This is an excellent model description paper from the team of Beijing Climate Center (BCC), which is a major climate modeling center in China and has been engaged in the development of climate/earth system models in the past decades, with excellent contributions to previous phases of CMIP. In this manuscript, the authors have documented the key technical details of the model which are crucial to the users of CMIP data of BCC models. The results are also useful to model developers as a reference for model development and improvement. In the manuscript, a comprehensive comparison of the historical simulations from middle- and high-resolution models of BCC is performed. The logic of the manuscript is well organized and helps model developers and users to know what kind of improvements can be achieved by developing a high-resolution version of the model.
resolution ocean-atmosphere coupled climate system model. While I believe that the manuscript can be accepted for publication in GMD, I still find some spaces for a further improvement, such as the inconsistent period of model and observations, lack of explanations on the improvement or backset in high resolution modelling etc. A moderate revision is needed.

Major comments: 1. Logic of the manuscript: It would be better to collect the descriptions of observational datasets in a new section before “Results”, instead of in each result subsection. The resolution of all the observational data used should also be marked. 2. How did you compare the low- and high-resolution data on a lat-lon map? Do you interpolate from low to high or high to low? and Why? Similar to the observational data, all the methods used should also be introduced and summarized before showing the results. 3. In Table 2 and related subsection, I think using the same period as CERES-EBAF product to evaluate the two-version models is better. What is the meaning of errors (how do you calculate it) in Table 2 and text? 4. I wonder why the period of 1971-2000 is used. According to the description of historical simulation of these two models (L356-359), they both ends at 2014 as recommended by CMIP6. So, using period of 1995-2014 should be better as more observational data are available. 5. Figure 7 uses a different color set to represent high and mid resolution models from Figure 2. For reading more easily, I recommend to make the color legend consistent throughout the manuscript. 6. L463-464: Can you explain why HR model improves the DJF precipitation in the SPCZ? Is it controlled by resolution or parameterization? Such kinds of information are very helpful to other model developers. 7. Figure 8: Here I think you should use the period of 2001-2014. 8: Figure 9: Large biases in Kuroshio extension and North Atlantic in higher resolution model should be marked and give possible reasons. I wonder whether this bias is resulted from the coarse resolution of observation, viz. the observation is “wrong” here due to its low resolution. 9: Figure 10: The color bar is weird. It is not easy to capture the relative magnitude, especially the areas with biases around zero value. 10: Why does HR model improve the TC density in western Pacific but not in the North Atlantic? Any explanations?
11: It is very interesting that the HR model can produce an excellent wind-pressure relation. Can you give a short physical explanation? 12: The color bar in Figure 17 should also be replaced by that or similar type used in Figure 16. 13: In Figure 18, the time series in (a)-(c) subpanels are not suitable for comparison. Maybe you can use probability density function to show the asymmetry and skewness of ENSO. 14. How weaknesses of observational data could influence the model evaluation, especially for the high-resolution result is recommended to be discussed. For example, low resolution SST data is unable to capture the SST gradient along the Kuroshio and Gulfstream extension regions, it would be unfair for high resolution models if you use low resolution data as observational metrics.

Minor comments:

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2020-284, 2020.