Authors' Response to Reviews of

A comprehensive assessment of in situ and remote sensing soil moisture data assimilation in the APSIM model for improving agricultural forecasting across the U.S. Midwest

Marissa Kivi, Noemi Vergopolan, and Hamze Dokoohaki
Hydrology and Earth System Sciences (HESS),

RC: ***Reviewers' Comment***, AR: Authors' Response, □ Manuscript Text

1. Reviewer #1

RC: ***Kivi et al (2022) is cited 25 times in the text! While this work is distinctly different to that work, perhaps a single mention at the start of each section, rather than every second and third level sub-section, would be adequate (such as the existing introductory paragraph in each major section).***

AR: Thank you for the comment. As per your suggestions, we reduced the number of citations to the Kivi et al., (2022) to only cases where it was necessary in describing a method or making a comparison with earlier works.

RC: ***As far as I can tell, and I am no expert in data assimilation techniques or applications, the work presented is solid and thoroughly describes the proper use of the techniques. The authors honestly presented when the technique improved certain predictions, when it made little difference, and when the model performance degraded. While I can see why the extension was required in Figure 4b, as it's only 3 of the results and the big picture is shown well with the box-and-whiskers, the extension can be placed in the Appendices and mentioned in the text as with free-model results.***

AR: Thank you for the comment. Since constraining uncertainties is a central idea to the application of sequential data assimilation (SDA) methods -as important as reducing model biases- we initially decided to present the both metrics to show the emphasize of this feature in SDA. If possible we would like to keep the figure in the way in which it is presented . If the reviewer believes that it is necessary to make the change, we will be happy to do so.

RC: ***Was the "free run" calibrated against any data, or was it just an ensemble of model runs with the parameters randomly assigned from prior distributions? Were they a single run with an arbitrary (or literature) set of values assigned to parameters? How do the modelled results compare to the "final" set of parameters after the full SDA? I did not get the appendices so don't know if this is covered.***

AR: Thank you for the comment. Model parameters for describing crop growth and development was adopted from Dokoohaki et al. (2022), while other priors such as initial soil water content were set based on a plausible range for each variable (presented in detail in Table A.2 Kivi et al., (2022)). Within the SDA framework, no parameter was directly adjusted/optimized and main role of SDA is to adjust model state variables such as soil moisture rather than model parameters. SDA attempts to nudge the model prediction of the state variables towards observation (remote sensing or sensor) and in an ideal scenario this helps with improving system representation (which can be assessed through evaluating other state variables such as crop yield, NDVI, drainage and etc as performed in this study).

Dokoohaki, H., Rai, T., Kivi, M., Lewis, P., Gómez-Dans, J.L. and Yin, F., 2022. Linking Remote Sensing with APSIM through Emulation and Bayesian Optimization to Improve Yield Prediction. Remote Sensing, 14(21), p.5389.

Kivi, M.S., Blakely, B., Masters, M., Bernacchi, C.J., Miguez, F.E. and Dokoohaki, H., 2022. Development of a data-assimilation system to forecast agricultural systems: A case study of constraining soil water and soil nitrogen dynamics in the APSIM model. Science of The Total Environment, 820, p.153192.

There are a number of questions that the work raises, that may be answered here (or later maybe).

RC: ***As the parameters are nudged successively with the SDA procedure, how much do they need to vary until the researcher considers that (a) their a priori range is incorrect, (b) that they cannot be considered a constant, or (c) that some process is too simplified or missing in the numerical model?***

Thank you for the comment. One main difference between SDA and classical calibration methods is that SDA nudges the model state variables and not the model parameters. But a model calibration is a complementary step that could be taken in addition to the SDA. For calibrating crop parameters in this study, we used regional model parameters provided by Dokoohaki et al., (2022) and then focused on quantifying the information contribution of different remote sensing data products in this study compared to gold standard sensor SDA. Overall this study attempts to understand the value provided by soil moisture remote sensing data products for improving the prediction of other processes (drainage, N leaching, crop yield and etc) where soil sensor is not readily available. Furthermore, one of the benefits of SDA is to actually correct the model state errors due to wrong parameterization and/or lack of process representation.

Dokoohaki, H., Rai, T., Kivi, M., Lewis, P., Gómez-Dans, J.L. and Yin, F., 2022. Linking Remote Sensing with APSIM through Emulation and Bayesian Optimization to Improve Yield Prediction. Remote Sensing, 14(21), p.5389.

RC: ***Further to the case of a non-constant model parameter, is there a pattern in the parameter adjustments, e.g., always too high in winter and too low in summer, that indicates a systematic misrepresentation within the model? Is there a form of a posteriori distribution of parameter values, e.g., normal versus log-normal versus bi-modal versus uniform, that may indicate systematic model or data errors?***

AR: Although there is the possibly of including model parameters in the SDA in addition to the model state variables, we didn't explore this avenue in this study. This was mainly due to the fact that in SDA we only have a single observation for nudging model state variables and including model parameters might have resulted in over-fitting (nudging many unknowns with only one uncertain remote sensing observation). However in Kivi et al., (2022) we explored this avenue with more confident soil sensor observations and we found that adjusting the SWCON model parameter (a parameter controlling water flow between soil layers) for the two assimilation layers, though marginally helpful, did not dramatically improve soil moisture estimates as compared with no change in parameters.

RC: ***Is there parameter bias (or trends?) associated with larger underlying groupings, such as soil texture, vertical layering (duplex, gradational, uniform), crop type, or management, that indicate model structure or data limitations?***

AR: This is a very interesting hypothesis. Even though we did not explore change in the parameters, we believe potentially using a hierarchical data assimilation framework and in a augmented system representation (states + parameters), there is a possibly of including site effects in the statistical model and exploring the relationship between the site effects and different environmental factors. However, the main limitation for this idea is the availability of high quality soil moisture data for a hierarchical data assimilation scheme.

RC: ***Given the use of numerical models is primarily predictive, i.e., what are future potential grain yields or nitrate loss or deep drainage under different management or climate conditions, which set of parameters (or reduced range) do researchers consider stable enough to make such computations?***

AR: Thank you for the comment. SDA can improve initialization conditions for forecast runs, such that, although we do not have future remote sensing observations, by assimilating the observations we know that we have a well represented the initial conditions needed for runs of short term (next days, weeks) forecasts.

RC: ***The References and citations need a lot of purely technical corrections.***

AR: All following technical comments were addressed.

RC: ***Following reference not cited in text: Akhavizadegan et al (2021); Archontoulis et al (2014, 2020); Balboa et al (2019); Crane-Droesch (2018); Das et al (2020); Dietze et al (2013); Dietzel et al (2016); Flathers and Gessler (2018); Guerif and Duke (2000); Hoffman et al (2020); Jeong et al (2016); Kang et al (2020); van Klompernburg et al (2020); Leng and Hall (2020); Li et al (2014); Martinez-Feria et al (2019); Pasley et al (2021); Puntel et al (2016); Shahhosseini et al (2021); Spijker et al (2021); Wallach et al (2021).***

AR: Thank you for the comment. The reference section was accordingly adjusted to reflect only cited references in the text.

RC: ***Following citations not in references: Lu et al (2019); Lu and Steele-Dunne (2019).***

AR: Thank you for pointing this out, this error was addressed.

RC: ***The two articles of Vergopolan et al (2021) are cited with (a) and (b) in the text, but not indicated as such in the references.***

AR: Thank you for the comment. This error was fixed in the references.

RC: ***Citation for Chakrabarti et al (2014) is misspelled in the text line 83 and 605.***

AR: Thank you for the comment. This error was fixed.