

Geosci. Model Dev. Discuss., author comment AC2
<https://doi.org/10.5194/gmd-2021-38-AC2>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC2

Lukas Strebel et al.

Author comment on "Coupling the Community Land Model version 5.0 to the parallel data assimilation framework PDAF: description and applications" by Lukas Strebel et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2021-38-AC2>, 2021

Scientific/Detailed Comments: Referee comments in bold and author answer non-bold.

Major Comments

The motivation for this study is weak. The authors briefly mention about the difference between online and offline DA (Ln 55), but they need to better motivate the coupling CLM5.0 with PDAF. Is it more for the standalone DA with CLM5.0 or for CLM5.0 within the TSMP framework? What new does PDAF bring? How does it reduce the number of core-hours or computation time compared to other offline DA? And, how it scales with increase in domain size and time period of simulation? This needs to be discussed clearly.

It is both for the standalone DA with CLM5, as shown in this study, and also the potential for future use in the complete TSMP framework. We intended to motivate both in the introduction, but we will improve this with comments below and comments from other referees. We did not perform computational performance comparison to other DA frameworks for this specific application. General scaling behavior of PDAF has been shown in Nerger et al. (2013) and Kurtz et al. (2016). Without extensive computational studies we do not want to discuss advantages and disadvantages of different DA frameworks in detail. Instead we want to focus on the specific implementation and application of a new coupling that can be used to perform DA with CLM5.

Kurtz et al. (2016) already presented the PDAF coupling to TSMP including CLM3.5. So, what is new in this study? I assume that there must be substantial work involved in developing the PDAF interface around CLM5.0 which has different software environment compared to earlier versions of CLM (e.g. CLM3.5). But it is not so clear in the current version of the manuscript.

The new developments in this study are the modifications to what Kurtz et al. (2016) presented. These modifications are necessary to interface with CLM5.0 which, as you mentioned, has a different software environment compared to earlier versions. We discuss the implementation and differences in section 2.3 in detail, but we will make the differences clearer and highlight the new developments more in the revised version.

Ln 85: This comes so suddenly. The authors need to provide better motivation to use single column model. The literature review is another weak part of the manuscript. The authors make no effort in presenting their results in context of

previous findings. Also, does the improvement in soil moisture also improves the surface energy fluxes. For LSMs, improvements need to be explored soil states as well as fluxes. And, a discussion section is missing.

We provide references to studies using single-point simulations (Ln 35-43) to motivate our choice of a single grid cell setup and also discuss other studies for the specific study site (Ln 79-85). However, we will make the motivation clearer in the revised version.

We provide context for this study in the literature review for single-point studies (Ln 35-43), for data assimilation in LSMs (Ln 59-73), for the specific software framework (Ln 73-77), and for the specific site (Ln 79-85). The first reviewer has already pointed out some additional studies that we will include in our literature review. We will also add more literature references and make the respective contexts clearer in the revised version.

We will include results and discussion for the changes to ET in the revised version. We included the contents of a discussion section in the conclusion section. We will make this clear in the revised version.

There is no README file or User manual to reproduce the results presented in this study, also please provide a web URL for Zenodo and cite this paper in the References. The upload should also include scripts for processing the figures and observation data for reproducibility.

We will create a README, and add scripts for processing and observation data in the revised version.

Minor Comments

Ln 10: Even tuned second generation LSMs can be "accurate", here maybe the authors want to imply that third generation LSMs better represent the key physical processes. Also, check in the rest of the manuscript.

Yes, we intended to stress the improvements in the representation of physical processes. We will modify the sentence to make this clear.

Ln 11: more? What type of data?

Various types, from new satellite products to new in-situ measurement stations, also new cosmic-ray and flux tower sites. We will add this information in the revised version.

Ln 15: Is this further development of PDAF or addition of new interface to connect PDAF with new models?

It is the addition of a new interface to connect PDAF with a new model, we will correct this sentence in the revised version.

Ln 34: common might not be the right word here.

Common as in 'often used' but we will change the sentence to make this clearer.

Ln 48-53: This paragraph needs to be rephrased (framework, external framework, within framework). It has just too many frameworks.

We will rephrase the paragraph with fewer 'frameworks'.

Ln 70: PDAF with joint state parameter update for CLM was also used in the

following study: Shrestha, P., W. Kurtz, G. Vogel, J.-P. Schulz, M. Sulis, H.-J. Hendricks Franssen, S. Kollet and C. Simmer (2018), Connection Between Root Zone Soil Moisture and Surface Energy Flux Partitioning Using Modeling, Observations, and Data Assimilation for a Temperate Grassland Site in Germany. JGR-Biogeosciences doi: 10.1029/2016JG003753

We will include this reference as another example of joint state parameter update with PDAF and clm3.5 in the revised version.

Ln 73: "In this study, we present the coupling of .."

We will rephrase the sentence.

Ln 93: Rephrase. "The paper ends with " is not appropriate.

We will correct the sentence in the revised version.

Ln 116: 1) variation methods, ...2) sequential methods

The reference we cite (Reichle 2008) calls them 'variational methods'.

Ln 125: Perturbation vector missing in Eq. 1, where y is generally the observation vector. It is discussed much later in Ln 146. What is the measurement error?

We will move the inclusion of the perturbation to the observation vector closer to Eq. 1. For simplicity, the measurement error is assumed be constant and set to a RMS of 2%. We will mention this in the revised version.

Section 2.3: There is always a discussion about older version, maybe the authors should discuss it before, and present their new formulation, rather than interchanging now and then. Maybe this would also highlight, what new work has been done.

We compare to the coupling with the older version of CLM three times in this section: 1) In the section about the difference in time stepping between CLM5 and TSMP. Here the comparison is not strictly necessary, but highlights that the approach to modify the driver is the same as before even if the software environment has changed significantly. 2.) To point to the changes in CLM5 hydraulic parameter calculations, which includes the new changes with the addition of soil organic matter. 3.) To mention that the more complex software environment motivates the modification of the existing CLM5 ensemble mode. Other comparisons in the section are not to older versions but to the framework of TSMP or PDAF specifically. We think it would be less useful to separate any comparisons, since they are mostly used to give context to new implementations. Nevertheless, we will highlight more clearly the new work that has been done in the revised version.

Ln 181: The "Figure 1" is not helpful, either improve or remove. Also, rephrase and elaborate the discussion.

We use 'Figure 1' as a visual aid to describe the structure of both the actual implementation and the paragraphs in the section. We will improve the figure by adding more context to it.

Ln 191: What is "CIME"?

CIME is the default cIm5.0 driver. We will add a definition for CIME to the revised version.

Ln 204: Maybe "clipping" ?

We will correct the sentence in the revised version.

Ln 218: Rephrase.

We will rephrase the sentences.

Ln 232: in Wüstebach , and Belgium ?

We will correct the sentence.

Ln 252: Explain the SWC unit.

We will add a definition for the volumetric soil water content.

Ln 305: "to overestimate SWC" or "wet bias in SWC"

We will rephrase the sentence.

Ln 333: What is variant here?

The online variant to differentiate from the offline variant of PDAF as discussed in the introduction and the next paragraph. We will clarify this in the text.

Figures: Add subplot numbers (e.g., a), b))

We will add subplot numbers.

Figure 2: "In the diagram NMLST means namelist, SIM means simulation process, HIST means history file output, PID means PDAF identification number." – this should as legend in Figure.

We will add a legend with the shorthand.

Figure 3 caption: red (solid line), light green (dotted line).

We will correct the formatting.