We thank the two anonymous referees for their valuable comments and constructive suggestions on the manuscript. Below, we explain how the comments and suggestions are addressed and make note of the revision we made in the manuscript.

Anonymous Referee #1

General comments:

- ***This study examines uncertainties in dust size distribution, AOD, and radiative forcing from size parameterization using WRF-chem model. They found strong biases in loading and radiative forcing aginst the best size parameterization (BIN8) when coarser size distributions (BIN4, MOD3, and MOD3_tuned) are used. As the global scale dust simulations using WRF-chem is still relatively rare, this study shows the importance of using proper size parameterization and there result would be beneficial to the readers. The manuscript is well written, but the following issues to be answered before publication.***

We thank the reviewer for a detailed review. Both text and figures are revised as the reviewer suggested.

Specific comments:

- ***The experiment setup needs to be clear. MADE/SORGRAM uses modal approaches with three lognormal modes and MOSAIC model uses bin approaches with 4 and 8 bins. Meanwhile, same dry and wet deposition schemes are applied for both aerosol schemes. Although the results of two bin models are similar, the difference between bin models and the modal model is huge, therefore authors made an additional run for the modal model (MOD3_tuned) by reducing emission by half. What is the purpose of showing both MOD3 and MOD3_tuned?***

The dust emission in climate models is commonly tuned to better reproduce the observations. The MOD3_tuned case is designed to use the modal size approach with tuning of dust emissions to simulate mean AOD comparable with the satellite retrievals over the dust source regions. Comparing the result of the MOD3_tuned case with that of the BIN8 case provides a more fair comparison between BIN8 and MOD3, when latter is configured using the more commonly adopted approach for simulating dust mass balance and radiative forcing given reasonable dust AOD against observations that are available. With the MOD3_tuned case constrained by satellite retrieved AOD, we also found that dust emission could differ by a factor of two from that used in BIN8 due to differences in the dust size representations.

- ***This experiment setup is rather confusing since the uncertainty comes not only from size parameterization as shown in title but also it comes from using different models. Also dry and wet deposition schemes are model dependent. The most explicit experiment would be changing size distribution only within one aerosol model either MADE/SORGRAM or MOSAIC. Or authors may need to discuss on the uncertainty of using different aerosol models.***

We agree that the comparison will be more straightforward if the bin and modal size approaches are applied in the same aerosol model. However, in version (v3.3.1) of WRF-Chem used in this study (also for the newest version (v3.5) of WRF-Chem), neither the MADE/SORGAM nor MOSAIC aerosol scheme is implemented with both size approaches. In this study, although the bin and modal size approaches are applied in two aerosol models (MADE/SORGAM versus MOSAIC), the two models have been made consistent in the mechanisms of dust emission, dry deposition, and wet deposition. The only difference between the two aerosol models is aerosol chemistry that may not play a significant role in simulating dust life cycle in this study. Therefore the difference between the two aerosol models in simulating dust mass loading and radiative forcing can be mostly attributed to their different size approaches (i.e., mode versus bin). It is now further clarified in the model description "*Although it would be more straightforward to compare the bin and modal size approaches in the same aerosol scheme, neither of the aerosol schemes is implemented with both size approaches in version (v3.3.1) of WRF-Chem used in this study (the same is true for the newest version (v3.5) of WRF-Chem). In order to treat the source and sink processes of dust consistently, both aerosol schemes are configured with the same aerosol emission and dry and wet deposition modules. The only difference between the two aerosol models is aerosol chemistry that may not play a significant role in simulating dust life cycle in this study. Therefore the difference between the two aerosol schemes in simulating dust mass loading and radiative forcing is mostly attributable to their different size representations (mode versus bin)*"

- ***This study used the GOCART dust emission. Discussion on the difference between WRF-chem and GOCART would be helpful as the GOCART model has been tested***

against various observations (Chin et al., 2002; Chin et al., 2009; Huneeus et al., 2011; Kim et al, 2013). Huneeus et al. (2011) have recently provided global estimates of GOCART on emission (3000 Tg/yr), loading (29.5 Tg), wet deposition (583 Tg), AOD550nm (0.035), MEE (0.6 m2/g), and lifetime (3.4 days). Since this study is also in global scale, a discussion that comparing WRF-chem with the global GOCART model is important. I am also curious why the WRF-chem model is so large emission (6000 Tg/yr) compared to GOCART and other global models that are below 3000 Tg/yr (size is less than 10 micron). Is it just tunable value or dust emission of 6000 Tg/yr is your best estimate?

This study used the empirical proportionality constant (tunable parameter) in the emission scheme provided by Ginoux et al. (2001). The performance of the dust scheme in the two aerosol mechanisms has been evaluated with surface, aircraft, and satellite observations in Zhao et al. (2010). The GOCART dust emission scheme was first coupled with the MADE/SORGAM and MOSAIC aerosol mechanisms in WRF-Chem by Zhao et al. (2010). The scheme in WRF-Chem has been used in several dust studies (e.g., Zhao et al., 2011, 2012; Chen et al., 2013; Kalenderski et al., 2013). The comparison between model simulated and satellite retrieved AOD over the dust source regions shown in Figure 2 also indicates a reasonable performance of WRF-Chem in modeling dust AOD. It is clarified now in section 2.3.1 as "*It was implemented and evaluated by Zhao et al. (2010) over North Africa and has been further evaluated and used over other regions (e.g., North America, East Asia, and the Arabian Peninsula) by previous studies for dust simulations with WRF-Chem (e.g., Zhao et al., 2011, 2012; Chen et al., 2013; Kalenderski et al., 2013).*"

Although the total dust emission mass flux simulated by WRF-Chem with the sectional size approach is 6000 Tg/yr, the emitted mass of dust particles with diameter sizes less than 10 μm is 4600 Tg/yr (summarized in Table 1). This value is near the higher end of the range (4313 Tg/yr by CAM and 3995 Tg/yr by SPRINTARS) reported in Huneeus et al. (2011). The value is about 50% higher than that of 3157 Tg/yr for year 2000 estimated by the GOCART model reported in Huneeus et al. (2011) and also that of ~3000 Tg/yr for 2000-2007 estimated by Chin et al. (2009) and Kim et al. (2013) using the GOCART model. Although this study uses a similar dust emission scheme as used in the GOCART model, the simulated meteorological fields (e.g., surface winds) and land surface conditions (e.g., soil moisture) could be different between WRF-Chem and the GOCART model. In addition, as found in this study, size representations can significantly impact dust modeling. With the constraint from satellite AOD over the dust source regions, the dust emission in WRF-Chem with the modal approach has to be tuned to ~2300 Tg/yr, which is 25% less than that simulated by the GOCART model (Huneeus et al. 2011). This further indicates the complication in comparing dust emissions from different models. A full investigation of the difference between WRF-Chem and GOCART in simulating dust emission is beyond the scope of this study. Now some discussion and comparison are added in the text "*The annual total global dust emission simulated by the three cases is about 6000 Tg yr^{-1}. The emitted mass of dust particles with diameter sizes less than 10 µm is 4600 Tg yr^{-1}, which is near the higher end of the range (3995-4313 Tg yr^{-1}) reported by Huneeus et al. (2011). This value is about 50% higher than 3157 Tg yr^{-1} estimated by the GOCART model for year 2000 (Huneeus et al. 2011) and ~3000 Tg yr^{-1} estimated by Chin et al. (2009) and Kim et al. (2013) for 2000-2007 using the GOCART model. Although this study uses a similar dust emission scheme as used in the GOCART model, the simulated meteorological fields (e.g., surface winds) and land surface conditions (e.g., soil moisture) could be different between WRF-Chem and the GOCART model. A full investigation of the difference between WRF-Chem and GOCART in simulating dust emission is beyond the scope of this study.*"

- ***Although authors wrote that the large differences in AOD and radiative forcing are due to the different size distribution (First paragraph in Page 19664, Table 2, and Figure 6), it is not shown how size distribution play a role for AOD and radiative forcing calculation. I would like to suggest comparing AOD and MEE (mass extinction efficiency) at 550 nm for different size bins. For example, if the ratio of PM2.5 to PM10 (in Page 19666, line 11-12 and Figure 7) is also calculated for AOD, what's there distribution?***

First, we need to clarify that the objective of this study is to investigate the impact of different size representations, not size distributions, on dust radiative forcing. Different size representations will result in not only different size distributions but also different mass and AOD.

Instead of showing the radiative forcing for different size bins, the radiative forcing efficiency (i.e., radiative forcing per unit AOD) is shown and can be used as an indication for the impact of size distributions on radiative forcing. The result shows similar dust radiative forcing efficiency for the same size distribution (MOD3 versus MOD3_tuned) regardless of the difference in AOD. More discussion about the dust radiative forcing efficiency is now added in section 4.3 "*Furthermore, different size representations also lead to significant differences in dust radiative forcing efficiencies (i.e., radiative forcing per unit AOD), also summarized in Table 2. The radiative forcing efficiency can be used as an indication of the impact of size distributions on radiative forcing. With similar dust size distributions (Figure 6), the MOD3 and MOD3_tuned cases simulate a similar dust radiative forcing efficiency regardless of their large difference in AOD. The two mode cases simulate the largest dust radiative forcing efficiency at TOA and the BIN4 case simulates the smallest value. The BIN8 case is in between. Although the mode size representation simulates larger dust radiative forcing at the surface than the bin size representation, the BIN4 case simulates the largest dust radiative forcing efficiency at the surface and in the atmosphere. The BIN8 case and the two mode cases simulate the smallest dust radiative forcing efficiency at the surface and in the atmosphere, respectively. Overall, the difference among the four cases in dust radiative forcing efficiency is a factor of 3 at TOA, a factor of 4.5 in the atmosphere, and a factor of 1.5 at the surface, which is smaller than the difference in dust radiative forcing.*" and in the conclusion section "*The radiative forcing efficiency is also analyzed to reflect the impact of size distributions on radiative forcing. With similar dust size distributions, the two mode cases simulate similar dust radiative forcing efficiency regardless of their large difference in AOD.*"

It is technically difficult to compare AOD for different size bins since version (v3.3.1) of WRF-Chem does not diagnose size resolved AOD and radiative forcing. However, the distribution of the ratio of $PM_{2.5}/PM_{10}$ in AOD should follow that in mass, although the magnitude may have some difference. For a particular size bin, AOD should be proportional to mass given the same optical property (i.e., refractive index). We agree that it would be nice to quantify the impact of the size distributions on MEE. However, again, it is technically difficult to compare MEE at different size bins in this study since it is not diagnosed in the WRF-Chem version that is used.

Minor Comments:

- ***Page 19658, Line 22: Is it two lognormal-modes or three lognormal-modes?***

It is two lognormal-modes. In WRF-Chem using the MADE/SORGAM aerosol scheme, dust is only simulated in the accumulation and coarse modes.

- ***Please check font size in figure 6 is legible.***

Improved.

Anonymous Referee #2

General comments:

- ***The paper "Uncertainty in modeling dust mass balance and radiative forcing from size parameterization" contrasts the sectional and modal representation of dust aerosol in the WRF-Chem model and how this influences important aerosol processes, such as deposition and the direct radiative effect. The work builds upon findings from Zhao et al. (2010) that highlighted the differences resulting from size distribution representation. The key findings are that the modal representation under-estimates coarse dust and over-estimates the fine dust aerosol, relative to the sectional schemes. This results in a longer dust lifetime, greater AOD and greater radiative impact relative to the bin schemes.***

  While this feels like a relatively small increment in extra knowledge upon the Zhao et al. (2010) research, the paper is mostly well-written and the figures generally clear. However, there are several points that need addressing relating to both the method and also the presentation prior to recommendation for publication.

Zhao et al. (2010) first raised the issue in modeling dust using a modal size representation with fixed standard deviation. However, this study investigates the impact of different size representations in simulating dust mass balance and radiative forcing quasi-globally. Therefore, we believe this study provides valuable knowledge for further understanding the uncertainty in modeling dust from size representations. We thank the reviewer for a detailed review. Both text and figures are revised as the reviewer suggested.

Specific comments:

- ***The MOSAIC model can be implemented as either a bin or modal aerosol scheme. This would make comparison between the two size representations much more transparent. The authors need to explain why they did not take this approach.***

We agree that the comparison will be more transparent if the bin and modal size approaches are applied in the same aerosol model. However, in version (v3.3.1) of WRF-Chem used in this study (also for the newest version (v3.5) of WRF-Chem), the MOSAIC aerosol model is only implemented with the bin size approach. In this study, although the bin and modal size approaches are applied in two aerosol models (MADE/SORGAM versus MOSAIC), the two models have been made consistent in the mechanisms of dust emission, dry deposition, and wet deposition. The only difference between the two aerosol models is aerosol chemistry that does not play a significant role in simulating dust life cycle in this study. Hence the difference between the two aerosol models in simulating dust mass loading and radiative forcing can be mostly attributed to their different size approaches (i.e., mode versus bin). It is now further clarified in the model description "*Although it is more straightforward to compare the bin and modal size approaches in the same aerosol scheme, neither of the aerosol schemes is implemented with both size approaches in version (v3.3.1) of WRF-Chem used in this study (the same is true for the newest version (v3.5) of WRF-Chem). In order to treat the source and sink processes of dust consistently, both aerosol schemes are configured with the same aerosol emission and dry and wet deposition modules. The only difference between the two aerosol models is aerosol chemistry that may not play a significant role in simulating dust life cycle in this study. Therefore the difference between the two aerosol schemes in simulating dust mass loading and radiative forcing is mostly attributable to their different size representations (mode versus bin)*"

- ***The GOCART dust scheme is well validated within the GOCART model. Is there a reference for validation of the dust scheme within WRF-Chem? I couldn't see one in Zhao et al. (2010) - was this the first implementation? It would be good to see some validation of this, especially since the estimated global emissions are 6000Tg/year –almost double that produced by the GOCART model (for 2000) and 5 times the median of the AEROCOM models (Huneeus et al., 2011).***

The GOCART dust emission scheme was first coupled with the MADE/SORGAM and MOSAIC aerosol mechanisms in WRF-Chem by Zhao et al. (2010). The performance of the dust scheme in the two aerosol mechanisms has been evaluated with surface, aircraft, and satellite observations in Zhao et al. (2010). The scheme in WRF-Chem has been used in several dust studies (e.g., Zhao et al., 2011, 2012; Chen et al., 2013; Kalenderski et al., 2013). The comparison between model simulated and satellite retrieved AOD over the dust source regions shown in Figure 2 also indicates a reasonable performance of WRF-Chem in modeling dust AOD. It is clarified now in the section 2.3.1 as "*It was* *implemented and evaluated by Zhao et al. (2010) over North Africa and has been further evaluated and used over other regions (e.g., North America, East Asia, and the Arabian Peninsula) by previous studies for dust simulations with WRF-Chem (e.g., Zhao et al., 2011, 2012; Chen et al., 2013; Kalenderski et al., 2013).*"

Although the total dust emission mass flux simulated by WRF-Chem with the sectional size approach is 6000 Tg/yr, the emitted mass of dust particles with diameter sizes less than 10 μm is 4600 Tg/yr (summarized in Table 1). This value is near the higher end of the range (4313 Tg/yr by CAM and 3995 Tg/yr by SPRINTARS) reported in Huneeus et al. (2011). The value is about 50% higher than that of 3157 Tg/yr for year 2000 estimated by the GOCART model reported in Huneeus et al. (2011) and also that of ~3000 Tg/yr for 2000-2007 estimated by Chin et al. (2009) and Kim et al. (2013) using the GOCART model. Although this study uses a similar dust emission scheme as used in the GOCART model, the simulated meteorological fields (e.g., surface winds) and land surface conditions (e.g., soil moisture) could be different between WRF-Chem and the GOCART model. In addition, as found in this study, size representations can significantly impact dust modeling. With the constraint from satellite AOD over the dust source regions, the dust emission in WRF-Chem with the modal approach has to be tuned to ~2300 Tg/yr, which is 25% less than that simulated by the GOCART model (Huneeus et al. 2011). This further indicates the complication in comparing dust emissions from different models. A full investigation of the difference between WRF-Chem and GOCART in simulating dust emission is beyond the scope of this study. Now some discussion and comparison are added in the text "*The annual total global dust emission simulated by the three cases is about 6000 Tg yr^{-1}. The emitted mass of dust particles with diameter sizes less than 10 μm is 4600 Tg yr^{-1}, which is near the higher end of the range (3995-4313 Tg yr^{-1}) reported by Huneeus et al. (2011). This value is about 50% higher than 3157 Tg yr^{-1} estimated by the GOCART model for year 2000 (Huneeus et al. 2011) and ~3000 Tg yr^{-1} estimated by Chin et al. (2009) and Kim et al. (2013) for 2000-2007 using the GOCART model. Although this study uses a similar dust emission scheme as used in the GOCART model, the simulated meteorological fields (e.g., surface winds) and land surface conditions (e.g., soil moisture) could be different between WRF-Chem and the GOCART* *model. A full investigation of the difference between WRF-Chem and GOCART in simulating dust emission is beyond the scope of this study.*"

- ***Other studies have performed bin-mode aerosol comparisons (e.g. Mann et al., 2012). While this study does not consider dust it may be useful to compare and contrast your findings in how a modal scheme influences the size distribution relative to a sectional scheme.***

More discussion about the background in bin-mode aerosol comparison is now added in the introduction "*The modal approach is also often further simplified (e.g., assuming a constant standard deviation of the log-normal distributions) and hence has biases in simulating aerosol size distributions (e.g., Zhang et al., 1999; Zhao et al., 2010). Herzog et al. (2004), using a box model, found that a modal approach simulates a difference of <20% in number concentrations and surface area density compared to a bin approach. Kokkola et al. (2009) found that sectional and modal approaches simulate significantly different stratospheric conditions perturbed by volcano eruption. Mann et al. (2012) compared the modal and sectional approach in modeling aerosol microphysics using a 3-D global offline chemistry transport model. They found differences in annual mean surface-level masses of sulfate, sea-salt, black carbon (BC), and organic carbon (OC) within 25% in nearly all regions. Although these studies have demonstrated some differences in modeling aerosols between modal and sectional approaches, none of them focused on dust and the impact of different size representations on simulating dust and its radiative forcing has not yet been investigated.*"

Mann et al. (2012) does not investigate the impact of size representations on dust simulations. It is difficult to do a direct comparison between this study and Mann et al. (2012). In addition, the uncertainties identified in this study depend to some degrees on the size distribution of emitted dust and σ_g. The results may also be somewhat sensitive to how processes such as dry and wet deposition are parameterized in the model. Some discussion between this study and Mann et al. (2012) is added now in the text "*The 3-mode approach (MOD3 and MOD3_tuned) retains more fine dust particles but fewer coarse dust particles compared to the 8-bin approach. The prescribed σ_g (i.e., with a constant value) is the main contributor to the bias of the 3-mode approach in* *representing dust size distribution (Zhao et al., 2010). This result is consistent with Mann et al. (2012), who found that the modal approach simulates lower number concentrations for larger aerosol particles than the bin approach. In addition, they also found that the result from the modal approach is sensitive to the fixed* σ_g.” and “*This study found a difference of 25% in dust mass loading between the simulations using modal and sectional size representations, which is consistent with what Mann et al. (2012) found for other aerosols. However, the difference in dust radiative forcing (up to a factor of 10) is much larger, indicating aerosol size representation may have much larger impact on modeling aerosol radiative forcing than aerosol mass.*”

- ***Radiative forcing is the change in radiative effect relative to pre-industrial conditions. This is not what is shown in the paper, therefore the title and relevant text throughout the paper require alteration.***

Radiative forcing is a metrics used to measure the response of the climate system to some perturbations. It is defined as the net change in the energy balance of the Earth system due to an imposed perturbation such as greenhouse gas, aerosol, cloud, and land surface properties, as defined in the fourth assessment report of the Intergovernmental Panel on Climate Change (IPCC AR4). It can be estimated based on changes in the energy balance between preindustrial and present day, or as in this study, calculated by the models as the difference in the radiative fluxes with and without the perturbation (dust).

- ***Often results and numbers are simply listed in the text (e.g. for mass loading, deposition, radiative effects). I recommend tabulating these to create a useful reference for the reader and then discussing the differences between aerosol distribution representations in the text. This will reduce the amount of unnecessary text and allow for a more involved discussion of the differences and the implications.***

We have reduced listing of numbers in the text. All the numbers related to dust AOD, mass balance (e.g., mass loading, deposition), and radiative forcing are already summarized in Table 1 and 2. The Tables are made more concise now. In the revised manuscript, values of the dust radiative forcing efficiency for all the cases are also summarized in Table 2. More discussion about the dust radiative forcing efficiency is now added in section 4.3 "*Furthermore, different size representations also lead to significant differences in dust radiative forcing efficiency (i.e., radiative forcing per unit AOD), also summarized in Table 2. The radiative forcing efficiency can be used as an indication of the impact of size distributions on radiative forcing. With similar dust size distributions (Figure 6), the MOD3 and MOD3_tuned cases simulate similar dust radiative forcing efficiency regardless of their large difference in AOD. The two mode cases simulate the largest dust radiative forcing efficiency at TOA and the BIN4 case simulates the smallest value. The BIN8 case is in between. Although the mode size representation simulates larger dust radiative forcing at the surface than the bin size representation, the BIN4 case simulates the largest dust radiative forcing efficiency at the surface and in the atmosphere. The BIN8 case and the two mode cases simulate the smallest dust radiative forcing efficiency at the surface and in the atmosphere, respectively. Overall, the difference among the four cases in dust radiative forcing efficiency is a factor of 3 at TOA, a factor of 4.5 in the atmosphere, and a factor of 1.5 at the surface, which is smaller than the difference in dust radiative forcing.*"

- ***The prescribed modal standard deviation is considered the main contributor to the bias in the modal representation relative to the sectional scheme. However, Fig. 6 suggests that the distribution produced in remote regions by the sectional schemes may not be able to be reproduced by a modal scheme even with a varying modal standard deviation. Has any research previously been done to determine if a three-moment implementation does give a modal scheme more skill? Would the extra degree of freedom negate any computation saving gained by using a modal scheme rather than a sectional scheme?***

The two-moment modal size distribution with prescribed standard deviation assumes that the shapes of size distributions would not change during the simulations, which may not be true. A three-moment modal approach can predict the standard deviation, which will definitely be more realistic. The three-moment modal approach has been used in some previous studies (e.g., Binkowski and Roselle, 2003). Binkowski and Roselle (2003) compared the size distributions simulated by a three-moment modal approach and a high resolution sectional approach. They showed that the two approaches agree much better than what is shown between the two-moment modal approach and 8-bin sectional approach in Figure 6. The extra moment will only add another three aerosol species for a 3-mode modal approach (i.e., total aerosol surface area for each mode) to the existing tens of aerosol species. Therefore, it may not significantly increase the computational time. Now the reference of Binkowski and Roselle (2003) is cited.

- ***What is the reason for using a "quasi-global" simulation? It would be interesting to know how the two representations affect the radiative effect in the polar regions in light of recent research (e.g. Lambert et al., 2013) especially as this is one of the metrics used to contrast the schemes.***

We agree it may be interesting to also investigate the impact of size representations on dust radiative effect in the polar regions. However, there are some technical difficulties to run global WRF-Chem (v3.3.1). Therefore, a quasi-global configuration is used instead. It may be worthwhile to conduct a similar study over the polar regions in the future. One way of doing this is to conduct regional simulations over the polar regions with boundary conditions provided by the quasi-global simulations. This method is actually being applied in one of our on-going works. It is clarified now in the model description "*The quasi-global configuration is used instead of global configuration due to some technical difficulties in running global WRF-Chem in v3.3.1. To our best knowledge, global WRF-Chem has not been used in previous published research. This study is the first to conduct a quasi-global simulation using WRF-Chem.*" and in the conclusion "*This study with quasi-global configuration also cannot assess the size impact on dust radiative effect over the polar regions, which may be of interest due to the potential role of dust in polar amplification of global warming (Lambert et al., 2013).*"

- ***Is the model co-sampled with MODIS and MISR data availability? Is Level 3 MODIS data used? Was any additional cloud screening and filtering applied?***

Yes. The model is sampled when the retrievals are available. Level 3 daily MODIS data are used. No additional screening and filtering is applied. It is clarified now in the observation description "*When compared to the MODIS and MISR retrieved AOD, the* *model results are sampled from 10 am - 2 pm for averaging and at the locations where retrievals are available.*"

Figures and Minor Comments:

- ***pg 19672, ln 25 - "focing"***

Corrected.

- ***Table 1 seems redundant, consider merging with Table 2***

Merged.

- ***Table 2 – Other than for emission, the separation of D < 10um and Total seems redundant. These could be stated in the caption to make the table more concise and legible.***

Revised as suggested.

- ***Fig. 4 and Fig. 5 - consider using a different color scale that highlights the relative differences (e.g. blue-white-red as in Fig. 7)***

Both Figure 4 and 5 use the color scale as Figure 7 now.