Dear Editor,

On behalf of the co-authors, I submit a revised version of the manuscript # gmd-2014-50 for your consideration of publication in GMD.

Changes we have made in the manuscript include:

1. The phrase "short ensembles" is added to the beginning of the title to highlight the key methodology.
2. The ensemble size and integration length of our simulations are explicitly mentioned in the abstract.
3. At the beginning of Sect. 3 we clarify that the second objective of the method evaluation is to quantify the length of spin-up time.
4. In the last section (lines 803-811 of the revised manuscript) we point out that in the time step sensitivity experiments, the 3-day ensembles do not reveal the same changes in the SPCZ as seen in the 5-yr simulations, probably because systematic changes in the SPCZ involve feedbacks from the large-scale circulation that can not sufficiently spin-up in just a few days.
5. Typo and reference corrections have been made at various places following the reviewers' suggestions.

A point-to-point response to the referees' comments is attached.

Thank you for your time. We look forward to your favorable decision.

Sincerely,

Hui Wan

Reply to Referee #1

Comments: The authors have proposed a method for tuning climate models, viz. replace the traditional serial-in-time long-term climate integration by representative ensembles of shorter simulations. The advantage of the new method lies in its efficiency, which is evidenced in its less computational cost, the dramatic reduction of the turnaround time in benchmark tests, sensitivity studies and model tuning exercises. The effectiveness of the new method is demonstrated by performing two experiments using CAM5. The first experiment focuses on the model cloud and precipitation sensitivity to the time step sensitivity, the second experiment focuses on the sensitivity of radiation balance to empirical parameters that are related to cloud microphysics and aerosols. The manuscript is technically well written and easy to follow. The authors show clear evidences that the efficiency of the ensemble method is useful for the development of high resolution, computationally expensive and complex climate models. I believe that the manuscript can be accepted for publication after only minor revisions.

Response: We appreciate the referee's comments and suggestions. Our responses to the specific points are listed below.

Comment: 1. Better to revise the title as “Ensemble of shorter simulations: An efficient method for discerning climate-relevant sensitivities in atmospheric general circulation models”

Response: In the revised manuscript the title is changed into "Short ensembles: An efficient method for discerning climate-relevant sensitivities in atmospheric general circulation models”

Comments:
2. P2175, L8-11: Another example is the sensitivity of monsoon precipitation to convection schemes. The traditional numerical experiments need several sets of computationally-expensive long-term integrations:

Chen, H. et al., 2010: Performance of the New NCAR CAM3.5 in East Asian Summer Monsoon Simulations: Sensitivity to Modifications of the Convection Scheme. Journal of Climate, 23, 3657-3675

Zhou T., and Z. Li, 2002, Simulation of the east Asian summer monsoon by using a variable resolution atmospheric GCM, Climate Dynamics,19:167-180

6. P2195, L23-28: Yes, the simulation of aerosol impacts on Asian monsoon tradition- ally takes long-time of integration and the integration should cover at least the whole monsoon season:

Song, F. et al. (2014), Responses of East Asian summer monsoon to natural and anthropogenic forcings in the 17 latest CMIP5 models,Geophys.Res.Lett.,41,doi:10.1002/2013GL058705

Response: In response to comments #2 and #6, corresponding sentences in the last paragraph of the manuscript are revised:

``For example, if one were interested in identifying how seasonal features such as the Asian summer monsoon responded to anthropogenic and natural forcings (e.g., Ganguly et al., 2012; Vinoj et al., 2014; Song et al., 2014), or to changes in model formulation (e.g., Zhou and Li, 2002; Chen et al., 2010), it might be possible to generate realizations of few-month simulations, and use ensemble averages to remove multi-year and multi-decade scale noise that would otherwise require hundreds of years of simulations."

Comment: 3. For Example-1, viz. the comparison of 30 and 4 minutes time steps, it would be better to add some comparisons to the observations. As a climate modeler, I am interested to the skills of two simulations: which one is more close to the observation? I understand that the satellite measurement may not be enough in time interval to provide the observational evidences, at least a discussion is needed. Or at least the results of reference simulation can be compared to the satellite cloud, as what has been done in previous papers of CAM5 evaluation.

Response: The theme of this paper is the ensemble method, rather than the performance of the CAM5 model. We thus believe the suggested comparison with observation is out of the scope of the present paper. A detailed characterization of the time step sensitivity in CAM5 will be reported elsewhere.

Comment: 4. Similar as Figure 2 and Figure 6, could you please add a figure of precipitation and examine the well-known double ITCZ bias? Nearly all climate modelers should have interests to this.

Response: Although the double-ITCZ problem is not yet well understood, previous studies have shown that the geographical distribution of tropical precipitation is strongly connected to the large-scale circulation, moisture, heat and momentum budgets, as well as model performance in regions away from the precipitation biases (e.g., Ma et al., 1996; GFDL, 2004; Hwang and Frierson, 2013). The time scales associated with these planetary-scale features are presumably on the order of months or longer. The 3-day ensembles presented in our paper are therefore not expected to be able to capture the ITCZ biases. Indeed, the 5-yr simulations indicated that a shorter (4 min) time step leads to a slight increase of precipitation in the SPCZ in boreal winter, while the 3-day ensembles does not reveal statistically significant differences in this region.

In the revised manuscript, we added the following sentences to Sect. 5 (``Conclusions and discussion”):

"(The strategy discussed in this paper using few day simulations certainly has limitations. It cannot be used as formulated here to investigate modes of climate variability or feedback mechanisms that operate on time scales of months to years, thus could not replace long-term simulations when long time scales are important.) For example, in the time step sensitivity experiments discussed in Sect. 3, while the 5 yr simulations reveal an increase of DJF precipitation in the South Pacific Convergence Zone (SPCZ) when time step is shortened (not shown), the ensemble simulations do not indicate statistically significant differences in this region. This is probably because systematic changes in the SPCZ involve feedbacks from the large-scale circulation which can not sufficiently spin up in just a few days."

References:

Ma, C. C., C. R. Mechoso, A. W. Robertson, and A. Arakawa (1996), Peruvian stratus clouds and the tropical Pacific circulation: A coupled ocean-atmosphere study, J. Clim., 9, 1635–1645.

GFDL (2004): The New GFDL Global Atmosphere and Land Model AM2–LM2: Evaluation with Prescribed SST Simulations. J. Climate, 17, 4641–4673. doi: http://dx.doi.org/10.1175/JCLI-3223.1

Hwang, Y.-T. and D. M. W. Frierson. Link between the double intertropical convergence zone problem and cloud biases over the Southern Ocean. Proc. Nat. Acad. Sci, 110. 4935-4940. 2013, doi: 10.1073/pnas.121302110

Comment: 5. The ensemble size of shorter simulation should be highlighted in both abstract and summary part. This may provide a useful guide for climate modelers who may follow your method in their studies.

Response: The abstract and conclusions are revised.

Reply to Referee #2

Comments: The manuscript describes a methodology for extracting the climatological response of CAM5 to "fast physics" parameters affecting clouds. The paper is well written and illustrates that one can reduce the computational requirements by an order of magnitude or more from the more standard several-year model integrations. The authors consider the responses among a large number of fields as well as a comparison to a 256 experiment exploration of parameter space among 16 model parameters. In all these cases, their method is shown to be quite effective. I have no substantial comments to make. Other than a few editorial corrections, I recommend this manuscript be published.

Pg 2176 Line 12 use "alternate" rather than "alternative"
Pg 2176 Line 26 observation[s]
Pg 2180 For me analyses I and II are the same, therefore I don't understand why they are separated.
Pg 2180 line 9 . . . I don't normally see 20 days as "extended". Perhaps the sentence could do without that qualifier.
Pg 2181 line 1. "It is worth noting that not only [are] these basic features . . ."

Response: We thank Dr. Jackson for the very positive feedback and the corrections. The manuscript has been revised accordingly. Regarding the difference between bullets #I and #II at the beginning of Section 3, we clarify in the revised manuscript that point I concerns the ensemble size, while point II concerns the spin-up time.

Reply to Referee #3

Comments: The authors proposed a strategy of using ensembles of shorter simulations to explore the responses of "fast physics" in GCMs to perturbations. It was shown that the ensembles of shorter simulations are able to produce results comparable to what is produced by traditional serial-in-time multiple-year GCM simulations, but at a fraction of the computational cost. The effectiveness of this strategy was demonstrated through two examples. One example showed the cloud and precipitation sensitivity to model time step. The other example examined the sensitivity of the TOA radiation balance to microphysics and aerosol related empirical parameters. The results of this study are useful to model development and evaluation community. The manuscript is well-written and easy to understand. I recommend the manuscript be published. Specific comments: 1. On P.2177, lines 8-9 "Shallow convection is parameterized as in Bretherton and Park (2009)"; lines 13-15 "The vertical transport of heat, ... is represented following the work of Park and Bretherton (2009)"; These two references should be swapped. 2. On P.2179, line 25, there is a typo. "simulaitions" should be "simulations".

Response: We thank the referee for the positive review. The typo and misplaced references are corrected in the revised manuscript.