Review of paper

'Improving operational flood ensemble prediction by the assimilation of satellite soil moisture: comparison between lumped and semi-distributed schemes

by C. Alvarez-Garreton, D. Ryu, A. W. Western, C.-H. Su, W. T. Crow, D. E. Robertson, and C. Leahy

Dear Authors, dear Editor,
I have reviewed the aforementioned work. My conclusions and comments are as follows:

1. **Scope**

The article is within the scope of HESS.

2. **Summary**

This study investigates the benefits of assimilating satellite-derived soil moisture data into ensemble streamflow predictions in sparsely monitored catchments at the example of the semi-arid 40.000 km² Warrego catchment in Australia. To this end, a conceptual hydrological model (PDM) is calibrated by streamflow observations at a single outlet gauge in a lumped and semi-distributed (7 subcatchments) configuration. For 3 components of the model (rainfall input for forcing, a conceptual store retention constant for parameters and the soil water storage for states), error models were formulated and their parameters found by calibration. The satellite observations were transformed to estimates of profile soil moisture, bias-corrected and then assimilated into the model predictions with an Ensemble Kalman filter approach. Model performance was then evaluated a) lumped vs. semi-distributed and b) unperturbed model vs. open loop ensemble vs. updated ensemble predictions (updated by data assimilation of satellite-derived soil-moisture estimates) by normalized RMSE (NRMSE), Nash-Sutcliffe efficiencies (NS), probability of detection (POD) and false alarm ratio (FAR). The authors' main conclusions are:
- The main limitation of the proposed ensemble model scheme was the too large ensemble spread of the open-loop model predictions
- Estimation of profile-average (here ~ 1 m depth) soil moisture from satellite data is difficult and involves large uncertainties
- Nevertheless, the data assimilation predictions outperformed the open-loop ensemble predictions, the data assimilation predictions with the semi-distributed model outperformed those based on the lumped model, and data assimilation predictions at uncalibrated gauges within the catchment outperformed the open-loop predictions.
The authors conclude with the recommendation to focus efforts on ensuring adequate hydrological models given the available data.

3. **Overall ranking**

The work is ranked **'major revision'**.

4. **General evaluation**

This is a thoroughly conducted study, where all of the (many) crucial assumptions that needed to be made on the way (e.g. the choice of error models or the profile soil moisture estimation) are discussed. However, I have several concerns about the explanatory power of the study:

- As the goal is to evaluate the additional value of assimilating remotely-sensed soil moisture data, it should be done with an 'as-good-as-possible' hydrological model, i.e. a model that has been set up making as good as possible use of the standard available data. I assume the author's last statement in the conclusions points exactly in this direction. So instead of evaluating the benefit of satellite-derived soil-moisture data assimilation (SM-DA) against a lumped model and a semi-distributed model that was calibrated with the outlet gauge only, it should be evaluated against a model calibrated on all 3 gauges. Also, the performance of the hydrological model could potentially be considerably improved by either improvement or calibration of the evapotranspiration (ET) module, especially so as this study focuses on soil moisture states, which are strongly influenced by the ET scheme. So please include a description of the ET module, explain whether it has been used in calibration or not (and if not consider using it for calibration), explain its influence on the quality of the model predictions, and also consider including it in your model error model. Then conduct the evaluation of SM-DA again with the 'as-good-as-possible' hydrological model.

- So far, the semi-distributed, non-ensemble model outperformed all others (NS = 0.77). So as a flood forecaster, given the choice of all presented model configurations (lumped/semi-distributed, open-loop/assimilation), I would choose the former (even though I recognize the additional benefit of a probabilistic prediction). It may well be that with the 'as-good-as-possible' hydrological model mentioned above, this may be even more the case. So for me the main message of the study is to focus on the set-up of the hydrological model and the associated error models rather than SM-DA, if better (and probabilistic) stream flow predictions are the goal. The authors have mentioned this in their study, but it should be stated more clearly. Also, for the study this means that before applying SM-DA to the model, the focus should be on improving the error models of the hydrological model (which, as the authors correctly state, is a highly underdetermined problem), until the performance of the open-loop model ensemble mean is comparable to that of the non-ensemble model. So far, it is considerably worse (NS 0.61 and 0.53 for lumped and semi-distributed model).

More specific points are (page/line):

- 10641/1-5: what is the spatial resolution of the satellite data? Also, it is not clear yet what the satellites actually observe (penetration depth etc.) → please include a reference to section 3.5
- 10642/14: Are the k-parameters really time-dependent, i.e. time-variable?
- 10645 pp.: Please justify in more detail the choice of your error functions, especially for k

Yours sincerely,
Uwe Ehret