

Geosci. Model Dev. Discuss., referee comment RC1  
<https://doi.org/10.5194/gmd-2021-401-RC1>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Comment on gmd-2021-401

Anonymous Referee #1

---

Referee comment on "Using a surrogate-assisted Bayesian framework to calibrate the runoff-generation scheme in the Energy Exascale Earth System Model (E3SM) v1" by Donghui Xu et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2021-401-RC1>, 2022

---

This work presents a surrogate-assisted framework for calibrating runoff relevant parameters in global-scale Earth System Models (ESMs). The large computation burden arisen from repeated simulations in calibration is alleviated by building fast-to-run PCE-based surrogate models of ESMs. It is concluded that the calibrated model obtains an improved performance compared to the one with default parameter values. In summary, the manuscript is generally well-written and may be eventually accepted after addressing the following comments:

-The title should be revised. In my opinion, uncertain quantification is different from calibration. How can one use a UQ framework to calibrate models? How about "Using a surrogate-assisted Bayesian framework to ..."?

-In table 1, the prior for  $q_{\text{drai,max}}$  is  $U(1e-6, 1e-1)$ . Why not use a logarithmic transformation for it? Otherwise, much more prior samples will be drawn from, e.g.,  $(1e-2, 1e-1)$ .

-Line 193 and Eq. (17), the authors should clearly present how they determine the values of sigma.

-In section 3.3, which criterion (e.g., the Gelman-Rubin R statistic [Gelman et al., 1995]) is used here to check the convergence of MCMC sampling? From Fig. 6 it can be seen that the posterior ranges are still relatively large, which gives the feeling that the MCMC chain has not totally converged.

-In section 5.2, the poor performance of PCE surrogate models in arid regions probably because of PCE's inability to approximate highly nonlinear functions (a well-known limitation of PCE) or/and the low signal-to-noise ratio in these regions. The authors should elaborate these to provide more informative results to readers. An alternative surrogate method for approximating highly nonlinear function is the deep neural networks.

-In Figs. 10a and 11, the simulated runoff time series with default parameter values could be even closer to the reference GRUN time series than the calibrated ones, giving the feeling that the calibration is not that satisfactory. Please further explain.

-Line 533, 'are estimated', Line 534, 'are run'?