Dear Referee #1,

We would like to thank you for your comments and suggestions to improve the quality of our manuscript "The Beijing Climate Center Climate System Model (BCC-CSM): Main Progress from CMIP5 to CMIP6" by Tongwen Wu et al. Following your suggestions, we added 5 further figures in the revised manuscript. So, all figures in the first manuscript are renumbered. We have also changed the order of presentation for some subsections in Section 4 (Results). The point-to-point responses to your comments are enclosed in the following.

Best regards,

Tongwen Wu and all co-authors

==

Response to Anonymous Referee #1

General comments:

In this article the authors describe the changes from the BCC-CSM1.1 model that was used in CMIP5 to the new BCC-CSM2 model that is now employed in CMIP6. They compare the "historical" simulations as specified by CMIP5 and CMIP6, respectively, of the old and new model. They show that the new model can simulate the mean climate as well as some modes of variability with some skill, and they point out a number of improvements in the new simulation compared to the old one. Obviously this article is meant to be a basic, citable documentation for the new CMIP6 simulations of the BCC-CSM2 model.

The documentation and analysis of the simulations is superficial. No attempt is taken to describe or analyze the effects of the single model improvements, or to attribute the identified differences between the two simulations to the introduced model changes. Therefore there is not much that the reader can learn from this article, except that the coupled model can reproduce the transient climate of the CMIP historical experiment with some skill. The authors point at future publications ("to be submitted") for other experiments, model resolutions or the QBO in the historical simulations.

Response:

Thanks for the relevant comments. Yes, the purpose of this manuscript is to document the transition of our model BCC-CSM from CMIP5 to CMIP6. We hope that it can be a reference for the different experiments of CMIP6 which are progressively available (or very soon for some specific runs) for the scientific community of CMIP6. So we compare here only the general performance between our old and new models, particular performance being investigated in detail in other specific papers. Nevertheless, we think that it is useful to add some more materials (including 5 new illustrations), which goes to the general sense of both

Reviewers.

L40: …Many climate models in the world have been developed since the IPCC-AR4 …The IPCC reports certainly motivated many groups to contribute climate projections for the assessment of future climates, for which these groups developed suitable models. But I do not see why IPCC-AR4 is pointed out as a special landmark along the path of the development of global climate models or Earth system models. Rather the initial coupled model inter-comparison project (CMIP) would deserve to be mentioned, and the way the CMIPs were developed by the community and the working group on coupled modeling (WGCM).

Response: We agree that CMIP and its promoter WGCM played an important role for climate modelling in the world. They should be more acknowledged than IPCC. The latter is however much more popular for a general public.

L71: … Its performance is presented in a separated paper (Wu et al., to be submitted). … Please check if references to "to be submitted" articles are allowed in GMDD. Better write "… Its performance will be presented in a separate paper. ...".

Response: Proposition relevant and accepted.

L78: Please include a figure that compares the profiles of layer thickness against the height of the layer for the L26 and L46 grids, using for example a simple log-p height definition. This would clarify how the vertical resolution has been improved.

Response: We added Figure 1 in the revised manuscript to show the profiles of layer thickness against the height for 26 vertical layers in BCC-CSM-1.1m and 46 vertical layers in BCC-CSM2-MR.

L81: 2.1 Atmospheric component BCC-AGCM. This is the main documentation for the new version of the atmospheric model. Therefore I would expect to find here basic information for all processes. This should include the numerical techniques used in the dynamics, the transport scheme and the physics. This basic information should be kept concise, so that most of the room can still be devoted to the subsections (a) to (d) for the changes compared to the preceding documented version.

Response: For the completeness of the documentation, we added more detailed descriptions on the model dynamics core, and all the models physics are summarized in Table 2.

L96: … considered …→… considers …
L100: … environment, The mass …→… environment, the mass…
L128:… at each model grid …→ …in each model grid cell …?
L138: …T_con → T_conv …

Response: Corresponding modifications are now included.

L144: "… Following the method above, the cloud fraction (C_deep and C_shallow),temperature (T_deep and T_shallow), specific humidity (q_deep and q_shallow) for the deep convective, shallow convective clouds can be then deduced sequentially…." Does the scheme allow the concurrent occurrence of shallow and deep convection in the same atmospheric column, or is only one type of convection allowed at any one time in a single column? If concurrent occurrence is allowed, how is the parametrized computation of deep and shallow convection split, and how is the necessarily sequential diagnostics of C, T and q in the shallow and deep updrafts organized?

Response: We have rewritten this part of the manuscript. In fact, shallow and deep convections can concurrently occur in the same atmospheric column at any time step. The shallow convection follows the deep convection, and it occurs at vertical layers where local instability still remains after the deep convection. In BCC models, the three moisture processes (i.e. deep convection, then shallow convection, and finally stratiform precipitation) is sequentially executed. That means, the model-box mean T and q are updated immediately after each process.

L151 and L153: These equations can be numerically unstable in the limit of C_deep+C_shallow!1, because of the division by (1-Cdeep-Cshallow) that is needed to obtain the unknowns T_ambient and q_ambient. Is this a practical problem, or rendered irrelevant by the multiplication factor (1-C_conv) in Eq. 2?

Response: We have added "if $\mathrm{C_{deep}} + \mathrm{C_{shallow}} > 1$, $\mathrm{C_{deep}}$ and $\mathrm{C_{shallow}}$ are then scaled to meet the condition $\mathrm{C_{deep}} + \mathrm{C_{shallow}} = 1.0$, and Cs=0." So we do not need to know the values of T_ambient and q_ambient to derive Cs using Eq. (10).

L157: : : : RH_abmient : : : ! : : : RH_ambient : : :
L164: I cannot find the publication by 'Kristj ÌA¸ansson and Kristiansen [2000]'. Do you mean 'Kristjánsson et al., 2000'? Kristjánsson, J. E., J. M. Edwards, and D. L. Mitchell (2000), Impact of a new scheme for optical properties of ice crystals on climates of two GCMs, J. Geophys. Res., 105(D8), 10063–10079, doi:10.1029/2000JD900015.
L202: : : : k = 1.18 _ 106 cmˆ □1 secˆ □1 : : : ! : : : k = 1.18 _ 10ˆ6 cmˆ □1 secˆ □1 : : :

Response: They are corrected now.

L206: Section 'd. Parameterization of gravity wave drag' This paragraph discusses drag by dissipating gravity waves originating from flow over orography or atmospheric sources. What about drag by blocking effects from unresolved orography? Are such effects, which sometimes are included in gravity wave parameterizations, considered in BCC-AGCM3-MR?

Response: The scheme of gravity wave drag generated from convective sources is that in Beres et al. (2004). The drag by blocking effects is still not involved in present version of BCC-AGCM3-MR.

L221:... This parameterization scheme of convective gravity waves can improve the model's ability to simulate the stratospheric quasi-biennial oscillation in BCC-AGCM3-MR. … This is a rather general statement. It is clear that non-orographic gravity waves make a significant contribution to the forcing of the QBO, and if the gravity waves are not resolved, then their effect needs to be parameterized for the simulation of the QBO, and tunable parameters can be used to improve the structure of the QBO. Is the CF parameter, which you tune, valid for all latitudes or only for equatorial latitudes, where the QBO exists?

Response: The convective fraction (CF) within a grid cell is an important parameter and is closely related to the deep convection process. It is tuned to obtain the right wave amplitudes. In BCC-AGCM3-MR, it is a constant and does not vary geographically. So it is valid for all latitudes where convection occurs.

L277: Is the "simple scheme about the surface albedo, roughness length, turbulent sensible and latent heat fluxes over rice paddies" documented, or is there a manuscript in preparation? If not, and if the scheme is indeed simple, you should include the documentation here.

Response: A new manuscript for this regard is indeed under preparation by our team in charge of land surface processes.

L348: …The preindustrial climate state of BCC-CSM2-MR is preceded by a more than 500 years piControl simulation following the requirement of CMIP6. …Which were the goals of the spin-up simulation for the piControl experiment? Which were the criterions for declaring the spin-up phase completed? It would be interesting to learn about these criterions.

Response: In the revised manuscript, we added some details about how the spinup was effectively accomplished. For this issue, we just followed the recommendations from CMIP project. Basically we check the steadiness of some globally-integrated quantities. We also added Fig.2 in the revised manuscript, showing energy balance at top of the atmosphere.and surface air temperature over the globe in the piControl simulation.

L.353: : : : the up-limit of the atmosphere : : : ! : : : the top of the model atmosphere : : :
Response: Modified.

L.362: …It means that the whole earth system in our models is very close to energy equilibrium. … For a transient period it is a bit difficult to judge from the similarity between TOA radiation fluxes of the model and the observations whether the model is

generally in a "good" equilibrium. How is the energy balance for the preindustrial control experiment? Here we know that the net energy flux at TOA should be zero except for fluctuations related to the internal variability of the coupled system. If the model system has energy leaks, as many climate models have, we should see this clearly in the stabilized piControl simulation. Such a leakage would have to be considered in the comparison of the transient TOA radiation fluxes of the model and the observations. My suggestion is that you discuss the net energy flux at TOA and the surface and maybe other quantities of interest of the piControl simulation before starting the discussion of the historical simulation. This could for instance be embedded in a new section that explains the tuning goals of the piControl experiment.

Response: Yes, we agree with the referee's reasoning. It is a bit difficult to judge from the similarity between TOA radiation fluxes of the model and the observations. We have modified the description. In addition, following referee's suggestion, we have added Figure 2 (in the revised manuscript) in Section 3 to show the time series of annual mean of net energy flux at top of the atmosphere and the global sea surface temperature from 600 years piControl simulation. It means that the whole system in BCC-CSM2-MR nearly reaches its equilibrium after 600 years piControl simulation.

L372: : : : These biases are reduced in BCC-CSM2-MR. : : : Which of the model changes discussed earlier cause the strong (and useful) changes in the tropical SW and LW cloud radiative forcing between the old and new model?

Response: Modified. In low latitudes between 30 °S and 30 °N, BCC-CSM1.1m shows excessive cloud radiative forcing for both shortwave and longwave radiations. These biases are reduced in BCC-CSM2-MR and may be attributed to the new algorithm in diagnosing cloud fraction especially convective cloud amount.

L379: : : : Biases of annual mean surface air temperature (at 2 meters) : : : Figure 2 shows the spatial patterns of the T2m bias. But first of all I am wondering how the transient global mean temperature is evolving from the stabilized pre-industrial mean temperature representative for 1850 to the present day. Please discuss first the global mean evolution before describing the pattern of the T2m bias near the end of the historical experiment. Further, it would be interesting to read your opinion on the contribution of the model changes to the observed differences between the models. Can you attribute the disappearance of the cold bias in the southern oceans in the new model compared to the old model to the changes in the model for turbulent fluxes over sea ice? Are specific changes of the land model important for the increased cold bias in east Asia and Siberia in the new model compared to the old model? For the land surface biases it would be valuable to know if these patterns are already present in AMIP simulations, where effects of oceanic biases are excluded.

Response:

(1) Following referee's suggestion, we advanced two sections (old Section 4.4 and Section

4.5) to the position just after Section 4.1. They become now Section 4.2 and Section 4.3. The former presents the transient global mean temperature evolving from 1850 to the present day, and the latter (new Section 4.3) presents climate sensitivity to CO2 increasing.

(2) As shown in Figure 2 (in the initial manuscript, and renumbered to Fig.5 in the revised manuscript), the disappearance of the cold bias in the Southern Oceans in BCC-CSM2-MR (compared to BCC-CSM1.1m) is quite certainly attributable to the new scheme of cloud fraction parameterization implemented in BCC-CSM2-MR, which largely improves the low-level clouds simulation between 40 °S to 60 °S over the Southern Ocean (not shown).

(3) Biases over land surface in the two coupled models are similar to each other. They are furthermore already present in AMIP. So biases over land surface (at least partly) come from the land surface model.

L386: annual mean precipitation What is the global mean precipitation in both models? Do you have any thoughts about the contribution of the changes in the deep convection scheme to the strong wet bias in the Maritime Continent?

Response: The global mean precipitation rates in BCC-CSM1.1m and BCC-CSM2-MR are 2.87 mm/day and 2.94 mm/day, respectively. A precipitation rate of 2.68 mm/day is the 1986-2005 mean observed precipitation from Global Precipitation Climatology Project (Adler et al., 2003). The excessive rainfalls in the Maritime Continent seem amplified in BCC-CSM2-MR, which is attributed to abundant stratiform precipitation which accounts 39% of total precipitation in BCC-CSM2-MR. That percentage was 35% in BCC-CSM1.1m.

L. 393: … The evaluation is done against climatology of ERA-Interim … NCEP dataset … From the text and the figure caption it seems rather that only NCEP is used and not ERA-Interim. Can you please clarify this?

Response： It was a mistake. In fact, ERA-Interim is used as a reference state instead of NCEP.

L421:…Given a much higher vertical resolution and an advanced parameterization of the gravity wave drag … Despite the more complete parameterization of gravity wave drag – now including gravity waves from atmospheric sources – the zonal mean zonal wind biases in the high latitudes of the stratosphere have increased near 10 hPa, where one would expect the main benefit from gravity wave parameterizations. Can you explain why there is no benefit from the improved GWD parameterization on the structure of the polar night jets? Did you attempt to tune the gravity wave drag to reduce errors in the stratospheric extratropical zonal circulation?

Response: We don’t know the exact cause for the zonal wind biases in the stratosphere of high latitudes in BCC-CSM2-MR. Maybe the lack of gravity wave drag generated by atmospheric blocking is the explanation. We expect to reduce the bias in next version by adding this process.

L427: … In Figure 6(b), the BCC-CSM2-MR simulations present a clear quasi-biennial oscillation of the zonal winds as observed. …The downward propagation … does not penetrate to sufficiently low altitudes. … Though the vertical resolution is increased, it is still too low to expect a QBO simulation down to at least 70 hPa, because the forcing from resolved waves cannot be adequately resolved. Therefore it seems like the QBO occurring in the new model must be dependent entirely or nearly entirely on the Beres parameterization. Has this scheme been tuned to obtain the QBO in the new model?

Response: Yes, we agree with the referee's arguments. The number of layers (46) in our model is certainly too low to resolve adequately vertically-propagated waves. But we have not tuned the Beres scheme that seems to work well in our new model. This issue deserves further studies in the future.

L470: : : : The most remarkable improvements of BCC-CSM2-MR appear in the boreal warm seasons, : : : To which model improvement do you attribute the strong improvement of SIE or SIC?

Response: It is hard to say. Those improvements may be partly contributed by many aspects of new model physics schemes such as turbulent flux over sea ice and ocean surfaces, cloud fraction, or atmospheric circulation improvement at high latitudes.

L495: : : : Our CMIP6 model can capture this warming hiatus. : : : The word "capture" suggests that the hiatus is a predictable climate feature that a coupled climate simulation can be expected to reproduce if the forcing is realistic and the model is "correct". Is this what you want to express? Maybe it is better to write for instance: "The historical simulation of the CMIP6 model shows a hiatus towards the end of the simulation that resembles the observed one." Do you have other ensemble members for the historical CMIP6 simulation, and if so do all members reproduce the hiatus of 1998-2013? The figure shows also that the CMIP5 simulation is significantly colder in the early decades than observations or the CMIP6 simulation. Later on, however, both simulations evolve by and large in a similar way. Can this be explained by the external forcing (volcanic aerosols?)? Do you have any insight that you can share?

Response: To address this comment, we need to point out three elements. (1) The phrase concerning the simulated warming hiatus has been changed in the revised manuscript as suggested. (2) As many CMIP diagnostic papers, we also used only one member in our manuscript. That member shows hiatus. Recently, two additional members of the historical simulation are also available. As shown in new Figure 4, two members show hiatus towards the end of the simulation. (3) Both BCC models seem too sensitive to volcanic aerosols. But we do not fully understand the behaviors of global mean temperature curves.

L509: …Observation-based NSIDC data are also plotted when available. … The caption for Fig. 11 reads: ": : : observations-based Hadley Centre Sea Ice and Sea Surface Temperature data set (Rayner et al., 2003)." Please clarify.

Response: Modified. In Figure 11 (in the initial manuscript, and renumbered to Fig.14 in the revised manuscript), the observed sea-ice extent are derived from Hadley Centre Sea Ice and Sea Surface Temperature data set. In Figure 8 (now renumbered to Fig.15), the observed seasonal cycles of sea-ice extent in (a) and (b) are based on the National Snow and Ice Data Center (NSIDC; Fetterer et al., 2002) data sets. In order to keep consistency, the NSIDC data in Figure 8 (now renumbered to Fig.15) is replaced by Hadley data.

L516: 4.5 Climate sensitivity to CO2 increasing Figure 12 following Gregory (2012) not only provides estimates for the ECS, but primarily provides information on the climate feedback. Comparing both models, the ECS is similar, but the feedback parameter is substantially different: BCC-CSM3-MR: ca. -1 W/m2/K; BCC-CSM1.1m: ca. -1.3 W/m2/K Thus the result that both ECS values are very similar results only because the initial 4xCO2 forcing is quite different: BCC-CSM3-MR: ca. 6 W/m2; BCC-CSM1.1m: ca. 7.5 W/m2/K Can you please comment on the origin of the large difference in the initial forcing?

Response: Yes, that's an interesting point. We added a paragraph in the revised manuscript concerning the 4xCO2 initial forcing, feedbacks and ECS. Due to changes of atmospheric profiles (temperature, water vapor and cloud), it is possible to have different forcing for a quadruple CO2, and even the radiative transfer is unchanged. Feedbacks operate certainly in different ways in the two models. What is interesting is that the final ECS converges between the two models.

L529: : : : abruptCO2 : : : ! : : : abrupt4xCO2 : : : ?

Response: Modified.