Geosci. Model Dev. Discuss., 2, C142–C145, 2009
www.geosci-model-dev-discuss.net/2/C142/2009/


Geoscientific Model Development Discussions



# *Interactive comment on* "Streamflow data assimilation for soil moisture analysis" *by* K. Warrach-Sagi and V. Wulfmeyer

**Anonymous Referee #1**

Received and published: 7 August 2009

This paper looks at assimilating streamflow data into a land surface model for the purpose of updating soil soil moisture. An observing system simulation experiment (OSSE) is devised to see if such a method is feasible.

I have several serious concerns with this paper which the authors need to amend before it is considered for publication. These matters are outlined below.

Firstly, the explanation of the method is not clear at all. Equations 1-5 do not really give the important aspects of the assimilation scheme, namely how the retrospective filter was implemented. The method would be very difficult to reproduce without many guesses and assumptions. For example it is said on page 560 that "too long timeseries are assimilated" - does this mean that you are assimilating more than one observation per update step?



Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

How large is the state vector? My reading of the paper suggests a vector with a length of at least 175680 (6 soil layers, 244 grid boxes, 60*2 half hour timesteps) and then add in the streamflow gauge points. If numerous observations are assimilated at once we need to add more model states (6*244). Is that right? If not, please explain the dimensions of the matrices and vectors. Please show, either mathematically or schematically, how the assimilation was done.

Why is the A matrix needed? The analysis error variance is propagated via the ensemble, it doesn't need any sort of error model to push the analysis variance to a new background variance. Why is a "tangent linear observation operator matrix" needed? Shouldn't the routing model equate the model quantities to the observation thus the H matrix is simply a set of 0's and 1's depending upon the time lag of elements in the state vector?

Why do you show histograms and spatial plots of soil moisture for t=0 in the figures? If this was a forecast system you would assimilate data up to the start of the forecast period and the final update would be used to initialize the COSMO-TERRA model. The final update is the one that should be shown.

However, what I would like to see is a timeseries of soil water content for a single layer (or if that is not explicitly calculated, for the soil column) over the period of the assimilation experiment. I would like to see the original control run and all the ensembles or at least the mean of the ensemble. This will give a much better idea of how the assimilation is functioning.

Page 553 line 22: "climate simulations rely on proper root zone soil moisture initialization". The wording used here should probably be changed. Climate is generally a boundary value problem, while weather and seasonal prediction is an initial value problems. The word "climate" is being used to refer to seasonal length cases; typically climate refers to numerous years.

Page 554 line 11: should probably read: "soil temperatures do not necessarily match



GMDD

2, C142–C145, 2009

reality."

P554 line 14: "It becomes a problem when improving parameterizations ..." What becomes a problem? This sentence could do with some clarification.

The text on page 554 raises a philosophical problem that does not seem to have been addressed by the authors. It is said that both Hess (2001) and Drusch & Viterbo (2007) performed assimilation to update soil moisture in NWP models. Apparently in both cases (I only read the Drusch & Viterbo paper) the forecast skill improved yet the simulation of soil moisture degraded in the assimilation. This suggests that either:

1) the structural or parameter deficiencies of the model overwhelm the initial value improvement or 2) incorrectly equating observations and model states

Indeed, Gupta et al., (1999) showed that when a land surface model was calibrated to produce very good simulations of soil moisture, the errors in surface energy fluxes (latent & sensible heat) increased dramatically. Ultimately it the the enrgy fluxes that are of most importance to NWP, not the actual soil moisture (as would be the case with streamflow forecasting).

The premise of the paper is to improve NWP by providing better soil moisture estimates via assimilation, but the literature cited by the authors suggests that unless the full NWP scheme is used, "improvement" in the soil moisture state may not necessarily aid forecasting.

Gupta HV et al., (1999) Parameter Estimation of a Land Surface Scheme Using Multi-Criteria Methods, JGR, 104 (D16), 19491-19503

With regard to Komma et al, (2008) it is stated: "However, they use a simple soil moisture model which is not complex enough as land surface model for atmospheric model". I would say that this comment is unfair. In the current paper tha authors only look at total soil water content, exactly the same as Komma et al. W-S and W do not mention how their observational operator works and what thresholds (if any) they apply to dis-



GMDD

2, C142–C145, 2009

Interactive Comment

tributing the updated soil moisture across the soil layers of TERRA-ML. Had the study looked at differences in evaporation or other surface fluxes it might be possible to say that this paper "go[es] a step further and study the potential of streamflow data assimilation for soil moisture analysis in a catchment, namely for initialisation of numerical weather prediction and climate models."

Figure 3: the data does not seem to add up. For the 260kmˆ2 Pforzheim basin we are told that 0.002kg/mˆ2 of runoff is applied and "no more runoff is assumed afterwards", thus the total volume of water is: 0.002 kg/mˆ2 * 260 *1000*1000 mˆ2 = 520,000 kg or 520 mˆ3 The data in Figure 3 shows that a peak of about 9 mˆ3/s is reached at about 33 hrs and finally goes to zero at about 60 hours. To simplify the calculation and make sure I am conservative in my thoughts, let us say that flow goes from 0 to 9 mˆ3/s in 25 hours and then back to 0 in another 25hrs. integrating this we get : 9 mˆ3/s * (25 * 3600) s = 810,000 mˆ3 of water. Have I missed something simple here?

It sounds like there was no difference in the precipitation inputs across all the ensembles, which means that eventually the ensemble will converge to the same answer once the initial soil moisture perturbation has dissipated. Thus as the simulation proceeds it will move closer and closer to the CONTROL simulation. Was different data used to drive each model ensemble - if not, why not?

The caption for Fig 6 says "The SWC of the is displayed for the initial time t=0 (5 May 1997) for the background, CONTROL and analysis assimilating streamflow from the CONTROL model simulation from Pforzheim from t=0 to t=48h with a 0.5 hourly timestep" However, the text on p562 says: "Figure 6 shows the spatial distribution of the SWC at time t=0."

There is a very large amount of additional explanation needed as well as a much clearer demonstration of the method and its success if this paper is to be published.

Interactive comment on Geosci. Model Dev. Discuss., 2, 551, 2009.





Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper