Dear Editor and Reviewer,

Thank you for taking the time to review our paper again. We fully understand your point and have conducted a sensitivity study for the water management components. We hope that the revision shown below meets your expectations.

Best wishes,
Naota Hanasaki (on behalf of authors)

*The authors have clarified most of the comments and concerns. However, one major concern remains: authors should include and analysis the uncertainties and sensitivity of the water management components. While large-scale hydrologic models have been developed and validated under the uncertainties and assumptions that holds at the large-scale (e.g. 1 deg resolution), when moving to hyper-res processes should be revised (including water management) and uncertainties should be understood. Since the focus of this paper is the implementation of water management at hyper-res scale and potentials to expand it to global scales, understanding of the uncertainties (and sensitivity) of the water management components at the hyper-res scale must be address prior to publication.*

> Thank you for your insightful words. Our response is shown below in detail.
> During the revision, we noticed that the script for making figures and tables for irrigation requirement (Figure 6, Figure S3, and Table 7) contained an error. That is the irrigation efficiency for the GLB simulation was identical to LOC by mistake. We have corrected the figures, the table, and relevant text. Please be informed that the conclusions are not affected by this correction.

*Previous comments and replies:*
*Review comment: This modeling exercise demonstrated how models with localized inputs perform better than with global inputs. It would be great if the authors could provide a sensitivity analysis of the different input datasets to identify which are the most critical for improved performance. I propose a validation analysis using a leave one input out approach. In this way, besides just reporting what we know is already expected (localized models perform better), this paper has the potential to actually inform the scientific community of which of the inputs for hyperresolution modeling we should be focusing on improving. Of course, all*

*of them are important, but ranking them would greatly value future work in this field. Is that crop data? Precipitation? Water use and withdraws, etc.*

*Author Reply: Thank you for this comment. For hydrological simulation, we have newly conducted a sensitivity test. Please see our response to Dr. Luka Brocca for the results. In short, the results indicate that the usage of local meteorological observation dominantly contributed to improving the performance. Similar sensitivity simulations can be done for other components, including irrigation water requirement estimation and dam operation, but we omitted them because we can easily expect earning the same conclusions.*

*Review Reply: I appreciate the authors efforts to include the sensitivity analysis on the precipitation products, as suggested by the other reviewer. I suggest moving it to the main body of the manuscript. However, the main objective of this paper is the implementation and assessment of water management at hyper-res scales and the potentials of expanding it to global scales, isn't it? The authors should include the sensitivity analysis of the water management components (irrigation and dam operation are great examples). Although the authors expect the same conclusions, in my opinion, the uncertainties in water management components or human influence processes are far much larger and must be understood before applying and using these models to simulate water management at the local scales. Quantify these uncertainties will provide confidence (or not) that such processes are well represented at H08 at the hyper-res scale. Furthermore, it can provide the scientific community with novel insights on what needs to be done or pathways forward to implement such water management components at global scales. Many papers have already quantified the uncertainties of precipitation data quality to hyper-res modeling, this manuscript has the unique opportunity*

Thank you for this comment. We have noticed the necessity of conducting some sensitivity tests for this study before the initial submission, but we have not implemented them because of the difficulty in designing such simulations. The difference between the localized (LOC) and the default (GLB) simulations is numerous, hence the number of combination of options/parameters can be easily exploded. Based on your advice, we have returned to a deep consideration what the most informative and efficient way of sensitivity analysis could be. Partly based on the previously added sensitivity test for natural hydrological components (i.e. Appendix C), we have newly added two elaborated sensitivity tests, one for irrigation water requirement and the other for dam operation. Because the structure of model is substantially different between the localized and the default global models hence the results are not always easy to interpret. Nevertheless, the discussion should be insightful for the scientific community, we believe. Last but not least, please accept the current style that is to place independent appendices for sensitivity analysis. We have tried to incorporate the sensitivity analyses in body for many times, but finally we abandoned it because it requires placing the descriptions of simulations and results in multiple sections (i.e. at least in the methods and the results sections, respectively) which makes the manuscript long, complex, and unreadable.

*Appendix D Sensitivity simulation on irrigation requirement*

*We have decomposed the differences between GLB and LOC simulations into three factors: adoption of (A) local crop calendar, (B) local cropping practices, and (C) local hydrometeorological conditions. The combination of factors is shown in Table D1.*

*Table D1 Sensitivity simulation on irrigation requirement.*

| *Simulation* | *LOC* | *SI1* | *SI2* | *SI3* | *GLB* |
|---|---|---|---|---|---|
| *Local crop calendar* | *Yes* | *No* | *No* | *No* | *No* |
| *Local cropping practices* | *Yes* | *Yes* | *No* | *No* | *No* |
| *Local hydrometeorological conditions* | *Yes* | *Yes* | *Yes* | *No* | *No* |
| *Model* | *LOC* | *LOC* | *LOC* | *LOC* | *GLB* |

*The sensitivity simulations SI1, SI2, SI3 were identical to LOC except for the aforementioned three factors. SI1 replaced the adjusted crop calendar for rice with the originally simulated one (Section 3.2.3). SI2 is same as SI1, but additionally replaced the localized settings of land use (Section 3.2.3, Figure 4) and the irrigation efficiency parameter (Section 3.2.3) with those of GLB. Here, crop calendar, land use, and irrigation efficiency are all essential in estimating irrigation water requirement but virtually solely available from governmental reports hence hard to obtain globally. We singled out crop calendar (i.e. separated SI1 and SI2) because it would be available in the future globally since some concrete methods have been proposed by utilizing the latest satellite remote sensing techniques (e.g. Kotsuki et al. 2015). SI3 is same as SI2, but additionally replaced the local meteorological data with the global ones, and the calibrated hydrological parameters with the default ones (Sections 3.1.2 and 3.2.1).*

*The results are shown in Figure D1. In general, the NSE scores for SI1, SI2, and SI3 are deteriorated in this order, and approaching to those for GLB. The apparent exception is the influence of local cropping practices (i.e. SI2). The irrigation efficiency of GLB is set at 0.35 uniformly while LOC for 0.45-0.8. The adoption of low efficiency inflates irrigation water requirement which drastically decreases the NSE score. The effect is canceled out when the global hydrometeorological conditions were applied which suppresses the total irrigation water requirement (Section 4.1.2).*

*In contrast to the sensitivity test for discharge shown in Appendix C, the influence of adopting local meteorological condition looks marginal. This is, however, largely due to the considerable drop in the NSE score caused by other factors. This sensitivity simulation highlights the importance of reliable local cropping practices. They can alter the shape, the phase, and the magnitude of simulation timeseries, hence highly influential in the results and performances.*

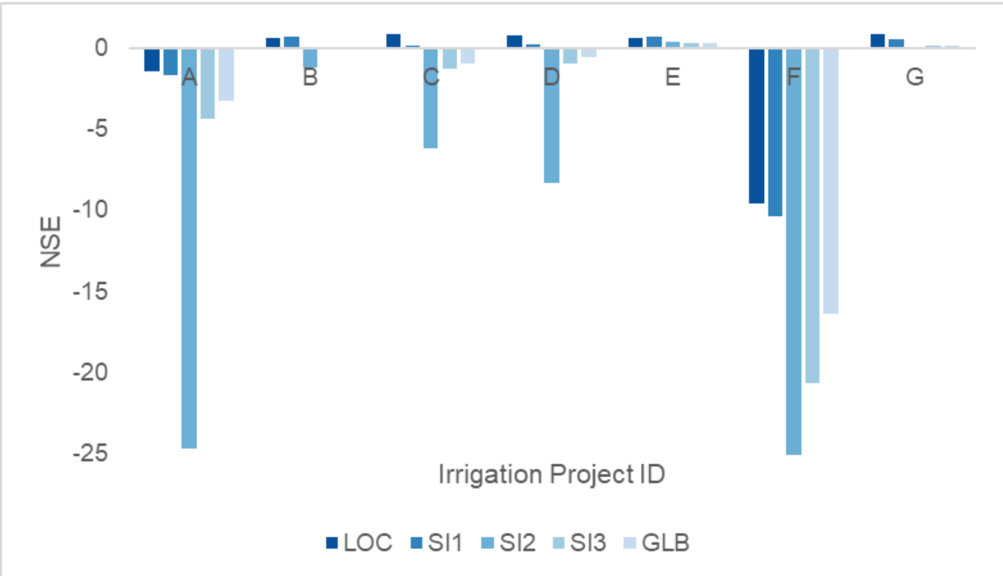


*Figure D1 The results of sensitivity analysis. See Table D1 for simulation runs. The NSE for Irrigation Project F SI2 is -98.4.*

*Appendix E Sensitivity simulation on reservoir operation*

*We have decomposed the differences between GLB and LOC simulations into three factors: adoption of (A) upper and lower storage rule curves (Figure 2a), (B) release curve (Figure 2b), and (C) local hydrometeorological conditions. The combination* *of factors is shown in Table E1.*

*Table E1 Sensitivity simulation on reservoir operation.*

| *Simulation* | *LOC* | *SD1* | *SD2* | *SD3* | *GLB* |
|---|---|---|---|---|---|
| *Storage curves* | *Yes* | *No* | *No* | *No* | *No* |
| *Release curve* | *Yes* | *Yes* | *No* | *No* | *No* |
| *Local hydrometeorological conditions* | *Yes* | *Yes* | *Yes* | *No* | *No* |
| *Model* | *LOC* | *LOC* | *LOC* | *LOC* | *GLB* |

*The sensitivity simulations SD1, SD2, SD3 were identical to LOC except for the three factors. SD1 removed the setting of upper and lower storage curves (Section 3.1.3, Figure 2). This setting allows the storage to fluctuate at any time between completely depleted and completely full. SD2 is same as SD1, but additionally removed the release curve. This setting fixes the release throughout a year at the mean annual inflow (see Hanasaki et al. 2018, 2006 for the complete formulations). Although these curves are essential in the actual reservoir operation, it is hard to obtain globally. The advances in satellite global reservoir monitoring would offer the opportunity to estimate the upper and lower storage curves for major reservoirs (e.g. Busker et al. 2019). Hence, we considered that the chance of earning the storage curves (SD1) is more likely than the release one (SD2). SD3 is same as SD2, but additionally replaced the local meteorological data with the global ones, and the calibrated hydrological parameters with the default ones (Sections 3.1.2 and 3.2.1).*

*The results for storage are shown in Figure E1. In general, the NSE scores for SD1, SD2, and SD3 are deteriorated in this order, and approaching to that for GLB. Note that SD3 and GLB do not completely agree because of other settings between LOC and GLB (e.g. water withdrawal and diversion). The removal of storage curves has the largest impact on the performance of storage simulation, which can be easily expected. The storage capacity of the dams is relatively small compared to their inflow which typically causes strong storage variations. Once storage rule curves are lost, the simulated storage showed considerable variations by reflecting errors in inflow and release.*

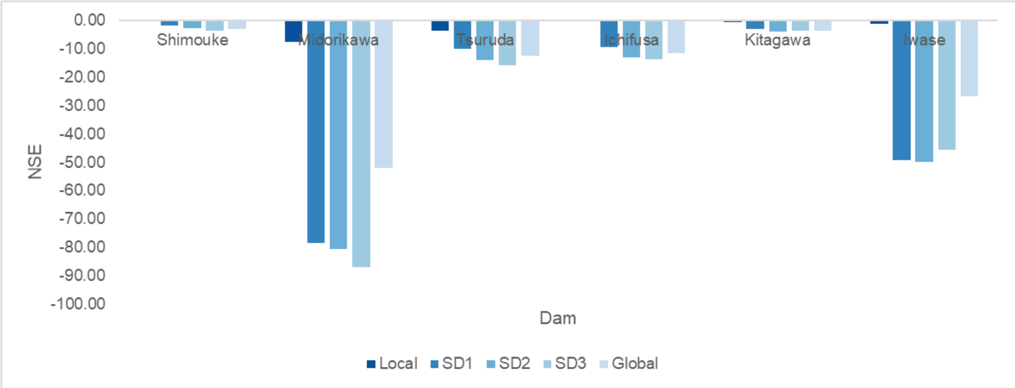


*Figure E1 Sensitivity analysis for reservoir storage. See Table E1 for simulation runs.*

*The results for release are shown in Figure E2. Again, the NSE scores for SD1, SD2, and SD3 are deteriorated in this order, and approaching to that for GLB. The impacts due to the adoption of the storage curves, the release curve, and the local hydrometeorological data are not well discernible except the case for the Midorikawa Dam and the Tsuruda Dam. For these two reservoirs, the adoption of local hydrometeorological data dominantly influenced the simulation performance.*

*Overall, specifying the key factors in the performance of reservoir operation simulation is harder than discharge (Appendix C) and irrigation requirement (Appendix D). This is likely due to the overall low reproducibility of the reality, in particular, the storage variation. Further fundamental improvement is needed in reservoir modeling to reproduce the operations in the past.*

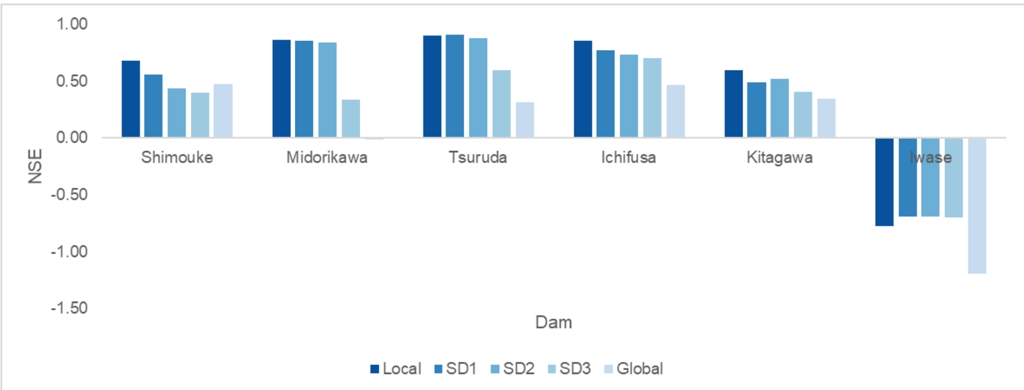


*Figure E2 Sensitivity analysis for release. See Table E1 for simulation runs.*

*References*

*Busker, T., de Roo, A., Gelati, E., Schwatke, C., Adamovic, M., Bisselink, B., Pekel,* *J. F., and Cottam, A.: A global lake and reservoir volume analysis using a surface water dataset and satellite altimetry, Hydrol. Earth Syst. Sci., 23, 669-690, 10.5194/hess-23-669-2019, 2019.*

*Hanasaki, N., Kanae, S., and Oki, T.: A reservoir operation scheme for global river routing models, Journal of Hydrology, 327, 22-41, 10.1016/j.jhydrol.2005.11.011, 2006.*

*Hanasaki, N., Yoshikawa, S., Pokhrel, Y., and Kanae, S.: A global hydrological simulation to specify the sources of water used by humans, Hydrol. Earth Syst. Sci., 22, 789-817, 10.5194/hess-22-789-2018, 2018.*

*Kotsuki, S., and Tanaka, K.: SACRA – a method for the estimation of global high-resolution crop calendars from a satellite-sensed NDVI, Hydrol. Earth Syst. Sci., 19, 4441-4461, 10.5194/hess-19-4441-2015, 2015.*