Biogeosciences Discuss., 10, C7528–C7532, 2013
www.biogeosciences-discuss.net/10/C7528/2013/


Biogeosciences
Discussions
Open Access

# *Interactive comment on* "Monte Carlo based calibration and uncertainty analysis of a coupled plant growth and hydrological model" *by* T. Houska et al.

**B. Scharnagl (Referee)**

benedikt.scharnagl@tu-bs.de

Received and published: 21 December 2013

## 1 General comments

Houska et al. present an application of a coupled soil hydrological and plant growth model to simulate winter wheat. The authors apply the GLUE methodology with three likelihood functions to calibrate the model using data sets originating from three different agricultural management practices. Based on these results, the authors identify structural deficits in the plant grow model. The discussion paper presents an interesting case study that fits well into the scope of this journal.



The discussion paper is well written, concise and comprehensible. However, it lacks some relevant information on the setup of the soil hydrological model and the observations used to calibrate it. In my view, this impedes a final assessment of the discussion paper. Further details are given in the specific comments below.

## 2 Specific comments

(2.1) Several aspects of the soil hydrological model need to be clarified. The authors say that the Richards equation was used to simulate one-dimensional water flow in 50 soil layers. However, what was the depth of the soil profile? Is the profile assumed to be homogeneous, and if so, can this assumption be justified by available information on soil type and texture? How thick is one such layer? The numerical simulation of soil water dynamics under natural boundary conditions requires a fine discretization of the simulation domain near the soil surface to accommodate for very large gradients in matric potential that may arise as a response to meteorological forcing. And what about root water uptake? How was it simulated? This is clearly not a trivial task for an annual crop and needs further specification.

(2.2) Several aspects concerning the soil water content observations need to be clarified as well. First, what was the measurement technique? Time Domain Reflectometry (TDR)? Gravimetric measurements? Were these measurements taken at a single location within a given plot or is it some sort of average of spatially distributed observations? Second, the authors state that measurements were taken at 33 days between 1992 and 1998. But the period that was actually used for model calibrations contain only 15 measurements. So why do the authors mention the 33 observations then? Third, the authors say that the measurements were taken at 0.15, 0.45, and 0.75 m depth. But the axes labels of Figs. 2 and 5 suggest that these measurements and corresponding simulations give the soil water content averaged over a 0.3 m depth

interval each. These are indeed two quite different quantities, especially near the surface where gradients in soil water content can become quit large. So, what is the actual depth interval of the observations?

(2.3) The authors mention that “behavioural parameter sets were found to be around 10% with respect to the measured value” (p.19523 ll. 7-8). But this is exactly what was expected because the threshold criterion used to distinguish a behavioural from a non-behavioural parameter set was set to 10% (p.19519 l.15). This essentially means that the GLUE algorithm worked in exactly the way it was intended to do. But it does not tell us anything about the goodness-of-fit or the ability of the model to describe the data.

(2.4) The GLUE method is known to be highly inefficient, essentially because the parameter sets are drawn randomly from the prior distribution and do not take into account the information gained so far from behavioural parameter sets obtained from previous model runs. This is especially true for high dimensional parameter spaces, which was corroborated by the present study. The authors point out correctly (p.19524 ll. 7-10) that, as an alternative, Markov chain Monte Carlo algorithms can be used to reduce the required number of model runs. But they incorrectly assume that these algorithms may be less robust with respect to finding the global optimal parameter sets. The DREAM algorithm mentioned by the authors implements a globally convergent search strategy that was especially designed to be capable to handle multiple local optima. Such an algorithm is very likely to perform much better in finding the region of the parameter space with behavioural models and in sampling this region adequately, leading to more robust uncertainty bounds of the parameters and the model predictions.



## 3 Technical corrections

(3.1) p. 19510 l. 9: The programming language that was actually used for model coupling needs not to be mentioned in the Abtract. It is a purely technical issue and not relevant to the results presented in this study.

(3.2) p. 19510 l. 12 and p. 19514 l. 14: It is not only the retention curve that was parameterized using the van Genuchten-Mualem model, but also the conductivity curve! A better expression would therefore be “and the van Genuchten-Mualem model of the soil hydraulic properties”.

(3.3) p. 19510 l. 20: “drawn from an equally distributed parameter space” should be replaced by “drawn from a uniform distribution”.

(3.4) p. 19511 l. 16: “parameters” instead of “parameter”.

(3.5) p. 19511 l. 22: “et al.” is missing in this citation.

(3.6) p. 19514 l. 13: It should be “the Richards equation”, not “the Richard’s equation”. It is named after Lorenzo A. Richards.

(3.7) p. 19514 ll. 18-19: “we used existing climate data” should be replaced by “we used meteorological data”. If they would not be existent, the data could not have been used anyway. And to be precise, these type of data are meteorological data, not climate data (which is some sort of long-term average of meteorological data).

(3.8) p. 19515 l. 22: The parameters that appear in brackets were not defined previously, nor is reference made to Tab. 1 where they are defined.

(3.9) p. 19516 l. 22: It is not the “probability distribution of the measurement errors” but rather the “likelihood function of the model residuals”. Measurement errors and model residuals are the identical if, and only if, the model would be perfect. But this is – almost by definition – never the case.

(3.10) p. 19517 l. 3: The minus sign before the summation should be removed.

(3.11) p. 19519 ll. 1-2: Is the Python package (with reference) really worth mentioning? I mean, sampling from a uniform distribution is a rather trivial task in any programming language and needs no further specification.

(3.12) p. 19522 ll. 9-10: The similarity of the results for the $n$ parameter of the current study and the study of Ippisch et al. cited by the authors is completely irrelevant. The optimal value of $n$ strongly depends on soil type and bulk density, among other soil properties. Ippisch et al. investigated a different soil that has – by chance – a similar $n$ value.

(3.13) p. 19523 ll. 22-23: The percent signs should be removed. The unit of soil water content (and also RMSE) in these studies was not percent, but given on a volume by volume basis.

(3.14) p. 19527 l. 27: It is not clear to me what the authors mean by "our implementation of iterative steps". What exactly was the iterative part in this study? Please add some explanatory information here. Or alternatively, rephrase this sentence.

(3.15) p. 19535: In my view, the figure might become much clearer if only the behavioural parameter sets would be plotted. I mean, what is the added value of showing the non-behavioural parameter sets? What information is this supposed to convey?

(3.16) p. 19536: For consistency, the model efficiency should be abbreviated with ″NSE" (see the legend).

---

Interactive comment on Biogeosciences Discuss., 10, 19509, 2013.