

## ***Interactive comment on “Sequential assimilation of satellite-derived vegetation and soil moisture products using SURFEX\_v8.0: LDAS-Monde assessment over the Euro-Mediterranean area” by Clément Albergel et al.***

**Anonymous Referee #2**

Received and published: 26 July 2017

### **General Comments:**

In the paper "Sequential assimilation of satellite-derived vegetation and soil moisture products using SURFEX\_v8.0: LDAS-Monde assessment over the Euro-Mediterranean area", the authors are motivated "to increase monitoring accuracy" of land surface variables such as soil moisture (SM) and leaf area index (LAI) over the European Mediterranean region. They use a global Land Surface Assimilation system called "LDAS-Monde" to assimilate both SM and LAI in experiments to assess the effects on the model land surface variables. The model results are compared to indepen-

C1

dent observations of river discharge, land evapo-transpiration and different agricultural statistics and measures.

I recommend that this paper undergo major revisions.

### **Specific Comments:**

The last few sections are not clearly organised or written. The abstract and introduction are very clear about the purpose of the paper, however, the sections from 3.4 onwards lack clarity and do not lead the reader to a direct interpretation of the results as stated in the abstract. The reader would have difficulty coming to the conclusions that are put forward in the abstract. I recommend that these sections be carefully re-written.

The title includes acronyms that should be spelled out in full if they really need to be used at all. Not everyone is familiar with the acronym "LDAS" or "SURFEX". Is it necessary to put a specific version number of "SURFEX" in the title?

In many places the RMSD and correlations computed are discussed in the same sentence and this creates confusion. It would be simpler to have two or more shorter sentences that are more explicit about which measure is being used for the comparison. I think that overall the authors have chosen brevity over clarity.

Also, the sign (positive or negative) of a change in the metric used is given without explaining what the change means in terms of the variables or the models. A physical interpretation of such results would be helpful.

Please rewrite the sentence in 535-536 "Most of the differences in RMSD are negative..." RMSD is a strictly positive or zero value. Are the differences in RMSD between two different data sets being compared? Could the authors please write two sentences that explain this point more explicitly? It appears to be an important point as it is going to "show the added value of the analysis".

The figures (details are given in the technical comments below) need work as well. For example, in Figure 8: What is N? You don't really need a legend for the red and

C2

green lines on each of the 6 month plots. Just define this in the caption. Panels need labels a, b,c and they need to be referenced as such in the caption. Please label the x axis with variable name and units. Most importantly, the y-axis is not a probability but a frequency of occurrence (the caption even says "histogram" which is correct). The integral of the probability distribution function should be equal to one by definition.

The last two paragraphs of section 3.4 are very unclear. The sentence "From Figures 13 d and e and Figure 13 one may notice...seasonal decrease and increase of the scores." does not make sense. Perhaps the authors need several sentences here that are more precise about which figures support which conclusions. Again, the RMSD and correlation should be discussed separately to make the evidence clearer. Line 555 contains another confusing sentence. "differences in RMSD and correlations are negative and positive: 70% and 79%. This just doesn't make any sense. Are there percent changes in a particular direction? If so, what are the implications for the model output or the physical system?

Section 4.1 is called "Can different data assimilation techniques improve the analysis?" I don't believe that this particular question has been answered in this section by the work presented here. I think that alternative methods are proposed and discussed but the actual results in the paper do not answer the question whether one techniques is better than another. If the section could be renamed, that would be more clear.

Technical Corrections:

Abstract: SM is defined but later in the main body SSM is used. Perhaps should just define SSM in abstract and stick to it?

Introduction: Define acronyms MODCOU and SAFRAN

Section 3.1 Should be "Consistency between the model and observations"

Need to state what is consistent with what. "Observations consistency over time..." is not clear. Do you mean that one set of observations are consistent with another

C3

observation set? or with the model output?

lines 61-63: should read "perform best for plant productivity...to used soil moisture and vegetation observations together to improve..."

line 91: WFDEI is defined in section 2 but should be defined here as well or instead of section 2.

lines 97-103 Sentence is much too long. Please break up into separate sentences.

line 115 CTRIP should be defined.

line 119 "detailed hereafter" should be "described in the following sections."

line 122 "They" what is it? Is it the model parameters in the previous sentence? Please be specific.

line 128-129 "net assimilation of CO<sub>2</sub>" Because the word assimilation is also used in the context of data assimilation, perhaps a different word could be used here? Like "uptake" or "intake"? I just think that using the word assimilation used in the 2 different contexts might confuse the readers.

line 132 "evaporation of"? Or should it read "evaporation due to (i) plant transpiration"?

line 140 What is "it"? Snow scheme or soil diffusion scheme?

line 140 "Richard's " should be "Richards' " (apostrophe after the s) and you need a reference: Richards, L.A., 1931. Capillary conduction of liquids in porous mediums. Physics 1, 318 – 333

line 143 Need a reference for the Brooks and Corey model.

line 187 The LSM is represented by the letter M, but that is not used until eqn. 5. Perhaps better to name M closer to eqn. 5 in the text.

line 237 Should "harmonies" be "harmonious"?

C4

line 297 Is "discharge" "river discharge"? If so please state this.

line 307 "model ability" should be "model's ability"

line 341-345 This is a long sentence and should be broken up. The last bit "...LAI (for SSM and LAI)." doesn't make sense to me, please clarify how LAI is for SSM and LAI? Please make sure that LAI is defined.

line 386 is also unclear with "data set is consistent over time" consistent with what exactly?

Section 3.3 title could be "Impact of the Analysis".

line 390 section 222 should be 2.2.2

Line 400. "Correlation (RMSD)" Please explain what RMSD is it the root mean square deviation, the difference or the sample standard deviation?

line 411 "good values" is vague. Do you mean "high correlation values"?

In the text, the differential terms such as delta (SSM)/ delta (LAI) are missing the superscript that is included in equation 9. Line 450 the lack of superscripts renders that term particularly unhelpful.

line 425 should be "higher than those" not higher compared to"

line 469 Should be "Jacobian's" not "Jacobians"

line 518 Where is "Eff." defined? I would change sentence to "greater than 0 and with 22 gauge stations reporting Eff greater than 0.5."

line 521 Change "superior" to "greater than" or use the mathematical symbol ">" in this paragraph.

line 521 Change to "(3 stations report a decrease > 0.05)"

line 532 Where is "open-loop" defined?

C5

Line 544 MTE needs to be defined.

Line 565 What is an "excessive Jacobian"?

Line 567 What is "They"? and what is the "force-restore version" version of what?

Line 586 which "model" and what is "It" in "It that accounts for the texture-based..."

Line 577 " system too reliant on the chosen forcing" might be better.

Line 573 "they exhibit" what is they?

Line 595 "elaborated methods" doesn't make sense.

Line 601 Again the term from the Jacobian matrix is missing sub or superscripts.

Line 609 Should be "Can better use of" not "Can a better use of"

Line 630 "too large" could be better as "such large"

Line 643 "suggest an added value on vegetation variables" is unclear. how do these variables add value and what exactly is the value added?

Line 652 should be "assimilating retrieved soil moisture"

Line 655 "Tb" needs "b" as a subscript.

Line 666 Better to write "at Meteo-France; it will account for "

Tables:

Table 1. Under "Model" what do DIF and NIT mean?

Figures:

Figure 1. typo "rigth" should be "right"

Figure 2: What does the shaded area represent? Should explain in the caption. Need full stop at end of sentence.

C6

Figure 3: The panels are very small. I think that all panels should be labeled a, b, c etc. and then referred to in the caption by letter. The top 6 panels appear to be for the median R values and the bottom is for a mean RMSD. This is not mentioned in the caption. What times are used in the creation of the median and mean? "Averaged values are reported..." which values are being averaged? In caption state that w\_2 is the second layer of soil.

Figure 4: Needs a label for the x axis. N is not defined in the caption but a number is given for N in each panel. The Jacobian elements need to match equation 9. There is a lack of superscript on the LAI variable. What are the solid blue lines in the histogram? Only the lines are defined in caption. Is there a vertical line drawn at 0.0? That should be stated because it is hard to see.

Figure 5: State which column is which and which row is which. "Rows from top to bottom represent averaged analysis increments for all months Feb, May, Aug and Nov from 2000-2012...."

Figure 6: The y axis is not labeled correctly. It should be latitude not 200001-201212. If that is a year and month, it should be in the title or caption. Captial "S" needed. Change to "whole period 2000-2012".

Figure 7: Panels need labels a, b,c and they need to be referenced as such in the panels.

Figure 8: What is N? You don't really need a legend for Red and Green on each of the 6 month plots. Just define in the caption. Panels need labels a, b,c and they need to be referenced as such in the panels. Label the x axis. y-axis is not a probability but a frequency of occurrence. Integral of the Probability function should be equal to one.

Figure 9: Panels need labels a, b,c and they need to be referenced as such in the panels.

Figure 10: In caption, please tell the reader what is Agreste?

C7

Figure 11: Panels need labels a, b,c and they need to be referenced as such in the panels.

Figure 12: The y axis is not labeled correctly. It should be latitude not 200001-201212. The multiple panels are very small and hard to see. I think that you could take the middle row and make it into a new figure. It is not about Evapotranspiration like the top and bottom rows. Please rewrite the second sentence. Be more explicit. For example: Maps of averaged taken over 2000-2012 of a) evapotranspiration...

Figure 13: Rewrite caption. Use full stops. For example: RMSD (a) and correlations (b) between analysed (modelled) ....Panels c and d show Carbon... Panels e and f compare...

Figure 14: Panels need labels a, b,c and they need to be referenced as such in the panels. What is the observation dataset being used? What is the red shaded area? Rewrite: " Monthly RMSD and correlation values between...."

---

Interactive comment on Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2017-121>, 2017.

C8