

Anonymous Reviewer #2

The paper discusses applying a system identification technique to find the responses of climate models to small perturbations using a single simulation. The objectives and the methodology are well described in sections 1 and 2. The method is then applied in section 3 to find how low cloud fraction and latent heat flux change globally in response to a +1K perturbation of the near-surface temperature over the Northwest Indian Ocean. The results are compared with the results of a step response simulation. The results and the potential applications of the technique are discussed in section 4.

I find the technique very interesting and I believe that it can be a powerful tool in studying the climate system and that it has various potential applications (as some described in section 4). However, the manuscript should be first improved in several ways before it is considered for publication:

- 1) It helps the reader and potential users of the technique if the paper presents a quantitative evaluation of how well the proposed technique works in the single example shown
- 2) Further discussions on how the range of the magnitude of the normalized sequences is chosen (Line 200) and how SNR and nonlinearity can be quantified are needed.
- 3) There are few other techniques proposed in the literature to characterize the linear response of climate models. They should be briefly mentioned in section 1.

We thank the reviewer for the careful consideration of our manuscript and the excellent comments. We address each of these points below after the details are elucidated.

Here are some details regarding 1-3:

- 1) Fig. 1 shows that the system identification technique underestimates (overestimates) the magnitude of the low cloud fraction (latent heat flux) response. The patterns of the response, which are simple and local, look similar. The errors in the magnitude and pattern (e.g. pattern correlation in the region of interest) should be quantified. That way how changing parameters such as the frequencies of the band- pass filtering affects the accuracy can be quantified. Among the two reasons offered for the discrepancies, contribution from nonlinearity can be eliminated by reducing the amplitude of the forcing in the step function simulation and using a pair of heating and cooling simulations. Finding the “true” linear response using the step function simulations against which the system identification technique can be verified might take some effort, but given that evaluating the performance of the technique is crucial, I think it is worth the effort to demonstrate how well the technique captures the linear response. As for the second reason, again, it would be nice to quantitatively show which range of band-pass filtering yields the best agreement with the linear response of the step function simulation at least

for this specific example. I am also wondering why the response to this specific region (Northwest Indian Ocean) and in these specific variables have been chosen.

We like the reviewer's suggestions and have included the results from two additional step response simulations to better explore the role that nonlinearity plays in the results. After conducting these simulations, we found that extracting the "true" linear response is surprisingly challenging. There are sufficient nonlinearities in the step response that quite a bit of additional work is required (indeed, this would probably be best as the subject of a separate paper) and would require methods that are beyond the scope of what we can explore in the present manuscript. We have added an additional Section (3.5) and two additional figures that provide the results of these simulations, as well as additional description that provides explanation of the difficulties in extracting this linear signal. We have included a description of several paths toward properly validating the system identification method. We thank the reviewer for identifying what our next paper will be about!

As for the region and variables, our choice was somewhat arbitrary. After we conducted our initial simulations, we looked at the results for a variety of regions and variables, identified some interesting features that were illustrative of our purpose (to explain the methodology of system identification), and showed them. The simulations certainly contain other interesting results that would likely work equally well for our purposes, but as this is primarily a Technical Note describing the methodology, we did not want to distract from our primary purpose.

2) How is the $[-1K,+1K]$ range is chosen for the normalized sequence? (Line 200). In problems concerning linear responses, a major difficulty is obtaining good SNR without violating linearity. Does this problem exist in this technique as well? If so, how the SNR and nonlinearity in depend on the magnitude of this range should be explored (and this likely vary across different regions), which requires quantifying the SNR and degree of nonlinearity. I am not suggesting that this should be done for this paper, but these issues should be (at least briefly) discussed.

The reviewer brings up several valuable points that we agree need to be discussed. We have added additional text to the conclusions section that addresses these concerns.

3) Other methods have been proposed in the literature which share closely some of the objectives of the current study. One example is the fluctuation-dissipation theorem (some recent examples: Gritsun and Branstator 2007, Ring and Plumb 2008, Cooper and Haynes 2011, Fuchs et al 2015). Another example is the Greens function method of Hassnazadeh and Kuang 2016. Note that these methods too can determine (at least theoretically) the full dynamic behavior of the system as they provide the linear response function of the model. I suggest to briefly mention other methods such as these two to better connect the current paper with the literature.

Gritsun, A. and Branstator, G., 2007. Climate response using a three-dimensional operator based on the fluctuation-dissipation theorem. *Journal of the Atmospheric Sciences*

Ring, M.J. and Plumb, R.A., 2008. The response of a simplified GCM to axisymmetric forcings: Applicability of the fluctuation-dissipation theorem. *Journal of the Atmospheric Sciences*

Cooper, F.C. and Haynes, P.H., 2011. Climate sensitivity via a nonparametric fluctuation-dissipation theorem. *Journal of the Atmospheric Sciences*

Fuchs, D., Sherwood, S. and Hernandez, D., 2015. An exploration of multivariate fluctuation dissipation operators and their response to sea surface temperature perturbations. *Journal of the Atmospheric Sciences*

Hassanzadeh, P. and Kuang, Z., 2016. The linear response function of an idealized atmosphere. Part 1: Construction using Green's functions and applications.

We thank the reviewer for pointing this out. We have now included additional discussion in the final section that puts this method in the broader context of linear methods to understand climate model behavior.

Minor comments:

Simulation details: it helps if some details about the model setup (e.g. horizontal and vertical resolutions in the atmosphere and ocean etc.) are included following line 178.

Agreed. Added.

End of line 264: “)” missing Line 315: Figure's label missing Fig 1a: dashed lines are visible on the pdf but do not show when printed

Thanks for the careful reading. We'll alert the production staff of the journal.

Line 205-215: z^A_i is uniform across each of the 22 regions (right?). But that means there would be discontinuities in the heat source profile at the boundaries of these regions. That might cause problems if the model uses spectral methods.

Yes, z_i^A are uniform across the 22 regions. We thank the reviewer for alerting us to the possible concern regarding spectral methods. We have added a mention of this to the manuscript.

We note that the reviewer raises an interesting, broader issue in climate model simulations – for example, consider the land-ocean discontinuities introduced by conducting fixed sea surface temperature simulations. Our familiarity with how climate model numerics behave in this case is insufficient for us to comment on it in the manuscript, but we are interested in exploring how this issue has been dealt with

in spectral models in the past, which could provide some insight into how our results may have been affected by discontinuities.

Is there a statistically significant difference between the climatology (eg. time mean latent heat flux or zonal wind field) of the pre-industrial control simulation and the climatology of the system identification (perturbed) simulation?

This is a good point. Please see the response to Reviewer #1 for a further discussion of statistical significance and how we have modified the manuscript to address this issue. We have also added a figure comparing inter-annual standard deviations of the climatology and perturbed simulations using each year as an independent degree of freedom (although we of course acknowledge that consecutive years may not actually be independent).

I am very curious to see how this interesting technique performs for a range of problems and especially for regions that excite teleconnection patterns, as the authors discussed in section 4. To evaluate the technique, it might be a good idea (for future) to use a simpler climate model (e.g. an atmospheric GCM with coarse grid) which is cheaper to run and can be used to explore more examples.

Great idea! We'll keep that in mind when we write our next paper.