Geoscientific
Model Development
Discussions

Interactive comment on "Multi-model simulations of springtime dust storms in East Asia: Implications of an evaluation of four commonly used air quality models (CMAQv5.2.1, CAMxv6.50, CHIMEREv2017r4, and WRF-Chem v3.9.1)" *by* Siqi Ma et al.

Anonymous Referee #1

Received and published: 10 June 2019

The authors describe and evaluate multiple dust simulation options available in WRF-Chem, CMAQ, CAMx, and CHIMERE using a springtime dust event in Northeastern China for comparison. The objective of the paper is to document model performance and to evaluate the sensitivity of the simulation results to key aspects of the dust emission schemes. While this discussion is useful, I find the authors have drawn conclusions about governing factors without sufficient evidence. This study attempts to assess multiple dust simulation capabilities embedded in multiple models while also

assessing the influence of dust source treatments. The paper is overall well written and organized, but there are too many degrees of freedom left unexplored to determine causality of the resultant dust concentration simulations.

I recommend revision of the manuscript, considering the following comments:

Major comments: The methodology, discussion, and results sections of this manuscript primarily focus on differences between the dust emission treatments used in each model simulation; however, the individual model descriptions provided in section 2.4 provide little to no information about the algorithms comprising these schemes. There really needs to be a succinct summary of the dust emission scheme equations discussed in this paper, either directly in the text or in the appendix section. Suggest using a model flow chart similar to the approach used in Darmenova et al. (2009) or LeGrand et al. (2019) for each dust emission scheme discussed and a symbology table.

The authors state that WRF v3.9.1 was used to generate the meteorological fields used to force all of the dust models discussed in the manuscript. This is confusing. WRF-Chem is an inline model. The dust emissions and airborne concentrations evolve simultaneously with the atmospheric conditions. In other words, the dust modules in the WRF-Chem assessments were likely subject to different environmental forcing conditions than those in the CMAQ, CHIMERE, and CAMx dust modules. Did the authors use the coupled WRF-CMAQ implementation as well? What was the output frequency of the WRF v.3.9.1 output (wrfout) files? This could potentially have significant influence on the results. Furthermore, are the CHIMERE and CMAQ dust modules configured to ingest windspeed (U) or friction velocity (u*)? The dust emission calculations described in this paper, with the exception of the WRF-Chem GOCART dust emission scheme, are calculated in terms of u*. Are the u* fields being ingested by the dust emission flux equations in WRF-Chem, CMAQ, and CHIMERE identical? If so, please add a figure showing the surface U and u* fields for a few time periods in the case study sequence. If not, please add a figure showing how they vary (especially if each model is doing its

own U to u* conversion) as this could be important for deciphering causative factors in model output discrepancies.

P13L4-7: The authors did not include the CAMx dust simulation in their in-depth analyses because the dust mask field required by the CAMx dust emission scheme did not include an erodible area in their region of interest. I'm confused by this reasoning. The dust mask and the dust source maps discussed for the other schemes in WRF-Chem (Figure 3) essentially serve the same purpose. Why test out different dust source fields in WRF-Chem but not the CAMx model? Claiming the paper includes an assessment of the CAMx model seems misleading to me. Recommend the authors either test the CAMx dust emission scheme with alternate dust source treatments similar to the exercise done for WRF-Chem, or remove the CAMx model and its discussion from the manuscript entirely.

P15L20-28: The strong dust emission magnitude from UoC and AFWA compared to GOCART in this study is somewhat unexpected given the findings discussed in the LeGrand et al. (2019) paper cited here. I don't think there's enough evidence to associate the excessive flux from the AFWA scheme with the saltation bin settings. I suspect these results may actually be related to the authors' use of the Pleim-Xiu (PX) land surface model (LSM) and Pleim (ACM2) planetary boundary layer (PBL) scheme. The U/u* conversion in the PX/ACM2 setting typically produces stronger u* values than NOAH LSM/PBL combos for equivalent U values. Operational agencies that use the AFWA dust emission scheme with the PX LSM frequently make use of the ustune tuning factor in the WRF-Chem configuration file to tone down u* values ingested by the scheme for this very reason. It would be interesting to see a time series plot of model estimated u* added to the time series plot in the appendix. If there is a strong sensitivity of dust emission scheme performance to LSM choice, it would be worth highlighting. Most other dust emission scheme assessment papers use the RUC or NOAH LSM.

The authors attribute over/under prediction of simulated dust conditions to dust emission scheme setting, but these conclusions are primarily based on comparison to daily

C3

average PM10 distributions. Simulated PM10 errors could also be due to issues with the atmospheric conditions (e.g., vertical mixing) and/or deposition/removal treatment. The validation methodology used for this study shows daily PM10 estimates are sensitive to the dust emission scheme configuration but does not provide enough evidence to confirm causality. This is especially important to note here given that multiple model frameworks are being used for this analysis.

Section 3.5: I don't understand the rational for scaling PM10 concentrations in the inter-model comparisons. Why scale the simulation output rather than the scaling the emission fluxes?

P26L13-14: The authors claim different algorithms for threshold friction velocity (FVT) resulted in significant differences in the simulated dust concentration and spatial distribution. This finding hasn't been demonstrated in this paper. The FVT treatments associated with each model haven't been introduced (again, need for model algorithm summary to guide discussion/conclusions). Recommend adding a figure of panel plots during the peak emission period showing simulated FVT estimates for a given grain size for each dust emission scheme - or - panel plots showing u*-FVT (U - FVT in the case of GOCART).

Minor comments: P4L4-6: I would not qualify this paper as the first comprehensive evaluation of dust models for East Asia. A single dust event case study is good for examination and discussion of how the dust models function under a given forcing condition, but an extended study period with several events would be needed to truly assess model performance.

P8L10: The AFWA scheme is adapted from the dust emission scheme originally described by Marticorena and Bergametti (1995), not GOCART. It would be appropriate to cite the LeGrand et al. (2019) paper here.

P8L19: The UoC coding error was not corrected in the public code distribution until the release of WRF-Chem v4.0. It is unclear here whether or not the authors manually

C4

corrected the coding error in their compilation of WRF-Chem v3.9.1. This was also mentioned by another community member on the forum.

P12L1: The G01 acronym hasn't been defined yet.

P12: The dust source map discussion is difficult to follow. Table 2 provides a good summary, but the labels (e.g., G01, K08, etc.) need to be introduced in the text.

P12L20-21: The authors mention they conducted a ground survey to assess representativeness of the erodibility field. What method was used for the ground survey? Is this something that was done subjectively or with in situ measurements (e.g., wind tunnel or PI-SWERL device)?

P12L20: Please list the NU-WRF version number used to acquire these fields.

Figure 3: Was the G01 1-deg field generated manually for this analysis or acquired from somewhere? The 1-deg field was replaced by the .25-deg field in the WRF-Chem repository back in 2012 and is no longer part of the standard WRF-Chem static dataset download. If it was generated manually, what process was used set the vegetation mask?

P14L1: Please describe the BELD3 dataset.

P14L3: USGS and MODIS land use datasets are brought up here with no context. Given how often land use datasets are brought up in the discussion section, the authors may want to consider listing which land use dataset was used to configure each of the model frameworks. Please also list the number of classes associated a particular dataset, since there are multiple versions of both the USGS and MODIS IGBP land use datasets.

Sections 2.4 and 2.5 in general: The overall model descriptions are somewhat vague. Additional model configuration information would be needed if others wished to replicate this study or the authors' methodology. Given the number of model frameworks used and the current paper length, it may be better to include specific model configu-

ration setting information in a supplement. This was also noted in the review forum by GMD editor David Ham.

P15L9: "Control" is too strong of a statement here. Suggest changing to distribution and intensity of modeled dust are sensitive to...

P16L8-10: This seems like an odd choice to me. Figures 5k&5L and 5Q&5R suggest these two treatments produce markedly different results.

P20L16-19: Dust concentration could be due to other factors outside of dust emissions (e.g., dispersion, mixing, deposition treatment, etc.). The causality statement here is too strong.

P2016-21: This discussion needs equations. See previous comment about emission scheme flow charts.

P24L1-11: Were these values tuned as well?

P25L7-8: These equations need to be provided. Is this error unique to this version of CMAQ? How is soil moisture integrated into the calculation of friction velocity threshold?

P26L6: Authors evaluated dust models, not dust emission schemes. There are too many free variables to isolate result outcomes to the dust emission schemes. Use of dust models here instead of dust emission schemes would make the language here consistent with the intro section.

P26L7: "four newly-introduced dust source maps in WRF-Chem" is a bit of an overstatement. The NU-WRF model some of these maps were obtained from is the NASA implementation of WRF-Chem.

P26L14-15: I agree with the authors that uncertainty associated with the U to u* conversion is likely an important source of error, potentially more so than minor differences in dust emission physics, but there are no discussions, figures or values presented in

this paper supporting this statement.

P26L24-25: Suggest changing to "All simulations performed best near the dust source areas and degraded in accuracy with downstream advection." This may indicate issues with transport, deposition, or forcing conditions if this feature is consistent across all model configurations. Were the WRF 3.9.1 meteorological fields assessed to ensure they captured the general storm evolution prior to being applied to the various dust models?

Typos: P17L15: Typo, might?