Comment on acp-2022-406
Anonymous Referee #1

Referee comment on "Global distribution of Asian, Middle Eastern, and North African dust simulated by CESM1/CARMA" by Siying Lian et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2022-406-RC1, 2022

Review of “Global distribution of Asian, Middle Eastern, and Saharan dust simulated by CESM1/CARMA” for ACP

This is an interesting modeling paper that focuses on the vertical distribution of dust and how that varies depending on the three main source regions listed in the title. The paper is overall well-written and interesting, but has a number of issues in comparisons against data and it is not clear what this paper contributes that is not already in the literature. The paper needs major revisions.

Main comments:

- Although this seems to be solid work (but see comments on methodology below), the rationale for this work is in my opinion not really articulated. Reading the abstract, I did not really learn anything that was not already in the (excellent) previous related paper by Froyd et al. (2022) and by other papers on the contributions of different dust source regions by Tanaka and Chiba (2006), Chin et al. (2007), and Kok et al. (2021). In other words, it’s not really clear to me what the reader can learn from this paper that is not already in the literature. For this paper to be published, that should be convincingly articulated in the abstract and other key parts of the paper.
- The processing of data and comparisons against measurements needs more rigor and detail:
My understanding is that the cruise data from (Mahowald et al., 2009) mentioned in 2.5 represent daily values, not monthly values as listed on line 194. Comparing this against an annually averaged climatology as you do opens you up to tremendous biases. For instance, the cruise data might be weighted towards a particular season (e.g., summer, which would bias the data high relative to an annual mean). Also, the daily concentration can vary by orders of magnitude so comparing against an annual mean will generate a very large representation error. At the least, these issues need to be discussed and realistic error bars need to be included in your comparisons, as done for instance in Huneeus et al. (2011). You also should match the season or month in which the cruise data are taken for a more meaningful comparison.

The “long-term” measuring station data is a misnomer as it is really only >3 days of data (see Mahowald et al., 2009, p. 251). So this data is subject to largely the same issues as the cruise data. Please explain that as well.

The comparison against the Atom dust data is one of the more novel parts of the paper as I think this hasn’t been done before. How exactly were the ATom data processed? Does each data point represent an average over some time period (e.g., an hour) or over some area? Were the different measurements simply averaged or are you showing median values? And here the same problems with representation errors apply, which should be discussed. And you should at least match the month or season of the measurement, rather than compare against an annual climatology (or maybe you did? If so, it’s not clear from the text and figure 4).

Please include basic statistics in your comparisons against data (Figs. 3b, c, 4c, 6b), including correlation coefficients and root-mean-square error.

Figure 5 has some issues. The legend in (a) is too small to read; please enlarge. The line styles are confusingly labeled in that there are two solid lines (one with symbols) and no dashed lines. The difference between the two simulated lines is also not clear from the figure caption. Most importantly, how is the comparison between measurements and simulations done and is this an apples-to-apples comparison? I assume a simple average was taken for all measurements within the specified latitude band? But then how was the simulation average constructed? Did you select the closest match in space for each measurement and then average those? Or did you take the average for the entire region in the monthly-averaged simulation? The latter would be problematic for some of the reasons above and the former would be more of an apples-to-apples comparison.

The comparison against AERONET data (Fig. 6) also needs more detail. First, some objective criteria should be provided that justifies picking the specific sites that were used. Second, since it’s more common to use AOD in the visible wavelength, could you provide a reason why you used AOD at 1020 nm? I assume because a larger fraction of the signal is dust at that wavelength? Third, since your comparison depends on the accuracy of the simulated AOD of other aerosols, please provide maps of the simulated dust and non-dust AOD to help interpret your results, for instance in the supplement.

Section 3.2: It seems clear that your model underestimates dust because it underestimates the various measurements of surface dust by ~10-60% (Figs. 3 and 4) and underestimates AOD in dusty regions by ~10-20% (Fig. 6). Why don’t you simply scale up the factor C on line 149, which as you note is an “arbitrary constant” anyways? Since your dust does not feedback onto the meteorology (at least, I think so; please specify this in section 2.1), you could simply do the rescaling offline.

Section 4: relating to my first comment, previous work by for instance Tanaka and Chiba (2006), Chin et al. (2007), and Kok et al. (2021) has studied how different source regions contribute to the global distribution of dust. The authors should discuss how their results differ from these previous works and what new information we learn from their study that we didn’t already know from those previous studies. Also, are the relative contributions of the different source regions supported by experimental constraints on source provenance of deposited dust?
Other comments:

- Doesn’t your source region in North Africa include the Sahel? It’s difficult to see in Figure 2. If so, then “Saharan” in the title and at many places in the paper should probably be changed to “North African”
- Relatedly, please specify the exact coordinates of the three source regions shown in Figure 2 for clarity.
- Lines 18-19: The statement about uncertainty due to shape and density seems a bit odd here. Although this statement is probably correct, the effects of dust shapes and density are not explored (or even discussed) in this paper. So I’d suggest either discussing this in the paper or removing it from the abstract.
- Lines 59-61: I think the statement here is not quite right. My understanding is that Kok et al. (2021) found that models simulate a ~65% contribution of North African dust and that this is an overestimate, not that models overestimate North African dust by 65%. Please correct.
- Line 104-5: Zender et al. (2004) is two decades old so the range of factor of four in dust burden in that study is not meaningful for the current state of knowledge. Please cite more recent work here and adjust the sentence accordingly.
- The statement is made in a few places in the paper (e.g., lines 118-9) that the model is “constrained” by the Atom measurements. But my reading of the paper is that the authors merely compare their simulations against the Atom measurements and don’t use those measurements to constrain the model. Please correct accordingly.
- The treatment of CARMA of two groups of particles, externally mixed sulfate particles and internally mixed “everything” particles (BC, condensed sulfate, organics, sea salt, dust) seems odd to me. As primary emitted particles, dust tends to be much less externally mixed than sulfates, which condense onto almost everything. Do you have a justification that this treatment is realistic, especially for dust? And given that dust is internally mixed in your model, what does the diameter of a dust particle mean (e.g., diameter < 4.5 um)? Does that really mean the diameter of the internal mixture that includes dust? Please provide some more details here so the simulations can be interpreted correctly by the reader.
- Line 138: please specify the wavelength for which these are the assumed optical properties and how those optical properties vary with wavelength / radiation band. Also, unless this is for ~400 nm or something, your dust is much too absorbing according to current understanding (e.g., Sinyuk et al., 2003; Balkanski et al., 2007; Di Biagio et al., 2019). I don’t think that matters for this study since you’re not looking at radiative effects, but it should be noted.
- Line 143: Please include a more appropriate and standard reference for this statement on the basic physics of dust emission
- Line 163: please elaborate on what you mean by “secondary activation” here
- Line 169: what do you mean by “solubility” here, exactly? Is this the assumed fraction of the dust particle mass that is soluble? If so, isn’t 0.2 much too high? Please include some justification for this particular model choice.
- Line 184-6: what diameter type did the PALMS instrument measure, exactly? If it’s based on time-of-flight, then I assume it’s the aerodynamic diameter? And I assume the model simulated geometric diameter? Did the authors make corrections to convert between the two diameter types? If not, why not and what sort of errors from
neglecting this do you expect?

- Line 209-210: This statement doesn’t make sense to me. Silt is >2 um diameter.
- Figure 9: For clarity, please note in the caption that this is a zonal average.
- Line 427-428: Do you have support for the statement that dust contributes only 0.04% and 3% to aerosol mass at 100 and 200 hPa, respectively? I didn’t see any results in the paper that directly address that. Perhaps you could include results of non-dust mass concentrations in the supplement.

References

- Sinyuk, A., Torres, O., Dubovik, O., 2003. Combined use of satellite and surface observations to infer the imaginary part of refractive index of Saharan dust.