Replies to the review comments

Dear S. Razavi:

We are very grateful to your comments for the manuscript. Based on your comments and requests, we have made extensive modification on the original manuscript, and the correction sections in the revised manuscript are marked with underlines for easy checking. Here below are our replies to your comments.

Referee #1:

This manuscript is on the interesting and important topic of multi-criteria sensitivity analysis. This kind of problem is commonly encountered by hydrologic modellers(and perhaps modelling community in general) that the sensitivity of different model responses (or summary metrics/criteria) to different model inputs can differ, and it may be non-trivial to come up with a unified sensitivity assessment that reflects the collective influence of a model input on multiple outputs. This process, however, should be carefully performed and interpreted as it might result in loss of information (sensitivity of each individual response may tell you an important side of the story in the intended context) or misleading assessment.

Overall, I found the method proposed in the manuscript interesting, but I have major reservations with this manuscript, outlined below. I believe substantial revisions are required. I should note that Razi Sheikholeslami and Shervan Gharari have helped me with this review.

Comment 1

First of all, the quality of writing (including English, grammar, organization, etc.) and more importantly equations and notation may be improved significantly. Also, some of the equations are repeated once or twice. Some of the notation might not have been chosen properly.

Reply to comment 1

Thanks to the comment of the Referee, we have improved some inappropriate places. Such as

In page 2 line 51, we changed ‘‘used to study’’ by ‘‘are used for studying’’.

In page 2 line 53, we changed ‘‘continuous’’ by ‘‘continuous domain’’.

In page 3 line 65, we changed ‘‘proposed’’ by ‘‘proposed a theory’’.

In page 3 line 71, we added ‘‘the’’ before ‘‘most’’.

In page 6 line 170, we changed ‘‘proposed for the importance measure of the multivariate output’’ by ‘‘proposed to analyze the uncertainty with multivariate output’’.

In page 8 line 209, we changed ‘‘lemma 2.2.1’’ by ‘‘Proportion 2.2.3 (I)’’.

In page 8 line 215, we changed ‘‘Proportion 2.2.1’’ by ‘‘Proportion 2.2.3 ’’.

In page 8 line 216, we changed ‘‘(i)’’ by ‘‘(II) ’’.

In page 9 line 217, we changed ‘‘(ii)’’ by ‘‘(III) ’’.

In page 9 line 218, we changed ‘‘(iii)’’ by ‘‘(IV) ’’.

In page 9 line 219, we changed ‘‘(iv)’’ by ‘‘(V) ’’.

In page 9 line 220, we changed ‘‘(v)’’ by ‘‘(VI) ’’.

In page 9 line 222, we changed ‘‘Point (i)’’ by ‘‘Point (II) ’’ and changed ‘Point (ii)’’ by ‘‘Point (III) ’’.

In page 9 line 223, we changed ‘‘Point (iii)’’ by ‘‘Point (IV) ’’ and changed ‘For (iv) and (v)’’ by ‘‘For (V) and (VI) ’’.

In page 11 line 273, we changed ‘‘these’’ by ‘‘which’’.

In page 14 line 326, we changed ‘‘can’t’’ by ‘‘cannot’’.
In page 14 line 327, we changed ‘‘during’’ by ‘‘when’’.
In page 20 line 416, we deleted ‘‘For SFDCE, it can be found’’.
In page 20 line 417, we added ‘‘the’’ before ‘‘same’’.
In page 20 line 420, we added ‘‘a’’ before ‘‘great’’.
In page 20 line 425, we deleted ‘‘which’’.
In page 21 line 433, we added ‘‘and’’ before ‘‘the following importance’’.

Comment 2

The main motivation of the study, as stated in the abstract, is to analyze the effect of model inputs on “correlated” multivariate output. However, the method proposed assumes that different outputs are orthogonal and their correlation is not accounted for anywhere in the formulations. I might be missing something, but it this is true, then the method doesn’t serve for purpose.

Reply to comment 2

Thanks to the comment of the Referee, I have already deleted the “correlated” word in revised paper.

Comment 3

My understanding is that the proposed method is nothing but a weighting approach that weights the Sobol’s indices for each individual model output based on that output variance. More accurately stated, the weight for the indices of an output is the square of the variance of that output. If my understanding is right, then the method might not possess much novelty. In other words, the proposed method is simply a “supposedly objective” weighting approach for different model outputs. In practice, this method results in sensitivity rankings that are the same as (or consistent with) the rankings based on the output with largest weight. The case study results of the manuscript confirms this.

Reply to comment 3

As pointed out by the referee, the proposed method is a weighting approach. But the weight is introduced from vector space, which considers the direction and the magnitude of the variance vector of the output space together. The method results in sensitivity rankings that are the same as the rankings based on Sobol index in some cases, but the values of sensitivity analysis the physical significances are completely different. And after the derivation, we can find that the extension of sobol index is a special case of the projection index.

Comment 4

4) Related to the above comment, the weighting approach of the proposed method might not be appropriate. The weighting is overly sensitive to the way that the different model outputs are normalized/became dimensionless (Equation 10 of the manuscript). If you use a different normalization approach, you might get an entirely different assessment. Also, dividing the values by their average (Equation 10) might not be a good strategy, as the spreads (variances) of different outputs might remain significantly different after the conversion, even of different orders of magnitude. A better strategy might be to standardize the outputs (dividing by standard deviation). But still you would have to deal with the issue raised in the previous comment.

Reply to comment 4

As pointed out by the referee, the weighting is overly sensitive to the way that the different model outputs are normalized dimensionless. I have tested it that not only the proposed index but the Sobol index also gets a different assessment using different normalization approach. And dividing by standard wrongly makes the different output equally important and some information lost. That is why we divide the values by their average instead of their variances.

Comment 5

5) Literature review: the manuscript does not provide the reader with the status quo. There is some literature review, but limited and unbalanced. First of all, I suggest the authors have a look at the following paper, which is a fully multi-criteria approach with minimal subjectivity:

Rosolem, R., Gupta, H. V., Shuttleworth, W. J., Zeng, X., & De Gonçalves, L. G. G. (2012). A fully multiple-criteria implementation of the Sobol' method for parameter sensitivity analysis. Journal of Geophysical Research: Space Physics, 117(7), [D07103].DOI: 10.1029/2011JD016355

Also, at the risk of self-promotion or self-propaganda, I'd suggest the authors have a look at the following papers. The fundamental question is that how variance itself and its decomposition can be meaningful for global sensitivity analysis. Of course, I am not arguing they are not meaningful, but asserting that there are caveats that need to be recognized and taken care of.

Razavi, S., and H. V. Gupta, (2015), What do we mean by sensitivity analysis? The need for comprehensive characterization of "global" sensitivity in Earth and Environmental systems models, Water Resour. Res., 51, 3070–3092 doi:10.1002/2014WR01652

Razavi, S., and Gupta, H. V., (2016), A new framework for comprehensive, robust, and efficient global sensitivity analysis: II. Application, Water Resources Research, 51,doi:10.1002/2015WR017559 full-text.

Razavi, S., and Gupta, H. V., (2016), A new framework for comprehensive, robust, and efficient global sensitivity analysis: I. Theory, Water Resources Research, 51,doi:10.1002/2015WR017558 full-text.

Also, derivative-based method and elementary effect method are essentially the same but with slightly different numerical implementations for the step size to calculate the numerical derivatives/elementary effects (lines 54-57). Please refer to Razavi and Gupta (2015 WRR).

Reply to comment 5

Thanks to the comment of the reviewer, we have learned more about usage of the fully multi-criteria approach, the difference between derivative-based method and elementary effect method and the general sensitivity analysis framework which called ''Variogram Analysis of Response Surfaces'' (VARS), and referred to these literatures in the introduction of the revised version.

Comment 6

The idea of using Polynomial Chaos Expansion (PCE) is divergent from the core of this manuscript, and may be confusing to readers. Please note that the only reason to use PCE instead of a full Monte-Carlo-based Sobol analysis is computational efficiency, which of course comes at the trade-off of losing accuracy. To me, computational efficiency is a wholly different story and does not go well with the core idea of the manuscript. Note that PCE is a metamodeling approach, and like any other meta-modeling approach, may have two major shortcomings. First is the quality of results depends on the quality of fit. So if the response surface is complex, PCE (polynomials) may not fit the response surface well, and as such, the results might be erroneous.

Second is metamodels are handicapped for high-dimensional problems, that is why they have typically been used in the literature only for problems with less than ~20 input variables. The authors may find more discussion on these two shortcomings in the following papers:

Razavi, S., B. A. Tolson, and D. H. Burn, (2012), Review of surrogate modelling in water resources, Water Resources Research,48, W07401 doi:10.1029/2011WR011527. 32pages.

Razavi, S., B. A. Tolson, and D. H. Burn, (2012), Numerical assessment of meta-modelling strategies in computationally intensive optimization, Environmental Modelling and Software, 34(0), 67-86

Reply to comment 6

According to referee's comment ,we add some interpretation and reference to illustrate the advantages and disadvantages of PCE in new revised manuscript.

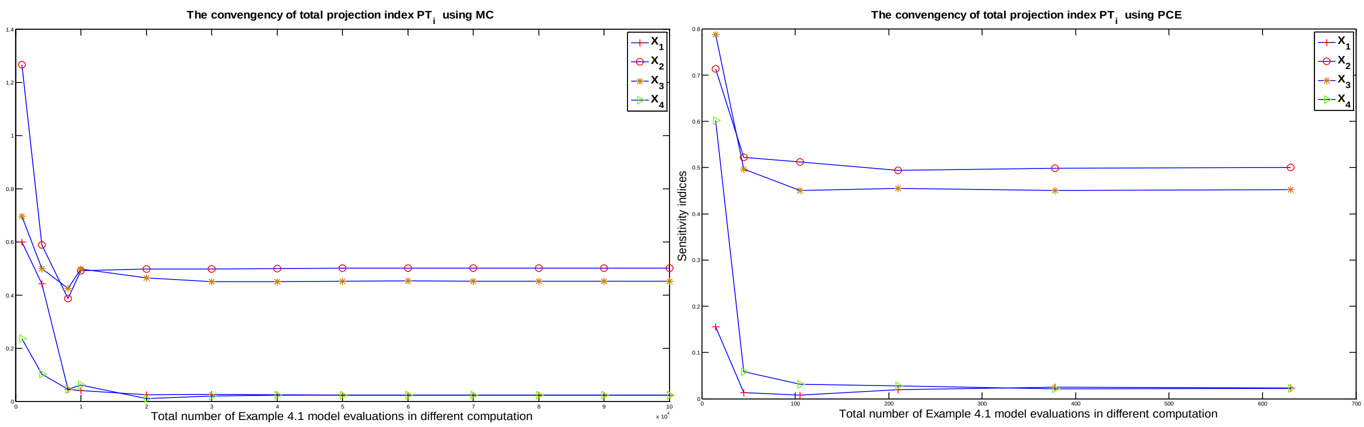
Comment 7

Related to the above comment, if the authors like to include the comparison of PCE and direct Monte-Carlo-based integration of Sobol's indices, it would require a more substantive analysis. Due to stochasticity of these algorithms, any comparison of this nature would require running multiple replicates of each method to account for sampling variability and randomness, and also, such analysis should be done at different computational budgets (different numbers of function evaluation) to ensure a fair and thorough comparison. The above Razavi et al. (2012 EMS) paper details the elements of such comparisons.

Reply to comment 7

Thanks to the comment of the reviewer, we had referred Razavi et al paper and added some substantive analysis. Such as

Fig.1.The total effect indices of multivariate output of the HBV model



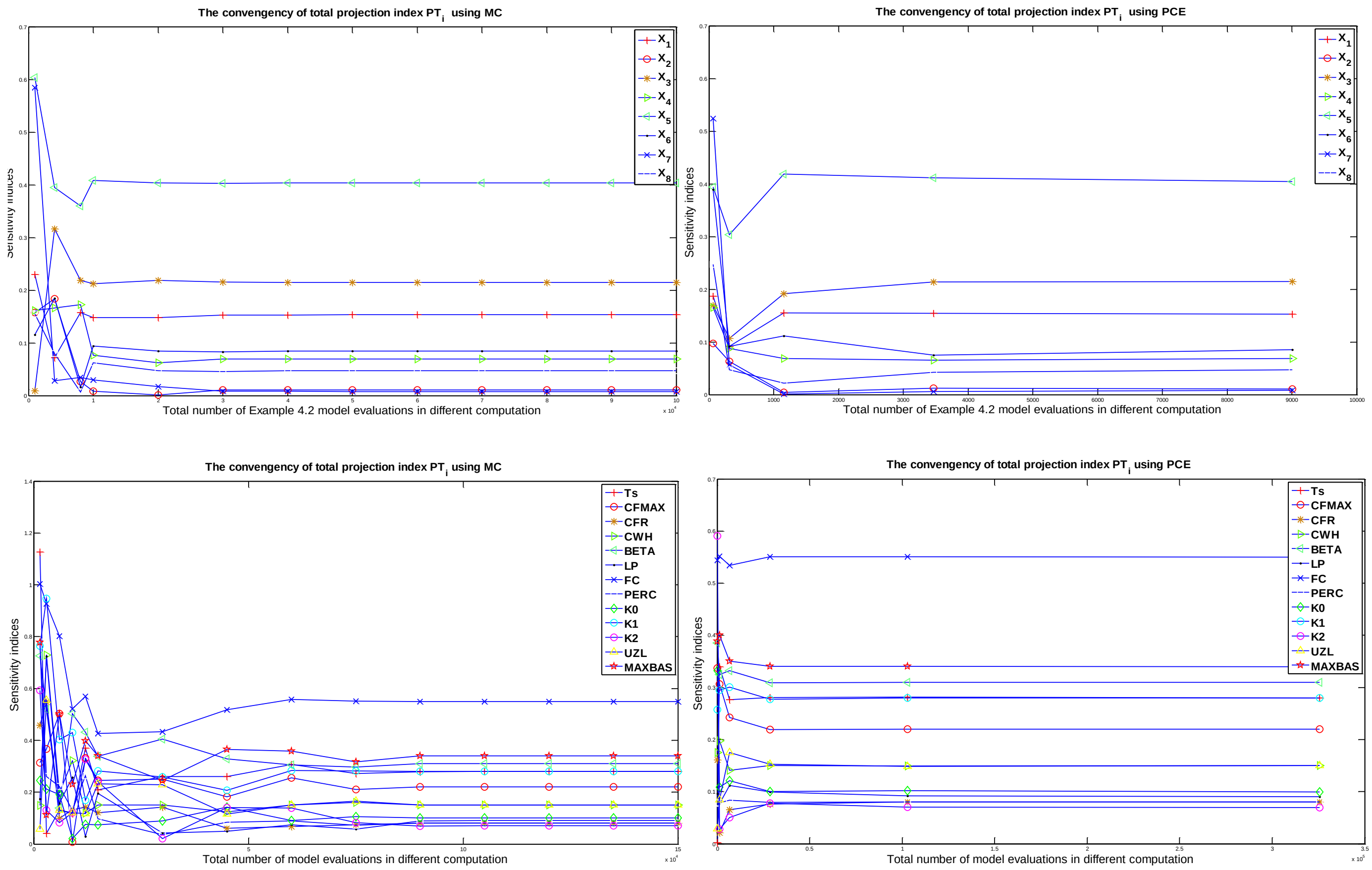


Figure 2. Methods average performance comparisons on the Example 4.1, Example 4.2 (over 30 replicates) and HBV case studies (over 10 replicates) at different computational budget scenarios.

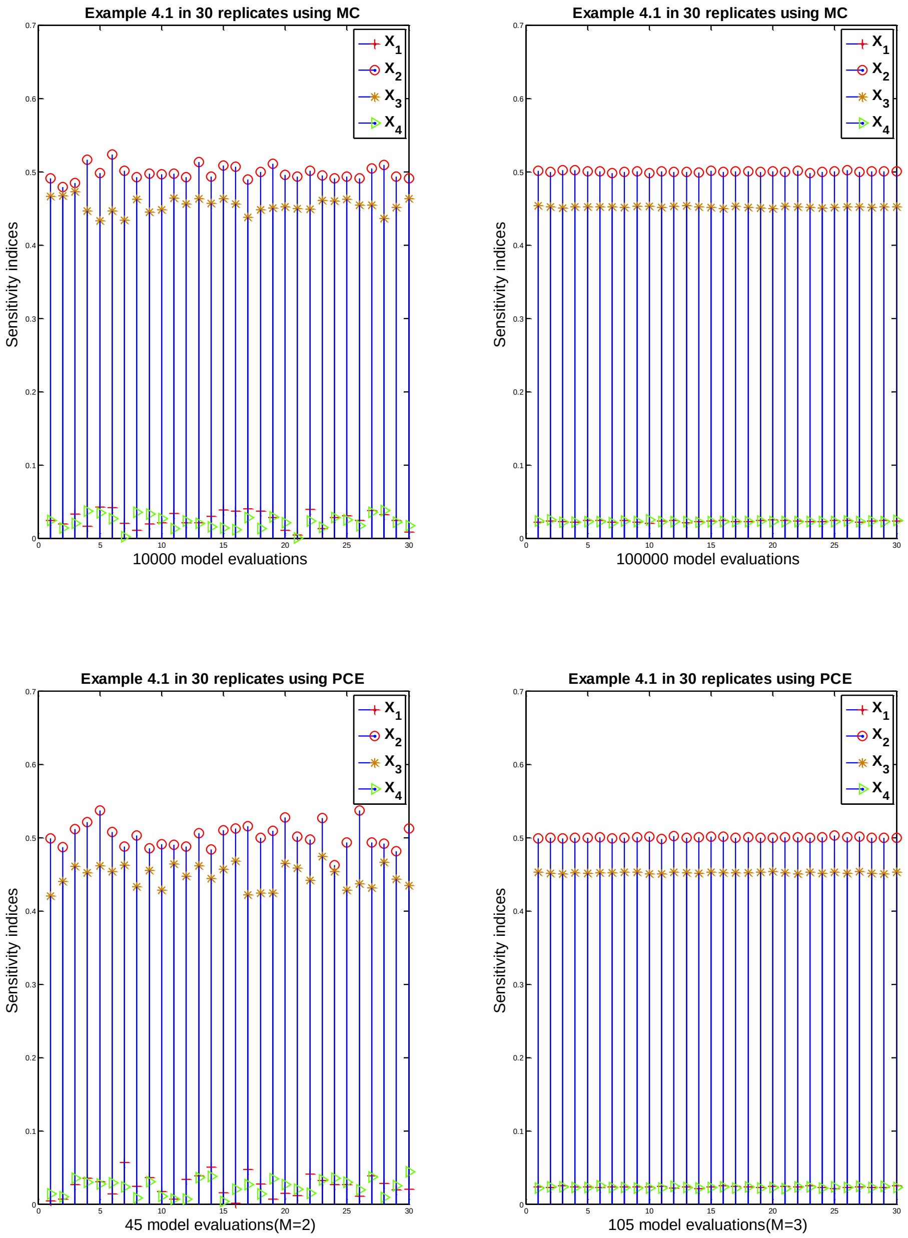


Figure 3 Methods performance comparisons on the Example 4.1.

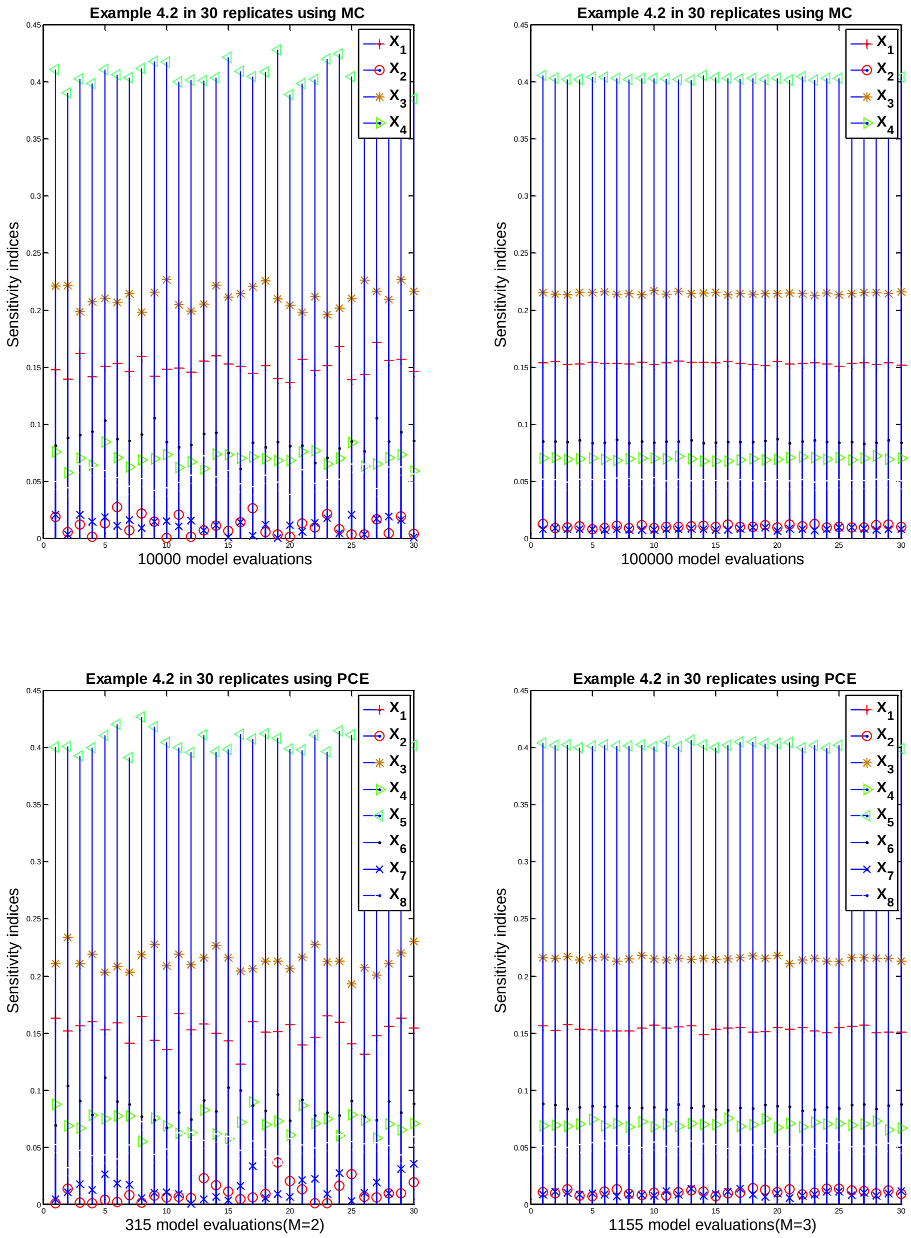


Figure 4.Methods performance comparisons on the Example 4.2.

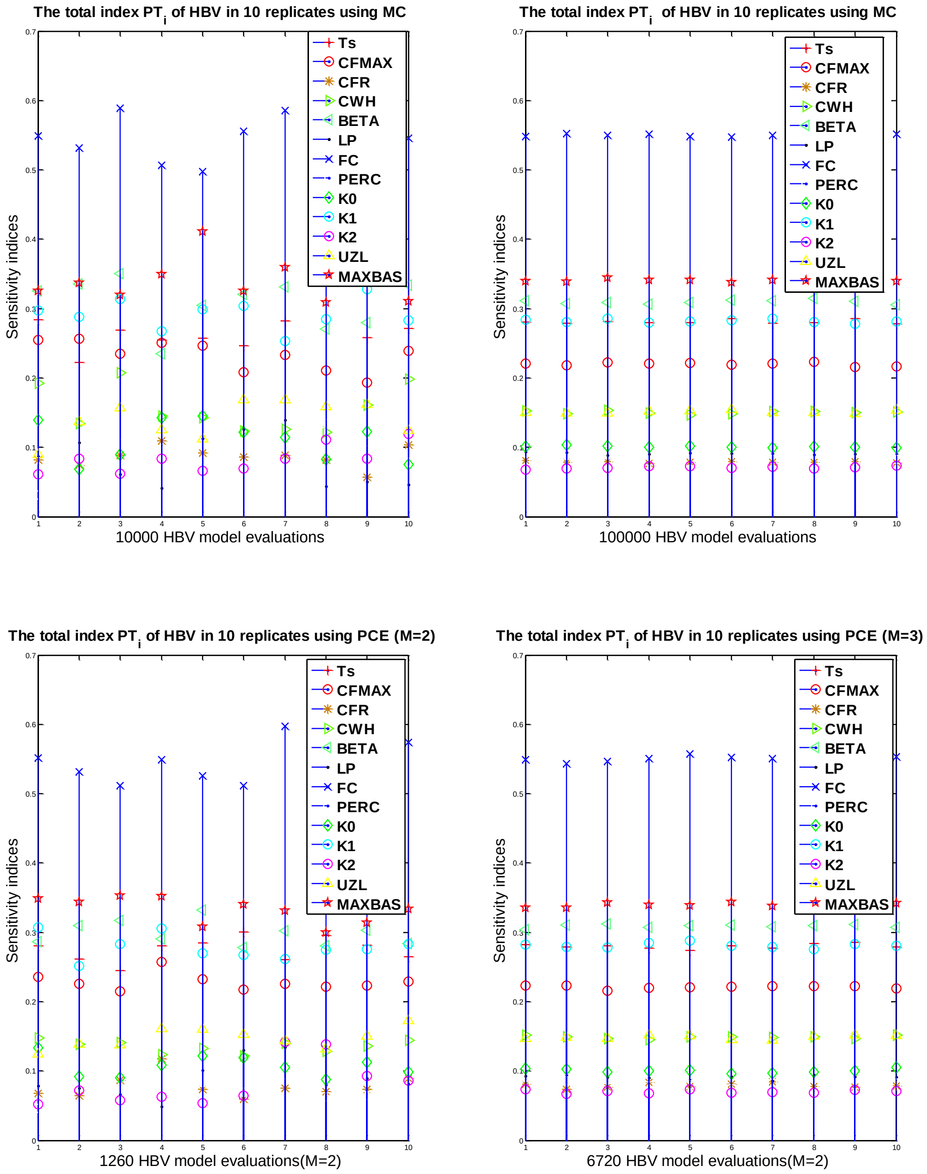


Figure 5.Methods performance comparisons on the HBV model.

Comment 8

This manuscript might be more suitable to be published in an Applied Mathematics journal. The relevance to the HESS community might not have been adequately established. The main (and probably the only) connection is the HBV case study. A little bit more work might be required to strengthen the connection. Lines 55-69, "it is recommended . . .", readers may wonder who recommends this. The authors?

Section 2 requires a lot of improvement. For example, do you need to have many 3rd level sub-sections such as 2.2.1?

Reply to comment 8

Thanks to the comment of the referee, the new index proposed in this manuscript mainly analyzes the multiple objective problems in hydrology models. Two numerical examples proof the accuracy of the index, and we apply it for HBV case to reduce uncertainty in parameter estimates through multiple calibrations. We added more discussion of HBV result in Abstract and literature review of multiple criteria of hydrological cases.

We have replaced “it is recommended . . .” by ‘‘And they recommended ’’.

And we have improved section 2 and deleted the repeated 3rd level sub-sections.