We thank the Anonymous Referee #1 for his/her constructive suggestions for improving the paper. Our specific reply is below with Referee's comments in *italics*.

1.) The discussion section states that this method is applied for the first time to a GCM. Jackson et al., Error reduction and convergence in climate prediction (Journal of Climate, 2008) chose a comparable approach to estimate parameters in a general circulation model (CAM3.1) in order to select an ensemble of plausible climate model configurations to narrow the range of simulated climate sensitivities. Jackson et al. include a relatively large number of observed variables in their cost function. In the here presented manuscript, the importance of the choice of the cost function is one of the main results. It would be great if the results of Jackson et al. could be discussed here and also be mentioned in the introduction.

Many thanks for pointing out the article by Jackson et al. (2008) which was not included in the references of the original manuscript. In Jackson et al. (2008), stochastic optimization is used for the search of optimal closure parameters in CAM3.1 climate model with respect to a metric based on multi-source observations. The search paths of the optimizer are used to infer about the parametric uncertainty.

While the approach employed by Jackson et al. is related to MCMC, there are also clear differences. We emphasize that our paper is the first in applying full MCMC in this context. A primary benefit of using the MCMC is that it allows rigorous estimates of the parameter posterior distribution. However, it is also true that this technique requires more model simulations than the optimization based methods (as those used in Jackson et al.). Our purpose is to demonstrate that it is feasible to use adaptive MCMC with large-scale models. More work on the efficiency of MCMC methods is needed, and is, indeed, being currently undertaken by our research team.

Irrespectively of how parameter uncertainty is estimated, the importance in choosing the cost function is emphasized both by us and Jackson et al. (2008).

In the revised manuscript we will refer to Jackson et al. (2008) in the Introduction, and discuss our results in light of Jackson et al. (2008) in the Discussion section.

2.) In the discussion, the authors mention the general problem of compensating errors in climate models. Fig. 4 shows that the model improved according to the net ToA radiative fluxes. This was likely achieved by reducing the skill in other climate variables related to the chosen parameters, e.g. cloud cover or precipitation. This could be tested and quantified very easy.

Many thanks for this suggestion. We performed this additional study and quantification for the cost function J-G+ZONAL. Qualitatively, the results are as follows. For SW radiation separately, the time-latitude cross section reveals an improvement as validated against the CERES data. For LW separately, the result is neutral. Thus, the improvement in the net TOA radiation (Fig. 4, original manuscript) is on average due to improved SW TOA fluxes. For precipitation (CMAP data; Xie and Arkin 1997) the result is also neutral. For total cloudiness we note that it is underestimated in the default model, and in the "best" model the underestimation is increased. In summary, some independent test quantities are improved and others are deteriorated, as anticipated by the Referee #1. This point will be discussed and illustrated with a new figure in section 4.3 of the revised manuscript.

3.) The conclusions remain very vague. I suggest to elaborate more on the importance of choice of the objective function, also in dependence of the choice of climate variables depending on the parameters to estimate. I assume that the influence of the two parameters related to precipitation (CPRCON/CAULOC) have a very small influence on the net ToA radiation budget. Here an objective function including precipitation would more probably lead to convergence. The choice of the objective function probably will depend on the parameter which should be estimated. The net ToA radiation only put constraints, depending on the chosen cost function, on one of the four parameters, which indicates that the objective function needs to be revised for future experiments.

First we generally agree with the Referee: the choice of the objective function is critically important for MCMC, and what we have presented in the manuscript is certainly not the "final solution" to this question. The discussion concerning the choice of the cost function will be improved in the revised manuscript (probably in Ch. 4.2 or 5). The following points will be considered:

- Comments will be made on which spatio-temporal structures to include in the cost function, based on the findings of sections 4.1 and 4.2. Specifically, a cost function including only a single term (J_G) does not work very well, because it is possible to get a single number "right" with quite different parameter combinations, and irrespective of biases in spatial and temporal structures. A cost function employing a very detailed spatiotemporal structure (J_XY and J_G+XY) is also problematic, because monthly values for individual grid points feature a lot of random variations. In order for a cost function to properly distinguish "good" from "bad" simulations, it is necessary that the systematic impact of parameter changes is relatively large compared to random variations. Use of zonal averages (i.e., J_ZONAL and J_G+ZONAL) suppresses the random variations, which leads to better parameter identifiability compared to J_XY and J_G+XY. For the specific case of the net flux at the TOA, accurate simulation of the annual global-mean value is highly important to avoid a climate drift in coupled AOGCM simulations. This is why we consider J_G+ZONAL as the best cost function of those tested so far. At the same time, it is worth pointing out that while the use of zonal averages is simple, it is hardly the optimal choice for the description of spatial structures in the cost function. E.g., use of EOFs, as in Jackson et al. is certainly an option we will study in subsequent works.

- It will be made clear that the definition of a cost function need not be based only on the TOA radiative fluxes (net, or even LW and SW separately). Indeed, in choosing the variables included in the cost function, care should be taken that the cost function is sensitive to those parameters.

Second, we would like to address some specific points raised by the Referee #1. It is not entirely true that cost functions based on the net flux at the TOA only provide constraints on one parameter

(CMFCTOP). (We admit that the cursory treatment provided in the original manuscript can easily give this impression). In the revised manuscript, we will discuss an analysis of a longer MCMC chain performed using the J_G+ZONAL cost function (see also our response to the last comment by Referee #2). This chain indicates that i) there is a clear preference for values of CAULOC higher than default, and ii) a preferred although broad range of values for ENTRSCV can be identified. For CPRCON, cost functions based on the net flux at the TOA do not work well. However, CPRCON has much more effect on the LW and SW fluxes separately than on the net flux (LW+SW). Thus we anticipate that a cost function using the LW and SW fluxes separately would work much better. As suggested by the Referee #1, precipitation is also quite sensitive to CPRCON, and so are middle and high cloud fractions. We agree with the Referee that there is a need to improve the definition of the cost function for future work, and that the cost function should preferably include other variables beyond radiative fluxes (cf. Jackson el al. 2008). These points will be discussed in the revised manuscript.