Reply on RC3

Yunxiang Chen et al.

Author comment on "Modeling of streamflow in a 30☐km long reach spanning 5 years using OpenFOAM 5.x" by Yunxiang Chen et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-313-AC3, 2022

[Reviewer comments]: This manuscript presents a detailed implementation of a 3D CFD model for a 30-km transect of the Columbia River. The model is calibrated with surface water elevations (SWE) and validated with SWE and velocity profiles. As part of this effort, an analysis of the model performance at three different time scales was performed: (1) short-term (2012), (2) middle-term (2013-2016), and (3) long-term (2018-2019).

The authors present a novel and practical approach to impose boundary conditions and calibrate roughness scales and shear stresses in large-scale 3D CFD models. After calibration and validation, the authors compared the new model results with previous modeling efforts in the study site. This analysis highlighted the importance of a detailed representation of 3D processes in channel flow modeling. Finally, the authors assessed the relative importance of dynamic and hydrostatic pressure along the reach, focusing on the potential implications for environmental and ecological functions in river systems.

Summary: Overall, this manuscript is well-written, clear, and represents an important contribution to the field of river hydraulics. The implementation approach is well-documented and reproducible. This will make for a significant contribution to GMD.

I include a commented pdf with editorial suggestions. In the following, I present two significant comments that require attention and a series of general observations.

[Responses]: We thank the review's comments and detailed suggestions in the attached pdf file. They are very helpful to improve the paper.

[Reviewer major comments 1]: The authors' argument for using a spatially-variable roughness (ks) is that this parameter is expected to vary for a complex river system with heterogeneous bedforms and high curvature. I agree with this argument. However, the selection of roughness regions seems guided by geometric convenience and data availability and not the spatial variations that are reasonable controls for ks. For instance, I would expect that a characterization of the different depositional environments and corresponding bedforms would provide a better guide for selecting roughness zones. The PNNL team has previously performed such classification for the study reach, which could be used. The main issue here is that given the complexity of this model, the authors are likely dealing with a non-unique solution, and the computational burden prevents them from analyzing this in detail. Some discussion in this regard would be valuable in this manuscript and essential to guide future model refinements. I suggest including some

discussion in lines 415-422.

On a related note, I wonder how sensitive is the spatial variation of stage and velocity to the calibrated ks within a roughness region. Again, the authors assess this sensitivity at the point scale, where the SWE observations are available, but the overall response may be insignificant.

[Responses]: We thank the reviewer's invaluable comments and suggestions. The reviewer is correct that we designed the model calibration procedure based on geometric convenience and data availability. Though it will be of great value to measure water stage at diverse locations (e.g., shorelines and thalwegs) and environments (e.g., gravels, dunes, mixed gravel and sand, vegetations), logistics and safety issues need to be considered in field surveys. For example, to install the water stage survey devices safely, we cannot walk to the depth water (>2 m) regions. Figure 1c shows the exact locations of each stage survey locations (red stars), and we can see that all these locations are very close (<100 m) to the river banks. Even though they are very close to the shorelines, the water depth at these locations could vary between 1 m – 4 m most of the time which can be observed in Figure 5c. Furthermore, devices deployed too close to the river centerline also impose dangers to normal human activities (e.g., boating, shipping, kayaking, etc.) and could be more easily destroyed by floating objects (woods) that more likely occur in the river center regions. Due to these reasons, the water stage deployment was guided by geometric convenience in this work. In addition, the water stage surveys were mostly conducted 10 years ago, therefore, the strategy of how to divide the riverbed into multiple roughness subregions is designed based on existing survey locations, but not guided by the characteristic river planform (sinuosity/curvature/narrow-width variations) and depositional environments. For future applications, however, it is recommended to first design the survey locations based on characteristic river planform and depositional environments (if logistics and safety are not an issue), and then apply the modeling framework presented in this paper. Relevant discussions on these issues have been added to lines 450-455 in the attached revised version of the paper.

Regarding the non-unique solution issue, a discussion on how to address this issue is added in lines 446-455. The key takeaway message is that the final model performance is controlled by both local optimal roughness estimation and local roughness adjustment. The first step is purely based on the error diagram and it will provide a unique roughness value for each location. The second step needs to adjust the roughness estimated from the first step based on (1) our physical understanding of how the characteristic river planform and depositional environments affect the water stage and (2) how much model more error at each location. As this is a case-by-case and trial-and-error procedure, it may not guarantee a unique solution. However, with a better design of the survey location distribution, it is likely to achieve a high model accuracy even without the second step. The higher the accuracy in the first step, the less important the second step. However, if the second step is a must, the latest machine learning approaches such as parameter learning and reinforcement learning could be potential directions to explore in the future.

To clarify the sensitivity of stage and velocity variations to roughness, a new section 4.1.3 and Figure 12 are added. Based on the discussions, we have the following answers: (a) the spatial variation of water stage is not sensitive roughness at locations near the river valleys, however, it is sensitive to the roughness at locations near the river banks; (b) the absolute value fo the water stage near the river valley regions vary 18% to 28% in response to changing roughness from 0 to 0.5 m, however, the water stage near the river banks experiences significantly more variations; (c) the spatial distribution of free surface velocity magnitude is not sensitive to non-zero (zero roughness is a toally different story) roughness, however, its maximum value could vary 25% to 30% when the roughness height increases from 0.025 m to 0.5 m. In short, the water stage near the banks and the velocity near the river valleys are very sensitive to roughness. By contrast, the water

stage near the valley and the velocity near the banks are less sensitive to roughness. More details of these discussions can be found in lines 456-489.

[Reviewer major comments 2]: The idea of analyzing the relative importance of dynamic and hydrostatic pressure is an excellent illustration of the importance of these models to gain a mechanistic understanding of exchange processes along the sediment-water interface. I commend the authors for including this analysis! However, I expect the conclusions regarding the exchange to be incorrect. The reason for my skepticism is that the exchange process is driven by gradients in the pressure distribution and not by its magnitude. In other words, I suggest that the authors revise this analysis and focus on the spatial variability of the pressure gradients.

[Responses]: We thank the reviewer's insightful questions. The reviewer is correct to point out that it is the pressure gradient at the riverbed that drives the exchange fluxes. The question is how to obtain the pressure gradient. Direct measurements of the pressure gradient at natural streambeds are challenging, therefore, numerical models are usually used to simulate the pressure gradients. The pressure (and gradient) at the riverbed is the summation of hydrostatic and hydrodynamic pressure. Depending on how to deal with the dynamic pressure, existing models that study surface-subsurface interactions can be categorized into three types: no dynamic pressure models, one-way coupled dynamic pressure models, and two-way coupled dynamic pressure models. The first type of model ignores the dynamic pressure. The second type of model solves the 3D hydrodynamics equations for the surface water and then uses the computed total pressure (the summation of hydrostatic pressure and dynamic pressure) as a boundary condition to drive the subsurface flow without considering the feedback of flow from the subsurface to the surface. The third type of model is similar to the second type but allows the flow from the subsurface to the surface. Though the third type of model can provide the pressure gradients at the riverbed, they are usually very computationally expensive and numerically unstable, and thus are currently only used at laboratory-scale and idealized flow conditions. A more commonly used approach is the one-way coupled model. In this approach, it first assumes the dynamic pressure gradient at the streambed is zero, and then solves the Navier-Stokes equation to obtain the magnitude of the dynamic pressure. In this paper, we also set a zero dynamic pressure gradient at the riverbed (see line 213). With the dynamic pressure at the riverbed, the one-way coupled model solves the subsurface flows (e.g., 3D Richards equations) by setting the total pressure at the riverbed as a boundary condition. This finally provides the pressure gradients (due to both hydrostatic and dynamic pressure) at the streambeds. Therefore, it is important to evaluate the relative importance of dynamic pressure to static pressure to understand how dynamic pressure affects the exchange fluxes. Actually, in another work, we have used the total pressure computed from the CFD model to drive a subsurface flow model and reported the impacts of dynamic pressure on exchange fluxes. To address the reviewer's concerns, we have added the above discussions in lines 555-580. The effects of dynamic pressure on exchange fluxes can also be found in Bao, J. et al., Modeling framework for evaluating the impacts of hydrodynamic pressure on hydrologic exchange fluxes and residence time for a large-scale river section over a long-term period. Environmental Modelling & Software, 2022, 148,105277.

[Reviewer general comments]:

p7, l 180: For clarity, the roughness elements directly resolved are larger than 1m in "the vertical direction." To be precise, you could include "the vertical direction" in the text since the horizontal resolution is much lower (20m).

[Response]: Yes, we have corrected this sentence at Line 185 in the revised version of the paper.

p9, l 206: the velocity components for the outlet seem to have the x- and y-direction components mixed

[Response]: Velocity is a three-component vector with the first, second, and third components denoting the velocity component along x, y, and z-direction. At the outlet, the velocity along x (east) and z (towards water surface) directions are zero, while the velocity along y-direction (north) is non-zero. Therefore, the velocity vector is (0, -uout, 0) and the equation at line 210 is correct.

p 11, l 255: To better illustrate the potential presence of systematic biases, I suggest plotting error vs. stage. For example, the bias for low SWE in Figure 9 will be more evident with this metric.

[Response]: To illustrate if using a water stage causes misunderstanding of the model performance (visual comparison, biase, mean error, mean absolute error, etc.), we add new subfigures in Figure 5 and Figure 9 by showing the comparison of water depth from the model and observations as well as their 1:1 plots. Both Figure 5d and 9d indicate that the model does not have systematic biases no matter using water stage or depth as an evaluation metric (see lines 267 and 385).

p 11, l 261 (and throughout the text): Using the tilde symbol "~" for value ranges is somewhat unconventional. I suggest using a dash "-"

[Response]: We thank the reviewer's suggestion. Yes, we replaced all "~" with a middle-sized dash to better represent the ranges but avoid confusing it with the minus symbol.

The authors use SWE to assess the model performance; however, this metric could be misleading, and I wonder if the water depth is a better alternative. In particular, when calibrating the roughness values, I expect the relative error in water depth to be a more reasonable measure of model performance.

[Response]: The response is similar to the comment in Line 255. Based on the new Figures 5c-d and 9c-d, we believe that using WSE and depth conveys an identical message in terms of the model's qualitative and quantitative performance. In addition, we provide the RME, RMAE, and RRMS in Table 2 to quantify the ratio of ME, MAE, and RMS to the average water depth at each location. These quantities are the same when using WSE and depth because the difference between WSE and depth does not contribute to the calculation of ME/MAE/RRMS and RME/RMAE/RRMS.

Labels and text in Figure 7 are hard to see.

[Response]: Yes, we have revised the figure to make it larger for visualization.

p 14, l 290: is it possible that the disagreement for high curvature results from using a constant roughness value for a region with varying depositional characteristics and bedforms?

[Responses]: We observed From Figure 12b,e that the non-zero roughness heights do not significantly affect the distribution of the velocity but mainly affect its maximum value. Figure 8 shows that the velocity distribution at E4/E11 is very different from the observed distribution. This disagreement is likely not caused by using a constant roughness. Instead, the poor performance at E4/E11 is likely attributed to the narrow width of the channels at E4/E11. As we use a uniform grid everywhere, mesh resolution may be too small at locations E4/E11. The small width of the channel may also affect the accuracy of the topography survey and boat-towed ADCP surveys, which eventually affects the accuracy of the comparison. Further studies on these issues may be necessary for the

future.

[Other responses]: Other comments not mentioned above but included in the pdf document are also addressed. Please refer to the color texts in the revised version for details.

The revised version of the paper revised based on both reviewers' comments can be viewed from the attached file.