

Geosci. Model Dev. Discuss., community comment CC3  
<https://doi.org/10.5194/gmd-2021-425-CC3>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Reply on RC1

Emilie Rouzies

---

Community comment on "How to perform global sensitivity analysis of a catchment-scale, distributed pesticide transfer model? Application to the PESHMELBA model" by Emilie Rouzies et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2021-425-CC3>, 2022

---

Dear Reviewer 1,

Thank you very much for the careful review and edits to the initial submission. Although the revision period is not over, we propose to start answering your detailed comments so as to keep the discussion as interactive as possible. All your comments and questions have been copied hereafter in *italic*. The revised manuscript will be provided in a second time so as to accommodate comments from all reviewers.

*Detailed comments*

**L21 p1 'simple enough to ensure flexibility': More explanation is needed here. This is vague and I am not sure what is meant by flexibility.** Here we mean that models used to support decision-making should be designed so that users could easily modify the code to integrate new physical processes and/or adapt the existing ones. "Flexibility" then refers to the structure of these tools that should be ideally simple enough to enable such evolutions. The sentence has been clarified in the manuscript.

**L30-31 p2 'catchment-scale model [...] afforded': Specify that this is spatially distributed models.** Yes, added as suggested.

**L61-73 p3: Also note the recent study of Smith et al. (2021).** Yes, added as suggested.

**Section 2.1: a presentation of the model parameters is missing. How many uncertain parameters that needs to be estimated are there? What are the different categories of parameters (e.g. soil, pesticide, vegetation etc., as I can read in Table 2). Parameters are only introduced much later in Section 2.5 (Table 2), which makes it difficult to follow Section 2.2 that describes the selection of the parameter values. The reference to Table 2 in the caption of Table 1 does not flow well.** The model setup section has been deeply modified so as to introduce a presentation of model parameters much earlier than in the original manuscript. The Table from Section 2.5 has been modified and comes earlier to introduce all model parameters and associated categories. A sentence has also been added in the text to specify the number of uncertain parameters involved in the sensitivity analysis.

**Section 2.2: Why performing the experiment on a virtual catchment and not a 'real' one ?** As mentioned in the text, the final targetted catchment for this study is the real La Morcille catchment. Figure 1 (left) depicts PESHMELBA meshing at this scale showing that such application results in a high number of landscape elements (>500). Conducting experiments at the full catchment scale would have drastically increased the computational cost of the analysis while turning difficult the interpretation of sensitivity analysis results considering that no such experiment has been conducted before. We then perform the experiment on a simplified case as a first try to get a clearer and simpler interpretation of the results both regarding methodological and spatial aspects.

**I understand that the simulation experiment considers the application of the fungicide at the beginning of the winter period. Is this realistic?** As pointed out, considering an application of fungicide at the beginning of the winter period is not very realistic. Actually, we suggest to remove all mention to "winter" period as the focus of this study is mainly methodological, based on a virtual case and realistic forcings. The chosen setup primarily aims at identifying influential factors on different physical processes integrated in PESHMELBA with a strong focus on lateral transfers of water and pesticides. We have then favoured a scenario with strong rain events since they result in both surface runoff and lateral saturated transfers in subsurface. The results of this study then provide general guidelines about the model behaviour but they should be further complemented with applications on each particular agropedoclimatic context of interest.

**Why performing the experiments over a 3-month winter period? This is a very short time period.** In this case study, PESHMELBA time step is 1h on dry periods and 30 minutes during or after rainfall events resulting in a high computational cost for a three-month simulation (2h per simulation on the cluster used to run the simulations). A longer time period was then not affordable for this first experiment. In addition, we chose a period characterized by high cumulative rainfall volume to make sure that the different physical processes simulated in PESHMELBA would activate during the simulation (we were mainly concerned with activation of surface runoff and lateral saturated exchanges). This way, the performed sensitivity can also be used as a consistency check on the model structure itself allowing to check different physical processes simulation. However, we remain aware that results from GSA highly depend on climatic conditions as precised in the conclusion of the manuscript. As mentioned, further researches may focus on other contrasted time periods to draw robust conclusions.

**A justification for the soil moisture initial condition (hydrostatic equilibrium L157) is missing.** An hydrostatic equilibrium has been chosen so as to provide the model with initial conditions as "neutral" as possible. We wanted the variables of interest to fully represent the dynamic of the catchment and not to include any non-physical warm-up period. To do so, another approach consists in running a warm-up simulation on a longer period but it would imply a high computational cost that could not have been afforded in this case.

**Section 2.3-2.4: I think that section 2.3 provides too many technical details that are not necessary to understand the methodology and analyses presented in the paper. The authors recognize themselves that this section could be skipped L183-184. My suggestion is to report only the main equations used to compute the sensitivity indices, while details on the derivation of these equations (that were taken from previous papers and that are therefore not really a contribution of this paper, if I understand correctly) can be moved in the supplements/appendix. I am mostly referring to the description of the Sobol' and HSIC methods, while I think that the description of the random forest method in Section 2.4 reads very well. The main equations and references of Section 2.3 can be combined with the summary of the GSA methods provided in Section 2.5, to provide the reader only with the information that are needed to**

**understand the methodology and the analyses, while avoiding unnecessary repetitions between Sections 2.4 and 2.5. In addition, I think that an overview of the methodology (why do you need to use the GSA methods?) is needed before introducing the specific GSA methods.** The section on method description has been fully reviewed as suggested. Section 2.3 and 2.5 have been merged and only the main equations relative to each methods now remain together with more practical interpretation of calculated indices. We have also added a justification for method comparison and an overview on the full methodology at the beginning of the section.

**Equation (17): The sensitivity index for a given input is the average of the first order indices estimated for the different model outputs, weighted by the outputs variance, am I correct? This paper aims to help applying these methods, therefore I think that interpreting the equations in simple (intuitive) terms, would improve readability and clarity. It is very nice to have the formal mathematical proof for the equation, but the proof does not have any practical implications and could be moved into the supplements/appendix (this is an example of how this section could be simplified, see my previous comment).** Indeed, aggregated sensitivity indices correspond to an average of Sobol' indices on each landscape unit weighted by local output variances. As suggested, the proof has been removed from the main text while a sentence has been added to qualitatively describe the formule for such indices.

**Only first order indices can be estimated for multidimensional outputs? In Figure 10 I see that also the total indices are calculated at the landscape scale. How was this done ?** The formulation from previous Eq. (17) can actually be applied to Sobol' indices from any order. We have clarified the text and have explicitly mentioned the calculation of first and total order indices in Section 2.6.

**Equation (24): If  $X_i$  and  $Y$  are not independent, the value of the dependence measure estimated for a given bootstrap resample (that is in a way obtained by randomly attributing values of  $Y$  to each value of  $X_i$ , if I understand correctly) will tend to be larger than the dependence measure estimated for the original non-bootstrapped sample? Why?** First, yes a bootstrap resample is indeed obtained by randomly attributing values of  $Y$  to each value of  $X_i$ . However, if  $X_i$  and  $Y$  are not independent, the HSIC value for such a bootstrap resample will be lower than the HSIC value for the original sample because the random resampling step breaks the existing dependence relationship. The p-value then will tend to zero.

**Section 2.4: The GSA workflow is not well explained in the text. In particular, the references to the sample sizes used are confusing. I read that 1000 points are used for PCE (L382), 4000 points for HSIC (L391), that 1000 points were derived from the 4000 points used for HSIC and that 1000 points are used for RF. It is only by looking at Figure 5 that I finally understood that these numbers are linked: 4000 points initially used for HSIC and then based on HSIC screening 1000 points are selected for all subsequent analyses. However, I am still a bit unsure why it is written L374 that 'a variance decomposition method was first used', isn't it HSIC?** First, a screening test is performed based on the statistical using HSIC from a 4,000-point LHS. Once influential parameters have been identified, a new 1,000-point LHS is generated with only influential parameters. On this new sample, Sobol, HSIC and RF indices are compared for ranking. This description has been explicitly integrated at the beginning of Section 2.4, when merging Section 2.3 and 2.5. with clearer references to sample sizes.

**L416 p17 '100 replications were used': Why using 100 replications for bootstrapping? 1000 bootstrap resamples are typically used (e.g. Archer et al., 1997; Yang, 2011).** Yes, indeed, we are aware that 1000 is a typical value for bootstrap

resamples. However, such value was not affordable for estimating HSIC measures in a reasonable computing time. We then preferred to use 100 replications for all the tested methods, even the ones with low computational cost. Justification for this value has been added in the text.

**Table 2: I believe that the LAImin and LAIharv are missing. The Table would also need to include an additional column that specifies at which spatial level the parameters are defined (e.g. soil horizon, plot/VFS). It took me a while and a bit of digging in the manuscript to get this information. I would also add the value of the standard scenario in Table 2, this would further improve readability.** As suggested, we added a column to Table 2 with spatial level definition and we also specified the values for the nominal simulation.

**Section 2.5: this section does not clearly explain that the vegetation parameters and hpond are considered for vineyard plots and VFSs separately. As already mentioned in my previous comment, I think that the parameter should be clearly introduced in Section 2.1, which would improve readability and clarity.** Yes, modified as suggested

**Section 3: As mentioned in my main comments, the manuscript lacks a discussion of the methodology and results with respect to previous studies, which could be highlighted in an additional discussion section.** As suggested, a discussion section has been added to comment on the global methodology and to put it into perspective in relation to previous studies.

**P463 'It is commonly stated that [...]'. This sentence needs to be better justified. A reference is missing (e.g. Wagener & Pianosi, 2019). It can also be that many parameters are influential, but have only a small impact on the output except for a few parameters (e.g. five or six) that dominate the output variability.** Indeed, the sentence is inaccurate. The screening step intrinsically does not allow to draw conclusions on the number of parameters that dominate the output variability. We propose to eliminate the sentence to avoid confusion and hasty conclusions.

**L566-568: Could you explain more why is it more costly to assess the sensitivity analysis at the local scale compared to the catchment scale? From Eq.17, it looks that anyway the catchment scale indices require the calculation of the local scale indices.** Indeed, in this case study we re-use the local scale indices to calculate the aggregated ones implying in this case no difference in computational cost. However, in its paper Gamboa et al. (2014) proposes an estimator for these aggregated indices that does not need the calculation of local indices. As local indices were calculated anyway in our case, we did not try such estimator but we mention it in the text since it seems very interesting to us, in the case the user does not want to compute local indices but directly the aggregated ones.

In addition, you also suggested to modify the Data and Code availability. In order to comply with GMD Code and Data Policy, two Zenodo repositories have been created to provide both PESHMELBA source code and data. The urls and DOI have been added to the 'Code and Data Availability' section:

- PESHMELBA software: <https://zenodo.org/record/6319769#.YinMV1TjKuk>

- Data and codes for sensitivity analysis:  
<https://zenodo.org/record/6319773#.YinMc1TjKuk>

