Dear Editor

We have addressed the reviewers' comments in the revised manuscript; our detailed responses to their comments are given below (reviewer's comments in italics).

Reviewer 1:

Overall Comments:

…The niche filled by the IGCM must be fully and clearly spelled out, starting in the abstract.

We thank the reviewer for this detailed review. The reviewer is correct and we now spell out in far more detail the existing and proposed user base for the IGCM4 in the 2nd and 3rd paragraphs of the revised introduction (lines 48-70 of revised manuscript):

'The rationale for such a model in the hierarchy of potential model codes is now addressed. Understanding key scientific questions related to climate and climate changes relies on understanding processes within the atmosphere, whose complex and nonlinear nature entails the use of global circulation models. However, understanding such complex processes in models is extremely challenging since unpicking processes within state-of-the-art climate circulation models can be extremely difficult given their complexity- especially when their computational demands are taken into account, leading to limits in both integration times and data storage.'

'Having said that, it is necessary for models to be complex enough to simulate the processes that are relevant to understanding a given question of interest. This is the niche which intermediate circulation models such as the IGCM occupies. This niche consists of models that are complex enough in terms of dynamical processes to represent a wide variety of processes from monsoonal circulations to extratropical storm tracks. However, their relative simplicity compared to state-of-the-art climate models that are employed by the Intergovernmental Panel on Climate Change (henceforth IPCC), enable process-level understanding to become more tractable because of (a) computational speed enabling long integrations or large ensemble members, and (b) flexibility and ease of use enabling the examination of idealised scenarios. Examples where the IGCM4 might be used are e.g.; conducting integrations of idealised perturbations to boundary conditions such as sea-surface temperature, topography, or continental distributions; conducting ensembles of multi-century integrations to collect robust statistics of small-amplitude responses to particular forcings".'

What is the goal of this paper?

1) Highlighting improvements?

The end of the abstract and lines 49-50 in the Introduction suggest showing the improvements made with respect to version IGCM3 are a goal. Improvements are thereafter not mentioned again and are never shown graphically. If this is truly a goal, "before and after" plots and a detailed discussion of improvements seen in the modeled basic-state climate would be needed. I recommend you simply drop "improvements" in favour of a solid description and validation of the IGCM4 as it is. The description could be accompanied by a table or bullet-point list that details the technical changes from IGCM3 to IGCM4 for those familiar with IGCM3.

We feel that it is necessary to describe the changes since the last published versions of the model to aid traceability from them. This model is not an entirely new code but the result of a series of ongoing improvements to the IGCM3 and IGCM3.1 configurations. Therefore describing such changes are necessary. We strongly feel that repeating the model description of Forster et al (2000) would be "re-inventing the wheel" given the similarities between the IGCM4 and IGCM3, and the fact that many numerical processes within the two models are identical.

Specific comments:

Introduction:

Provide a thorough motivation for a model like the IGCM within a more detailed explanation of the hierarchy- of-models concept. What are the strengths and limitations of a simpler model? What does it contribute that state-of-the-art models cannot and who would benefit from such a model?

We have now done so in detail in the introduction in lines 48-70 (also see response to above point).

Ending the Introduction with the customary reader-orientation statement outlining which topics are covered in each section would be helpful.

We have now included such a statement.

Section 2:

This should be a complete, thorough description of the model configuration and ALL its physics. Simply referring readers to the Appendix from Forster at al. (2000) leaves too many gaps. Include all relevant information and highlight updates as appropriate.

We think that complete documentation is not feasible within the scope of this journal, which is why we have concentrated on describing changes since the last IGCM version: our model description, in addition to the material it cites, forms a complete description. No model description paper of any complex model can be completely independent. If this was the case, as well as repeating the description contained in Forster et al (2000) this paper should also have to give details of the advection and filtering schemes detailed in Hoskins and Simmons (1975) for example. It is certainly true that where necessary information is not included, it should be cited- and we do so. Forster et al (2000) contains a comprehensive well-written description of IGCM physics, and we contend that repeating it is not necessary.

We note that similar limitations are also inherent in other papers. For instance, see the description of the UK Met Office Hadley Centre's HadGEM1 model paper (Johns et al J. Climate 2006; http://dx.doi.org/10.1175/JCLI3712.1). This paper describes model performance, and also describes how HadGEM1 is different to its predecessor HadCM3. However, it does not repeat the HadCM3 description because that is already contained within the peer-reviewed literature. The same principle holds true for the present paper.

2.1 If three "standard configurations" are named, results from all three should be shown. It seems to me from all your graphs that T42 is the standard horizontal resolution. It can be mentioned that T170L20 is an additional possibility which has been tested.

This is a good point: we now describe the T170 as an additional possibility in the text in section 2, leaving T42L20 and T42L35 as the two standard configurations.

2.2 This does not warrant a separate section; parallelisation is standard in GCMs.

We have now merged sections 2.1 and 2.2

The pet name of the UEA's cluster is of rather less interest than the machine architecture would be! Which compiler do you use, what is your timestep, and, consequently, what is performance speed not only on multiple processors (which you give), but also on a single processor for comparison? Are production benchmarks available upon request or on a website? Of the three websites given in the text, two are the same and none relates explicitly to the IGCM4. Listing reference websites should come at the end, not replace proper descriptions.

We have now rewritten this section giving details of compiler, timestep, and merged and moved the website reference to the reference list.

2.3 Rewrite as a complete description of the IGCM4 model as is. It is worth making a sub-section for surface processes (including the new snow and sea ice) and a separate one for convection.

We now include a description of the convection scheme in section 2.3 since this has not been described before in one paper. Incremental changes to the snow scheme from Forster et al (2000) are described in section 2.2.

2.4 Name all radiating species. In keeping with my "sell the model" message, this is an opportunity to point out how easy it is to modify the concentration of any species via namelist in the rundeck, for anyone envisaging climate sensitivity experiments.

We have now added a list of the radiatively active species in section 2.3, and also state that they are easily modified in the model namelist.

Discussion of aerosol treatment (or of the compensation for its absence) belongs in this section, not in the Conclusions (lines 339–347).

We have moved the treatment of the aerosol to the end of section 2.3.

2.5 This sub-section describes what is in my view one of the biggest improvements to the model. It would be desirable to find the same level of care and detail in the other model description paragraphs.

We have carefully described the stratosphere of the model since this is by far the most significant structural model development since IGCM4, and hence required more benchmarking.

Section 3:

This section is the most disappointing. Rather than illustrating model strengths by presenting the sort of robust statistical analyses enabled by very long runs, or by a large ensemble of runs, you present a few graphs from 50 years of the most basic model setting. It is unclear to me why you show a T42L20 model at all, when only a 50-year integration is presented. Furthermore, once the surface fields are validated with the T42L20 version, you switch to T42L35 for the zonal-mean fields, before finally adding on the q-flux ocean for a brief mention in the Conclusions. I agree that results from a coupled full ocean warrant their own paper, but given how important even a slab ocean is for achieving correct stratospheric vortex driving (see Winter and Bourqui, 2011), and how comparatively inexpensive it is to run, such an integration would constitute a valuable part of the Results section.

I recommend that each of the fields you wish to validate be shown for each of three model versions: T42L20, T42L35, and T42L35 coupled to the slab ocean. Do at least 100 years (as you have done already in the T42L35 setting), with the running times for each one given in the Configuration subsection. The depth of the slab ocean and length of spin-up time required should also be stated. The variability of each configuration could be illustrated with graphs of the variances of selected fields.

We now describe results from 200-year long integrations of the T42L35 configuration of the IGCM4 and 100-year long integrations of the T42L20 configuration.

We have not performed simulations of a slab model for this paper because although one effect of a slab ocean is to change the characteristics of model interannual variability (as shown by Winter and Bourqui 2011), the nature of such changes will depend on the depth of the slab, and how this depth changes seasonally and geographically: for instance in the North Atlantic Ocean the effective mixed layer depth changes from 50 m during summer to 500 m in winter. Moreover, the dynamic influence of the atmosphere on the ocean will also depend on the effective mixed layer depth of the ocean, or depth of the slab, as shown by O' Callaghan et al, Geophys. Res. Lett. (2014), DOI: 10.1002/2014GL062179 (henceforth OC14).

Because interannual variability is sensitive to slab ocean depth, and the IGCM has a constant slab depth, rather than one that varies seasonally and geographically, we have not discussed interannual variability in this paper. However, such a topic would be a source of useful research in the future for a configuration of the IGCM that had such a varying slab ocean model.

We have inserted the two paragraphs above in the results section to explain our rationale for not describing slab-ocean integrations in this paper (lines 381-394).

Use contours of the "truth" fields overlaid on shaded model fields; it allows much easier comparison than seeing the plots side-by-side does. I expect that the surface temperature will not change between the three IGCM4 versions, but perhaps precipitation will: this will be interesting to analyse. In the zonal mean, it will be interesting to see the difference the lid height makes to the wind field. What role does the GWD play here? What difference does the slab ocean make between versions having the same lid height? I am aware that the manuscript is categorized as a "model description paper" to be published in a model development journal, but readers and future users will find value *in seeing at a glance which version best lends itself to whatever climate study they intend to perform.*

We have now produced plots comparing both L20 and L35 models against reanalyses (Figures 7 and 8) that show model contours with model bias shaded contours. We believe this is the best way to show the model biases in the context of the simulated climatologies without making the plots too hard to read because of their complexity.

The level of detail given to the annual count of stratospheric sudden warmings is inconsistent with the very sparse treatment of everything else, and is not particularly well backed up by graphs.

We have given more detail to those aspects of the model which are new, i.e. SSWs.

Section 4:

New concepts such as climate sensitivity should not be introduced in the Conclusions. This can be addressed in Section 3, with a short definition of the term and references to Joshi et al. (2003). Similarly, the treatment of aerosols should be addressed in the model description section.

We have renamed this section “Climate Change and Energy Balance” since this best describes this section. We have moved the description of aerosols to section 2.3.

Bibliography:

Check that order is alphabetical throughout. You cite three papers from the 1990s using early versions of the IGCM to study Mars and zero papers done in the last five years using late versions of the IGCM coupled to a slab ocean to better study the Earth system!

We have reordered the references, and inserted references to Winter & Bourqui, J. Climate, 24, 5397–5415 (2011) and Winter & Bourqui, Geophys. Res. Lett., 38, (2011) doi:10.1029/2011GL047011.

Technical comments:

Line 86: When the truncation wavenumber is 42, then 2.8 degrees represents the grid spacing of the transform grid. It is not the resolution of the model; the resolution is half of one of those 42 smallest waves, or about 4.3 degrees.

We have replaced the text with “(having a 128 x 64 horizontal grid)” to make it clearer.

Lines 183-188: What is the reference for this new GWD scheme? Is it still Lindzen (1981)?

We have added the text “The IGCM4 scheme is as above,…” to clarify the description.

*Line 190: **3** Å~ **10−**6 or 3 ppmv but not both!*

We have corrected this error and simply state the value of 3 ppmv.

Lines 213-215: Reference needed.

We have added the reference to Eyring et al (2010).

Line 219: Reference needed.

We have added the reference to the CH_4 scenario (Holmes et al 2013).

Line 222: “Results” is perhaps not the best term, since no experiment is being performed. “Model validation” would be more accurate.

We have changed the section to “Model Evaluation” (evaluation being a less judgemental term than validation)

Lines 225-226: Explicitly state which climatologies are used.

We now state the source of the climatology.

Lines 240-241: This is inconsistent with the figure caption and the text printed on the figure itself. See comments under Figures.

We have changed the wording.

Figures:

All figures: more complete captions. Data sources should be explicitly named and there's no harm in repeating the data sources from one figure to the next. Which years from the reanalysis datasets are used, and why?

We have now added data sources to the figures, and state explicitly the years of the reanalysis datasets. These years are used to provide a climatology that is long enough to even out climate variability.

1. Not sure a figure is warranted here, especially when it gives the impression that L20 levels are indexed 16:35. Perhaps a table with exact pressure values as well as approximate heights in km?

We have changed the graph so that both L20 and L35 model layers start at "layer 1". We have kept the figure as we feel that a figure gives the best visual representation of where the model resolution is better than a table.

2. Colourbars! Or much larger and bolder labels on the curves

We have added bolder and larger labels on the curves to Figure 2.

3. The "observations" are in fact reanalysis data; give complete sources in the figure captions. See my comments above for combining reanalysis contours with IGCM shading, and for showing all three of the IGCM4 settings discussed in the manuscript.

We have changed both figure and figure caption. We have used one IGCM configuration out of two in order for ease of comparison here. We have only overlayed reanalysis and IGCM for the zonally averaged plots (Figures 7 and 8) since we feel to do so for Figures 3-6, which already have land-sea boundaries marked on them, would make them too "busy" and hard to read: instead we show difference maps for greatest clarity.

The bias over northern Africa in both seasons, Australia in DJF, and high-latitude North America in JJA is widely greater than 10 K, in contradiction with the text. Instead of shrugging this off as "reasonable compared to CMIP5", it would be more helpful to have a more careful analysis of possible reasons. How much is this related to the simplicity of the soil scheme and the ensuing surface humidity errors? What does this bias do to convection?

We now give more specific reasons for the discrepancy during DJF and JJA in North Africa and DJF in Australia in the text. The absence of dust will raise the surface temperature in these areas: we comment on this in section 3.1 of the revised manuscript (lines 285-287).

We now note that both ice caps display too large a seasonal cycle, which we attribute to the simplicity of the snow scheme in the model, which has no facility for changing density or conductivity when snow is compacted into ice. This could be a source for future model improvement.

4 and 5. These are confusing, in part because of inconsistencies between the caption and the text, especially the top right panel, described as "difference between IGCM4 and observations" in the caption but not in the text or in the small title under the panel itself. The centre and right bottom panels are overkill – why not just show the variance of the model precipitation? The top left panel contains the colour green and the colourbar at the bottom does not – are all panels really on the same scale?

We have changed both figures and captions. The bottom row of subfigures are useful to show the spread within the CMIP5 ensemble. Just showing the variance, but not directly showing the bias of the multi model mean +/- one standard deviation (subfigures d,f), makes it more difficult to compare the bias of the IGCM4 (subfigure c) in the context of other CMIP5 GCMs.

6. Provide a complete caption.

We have provided a new caption for this figure.

7 and 8. This is minor, but Figure 2 has North on the right at the bottom of every panel, and these Figures have North on the left on the bottom panels only. As a kindness to the reader, you could draw attention to this reversal. (My first impression was that the figures had accidentally been mislabelled.) Otherwise these in particular should be shown for all settings of the IGCM4 that you discuss, with ERA-40 contours overlaid. I am interested to see whether the boreal winter tropospheric jet axis shifts slightly in the meridional between the high-top and the low-top versions.

We have reversed the axes to be consistent with Figure 2. We now describe both L20 and L35 configurations of the IGCM separately since their vertical extent is where they differ the most. We also overlay contours of IGCM zonal wind and temperature on colours of IGCM4-reanalysis difference here since for these plots it makes to show where the biases are (e.g. in the southern hemisphere tropospheric jetstream).

9. The caption says this is for DJF and the graph title says DJFM. Which is it? It should be the same months as the zonal-mean wind and temperature plots. The text says the "zonally asymmetric component of the circulation is apparent" (line 295); it would be more apparent if you indeed plotted only the zonal anomaly of the geopotential height. If you were to show this ***Z0*** *field at the 200 hPa level as well, it could nicely tie in with your later discussion of the stratospheric vortex variability, since the wave-1 and wave-2 stationary patterns that mainly force the stratospheric circulation would stand out well.*

We now describe both L20 and L35 configurations of the IGCM at both 500 hPa and 200 hPa. The figure now displays eddy fields to show the wave patterns that force the stratospheric circulation. The captions have been fixed to display "DJF".

10. Here it will be very interesting to see the differences between versions of the IGCM4. I'm not sure what is contributed by differentiating between splitting and displacement events, since the manuscript does not go into the mechanisms required for either or address why the model "misses" when it does.

Splitting and displacements events have different effects on the surface and ocean and are described elsewhere (OC14), which is they their differences are not detailed here.

11. A more complete caption here; presumably these are annual, global averages?

We now state that in the caption that this is "Annually averaged net downward zonal surface energy imbalance".

References:

Bracegirdle et al. (2013), JGR-Atmospheres, vol. 118, pp. 547–562

Forster at al. (2000), Clim. Dyn., vol. 16, pp. 833–849

Joshi et al. (2003), Clim. Dyn., vol. 20, pp. 843–854

Winter and Bourqui (2011), Geophys. Res. Lett., vol. 38, 10.1029/2011GL047011