Hydrol. Earth Syst. Sci. Discuss., 10, C1398–C1406, 2013
www.hydrol-earth-syst-sci-discuss.net/10/C1398/2013/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Resolving structural errors in a spatially distributed hydrologic model" *by* J. H. Spaaks and W. Bouten

J. H. Spaaks and W. Bouten

jspaaks@uva.nl

Received and published: 3 May 2013

1 Reply to general comments from Referees #1 and #2

We are grateful that the Referees consider our paper 'innovative', 'interesting', 'novel', 'of high relevance to the readers of HESS', 'compelling' and 'very well written'. The Referees' main criticism is related to whether the approach will work on real-world cases, in particular with regard to data abundance, data quality, unknowable boundary fluxes, and soil heterogeneities. Both Referees consider this their main criticism of the paper. Below, we will discuss our motives for designing the paper as we did.

We have tried to be very conscious about how we communicate the paper's main message to the reader—which elements of our 'story' should be included and what was better left out. We feel that the reader is best served by having one main message ('SODA's state updating patterns are more helpful in diagnosing model structure errors than are SCEM-UA's simulation-observation residuals'). To avoid obscuring this main message, less important messages should be avoided, if at all possible. This was in fact the primary reason for our use of artificial data in this study: had we used real data instead of artificial data, we could not have conclusively shown the potential of analyzing state updating patterns, because the truth would not be known and our message would be obscured by issues relating to data quality (including incommensurability, measurement error, etc.), data abundance relative to process heterogeneity, and uncertainty associated with for instance the estimate of water balance. This is because the results would still require *interpretation* on our part, as opposed to less subjective *clarification* that is needed when using artificial data. Since our interpretation would likely be different from someone else's, there would be great opportunity for the discussion to get bogged down in the specifics of one particular hillslope, wheras we were hoping to stimulate a discussion about the relative merits of SCEM-UA and SODA.

Because of these considerations, we do not intend to show any results based on real world data in the revised version of the manuscript. However, the Referees' comments have made it clear to us that we have painted too positive a picture with regard to the immediate possibilities of analyzing state updates. For the revised version of the manuscript, we will emphasize that this method is not the end-all and be-all of model diagnostics, and that much work remains to be done before we know how to usefully apply the method within the context of real-world experiments. We also propose to better explain our choice for artificial measurements in the discussion and to include additional simulation work, in which we decrease the number of observation locations. The results of this additional simulation will then serve to illustrate a discussion about the Referees' questions, specifically:

1. what happens if there are less measurements?
2. what happens if the measurements of pressure head are not taken at exactly the location where model structural error is introduced?

Referee #1 further expresses interest in the effect of multiple, interacting model deficiencies on the feasibility of the proposed approach, and suggests that this may be pursued in a future paper. We agree with the Referee that this is an important subject, and unfortunately also a complicated one. Having multiple interacting contributors to state updating will feature in the additional simulation we do for the revised version of the manuscript. This is because when the measurements of pressure head are not taken at the location where model structural error is introduced, state updates can not be performed at the locations of the hotspots anymore, but may occur some distance downslope. This means that the state updating pattern that is performed at the location away from the hotspot is in fact different from the state updating that would be performed had pressure head been observed at the hotspot. State updates can thus no longer be interpreted directly as additional water loss to the bedrock, but must instead be interpreted as the local effect on pressure head of failing to remove additional water some distance upslope. When pressure heads are not observed at the hotspots, state updates thus become a mixture of two signals: (1) the structural error introduced at the (unobserved) hotspots, which previously could be interpreted directly as additional water loss to bedrock; and (2) the altered dynamics of flow in the part of the domain between the current (observed) location and the (unobserved) hotspot. In the revised version of the manuscript, we will present and discuss simulation results that illustrate the above.

In addition to the questions raised above, Referee #2 also asks: How does analyzing the state updates help with real-world field experiments, in which boundary fluxes are

unknown? Will [SODA] still be able to improve the model structure if the error-free perfect "truth" is not known? These are questions that are very relevant for practical application of the method. With regard to unknowable boundary conditions, our results (Fig. 8–10) show that the correct parameter values were identified, and that the associated state updates provide insight into where sink hotspots are located. It is important to note that these locations were indeed 'unknown', in the sense that we did not explicitly specify them during calibration of the model with SODA. The fact that the hotspot locations were correctly identified is just due to the fact that any process associated with a noticable (as in 'measurable') effect on the value of the system's state (in this case: pressure head) will leave its signature in the space-time pattern of state updating (for example, Fig. 9). In the manuscript we showcase this for a sink, but it does work fundamentally the same for any process.

With regard to the 'error-free' aspect: whether or not an analysis of state updates can help to improve the model structure depends on whether the effect of any particular missing process is noticable in the observations of the system's state values. For example, if the measurement uncertainty is very large, or the process has only a small effect on the observed value of the system's state, an analysis of state updates may not yield any useful insights (but note that this not so much a shortcoming of the method; other methods would be equally hard-pressed to make sense out of so little useful information). A similar argument can be made for the measurement interval (time) or measurement spacing (space) relative to the scale of the process of interest: if the system's state is affected by a process that is not represented in the model, but that only works over short ranges (in time or space, respectively), the observations of the system's state may not show any evidence of the existence of this process if the measurement density is low. Again, this is not so much a shortcoming of the SODA method, but simply results from the fact that even powerful methods of analysis cannot *create* their own information—they can only *use* information contained in measurements. If the measurements do not contain any such information, then analysis is hopeless anyway. In the revised version of the manuscript, we will elaborate on these issues.

The remainder of this reply covers Referee #2's specific comments in detail.

2 Specific comments

1. (comment) Title is a little general

   (reply) We agree with the Referee that the title may be too general. In the revised manuscript, we will include some more specific terms in the title in order to clarify what the paper is about.

2. (comment) Fig. 1 not very meaningful, can easily be described in words

   (reply) For the revised version of the manuscript, we will remove this figure.

3. (comment) P. 1821, L 9: scale triplet?

   (reply) We mean to refer here to to the three aspects of scale: spacing, extent, and support, as per Blöschl and Sivapalan (1995) and Western and Blöschl (1999). In the revised version of the manuscript we will include these references.

4. (comment) P. 1821, L 27: being right for the wrong reasons is not even acceptable in the first case!

   (reply) Whether it is acceptable or not to be right for the wrong (or unknown) reasons depends on the purpose of the modeling. In the manuscript, we distinguish between modeling for making predictions on the one hand and modeling as a tool for testing and improving our understanding of hydrological systems on the other. Predictions tend to be needed on short notice, so there is generally no time to

   be right for the right reasons. We agree with what the Referee implies though: of course there are risks involved with the application of models that are right for the wrong reasons (for instance when conditions such as landuse or climate cannot be considered static). But the alternative is to have no predictions at all until we are right for the right reasons (which is an unattainable concept anyway, since all models, no matter how sophisticated, are simplicfications and therefore wrong to some extent).

5. (comment) P. 1821, L 28-29: "..our understanding of how hillslopes and watersheds function".

   (reply) For the revised version of the manuscript, we will adapt the text according to the Referee's suggestion.

6. (comment) P. 1824, L 28: what does "assimilating observations" mean exactly? Unclear

   (reply) In the revised version of the manuscript, we will explain what we mean by "assimilating observations": adjusting the system's state value every time a new observation becomes available.

7. (comment) P. 1826, L 10: the terms "forward model" and "inverse model" are a little confusing; maybe use perfect/reference model and simplified model?

   (reply) The terminology is indeed a bit confusing. In the revised manuscript we will change the text in accordance with your suggestion; instead of "forward model" we will use "reference model", and instead of "inverse model" we will use "simplified model".

8. (comment) P. 1828, L 19: How were the parameters listed in Table 1 chosen? Random? Taken from another study?

(reply) The parameters and the geometry of the soil domain were taken from Luisa Hopp's work about virtual experiments with HYDRUS on the Panola hillslope (Hopp and McDonnell, 2009). However, we altered a few things to suit our purposes. For example, we made the soil domain thicker, we decreased the spatial resolution of the grid, and changed the hydraulic conductivity. In the revised version of the manuscript, we will refer to Hopp and McDonnell's paper.

9. (comment) P. 1828, L 20: provide more details here how the initial state was generated; zero pressure head at the soil-bedrock interface, i.e. lower boundary, and then in the soil column above?

   (reply) The reference level of pressure head is the soil-bedrock interface, i.e. the lower boundary of the domain. A pressure head of 0 was assigned to all nodes at the lower boundary of the domain at t=0. Nodes in the upper 4 layers were assigned a pressure head of $-z$, in which z is the vertical distance from a given node to the lower boundary of the domain at a particular X,Y location. With this initial state, SWMS_3D was run until t=96 h. During this period, soil water was redistributed due to hydraulic head differences. The slope of the domain, convergence of flow due to varying soil depth, as well as water removal from the domain at the sink hotspots were the driving factors in this redistribution. The pressure head pattern at t=96 h was then saved to file and served as a starting point for all further simulations. In the revised version of the manuscript, we explain in more detail how the pressure head was initialized.

10. (comment) P. 1829, L 6: insert here the total simulation time of 216 h

    (reply) In the revised version of the manuscript, we will mention the total simulaton time.

11. (comment) P. 1829, L 20-21: note that variable 1 is not measurable in the field whereas the other two can be measured (although not completely, error-free,...)

    (reply) Please refer to the previous section for our rationale for using artificial data.

12. (comment) P. 1831, L 17: When convergence has not been...

    (reply) In the revised version of the manuscript, we will change “while” to “when”, as per the Referee's suggestion.

13. (comment) P. 1833, L 9: i.e. the variables that are used in the first two OF describe integrated hydrologic response whereas the third variable is spatially-distributed hydrologic information

    (reply) In the revised version of the manuscript, we will add the text suggested by the Referee in order to emphasize the difference between integrated responses on the one hand, and the spatially distributed response on the other.

14. (comment) Fig. 2 and 4: I would combine those figures, i.e. indicate the location of hot spots with high leakage to bedrock on the soil depth map. Fig. 4 doesn't really provide additional information, and relating soil depth and hot spots to pressure head patterns in Fig. 5 and others becomes easier

    (reply) In the revised version of the manuscript, we will merge the soil depth figure (Fig. 2) with the figure showing the spatial distribution of the sink parameter (Fig. 4).

15. (comment) P. 1836, L 27-28: “residuals cannot be interpreted as being new and local” – can you explain this better? Meaning unclear

    (reply) We have included an explanation about this in our reply to Referee #1 (item 29). In the revised version of the manuscript, we will provide a more elaborate explanation about ‘new and local’ residuals, any why this is important within the context of model improvement.

16. (comment) P. 1838, L 18-20: I would agree! Except that in the real world we wouldn't have the perfect, error-free observation. How does this concept then work?

    (reply) Please refer to the previous section for our rationale for using artificial data.

17. (comment) P. 1839, L 16: "installing a more precise measurement device" sounds much easier than this actually is; there is always the problem of density of installations and the representative volume of a measuring device

    (reply) Please refer to the last paragraph of the previous section for our remarks on measurement scale relative to the scale on which the process of interest acts.

3 References

- Blöschl, G. and Sivapalan, M.: Scale issues in hydrological modelling—a review, Hydrol. Process., 9, 251–290, 1995.
- Hopp, L. and McDonnell, J. J.: Connectivity at the hillslope scale: Identifying interactions between storm size, bedrock permeability, slope angle and soil depth, J. Hydrol., 376, 378–391, 2009.
- Western, A. W. and Blöschl, G.: On the spatial scaling of soil moisture, J. Hydrol., 217, 203–224, 1999.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 1819, 2013.