Hydrol. Earth Syst. Sci. Discuss., 12, C3476–C3481, 2015
www.hydrol-earth-syst-sci-discuss.net/12/C3476/2015/

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Comparing the Ensemble and Extended Kalman Filters for in situ soil moisture assimilation with contrasting soil conditions" *by* D. Fairbairn et. al.

Anonymous Referee #1

Received and published: 6 September 2015

This manuscript presented the assimilation of near-surface soil moisture into the ISBA model at 12 SMOSMANIA sites. The SEKF and SEnKF assimilation methods are compared in a series of synthetic and real-data experiments. In the latter, in situ observations are assimilated, and evaluated against in situ observations in the deeper soil layers. An ensemble bias correction method is applied for the EnKF to correct biases introduced by the ensemble. The topic is of interest, and the authors have undertaken a lot of work to understand the inner workings of the two assimilation methods, and specifically how each responds to non-linearities in the ISBA model. However, I have some major concerns regarding the design of the experiments that must be addressed

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

prior to assimilation. Some of the presented details of the assimilation are also confused.

MAJOR:

The ensemble-bias correction method may have done more harm than good. The WG2 bias of 4 mm was fairly small in the first place, and while the ensemble bias correction method did reduce this soil moisture bias, it has introduced very large biases in the fluxes in Figure 2. For many users the fluxes are of more importance than the states. Since the problem arises largely when the ensemble perturbations generate model states outside of the usual bounds of the model (<wilt, >field capacity), it may be better to limit the ensemble members to not go outside these bounds, particularly for the field capacity (if there is no mechanism by which the model would dry the soil moisture below this point).

The synthetic experiments are incorrectly designed. The observations were generated by running the model (single member?) with perturbed precipitation, then adding a random error. The same model is then used in the synthetic assimilation experiment, and if I understand the manuscript correctly (this is not totally clear), the same precipitation perturbations were used to perturb the ensemble in the assimilation experiments. Hence the same precipitation perturbations are used to represent errors in the observations (for the generation of the synthetic obs), and then to represent errors in the model (for the assimilation). Please review the literature on synthetic experiments to redesign these experiments. The use of an observation error just 10% the size of that in the real experiments also limits the relevance of these experiments.

The manuscript assigns all of the difference in results between the sites to differences in soil class, however other differences between the sites that may affect the assimilation are not accounted for (differences in climate!). Also, it looks like the conclusions regarding soil class are made from comparing just two sites. The statements about the role of soil class need to be greatly de-emphasized (including removed from the title

HESSD

12, C3476–C3481, 2015

and abstract).

The evaluation statistics need refining, and better description: -P7367: Edit these equations to make a clear distinction between the assimilated observations, and the observations used for evaluation. -The fact that the evaluation observations are from the same network as the assimilated observations (but at a different depth) needs to be very prominently acknowledged in the text. This must include a discussion of any potential dependence between the assimilated and evaluation observations. -The ACC presented here is the correlation, not the anomaly correlation. Change all reference to anomaly correlation to correlation, and include a true anomaly correlation in the evaluation. -P7367 implies that the observations used for the evaluation were CDF-matched. This should have been clearly stated with the introduction of the evaluation metrics. There is also some discussion of the introduction biases in the RMSE due to inconsistencies in time period – why not use consistent time periods? -Also, presenting biases to the in situ data (or RMSE, which is not bias robust) for the different experiment is not very informative since the in situ data were arbitrarily rescaled. It is more usual practice to rescale each experiment separately to the in situ observations before comparing them for evaluation (thus removing the mean difference between each experiment and the in situ data).

Before proceeding to investigate the role of model physics in generating the bias, confirm that the perturbation time series is mean-zero (see comments below: re the precip perturbations). Include a note in the text that this has been confirmed.

MINOR:

I was several pages in before I realized that in situ observations were assimilated in this study. This needs to be stated clearly in the abstract and introduction.

P7355, L6. Soil moisture assimilation is not the “main objective of DA”. Many other variables are assimilated. Please rephrase.

HESSD

12, C3476–C3481, 2015

P7355, L35. Uncorrelated with what?

P7355, L27: change “with an NWP model” to “with an NWP model at Meteo-France”

P7356, L2: specify that the SEKF at ECMWF assimilated screen-level variables, and not soil moisture.

P7357, L23: Remove Calvet and Noilhan reference if they did not do CDF-matching (i.e., match the full CDF, not just the mean and variance)

P7359, last paragraph: it is not clear here why only grassland was used. Is this the land cover at all of the SMOSMANIA sites?

Include a map or table with the SMOSMANIA site locations.

P7361, L19: CDF-matching also matches the higher order moments. Please rephrase.

P7363, L5: H does not equal [1 0] for the SEFK, since there is also a time integration. Please edit.

P7363, around L10: Please make it clear here that the SEnKF could handle horizontal error correlations, while this is much harder to do in the SEKF. This is a major difference between the two methods.

Equation 4 needs time indexes for x and Delta x and M. The l superscript on Delta x is also not defined.

State in the text that equation 4 requires an extra model run for each element in the state update vector.

State at what time the analysis update is made (at the start or end of the perturbed runs for equation 4).

P7363, L19: kl superscript is not defined.

P7365: A Jacobian > 1 does not itself indicate a non-linear model. This is implied in several places, please rephrase.

HESSD

12, C3476–C3481, 2015

P7366: the presentation of the SEnKF needs equations for the analysis update (not just \del xˆa. Also the j subscript is not defined.

P7368: Since both assimilation methods are tuned using the same data as presented for the evaluation (rather than independent data) it must be prominently acknowledged that the presented results will not necessarily generalize to sites where data are not available for calibration.

L7368, P20. I can't make sense of the discussion here. The ensemble spread is used to represent the background error. This will include estimation of forcing and model errors if the ensemble is appropriately perturbed.

Also, here(and elsewhere): The references to specifying or tuning the error covariance matrix for the SEnKF is misleading, as this the matrix itself is not specified (or even estimated) for the En methods. Please rephrase throughout to avoid referring to the matrix for the En methods (and replace with something like "ensemble spread").

L7369: The precip was perturbed using Gaussian noise with standard deviation of 50%, I assume these were additive perturbations. It is not stated over what time period of precip the 50% is taken. It might be worth rethinking this – I'm concerned that the long term mean of these perturbations may not be zero, since the standard deviation of the added noise is not stationary. A lognormal multiplicative perturbation would be more appropriate. Also 50% of precip is a very large perturbation.

Equation 14: introduce notation to distinguish between the bias corrected and original xˆb

P7371, L6. I wouldn't say that these biases are unexpected (only unexpected / or inconsistent under the linear assumption). The model is non-linear and soil moisture is bounded, so these biases are to some extent expected.

Split Table 2 into separate tables for synthetic and real-data experiments.

HESSD

12, C3476–C3481, 2015

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 12, 7353, 2015.