Reply to anonymous referee 2:

We would like to thank the referee for their review of this manuscript and their constructive comments. Below is a response to each comment (the referee's comments have been included in italics).

"• What is the temporal resolution of CCDAS?"

The temporal resolution of CCDAS in the version used here is daily.

"• page 667 equation 3 and 4: How are the carbon pools fed. This should be made clear as well as the dependence on pools sizes, NPP and beta parameters (eq 7)."

As described in Rayner et al. (2005), input to the fast pool is parameterised by the annual course of LAI for deciduous PFTs and the constant fraction of the leaf carbon pool for evergreen PFTs.
The relationship between the slow respiration flux and long term mean NPP determines the overall carbon balance.

"• page 668 line 21ff: Does this hold for all kind of experiments or only for the Gaussian one?"

This holds for all experiments.

"• section 2.3.3: The transformation should be given as formula, otherwise one of the core aspects of this work is not reproducible."

We have included the equations explaining the transformations in the paper.

"It would be also good the have a more descriptive discussion of the implications of the different transformations (and other methods) on the interpretation of the results, given the underlying Bayesian paradigm. How do the different methods influence the interpretation of the posterior results as a joint probability? Is there a justification to prefer one method or are the proposed transformation purely pragmatic solutions?"

We have added the following paragraph to the manuscript to describe the differences in the transformations: "The essential difference between the three approaches is the form of the prior (and thus posterior) pdfs in (physical) parameter space. Both the constrained and the penalty function approaches produce a prior of Gaussian shape inside the bounds/the non-penalised region. Outside, the constrained approach produces a zero probability while the penalty approach produces a non-zero probability consisting of a gradual reduction of the Gaussian probability with increasing distance from the bounds. The double bounded transformation approach produces a zero probability outside the bounds and a non- zero but non-Gaussian probability within the bounds."

Obviously, we recommend (as stated in the manuscript) the parameter transformations as the preferred method, however there is no objective criteria to prefer either the quadratic or the logarithmic transformation.

"• page 671 line 27ff: Is there a reason why to chose those values and no others."

We chose a value of 10^4 for the penalty factor as it is on the same order of magnitude as the cost function at the previous minimum we had found. We need a value large enough to affect the optimisation but not so large that it dominates. We chose the sensitivity value to be 4 as it needs to be even, as mentioned in the text, but we did not want 2 as we want to have a non-constant second derivative.

"• page 672 line 7ff: Were the 5 starting points for the different experiments the same or did they also change within the experiments? And why were they changed by 10"

One of the starting points was the same for each of the experiments (that which used the prior parameter values). The other 4 optimisations for each experiment had different starting points. The starting points were varied by +/- 10% to allow some variation to the starting points but not so much that they were too far away from the prior parameter values.

"• Fig 2: Not necessary"

We agree and this has been removed.

"The authors state that the unsuccessful experiments are of little use. But in the context of this work, where strategies to avoid non-physical parameters are investigated, an understanding of the failure of the methods would help to interpret the results and increase its relevance for other similar studies. One method does not converge, one method and at least 1 out of 5 starting points of the other methods still yield non-physical parameter values."

The manuscript includes now a more complete description of the unsuccessful experiments. In fact, for the penalty experiments, all five optimisations did converge and found a minimum in a mathematical sense (zero or at least close to zero gradient) but still the parameter that had been selected for bounding (parameter 18) had a non-physical value and therefore we discarded this method. For the constrained experiments, four out of the five optimisations failed to converge because of internal overflow problems within the optimiser and the fifth one (started from the default parameter values) stopped because of reaching the maximum number of iterations (5000, which is about 10 times more than the average number of iterations for the parameter transformation experiments). Also, for this optimisation, the selected parameters for bounding were exactly at their limits, which at least in the case of 0 for a beta parameter does not make sense.

"Interestingly these last cases show a smaller cost compared to the successful experiments. This could be interpreted such that the optimal solution is only found

with non-physical parameters. Maybe the application of a parameter transformation is not the one solution to the problem."

It is possible that the global minimum is within the non-physical space because the model is highly non-linear with a complex, 19-dimensional parameter space and from a purely mathematical point of view a smaller minimum can be found outside the physically meaningful parameter space. However, this does not constitute a solution for our optimisation problem. Another possible reason for finding a minimum outside the physically valid parameter space is that the model, as it stands, is missing or does not fully describe a relevant process and therefore the optimisation has to compensate for this missing process by choosing non-physical parameter values. One way to resolve this could be further model development and include missing processes in the model formulation. But this is not always feasible and therefore we use the current model formulation with parameter transformations to ensure physically meaningful parameter values.

"For sake of completeness, I also suggest to add the penalty and constrained cases to table 3."

The penalty and constrained parameter values have been added to Table 3.

"As already mentioned by the authors themselves, the results of Koffi et al. (2012), that the parameter transformation change the results is not found here. What is the authors view on this. Why is this the case for this relatively similar systems? What can be learned from this discrepancy?

While Koffi et al. (2012) use a relatively similar system there are some important differences to our system that prevented us from discussing the different behaviours in our manuscript. Firstly, they use in their experiment the full CCDAS including the photosynthesis and autotrophic respiration processes in the optimisation. Hence they are optimising 57 parameters. We use here a simplified version of CCDAS, which only includes the heterotrophic respiration and carbon balance processes in the optimisation and optimises altogether 19 parameters, a factor of 3 less parameters resulting in a factor of 3 less dimensions in the parameter space. Secondly, in the experiments of Koffi et al. (2012) the optimisations did not converge to a minimum, they have stopped the optimisation iterations after a certain reduction of the cost function value without obtaining a near zero gradient. In our experiments for this manuscript here, all the optimisations with parameter transformations have converged to a minimum with a final gradient approaching zero.
We have clarified the differences between the two studies in the manuscript.

"• Figure 7 could be omitted. If I interpret this correctly, the differences are several orders of magnitudes smaller then the correlations themselves. It would be enough to only mention this in the text. Otherwise, this needs to be discussed in more detail."

Figure 7 has been omitted; the differences are indeed several orders of magnitudes smaller then the correlations.

"• The discussion of the convergence (figure 4) should be extended. Why do the different methods converge differently fast. How do the different starting points behave? Are their robust interpretation of the convergence behaviour?"

Thank you for this suggestion, we have picked up this point and added a discussion on this in the manuscript. In brief, the different methods and also the minimisations from different starting points converge differently as they are solving different problems. Each change in the formulation of the cost function results in a different optimisation problem. When an optimisation begins at a different starting point in the control space, it follows a different trajectory to find a minimum.

"• It would be good to add to figure 5 also NEP obtained with the prior parameters and also that of CONS and PEN. Then the differences should be interpreted. Otherwise figure 5 does not provide additional and necessary information."

Our terrestrial biosphere model assumes a balanced carbon budget over the simulation period (i.e. a long-term mean NEP value of 0) for the prior parameter values. A dedicated parameter (beta) scales the product of the size of the slow decomposing soil carbon pool and its turnover time to adjust for the terrestrial sources and sinks of CO_2. This has already been described in the original CCDAS paper by Rayner et al. (2005). Therefore, we don't think it makes sense to include prior NEP in figure 5. Furthermore, as the parameter values from the CONS and PEN experiments are reasonably similar to the values from the parameter transformation experiments, the plots of NEP are imperceptibly different (around 2-4%) so these will not be added.

"• Table 2: For completeness I suggest to add the values obtained with the prior parameters as well."

Whilst we don't think it would make sense to include the values obtained with the prior parameter values in Table 2, we have included the cost function value obtained from the prior parameter values in the text.

"page 664 line 19: Which transport model is used?"

The transport model is TM2, it is stated later in the text. We have included this at the first mention of the transport model.

"page 665 line 11: "NPP parameters" should only be NPP?"

This has been changed.

"page 668 line 5: This should be table 3"

This has been corrected.

References

Koffi, E. N., Rayner, P. J., Scholze, M., & Beer, C.:Atmospheric constraints on gross primary productivity and net ecosystem productivity: Results from a carbon-cycle data assimilation system. Global Biogeochem. Cycles, 26(1), GB1024. AGU. doi:10.1029/2010GB003900, 2012.

Rayner, P. J., Scholze, M., Knorr, W., Kaminski, T., Giering, R., & Widmann, H.: Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS). Global Biogeochem. Cycles, 19(2), GB2026. AGU. doi:10.1029/2004GB002254, 2005.