Anonymous Referee #2

We thank the valuable comments, whose responses follow:

Major Revisions:

1.1 There is no description of the pre-industrial control run.

Reply:

Our group has prepared two manuscripts describing describing the results using piControl and abrupt 4 regarding the atmospheric component has been submitted to the GMD and is under review for GMDD (Capistrano et al., 2018). The manuscript describing the results regarding the oceanic component is in preparation and will be submitted soon (Nobre et al. 2018). We have included a summary description of the pre-industrial control run in the revised manuscript (Page 8, Lines L20-L23).

In order to advance the results regarding the analysis for the ocean component of the piControl run, we describe some of the analyses of Nobre et al. (2018) below.

Fig. 1 shows the variation of the global averaged (a) near-surface air temperature (SAT), (b) sea surface temperature (SST), (c) sea surface salinity (SSS), (d) net of the radiation at the top of atmosphere (TOA), (e) net of the ocean/atmosphere heat flux and (f) the maximum strength of the Atlantic Meridional Overturning Circulation (AMOC), over 1000 years of simulation. This period includes the 13 years of coupled ocean-atmosphere spin-up used by the Historical simulation and the remaining 156 years of simulation that the piControl runs in parallel with the Historical simulation. The red dash-dot lines show the linear trend of each time series. With the exception of the SSS, all the variables show a fast adjustment to stable conditions during the first 13 years of the piControl simulation, which is the period used as the spin-up period for Historical simulation. With the exception of the SSS, the piControl simulation maintains stable conditions. Over this stable period, the mean SAT is 13.286 ºC with a linear trend of 0.0091 ºC (100 yr)$^{-1}$ and the mean SST is 19.57 ºC with a linear trend of 0.0084 ºC. Although the SAT and SST trends are statistically significant at 95%

level, they are small and comparable to other state-of-the-art models (e.g. NorESM1-M presents a linear trend of 0.0074 ºC (100 yr)$^{-1}$ and 0.0062 ºC (100 yr)$^{-1}$ for SAT and SST, respectively, and NESMv3 presents a linear trend of 0.0021 ºC (100 yr)$^{-1}$ and 0.0073 ºC (100 yr)$^{-1}$ for SAT and SST, respectively), therefore indicate stable variation for both variables. The mean SSS is 34.07 PSU with a linear trend of -0.017 PSU (100 yr)$^{-1}$. The SSS linear trend is significantly higher than simulated by NorESM1-M that has a linear trend of -0.00006 PSU (100 yr)$^{-1}$ and NESMv3 that has a linear trend of -0.008 PSU (100 yr)$^{-1}$. However, it should be noticed that there is stabilization after 500 years of simulation, since the linear trend for the first 200 years is -0.09 PSU (100 yr)$^{-1}$ and for the last 500 is -0.004 PSU (100 yr)$^{-1}$. This SSS long-term adjustment can be related to the ocean drift.

The net radiation at TOA has a mean value of -4.33 W m^{-2} and a linear trend of -0.0090 W m^{-2} (100 yr)$^{-1}$. The linear trend is comparable to other state-of-the-art models (e.g. NorESM1-M presents a linear trend of -0.0038 W m^{-2} (100 yr)$^{-1}$ and NESMv3 presents a linear trend of -0.0041 W m^{-2} (100 yr)$^{-1}$). The negative average value means that the atmosphere is losing heat to the outer space and can be an explanation for the weak warming observed in the Historical simulation (Fig. 2 (manuscript)). Although the negative bias of the net radiation at TOA is significant, there is an important improvement from the previous version (BESM-OA2.3) that presents a negative net radiation at TOA of roughly -20 W m^{-2} (Marcus Bottino, personal communication, 2018). Nevertheless, the mean value of the net radiation at TOA is still away from zero (optimal value), when compared with NorESM1-M that presents an average value of 0.086 W m^{-2} and NESMv3 that presents an average value of 0.17 W m^{-2}. The global net of the ocean/atmosphere heat flux has a mean value of 0.94 W m^{-2}, which is slightly away from zero (optimal value) than NorESM1-M (0.122 W m^{-2}) and NESMv3 (0.31 W m^{-2}, for Earth surface). The net of the ocean/atmosphere heat flux presents a linear trend of -0.018 W m^{-2} (100 yr)$^{-1}$, which is an order of magnitude higher than NorESM1-M (-0.004 W m^{-2} (100 yr)$^{-1}$) and NESMv3 (-0.006 (100 yr)$^{-1}$, for Earth surface). The positive values indicate that the ocean is gaining heat from the atmosphere throughout the simulation, but since it has a negative trend, the ocean heat content is diminishing during the integration. The small trends (although statistically significant at 95%) of the net radiation at TOA and net of the ocean/atmosphere heat flux show that the model has achieved stable conditions for the heat flux after the initial adjustment. The AMOC has a maximum mean value of 13.38 Sv and a linear trend of -0.11 Sv (100 yr)$^{-1}$ at 28 ºN. However, it can be noticed that there is a sharp decrease of the AMOC in the period 80-170 years and then its strength recoveries. The average AMOC strength measured by the project RAPID at 26.5 ºN is 17.2 Sv (McCarthy et al., 2015), which is slightly higher than simulated by BESM-OA2.5 piControl. The AMOC strength at 26.5 ºN simulated by NorESM1-M and NESMv3 are ~31 Sv and 14.8 Sv respectively. The linear trend for NorESM1-M and NESMv3 are -0.12 Sv and -0.22 Sv respectively. The AMOC negative linear trend simulated by BESM-OA2.5 is very similar to the linear trends given by NorESM1-M and NESMv3. However, this negative linear trend indicates that the model is still drifting throughout the piControl simulation. Despite the surface quantities (except SSS) and heat flux in the air-sea interface and at TOA indicate stable conditions, the model still drifts in the ocean.

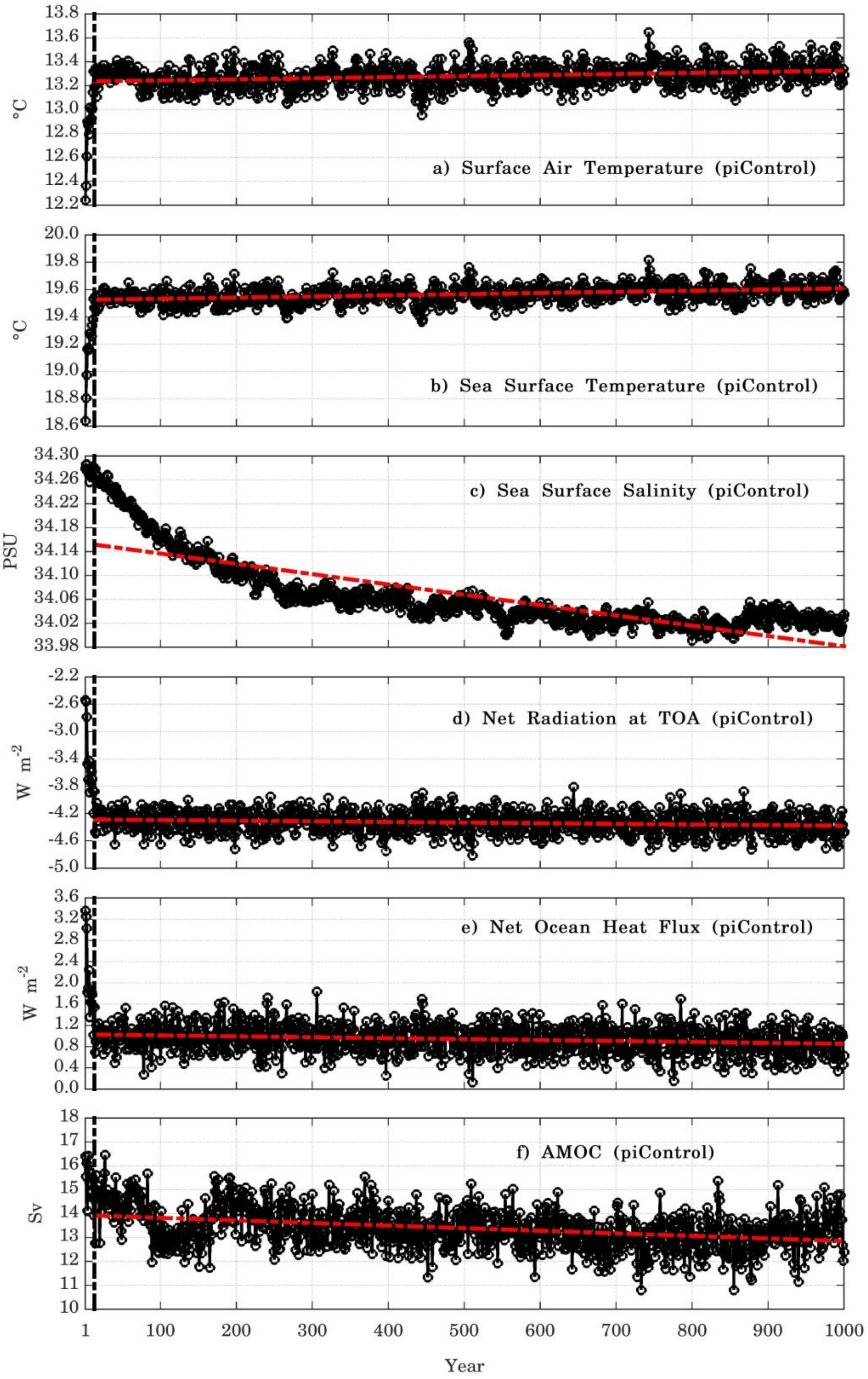

Figure 1 – Annual average time series for the global average near surface air temperature (a), sea surface temperature (b), sea surface salinity (c), net of the radiation at TOA (d; positive values indicates that the atmosphere is warming), net of the ocean/atmosphere heat flux (e; positive values indicates that the ocean is warming) and the maximum AMOC (f; at latitude 28º N and depth of 800 m), simulated by the piControl from the beginning of the simulation up to 1000 years. The red dash-dot lines show the linear trend of each time series, starting from the year 14th of the simulation (vertical black dash-dot lines). All the trends are statically significant at 95% of confidence level.

To further evaluate the piControl ocean drift, it is computed the depth-time Hovmöller diagrams of global mean ocean temperature and salinity departures from their respective initial conditions (Fig. 2). By initial conditions we mean the value of the first year after the model adjusts, in this case, the 14th year. The ocean warms in the sub-surface waters, between 150 and 350 m depth, and in deeper waters, from 1500 m up to the ocean floor. At 1000 years of simulation, the higher warming occurs below 4000 m depth, where the waters warm ~0.8-0.9 ºC relative to the initial values (Fig. 2a). There is a cooling of 0.2-0.3 ºC between 500-1500 m depth that starts to narrow throughout the simulation up to 600 years. From 600 years of simulation the whole global ocean warms with intermittent signals only on the surface. The ocean salinity slightly increases below 1000 m depth up to the ocean floor and at 1000 years of simulation the increase is ~0.06 PSU comparing with the initial values (Fig. 2b). Above 1000 m depth there is a significant freshening of the ocean waters, with the surface waters salinity decreasing up to 0.28 PSU in 1000 years of simulation when compared with initial values. Such results indicate that the global ocean is still drifting from its initial conditions.

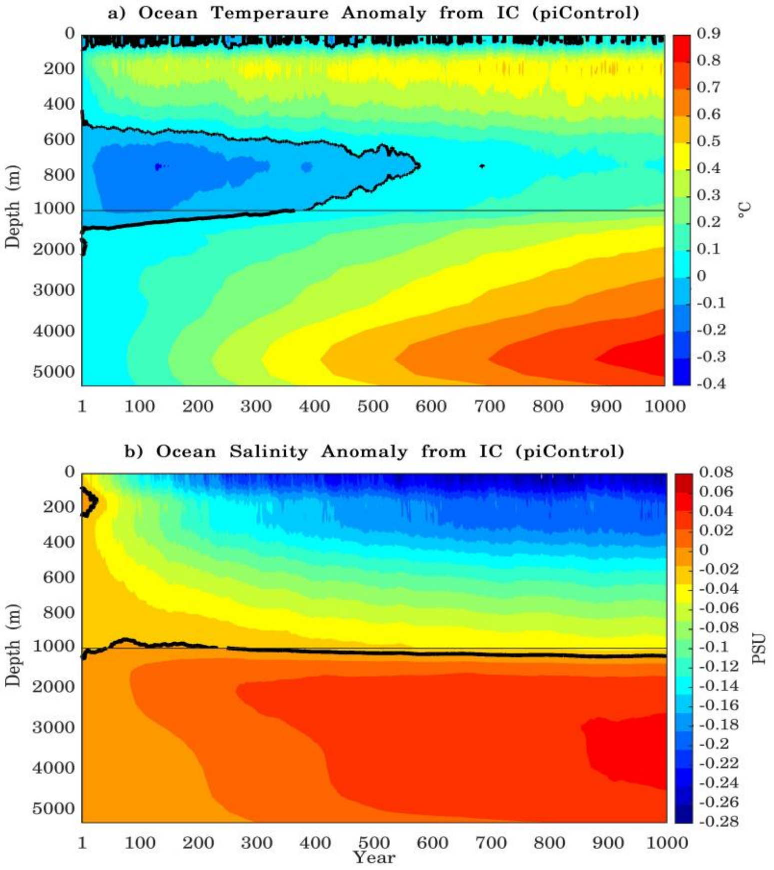

Figure 2 – Depth-time Hovmöller diagrams of global average ocean (a) temperature and (b) salinity anomalies from the respective initial conditions (IC). Here the initial conditions are taken from the 14th year, therefore after the 13 initial years of the adjustment. The diagrams are based on annual average time series by the piControl simulation over the period 1000 years. The thick black line represents the zero contours. Note that the vertical scales are different above and below 1000 m,

The linear trends are obtained through a linear regression of each respective time series and the statistical significance is tested through the Mann-Kendall test, considering statistically significant at 95% of confidence level. The number of degrees of freedom in the test is adjusted due to the autocorrelation of the time series.

1.2 In particular, the coupled model simulation presented here is done after an extremely short spin-up phase (just 13 years after one cycle of CORE forcing!). Experience tells us that models tend to drift after changing from stand-alone conditions to fully coupled. BESM could be an exception, but it could be that many features seen in this paper are due to incomplete adaptation.

Reply:

We agree with your observations, and the spin-up process is better described below. After the ocean stand-alone run forced with CORE fields (71 years), a spin-up run of the fully coupled model in a previous version is done for 100 years. The atmosphere and ocean state at the end of this 100 years long integration is used as the initial conditions for the piControl run. Moreover, land ice albedo and cloud microphysics were changed from the initial spin-up run (exp178) to the piControl simulation. The ocean model configuration remains identical on both versions. Therefore, regardless of the coupled model having being integrated for 100 years, in this case, it is assumed only as spin-up the first 13 years period of the piControl simulation.

We recognize that long spin-ups are desirable. In fact, the deep layers of the ocean are still adjusting after 1000 years. However, through previous tests, it was observed that the surface layers of the ocean model tends to adjust in the first 13 years for surface variables, as SAT, SST and net radiation at TOA. As discussed above (point 1.1), the piControl simulation shows a fast adjustment to stable conditions in the first 13 years of the simulation that can be inferred through the variation of the SAT, SST, the net radiation at TOA and the ocean/atmosphere heat flux. After this fast adjustment the model reaches stable conditions, given by the linear trend of each variable. For this reason, it is assumed that the model was prepared to initiate a coupled Historical simulation.

This information has been included in the revised manuscript (Page 8, Lines L17-L24).

1.3 It should be documented how the 3d ocean temperature and salinity fields evolve. How does the integrated ocean-atmosphere heat flux evolve? Is there any energy imbalance that could (partly) explain the weak warming in the 20th historical simulation?

Reply:

This is an important issue. Thank you for comment it. This topic has been included in the revised manuscript (Page 12, Lines L17-L23), as well as the Fig. 3 (Page 55). Similarly to the piControl simulation, the net radiation at TOA has a negative bias and net of the ocean/atmosphere heat flux has a positive bias (Fig. 3). The net radiation at TOA has a mean value of -4.20 W m^{-2} and the net ocean/atmosphere heat flux has a mean value of 1.16 W m^{-2} in the first 50 years. Throughout the simulation, the net radiation at TOA becomes less negative due to the increasing CO_2 on the atmosphere and consequential increasing atmospheric heat content. Part of this heat is transferred into the ocean as positive net of the ocean/atmosphere heat flux increasing indicates. Negative values of net radiation at TOA means that the atmosphere is losing heat to the outer space during the simulation, which is likely the reason for the weak warming observed in the Historical simulation (Fig. 2; manuscript).

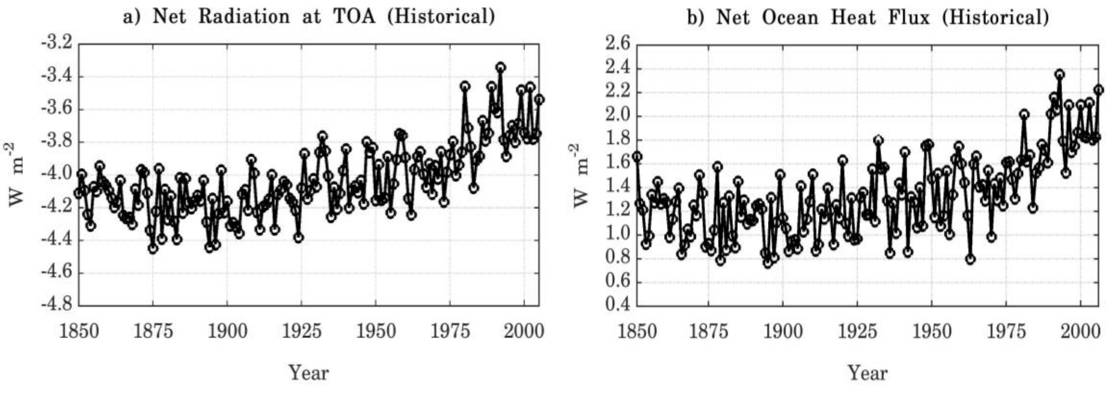

Figure 3 – Annual average time series for the global average net of the radiation at TOA (a; positive values indicates that the atmosphere is warming) and net of the ocean/atmosphere heat flux (b; positive values indicates that the ocean is warming), simulated by the Historical run over the period 1850-2005 (156 years).

The depth-time Hovmöller diagrams of global mean ocean temperature and salinity departures from their respective initial conditions simulated by the Historical run is shown in Fig. 4. Here initial conditional means the value of the first year of simulation, in this case, the year 1850. The prominent warming occurs from the surface up to 400 m depth (Fig. 4a). This warming is more significant at the end of the simulation (~0.6 ºC comparing with initial conditions) and is likely to be related to the global warming of the planet and consequential increasing heat flux from the atmosphere into the ocean. In deeper waters, from 1500 m up to the ocean floor, there is a weaker warming, indicating that the ocean is gaining heat mainly in the upper layers. Between 500-1500 m depth, it is observed a cooling tendency respective to initial conditions. The ocean salinity slightly increases below 1000 m depth and from 1935 the increase reaches 0.04 PSU between 1500 and 3000 m depth compared with the initial values (Fig. 4b). Above 1000 m depth there is a significant freshening of the ocean waters, with the surface waters salinity decreasing up to 0.18 PSU at the end of the simulation. Such tendency can mean that the ocean is still drifting from its initial conditions in the Historical simulation.

This topic has been included in the revised manuscript (Pages 19-20, Lines L17-L28), as well as the Fig. 4 (Page 70).

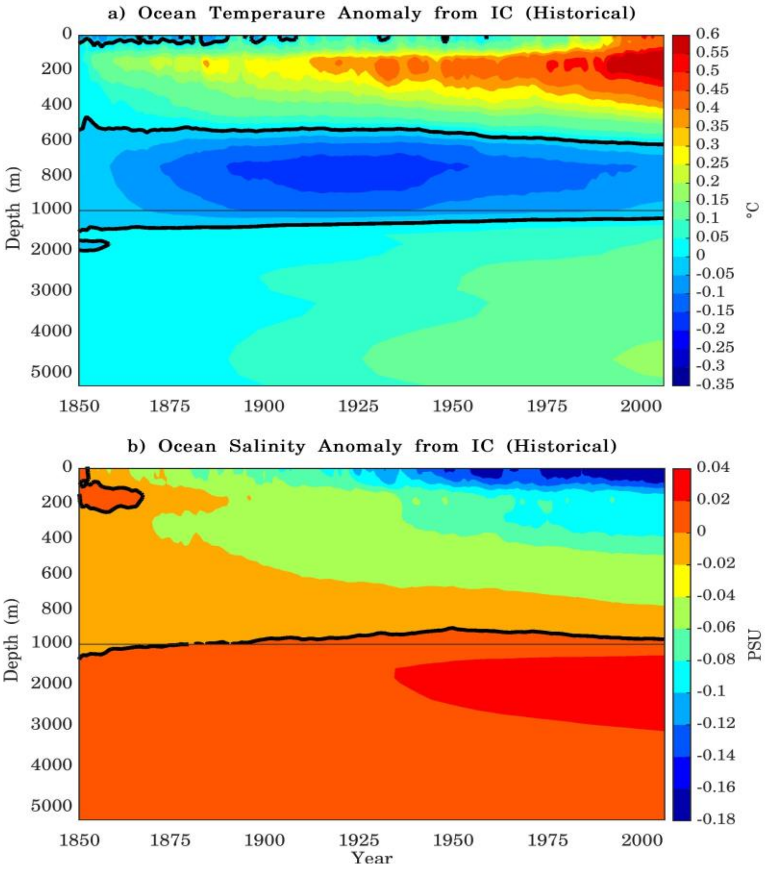

Figure 4 - Depth-time Hovmöller diagrams of global average ocean temperature and salinity anomalies from the respective initial conditions (IC). Here the initial conditions are taken from the 1th year (1850). The diagrams are based on annual average time series simulated by the Historical simulation over the period 1850-2005 (156 years). The thick black line represents the zero contours. Note that the vertical scales are different above and below 1000 m.

2. A documentation of a newly developed couple model should include an estimate of climate sensitivity. The analyses for that should be standard procedure.

Reply:

As mentioned in point 1 above, the group has prepared a manuscript that analyzes the piControl and abrupt 4×CO2 simulations. In this manuscript, it is evaluated the climate sensitivity of the model by analyzing the response of the atmosphere in the piControl and abrupt 4×CO2 simulations (Capistrano et al., 2018). For this reason, the climate sensitivity of the model is not addressed in the revised manuscript. Nonetheless, since it is relevant information, a summary has been included in the revised manuscript (Pages 8-9, Lines L24-L4).

Capistrano et al. (2018) estimates that BESM-OA2.5 has an equilibrium climate sensitivity of 2.96 ºC for the abrupt 4×CO2 experiment. This value is within the range from 2.07 to 4.74 ºC that has been computed for 25 CMIP5 models and close to the ensemble averaged value (3.30 ºC).

3.1 The discussion of the evaluation with observations/reanalyses is descriptive, but does most often not discuss if the biases are acceptable, larger/smaller than in other models, or which consequences come with them.

Reply:

We agree and discussions comparing BESM biases with other CMIP5 models were included in the manuscript: Page 14 (Lines L2-L13), Page 16 (Lines L18-L22), Page 17 (Lines L8-L10) and Page 17 (Lines L21-L23).

3.2 For example, if we assume that including aerosols into the BESM influences the historical simulation in a similar way as in other models (Figure 2), BESM would severely underestimate global warming in the last century.

Reply:

We agree, and a disclaimer has been included in the revised manuscript (Page 12-13, Lines L23-L3).

3.3 This should be a matter of concern and lead the authors to look for the origin of this discrepancy. Or is the plan that everything will be better in the next generation of BESM, as somehow implied in the conclusions?

We agree with your observation. However, models can respond in different ways to external forcing, therefore, in the near future, the aim is to carry out a numerical experiment in which the model is forced with observed estimate of aerosol concentration (as read-in field) in order to address to what extension BESM is impacted.

4. The quality of the figures should be improved. At least in my pdf version I could hardly decipher axis and contour labels.

Reply:

Thank you for pointing out this issue. All figures have been improved.

Minor Revisions:

Abstract, ln 6: "validation" would mean that you have some measures for when a model is valid, better use "evaluation"

Reply:

"validation" has been replaced by "evaluation".

Page 3, ln 6 and following text. The authors say that for an ESM there needs to be an interactive biogeochemical module. But it is not explained later, if or what kind of biogeochemistry model is included and if there are other publications planned on this aspect.

Reply:

The model used in this work is an ocean-atmosphere-biosphere coupled model, indicated in its name BESM-OA version 2.5. Therefore, the current version of the model is not a full Earth System Model, as disclaimed in the manuscript (Page 3, Lines L20-L23). We acknowledge that an Earth System Model is a comprehensive model that includes all Earth system components, including biogeochemical cycles (e.g. dynamic carbon cycle processes) and cloud-aerosol-chemistry processes. Although the main aim of our group is to build up such a model, at the moment the model is a coupled ocean-atmosphere climate model, without the representation of either biogeochemical cycles or cloud-aerosol-chemistry processes. Nonetheless, the name Brazilian Earth System Model was chosen in order to avoid a future change of the model's name on its transition from an ocean-atmosphere coupled model to an Earth System model. For this reason, in the acronym BESM was always added the letters OA, which stand for ocean-atmosphere coupled model. Therefore, at the moment, there is any evaluation of the biogeochemical cycles and their interaction with other Earth system components.

For the next version of the model, the group has been working on activate the biogeochemical model (TOPAZ) within the MOM5 in order to simulate biogeochemical cycles in future simulations. Currently, it is running the stand-alone spin-up of the ocean and biogeochemical models. This information has been included in the revised manuscript (Page 4, Lines L2-L4).

Page 4, ln 1: do you mean interactive aerosols and chemistry or just the ability to use them as read-in fields?

Reply:

At the moment, the aim is to be included as read-in fields only. For clarity, this information has been included in the revised manuscript (Page 4, Lines L1).

Ln 11: "seamless predictions" is it really used for weather predictions (at which resolutions?) and do you have to apply problem-specific parameterizations?

Reply:

Thank you for pointing out this issue. We apologize for the information not being clear. BESM-OA2.5 has not being used for either on short-range and medium-range weather forecasting (<10 days) or climate prediction on a seamless prediction framework. Seamless prediction framework is an aim that the group would like to pursuit in the future. BESM-OA2.3 is used on extended-range weather forecasting (>10 days) and for seasonal prediction.

The sentence about seamless predictions has been deleted: Page 4 (Lines L15-L18).

Page 6, ln 18ff. How does the system behave after switching from CORE to fully coupled. I doubt that the spin-up is long enough.

Reply:

This topic is addressed in more details above, in point 1.1 and 1.2 of major revisions.

Page 15: SSTs seem to be generally too warm. Is that also true for surface air temperature in the control run or in the beginning of the historical? Wouldn't that call for some tuning exercise, e.g. looking into cloud parameterizations?

Reply:

The surface air temperature (SAT) biases for piControl (Fig. 5) and Historical (Fig. 6) simulations are generally negative over the ocean, with exception of the warm SST biases along the western coast of Africa and the Americas; a common problem of CMIP5 fully coupled model runs (Wang et al., 2014).

Indeed, adjustments of atmospheric model heat fluxes, as those affected by cloud parameterizations is a current object of research in BESM development team.

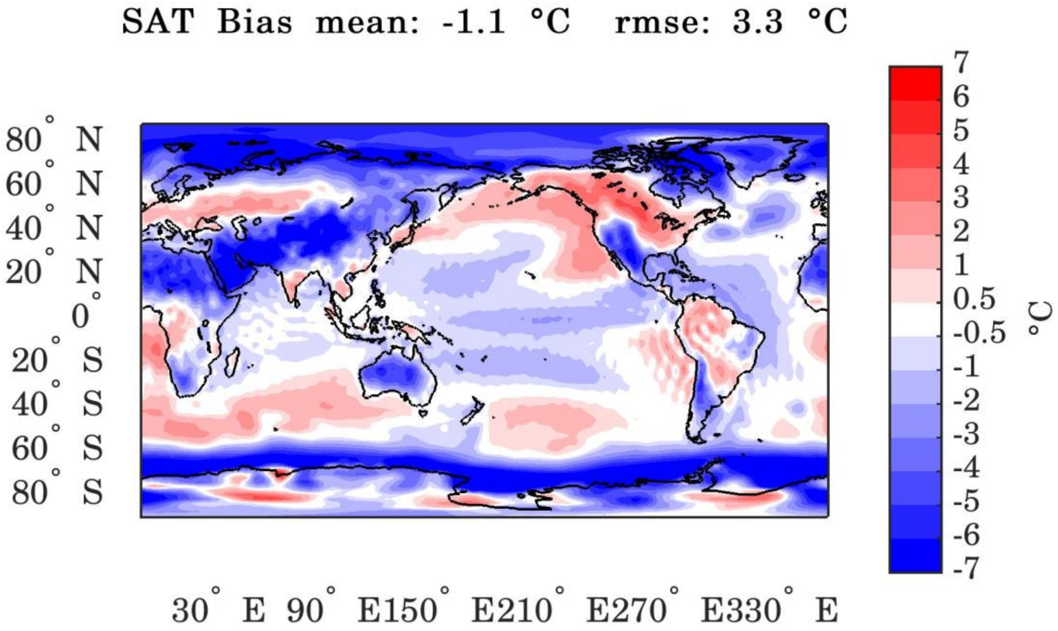

Figure 5 - Spatial map of annual mean surfasse air temperature (SAT) bias of BESM-OA2.5 (piControl) relative to 20CRv2. The averages values are computed over the periods 971–1000 (for BESM-OA2.5) and 1871–1900 (for 20Crv2). Units are in C.

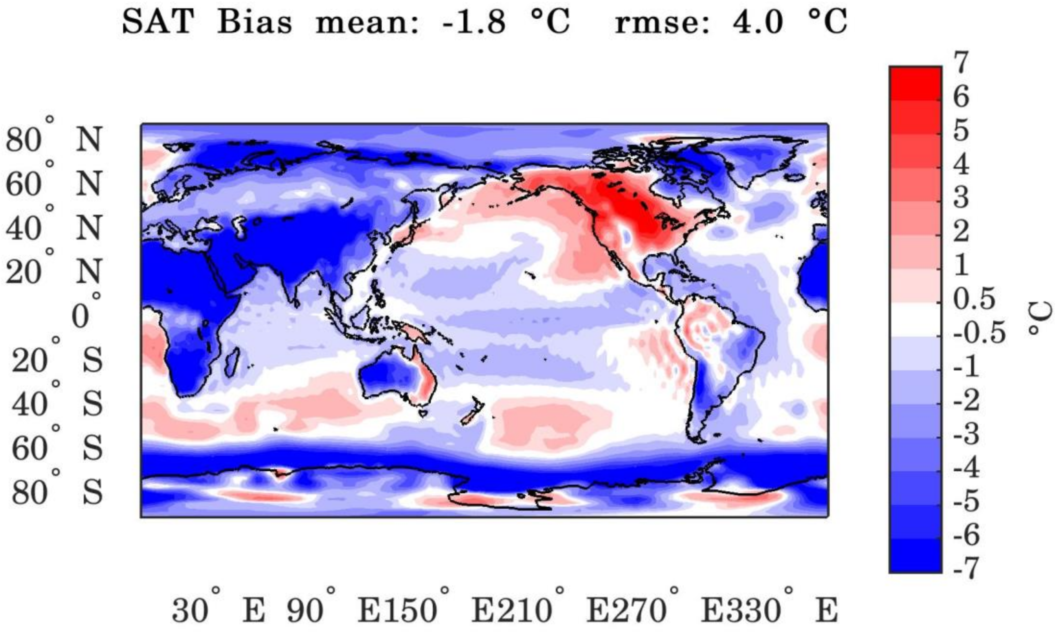

Figure 6 - Spatial map of annual mean surfasse air temperature (SAT) bias of BESM-OA2.5 (Historical) relative to 20CRv2. The averages values are computed over the periods 1871–1900 (for BESM-OA2.5) and 1871–1900 (for 20Crv2). Units are in C.

Page 17, ln 21ff. How does AMOC look in the control run, any drift?

Reply:

This topic is addressed in more details above, in point 1.1 of major revisions. There is a drift but is comparable with other state-of-the-art models.

Bibliography

Capistrano, V. B., Nobre, P., Tedeschi, R., Silva, J., Bottino, M., da Silva Jr., M. B., Menezes Neto, O. L., Figueroa, S. N., Bonatti, J. P., Kubota, P. Y., Reyes Fernandez, J. P., Giarolla, E., Vial, J., and Nobre, C. A.: Overview of climate change in the BESM-OA2.5 climate model, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-209, in review, 2018.

Jones, G. S., Stott, P. A. and Christidis, N.: Attribution of observed historical near-surface temperature variations to anthropogenic and natural causes using CMIP5 simulations, J. Geophys. Res. Atmos., 118(10), 4001–4024, doi:10.1002/jgrd.50239, 2013.

Wang, C., Zhang, L. and Lee, S.: A global perspective on CMIP5 climate model biases, Nat. Clim. Chang., 4(3), 201–205, doi:10.1038/NCLIMATE2118, 2014.