Review on the “Uncertainty in modeling dust mass balance and radiative forcing from size parameterization" by C. Zhao et al.

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 befor publication.

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? 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.

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?

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?

Minor Comments:
Page 19658, Line 22: Is it two lognormal-modes or three lognormal-modes?
Please check font size in figure 6 is legible.