Reply to the editor:

many thanks again for taking your time to handle our manuscript. Frankly speaking, we are a bit puzzled by your persistent decision asking for major revisions and redoing the identical twin experiments for the following reasons and would request you to reconsider your decision.

First, the Referee #2 classified our manuscript in three out of the four categories as 'excellent' and in the fourth category as ‘good'. Furthermore, she/he did not ask for any additional experiments but for 'incorporating some of the discussion in your review response' into the manuscript. We exactly did that in our last revision.

Second, we disagree that the identical twin experiment has not worked. As explained in our response, due to the high complexity and non-linearity of LPJ-GUESS, we expect equifinality issues, and that is exactly what we encountered with the identical twin experiments (as explained in the response as well as in the manuscript). We do see, however, when comparing the prior and posterior time series of the simulated CH4 emissions with the simulated observations that the posterior time series does capture the time series of the simulated observations much better than the time series using the prior parameter values (see Figure 1 here and also Table 4 in the manuscript listing the prior and posterior cost function values). Hence, we believe there is value in the identical twin experiments. This is also acknowledged by the referees.

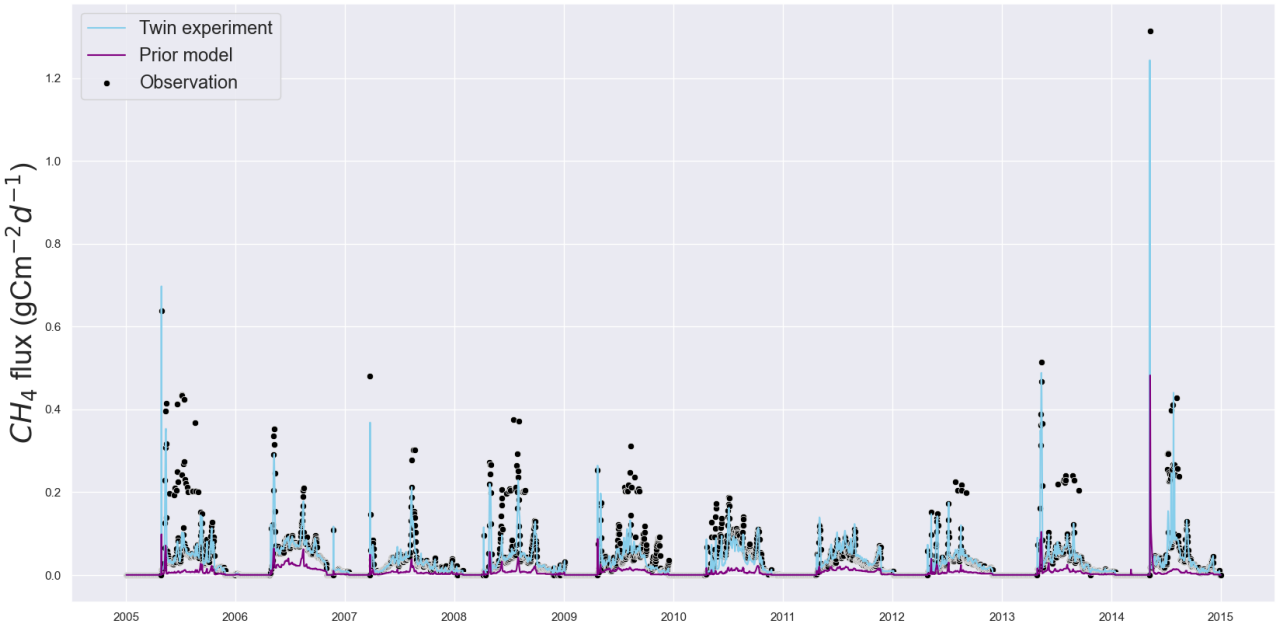

Figure 1: Time series of simulated observations together with prior and posterior CH4 emission values from one of the twin experiment. The prior is shown in purple, posterior in blue and the twin observation in black

We also note that our study is not the only one encountering equifinality issues (or more generally problems) in retrieving the true parameter values from identical twin

experiments. Many similar studies working with high-dimensional Bayesian problems and non-linear models both in the Earth Sciences but also other fields have reported similar equifinality and parameter correlation issues, e.g. Hargreaves and Annan (2002), Apte et al. (2008), Annan and Hargreaves (2010), Long et al. (2010), Santaren et al. (2014), Lamminpää et al. (2019), just to name a few. Here, especially the study by Santaren et al. (2014) is comparable to ours as it also concerns model parameter optimisation against greenhouse gas flux observations, but it should be noted that we have achieved a much better parameter retrieval with the methodology employed in our study.

Considering your suggestion in your previous letter to investigate if the twin experiment can recover the true parameter values in a simplified, low-dimensional problem, we have conducted two sets of twin experiments with different sets of parameters. In the first set of parameters, we included the parameter CH4/CO2, and in the second set, we included the parameter lambda_root. These are the two parameters that are offsetting each other in the optimisation process. The results are given below in Figures 2 and 3. From these figures, it is clear that our algorithm is well capable of retrieving true parameters in a low-dimensional problem.

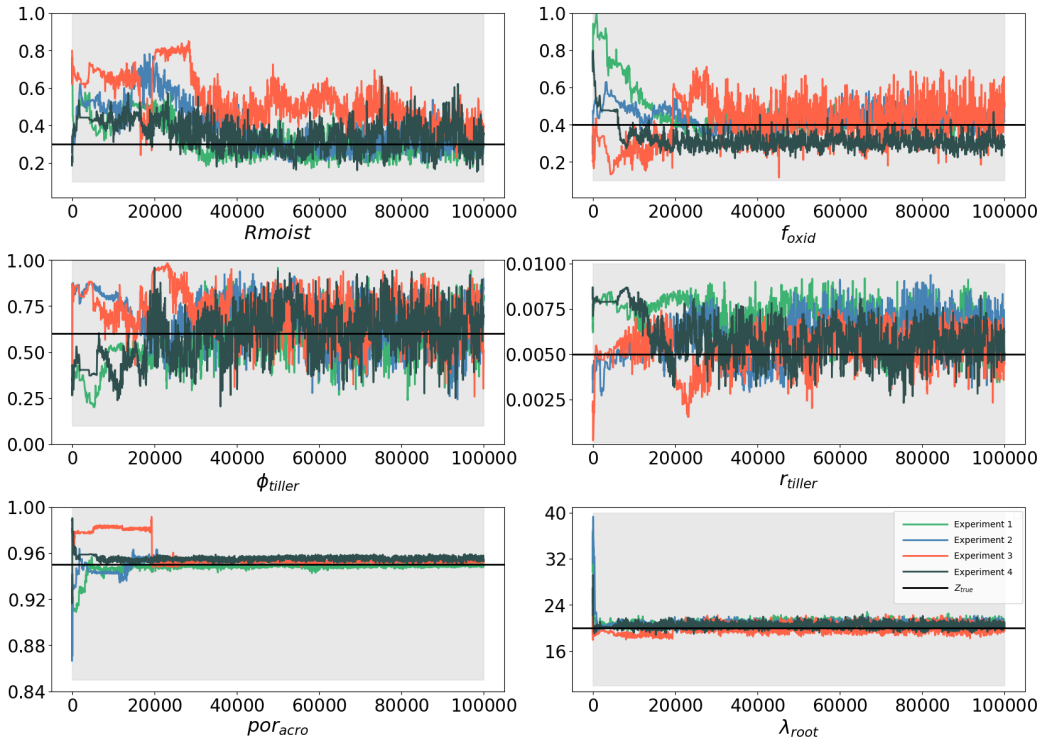

Figure 2: Four sets of twin experiments with six (incl. lambda_root and excl. CH4/CO2) out of the eleven parameters (optimised in the manuscript) started at different points in the parameter space. The straight black line in each panel indicates the true parameter values used to generate the identical twin observations.

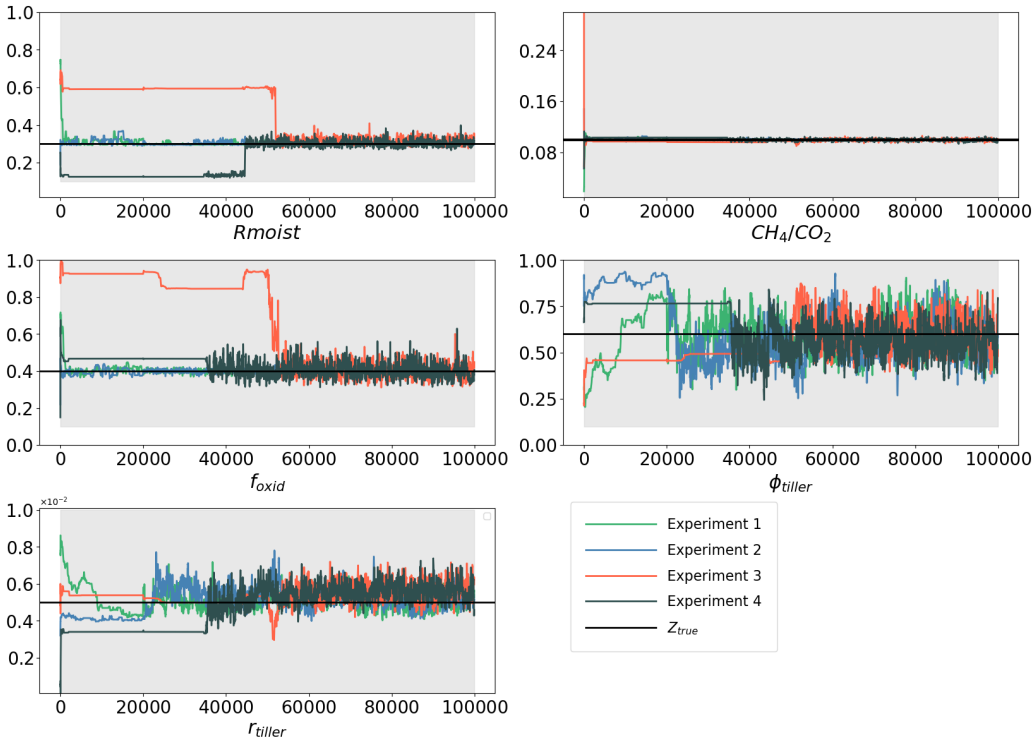

Figure 3: Four sets of twin experiments with five (incl. CH4/CO2 and excl. lambda_root) out of eleven parameters (optimised in the manuscript) started at different points in the parameter space. The straight black line in each panel indicates the true parameter values used to generate the identical twin observations.

Finally, regarding your suggestion to include observational uncertainty in the identical twin experiments, we have already incorporated observational uncertainty to estimate the cost function.

Best regards
Jalisha

References:

1. Annan, J. D., and J. C. Hargreaves. "Efficient identification of ocean thermodynamics in a physical/biogeochemical ocean model with an iterative Importance Sampling method." Ocean Modelling 32.3-4 (2010): 205-215.
2. Apte, A., C. Jones, and A. M. Stuart. "A Bayesian approach to Lagrangian data assimilation." Tellus A 60.2 (2008): 336-347.
3. Hargreaves, J. C., and J. D. Annan. "Assimilation of paleo-data in a simple Earth system model." *Climate Dynamics* 19 (2002): 371-381.
4. Lamminpää, O., et al. "Accelerated MCMC for satellite-based measurements of atmospheric CO2." Remote Sensing 11.17 (2019): 2061.
5. Long, K. J., ,S. E. Haupt, and G. S. Young. "Assessing sensitivity of source term estimation." Atmospheric environment 44.12 (2010): 1558-1567

6. Santaren, D. et al. "Ecosystem model optimization using in situ flux observations: benefit of Monte Carlo versus variational schemes and analyses of the year-to-year model performances." Biogeosciences 11.24 (2014): 7137-7158.