Reply on RC2

The manuscript by Wang et al. presents a framework for determining data worth of soil moisture measurements. The framework uses Gaussian process regression (GP) to replace unsaturated flow models; GP is combined with EnKF to evaluate the prior, post-posterior, and posterior (regarding potential new data) distributions of variables of interest (i.e., soil moisture averaged for varying portions of the soil column). The change of distributions was summarized using three indices to determine data worth. The framework was demonstrated using three soil columns from ISMN using several test cases to illuminate the roles of prior data length, observation noise, and combinations of potential new data.

Overall the manuscript is clearly organized and results are thoroughly described and discussed. Data worth analysis in a model-free framework (using machine learning) is novel. I therefore recommend it to be published in HESS after minor revisions. Below are detailed comments, most of which are intended to improve clarity and generalizability of the presented framework.

Some discussion is needed to support the conjunctive use of GP and EnKF. EnKF is very commonly used with deterministic models such as Hydrus for data assimilation and propagation of uncertainty in time. GP is capable of assimilating newly available data by simply training the model again once new data is available and calculating the mean and covariance of a variable of interest. Given this, it seems that the data-worth framework can be done using GP alone, without EnKF. Some discussion is needed to help readers understand the framework design. For example, what is the role of EnKF in improving accuracy of data worth estimation or enabling the framework to be used for machine learning algorithms other than GP? Will EnKF result in covariance matrices different from those calculated by GP?

Answer:

We thank the reviewer for the constructive suggestions. In fact, the necessity of the conjunctive use of GP and EnKF has been discussed in detail in our previous studies (Wang et al., 2021; Wang et al., 2021). As stated in (Wang et al., 2021), on the one hand, the fusion of EnKF can effectively reduce the risk of unreasonable spatio-temporal interpolation in GP models, ultimately enhancing the robustness of such purely data-driven models; On the other hand, by combining with Kalman update, the forecast cross-covariance ($\boldsymbol{C}_k^f \boldsymbol{H}^T$) between the state ($\boldsymbol{Y}_k^f$) and the predictions corresponding to available observations ($\boldsymbol{H}\boldsymbol{Y}_k^f$) in Eqs (6-7) constrained the otherwise high error covariances of state variables at unobserved depths, which resulted in a significantly reduced uncertainty for this hybrid method relative to GP alone. To keep this manuscript more focused, we finally decided to only add a brief explanation of the conjunctive use of GP and EnKF (please see Lines 160-165), without adding extra cases with GP alone in the revised manuscript.

References:

Wang, Y. et al., 2021. A nonparametric sequential data assimilation scheme for soil moisture flow. Journal of Hydrology, 593: 125865.

Wang, Y., Shi, L., Zhang, Q. and Qiao, H., 2021. A gradient-enhanced sequential nonparametric data assimilation framework for soil moisture flow. Journal of Hydrology, 603: 126857.

It is unclear from my reading (1) what depth(s) are used as prior data for training GP, (2) what specific depth(s) are considered for potential data (theta_s, m, d), and (3) do the depths for prior data and potential new data overlap?

Answer:

We are sorry for the confusion caused by our unclear description. (1) Prior data for training GP includes the soil water content at all observed depths during the prior stage (from $t=1$ to $t=T_p$), i.e., z=0.08, 0.15, 0.30, 0.45, 0.60, and 0.90 m at Falkenberg, z=0.05, 0.10, 0.20, 0.50, and 1.00 m at Cape, and z=0.05, 0.10, 0.50, and 1.00 m at the DAHRA. (2) The depth of the potential soil moisture data is different in different test cases. For example, as listed in **Table 2**, the potential data in TC1-1, TC2-1, TC3-1, TC4, TC5, TC6, and TC7 refers to soil moisture in the surface layer (θ_S), i.e., z=0.08 m at Falkenberg and z=0.05 m at Cape and DAHRA. (3) The depths of the prior and potential new data in this study partially overlapped due to the limited depths of observations under real-world circumstances. For example, TC1-1 at Falkenberg used soil moisture observations taken at six depths (z=0.08, 0.15, 0.30, 0.45, 0.60, and 0.90 m) over the first 80 d as prior data, and the generated soil moisture at z=0.08 m over the last 20 d as potential data. We have added the relevant descriptions in the revised manuscript (please see Lines 340-355).

Line 305 - how is the noise level used in the computations? For GP or for KF? Is the noise level specified or estimated when training GP?

Answer:

Thank you for your carefully reading. Noises from soil moisture observations are considered in both GP and EnKF in this study. At any time step $t=k$ during GP modelling, the observed time series from t=1 to (k-1) are corrupted by the prescribed observation noises satisfying Gaussian distribution to obtain N sets of training data. Subsequently N sets of GP models are constructed independently, to generate $\boldsymbol{Y}_k^f = [\boldsymbol{y}_{k,1}^f, \boldsymbol{y}_{k,2}^f, \dots, \boldsymbol{y}_{k,m}^f, \dots, \boldsymbol{y}_{k,N}^f]^T$ in Eq. 7 (please see Lines 175). In the analysis stage of EnKF, the real-time observation $\boldsymbol{d}_k^{obs}$ perturbed by the specified noise was assimilated via Eq. 8 (please see Lines 230-235). Considering the difficulty of determining the observation noise under real-world circumstances, the noise level is artificially specified in this study. We have added the relevant explanations in the revised manuscript (please see Lines 315 and 369).

Line 375 - I don't have a specific comment here, but would like to highlight that the difference in data worth between a physical model and a machine learning model is very interesting and a key contribution of this study.

Answer:

Thank you for your valuable recognition. Considering that the data-worth analysis in physical models has been discussed in detail in our previous study (Wang et al., 2018), this study did not add the corresponding test cases, but directly compared the findings in (Wang et al., 2018) with the results of the proposed NP-DWA. As you mentioned, the comparison in data-worth between physical and machine learning models can help modelers better understand the impact of the ways data being utilized on its worth. We have accepted your suggestions and further highlight this difference in the revised manuscript (please see Lines 20-30, 420-425, 565-575, and 640-650).

References:

Wang, Y. et al., 2018. Sequential data-worth analysis coupled with ensemble Kalman filter for soil water flow: A real-world case study. Journal of Hydrology, 564: 76-88.

Predictions (variables of interest) considered in this study are depth-averaged soil moisture. It would be of more interest to the broader hydrologic community to discuss the potential application of the presented framework for other types of predictions, e.g., ET, infiltration.

Answer:

Thank you for your constructive suggestions. This study used the GP model to reconstruct the nonlinear relationship between multiple variables (including time, depth, precipitation, and air temperature) and soil moisture. Therefore, the expected data-worth of future monitoring programs regarding the estimation of depth-averaged soil moisture can be evaluated. Our future study will further discuss the application of the NP-DWA framework for other types of predictions, e.g., ET, infiltration.

Line 365: “matrixes” should be “matrices”

Answer:

Thank you for your carefully reading. We have revised “matrixes” to “matrices”.