To Referee #1:

Dear Referee #1:

Thank you very much for your valuable comments and suggestions. We have revised the manuscript according to your comments and suggestions. We hope that the manuscript has been improved and is now acceptable for publication.

The detailed responses to your comments follow.

Yours sincerely,

Tomomichi KATO

****

General Comments:

1. **This study investigates the use of environmental observations to improve the estimation of the parameters of a land surface model over an African savannah site. Although there is a need to investigate data assimilation issues in view of the use of future satellite sensors able to observe land surfaces at a high spatial resolution, the subject and methods of the study are not new. The authors should clearly demonstrate the added value of this study. The introduction should be revised in order to better describe the focus of this study. Generally, the model and CCDAS descriptions are too elusive. Re-ferring to past publications is not enough and more details must be given in the text so that the reader can understand the protocol used by the authors. The model description given in the Supplement should be completed by an Annex to the main manuscript part including the definition and the role of the parameters listed in Table 2.**

   Answer: Thank you for your comments. We have modified the description of CCDAS to include several central parts of the model formulations, e.g. the calculation of FAPAR. The 24 parameters to be optimized are now explained in the Method section. However, we still keep the larger part of the model description in the Supplement material, because otherwise the main manuscript would become too long.

2. **The result Section is very short (only 2 Tables and 5 figures) and this gives the impression that the data were not completely analyzed, and modeling issues not completely addressed. The results indicate that the maximum plant-available soil moisture Wmax is the more sensitive parameter. This is not a surprise. While ecophysiological parameters can be derived from an analysis of the literature, Wmax is a local property resulting from the soil characteristics and the vegetation type. Wmax is not directly observable and constitutes the most uncertain parameter in land surface models. This raiseses the question of the usefulness of the complex optimization procedure used in this study. One could have tried to search for physiological parameter values using local observations or a literature review, and optimize Wmax, only, or together with one or two key parameters of photosynthesis and plant growth. Trying to optimize 24 parameters for 2 PFTs at the same time may be disinformative for most of them.**

Answer: We have moved the discussion of the posterior parameter error-covariance matrices (Tables 3, 4, and 5) from the supplemental information to the main manuscript text as we agree that this is an essential part of the analysis of the results. We added the discussion in second paragraph of subsection 4.1., as follows,

> "For all three experiments, the posterior error covariance matrixes of the 24 parameters show values of less than 0.1 for the error covariances of $W_{\max}$ with all other parameters (Tables 3, 4, and 5), suggesting that $W_{\max}$ can be independently constrained by LHF and FAPAR observations."

and in the fourth paragraph of subsection 4.1.,

> "Focusing on the $V_{\max}^{25}$ parameter, interestingly, Experiment 3 shows a fairly high negative error covariance with the respective $f$ci parameter (ratio of CO_2 concentration inside the leaf tissue to the outside concentration) of -0.43 for C3 trees and -0.25 for C4 grass as shown in Table 5. While the posterior $V_{\max}^{25}$ value of 34 μmol $m^{-2}s^{-1}$ for PFT 2 in Experiment 3 is

smaller than that in Experiment 2 (78 μmol $m^{-2}s^{-1}$), and thus should lead to lower GPP in Experiment 3 as compared to experiment 2, GPP is in fact in Experiment 3 larger than in Experiment 2. This seems to be caused by the larger f_{ciC3} value, which increases the CO_2 uptake of plants by photosynthesis, to some extent."

In terms of W_{max} optimization, we have carefully considered your suggestion of the advantage of doing two step optimization: first doing the optimization only on W_{max} and being followed by the optimization on the 24 parameters. Despite a great reduction in relative uncertainty by W_{max} (Fig. 2), this relatively large prior value of W_{max} (1500 mm) would not have a strong impact on other parameter optimizations as indicated by a calculation of impact on cost function, which is done by changing prior W_{max} from 1500 mm to 500 mm in response to your particular comment No.10. Also, the adjoint method of calculating the steepest slope in the cost function in this study allows us to optimize a large number of parameters simultaneously in theory. So, this implies that the results should be the same no matter which we take two step optimization or single step optimization of the 24 parameters, as long as the prior and the uncertainty are not changed. Therefore, in our view the data assimilation scheme that we have applied here is to work indeed an appropriate way to obtain the optimized value of 24 parameters, including W_{max} or other large number of key parameters at the same time.

Particular Comments:

1. **Figures: Fig. 1 is never mentioned in the text. Figure 5 is cited before Fig. 2. Figure 2 is difficult to interpret (too many points). It should be split into several sub-figures**

Answer: Figure 1 is quoted in the text now. Figure 2, 3, and 4 are revised with a clearer color and symbol coding, and the previous Figure 5 has been changed to Figure 2 now.

2. **P. 3617 (abstract): the sentence "The closest agreement is found for each observed data stream when only the same data stream is assimilated" is unclear. Moreover, the last sentence of the abstract is confusing as LHF and satellite FAPAR are not measured at the same spatial scales.**

   Answer: We revised it in the abstract as a follow,

   > "On the other hand, the closest agreement of the model simulations with one of the data streams, is found when only the same data stream, i.e. LHF or FAPAR, respectively, is assimilated"

3. **Sect. 1 (Introduction): The open literature concerning the data assimilation related to vegetation variables is not completely described. Nice references can be found in Biogeosciences and in other journals. The vision the authors have of data assimilation is a bit restrictive. This paper considers the optimization of model parameters, while in monitoring systems data assimilation consists in integrating observations into models in order to continuously update the modeled state variables (e.g. Barbu et al. 2011).**

   Answer: We added several references concerning recent progresses in data assimilation as follows,

   > "Recently, Barbu et al. (2011) have applied the Simplified Extended Kalman Filter to assimilate the Soil Wetness Index (SWI) and satellite-derived Leaf Area Index (LAI) in order to optimise the ISBA-A-gs land surface model for French grasslands. They could achieve a significant improvement in the root-zone soil water content of around 13% as compared to results from the prior model. Applying the same assimilation system for grasslands and croplands in France, Clavet et al. (2012) also found that the maximum available soil water capacity has a large influence on the correlation between the model and the agricultural statistics."

4. P. 3619, L. 5: “biososphere” ?

Answer: It was corrected.

5. P. 3619, L. 7 and P. 3620, L. 7: again, LHF observations are local, and the rationale for trying to merge them with low resolution variables has to be explained/justified.

Answer: That can be explained by the aerodynamic fetch of the eddy covariance method. Eddy flux measurements are influenced by a basal source area, which is determined by the aerodynamic conditions and the measurement height. Roughly speaking, a large fraction (more than 80 %) of the total measured flux originates from a lateral extent to, at maximum, 100-200 times the length of the vertical distance between the 3D sonic anemometer and the zero displacement height (*d*) towards upwind direction. At the Maun site, a previous study by Veenendaal et al. (2004) calculated the 90% fetch distance to be 520 m for daytime data in March 2000. We added the following explanation,

> “Eddy flux measurements are influenced by contributions from a basal source area, whose size and position is varying depending on the aerodynamic conditions: wind direction, friction velocity, atmospheric stability, etc. At this site, Veenendaal et al. (2004) estimated the mean 90% fetch distance of the installed eddy correlation measurement system to be 520 m for daytime data in March 2000 by a footprint analysis according to Schuepp et al. (1990) and Kolle and Rebmann (2002). Assuming that a fetch radius of 520 m can be enlarged occasionally depending on aerodynamics conditions, the possibly larger maximum fetch area is comparable to the footprint of the FAPAR measurement as a rectangle with length and width of 4500 m (see details in FAPAR data in

below)."

6. **P. 3621, L. 12: LHF: what about the closure of the energy balance for this site ? It should be mentioned here that the energy balance is not closed for this site.**

   Answer: We already mentioned about the closure of energy balance in the last paragraph of subsection 2.3.

7. **P. 3622 (top): what is the spatial resolution of the FAPAR product ?**

   Answer: We use the average value of 3x3 pixels from original FAPAR data with a spatial resolution of 1.5 km, as it is mentioned in the fourth paragraph of subsection 2.1.

8. **P. 3622, L. 25: what is the time step of the BETHY model (hourly ?, daily ?). Two or three more sentences are needed in order to describe this model further. Since FAPAR is assimilated, it should be made clear how FAPAR is simulated by the model.**

   Answer: We revised as a follow.

   "BETHY calculates the energy balance (including LHF), photosynthesis (including FAPAR) and autotrophic respiration on an hourly time step, and phenology, hydrology and heterotrophic respiration on a daily time step using the before-mentioned climate input data (Fig. 1)."

9. **P. 3623, Sect. 2.3: More details should be given on the assimilation algorithm. In particular, the difference between prognostic and diagnostic variables should be made clear. The very large bias between the observed and prior FAPAR is a big issue. The assimilated variables should not be biased too much.**

Answer: First, we added a brief explanation of the assimilation algorithm in the last paragraph of subsection 2.2.

> “Differences between simulated LHF and FAPAR values and the observed data are minimized by optimizing model process parameters. Here we only briefly summarise the main methodological aspects. There are two modes in our data assimilation process: a calibration mode and a diagnostic/prognostic mode. In calibration mode, the optimal parameter set is derived from FAPAR and LHF observations by propagating the observational information in an inverse sense through a chain of models. The mismatch of modeled values to observations is defined as a cost function as explained in the following subsection 2.3, and model parameters are then calibrated through iterative parameter adjustment (using the gradient information provided by the adjoint model) until the cost function reaches a minimum. In diagnostic/prognostic mode, the quantities of interest (i.e. LHF and FAPAR) and their uncertainties can be calculated from the optimized parameter vector and its uncertainty as derived in the calibration mode. When the model is run in diagnostic mode the quantities of interest are calculated for the same time as the assimilation window whereas in prognostic mode they are calculated fo a time period outside the assimilation window. For detailed information on the CCDAS methodology we refer to Kaminski et al. (2003), Scholze et al. (2003), Rayner et al. (2005), and Scholze et al. (2007).”

We do not consider the difference between the observed and the prior FAPAR as a bias because there is no systematic reason for this difference as commonly anticipated when referring to bias. This difference is rather reflecting our prior knowledge for the phenology scheme in this case. We employ here an un-tuned phenology scheme with prior parameter values we consider reasonable by expert knowledge. These prior phenology parameter values have not ‘seen’ any observational data in order to be able to assimilate the FAPAR data. Any previous tuning of the phenology scheme with remote sensing data before the assimilation of FAPAR would not be independent of the assimilated FAPAR data.

10. **P. 3624, Sect. 2.4: the considered biomes includes 2 PFTs (trees and C4 grass), it should be made clear which parameter values apply to the 2 PFTs or not. It seems to me that most parameters should display contrasting values from trees to C4 grass. I was not able to see these differences in Table 2. The prior Wmax value in Table 2 (1500mm) is completely unrealistic. Why using such a value ? I have the impression that the authors prescribed unrealistic parameter values on purpose in order to show a dramatic impact of the CCDAS. The prior parameter values should be based on published standard values of these parameters.**

Answer: We have already indicated which parameters are used for PFT 2 or 10 only, or for both in the second row of Table 2. In addition, we added the following sentence at the beginning of section 2.4.

> "8 parameters are commonly used for both PFT 2 and 10, while the remaining parameters are used either for PFT 2 or PFT 10 only (as shown in the second row in Table 2)."

As we discuss in section 4.1., previous studies have shown that the rooting depth should be deeper than, at least, 1.0 m. Therefor we have chosen 1.5 m as a prior value, which is close to the 1.44 m suggested by Schenk and Jackson (2002). In general, plants would grow their roots deeper to survive water-limited time periods in arid environments. However, these literature values can also be regarded as "potential" maximum values of rooting depth by a few well-adapted plant species, which don't necessarily correspond to "effective" maximum values, and as such representing the mean behavior on water availability by the major dominant species. So we understand that the 1500 mm is not an unrealistic prior value. The lower optimized $W_{\max}$ (86 to 332 mm) is also acceptable given the heterogeneity in rooting depth. Nevertheless, we demonstrate here with a brief calculation the effect of different prior values on the cost function at the end of optimization from the posterior values. As mentioned already, W_{max}, which shows the highest reductions in relative uncertainty (Fig. 5), has a prior value of 1500 mm. Now consider changing the prior value from 1500 to 500 mm, and everything else being the same, its impact on the cost function J (ΔJ) against the change in the cost function from prior to posterior can be estimated for the three experiments as follows:

Exp1: ΔJ from lower W_{max} prior is 0.60 (while J changes from 470 (prior) -> 313 (posterior: the reduction is 157)).
Exp2: ΔJ from lower W_{max} prior is 0.81 (while J changes from 1825 (prior) -> 32 (posterior: the reduction is 1793)).
Exp3: ΔJ from lower W_{max} prior is 0.77 (while J changes from 2295 (prior) -> 908 (posterior: the reduction is 1387)).

Thus, the contribution of a change in W_{max} prior from 1500 to 500 mm is small and does not make a large difference in the cost function J, suggesting that the results would be similar with a lower Wmax prior value. This is mainly due to the large uncertainty of Wmax (1500mm).

We also mention here that W_{max} is actually divided into several W_{max} s for each PFT separately in the optimization and re-distributed such that the average W_{max} s of the grass PFT is 30% of the average W_{max} s of the tree PFT, according to their cover fraction. In this study, PFT2 (tree) and PFT10 (C4 grass) cover 0.7 (frac2) and 0.3 (frac10) of ground surface, such that W_{max}_tree and W_{max}_grass are set as:

W_{max}_tree = W_{max} / (frac2 + 0.3 * frac10) = 1500 / 0.79 = 1899
W_{max}_grass= W_{max} / (frac10 + frac2 / 0.3) = 1500 / 2.633 = 570

as prior values. According to the above-mentioned literature we believe that these values are realistic prior assumptions. We added the explanation for this individual separation of W_{max} for grass and tree PFTs in the main text in subsection 4.2.

> “$W_{\max}$ is divided into several $W_{\max}$s for each PFT separately in the optimization and re-distributed such that the average $W_{\max}$ of the grass PFT is 30% of the average $W_{\max}$ of the tree PFT according to their cover fraction. In this study, PFT2 (tree) and PFT10 (C4 grass) cover a fraction of 0.7 (frac2) and 0.3 (frac10) of ground surface, such that prior $W_{\max}$ (tree) and $W_{\max}$ (grass) are set as follows; $W_{\max}$ (tree) = $W_{\max}$ / (frac2 + 0.3 * frac10) = 1500 / 0.79 = 1899 mm, $W_{\max}$ (grass) = $W_{\max}$ / (frac10 + frac2 / 0.3) = 1500 / 2.633 = 570 mm.”

11. **P. 3625, L. 19: the substantial reduction in the Wmax value is mainly due to the unrealistic prior value.**

Answer: We already addressed this comment above (No. 10).

12. **P. 3627 (top): why is there a time lag between the observed and simulated maximum FAPAR in Fig. 3 ?**

Answer: Indeed the maximum FAPAR is slightly delayed in the simulation. However, LHF and GPP show an adequate maximum period in the simulations as compared with the observations. So we think that the delay in maximum FAPAR is due to the phenology scheme

13. **P. 3627, L. 17: Wmax=332mm is not a small value (see for example Calvet et al., GMD, 2012). “General belief” is not a proper reference.**

Answer: We revised the sentences of first paragraph in section 4.1 as follows,

> “$W_{\max}$ is consistently constrained to a relatively small value in all three experiments (86 to 332 mm) compared to the prior value (1500 mm), which is close to the reported rooting depth in such dry conditions. For example, Schenk and Jackson (2002) suggested that dry tropical savannas have on average a rooting

depth of 1440 mm containing 95% of the total ecosystem roots. In fact, Veenendaal et al. (2008) showed that the tall and short mopane trees rooted at least deeper than 1.0 m by field measurements at the Maun site. However, they also indicated that the total root density of both mopane types as well as the fine root density of short mopane were concentrated in the upper soil fraction up to 200 mm depth. Moreover, Calvet et al. (2012) performed a sensitivity analysis on the ISBA-A-gs land surface model and showed that the median value of maximum available soil water capacity content was estimated to be around 129 mm. With this value they could simulate the interannual variability in the productivity reasonably well for both C3 crop and C3 grassland in France, where the climate is more humid and the plant productivity suffers less to drought than at the Manu site. This suggests that the active layer for soil water uptake would be in the shallow soil layer, supporting the smaller W_{max} values from our optimisations as compared to the prior value."

14. P. 3629, L. 8: "paramters" ?

Answer: It was corrected.

15. P. 3629, L. 10: does it mean that the functional relationship between Vmax and fci is not sufficiently accounted for by the BETHY model (at least for trees) ?

Answer: Thank you, this was a bit unclear in the original manuscript. What we wanted to say is as follows. There are three facts: 1) in experiment 3 there is a negative error covariance between V_{max}^{25} and f_{ci} for the tree PFT, which is not present in experiment 2. 2) the posterior V_{max}^{25} value is smaller in experiment 3 (34 μmol $\mathrm{m}^{-2}\mathrm{s}^{-1}$) than in experiment 2 (78 μmol $\mathrm{m}^{-2}\mathrm{s}^{-1}$), which should lead to a smaller simulated GPP in experiment 3 compared to experiment 2. 3) But in fact GPP is simulated larger in experiment 3 than in experiment 2.

To explain the reason for larger GPP in ex. 3 than in ex. 2 even under a smaller V_{max}^{25} value, we think that the higher posterior f_{ci} value for the tree PFT counterbalances the negative effect by the smaller V_{max}^{25} value on photosynthesis as shown by the negative error covariance of f_{ci} with V_{max}^{25}. We revised the sentence in the fourth paragraph of subsection 4.1., as follows,

> "Focusing on the V_{max}^{25} parameter, interestingly, Experiment 3 shows a fairly high negative error covariance with the respective $f$ci parameter (ratio of CO_2 concentration inside the leaf tissue to the outside concentration) of -0.43 for C3 trees and -0.25 for C4 grass as shown in Table 5. While the posterior V_{max}^{25} value of 34 μmol $m^{-2}s^{-1}$ for PFT 2 in Experiment 3 is smaller than that in Experiment 2 (78 μmol $m^{-2}s^{-1}$), and thus should lead to lower GPP in Experiment 3 as compared to experiment 2, GPP is in fact in Experiment 3 larger than in Experiment 2. This seems to be caused by the larger $f\text{ci}_{C3}$ value, which increases the CO_2 uptake of plants by photosynthesis, to some extent."

16. **P. 3631, L. 9-21: Problems in assimilating FAPAR may be caused by the way this quantity is simulated by the model. More details have to be given about the simulation of FAPAR and about the radiative transfer model used in BETHY. Please explain why the joint assimilation of LHF and FAPAR degrades the LHF score that much. Please explain why the assimilation does not significantly improve the GPP simulation.**

Answer: First, we added a more detailed description of the FAPAR calculation in 2.2 as follows,

> "FAPAR is calculated as the vertical integral of absorption of photosynthetically active radiation by healthy green leaves divided by the difference between the incoming and outgoing radiation flux at the top and bottom of the canopy (Knorr et al., 2010). This integration is carried out by a two-flux scheme, which takes into account soil reflectance, solar angle and amount of diffuse radiation. Equating satellite and model

> FAPAR means that given the same illumination conditions, the same number of photons enter the photosynthetic mechanism of the vegetation, even if some of the assumptions differ between BETHY and the model used to derive FAPAR (Gobron et al., 2000)."

As we show in the manuscript assimilation of only FAPAR results in simulated FAPAR very close to the observed FAPAR (experiment 2; Fig. 4). This proves that both FAPAR calculation and its assimilation system work very well to fit the simulation to the observation. On the other hand, the joint assimilation of LHF and FAPAR (experiment 3) cannot bring the lowest value in cost function of posterior LHF, among three experiments. We think that unfortunately both data sets and the model formulation are still inconsistent, and therefore it's impossible for optimizing the parameters for both data streams to be well fitted at this moment. Consequently, simulation of GPP in ex. 3 by the optimized 24 parameters also could bring the lowest RMSE against the observation among three experiments. Although GPP is not a target for assimilation, this is something different from our expectation that the LHF could contribute positively to optimizing the photosynthesis-related parameters to fit the simulated GPP in exp 3 better than in exp 1. On contrary, we can say that this inconsistency is one of the most remarkable points in this study that there is still large space to reduce potential bias or error in assimilated both observations and to improve the model performance.

17. **P. 3632, L. 15: The FAPAR scale issue is even more acute for SMOS (spatial resolution of about 40km): : : However, a number of authors have shown that low resolution products can correlate quite well with local in situ soil moisture observations.**

Answer: We think that the simultaneous assimilation of multiple data streams is the only way forward to better understand both model formulation and the observations. It is important to explore the consistency in the observational data with the model. This provides important information for improving the quality of both model and observations. We put the following part in the conclusion,

> “Despite of SMOS’s lower spatial resolution (35-50 km)”