Reviewer specific comments:
Recent literature on spatial calibration of LSM is missing from Introduction section.
Reply:
We agree that recent literature on spatial calibration of LSMs is missing in the Introduction section. However, there are not many studies where land surface model parameters are estimated. We already write that "… only a limited number of studies have dealt with calibration and sensitivity analysis of the energy and hydrology parameters in LSMs" (p3, line 15-16). Because we apply CLM in point-scale mode, the Introduction section put emphasis on the point-scale LSM calibration, however we agree that especially (Demirel et al., 2018; Mendiguren et al., 2017) is relevant, and we will include those in the Introduction section.

We will include the suggested papers as references:

(Tangdamrongsub et al., 2017) at page 2 line24

(Demirel et al., 2018) at page 2 line 25

(Mendiguren et al., 2017) at page 3 line 24

(Lane et al., 2021) at page 2 line 28

Reviewer:
-Page 7 Line 11: "1000 years" please explain how?

Reply:
1000 years is a somewhat arbitrary number to ensure that the model had reach a quasi-equilibrium. According to "CLM5 userguide" CLM5.0-BGC-Crop needs at least 1000 years of spin-up. Our CLM model does not include ecosystem carbon, however is takes approximately 150 years of spinup for soil temperature to reach equilibrium.

Reviewer:
-P7L14: "final simulations" why only before final simulations and not during calibration? Please explain more..
Reply:
We included four years of spin-up preceding each and every simulation in the calibration. We see that this is not clearly written, and we will rewrite the sentence.

Reviewer:
-P8L5 to11: this paragraph should be moved to the section 2.3 describing calibration approach to avoid repetition.
Reply:
Ok, we will do that.

Reviewer:
-P8L27: "Focus was given to a set of 30 time-invariant model parameters." Apparently no sensitivity analysis was applied? Why?
Reply:
John Doherty (personal communication) recommends Highly Parameterized Inversion, were most parameters are included in the calibration. The regularization approach will keep the insensitive parameters at their preferred values. We are aware that the calibration time could be reduced if removing some parameters from the calibration,

however we were interested in studying how much the parameter values deviated from the look-up table values after regularization.

Reviewer:
-In a calibration framework it is essential to apply SA first to reduce search dimension. May be some of the 30 parameters have zero influence on the objective function? Did you utilize PEST's local sensitivity analysis option?
Reply:

We did do a local sensitivity analysis with PEST (see figure below). However as explained in the former question we decided to include all parameters in the calibration.

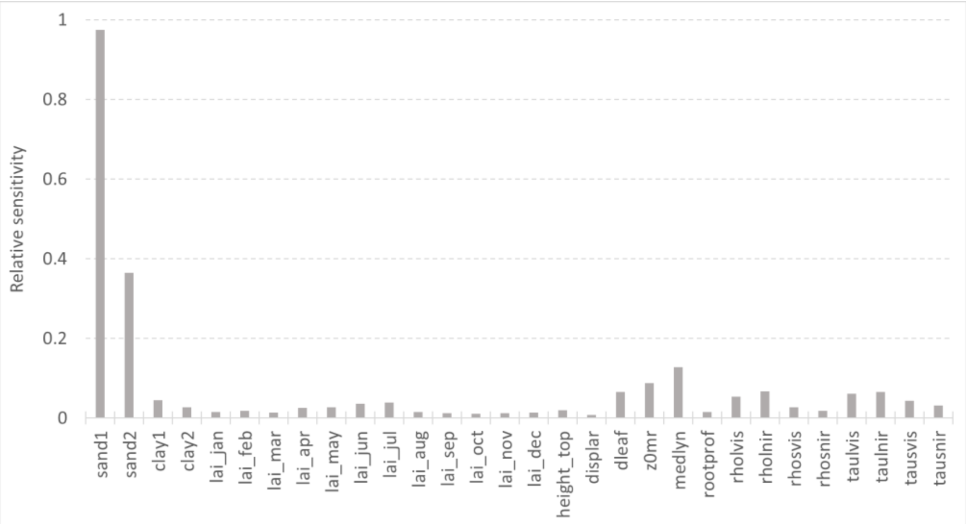

Figure 1: Local sensitivity analysis from PEST. The analysis has been done on initial parameter value of LAI in March, September, October, November and December of 0.5 instead of 1, because log-transformation of those parameters would reveal a sensitivity of zero.

Reviewer:
-Section 2.3: The reader can be curious about several details of the calibration framework.
Reply:
We agree that readers could be curious about the details of the calibration framework. We have answered your questions in the following and we will incorporate the information on the calibration in the manuscript.

1)what was the user defined maximum number of iterations for such a sophisticated mode?
Reply:
Yes, maximum number of iterations were defined as 50 iterations. The only scenario reaching this were scenario F.

Reviewer:
2)computer runtime statistics and cluster properties (logical processors, ram capacity, intel/amd etc)
Reply:
One model run took about 10 minutes on a Linux server (Intel Xeon Gold 6148 processor, 20 cores, 380 GB RAM).

Reviewer:

2)Pest has three search algorithms LM, SCE-UA and CMAES. Can "Tikhonov regularization" be used together with one of these search algorithms?

Reply:

We chose to apply the gradient-based nonlinear Gauss-Marquardt-Levenberg method implemented in PEST, were the calculation of finite–difference derivatives are used in the inversion process. We did that because those often use fewer models runs that alternative optimization techniques (Doherty, 2015). We introduced Tikhonov regularization to honor the observed parameters values as prior knowledge. If using the global optimizer of CMAES_P (also implemented in the PEST suite) we could have used "pseudo-regularization" (Doherty, 2018) were credence to parameters derivatives is done by weight adjustments.

Reviewer:

3) sharing PEST control file ".pst" in appendices (or supplementary) can be good for this open access journal.

Reply:

Ok, we will share the .pst file either as appendix or supplementary.

Reviewer:

-only eq 10 is bias insensitive metric. Why the authors did not choose a spatial metric focusing on patterns of fluxes in growing season? Evaluating hourly (unstable) fluxes can be misleading. Instead evaluating monthly patterns of SWC, AET, SM can be a robust guide for the model. Fig 2-3-4 are showing only temporal aspects of the fluxes/states but this kind of finite element based LSMs can provide maps outputs. The authors should show also some map results. Looking at only time series can be boring.

Reply:

As we apply CLM in point-scale mode it is not possible to include a spatial metric or showing the results at maps.

We write that: "The target of the applied calibration approach is the dynamics of the 24-hour cycle of hourly observations rather than the seasonal energy and water balance components." (page 21, line 4) Therefore, we did not include the monthly and seasonal patterns.

Reviewer:

-why Pareto approach was not used for multi-objective calibration to avoid dominating solutions. Pareto DDS algorithm (available in Ostrich) could offer multiple non dominating solutions. PEST doesn't include this algorithm yet.

Reply:

PEST is one of the most well-developed inversion and parameter uncertainty software programs. However, we agree that it would have been an opportunity to use the Pareto DDS algorithm (in Ostrich). Obtaining multiple non dominating solutions would show if the look-up table parameters values were inside the parameter values interval of those solutions. However, the single parameters value set of minimum error variance obtained in our approach, somehow correspond to the method of using a single-set look-up table values as generally done in LSMs, so that those two parameter set can be directly compared.

Literature:

Demirel, M. C., Mai, J., Mendiguren, G., Koch, J., Samaniego, L., & Stisen, S. (2018). Combining satellite data and appropriate objective functions for improved spatial pattern performance of a distributed hydrologic model. *Hydrology and Earth System Sciences*, *22*(2), 1299–1315. https://doi.org/10.5194/hess-22-1299-2018

Doherty, J. (2015). *Calibration and Uncertainty Analysis for Complex Environmental Models - The PEST book*.

Doherty, J. (2018). *PEST - Model-Independent Parameter Estimation - User manual Part I*.
Lane, R. A., Freer, J. E., Coxon, G., & Wagener, T. (2021). Incorporating Uncertainty Into Multiscale Parameter Regionalization to Evaluate the Performance of Nationally Consistent Parameter Fields for a Hydrological Model. *Water Resources Research*, *57*(10), 1–19. https://doi.org/10.1029/2020WR028393
Mendiguren, G., Koch, J., & Stisen, S. (2017). Spatial pattern evaluation of a calibrated national hydrological model - A remote-sensing-based diagnostic approach. *Hydrology and Earth System Sciences*, *21*(12), 5987–6005. https://doi.org/10.5194/hess-21-5987-2017
Tangdamrongsub, N., Steele-Dunne, S. C., Gunter, B. C., Ditmar, P. G., Sutanudjaja, E. H., Sun, Y., Xia, T., & Wang, Z. (2017). Improving estimates of water resources in a semi-arid region by assimilating GRACE data into the PCR-GLOBWB hydrological model. *Hydrology and Earth System Sciences*, *21*(4), 2053–2074. https://doi.org/10.5194/hess-21-2053-2017