Reviewer2 (R2) specific comments (C):

R2_C1:

Denager et al. implemented multi-objective calibration of point-scale CLM5 using several types of flux/states observations of LE, H, recharge (q) and SWC from the Danish hydrological observatory HOBE. This topic of constraining model parameters against multi-source observations is quite relevant to the HESS journal, and it can be a valuable contribution to the community after addressing my following comments listed below. Additionally, the paper is clearly written and well-referenced; some parts require revisions, as follows. English typos should be double-checked, and some textual suggestions are further provided at the end. In the figures, I often can't distinguish individual scenarios. Differences between individual calibration scenarios are not clearly depicted. Please, improve the readability of the figures.

Reply:

To improve the readability, Figure 2, 3 and 4 will only include observed, scenario X and scenario D. Scenario A and E will be removed from the figures. Figure 1 will be changed to a table. See reply on R2_C6.

R2_C2:

1) Title requires modification. It needs to be clear from the title that the calibration is for one point-scale site.

Reply:

"Local-scale" will be added to the title such that it will read:

Local-scale multi-objective calibration of the Community Land Model Version 5.0 using in-situ observations of water and energy fluxes and variables

R2_C3:

2) The abstract should be more concise and to the point, highlighting concrete results of the present study, and quantifying the results. So, please remove/rewrite too general statements, which are probably better suited for discussion of the results or conclusions. E.g. I suggest removing "Furthermore, reliability of the optimized model parameters can be estimated by statistical measures such as identifiability and relative error variance reduction. As in most other eddy covariance studies, closure of the land surface energy balance is not achieved on observation data." The following statement, "The fact that CLM5 is not capable of matching sensible heat, not even with advanced parameter optimization of model parameter values, suggests that the lack of energy closure is due to biases in the sensible heat flux" is probably also not the most suitable one for the abstract. Instead, I would like to know from the abstract, which of the considered variable was most useful in improving the process representation. Did calibration of one variable improve the model's predictive skill of another (uncalibrated) variable? Which one? Also, an abstract should mention at which site (i.e., agricultural field observatory in Denmark) the CLM5 is established.

Reply:

The abstract will be rewritten according to the suggested improvements. The abstract will not focus on the energy balance but focus on calibration target variables and parameters. See suggested new abstract below:

Abstract: This study evaluates water and energy fluxes and variables in combination with parameter optimization of the state-of-the-art land surface model Community Land Model version 5 (CLM5), using six years of hourly observations of latent heat flux, sensible heat flux, groundwater recharge, soil moisture and soil temperature from an agricultural observatory in Denmark.
The results show that multi-objective calibration in combination with truncated singular value decomposition and Tikhonov regularization is a powerful method to improve the current practice of using look-up tables to define parameter values in land surface models. Using measurements of

turbulent fluxes as target variable, the parameter optimization is capable of matching simulations and observations of latent heat, especially during the summer period, while simulated sensible heat is clearly biased. Of the 30 parameters considered soil texture, monthly LAI in summer, stomata conductance and root distribution have the highest influence on the local-scale simulation results. The results from this study contribute to improvements of the model characterization of water and energy fluxes. The study underlines the importance of performing parameter calibration using observations of hydrologic and energy fluxes and variables to obtain optimal parameter values of a land surface model.

R2_C4:
3) Second half of the Introduction should clearly point out the research gap and your contribution to filling it in. Clearly stating the novelty of your manuscript somewhere in the last two paragraphs of the Intro.
Reply:
We will include the following paragraph in the introduction:

"In this study, we evaluate in-situ water and energy fluxes and variables at an agricultural field site in Denmark using the state-of-the-art LSM Community Land Model version 5 (CLM5) coupled to the optimization code PEST (Doherty, 2015). In most previous research, LSMs are not calibrated and instead use lookup tables to define parameter values. Here we identify values of important parameters in an LSM using multi-objective calibration in combination with regularization to improve the simulation of the hydrological processes."

Furthermore, we will improve the overall language of the introduction such that the novelty of the study stands out more clearly.

R2_C5:
4) Regarding the experimental design (Page 8, Line 8), please be consistent; earlier, you mention four variables (LE, H, recharge (q) and SWC); here, you mention six different observation data sources. Which one is correct, then? Please, synchronise, otherwise it is confusing. Table 1 already includes the calibrated parameter values, It is not clear how these parameters were identified when Table 1 was first introduced. From Table 1, it looks like you calibrated sand and clay contents directly. Was Clapp-Hornberger exponent B also part of the calibration process? As it is not part of Table 1. Please, clarify.
Reply:
Seven different observation data sources are used in the study, LH, H, q, SWC, Sout, Rn and (in the revised manuscript also) Tsoil. This will be stated clearly in the revised manuscript.
Clapp-Hornberger exponent B is not a calibration parameter. This will be clarified in the manuscript. The Clap-Hornberger B exponent is inherently defined in CLM5 from pedo-transfer functions of percentages of sand, clay and organic matter.

R2_C6:
5) Figure 1 is a rather set of tables than a figure. Increase the font and readability of the Table.
Reply: Thank you for the comment. We have discussed this internally earlier. In the revised manuscript we will change it to a table. To improve the readability of the table, we will also clean up in the metrics by putting some of them in supplementary materials. We are open for other suggestions for improving the readability.

R2_C7:
6) Why the calibration of LE (scenario A), does not improve the climatology of LE during March at all? (see Figure 4, please clarify)
Reply:
Actually, the calibration of LE (Scenario A, D and E) does slightly improve the climatology of LE during March compared to the control run (Fig 4a).

R2_C8:
7) How is it possible that the calibrated sand and clay values have such a large spread among scenarios? Sand[%] and Clay[%] could probably be well estimated by field measurements which you have available. I would instead calibrate some parameters which can not be measured in the field.
Reply:
We are interested in estimating the hydraulic properties of the soil in the form of the retention and hydraulic conductivity functions. However, in CLM5 it is not possible to specify these functions directly. Instead, %sand and %clay are used for estimating the Clapp-Hornberger exponent B and therefore we consider %sand and %clay as calibration parameters regardless of the values they may have from field measurements.

R2_C9:
8) It might also be interesting to see the scenarios aggregated into monthly seasonal values in addition to the diurnal climatology.
Reply:
We certainly take this as a legitimate suggestion. However, to constrain the study we have chosen to focus on the diurnal variations and not the seasonal values. As our observation data is exclusive in the way that we have hourly observations available for a long time period, we choose to focus on the diurnal variations.

R2_C10:
Data availability: under the provided link, data can not be easily found. Also, the processing codes are not available.
Reply: If interested parties need help in locating the data at the provided link the corresponding author can assist in this. The processing codes can also be made available by the corresponding author. This will be stated in the manuscript.

R2_C11:
Textual suggestion:

Page 2, Line 13: practice is to use => practice to use

Page 3, Line 13: list of LSM is too short, why not be more extensive here, include some more operationally used LSMs.
Reply: The list of LSMs will be expanded in the revised manuscript.

Page 3, Line 25: correct parenthesis around the reference.

Page 4 Line 6: few => a few

Page 4 Line 13: observations are available => observations available

Page 4 Line 16: combine => combines

Page 5 Line 16: were => was

Page 5 Line 23: of => between

Page 6 Line 11: reach => reaches

Page 6 Line 24: leaf => leaves

Other textual English improvements should be double checked as well.

Reply: All these suggestions will be corrected