Response Letter to Reviewer #2

Title: A Hydrological Emulator for Global Applications – HE v1.0.0

Journal: Geoscientific Model Development

We would like to thank the referee for their detailed review of our manuscript and their positive feedback, constructive suggestions and criticisms. The responses to the Referee's comments are shown in blue font below. All the line numbers indicated refer to the main text of the revised manuscript (clean version without tracking changes).

Reviewers' Comments to Author:

In this study, the authors use a simple hydrological model "abcd" to emulate the behavior of more complex models (e.g. VIC). They modify the abcd model by including the baseflow index to better represent the partition of total runoff into direct runoff and baseflow. They present a lumped and a distributed version of the model, which are calibrated using the GA technique. They apply the model on global scale and compare the results against VIC simulations. Based on the results, they provide recommendations on the use of different versions of the model. Although the model used is not new and the concept of simplified emulator is an established one, however, the global-scale application of the model and its assessment over multiple basins across globe make it an interesting study. A simple and computationally efficient emulator that can work well on global scale is useful for several applications.

I think the manuscript at its current stage needs some more work. Some additional analyses need to be added. Therefore, I suggest moderate revisions for the manuscript before it is accepted in GMD. Following are my comments:

Comment 1: How reliable are the VIC simulations? Calibration of VIC can significantly change its streamflow outputs. So what type of simulations are used in this case for the comparison purpose? How were the soil and vegetation parameters calibrated/selected? All these points need to be discussed in greater details.

Response 1: The VIC runoff product (Hattermann et al., 2017; Leng et al. 2015) is used as a benchmark in this study, and its use is merely to demonstrate the capability of the hydrological emulator (HE) developed in this work to mimic complex global hydrological models (GHMs). Despite the potential bias in the VIC product, it does not affect the key findings of this work about the capability of the HE. We have added detailed descriptions about the VIC simulations in lines 186-206:

"*The VIC runoff product here is a global simulation with a daily time step and spatial resolution of 0.5 degree for the period of 1971-2010, and the VIC daily runoff is aggregated to monthly data to be consistent with the temporal scale of the "abcd" model. The VIC model settings used in this study are based on the University of Washington VIC Global applications (http://www.hydro.washington.edu/Lettenmaier/Models/VIC/Datasets/Datasets.shtml). The sub-grid variability of soil, vegetation and terrain characteristics are represented in sub-grid area-specific* *parameter classifications. Soil texture and bulk densities are derived by combining the World Inventory of Soil Emission Potentials database (Batjes, 1995) and the 5-min digital soil map of the world from the Food and Agricultural Organization (FAO, 1998). Based on the work of Cosby et al. (1984), the remaining soil properties (e.g. porosity, saturated hydraulic conductivity and unsaturated hydraulic conductivity) are derived. Vegetation type data are obtained from the global land classification of Hansen et al. (2000). Parameters including the infiltration parameter, soil layer depths and those governing the baseflow function were calibrated for major global river basins and transferred to the global domain as documented in Nijssen et al. (2001b), based on which Zhang et al. (2014) and Leng et al. (2015) conducted additional calibrations in the China domain. In this study, the VIC model was forced by WATCH climate forcing at the daily time step Weedon et al. (2011), based on the calibrated parameters from Nijssen et al. (2001b), Zhang et al. (2014) and Leng et al. (2015). The simulated runoff used in this study has recently been validated globally within the framework of the Inter-Sectoral Impact Model Intercomparison Project and shows reasonable performance compared to other hydrological models (Hattermann et al., 2017; Krysanova and Hattermann, 2017)."*

Further, we compared the VIC product to other products to corroborate its appropriateness. The comparison is presented in lines 219-229:

"The VIC runoff product also compares well to other products (see Fig. S1, S2), including the UNH/GRDC runoff product (Fekete and Vorosmarty, 2011; Fekete et al., 2002) and the global streamflow product (Dai et al., 2009). The scatterplot pattern of the VIC long-term annual runoff product vs. the streamflow product matches well with that of the UNH/GRDC runoff vs. the streamflow product (streamflow is transferred to the same unit as runoff via dividing by the basin area). Further, the correlation coefficient of the VIC and the UNH/GRDC long-term annual runoff is as high as 0.83 across the global 235 basins. This suggests the reasonability of VIC runoff product, because the UNH/GRDC runoff is calibrated with the GRDC observations. At the same time, the discrepancies between the VIC runoff products and the streamflow products (Fig. S2) may be attributed to human activities, such as reservoir regulations and upstream water withdrawals, which are not embedded in the runoff but reflected in the streamflow."

References:

Batjes, N., 1995. A homogenized soil data file for global environmental research: A subset of FAO, ISRIC and NRCS profiles (Version 1.0), ISRIC.

Food and Agriculture Organization (FAO), 1998. Digital soil map of the world and derived soil properties. Land and 410 Water Digital Media Series 1, CD-ROM.

Hansen, M., DeFries, R., Townshend, J.R., Sohlberg, R., 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. Int. J. Remote Sens., 21(6-7): 1331-1364.

Leng, G., Tang, Q., Rayburg, S., 2015. Climate change impacts on meteorological, agricultural and hydrological droughts in China. Global Planet. Change, 126: 23-34.

Cosby, B.J., Hornberger, G.M., Clapp, R.B. and Ginn, T., 1984. A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. Water resources research, 20(6), pp.682-690.

Krysanova, V., & Hattermann, F. F. (2017). Intercomparison of climate change impacts in 12 large river basins: overview of methods and summary of results. Climatic Change, 141(3), 363-379.

Hattermann, F. F. et al. (2017). Cross-scale intercomparison of climate change impacts simulated by regional and global hydrological models in eleven large river basins. Climatic Change, 141(3), 561-576.

Nijssen, B.N., G.M. O'Donnell, D.P. Lettenmaier and E.F. Wood, 2001: Predicting the discharge of global rivers, Journal of Climate, 14(15), 3307-3323

Weedon, G. et al., 2011. Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century. J. Hydrometeorol., 12(5): 823-848.

Zhang, X. (2014). A long-term land surface hydrologic fluxes and states dataset for China. Journal of Hydrometeorology, 15(5), 2067-2084.

Comment 2: How did VIC perform in the extreme climate regions, for example, in snow dominated catchments? This issue needs to be addressed properly. Maybe you can explore the following cases: Case-1: If both emulator (E) and model (M) are matching the observations (O) well then that's great. There could be some sub-cases for this case: (i) Both emulator and model match the observations well but from different directions (M – O – E). For example, they might have opposite (positive/negative) bias errors but the absolute values of the errors could be close. (ii) The model is matching the observations well and the emulator is matching the model well, all in one direction (O – M – E). (iii) The emulator is matching well both the model and the observations, but in different directions (M – E – O). Case-2: If none of them are matching the observations well but their own outputs match each other, then too, I think an emulator is serving its purpose in a way (although not quite useful). Case-3: If the emulator is matching the observations well but the model isn't then that's an interesting finding. Case-4: If the model is matching the observations well but the emulator isn't then there is a problem. Therefore, this needs to be explored in greater depth.

Response 2: We thank the referee for the detailed comments regarding the comparison between the VIC model and the hydrological emulator (HE). The essential point of this work is to deliver an open-source and easy-to-use hydrological emulator that can be used for emulating global hydrological models (GHMs) of interest. VIC is used as an example GHM in this study to demonstrate the capability of the HE to emulate complex and computationally expensive GHMs (see also Response 4 in the Response letter to Reviewer #1). Exploring the sources of differences between the performance of the VIC and the HE is outside the focus of this work, and it would be incorporated in our future work.

Comment 3: At seasonal time scales, the model performance is expected to be better. It would be crucial to also check the results on daily time scale. Maybe you can produce a set of time series plots, scatter plots, and spatial contour plots for daily level, as done for the seasonal case.

Response 3: We agree with the reviewer on the better performance of monthly time scale than that of daily, however, due to the needs of high computational efficiency for the hydrological emulator (HE), it is simulated at monthly time step and a daily time series comparison is not even feasible in this case.

Comment 4: Figure 3: Any idea why there are those biases in the lower streamflow values? Is there any location-specific pattern of these biases?

Response 4: From the uncertainty analysis we added in Section 3.5 (see also Figure 7), it shows basins like Congo and La Plata are not as robust as other basins to changes in parameters – slight changes in parameters may lead to large changes in runoff estimates. Then the uncertainty in the calibrated parameters for the two basins may lead to large bias in the simulated runoff. We have added discussions in lines 415-421:

"*The uncertainty analysis indicates that most basins are robust to changes in parameters, other than the Tocantins, Congo and La Plata (Fig. 7). In other words, for basins Congo and La Plata, slight changes in parameters may lead to large changes in runoff estimates. Then the uncertainty in the calibrated parameters for the two basins may lead to large bias in the simulated runoff, which may more or less explain why modelled runoff for the two basins tend to have higher biases than other basins (Fig. 4)*."

Comment 5: Line 113: Which one's the other parameter you adopt the value of?

Response 5: We adapt two snow-related parameters from literature, and make the other one – snow melt coefficient – tunable during the calibration process, it has been clarified in lines 139-143:

"*in order to enhance the model efficiency with as least necessary parameters as possible, instead of involving three tunable snow-related parameters in the calibration process, we set the values for two of the parameters (i.e., temperature threshold above or below which all precipitation falls as rainfall or snow) from literature (Wen et al., 2013) and only keep one tunable parameter m – snow melt coefficient* ($0 < m < 1$)."

References:

Wen, L., Nagabhatla, N., Lü, S., Wang, S.-Y., 2013. Impact of rain snow threshold temperature on snow depth simulation in land surface and regional atmospheric models. Adv. Atmos. Sci., 30(5): 1449-1460.

Comment 6: Line 200: Did you try different weights on the two objectives?

Response 6: No, we use the same weights for the two objectives as we believe the two objectives are equally important.

Comment 7: Line 290: In order to do a fair comparison, VIC and the two versions of the models should be run on the same computer, preferably with good configuration.

Response 7: Yes, we have clarified this in lines 365-367:

"*Note that all of the simulations here are conducted on the Pacific Northwest National Laboratory (PNNL)'s Institutional Computing (PIC) Constance cluster using 1 core (Intel Xeon 2.3 GHz CPU) with the same configuration*."

Comment 8: Figure S1: I am not sure if you can say that all of them are comparing well. The discrepancies/mismatches should be clearly discussed in the manuscript. You are only showing the correlations here. What about the bias error?

Response 8: We thank the referee for the concern. Figure S1 is to illustrate the relationship of VIC and UNH/GRDC runoff product with streamflow measurements at gauge stations, and the similar scatter patterns between the upper and lower panel indicates the similarity of the two runoff products. This analysis is to reveal the appropriateness of the VIC runoff product as a benchmark product in this work. The discrepancies between runoff products and streamflow measurements are induced from the ignorance of river routing, reservoir regulations and upstream water withdrawals in the simulated runoff products. This has been recognized in the main text (lines 226-229):

"*At the same time, the discrepancies between the VIC runoff products and the streamflow products (Fig. S2) may be attributed to human activities, such as reservoir regulations and upstream water withdrawals, which are not embedded in the runoff but reflected in the streamflow.*"

However, exploring the sources and magnitudes of the discrepancies among them is outside the focus of this study.

Comment 9: X-axis marks are missing for the first two subplots of Figure S1. Use same axis for the scatter plots in Figure 2.

Response 9: The X-axis marks for Figure S1 have been fixed. For the previous Figure 2 (currently Figure 3 in the revised manuscript), we use different axis for total runoff, direct runoff and baseflow is because they have different magnitude, and this may make the figure and scatter points more discernable.

Comment 10: My comments about the manuscript: Writing: The manuscript is very well written. I don't have any suggestions on this part. Figures: Figures look good. Increase the legend in Figure 3. Tables: Table 2 can go to the supplementary materials.

Response 10: We have increased the legend in the previous Figure 3 (currently Figure 4 in the revised manuscript) and moved Table 2 to the supplementary materials as suggested.