Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2019-10-RC1, 2019

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Spatially distributed sensitivity of simulated global groundwater heads and flows to hydraulic conductivity, groundwater recharge and surface water body parameterization" *by* Robert Reinecke et al.

Anonymous Referee #1

Received and published: 23 April 2019

The manuscript presents the results of a Sensitivity Analysis (SA) applied to a global groundwater model (G^3M). As an expert of SA more than groundwater modelling, I will focus my comments on the SA aspects of the work, and leave it to other reviewers to comment on the groundwater modelling side. From such perspective, I think the manuscript contributes an interesting demonstration of the usefulness of SA for model testing and evaluation, as well as some methodological advances on how to make spatially-distributed global models tractable by SA. I thus think the manuscript should be considered for publication in HESS, although I would suggest that a major revision is

Printer-friendly version

Discussion paper

needed in order to first improve its clarity. Many key points are confusing in the current version and grammar and mathematical notation would benefit from an overall revision.

MAJOR POINTS

[1] The working principles of the global groundwater model G^{3}M should be explained in more details, otherwise I find it very difficult to fully understand the SA set-up and results.
For example, on P. 2 L. 32, it would be good to expand a bit on the connection between G^{3}M and WaterGAP (which variables are exchanged from one model to the other and why this is important to improve WaterGAP predictions). In Section 2.1 I would insert a schematic figure of the links between the key variables of the G^{3}M model (h, Q_{sub}, B_{swb}, etc.) - it would makes it easier to follow Eq. (1),(2), etc. and understand the role of the input data and parameters subject to the sensitivity analysis.
Also, the mathematical descriptions in Sec. 2.1.3 to 2.1.5 is a bit messy and possibly incomplete. Variable h in Eq. (1) is undefined and there is no further equation or description of how it is calculated. Variable Q on L. 17 does not appear in Eq. (1) (unless by Q the authors actually mean Q_{swb}). Many sentences in Sec. 2.1.3 are rather unclear - see more specific comments in the last section of my review.

[2] The ultimate goal of the SA should be more clearly stated.
The Introduction ends with the statement (P. 3 L. 10): "The derived global maps show, for the first time, the sensitivity and parameter interactions of simulated hydraulic head and groundwater-surface water flows in the simulated steady-state global groundwater system to variations in these uncertain inputs." Still, this does not clarify what are these maps useful for. Will they serve to set priorities for improvement of "input data"? Or to decide which parameters should be calibrated and which can be set to default values? Or maybe as a "sanity check" test, i.e. to prove that the dominant parameters are as expected for each particular output in each particular region? This needs to be clarified. At present, the manuscript Introduction only states which "sensitivities of

the model are explored" (P. 3 L. 3) but it does not say what research questions this exploration is meant to answer.

[3] One of the key ideas that make SA applicable to such a spatially complex model, is the use of Global Hydrological Response Units (i.e. groups of hydrologically similar cells to which the same parameter perturbations can be applied) as described in Sec. 2.2.3. This is a simple but effective approach that could be of interest to a wide audience of modellers who deals with large-scale models and may confront similar problems when applying MC simulations or SA. Hence it should be mentioned in the Introduction or it may go unnoticed by interested readers. Also, it would probably be good to recognise that similar approaches have been used before, for example (if I get this right) by Hartman et al. 2015 (A large-scale simulation model to assess karstic groundwater recharge over Europe and the Mediterranean, GMD).

[4] Reliability and meaningfulness of the SA results.
The authors say that many model runs needed to be discarded from the SA because the simulation outputs were unreasonable, and that the sensitivity indices for many cells were not reliable because estimation errors were too large. I think these two issues are very important as they may undermine the usefulness of the entire analysis. As such, they need to be explained and discussed more clearly. Specifically:
P. 9 L. 4: "A converged simulation does not necessarily constitute a valid result for all computed cells. Numeric difficulties based on the model configuration (due to the selected parameter multipliers) may lead to cells with calculated h that are unreasonable." This needs further clarification. First, how is an unreasonable value of h defined? Second, what do you do with simulation runs that provide unreasonable h values? Do you retain them in the SA? If so, sensitivity estimates may be affected by simulation results that you consider unreasonable. Is this acceptable?
Table 2 (last column) and Table 3 (footnote): majority of the cells do not provide

”reliable” sensitivity estimates (i.e. CIs of sensitivity estimates are overlapping). Again, the criterion by which CIs are deemed overlapping and hence sensitivity estimates are considered unreliable needs to be explained more clearly. Appendix 1 covers the topic but is very concise and not very clear. The sentence on P. 22 L. 1 seems to suggest that the ’reliability criterion’ is based on the fact that the CI be smaller than 15% of the calculated $\mu*$ of the first rank - if so, where is the 15% threshold coming from? And is this criterion really related with the fact that CIs overlap? I suppose one could have CIs of P2 and P3 that overlap even if each of them is smaller than 15% of P1.
Figure A1 does not help clarifying the matter. The 15% threshold does not appear in there, and many other things are confusing. For example, in P2 why the text ”CI 95%” only refers to half of the CI (instead of all the CI)? In P3, why σ ’comes out’ of an arrow starting from the CI of $\mu*$? Please clarify

Last, are the SA results really useful if so many cells provide unreliable results? This is difficult for me to say given that, as pointed out in [2] above, the ultimate goal of the analysis is not totally clear. For example, if the ultimate goal was to identify the 2-3 key controls of the model behaviour in each different region, then an overlap between the CIs of the first and second ranked parameters would not be too much of the problem: the key message of which are the 2 most important parameters would still emerge clearly from the SA.
So I am not suggesting that the SA results presented here are not useful - I just think the manuscript should clarify better what can and what cannot be inferred from such results, and what the implications are for the future improvement or use of the model. At present, it sounds a bit like the authors produced SA maps and draw some conclusions, then checked the CIs and realised most of the regions in those maps are actually unreliable. This is unconvincing. I would approach the issue from another angle: given the questions you wanted to answer, is still possible to answer them despite the overlapping CIs?

MINOR POINTS

P. 2: ”Global-scale hydrological models have recently moved to include these processes by implementing a gradient-based groundwater model approach (de Graaf et al., 2015; Reinecke et al., 2018).”
It would be good to be a bit more specific here. How many gradient-based groundwater models are currently available at the global scale? One (to which both cited papers contributed) or two (one developed by de Graff et al 2015 and a different one by Reinecke et al 2018)? And if the Reinecke model cited here is (as I guess) the G^3M model that is then analysed in this paper (as introduced on L. 26), then it would be good to clarify the point. If G^3M is the only (or one of the two) global model currently able to simulate global groundwater heads and flows, then the relevance of this manuscript is higher than the manuscript currently communicate.

Throughout the manuscript there is some confusion around the difference between ”input data” and ”parameters”. I understand that ”input data” essentially refer to the GLHYMPS dataset, of which two versions (1.0 and 2.0) are tested and compared (point (1) on P. 3 L.3). However, such input data are used to estimate the hydraulic conductivity K, which is also one of the parameters that are later made randomly vary in the Monte Carlo experiments. Therefore there is some overlap between the two concepts (input data = parameter in the case of K, if I get this right?). This is difficult to grasp if the authors do not clarify the point. Again, having a schematic of the key relationships between variables would probably help here.

P. 6 L. 22: ”Based on previous experiments...” I think it would be good to add some more information about the selection of the 8 parameters to be subject to SA. Mentioning ”previous experiments” is too vague. How many other parameters are there in the model that are held fixed? What did these previous experiments show

that made you choose those 8 in particular? Also, if the SA is conducted by varying the parameter multipliers, then the choice of the baseline parameter values that are perturbed by the multipliers may be critical. How were this baseline values estimated?

P. 2 L. 30: "the Morris method does not provide the variance decomposition"
The sentence suggests that not providing the variance decomposition is a problem per se, but I do not understand why that should be the case. Many global SA methods (e.g. regional SA, density-based methods, etc.) do not provide variance decomposition as they define output sensitivity based on other principles than "contribution to output variance" - yet they can perfectly fit their purpose. So I think this sentence is misleading and should be revised or removed.
P. 6 L. 10-15: I suppose you must have used the (most common and most sensible) implementation of the Morris method where sensitivity $*$ is given by the mean absolute value of the Elementary Effects. Still neither Eq. (6) or the text mention using absolute values. Please clarify.

P. 15: "The number of clusters was determined based on the feasible number of model evaluations"
P. 8 L. 2: "With seven parameters per GHRU plus the ocean boundary, 10,000 base points were sampled in total (Campolongo et al., 2007) and optimized using Ruano et al. (2012). The experiment resulted in 1848 simulations"
This is very confusing. How is the number of clusters (n) related to the total number of model evaluations (N)? I would think: $N = r \times (n \times 8 + 1)$, where 8 is the number of parameters in Table 1 (hence $n \times 6$ gives the total number of multipliers to be sampled in the application of the Morris method) and r is the number of Elementary Effects for the Morris method. However, as the authors use n=6 clusters and N= 1848 model evaluations, I cannot figure out a feasible value for r! This needs to be explained more clearly. Also, I do not understand what the term "base points" refer to? It clearly

cannot be the number r of points from with OAT perturbations are applied, so what is it?

TYPOS AND GRAMMAR

P. 1 L. 4: ”the reliability of model outputs is limited by a lack of data as well as model assumptions required due to the necessarily coarse spatial resolution.” Something not right with this sentence, maybe better: ”the reliability of model outputs is limited by lack of data and by uncertain model assumptions that are necessary due to the coarse spatial resolution.”
P. 1 L. 14: ”numerical difficulties”. Unclear. Is it a problem of numerical instability? Or what else? ”difficulties” is not a technical term.
P. 2 L. 26: ”an application of .... with the Global ...” should be ”an application of ... to the Global ...”
P. 2 L. 29: ”sensitivity parameters” should be ”sensitive parameters”
P. 3 L. 15 and L. 23: ”to be coupled with WGHM”.... ”computed by WaterGAP 2.2c”. I suppose WGHM and WaterGAP 2.2c are the same model, if so please use one name, otherwise explain the difference.
P. 3 L. 25: ”to include it into a stead-state model represents a natural equilibrium” Something missing/wrong in this sentence, please revise.
P. 3 L. 26: ”shown in Fig. 2(a)”. Figure 2 is cited before Figure 1, which is odd. Maybe change the Figure order?
P. 4 L. 16-17: ”The in- and outflows are described similar to MODFLOW as flows from the cell Q, thus losing and gaining surface water bodies (lakes, wetlands and rivers) are described as” Very unclear. ”from the cell Q” seems to suggest that ”Q” is the index of the cell, which from the subsequent equation clearly is not. Also, it is unclear here if the spatial unit of interest is a grid cell (as in the text) or a surface water body (as in the equation) - if there is a difference between the two? Last, the subject of the sentence changes from ”flows” to ”surface water bodies” but the subsequent equation defines

(again) flows (Q_{swb}) not water bodies. So maybe rephrase as "gains and losses from surface water bodies (lakes, wetlands and rivers) are described as" (?)
P. 4 L. 21: "For lakes (including reservoirs) C_{lak} and wetlands C_{wet}, C_{swb} is estimated ...". Unclear what is the difference between one variable and another. Maybe the authors mean: " For lakes (including reservoirs) and wetlands, the conductances C_{lak} and C_{wet} are estimated ..."
P. 4 L. 15-30: "To account for that we assume ... the river is the sink for all the inflow to the grid cell ... that is not transported ...". Very long, convoluted sentences that can be hardly followed - please revised.
P. 5 L. 22: "the sensitivity of .... caused by variability" should be: "the sensitivity of ... to the variability"
P. 5 L. 23: "The results are then compared to the variability in parameters of the Monte Carlo experiments." Unclear. I suppose what can be compared is the variations of outputs, not the variability of inputs. Hence the sentence should sound something like: "The results are then compared to the effects of parameter variability, as quantified by the Monte Carlo experiments."
P. 6 L. 5: "(Sect. 2)". Circular reference: this is actually Section 2!
P. 6 L. 10: "model evaluation responses". Unnecessarily confusing. I would just say: "model executions". P. 6 L. 19: "To achieve that, $\mu*$ and σ_i are presented as ranked parameters". This is not understandable. What do the authors mean by "ranked parameters"? Please clarify.
P. 6 L. 22: "we identified eight uncertain model parameters presented as multipliers in Table 1". Again, unclear. I guess this means that eight uncertain parameters were selected for the SA, and the analysis was performed by multiplying each parameter by a random multiplier (sampled from ranges specified in Table 1). If so, please clarify.
P. 7 L. 6: "(Sect. 2.1.4)" Possibly wrong reference? E is defined in Sec. 2.1.3.
P. 10 L. 3: "3.2 Monte Carlo experiments". I find it a bit confusing that the results reported here are sometimes referred to as "Monte Carlo experiments", some other times as "Morris method". I would choose one term (possibly the latter, as it is more precise)

and stick to it.
P. 10 L. 11: ”the groundwater equation”: which equation? please clarify. P. 11 L. 20: ”... are determined by h in relation to E_{swb}.” Vague. Explain what the relationship is.
P. 11 L. 21: ”the the surface”: remove one ”the”
P. 11 L. 22: ”independent of”: should be ”independently of”
P. 17 L. 3: ”as most sensitive” should be ”is most sensitive”
P. 18 L. 15: ”grid cells with either zero sensitivity value” This is strange. If the sensitivity estimate is exactly zero, that should suggest there must be some calculation error.
P. 18 L. 22: ”the evaluated model is a numerical model and thus behaves differently for different parameterizations” This sentence does not mean much. Every model subject to SA is a numerical model and all models behave differently if one changes the parameter values. Please clarify what you mean to say here.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-10, 2019.

Printer-friendly version

Discussion paper