Hydrol. Earth Syst. Sci. Discuss., 9, C3992–C4000, 2012
www.hydrol-earth-syst-sci-discuss.net/9/C3992/2012/

Hydrology and Earth System Sciences Discussions

Interactive comment on "A dual-pass data assimilation scheme for estimating turbulent fluxes with FY3A data" *by* T. R. Xu et al.

Anonymous Referee #2

Received and published: 24 August 2012

This paper describes experiments assimilating IR-derived land surface temperature observations into the Common Land Model at 4 sites in China. A dual pass filter is used, in which the first pass updates the model parameters, and the second pass updates the model soil moisture. The results are compared to surface flux observations obtained with two measurement systems. This is a an interesting and highly relevant area of research. In particular, the assimilation of surface temperature observations has not received sufficient attention to date, and the use of observations to refine the model parameters as well as the state vectors is very promising. However, I have some major reservations regarding the work presented, including some concern regarding the author's understanding of the land surface model used, the originality of the work, and the appropriateness of the assimilation technique and analysis of the results. I

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

recommend that it only be publishable if these concerns can be addressed by major revisions. Please see below for specific comments.

* MAJOR:

* Equations 12-15, describing the CoLM model have obvious inaccuracies:

Equations 12 and 13 describe a surface layer that has no capacity to store heat (right hand side equals zero), and hence cannot have a temperature. However, these experiments have assimilated surface temperature observations, and equation 16 describes the temperature of the model surface layer. Something here is very wrong. I am not familiar with CoLM, but my guess is that equations 12 and 13 became out-dated when the big-leaf model was introduced in 2004 (P8512, L5).

Equations 15 describes a model with no soil surface or canopy resistance. While the absence of surface resistance for bare soil evaporation in CoLM is discussed later in the paper (P8510, L20), CoLM certainly includes a canopy resistance term for transpiration (which is far more relevant than the soil surface resistance, discussed on P8510, since the experimental sites are vegetated).

Please correct the above errors, check *all* details of the model carefully for other errors, and specify in the text which version of the CoLM model was used.

* Lack of originality

This work is very similar to two previous papers by the same authors (Xu et al 2011a, and Xu et al 2011b) - the differences are limited to tweaks to the assimilation method and selection of the assimilated data set. It is difficult to judge the originality of this work (which I have some serious doubts over), without a better comparison between this work and the previous studies. Please include some discussion motivating why the changes from the previous studies were necessary, and then comparing these results to previous results (i.e., whether the changes to the methods yielded the desired improvements).

HESSD

9, C3992–C4000, 2012

For example, one of the main differences between the different papers is the use of different LST data. However, the error estimates (correlations and RMSE) for the FY3A data in Section 3.2 indicates less accurate observations than the LST data sets previously assimilated in Xu et al 2011a and Xu et al 2011b. Why was this data used then?

* Analysis of results

Figure 3 shows a bias between the assimilation and simulation results, and P8510, L10 states that the assim has corrected the biases (and the reduction in RMSE will mostly be due to this reduction in bias). However, the assimilation system used is bias-blind (i.e., designed only to correct random errors). If the intention is to correct biases, then you must design the assimilation system to do this - see de Lannoy et al 2007. Also, given the differences in scales (and lack of closure) in the evaluation data set, the bias between the model and the observations could easily be due to representivity / definition differences between the two - that is, you cannot assume that it necessarily indicates that the model is wrong. Please repeat the analysis focusing on more suitable statistics (correlation is the obvious choice). Same comment for soil moisture results in Section 5.3

De Lannoy, G.J.M., Reichle, R.H., Houser, P.R., Pauwels, V.R.N., Verhoest, N.E.C. (2007). Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter. Water Resources Research, 43, W09410, doi:10.1029/2006WR00544.

The use of "analysis error" (the spread of the ensemble after the analysis update) to evaluate the assimilation success is not appropriate. The AE represents the EnKF estimate of the accuracy of the analysis, however it is a very large assumption to assume that this provides a reasonable estimate of the true analysis error. It is more a reflection of the assimilation set-up than the true accuracy of the assimilation output (e.g., the AE is decreased by the assimilation by design, so showing that the assimilation decreases AE is redundant). An analysis of the background departure (obs-model) would

HESSD

9, C3992–C4000, 2012

be better - while reduced background departures does not necessarily indicate an improvement, it does show that the assimilation has improved the ability of the model to forecast future observations.

* Clarity of methods

In general, the expression in the paper is very good although there are a number of phrases that are unclear/awkward. The impact of the paper would be enhanced by a careful editing of the language. Of more concern, the methods applied are not very clear. Specifically, the timing of the parameter update pass and the state update pass of the assimilation is not made clear, and the model parameter sensitivity analysis method is difficult to interpret. Also, soil moisture is updated from surface temperature observations (although I couldn't find reference to which layers are updated - or even how many layers the model has). This is not the obvious choice of update vector, and needs to be justified.

MINOR:

Title: Most people are not familiar with FY3A. Specify the type of data in the title instead.

P8495, L10: Kanemasu et al is not listed in the references.

P8496, L15: Is the argument backwards here? I would expect the dominant dependency is for land surface temp to control humidity, not the other way around.

P8496, L20: Change "biases" to "errors"

P8497, L5: Change "obvious improvements were found" to "greatest improvements in the land surface fluxes were found"

P8497, L15: "The results showed that soil moisture play more important roles than soil temperature in turbulent flex prediction". This statement is far too broad - my understanding is that the cited study found that, for their experimental set-up for assimilating

HESSD

9, C3992–C4000, 2012

surface temperature, updating soil moisture produced better results than updating soil temperature. Please re-phrase.

P8497, L20: change "proved" to "obtained"

P8497, L25: "It is physically unreasonable to optimize both model states and model parameters at the same temporal scales". Please provide references, or more justification for this statement.

P8498: Please include information here on what type of satellite FY3E is, and what type of observations are being assimilated.

P8498, L10: I don't think "EC" or "LAS" have been defined in the main body yet.

P8499, L5: Delete sentence "With the optimal model parameters and soil moisture, we assume the optimal turbulent fluxes are predicted." . This assumption is almost certainly incorrect, and not really necessary. Also the expression in the paragraph above is quite difficult to understand.

P8499, L25: change "first-guest" to "first guess"

Equation 3: I would remove equation 3, since it is generalised by equation 4. Also, Q is not indicated as a matrix in equation 3.

P8500, L10: Please provide the actual soil moisture errors used in the text (in addition to the reference). Also provide all details on the generation of the ensemble - which forcing / params / model states are perturbed, and by how much.

P8500, L10: parameter uncertainties of 10\% of the default seems pretty naive to me. Please improve on this, or provide more details from the reference as to why this choice is justified. At P8509, L10 it is highlighted that this approach is unreasonable, since it leads to an unrealistically large range in one parameter.

P8501,L15: Replace "proved" with "demonstrated"

HESSD

9, C3992–C4000, 2012

Section 2.1 Please indicate in this section exactly which variables are in the state update vector. This should also be clearly stated in the introduction and abstract.

The mechanism of the dual pass algorithm is not clear. In Figure 1, specify what "long time" and "short time" refer to, and somewhere (possibly by completely redesigning Figure 1) indicate how, how often, and when the updated parameters from pass 1 are used in pass 2 (i.e., it is not clear whether the system runs over the time period twice -once for params, once for state vector- or whether the params from the last period are used - or some combination of these options).

P8503, L20. indicate whether monthly, annual, or longer-term MODIS LAI averages have been used.

P8504, L25: How have the results from using EBR to assess the energy-imbalance? The numbers are given, but this information is not analysed at all.

P8504, L25: Give details (in a few words), describing the method of Yang and Wang

P8505, top line: why was the soil heat flux not obtained at the site?

P8505, L10: Please explain why the night time fluxes were excluded.

P8506, L10: If this is possible, a very brief (two sentences) description of the local split window method would be useful here. Also, it not totally clear which FY3A observations were used to get LST - was it just two IR channels?

P8507, top line: 1-5 K is a very large range. What were the exact values used for R?

P8507, L5, and L8513, L5: : delete "as we all know"

P8507, L15: fix "effect contribution"

P8508, equation 20. State what V is.

P8508, equation 21: It is not clear what the subscripts indicate here. Is the S for some combination of i_1, i_2, ... or should it read simply S_i = ... for each i.

HESSD

9, C3992–C4000, 2012

P8508, equation 22: State what the tilda indicates.

P8508: why was a different method used to identify the parameters to be updated than was used in previous studies? How do the results differ? And why?

P8509, L5: There are only 4 sites in this study. Please repeat the sensitivity analysis at all sites.

Figure 3: These plots are far too small to be easily interpreted. Please increase the size to ensure all features are clear.

Section 5. All results should be presented for the full time period for which the experiment is covered. Otherwise it looks too much like the evaluation period has been tailored to give the best results. If there are significant gaps in the evaluation data, then also include some statistics for the number of evaluation data used in each analysis.

P8509, L10, L5 : Please use the full parameter name in the text (all occurrences)

P8509, L5: It is not clear in the text what range of values were used for z0m and displa.

P8509, L10: " The hhti parameter means 1/2 point of high temperature inhibition function". This phrase is not clear.

P8510, L5: This is the day of year, not the Julian Day. Replace all instances of "Julian Day" with "Day of year"

Section 5.1: Have the parameters been optimised for the results presented in this Section?

P8510, L20: The last sentence on this line is not at all clear.

P8511, L15: Replace "ration" with "ratio"

P8511, L20: "daily averaged EF is obtained by averaging the data from 10:00 to 15:00". It is not a daily average if it is based on only a 5 hour period. Please present the data for the full diurnal cycle.

HESSD

9, C3992–C4000, 2012

P8512, top line: Rephrase "The representative of EC is usually 1 km". Same line "the model usually predict turbulent fluxes at a larger scale" - replace this with a precise statement indicating the scale at which the model has been run (this information should also have been given with the model description).

P8512: stating that LAS "avoids the energy imbalance problem" is misleading. It does not measure all terms of the energy balance, and so the balance cannot be tested. This does not "avoid" the problem, it just means that the problem cannot be detected.

Section 5.2 Please provide a brief analysis of the agreement between the LAS and EC observations, as this will help demonstrate the accuracy of these observations.

L8513, L5 "need to be retrieved using ..". This sentence is unclear, and possibly not justified (the dual-pass assimilation does not *need* to be done)

P8513, L15: If I understand correctly, it is argued here that the bias is the result of an incorrect porosity in the model. Shouldn't this have been identified by the parameter optimisation. If it wasn't, why not?

P8513, L25: it is not quite clear what has been done here. Were the "stable parameters" (temporal and spatial average) used in place of the default?

Figure 8 -explain what the error bars are. Also provide some explanation in the text regarding the results for binter (which is only changed once).

P8514, L15. State what the "short" and "long" time scales are.

Figures and Tables

Where observations are included, state which observations they are (assimilated, evaluation (LAS or EC), etc)

Figure 6 - what is the time period plotted? Indicate in the caption what variable is plotted.

HESSD

9, C3992–C4000, 2012

Some figures have T_G, which is not defined anywhere, and T_S is used in the tables. Please be consistent.

Table 4 and 5 - use the same notation for assimilation / no assimilation

Combine Tables 4 and 6 - then the relative impact of updating the state variables, and updating the parameters can be assessed.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 9, 8493, 2012.