Hydrol. Earth Syst. Sci. Discuss., 9, C4577–C4591, 2012
www.hydrol-earth-syst-sci-discuss.net/9/C4577/2012/

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.

T. R. Xu et al.

xutr@bnu.edu.cn

Received and published: 6 October 2012

Response to reviewer #2 comments on "A dual-pass data assimilation scheme for estimating turbulent fluxes with FY3A data"

Thanks for your comments and recommendations to help us improving our manuscript and organizing our paper. Please find below our responses:

Major Comments:

Comment: 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 exper-

iments 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.

Response: Thanks for your comments. In CoLM, the two big-leaf model was introduced, and we just summarize the model in this manuscript. In this study, the latest version of CoLM was used to construct the data assimilation scheme (Dai et al., 2003, 2004). We will correct the description of the model in the revised version. For equation 12, the soil heat flux that produces soil surface temperature is put in the left hand side, and we will revise it. In CoLM, the canopy heat storage is not considered, and the canopy temperature can be calculated using the longwave radiation from canopy. For equation 14-15, we also summarize the sensible heat flux and latent heat flux calculations, and we will revise it. In short, we will check all details of the model descriptions, and make it more clearly.

Comment: 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). 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?

Response: Thanks for your advices. This manuscript looks similar with Xu (2011a, b), but they are different actually. The main novelty is the data assimilation technique. In Xu (2011a), the data assimilation scheme only updates soil moisture to improve turbulent flux predictions by using MODIS data. However, turbulent flux predictions are affected not only by model states but also model parameters. Therefore, in Xu (2011b), the model states and parameters are simultaneously optimized by using GOES data. In Xu (2011b), they have compared the results with Xu (2011a). Since the model states and parameters vary at different temporal scales, it's more reasonable to optimize the model states and parameters with two loops than optimize them simultaneously. That's why we develop the dual-pass data assimilation technique in this study. Actually, both the data assimilation schemes, in Xu (2011b) and in this study, can produce significant improvements in turbulent flux simulations and both of them are excellent schemes. In this study, we can compare the results produced by these two data assimilation scheme. In this study, FY3A LST data were assimilated into a land surface model for the first time to improve model predictions. FY3A satellite has been launched in the year 2008 by the Chinese government. More satellites like FY3A will be launched in the next time. The aim of assimilation data from these satellites is to open up a new remote sensing data source for land surface process researches.

Comment: 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 biasblind (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 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.

Response: Thanks for your comments. In this study, we assume the model biases are mainly from the uncertainty of model parameters. The first pass of the dual-pass data assimilation scheme is used to optimize these parameters, like parameter calibration. The EC and LAS system represent different spatial scales, EC system can represent one remote sensing pixel, and LAS system can represent more than one pixel. For LAS, the source areas were calculated and compared with pixels they covered. To avoid evaluation error from one data, two data sets (LAS and EC) are used to maintain the objectivity of the evaluation results. The analysis error (AE) used in this study is used to show the data assimilation scheme can reduce model uncertainties with the assimilation remote sensing data. Except for RMSE and AE, the correction and an analysis of the background departure (obs-model) will be used to evaluate the data

assimilation results.

Comment: 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.

Response: We agree with your suggestion. We will edit the language carefully to enhance the impact of the paper. We will revise the description of the methods, the dual-pass data assimilation technique in Section 2. Ten soil layers are designed in CoLM, and the data assimilation scheme updates all these layers. We will describe these things in the revised version.

Minor Comments:

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

Response: We accept your suggestion, and change FY3A to remote sensing data.

Comment: P8495, L10: Kanemasu et al is not listed in the references. Response: Thanks, we will list it in references.

Comment: 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.

Response: Thanks, we will revise it.

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

Response: Thanks, we will correct it.

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

Response: Thanks, we will change it.

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

Response: Thanks, we will re-phrase it.

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

Response: Thanks, we will change it.

Comment: 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.

Response: Thanks for your advice. Since the vegetation parameter change not very fast, we assume the vegetation parameters change once per week. Soil moisture can be changed every day caused by the evapotranpiration, rainfall, and so on. Then, soil moisture was optimized once the satellite data was available (daily). Thus, the model states and parameters vary at the different temporal scales. We will justification for this statement in the manuscript.

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

Response: Thanks, we will add some information of FY3A here.

Comment: P8498, L10: I don't think "EC" or "LAS" have been defined in the main body yet. Response: Thanks, we will define EC and LAS here.

Comment: 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.

Response: Thanks, we will delete it.

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

Response: Thanks, we will correct it.

Comment: 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.

Response: Thanks, we will summarize equation 3 and equation 4, and clarify Q is a matrix.

Comment: 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.

Response: Thanks for you advice. We will provide the actual soil moisture errors used in the text. We will add some contents about how the ensemble generated in our data assimilation experiment.

Comment: P8500, L10: parameter uncertainties of 10n% 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.

Response: Thanks, we define 10% according to Mölders (2005). At the same time, the range of each parameter was also given that they must vary within this range, and will

be listed in table 2.

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

Response: Thanks, we will change it.

Comment: 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).

Response: thanks, we will redesign Figure 1 with information that how the dual pass data assimilation scheme works. The long time refer to the model parameters will be optimized weekly, and the short time refer to the model soil moisture will be optimized daily. The model parameters optimized in last week (first pass) will be used to the model in the next week.

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

Response: Three year averaged (2008-2010) MODIS LAI has been used in the experiment, and we will indicate it.

Comment: 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.

Response: To indicate the objectivity of the EC data, we list EBR of EC data at each site. We want to show that the turbulent flux observations should be lager than EC measurements, and we will add this information to the result analysis.

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

Response: Thanks, The surface soil heat flux is calculated using multi-layer soil temperature and soil moisture observations that proposed by Yang and Wang. (2008).

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

Response: Thanks, the method to calculate surface soil heat flux proposed by Yang and Wang. (2008) need multi-layer soil temperature and soil moisture observations, and the accuracy of surface soil flux is high with more observation data. Since soil moisture was only measured at 4 and 20cm at BJ site, the surface soil heat flux is not calculated at this site.

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

Response: Thanks, not the night time fluxes were excluded, but the footprint values were excluded. We will revise it.

Comment: 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?

Response: Thanks, the VIRR instrument has 10 channels ranged from 0.58 to 12.5 μm. The 4th (10.3 ï¡đ 11.3μm) and 5th (11.5 ï¡đ 12.5 μm) are infrared channels little water vapor absorption, and they have a nominal spatial resolution of 1km$\times$1km at the nadir. With these two infrared channels, land surface temperature is obtained using a local split window method (Becker and Li, 1990): Ts=A0+PâĂć(T4+T5)/2+MâĂć(T4-T5)/2 Where Ts is the FY3A LST, A0 is a constant, T4 and T5 mean the brightness temperature of the 4th and 5th channel, P and M mean the function of land surface emissivity.

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

Response: Thanks, the RMSE between ground-measured temperature and FY3A LST are used for observational error, the values are list in Figure 2.

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

Response: Thanks, we will delete it.

Comment: P8507, L15: fix "effect contribution"

Response: Thanks, we will change “effect contribution” to “contribution”

Comment: P8508, equation 20. State what V is.

Response: Thanks, V(Y) means the variance of Y which is the model output.

Comment: 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.

Response: Thanks, the equation 21 should be changed to Si=Vi/V Where Si is called first-order sensitivity index, the subscript i indicates that Si is the index of xi.

Comment: P8508, equation 22: State what the tilda indicates.

Response: Thanks, V$\sim$i means the variance of model output which is not induced by xi.

Comment: 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?

Response: In previous studies, Gaussian error Propagation (GEP) is used to identify the sensitive parameters and in this case we used Extended Fourier Amplitude Sensitivity Test (EFAST) to do the same work. Compared with the former case, EFAST is a kind of global sensitivity analysis method but GEP is a local sensitivity analysis. Global sensitivity analysis is capable to capture the influence of the full range of variation of each input factor (Saltelli et al. 2000). Since Common Land Model is a complex non-linear model, using global sensitivity analysis will be better to identify the sensitive

parameters.

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

Response: Thanks, we will repeat the sensitivity analysis at all sites.

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

Response: Thanks, we will increase the size of plots in Figure 3.

Comment: 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.

Response: Thanks, we will show the data assimilation results for the full time period.

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

Response: Thanks, we will use the full name of the parameters in the text.

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

Response: Thanks, the canopy height of each site will be given in the revised manuscript.

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

Response: Thanks, the hhti parameter means photosynthetic stress high temperature.

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

of "Julian Day" with "Day of year"

Response: Thanks, we will replace “Julian day” with “Day of Year”.

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

Response: Thanks, the parameters have been optimized using the data assimilation scheme, including in this section. The results are presented in section 5.3.

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

Response: Thanks, this sentence means “ If the resistance is not taking into consideration in a land surface model, the latent heat flux will reach a very high value and the land surface will dry up quickly caused by the evaporation. Then, the dry land surface will lead to low latent heat flux values.”

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

Response: Thanks, we will change it.

Comment: 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.

Response: Thanks, we will present the data for the full diurnal cycle.

Comment: 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).

Response: Thanks, the spatial scale of land surface model and data assimilation scheme depends on the representative of ground-measured meteorology data (in-situ), land surface parameters and remote sensing data (1KM). Then, the spatial scale will

be larger than 1KM. we will revise the descriptions.

Comment: 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.

Response: Thanks, we will revise it.

Comment: 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.

Response: Thanks, we will provide a scatter plot of sensible heat flux between LAS and EC, and give a brief analysis.

Comment: 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)

Response: Thanks, we will revise this sentence to “Thus, this section presents the retrieved parameters from the dual-pass data assimilation scheme.”

Comment: 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?

Response: Thanks, a new accurate soil texture map with the spatial resolution of 1KM provided by Shangguan et al (2012) was used in this study. The porosity can be calculated using the soil texture data. Thus, the soil porosity is not optimized in this study. BJ site is located in Tibet plateau, and the soil texture data need more detailed investigation and application to land surface modeling.

Comment: 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?

Response: Thanks, the stable values at the four sites are averaged to one value (temporal and spatial average) and used in place of the default. We will rephrase it.

Comment: 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).

Response: Thanks, the error bars means the uncertainty of the retrieved parameter. In this study, it means parameters maximum and minimum value of ensemble member. After a time period, the binter tends to a stable value, and the variance is very small.

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

Response: Thanks, the long time means one week, and short time mean one day.

Figures and Tables

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

Response: Thanks, we will add information of this.

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

Response: Thanks, the time period was the same as in section 5.1.

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

Response: Thanks, we will revise it.

Comment: 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.

Response: Thanks, we will combine table 4 and 6.

Please also note the supplement to this comment:
http://www.hydrol-earth-syst-sci-discuss.net/9/C4577/2012/hessd-9-C4577-2012-supplement.pdf

---

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