General Comments

Colleoni el al. develop a differentiable hydrological modeling framework, smash v1.0. They demonstrated good performance of the smash after conducting spatially distributed calibration and regionalization by learning the relationship between physical descriptors and parameters, with median evaluation metric KGE > 0.6. They tested their model framework in two configurations: (1) CONUS at 3km spatial resolution and (2) France watershed at 500m spatial resolution. Both applications show significant improvement in the model performance after using the regionalized parameter values. This study represents a significant contribution to hydrological modeling as it addressed a critical gap in hydrological model calibration, such as the traditional calibration method requiring tens of thousands of simulations with perturbed parameter values, which are computationally expensive. Please find my comments in the following. I recommend a revision before publication.

Major Comments

The description of the hydrological model is not clear. Could the authors provide a brief description of what processes are included in the hydrological model? If there is no space in the main text, the authors should consider adding it to the supplementary text.

What is the assumption of the routing model in smash? There exist different types of routing models. For example, large-scale 1D river routing models are commonly applied at relatively coarse spatial resolution with the assumption that each grid cell contains a representative channel. A 2-dimensional routing model that solves kinematic, diffusion, or dynamic wave routing at high spatial resolution, such as less than 1km, does not need such an assumption. However, it may not be appropriate to apply such 2-dimensional routing model at coarse spatial resolution. Which type of routing model was implanted in smash? It has been applied at 3km resolution for CONUS and 500m resolution for France application. It would be helpful if the authors could elaborate on the routing method.

Could the author report the computation time for the regionalization simulations vs. an individual forward run? This can help the readers to better understand the benefits of differentiable models. For example, in typical hydrological model calibrations, one has to run the model many times, e.g., 1000, with perturbed parameters, to obtain the optimal parameters. I think the differentiable model requires much fewer simulations to get the optimal parameter values, making its application to large scales more feasible.

There is no description of the calibrated parameters and their possible range. In addition, how are the parameters perturbed to solve the adjoint of the forward model in Figure 2 is

not clear. Are the parameters perturbed at watershed level or are they perturbed at grid level within each watershed?

Specific Comments

Line 61: Please provide the full name of MPR.

Line 60 – Line 65: There are more large sample studies in terms of calibration of spatially distributed models using efficient statistical approaches. For example,

Yang, Y., Pan, M., Beck, H. E., Fisher, C. K., Beighley, R. E., Kao, S. C., Hong, Y., and Wood, E. F.: In quest of calibration density and consistency in hydrologic modeling: distributed parameter calibration against streamflow characteristics, Water. Resour Res., 55, 7784–7803, 2019.

Xu, D., Bisht, G., Sargsyan, K., Liao, C., and Leung, L. R.: Using a surrogate-assisted Bayesian framework to calibrate the runoff-generation scheme in the Energy Exascale Earth System Model (E3SM) v1, Geosci. Model Dev., 15, 5021–5043, https://doi.org/10.5194/gmd-15-5021-2022, 2022.

Hirpa, F. A., Salamon, P., Beck, H. E., Lorini, V., Alfieri, L., Zsoter, E., and Dadson, S. J.: Calibration of the Global Flood Awareness System (GloFAS) using daily streamflow data, J. Hydrol., 566, 595–606, https://doi.org/10.1016/j.jhydrol.2018.09.052, 2018.

Line 98: What type of mesh does smash support? Structured or unstructured?

Line 114 – Line 129: Are the snow, hydrological, and routing operators process-based functions or data-driven operators?

Line 159: Is equal weight used for w_g? Please specify how w_g was estimated for each gauge. In addition, how j_reg was estimated?

Line 215: Which year was chosen?

Line 221: Please specify how the computation times were evaluated. According to my experience, it is unrealistic for a model to take only 0.1 seconds to run 365 time steps for 3846 grid cells. This relates to my previous comments: are the snow, hydrological, and routing operators process-based models?

Line 226: The better scaling in the inverse run is likely due to it being more computationally expensive than the forward run since the inverse run requires more computations.

Line 229: Why the routing scheme not be parallelized? There are a lot of global river routing models, and 2-dimensional hydrodynamic models are well parallelized.

Figure 5: Please add longitude and latitude to the figure. Also, consider adding a colorbar to show what is plotted.

Line 268: Which 10 years are used for “warm-up”?

Line 270: What parameters are calibrated?

Line 275 – Line 280: Can you also report the default forward model performance?

Line 279: Please show the calibrated parameter values. Also, for the spatially uniform parameters, are they spatially uniform for the whole CONUS? Or are they only spatially uniform for each basin?

Line 279: I found the performance between spatially uniform and spatially distributed simulations is very close.

Line 341: For the non-inertial shallow water models, do you mean diffusion wave routing?

Line 342- 344: I wonder if the authors can give some comments on how to implement the differentiable model to more complex land surface models or hydrological models?

Line 368: Why change the evaluation metric from KGE to NSE in the application of Frace?

Line 403: I agree with the authors that the smash framework represents an advancement in data assimilation in the hydrological model. However, I disagree that the smash framework represents an advancement in the hydrological model. Based on the method section, smash implements simplified hydrological processes.