Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-589-RC2, 2019

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Observation operators for assimilation of satellite observations in fluvial inundation forecasting" *by* Elizabeth S. Cooper et al.

Anonymous Referee #2

Received and published: 17 January 2019

The paper compares three observation operators in the framework on Ensemble Kalman filtering when SAR observations are assimilated in a hydrodynamics model. This is very interesting piece of work and clearly deserves to be published in HESS as it deals with the use of innovative and rich data and aims at correcting model parameters and state to improve flood forecasts. The paper is very well written, the state of the art well described and the results presented in a clear manner in the context of synthetic experiments when the control vector is defined as the state vector or augmented with the model friction coefficient. However, a number of issues should be address before the paper can be considered for publication. Some details on the data assimilation algorithm, especially the cycling should be added and some clarification

on the observation operators should be given. The major comments are listed below.

Introduction

- References from Oubanas should be included in the state of the art when referring to hydrodynamics data assimilation, with remote sensing data.

Oubanas, H. : Variational assimilation of satellite data into a full saint-venant based hydraulic model in the context of ungauged basins. Ph.D. thesis, Institut National des Sciences Appliquées de Toulouse (2018)

Oubanas, H., Gejadze, I., Malaterre, P.O., Durand, M., Wei, R., de Moraes Frasson, R.P., Domeneghetti, A. : Discharge estimation in ungauged basins through variational data assimilation : The potential of the swot mission. Journal of Hydrology pp. 2405–2423 (2018)

Oubanas, H., Gejadze, I., Malaterre, P.O., Mercier, F. : River discharge estimation using variational data assimilation involving the full saint-venant model and synthetic swot-type observations. Journal of Hydrology pp. 405–430 (2017)

- Front like observations were assimilated in the framework of wildfire propagation. Coordinates (x- and y-) of along edge markers were used as observation, avoiding the non-gaussian issue of binary data (burn or unburned area, dry or wet area). This work should be cited in the references as it proposes an alternative to the 3 options presented here.

Rochoux, M. : Vers une meilleure prévision de la propagation d'incendies de forêt : évaluation de modèles et assimilation de données. Ph.D. thesis, Ecole Centrale Paris (2014)

Rochoux, M., Collin, A., Zhang, C., Trouvé, A., Lucor, D. : Front shape similarity measure for front position sensitivity analysis and data assimilation. ESAIM : Proceedings and Surveys 63, pp. 215-236.

Rochoux, M., Ricci, S., Lucor, D., Cuenot, B., Trouvé, A. : Towards predictive data-driven simulations of wildfire spread - part 1 : Reduced-cost ensemble kalman filter based on polynomial chaos surrogate model for parameter estimation. Nat. Hazards and Earth Syst. Sci. 14(11), 2951–2973 (2014)

Part 2

- ETKF algorithm should be presented in more details even though this is a classical algorithm and references are given. The choice of the perturbation matrix is essential in this deterministic filtering algorithm and for the present paper to be self-dependent, a short description of how this is done should be included. - Please justify why using a deterministic filter ETKF instead of a stochastic EnKF?

Part 3

- Part 3.1 should include Figures 7 and 9. - Assuming that the water level is constant perpendicularly to the flow is essential for the second observation operator hnp. While this was mentioned at the end of 3.2 relating to other published papers, this should be mentioned earlier when presenting the operator along with the related difficulties (such as finding the nearest point). - For the backscatter approach, operator hb associates either md or mw to the model equivalent at the observation point. This means that operator hb only returns 2 possible values that are compared to the entire range of backscatter values. As a consequence, any wet pixel in the model state (for instance different WLO values for different members) would return the same equivalent, and the difference between members is lost. I feel, we are losing information in the ensemble here. Yet, I may be missing a point here, so please clarify.

Part 4

- While I am not questioning the use of the Clawpack model in this work, I am curious to know why the authors did not use a community model such as LISFLOOD, MIKE or TELEMAC. Some details on Clawpack model may be included here again for the pa-

per to be a little more self-dependent: is it a full 2D model? (is the water level constant perpendicular to the flow direction ?) How are the limits of the simulation domain prescribed? (solid boundaries?) - The ensemble construction relies on the perturbation of a true inflow with additive time correlated signal, assuming that the correlation length is large. Why not simply using a scalar additive perturbation constant over time, as it basically comes down to the same result? - Cycling of the analysis requires explanations on how the friction coefficient is updated along the analysis. First, it is not clear to me whether the analysis is carried out at an instant of observation or over an assimilation window (like a smoother would be). Secondly, it is mentioned that the friction coefficient is drawn from a normal distribution with mean different from the true friction coefficient. But how is the analysed value of the friction used for the following cycle? Is the friction drawn from a normal distribution with a mean equal to the mean analysis? How about the standard deviation? Is there any inflation on the model parameter to avoid ensemble collapse? Is the corrected value of the friction kept persistent for the forecast? I suggest adding a scheme to properly explain the ensemble cycling in part 4.3. - In the synthetical observations part 4.4, I understand that given the WLO in a flood edge pixel, a backscatter value is drawn from a normal distribution centred in md or mw. Why bother computing the Gaussian fit and new Gaussian values when these observations are going to be compared to binary values (equivalent model state values that are either md or mw)? - The location of the observation is not clear to me: in 4.4, it is said that the flood edge is defined to be the elevation at the first ‘dry’ pixel encountered when moving in a perpendicular direction from the centre of the channel along one of the defined cross sections then it is said that two observations per flood edge are considered. Please clarify and locate the observations in Fig 3. - I couldn’t find information of observation frequency while it is mentioned that the assimilations are carried out every 12 hours. This goes back to my previous question on instantaneous or time window assimilation.

Part 5

- I suggest adding rank diagram to check the validity of the ensemble with regards to the observation. This would be a starting point to identify cases where all members WLO are lower than observation as this case leads to a problematic zero correction in the analysis. This would allow for correcting the ensemble (and its spread) beforehand applying data assimilation while being aware of a problem. - The computation of the ensemble mean at the flood edge illustrated in Fig 8 causes a negative effect of the analysis because the members that are shallower than the observation are associated with a zero WLO at the flood edge, thus not contributing to the mean computation. I regret that no solution was proposed in the paper. A suggestion would be to compute the mean WLO at the centre of the river and assume it is constant perpendicularly to the flow. This assumption is made already for the second operator solution. - I have doubts about Figure 9: the observation is located at the true flood edge, where the innovation is computed. I guess the arrow in the observation space should be translated on the left, above the flood edge. Plus, md and mw are mixed in the right hand side legend. - The results for hb are satisfying while difficulties occur when all members are shallower than the truth or reversely. I regret the authors did not propose an alternative to this while being aware of it. I suppose that in a real case scenario, this situation may occur depending on how the ensemble is generated. Thus, I suggest adding ensemble validity check as well as reconsidering the computation of the model equivalent that binaiyly returns md or mw indistinctly of the water level value.

Part 6

Locating the flood edge seems a difficult task for a real case with a randomly shaped flood surface, also, locating the nearest wet pixel is a complex task in non-idealized cases. I suggest to investigate this topic that is central to all 3 observation operator proposed here.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-589, 2018.