Hydrol. Earth Syst. Sci. Discuss., 12, C1530–C1540, 2015
www.hydrol-earth-syst-sci-discuss.net/12/C1530/2015/

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Inverse modelling of in situ soil water dynamics: accounting for heteroscedastic, autocorrelated, and non-Gaussian distributed residuals" *by* B. Scharnagl et al.

B. Scharnagl et al.

benedikt.scharnagl@ufz.de

Received and published: 15 May 2015

We thank Thomas Wöhling for the insightful comments. We have addressed most of them separately below. The sectioning and numbering of our replies follows those of the original comments. Changes we made to the discussion paper in response to the comments are indicated by italic font.

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

1 Major comments

1. This comment summarizes the scope and novelty of the discussion paper. We think that a reply is not needed here.

2. This comment summarizes the methodology used in the discussion paper. We think that a reply is not needed here.

3. We do not agree with the conclusions drawn here by the referee. In our view, the AR(1) scheme is effective even if the autocorrelation coefficient ϕ approaches one. This is because the effect of large values of ϕ, as described by the referee, is balanced out by the likelihood model (Eqs 14, but also Eqs. 15 and 16). An overly large value for the autocorrelation coefficient means that the decorrelated residuals ν_i are divided by a too small standard deviation $\sqrt{1-\phi^2}$. As a result, the standardisation of ν_i carried out in Eq. (14) leads to very high values of the standardised decorrelated residuals, which in turn results in a small value of the corresponding likelihood. In principle, the AR(1) scheme can therefore also handle the case ϕ close to one because it makes the posterior very small. However, evidence in the hydrological literature (Wöhling and Vrugt, 2011; Evin et al., 2014) indicates that the AR(1) scheme seems to be particularly sensitive to model errors (of any kind) if the autocorrelation coefficient approaches one, which eventually may result in physically meaningless parameter estimates. *We added a paragraph to the revised discussion paper in which this aspect of the AR(1) model is explained, following the arguments given above. We think that this additional paragraph will help to better understand this issue and will reduce the chance of potential misconceptions.*

4. This comment relates to the previous one. We do not agree with the referee in that the AR(1) scheme, which is implemented in Likelihood 2, is essentially ineffective. We think that a modified likelihood that does not account for autocor-

HESSD

12, C1530–C1540, 2015

Interactive Comment

relation in the residuals, as proposed by the referee, would not be meaningful. This is because the resulting residuals would not fulfil the assumption of mutual independence of the residuals, which would be inherent to the formulation of such a likelihood model. As we also argued in the discussion paper, we consider parameter estimates obtained from likelihood models that are based on unrealistic and unjustified assumptions as not meaningful from a statistical point of view.

5. This comment relates to the previous two. Again, we do not agree with the referee in that the application of an AR(1) scheme would actually not be desirable. In our view, it is a powerful method to explicitly consider autocorrelation in the residuals, which has shown good performance in manifold applications in many fields of science. The referee argues that the stationarity assumption inherent to the AR(1) scheme is not likely to hold in practice because modelling errors (leading to autocorrelated residuals) are strongly event-driven. The same fundamental criticism can also be found in Doherty and Welter (2010). In our view, the stationarity assumption is indeed not expected to hold exactly, but it might be approximately true such that the AR(1) model may be useful in practice. In the discussion paper, we essentially tested this hypothesis and showed that the (modified) AR(1) model indeed worked well in the present study, which provides evidence about the usefulness of the AR(1) model in practice. We therefore strongly disagree with the referee's argument that our argumentation "is misleading or even wrong". Note that the analyses done here does not exclude other, more complex formulations of the likelihood function, which further relax the assumptions on stationarity. Our contribution should rather be viewed as a crucial step into this direction. *In the revised discussion paper, we included a short paragraph in which we briefly discuss the issue of the stationarity assumption of the AR(1) model.* As we also argued in the previous two replies, and contrary to the referee's opinion stated in this comment, we do not think that there is a "fundamental problem in the implementation of the likelihood". We consider our statement regarding the modified

AR(1) scheme that was cited by the referee as still valid.

6. The first sentence of the comment points at a potential misconception, which we like to resolve first. Likelihood 3 results in correct posterior distributions because the resulting residuals fulfilled all the underlying assumptions regarding the statistical properties of the residuals. We also like to make clear that the standardised decorrelated residuals are not "closer to normality", as assumed by the referee, but closer to a skewed t-distribution. In the following we reply to the five sub-comments given by the referee. Firstly, as we also pointed out in the previous three replies, we do not share the referee's view that the AR(1) model "does not provide the correct statistical framework". In fact, we are convinced that it does. The practical results presented in the discussion paper illustrate the usefulness of this statistical framework. Secondly, we agree in that the estimated parameters of the likelihood model are also conditional on the process model (besides the observations). The dependence of the parameter estimates on the structure of the process and likelihood models is usually not explicitly stated in the formulation of the inverse problem. In our opinion, however, this dependence is unavoidable and not a potential weakness of the present approach. Thirdly, we do not think that the numbers of degrees of freedom in the likelihood model is too large and we dismiss the statement that it is "extremely large". In fact, every single parameter has a specific and well-defined function within the likelihood model. As we mentioned in the discussion paper (p. 2178, ll. 15-20), the identifiability of the likelihood model parameters was very good, which adds confidence in that the likelihood model is not overparameterized, as suggested by the referee. Please note that we now include scatterplots of the estimated statistical parameters as was suggested by Referee 1 (see our reply to comment 9 of Referee #1). Fourthly, we feel that a detailed analysis of the contribution of the various sources of modelling errors to the overall uncertainty, as suggested by the referee, would indeed be an interesting aspect that is well worth further consideration. This,

however, is clearly beyond the scope of the current study, which focuses on the formulation of an appropriate likelihood model. Fifthly, in contrast to the opinion of the referee, we consider the expressions “statistically correct” and “invalid” as appropriate in the given context and insist on keeping them in the text.

7. We assume that the referee refers to Figure 7 (instead of 3) here. It is important to realize that the “weighting” in Likelihood 1 is not correct from a statistical point of view. This is because the assumptions of independent, homoscedastic, and Gaussian distributed residuals that underlie this likelihood model are not fulfilled (see Figure 3). This is also why we think, contrary to the referee’s suggestion, the parameter estimates obtained with Likelihood 1 should be rejected. As opposed to what Referee #2 suggests here, these parameter estimates are not rejected because of the misfit they cause (which is similar to that obtained with Likelihood 3) but because of the striking inability of Likelihood 1 to correctly describe the actual statistical features of the residuals. We further think that “setting the variance in Likelihood 1” to a larger value, is not justified from a statistical point of view and is merely a subjective choice and a trick to artificially increase uncertainties. Furthermore, it would simply result in an even worse description of the actual statistical features of the residuals. The so-obtained uncertainty bounds would be “more realistic” (based on intuition or whatever makes up our idea of being “more realistic”), but would lack an objective statistical foundation.

8. We fully agree with the referee in that the ROSETTA pedotransfer function may give biased results that are of little value to predict field-scale water dynamics. The problem of biased prior information was considered and discussed in Scharnagl et al. (2011), where the prior distribution was derived and applied for the first time. The same authors also show that the inverse approach was robust against bias in the prior distribution. Furthermore, as already outlined in the previous reply, we do not consider the parameter estimates obtained with Likelihood 1 as “statistically incorrect” because they differ more strongly from the prior estimates.

The reason for our rejection of this parameter estimate is that the underlying assumptions of Likelihood 1 were not fulfilled (even not approximately).

9. Of course, the development and use of an adequate likelihood model, which is the topic of the present study, does not replace the need to revise and improve the process model – if necessary. We think that both steps – revision of the likelihood and process models – should ideally go hand in hand. But in the end, we will almost never be able to formulate an environmental model which makes residuals independent, identically, and Gaussian distributed and therefore some effort should be spent on revising the likelihood function as well. Therefore both steps are needed. Basically the referee asks us to shift the topic of the discussion paper to an entirely different direction and we do not see what this remark contributes to improve the manuscript. In our view, an analysis of the structure of the residuals obtained with an incorrect likelihood model could possibly make us draw wrong conclusions about the inadequacies of the process model. The use of a correct likelihood model is basically a prerequisite for further analysis and possible revisions of the process model.

10. Firstly, we had no “preference to constrain the parameter posterior to an area within the prior”. See also our replies to comments 7 and 8. Secondly, we fully agree with the referee in that observations of different state variables may help to parameterize the model. Unfortunately, in the present case study, there is no data of different type available on which the predictive model could be tested. Thirdly, an analysis of the contribution of the various sources of model error to the overall predictive uncertainty would be very interesting, we feel that this is clearly beyond the scope of the present study. See also our reply to comment 6.

11. Firstly, we chose this particular structure of the methods section deliberately. It puts the development of the likelihood models, which is the focus of the present study, in a prominent position, directly after the description of the process model.

HESSD

12, C1530–C1540, 2015

We personally prefer this structure over the more classic structure proposed by the referee. Secondly, the introduction to Section 2.3 introduces important concepts on which the further development of the three likelihood models is based. We feel that the definitions and explanations of these fundamental concepts given here are needed.

12. We feel that the introduction section of the discussion paper already gives a concise formulation of “research gaps and the corresponding objectives of this study”. However, we agree with the referee in that the aims of the study could be stated more explicitly. *In the revised discussion paper, we included two sentences in the last paragraph of the introduction section that explicitly state the aims of the study.*

13. We feel that this comment basically summarizes previous comments made by the referee. Please refer to our replies above.

2 Other comments

1. We refer to the results obtained by Wöhling and Vrugt (2011) in their run D4, which is the only likelihood model that accounts for autocorrelation and uses both water content and matric head observations simultaneously. We consider our citation as correct but we agree with the referee in that it is not precise enough and may lead to confusion. *In the revised discussion paper, we clarified this point. Additionally, we mention the results obtained by Wöhling and Vrugt (2011) for the case that uses a likelihood model that neglects autocorrelation (where a simultaneous fit was obtained).*

2. See our replies to comments 3 to 5.

HESSD
12, C1530–C1540, 2015

3. The minimum pressure head is occasionally reached in the model simulations. The setting we chose for this parameter is a commonly used value. The consequence is that the numerical model switches to a head boundary condition when this minimum pressure head is reached. *More information on this is contained in Scharnagl et al. (2011), and we will refer to this reference in the revised discussion paper.*

4. Readings from a nearby piezometer suggest that the groundwater table is about 3 m below the surface. A detailed justification for choosing a constant pressure head as the lower boundary condition is provided in Scharnagl et al. (2011), as also stated in the discussion paper. This choice is certainly prone to errors, and can be debated. It is motivated mainly by a lack of information concerning the lower boundary and a failure of other boundary conditions tested before (see Scharnagl et al., 2011).

5. We consider the two terms “likelihood model” and “likelihood function” as being closely related and think they can be used interchangeably in the manuscript. The likelihood function is, of course, the result of a statistical model. We prefer the term “likelihood model” in many cases because it emphasizes its nature of being based on statistical assumptions and it is thus an entity which requires careful examination and possibly needs a revision of its structure – as it is the case for the process model.

6. This comment closely relates to Major Comment 6. Please see also our reply to this comment. Since the sentence does not imply that the parameter estimates are independent of the choice of the process and likelihood models, we prefer to keep it as it is.

7. This aspect is treated in the discussion paper (p. 2180, ll. 12-26). As we argue, the next step should be a revision of the process model. This is exactly because

the results obtained with Likelihood 3 point at a critical deficit of the process model (a systematic overestimation of large water contents).

8. Firstly, we also tested a simple linear model to account for heteroskedasticity, as used for example in Schoups and Vrugt (2010). Unfortunately, with this simpler and more parsimonious model we obtained worse results, indicating that the simple linear model is not adequate. *In the revised discussion paper, we mention that a simple linear model did not perform well in a preliminary study.* Secondly, the variance of the residuals cannot be simply derived from the observations because it also depends on the model prediction. The variance of the residuals is unknown a priori. Treating this unknown as an addition parameter that needs to be estimated jointly with the other unknown parameters is a common approach in the statistical literature. With respect to a validation effort, we refer to our reply to comment 2 of Referee #3.
9. Absolutely right. We are grateful for pointing to the wrong equation given in the discussion paper. It should indeed be Eq. (12).*We changed that accordingly.*
10. Indeed. Sounds a bit odd. *We rephrased this sentence.*
11. We used our own implementation of the DE-MC algorithm, which adequately considers the prior distribution of the parameters. The prior does not only control the sampling of the initial population but is part of the posterior likelihood as clearly pointed out in the manuscript by Eq. (22).
12. We are not really sure what the referee means here with “fit of likelihoods”. We usually refer to the parameter estimate with maximum posterior density as the “best” prediction and think that this is clearly stated.
13. The interpretation of quantile-quantile plots is subjective to some degree. This is also because quantile-quantile plots emphasize the extreme deviations between

the theoretical and the actual distribution of the (decorrelated) residuals. These plots tell a lot on how well the tailing behavior of the actual distribution is reproduced by the theoretical model, but not so much about if the overall shape of the distribution was reproduced. We consider the agreement of the decorrelated residuals with the skewed t-distribution obtained with Likelihood 3 (Figure 6f) as better because only very few observations deviate substantially, while the deviations are more systematic in case of the Gaussian assumption (Figure 3c). *In the revised manuscript, we briefly discuss this issue.*

14. We think both the parameter estimates and the corresponding model predictions are rather meaningless. If the parameter estimates are meaningless, the corresponding model predictions are as well. And vice versa.

15. As we explained analytically (p. 2167, ll. 14-18), the classical AR(1) essentially always introduces bias. The only exception to this would occur if the expectation of the residuals is exactly zero, which is never the case in practice. We therefore consider our statements as valid.

16. We agree on that the value of the autocorrelation coefficient ϕ is very large for Likelihood 3. But this is exactly what can be expected (and also shown analytically) for the case of a time series of observations of high temporal resolution.

17. We consider these statements as adequate. See our replies to comment 6.

18. We fully agree with the referee in that tight uncertainty bounds do not point at a correct (accurate rather than precise!) process model. It is, however, not the larger number of degrees of freedom in Likelihood 3 compared to Likelihood 1 that makes the uncertainty bounds tight. In fact, they are even slightly larger than in the case of Likelihood 1 (Figure 7).

19. We still consider this statement as valid. Using a likelihood model that neglects autocorrelation in the residuals means that we assume independence of the

residuals, which would only be the case if the process model would be perfect. This occurs in studies using synthetic data, but it hardly ever occurs in the real world.

20. We feel that this comment is basically a repetition of previous comments. Please refer to the corresponding replies.

21. The uncertainty bounds shown in Figure 8 are the uncertainty bounds of the estimated soil hydraulic properties. The actual observations are not expected to lie within this bounds with stated probability. We chose to not include the actual observations in Figure 8, because otherwise this figure would be too crowded. The differences between the hydraulic properties obtained with Likelihood 1 and 3 would then no longer be discernable.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 12, 2155, 2015.

HESSD

12, C1530–C1540, 2015

Interactive Comment