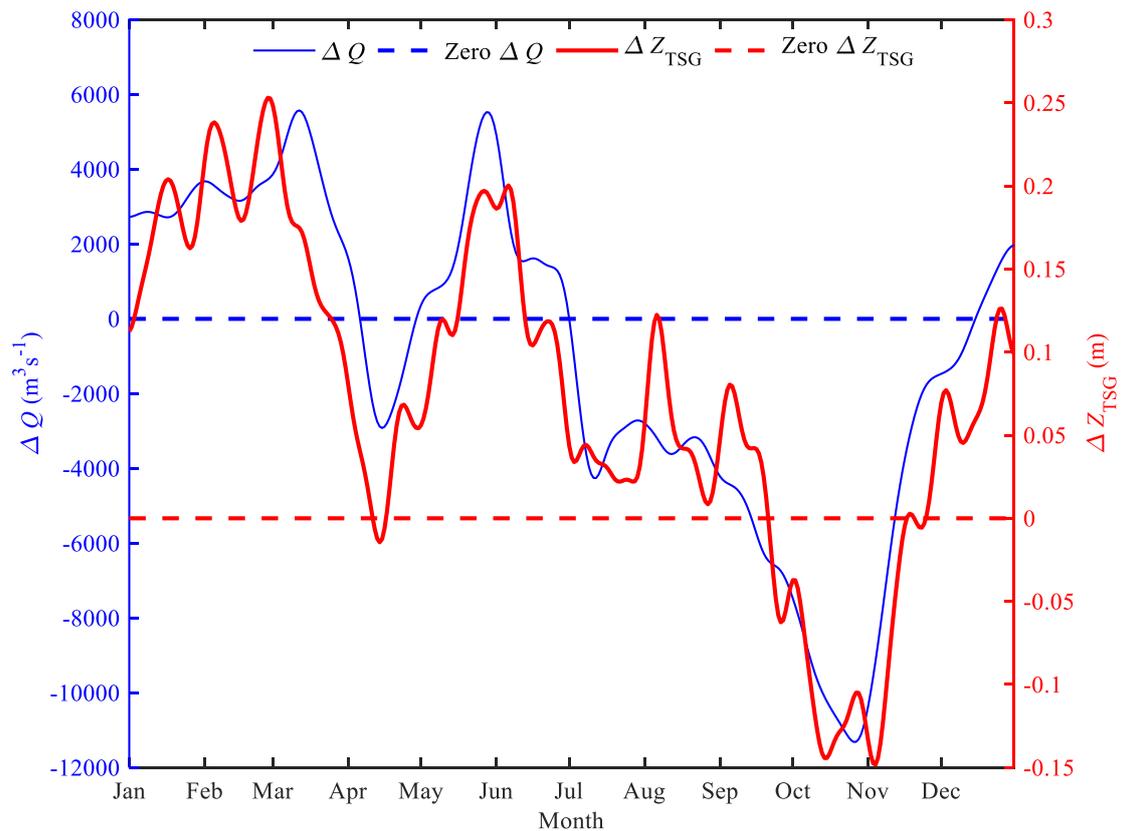


Figure R7. Alterations in river discharge and water level observed at DT and TSG, respectively, during the post-TGD period relative to the pre-TGD period over the climatological year. The daily averaged river discharge and water level were smoothed using a moving average filter with a span of 30 days.

#### References:

Cai, H., Savenije, H. H. G., Jiang, C., Zhao, L., and Yang, Q.: Analytical approach for determining the mean water level profile in an estuary with substantial fresh water discharge, *Hydrol. Earth Syst. Sci.*, 20, 1177-1195, <https://doi.org/10.5194/hess-20-1177-2016>, 2016.

Kastner, K., Hoitink, A.J.F., Torfs, P.J.J.F., Deleersnijder, E., Ningsih, N.S.: Propagation of tides along a river with a sloping bed. *J. Fluid Mech.*, 872, 39-73, <https://doi.org/10.1017/jfm.2019.331>, 2019.

Kukulka, T., and Jay, D. A.: Impacts of Columbia River discharge on salmonid habitat: 2. Changes in shallow-water habitat, *J. Geophys. Res.*, 108(C9), 3294, <https://doi.org/10.1029/2003JC001829>, 2003.

Savenije, H.H.G., 2012. *Salinity and tides in alluvial estuaries*. completely revised 2nd edition. Available from <http://www.salinityandtides.com> [Accessed 11 June 2022].