Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2015-544-AC1, 2016

Hydrology and Earth System Sciences Discussions
Open Access

HESSD

Interactive comment

Interactive comment on "A Bayesian Consistent Dual Ensemble Kalman Filter for State-Parameter Estimation in Subsurface Hydrology" *by* B. Ait-El-Fquih et al.

B. Ait-El-Fquih et al.

boujemaa.aitelfquih@kaust.edu.sa

Received and published: 30 March 2016

- **Reply to Referee #1** -

We would like to thank the referee for his/her valuable and constructive comments. These greatly improved the quality of our manuscript and helped us clarify several points, which we gratefully acknowledge. All the referee's comments were considered in the revised manuscript and detailed replies are given below.

General evaluation

C. *This paper is of potential interest to HESS. In general, the paper is well written and organised. The results support the proposed improved methodology. I have only one major concern. This is the difference with the paper of Gharamti et al., 2015 in Journal of Hydrology. I understand that the methodology is already introduced there, and that now the mathematical-statistical basis of it is improved. In addition a new rigorous synthetic study was carried out. The authors should exactly point out what is new in this paper and motivate why this warrants a new publication. If answered satisfactory, the paper can be published with minor revisions.*

R. We thank the referee for acknowledging our work and for the comment he/she raised. The proposed Dual-EnKF$_{\mathrm{OSA}}$ in this work results from the generic algorithm of Section 3.1 by applying two random sampling properties (see Appendix A) under the Gaussian assumption. At each assimilation step of the Dual-EnKF$_{\mathrm{OSA}}$, the observed data are used three times through Kalman-like updates: twice in the smoothing step (one for smoothing the previous state as in Eq. (27) and one for updating the parameters as in Eq. (28)), and once in the "forecast" step to compute the analysis of the current state as in Eq. (35).

The work in *Gharamti et al. (2015)* follows a similar approach and applies the same two random sampling properties on the generic algorithm of Section 3.1, but under the following assumption (beside the Gaussian assumption)[1]:

$$p(x_n|x_{n-1},\theta,y_n) = p(x_n|x_{n-1},\theta), \tag{1}$$

which is based on the fact that given the previous state, x_{n-1}, and the parameters, θ, the current state, x_n, is independent of its observation, y_n. Following this assumption,

[1] Refer to Eq. (16) in *Gharamti et al. (2015).*

the "forecast" step of the generic algorithm which is composed of a Bayesian step (24) followed by a propagation step (23), reduces to the propagation step (23). As stated in page 445 of *Gharamti et al. (2015)*, the assumption (1) has been adopted to compute the analysis pdf, $p(x_n|y_{0:n})$, from the joint smoothing pdf, $p(x_{n-1},\theta|y_{0:n})$, based on Eq. (9) (or Eq. (23) in our manuscript), by avoiding the use of the computationally demanding term $p(x_n|x_{n-1},\theta,y_n)$ and replace it by the more easily sampled state transition pdf, $p(x_n|x_{n-1},\theta)$. Here, we propose a more efficient approach (see pages 10-11 and Appendix B) to directly sample from the analysis pdf without explicitly computing $p(x_n|x_{n-1},\theta,y_n)$ and without the need of any additional assumption.

Now, when applying the two random sampling properties (under the Gaussian assumption), the Joint-EnKF$_{\text{OSA}}$ in *Gharamti et al. (2015)* shares the same smoothing step as the Dual-EnKF$_{\text{OSA}}$, which, as mentioned above, involves two Kalman-like updates. However, in contrast with the Dual-EnKF$_{\text{OSA}}$, the Joint-EnKF$_{\text{OSA}}$ does not involve a Kalman-like update in the "forecast" step (because of the omission of the Bayesian step (24) from the generic algorithm). In terms of computational cost, these algorithms have almost the same cost as they require the same number of model runs; the only difference is one Kalman update for each member which is generally computationally not consequent compared to the cost of integrating the model. This further allows to explicitly put in context the conditions under which the (heuristic) steps of the standard Dual-EnKF can be derived in a Bayesian setting.

To summarise, the proposed Dual-EnKF$_{\text{OSA}}$ is more general than the Joint-EnKF$_{\text{OSA}}$ of *Gharamti et al. (2015)*, inasmuch as it involves one more Kalman-like update. This was made clearer in lines 332-343 of the revised manuscript. Moreover and as a way of illustrating the difference between the two schemes, we have included additional experiments results using the Joint-EnKF$_{\text{OSA}}$. We showed that the proposed Dual-EnKF$_{\text{OSA}}$ slightly outperforms the Joint-EnKF$_{\text{OSA}}$ ($\sim 5\%$ better accuracy). We further reported the average ensemble spread, as estimated by the Joint-EnKF$_{\text{OSA}}$, for both states and parameters in Table 1.

Detailed comments

C1. *L40: “have been proposed” instead of “has been proposed”.*

R1. Done. Thank you.

C2. *L64: “was given” instead of “was carried”*

R2. Done. Thank you.

C3. *L99-L101: Rephrase.*

R3. We have rephrased the sentence which now reads as follows: “Our goal is to derive a new Dual-EnKF-like algorithm that retains the separate formulation of the state and parameters update steps, within a fully Bayesian framework.”

C4. *L109-L110: Change to: “(...) various experiment settings and observation scenarios.”*

R4. Done. Thank you.

C5. *L122: this should be t(n-1) to t(n)?*

R5. Based on our notation in system (1), our statement in L122 is correct: “$\mathcal{M}_n$ *is a*

Printer-friendly version

Discussion paper

nonlinear operator integrating the system state from time t_n to t_{n+1}".

C6. *L203-L206: This was not found in Song et al. (2015, VZJ). There Dual EnKF performed worse, and only a rigorous Restart EnKF gave better results. Reformulate.*

R6. We thank the referee for pointing this out. In fact, we have read the article by *Song et al. (2015)* and we were aware of this study prior to our submission. The authors tested the use of the Confirming Step EnKF, the restart EnKF and a modified variant of the restart EnKF. Unlike the Joint and the Dual EnKFs, these filters only update the parameters using a Kalman-type analysis. The state ensemble members, on the other hand, are obtained after integrating the model. Comparing the Confirming-Step-EnKF of *Wen and Chen (2006)*, which we cite in the manuscript, to that of *Song et al. (2015)* may not be that straightforward considering the differences between their models. On the one hand, *Wen and Chen (2006)* worked with a reservoir simulator and *Song et al. (2015)* used an unsaturated flow problem. It is possible that the efficiency of the Confirming-Step becomes more pronounced in nonlinear reservoir systems and strongly heterogenous subsurface formations, which is not the case in *Song et al. (2015)*. To avoid the confusion, we have removed the confirming-step EnKF reference from the discussion and reformulated the sentence.

C7. *L377-L378: The pumping rate is unfortunately unrealistic low. It would have been better if the authors would have worked with a more realistic case.*

R7. We fully understand the concern of the referee. The choice of the well pumping rates was based on contributions from the initial head values in addition to the present recharge. Starting from a uniform initial hydraulic head $h_0 = 15$ m, the recharge and the pumping rates eventually yield heterogeneous spatial distribution of h, varying between

~ 13 to 20 m, as shown in Figure 3. Increasing the pumping rates in our setup may, however, cause clear suction of the groundwater (negative head) and the dynamics will be mainly dominated by this forcing term. The considered rates create enough variability in space and time to test the assimilations schemes. Thank you.

C8. *L587-L592: I do not see many differences and these are probably related to the initial conditions after the assimilation phase. Reconsider this text part.*

R8. We are not sure what the referee exactly mean by the initial conditions at the end of assimilation. In fact, after the assimilation phase we simply continue the simulation in a prediction mode without assimilating any data. Thus, we don't tamper or impose any changes on the hydraulic head obtained after assimilation. In terms of the differences, we disagree with the referee because the Joint-EnKF shows a clear overestimation of h at the control well unlike the Dual-EnKF and more importantly the proposed Dual-EnKF$_{\mathrm{OSA}}$. To illustrate, we evaluated the absolute bias during the prediction phase for the three schemes. We plot the resulting curves in Figure 1 below. As shown, the bias suggested by the proposed scheme is the smallest and around 0.5 m less than that of the Joint-EnKF. To support our argument, we have included this figure in the manuscript and we interpreted the results accordingly. Thank you.

C9. *Caption Figure 1. Change to: "The reference log-conductivity field was obtained (...)"*

R9. Done. Thank you.

C10. *Caption Figure 9. Why does AAE not decrease for joint EnKF and dual EnKF* *for small observation errors? Please comment.*

R10. The referee is raising an interesting point here. The performance of the Joint-EnKF and the Dual-EnKF clearly degrades when the observation error decreases, which might seem as counterintuitive. However, the errors only increase from 1.05 to 1.06 which is not significant. When the observational error decreases further to 0.1 m, the AAE of conductivity decreases to approximately 1.04. To test this further, we examined the performance of the filters with even smaller observation error, i.e., 0.05 m. We updated the figure in the revised manuscript, accordingly. We notice that the performance of the Joint-EnKF and the Dual-EnKF continue to improve reaching an AAE of around 1.01. Regarding the performance for measurement errors between 0.1 and 0.15, this could possibly be due to statistical fluctuations related to this particular case and setup.

C11. *Caption Figure 9: "are obtained" is not correct.*

R11. Rephrased. Thank you.

C12. *Caption Figure 10. Why do you use lines in the figures? The legend is not clear.*

R12. The figure and its legend have been updated following the referee's comment.

Please also note the supplement to this comment:
http://www.hydrol-earth-syst-sci-discuss.net/hess-2015-544/hess-2015-544-AC1-supplement.zip

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2015-544, 2016.

Printer-friendly version

Discussion paper

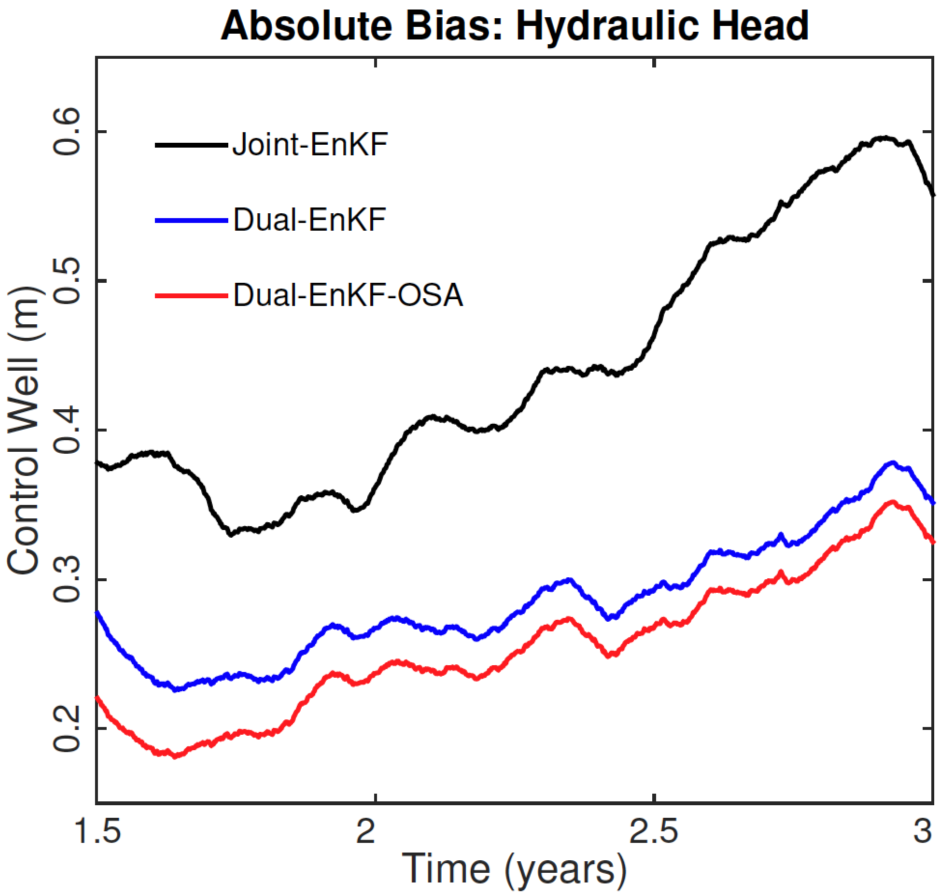

Fig. 1. Absolute bias resulting from the three assimilation schemes during the prediction phase at the control well.

Printer-friendly version

Discussion paper