Dear referee reviewer #1:

Many thanks to your insightful comments and valuable suggestions on our manuscript "Multi-model simulations of springtime dust storms in East Asia: Implications of an evaluation of four commonly used air quality models (CMAQ v5.2.1, CAMx v6.50, CHIMERE v2017r4, and WRF-Chem v3.9.1)". After careful discussions with other co-authors, we have carefully revised our manuscript, and written this point-to-point response letter here.

The marked-up manuscript with revisions has been presented at the end of this file. The revised or added contents are listed as follows (words in red are the responses):

1. 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.

Response: We want to thank the reviewer for the constructive and insightful advice. We realize that a succinct summary of the dust emission scheme is needed in our manuscript. According to your suggestion, we further introduced the dust emission schemes (such as the algorithms of dust flux and relevant parameters) used in each air quality model in Section 2.4. The differences between the dust schemes are also described. The flow charts for all dust emission schemes including equations, relevant literature and required input parameters, as well as variable lists are provided in the Supplementary file (Fig. S1~S6 and Table S1~S6).

2. 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.

Response: Thanks your hard works and these valuable comments. We found that this sentence was not correctly expressed in this part of the manuscript. WRF v3.9.1 was used to generate the meteorological fields. The output was used to drive the air quality model of CHIMERE, CMAQ and CAMx. As to WRF-Chem, it is an inline model and dust emissions are calculated simultaneously with the atmospheric fields. Therefore, we cannot say that it is driven by the WRF meteorological fields. In the manuscript, it is revised as "The Weather Research and Forecasting (WRF) model version 3.9.1 was used to conduct the meteorological simulations, then to provide the hourly meteorological output fields to drive the air quality models of CHIMERE, CMAQ and CAMx while the chemistry module of WRF (WRF-Chem) was conducted simultaneously with the meteorological fields."

We don't use the coupled WRF-CMAQ implementation in this study. The output frequency of the WRF v.3.9.1 file for these four models is one hour.

Among these 4 air quality models, the u_* fields calculated by WRF is only used in WRF-Chem model for dust flux calculation while other 3 models implement u_* calculation independently. In CHIMERE v2017r4 model, u_* is calculated according to $u_* = \frac{\kappa u_{10}}{\ln\left(\frac{10}{z_0}\right)}$, where u_{10} is wind speed at 10m, z_0 is roughness length and κ is the Karman constant. In addition, the equation of Weibull distribution ($p(|u_{10}|) = \frac{k}{A}\left(\frac{|u_{10}|}{A}\right)^{k-1} e^{\left[-\left(\frac{|u_{10}|}{A}\right)^k\right]}$, where k is a dimensionless shape parameter and A is modeled wind speed, in meters per second.) is introduced for wind speed adjustment (Cakmur et al., 2004; Pryor et al., 2005). The friction velocity in CMAQ v5.2.1 was calculated based on an updated dynamic relation ($\frac{z_0}{h} = \begin{cases} 0.96\lambda^{1.07} & \lambda < 0.2 \\ 0.083\lambda^{-0.46} & \lambda \geq 0.2 \end{cases}$, where h is height of solid element, λ is total roughness density) to calculate the surface roughness length relevant to small-scale dust generation processes (Foroutan et al., 2017). The u_* in CAMx v6.50 is calculated according to the equation described in Louis (1979) expressed as $u_*^2 = \frac{\kappa^2 u^2}{\ln(z/z_0)^2} F_m(z/z_0, Ri_B)$, where F_m is a term as the function of Richardson number Ri_B. In addition, it also limits the minimum maximum value of the friction velocity to 0.4 m s^{-1}.

All the equations for u_* are presented in the flow charts of supplement file and the U_{10}, as well as u_* variations in each model during the dust episode are showed in Fig. 1 below (Fig. S11 in Supplementary Information). It reports that the u_* variations of WRF-Chem and CHIMERE are quite similar with averaged value of 0.60 m s^{-1}. In comparison, the u_* from CMAQ presents much lower with mean value of 0.41 m s^{-1}. It means that the introduction of a dynamic roughness length term in CMAQ results in lower friction velocity. This could be one of the reasons of the underestimation in CMAQ. As to the u_* in CAMx, it is the lowest among the values with mean value of 0.34 m s^{-1}because of the maximum limitation of 0.4 m s^{-1}, so that CAMx performed no dust emissions as the u_* was lower than u_{*t} during the episode.

The discussion about influence of different u_* is also presented in the Section of 3.6 in the revised manuscript.

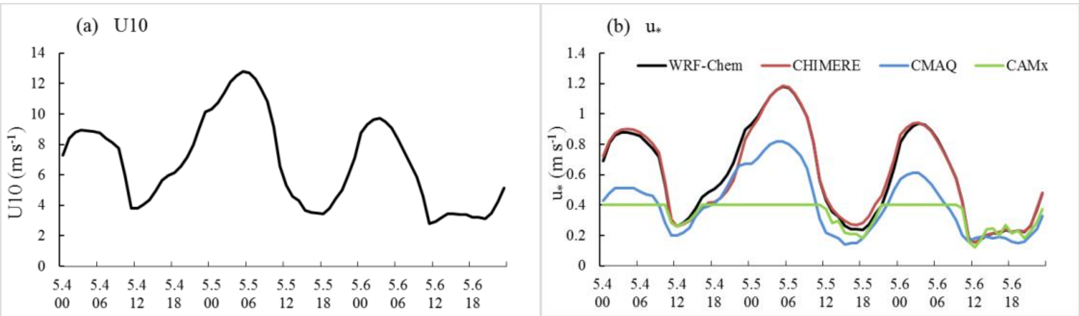


Figure 1. Time series of U10 (a), and u_* (b) from WRF-Chem, CHIMERE, CMAQ and CAMx in Changchun City during the dust episode.

References

Cakmur, R. V., R. L. Miller, and O. Torres: Incorporating the effect of small-scale circulations upon dust emission in an atmospheric general circulation model, J. Geophys. Res., 109, D07201, doi:10.1029/2003JD004067, 2004.

Foroutan, H., Young, J., Napelenok, S., Ran, L., Appel, K. W., Gilliam, R. C. and Pleim, J. E.: Development and evaluation of a physics‐based windblown dust emission scheme implemented in the CMAQ modeling system, J. Adv. Model. Earth Syst., 9(1), 585‐608, doi:10.1002/2016MS000823, 2017.

Louis, J. F.: A parametric model of vertical eddy fluxes in the atmosphere. Boundary-Layer Meteorology, 17(2), 187-202, doi: 10.1007/bf00117978, 1979.

Pryor, S., J. Schoof, and R. Barthelmie: Empirical downscaling of wind speed probability distributions, J. Geophys. Res., 110, D19109, doi: 10.1029/2005JD005899, 2005.

3. 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.

Response: Thank you very much for your helpful suggestion. This question was also mentioned by reviewer #2. After discussion with all authors, we had implemented the CAMx model for further simulations. For the implementation of the dust emission scheme in CAMx, we select the seasonal dust source map (G12_0.1_seasonal in the manuscript) to replace the original dust mask file as this dust source map had the best performance among those source maps in the WRF-Chem model. The values in source map file were changed to 1 when the erodible fraction > 0 to fit the format of the dust mask file. Similarly, the performance of the CAMx dust simulation during the dust episode is analyzed and evaluated, and the result is showed in Section 3.5 of the manuscript. The daily averaged PM_{10} distribution on May 5th, 2015 is presented in Fig. 2b. It shows that the daily PM_{10} concentration simulated by CAMx ranged from 0 to 30 μg m^{-3} with high value area in the southwest part of the simulated domain, and there was no dust emitting from any erodible area in NEC. A control simulation without dust emission was also conducted and the PM_{10} pattern is same with Fig. 2b. It means that no dust emission at all and CAMx model failed to reproduce this dust episode occurred in NEC.

Considering the dust mask had been changed and the erodible areas were included in model, the failed simulation of CAMx might result from the lower value of friction velocity. In the dust model of CAMx, the friction velocity is limited to a maximum value of 0.4 m s^{-1}, making it keep a low level comparing to the values of other models (Fig. 1). It was difficult to exceed u_{*t} which was generally larger than 0.4 m s^{-1} (Fig. 5), so no dust emission occurred. Therefore, this limitation value was subsequently removed and the simulation was conducted again. The distribution of simulated PM_{10} without the u_* limitation was presented in Fig. 9c. It shows that the dust was mainly from western Jilin Province near the Songnen sandy land and transported westward. This pattern could be also observed from ground observations (Fig. 2e in manuscript). However, there was no dust emitting from Horqin sandy land. Simulated PM_{10} concentrations were generally lower than the observations with about 120 μg m^{-3} in source areas and 10~50 μg m^{-3} in the transported areas. Compared with the simulation with u_* limitation, this result was obviously improved which indicated that the limitation value of u_* in CAMx needs further adjustment to improve its performance over the areas other than barren and sparsely vegetated area.

The analysis and discussion on the CAMx result were in Section 3.5 and 3.6 of the manuscript.

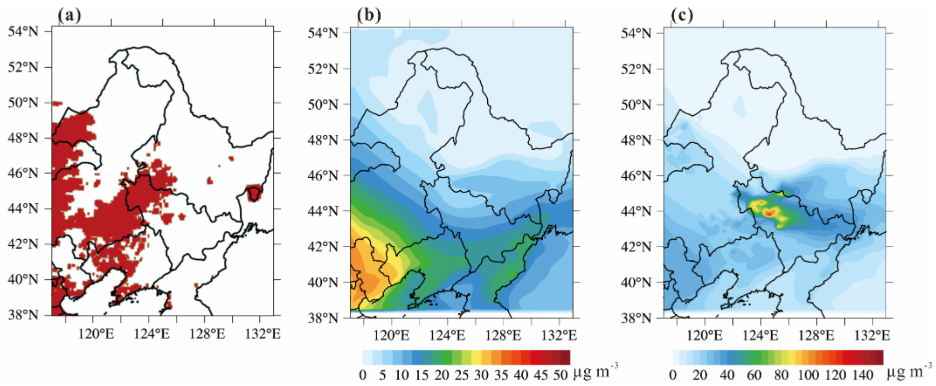


Figure 2. The substituted dusk mask (a) and daily mean PM_{10} distributions in NEC on May 5th, 2015 (b) using CAMx model.

4. 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.

Response: Thank you very much for your helpful comment. When we were preparing the meteorological input files, we firstly evaluated the simulation performances of WRF surface windspeed between the LSM of PX and Noah. The result showed they had same correlation coefficient (0.8) and close RMSEs (1.52 m/s for Noah-MP scheme and 1.61 m/s for Pleim-Xiu scheme). These comparisons showed close results between two schemes, however, the errors of Noah scheme had larger standard deviation showing higher dispersion than PX scheme. Considering this and the previous study by Zhang et al. (2015) which showed a good performance of PX scheme in the same research area, we finally choose PX scheme.

Later, according to this comment, we further conducted the dust emission simulation with the LSM of Noah and find that the simulated dust concentrations are much lower than those derived from PX scheme (Fig. 3). We further compared u_*, U10 and surface soil moisture calculated via PX and Noah scheme and the temporal variations of them are provided in Fig. 3. It shows that the variation curves of u_* calculated by PX and Noah scheme are quite similar. It does not present a stronger U-u_* conversion in the PX/ACM2 setting than in Noah scheme over the research area at this time. By contrast, the Noah surface soil moisture shows larger difference, with values 93.6% higher in Changchun City and 29.6% higher in the NEC area (Fig. 4 and Table 1). Moreover, the soil moisture curve with two LSM schemes are quite different. These discrepancies may result in the differences of estimated dust emissions. The lower soil moisture (which makes smaller threshold friction velocity) simulated by using PX scheme could be the reason of the stronger dust emission magnitude from PX compared to Noah LSM scheme. However, the dust emission simulated by AFWA scheme is considerably higher than that by GOCART no matter which LSM is used. Therefore, we think the differences of wind velocity and soil moisture between PX and Noah scheme could lead to the dust emission discrepancies, but it is hard to explain the differences of dust emission magnitude between

GOCART and AFWA scheme.

Many researches indicate higher dust emission simulated by GOCART than that of AFWA scheme over the areas like Mediterranean, West Asia and southwest Asia (Flaounas et al., 2017; Nabavi et al., 2017; LeGrand et al., 2019) which are quite different from our study. So we try to find out the reasons for large discrepancies between different dust schemes under the same meteorological condition. The saltation bin configuration which influences the dust mass distributions described in LeGrand et al. (2019) maybe have impact on the dust emission and concentration. One another explanation of the over-prediction of the AFWA dust concentration is that AFWA scheme considers vertical dust flux only related to the clay content which results in a higher vertical-to-horizontal dust flux ratio (Kang et al., 2010; Rizza et al., 2016; Rizza et al., 2017). We have added this into Section 3.2 of our new revised manuscript. Meanwhile, we also add the discussion about the LSM influence on the dust simulation in the same part.

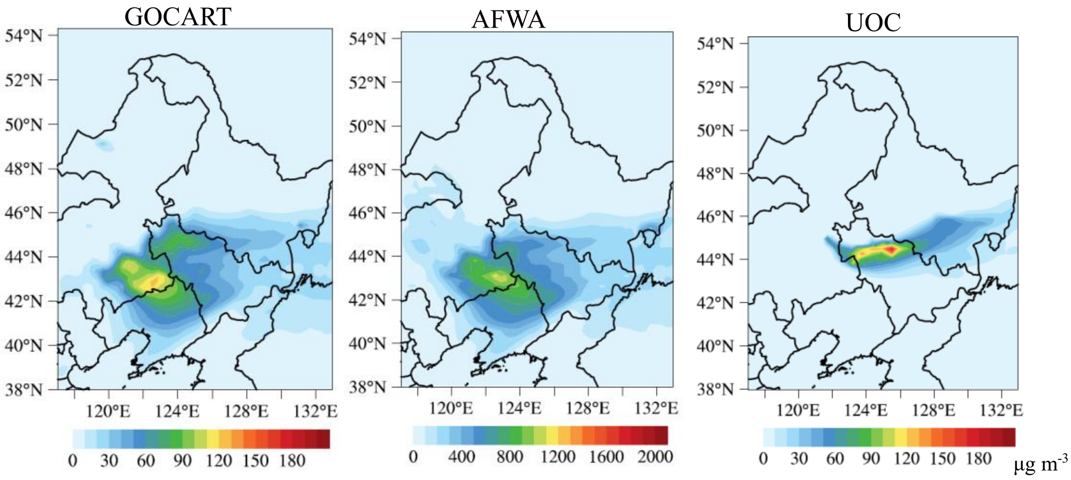


Figure 3. Daily mean PM_{10} distributions in NEC on May 5th, 2015 using GOCART, AFWA and UOC_Shao2004 with LSM of Noah.

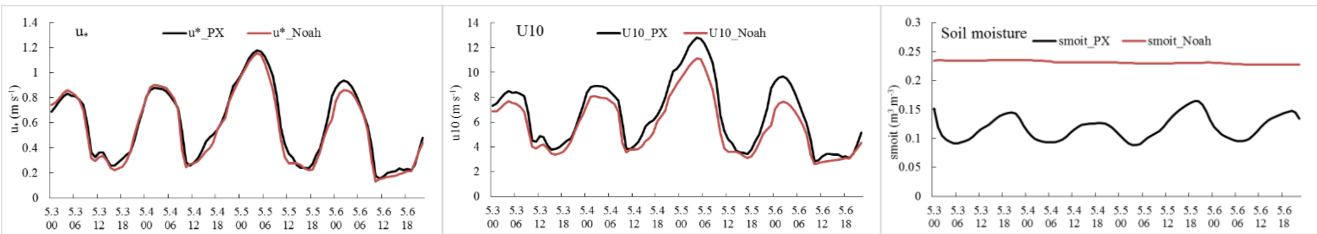


Figure 4. Time series of u_*, U10 and surface soil moisture simulated via LSM of PX and Noah in Changchun City during the dust episode.

Table 1. Mean u_*, U10 and surface soil moisture simulated via LSM of PX and Noah in the research area of NEC

	PX	NOAH
U10 (m s^{-1})	5.10	4.66
u_* (m s^{-1})	0.51	0.46
soil moisture (m^3 m^{-3})	0.27	0.35

Reference

Flaounas, E., Kotroni, V., Lagouvardos, K., Klose, M., Flamant, C., and Giannaros, T. M.: Sensitivity of the WRF-Chem (V3.6.1) model to different dust emission parametrisation: assessment in the broader Mediterranean region, Geosci. Model

Dev., 10, 2925-2945, https://doi.org/10.5194/gmd-10-2925-2017, 2017.

Kang, J., Yoon, S., Shao, Y. and Kim, S.: Comparison of vertical dust flux by implementing three dust emission schemes in WRF/Chem, J. Geophys. Res. Atmos., 116(D9), doi:10.1029/2010JD014649, 2011.

LeGrand, S. L., Polashenski, C., Letcher, T. W., Creighton, G. A., Peckham, S. E., and Cetola, J. D.: The AFWA dust emission scheme for the GOCART aerosol model in WRF-Chem v3.8.1, Geosci. Model Dev., 12, 131-166, https://doi.org/10.5194/gmd-12-131-2019, 2019.

Nabavi, S. O., Haimberger, L. and Samimi, C.: Sensitivity of WRF-chem predictions to dust source function specification in West Asia, Aeolian Res., 24, 115–131, doi:10.1016/j.aeolia.2016.12.005, 2017.

Rizza, U., Barnaba, F., Miglietta, M. M., Mangia, C., Di Liberto, L., Dionisi, D., Costabile, F., Grasso, F. and Gobbi, G. P.: WRF-Chem model simulations of a dust outbreak over the central Mediterranean and comparison with multi-sensor desert dust observations, Atmos. Chem. Phys., 17(1), 93, doi:10.5194/acp-17-93-2017, 2017.

Rizza, U., Anabor, V., Mangia, C., Miglietta, M. M., Degrazia, G. A., and Passerini, G.: WRF-Chem Simulation of a saharan dust outbreak over the mediterranean regions, Vol. 38, Special Edition, 330–336, Ciência e Natura, doi:10.5902/2179460X20249, 2016.

Zhang, X., Zhou, Q., Chen, W., Wang, Y. and Tong, D. Q.: Observation and modeling of black soil wind-blown erosion from cropland in Northeastern China, Aeolian Res., 19, 153–162, doi:10.1016/j.aeolia.2015.07.009, 2015.

5. 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 average PM_{10} distributions. Simulated PM_{10} 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 PM_{10} 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.

Response: Thanks for your valuable comment. In this study, we focus on evaluating the performances of different dust models. Compared their spatial-temporal distributions and calculated the correlations, biases and errors with observational data. For this purpose, hourly ground-based monitoring PM data were used. In Section 3.2~3.5, the distributions of daily mean PM_{10} concentration were presented, in order to provide overall simulation situations (such as dust coverage and concentration level) of four air quality models. Preliminarily evaluated the dust performance of each model. Then the inter-model comparisons were conducted basing on the hourly data. We conclude the relative advantage and accuracy of dust performances of the air quality models. The error analysis and simulation assessment are both focus on the dust emission models. And indeed, the influence of atmospheric conditions and wet/dry deposition on the dust simulation are not conducted in this study, and we have mentioned this in the last paragraph of Section 3.6 in the manuscript. Thank you again for the helpful suggestion and the study on the impacts of atmospheric condition and deposition will be conducted in the future.

6. Section 3.5: I don't understand the rational for scaling PM_{10} concentrations in the inter-model comparisons. Why scale the simulation output rather than the scaling the emission fluxes?

Response: Thanks again. The PM_{10} concentrations simulated by different models varies from 10^0 to 10^4 ug/m^3 which means considerable discrepancies. The bias and error between simulated and observed PM_{10} concentration differ widely as well. It is very hard to plot all of the simulated concentrations or validation results in one figure and inconvenience to conduct the evaluation on all of the simulations at the same time. Therefore, we use scaling factor to adjust the outputs of WRF-Chem and CMAQ (as well as CAMx) to make them have similar concentration level with others. The scaling factor could help us to better understand the advantages and disadvantages of the model performance on dust simulation, rather than improve the accuracy of dust modeling at this time.

The evaluation in this study indicates tuning factor for dust flux is needed to improve model performance and puts forward an improving direction of dust simulation in present air quality models. Brief discussion and explanation about this are added in the parts of comparison and summary. The further work on tuning and localization of the dust models will be conducted basing on this study.

7. 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).

Response: Thank you very much for your helpful suggestion. We have calculated the FVT in each dust emission model according to the algorithms described in relevant references and source codes in models. The FVT variation used in each dust emission model during the dust episode is showed in Fig. 5. We have added this into Supplementary Information and also have discussion about its influence to dust emission in Section 3.6 of the manuscript.

From the figure we find that the overestimation of WRF-Chem AFWA could in part be explained by the lower FVT comparing to those used in other dust schemes. The variation of UOC FVT fluctuated widely and presented highest FVT peaks, but during the dust episode (5 May, 2015) it kept lower values making it in the middle level of FVT. This may be part of the reason of the overestimated PM_{10} concentration. However, the FVT of GOCART has the lowest value while its simulated dust emission is in the lower level. As we know, the FVT is very important in the dust emission calculation, however, it needs further adjustment and improvement as the air quality at present have difficulty in calculating FVT properly in some areas such as northeastern China. And this is also our next step work to implement field study and measurement of FVT, adjust and localize the FVT algorithm and the relevant parameters.

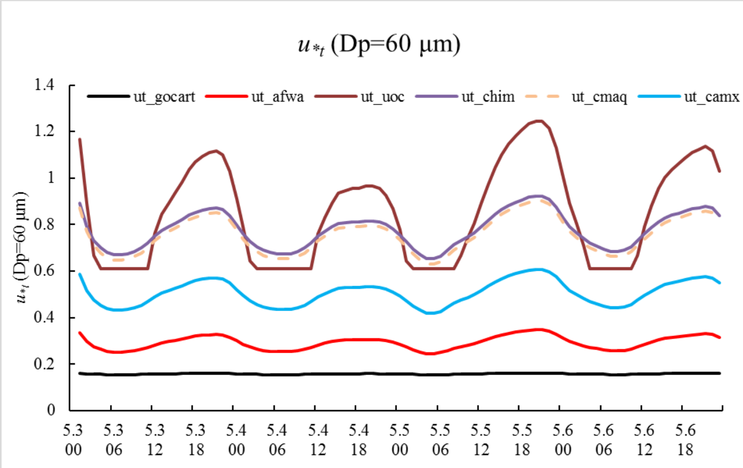


Figure 5. Time series of u_{*t} in from dust emission model of GOCART WRF-Chem, AFWA WRF-Chem, UOC WRF-Chem, CHIMERE, CMAQ and CAMx in Changchun City during the dust episode.

Minor comments:

1. 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.

Response: Thank you for the helpful comment. This study is not the first comprehensive evaluation of dust models for East Asia, but we are sure it is the first evaluation of the dust modules in air quality models (rather than climate model (Uno et al., 2006), global model et al.). In order to keep in rigorous, this sentence is now revised as "Here we present a comprehensive evaluation of multi-model simulations of windblown dust emissions in air quality models during a dust episode in East Asia…" Now we are working on the extended evaluations of the dust emissions in air quality models, such as the evaluation on the performance of CHIMERE during an autumnal dust episode (Ma et al., 2019).

Reference

Ma, S., Zhang, X., Gao, C., Tong Q., Xiu, A., Zhao, H and Zhang, S. Simulating Performance of CHIMERE on a Late Autumnal Dust Storm over Northern China[J]. Sustainability, 11(4). doi: 10.3390/su11041074, 2019.

Uno, I., Wang, Z., M. Chiba, Y. S. Chun, Sun Ling Gong, Yukari Hara, E. Jung et al. "Dust model intercomparison (DMIP) study over Asia: Overview." Journal of Geophysical Research: Atmospheres 111, no. D12 (2006).

2. 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.

Response: It had been revised as "The AFWA dust scheme is a modified version of Marticorena and Bergametti (1995) dust scheme developed by the Air Force Weather Agency (LeGrand et al., 2019)."

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

Response: The error in UOC source code has been corrected before we conduct the simulations in this study. But it seems that we didn't make it clearly expressed. After revising, it is now expressed as "Note that the last term in the saltation flux formula in UOC source code is expressed as $(1+(\frac{u_{*t}}{u_*})^2)$ by error in WRF-Chem before the version of 4.0. In this study, it has been changed into $(1+\frac{u_{*t}}{u_*})^2$ in WRF-Chem version 3.9.1 according to the description in Shao et al. (2011)."

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

Response: The explanation of the way we name the dust source map is provided into manuscript and it is described as "In this study, we named this source map as G01_0.25 according to first author and published year of relevant literature and its spatial resolution.".

5. 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.

Response: The basic information and name of each source map are now introduced in the manuscript and showed as "The source map with resolution of 1°×1° developed by Ginoux et al. (2001) (namely G01_1.0) basing on the same method with G01_0.25 was obtained from the homepage of Dr. Paul Ginoux (https://www.gfdl.noaa.gov/pag-homepage). Except G01_0.25 and G01_1.0, the other source maps were obtained using satellite observations. The map with resolution of 0.25°×0.25° provided by Koven and Fung (2008) (K08_0.25 hereafter)", "The source map with only natural origins was named G12_0.1_natural while that with both natural and anthropogenic origins was named G12_0.1_ant+nat." and "The dust source map obtained from the NASA-Unified WRF (NU-WRF) version 7 (Kim et al., 2014) was similar to the former source map but divided into four seasons, therefore it was named as G12_0.1_seasonal in this study".

6. 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)?

Response: Actually, it is in-situ measurements but not the ground survey. In spring of 2013, we conducted a field in-situ observation for wind erosions over the cropland experimental station (44° 12′ N, 125° 33′ E) with black soil in Northeastern China. The SENSIT sensor (Model H11-LIN, Sensit, Co.) and wind velocity sensors (5 heights, 0.2m, 0.5m, 1.0m, 2.0m and 5.0m, model 010C, Met One Instruments Inc.)were adapted for observing the threshold friction velocity and the comment value is 0.37m/s at the surface ground.

Reference

Zhang, X., Zhou, Q., Chen, W., Wang, Y. and Tong, D. Q.: Observation and modeling of black soil wind-blown erosion from cropland in Northeastern China, Aeolian Research, 19, 153-162, doi:10.1016/j.aeolia.2015.07.009, 2015.

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

Response: Now it is described as "The dust source map obtained from the NASA-Unified WRF (NU-WRF) version 7".

8. 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?

Response: The file of source map G01_1.0 with format of NetCDF was obtained from the homepage of Dr. Paul Ginoux (https://www.gfdl.noaa.gov/pag-homepage). And converted to ASCII format (.bin) by using write_geogrid.c in WPS. We find that the erodible fraction distribution of it shows more reasonable than the default one in WRF-Chem v3.9.1, therefore, it is introduced and the evaluation of its performance is conducted in this study. The data source is now given in the manuscript.

9. P14L1: Please describe the BELD3 dataset.

Response: The BELD3 dataset is explained in the manuscript as "when the Biogenic Emissions Landcover Database version 3 (BELD3) dataset was used during regional simulation in the USA."

10. 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.

Response: The USGS and MODIS land use datasets in this study are the default land use and soil category datasets which provided as part of the WPS static data file. They were obtained from http://www2.mmm.ucar.edu/wrf/src/wps_files/. These default data were remapped via the geogrid program of WPS. We find that the content in P14L3 could not be clearly expressed, so it is revised to "However, the grassland fraction was not taken into account in when using the USGS or MODIS land use dataset according to the source code of CMAQ, which may lead to an underestimation of dust emission." The selections of land use data are described in Section 2.5 and detailed information provided in the Supplementary Information.

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

Response: Thanks for your kindly suggestion. According to the advice of Editor David Ham. We had upload all the model source codes that we used, the namelist files, the sources of the input data (such as the meteorological reanalysis data and erodible fraction data), as well as pre- and post- processing script we used to the Zenodo website, and we had added the link of https://doi.org/10.5281/zenodo.3376774 in our new revised manuscript.

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

Response: This has been revised to "As mentioned above, we found that the distribution and intensity of modeled dust aerosols were sensitive to the dust source maps in use."

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

Response: Although the dust source map of G12_0.1_ant+nat and G12_0.1_seasonal (Fig. 5k&5L and 5Q&5R in manuscript) were obtained via the same methodology, G12_0.1_seasonal is divided into four seasons which makes it able to present the temporal variation of the erodible fraction. It is more reasonable than the constant field (G12_0.1_ant+nat) and makes obvious different between the distribution of G12_0.1_ant+nat and G12_0.1_seasonal (Fig. 3e&3f). Therefore, G12_0.1_seasonal is chosen for the next step in the evaluation and the reason is explained in the manuscript as "Considering that the source maps of G12_0.1_ant+nat and G12_0.1_seasonal were obtained via the same methodology but the latter one provides seasonal divisions making it more reasonable and closer to the actual environment, G12_0.1_seasonal was chosen for the next step in the evaluation."

14. 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.

Response: This sentence is revised to "When considering that the dominant soil textures were loam and clay loam in NEC, this explained the reason for the underestimation occurred in CMAQ compared with WRF-Chem and CHIMERE."

15. P20L16-21: This discussion needs equations. See previous comment about emission scheme flow charts.

Response: The equations for calculating α (sandblasting efficiency) are provided in the flow chart of Supplementary Information and we have explained about this in the manuscript.

16. P24L1-11: Were these values tuned as well?

Response: Yes, the values in this part (discussion of Fig. 12 in manuscript) are also tuned. The comparison and analysis in Section 3.6 are both based on the tuned outputs of the air quality model. The explanation about this is implemented in the beginning of Section 3.6

17. 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?

Response: This part focuses on the discussion of CHIMERE output. The soil moisture is used to calculate of soil moisture correction, and then multiplies the dry/smooth threshold friction velocity to obtain the actual threshold friction velocity. The equation of soil moisture correction is showed as

$$H_w(w) = \begin{cases} \sqrt{1 + 1.21(w_g - w_g')^{0.68}} & w \geq w' \\ 1 & w < w' \end{cases}$$

where w_g is gravimetric soil moisture, w_g' is moisture without effect on capillary forces. The unit conversion from volumetric soil moisture w_v (m^3 m^{-3}) to gravimetric soil moisture w_g (kg kg^{-1}) are different in different models. Such as $w_g = \frac{w_v \rho_w}{(2.65 - 0.15clay)(1 - porosity)}$ in WRF-Chem v3.9.1, and $w_g = \frac{w_v \rho_w}{(2650 \times (0.511 + 0.126 \times (soiltxt1 + soiltxt2)}$ in CMAQ v5.2.1. However, the conversion $w_g = 100 \times w_v$ in CHIMERE v2017r4 is obviously unreasonable. Later, we also find this error exist in the previous version of CHIMERE (such as CHIMERE v2016a1). The equations are provided in the flow charts in Supplementary Information.

18. 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.

Response: It is revised to "In this study, we quantitatively evaluated the performance of several physically-based dust emission models in…."

19. 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.

Response: Thanks for your valuable comment. We find that these statements are inaccurate here, therefour we revise it as "Four dust schemes and four additionally-introduced dust source maps in WRF-Chem v3.9.1…"

20. 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.

Response: Thank you very much for your helpful comment. We found that we omitted the analyses about u_* in the discussion part. On the basis of the results in the second major comment, we compared the u_* in different model in P28L22 and P29L14 when analyzing the results of CHIMERE and CMAQ. It shows as "Moreover, different algorithms for u_{*t} resulted in significant differences in the simulated dust emission. For instance, the variations of u_* in CHIMERE and WRF-Chem are similar (Fig. S11a, b) during the dust episode, however, u_{*t} presented large discrepancies from one model to another (Fig. S12). The value of u_{*t} in CHIMERE is in the upper level of the six kinds of u_{*t} used in the dust emission models while that of AFWA is much lower. This could also be one of the reasons of the overestimation in WRF-Chem AFWA." and "In addition, using the dynamic roughness length term when calculating u_* in CMAQ led to lower u_* (with mean value of 0.39 m s^{-1}, Fig. S11c) comparing to those in WRF-Chem and CHIMERE (0.58 m s^{-1}). This would be another reason for its underestimated results.", respectively. The u_* algorithm in CAMx was discussed in Section 3.5 of the manuscript (and in major comment 4). The figure of u_* variations in each air quality model is provided in Supplementary Information (Fig. S11).

21. 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?

Response: Thanks for your helpful suggestion. It is revised to "All simulations performed best near the dust source areas and degraded in accuracy with downstream advection." The meteorological fields of WRF v3.9.1 has not been test in this research area before, but we have evaluation the surface wind simulated by it in this study and it presented acceptable result (Section 2 in Supplementary Information).

22. Typos: P17L15: Typo, might?

Response: It has revised as "This difference might have arisen because the KOK scheme was mainly built on fragmentation theory"

In addition, besides the major and minor questions as mentioned above, we conducted reproducibility tests of our simulations in this study, and found an error in the WRF-Chem section. We have run the WRF-Chem model with UOC_Shao2011 scheme and it showed that PM_{10} simulated by UOC_Shao2011 presented little similarity with the observations (Fig. 6). Considering its unreasonable results, then we selected UOC_Shao2004 for the subsequently simulations. It means that the used dust scheme in original manuscript is UOC_Shao2004. However, the texts expressed the used scheme as UOC_Shao2011 by wrong. Therefore, we corrected this into UOC_Shao2004 in the manuscript.

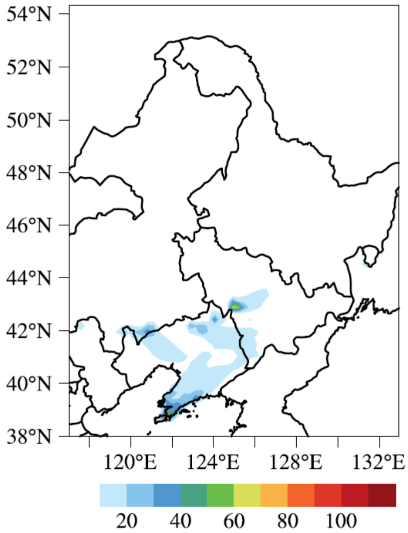


Figure 6. Daily mean PM_{10} distributions in NEC on May 5th, 2015 using WRF-Chem UOC_Shao2011.

We have tried our best to improve the manuscript and made changes in the manuscript. These changes will not influence the content and framework of the paper. And here we do not list the changes but marked in the revised manuscript. We thank the reviewer again for the constructive advice that have helped us to improve our manuscript.

All in a word, via these evaluation works, we hope to do some contributions to the community for enhance the dust forecast ability on regional scale in air quality models.

Dear referee reviewer #2:

We are very grateful for your constructive comments and suggestions on our manuscript "Multi-model simulations of springtime dust storms in East Asia: Implications of an evaluation of four commonly used air quality models (CMAQ v5.2.1, CAMx v6.50, CHIMERE v2017r4, and WRF-Chem v3.9.1)". After carefully discussions with other co-authors, we have revised our manuscript according to this point-to-point response letter.

The marked-up manuscript with revisions has been presented at the end of this file. The point to point responses are listed as follows:

1. Title. To better reflect the content of the article, I suggest to change the title to "Multi-model simulations of a springtime dust storm over Northeast China: Implications of an evaluation of three commonly used air quality models (CMAQ v5.2.1, CHIMERE v2017r4, and WRF-Chem v3.9.1)" due to the following reasons: 1. Only one dust storm is considered. 2. Referring to East Asia is misleading as only a comparably small subregion of East Asia is considered. East Asia in contrast comprises two of the earth's major deserts, Taklamakan and Gobi, which are not at all subject of the study. 3. The evaluation of CAMx is limited to identifying that practically no emissions can be produced from within the model domain due to the MODIS based desert mask applied in the emission scheme which precludes emissions from regions that are not barren or sparsely vegetated. While this is an important conclusion, no further evaluation of CAMx is presented and thus the present title is misleading. Considering that some efforts where made to adjust the emissions of other models, it would have been interesting to see results from CAMx after expanding the mask to include other landcover types, but I understand that this might be beyond the scope of the study. In that case I recommend to simply adjust the title, to clarify in the abstract that three models are evaluated and (as before) to discuss in the main text why CAMx is not one of them.

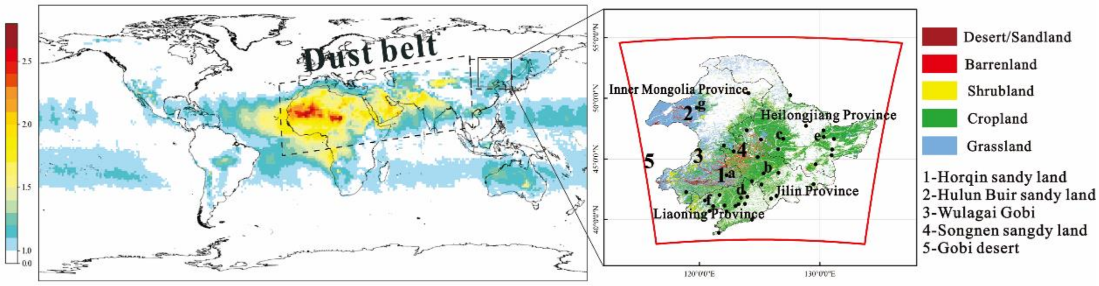
Response: We want to thank the reviewer for the constructive and insightful advice. According this comment and the suggestion from Reviewer 1. The further simulation and analysis of CAMx were implemented, and provided in Section 3.5 and the supplement file of the manuscript. As the dust mask used in CAMx showed no coverage in NEC area, the seasonal dust source map (G12_0.1_seasonal) was adapted to replace the original dust mask file as it had the best performance among those source maps in the WRF-Chem model. Therefore, the title is changed to "Multi-model simulations of a springtime dust storm over Northeastern China: Implications of an evaluation of four commonly used air quality models (CMAQ v5.2.1, CAMx v6.50, CHIMERE v2017r4, and WRF-Chem v3.9.1)"

2. Page 1, line 27. "to simulate dust storms in East Asia" should read "to simulate a dust storm over Northeast China", see above

Response: It is revised as "This study applies and evaluates four widely used regional air quality models to simulate dust storms in Northeastern China."

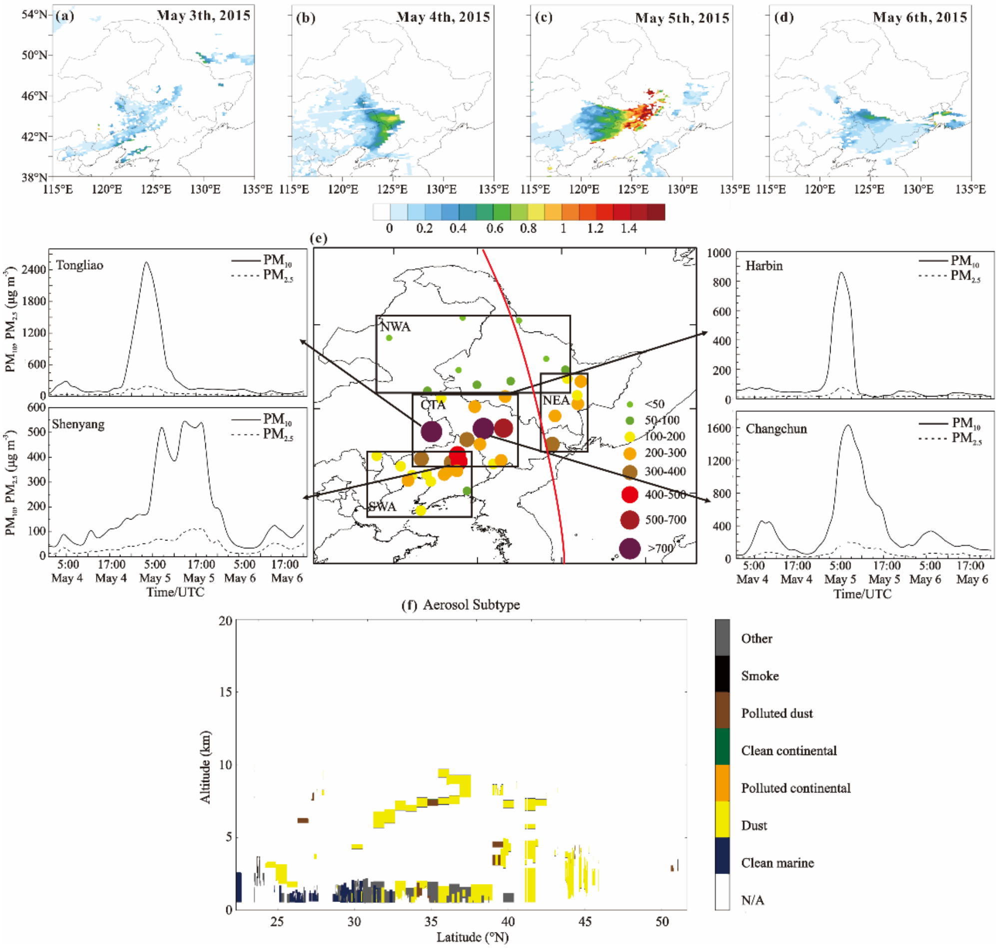
3. Page 5, line 8. The line in Fig. 1 is hardly recognizable as being dark blue.

Response: Thank you for your reminding. We find that the color blue didn't present well when converting to PDF format. So we changed it to color red and the revised figure is showed below.



4. Page 7, Fig. 1. Please label the regions with CTA, NWA, NEA and SWA. The CALIPSO path can hardly be identified as being blue as mentioned in the caption.

Response: The labels of subareas of CTA, NWA, NEA and SWA were added on the figure and the CALIPSO path color was changed to red. The revised figure is showed below.



5. Page 8, line 20. "would" or "did"?

Response: It has revised to "This revision could increase the saltation flux by a factor of 2 or more." According to the description in LeGrand et al. (2019).

Refference

LeGrand, S. L., Polashenski, C., Letcher, T. W., Creighton, G. A., Peckham, S. E., and Cetola, J. D.: The AFWA dust emission scheme for the GOCART aerosol model in WRF-Chem v3.8.1, Geosci. Model Dev., 12, 131-166, https://doi.org/10.5194/gmd-12-131-2019, 2019.

6. Page 10, line 11. "omitting the effect of soil moisture" should read "omitting the term supposed to account for the effect of soil moisture"

Response: Thank you for your very helpful suggestion. This sentence is revised as "The major modifications were omitting the term supposed to account for the effect of soil moisture on dust emission".

7. Page 10, line 13. "maximum", not "minimum"

Response: Thanks for your comment, it should be "maximum" and had been revised in our new manuscript.

8. Page 17, line 5. In Figs. 2 and 6 it is not possible to identify trajectories, not even the approximate direction of the outflow from the source regions can be identified in Fig. 2, it is therefore hard to make this comparison. Neither do I expect any difference, as the WRF wind fields should be quite realistic.

Response: Thank you very much for pointing out this problem. We find that this part could not properly expressed. According to the observations, the large areas of NEC such as northern Liaoning, Jilin and eastern Heilongjiang Province were influenced by this dust episode while the simulated results of CHIMERE did not show an obvious impact on northeastern NEC (eastern Heilongjiang Province). This is one of the differences between CHIMERE simulation and observation. It should be simulated and observed patterns rather than trajectories. Therefore, we revised this sentence to "The simulated dust showed its impact on the eastern areas like Jilin and northern Liaoning Province (Fig. 6a~c), while northeastern NEC (such as eastern part of Heilongjiang Province) were also observed to be influenced by this dust episode (Fig. 2).".

9. Page 17, line 11. "The most striking feature of the model results was their concentration" should read "The most striking discrepancy between the model results was in their concentration level" or similar.

Response: It has revised to "The most striking discrepancy between the model results was in their concentration level.".

10. Page 17, line 15. "might" should be deleted

Response: Thanks again. This is revised to "This difference might have arisen because the KOK scheme was mainly built on fragmentation theory".

11. Page 18, line 7. The sentence "With further comparison ..." needs rephrasing

Response: It is revised as "Comparing to observations and simulated results of WRF-Chem and CHIMERE, the dust simulated by CMAQ was only short-distance transported southeastwards to….".

12. Page 19, Eqs. (1) and (2) and Section 3.4 in general. Please define all variables and make sure to use units (e.g. for the p limits in Eq. (2)). The discussion would benefit from some revision because it is hard to follow what is used in (a) the different models (b) the literature cited and (c) in the present study. E.g., it would help to mention both, model name and the related citation next to each other where applicable and make use of active voice.

Response: The definitions of variables and parameters used in equations were explained more detailed in the manuscript. Moreover, we also named the methods described in literature or used in models: the formula described in Lu and Shao (1999) is named as LS99 and a version of LS99 modified by Kang et al. (2011) and introduced in CMAQ since version v5.2 by Foroutan et al. (2017) is called F17. The formula involving p for calculating α described in Shao (2004) is named as S04. This part is now revised as "The formula is expressed as follows:

$$\alpha = \frac{F}{Q} = \frac{C_\alpha g f \rho_b}{2p}(0.24 + C_\beta u_* \sqrt{\frac{\rho_p}{p}}) \quad (1)$$

where f is the fraction of fine particles contained in the soil volume, p is plastic pressure, in the range of 10^3~10^7 N m^{-2} (Gillett, 1977; Callebaut et al., 1985; Rice et al., 1997), ρ_b and ρ_p are the bulk soil and soil particle densities with unit of kg m^{-3}, g is the gravitational constant in m s^{-2}, u_* is friction velocity in m s^{-1}, and C_α and C_β are constants. Here the formula described in Lu and Shao (1999) is named as LS99 and a version of LS99 modified by Kang et al. (2011) and introduced in CMAQ since version 5.2 by Foroutan et al. (2017) is called F17. The formula involving p for calculating α according to Shao (2004), namely S04, can be described as:

$$\alpha = c_y \eta_{f,i}[(1-\gamma) + \gamma \frac{p_m(d_i)}{p_f(d_i)}] \frac{g}{u_*^2}(1 + 12u_*^2 \frac{\rho_b}{p}(1 + 14u_* \sqrt{\frac{\rho_b}{p}})) \quad (2)$$

Where $p_m(d_i)$ and $p_f(d_i)$ are respectively the fully and minimally disturbed dust fraction in bin d_i, and $\eta_{f,i}$ is the fully disturbed dust fraction. ρ_b=1000 kg m^{-3} is bulk soil density. γ is a function specified as $\gamma = \exp[-(u_* - u_{*t})^3]$ where u_{*t} is threshold friction velocity. c_y is a dimensionless coefficient which is set to be 1×10^{-5}, 4×10^{-5}, 5×10^{-5}, 3×10^{-4} for different soil textures and locations in Shao (2004); then, values of soil plastic pressure p in the range of 10^2 to 10^4 N m^{-2} were obtained via matching with observed dust flux and friction velocities. This formula is now used in WRF-Chem v3.9.1."

13. Page 19, Eq. (2). Unless u_* is about 1, the RHS has a discontinuity at p = 3e4 N m^{-2}. The two cases on the RHS are limiting expressions for large and small p, it seems to be problematic to apply them on adjacent p intervals, and not use the full expression for intermediate values of p. Where does this distinction of cases and the threshold of p = 3e4 N m^{-2} come from?

Response: The threshold of p = 3e+5 N m^{-2} (however, it was miswritten as 3e+4 N m^{-2} in the manuscript) was from the description of Shao et al. (2004) "For p >3e+5 N m^{-2}, σ_m becomes negligibly small (<0.1) under normal wind conditions, implying that saltation bombardment is insignificant in such circumstances and aggregates disintegration is the main mechanism for dust emission." Nevertheless, the reference doesn't clearly present the threshold value of the equation and no explanation about how the latter part of the equation ($\alpha = 168c_y[\eta_{mi} + (1-\gamma)\eta_{ci}]\left[\frac{1000}{p}\right]^{\frac{3}{2}} u_* g$) established is provided, and it has not been used in the air quality model. We read the reference again and found that equation 2 in the manuscript was improperly introduced. The equation 6 in Shao et al. (2004) which showed as

$$\alpha = c_y \eta_{f,i}[(1-\gamma) + \gamma \frac{p_m(d_i)}{p_f(d_i)}] \frac{g}{u_*^2}(1 + 12u_*^2 \frac{\rho_b}{p}(1 + 14u_* \sqrt{\frac{\rho_b}{p}}))$$

is used in WRF-Chem model and it doesn't need piecewise p values for calculation. Therefore, this part was revised and the discussion about the threshold of p is removed. Now it is shows as "Note that the fitted c_y and p defined above could only be used in S04 and not in LS99 and F17 with different physical parameters. For example, the fitted value of 5000 for p (silty clay loam) in Table 3 of Shao (2004) was used as p of sand in Kang et al. (2011). To correct the overestimated p used in the vertical flux calculation of LS99, Kang et al. (2011) reported that a modified C_α was recalculated based upon c_y (which is used in S04). However, to our knowledge, no method based on physical evidence is available to complete this conversion. Moreover, the source code of Shao_2004 in WRF-Chem only uses prescribed values $p = 3 \times 10^4$ and $c_y = 1 \times 10^{-5}$

without considering the soil textures. As both of their values varied widely over soil types and locations, the mismatch in part of the study domain would lead to difference in magnitude, no matter in CMAQ or WRF-Chem."

14. Page 21, line 13 to 15. Please mention that you use the AGO scheme

Response: Thank you for your helpful suggestion. It is revised as "here only the outputs simulated by AGO scheme with *ierod*=3 (mixed USGS and MODIS) were chosen for further validation."

15. Page 20, line 19. The sentence "Furthermore, ..." needs rephrasing

Response: It is rephrased as "Furthermore, comparing to the sandblasting for the clay and clay loam, the dust originated from aerodynamic entrainment (which was not taken into account by the present dust model) was significantly constituted up to 28.3% and 146.4%, respectively"

16. Page 22, line 3. Please mention that you analyse hourly values.

Response: This sentence is revised as "…and normalized mean error (NME)) for the hourly data of 12 simulations and observations at 40 ground-based monitoring sites in NEC were calculated…."

17. Page 22, line 9. "was shown" should read "is shown"

Response: It is revised to "is shown"

18. Page 22, line 10. "abscissa" should be replaced by "distance to point OBS"

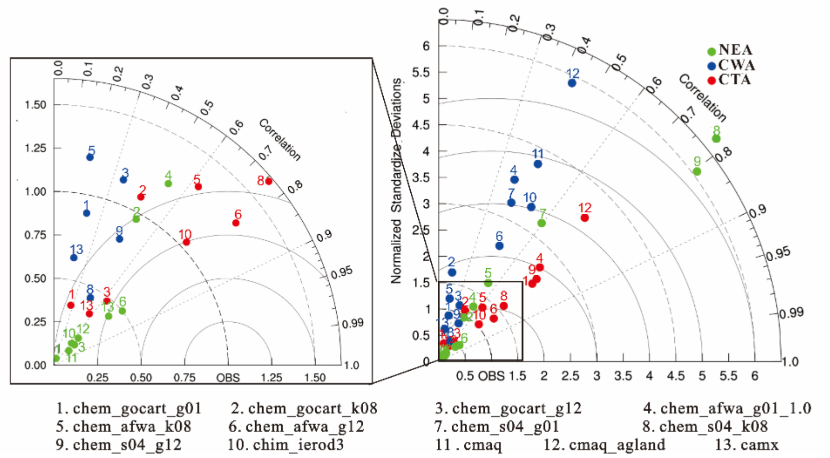
Response: It has been revised to "…while the RMSE (distance to point OBS) measures differences between the modeled and observed PM_{10}…."

19. Page 23, line 6. "Thus, NSD..." should read "Thus, NWA..."

Response: It is revised to "Thus, NWA was not included in the Taylor diagram."

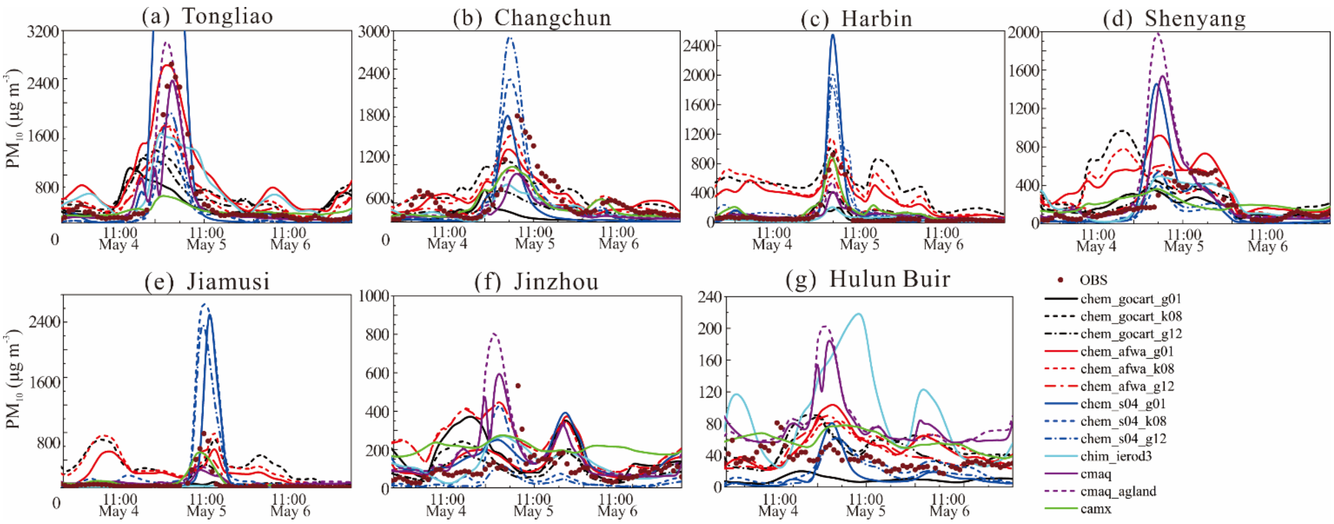
20. Page 23, Fig. 9. The colours for NEA and CWA are hard to distinguish

Response: Thank you for this helpful comment. The color of NEA was changed from purple to green and the revised figure is showed below.



21. Page 24, Fig. 11. It might be worth to enlarge the figure and refine the colours

Response: This figure had been enlarged and we also carefully refined the colors. It now shows as:



22. Page 26, line 25. "best near" should read "perform best close" or similar

Response: It is revised as "All simulations performed best near the dust source areas and degraded in accuracy…."

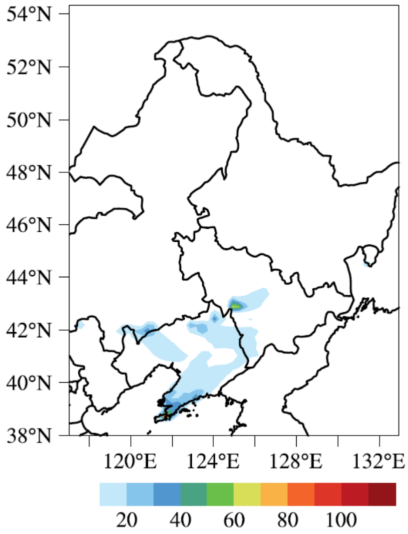
23. Page 27, line 3. This clearly is not related to the resolution but simply a matter of allowing emissions from areas not classified as desert or sparsely vegetated by refining the landcover mask.

Response: Thank for your very helpful suggestion. This sentence is revised as "A dust mask including dust emissions from regions not classified as "barren or sparsely vegetated" in CAMx should be developed by refining the land cover mask in future works."

24. Page 27, Author contributions. Please make sure that the order of the initials of each contributor is consistent with the author names on the title page

Response: Thank you for your valuable suggestion. The part of Author contributions is now revised as "MS, XZ and CG performed the majority of the source code reconfiguration of WRF-Chem, CHIMERE, CMAQ and CAMx, and initially designed the numerical simulations to carry them out. DQT, AX, WG and CX provided help for the simulation designation. LH provided support for conducting the CAMx model. HZ and SZ provided advices on the selection and usage of observational data. MS, XZ and DQT led the analysis of the simulations, and SIE, XW, XL and MD provided professional advices. SM and XZ wrote the paper and all authors read, revised, and approved the final manuscript."

In addition, we conducted reproducibility tests of our simulations and found an error in the WRF-Chem section. We have run the WRF-Chem model with UOC_Shao2011 scheme and the daily mean PM_{10} on May 5th, 2015 simulated by UOC_Shao2011 is provided below. It showed that spatial pattern and concentration level had little similarity with the observations. Considering its unreasonable results, then we selected UOC_Shao2004 for the subsequently simulations. It means that the actually used dust scheme in manuscript is UOC_Shao2004. However, the texts expressed the used scheme as UOC_Shao2011 by wrong. Therefore, we corrected this into UOC_Shao2004 in the manuscript.



Thank the reviewer again for the constructive criticisms that have helped us to improve our manuscript. We have tried our best to improve the manuscript and made changes in the manuscript. These changes will not influence the content and framework of the paper. And here we do not list the changes but marked in the revised manuscript.

All in a word, via these evaluation works, we hope to do some contributions to the community for enhance the dust forecast ability on regional scale in air quality models.

Multi-model simulations of a springtime dust storm~~s~~ over Northeastern China~~in East Asia~~: Implications of an evaluation of four commonly used air quality models (CMAQ v5.2.1, CAMx v6.50, CHIMERE v2017r4, and WRF-Chem v3.9.1)

Siqi Ma[1,2], Xuelei Zhang[1,3], Chao Gao[1,2], Daniel Q. Tong[3], Aijun Xiu[1], Guangjian Wu[4,5], Xinyuan Cao[1,2], Ling Huang[6], Hongmei Zhao[1], Shichun Zhang[1,7], Sergio Ibarra-Espinosa[1,8], Xin Wang[9], Xiaolan Li[10,11], and Mo Dan[12]

[1]Key Laboratory of Wetland Ecology and Environment, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130102, China
[2]University of Chinese Academy of Sciences, Beijing 100049, China
[3]Center for Spatial Information Science and Systems, George Mason University, Fairfax, VA 22030, USA
[4]Key Laboratory of Tibetan Environment Changes and Land Surface Processes, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100101, China
[5]CAS Center for Excellence in Tibetan Plateau Earth Sciences, Beijing 100101, China
[6] School of Environmental and Chemical Engineering, Shanghai University, Shanghai 200444, China
[7]Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, NC 27695, USA
[8]Department of Atmospheric Sciences, Universidade de São Paulo, São Paulo, SP, Brazil
[9]Key Laboratory for Semi-Arid Climate Change of the Ministry of Education, College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China
[10]Institute of Atmospheric Environment, China Meteorological Administration, Shenyang 110166, China
[11]School of Meteorology, University of Oklahoma, Norman, OK 73072, USA
[12]Beijing Municipal Institute of Labor Protection, Beijing 100054, China

Correspondence: Xuelei Zhang (zhangxuelei@neigae.ac.cn); Daniel Q. Tong (qtong@gmu.edu)

Abstract: Mineral dust particles play an important role in the Earth system, imposing a variety of effects on air quality, climate, human health, and economy. Accurate forecasts of dust events are highly desirable to provide early-warning and inform decision-making. East Asia is one of the largest dust sources in the world. This study applies and evaluates four widely used regional air quality models to simulate dust storms in ~~East Asia~~Northeastern China. Three dust schemes in the Weather Research and Forecast with Chemistry (WRF-Chem) (version 3.9.1), two schemes in CHIMERE (version 2017r4) and CMAQ (version 5.2.1), and one scheme in CAMx (version 6.50), were applied to a dust event during May 4th~6th, 2015 in Northeastern China. Most of these models were able to capture this dust event, except CAMx which has no dust source map covering the study area, hence ~~is excluded from subsequent analysis~~another dust source mask map was introduced to replace the default one for the subsequent simulation. Although these models reproduced the spatial pattern of the dust plume, there were large discrepancies between predicted and observed PM_{10} concentrations in each model. In general, CHIMERE

had relatively better performance among all simulations with default configurations. After parameter tuning, WRF-Chem with the AFWA scheme using seasonal dust source map from Ginoux et al. (2012) showed the best performance, followed by WRF-Chem with UOC_~~Shao2011~~ Shao2004 scheme, CHIMERE, and CMAQ. The performance of CAMx had significantly improved by substituting the default dust map and removing the friction velocity limitation. This study suggested that the dust source maps should be carefully selected on regional scale or replaced with a new one constructed with local data. Moreover, further study and measurement on sandblasting efficiency of different soil types and locations should be conducted to improve the accuracy of estimated vertical dust flux in air quality models.

Key words: Air quality models, Dust, Forecast, Evaluation, Asia, Northeastern China

1. Introduction

Wind-blown dust is typically emitted from areas with dry, erodible surfaces, such as desert, Gobi and cropland, during high wind periods. It exerts significant effects on air quality (Giannadaki et al., 2014), atmospheric visibility (Mahowald et al., 2007), human health (Goudie, 2014; Zhang et al., 2016; Tong et al., 2017), ecosystem (Jickells et al., 2005; Schulz et al., 2012) and climate (Prospero and Lamb, 2003). Depending on the extent to which human activities are involved, dust emissions can be classified as natural or anthropogenic. Natural dust emissions are activated by wind from undisturbed surface in arid or semi-arid areas, such as the Sahara Desert in North Africa (Formenti et al., 2011), alluvial plains and deserts in West Asia and Central Asia (Cao et al., 2015; Xi and Sokolik, 2015), deserts and sandy lands in East Asia (Laurent et al., 2005), various desert landforms in Australia (Revel-Rolland et al., 2006) and deserts in the southwest USA (Gillette et al., 1996; Zhao et al., 2012). Anthropogenic dust emissions are either activated by mechanical forces (tilling, mining, etc), or by wind at surface disturbed by human activities. Agricultural activities that disturb the soil surface (such as tillage and reaping) can greatly increase the frequency and intensity of wind-blown dust (Zender et al., 2004; Guan et al., 2016). The erodible potential of farmlands depends strongly on agricultural management practices, such as timing of cropping and grazing, and soil conservation measures (Munkhtsetseg et al., 2017). A modeling study by Liora et al. (2016) showed that anthropogenic dust contributes approximately 10% of total PM_{10} emissions in Europe. Using remote sensing observations, Ginoux et al. (2012) estimated that anthropogenic wind-blown dust sources account for 75% of emissions in Australia and 25% globally. Wind-blown dust emissions from cropland is of global importance (Mendez and Buschiazzo, 2010; Singh et al., 2012; Wang et al., 2013; Chappell et al., 2014; Xi and Sokolik, 2016).

Numerical dust models are often used to assess the magnitude of wind-blown dust emission and to predict its effects on air quality and climate. Several dust schemes to estimate the dust flux into the atmosphere and other relevant parameters have been proposed in the past twenty years, and some have been coupled with air quality models, such as WRF-Chem (Kang et al., 2011; Su and Fung, 2015; Flaounas et al., 2017) , CMAQ (Wang et al., 2012; Foroutan et al., 2017), CAMx (Klingmuller et al., 2018), CHIMERE (Menut et al., 2013; Mailler et al., 2017), ALADIN-SURFEX (Mokhtari et al., 2012), LOTOS-EUROS (Manders-Groot et al., 2016), EMEP MSC-W (Simpson et al., 2012), NAQPMS (Li et al., 2012) and CUACE/Haze (Wang et al., 2015). These models are widely used to study the air quality and climate effects of dust emissions. Most model applications, however, only adopt dust schemes designed for natural wind-blown dust from arid areas. It is unclear how well these models perform in areas with active agricultural operations.

The selection and usage of dust emission schemes and their input datasets are very important in establishing reliable air quality prediction. The evaluation and validation of dust emission schemes and relevant datasets in different air quality models on a continental scale has been carried out for East Asia (Dong et al., 2016), West Asia (Nabavi et al., 2017; LeGrand et al., 2019), North America (Foroutan et al., 2017), Europe, Northern Africa and the Middle East (Menut et al., 2013; Flaounas et al., 2017; Rizza et al., 2017). Nevertheless, evaluation of the regional performance of different air quality models and dust schemes remains inadequate.

Many previous multi-model evaluation studies focused on the climatic implications of different dust schemes at both global and regional scales. A comprehensive evaluation of 14 global aerosol models reported that the estimated dust emissions in Asia vary widely, ranging from 27 to 873 Tg/year (Huneeus et al., 2011). Two different dust emission schemes in EMAC (ECHAM5/MESSy2.41 Atmospheric Chemistry model) were shown to produce similar atmospheric dust loads in North Africa, but differ considerably in Asia, the Middle East and South America (Astitha et al., 2012). Ridley et al. (2016) reported that the global simulated aerosol optical depth (AOD) may vary by over a factor of 5 among four global models, and that dust emissions in Africa are often overestimated at the expense of emissions from Asia and the Middle East; in addition, dust was removed too rapidly in most models. On the regional scale, Todd et al. (2008) showed that the simulated dust flux and concentration from five models differed by at least one order of magnitude during a 3-day dust event over the Bodélé depression in northern Africa. Evan et al. (2015) demonstrated a power law relationship between modeled dust emission frequency and dust emission intensity in four regional models for North Africa. Huneeus et al. (2016) evaluated five dust forecast models during an intense Saharan dust outbreak affecting western and northern Europe in April 2011, noting that all models were better at predicting AOD than near-surface dust concentration over the Iberian Peninsula and tended to underestimate the long-range transport of dust. An evaluation of eight regional dust models with various dust emission schemes and other configurations was conducted for East Asia by Uno et al. (2006). Their results demonstrated that the models could correctly capture the major dust onset and cessation timing, but the maximum concentration of each model differed by a factor of 2-4. Several other studies have focused on the assessment of one or more dust schemes in the

WRF-Chem model, over regions such as the Mediterranean (Flaounas et al., 2017), Middle East (Prakash et al., 2015), and Central and East Asia (Darmenova et al. 2009; Xi and Sokolik, 2015). Dust modeling requires sufficient parametrization, high-quality input data and practical tuning techniques to enable results to best match observations (Basart et al., 2012; Flourous, 2017).

Accurate forecasts of dust emissions and transport are demanded globally by society, to address many health and economic issues, especially air quality. Here we present a~~the first~~ comprehensive evaluation of multi-model simulations of windblown dust emissions <u>in air quality models during a dust episode</u> in East Asia, using a number of dust emission schemes with four state-of-the-art air quality models. East Asia is one of the world's largest dust sources, contributing about 30% of total global dust loading. This study focuses on Northeastern China, a unique dust source region with varying land use types, including deserts, semiarid land, and croplands. In addition, this region is known for diverse soil texture and organic content. The Northeast China Plain, the nation's breadbasket, is made of soil with abundant organic matter. Dust storms originated in this region are often called "Black sandstorm" (Zhang et al., 2015). There are also areas with saline-alkali soil on the western side of this region, giving dust storms the white color ("white sandstorms"). This region is also known to experience high wind during springtime. All of these characteristics present challenges for numeric models to predict dust storms, making Northeastern China an ideal region for assessing the capability of dust models. We choose four air quality models (CMAQ v5.2.1, WRF-Chem v3.9.1, CHIMERE v2017r4 and CAMx v6.50), each configured with a selection of dust emission schemes and source maps to simulate a well-observed regional pollution event with strong dust influence. Detailed description of the study region, model configuration and dust schemes are presented in Section 2. Comparisons of dust schemes and dust source maps are described in Section 3. Results of model simulations and verification with the ground-based and satellite-based observations are presented and discussed in Section 4. We conclude in Section 5.

2. Model configuration, observations and methods

2.1 Study area and model domain

Northeastern China (NEC) (38°42′–53°33′ N and 115°31′–135°2′ E) is located at the eastern end of the northern hemisphere dust belt. This area covers about 1.47 million square kilometers, accounting for about 15% of the Chinese land area (Fig. 1). NEC has a semi humid continental climate with prevailing westerly winds throughout the year. A major grain production region in China, NEC includes the alluvial Northeast China Plain with farmlands characterized by mollisol (Udolls, USDA Soil Taxonomy, or Black Chernozem, Canadian soil classification). Due to the long cold season and strong spring winds, the exposed cropland is vulnerable to wind erosion (Dickerson et al., 2007). Two of the four major Chinese sandy lands, the Horqin and Hulun Buir sandy lands, are located in the western NEC while several other sandy/barren land regions are located in the central area and surrounded by cropland. The Gobi Desert between China and Mongolia is located to the west of NEC.

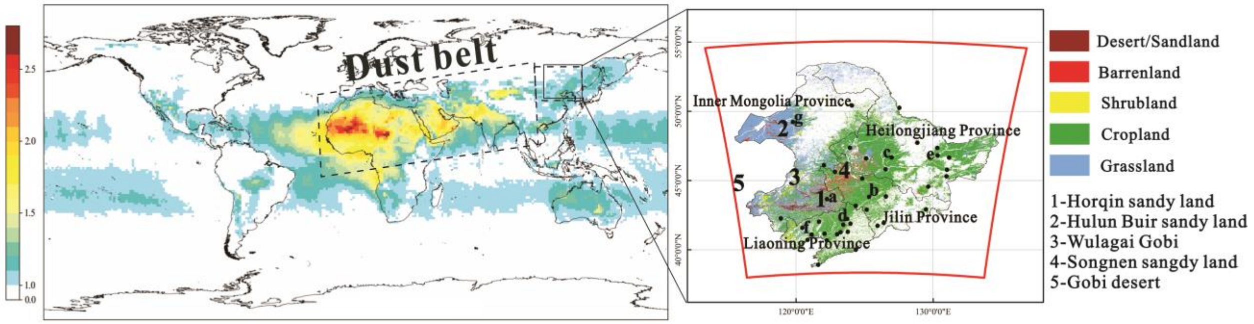


Figure 1. The global aerosol index distribution and the dust belt location (dashed rectangle) as described in Varga (2012), and the geographical coverage of the NEC domain on the right. Dots in the NEC domain represent monitoring sites (Labels a~g indicates the monitoring sites at Tongliao, Changchun, Harbin, Shenyang, Jiamusi, Jinzhou and Hulun Buir.)

The model domain centers on 46.715 °N, 125.081 °E and is defined on a lambert conformal projection. The true latitudes of the domain are 30 °N and 60 °N, and composed of 60×73 grid cells with a horizontal grid resolution of 25 km × 25 km and 30 vertical levels. The domain covers the whole of NEC (as shown by the dark blue line in Fig. 1). The initial and boundary fields were obtained from the final (FNL) operational global analysis data of the National Center for Environmental Prediction (NCEP) with a horizontal resolution of 1 °× 1 °, updated every 6 h (http://rda.ucar.edu/datasets/ds083.2).

2.2 Observational data sources

Air quality monitoring data were acquired from the national air quality history database (http://beijingair.sinaapp.com), which contains hourly data and information from China's national environmental monitoring center. The hourly monitoring data used in this study include PM_{10} and $PM_{2.5}$ concentrations in 40 cities of NEC (Fig. 1) for the time period from May 3rd to May 7th, 2015. Deep Blue Aerosol Optical Depth (AOD) data were obtained from the MODIS-Aqua with a resolution of 10 km×10 km from the archive of NASA Level-1 and Atmosphere Archive & Distribution System (LAADS) (https://ladsweb.modaps.eosdis.nasa.gov/archive). In addition, data from the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) satellite were used to investigate the vertical distribution of transported dust particulates.

2.3 The Springtime Dust Episode

The dust event on May 5, 2015 was selected for model evaluation in this study from examining the time series of observed

PM_{10} concentrations. Satellite images indicate that large areas of central NEC were covered by higher AOD during the May 5th dust event when compared to the preceding two days (Fig. 2a~c). The mean AOD in central NEC quickly increased from 0.6 on May 4th to >1.0 on May 5th, while AOD in other regions was relative lower (<0.3). This indicated that the event was not caused by long-distance transported dust from western China and Mongolia, but instead was a locally generated event. Meanwhile, the vertical distributions of aerosol subtypes derived from CALIPSO observations indicated that dust was distributed in the atmosphere below 1 km in NEC, acted as the primary pollutant on May 5th.

PM concentrations observed from 40 air quality monitoring sites over NEC were used to analyze the spatiotemporal distribution of the dust plumes. The spatial distribution of daily PM_{10} concentrations during May 5th, 2015 was consistent with the retrieved AOD (Fig. 2d). This event originated in the region around Tongliao (such as the Horqin sandy land and saline-alkali soil) on 20:00 UTC of May 4th, and lasted for nearly 17 hours.

To facilitate comparison and evaluation of these air quality models, we divided the study region into four areas according to PM_{10} levels: heavy dust central area (CTA), northwest moderate dust area (NWA), and two light dust areas in the northeast (NEA) and southwest (SWA) (Fig. 2e). Daily concentrations of PM_{10} at the central sites, such as Tongliao and Changchun, exceeded 700 μg m^{-3}. The concentrations of PM_{10} in NE and SW ranged from 100 to 500 μg m^{-3}. All these values considerably exceeded the PM_{10} level-2 concentration limits (150 μg m^{-3}) of the NAAQS (National Ambient Air Quality Standard). On May 5th, the ratio of $PM_{2.5}/PM_{10}$ was 0.14, indicating that the particulate matter was dominated by coarse dust particles (Tong et al., 2012), consistent with the aerosol subtype observations of CALIPSO (Fig. 2f).

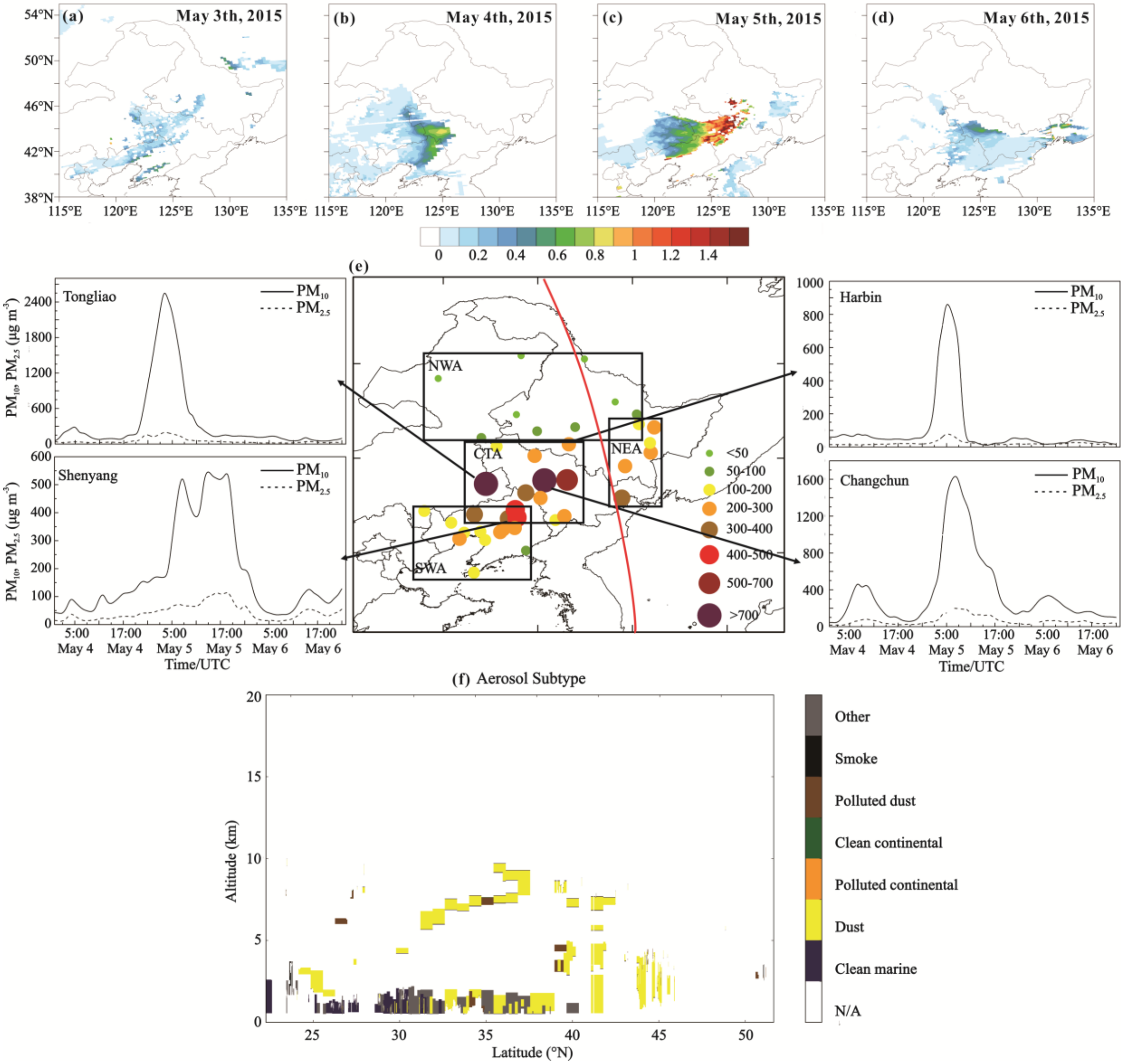


Figure 2. Satellite and ground observations of the May 5, 2015 dust event over Northeastern China: (**a**)–(**d**) Daily MODIS aerosol optical depth (AOD) at 550 nm before, during and after the storm; (**e**) daily mean PM_{10} concentrations (μg m^{-3}) measured at four ground sites (Tongliao, Changchun, Harbin and Shenyang) on May 5th, 2015 and hourly PM_{10} (solid line) and $PM_{2.5}$ (dashed line) variations in during the dust period. The blue line in (**e**) indicates the path of the CALIPSO satellite, and CALIPSO aerosol subtype between 4:30 and 4:43 UST (**f**).

2.4 Description of air quality models and dust schemes

This study focuses on comparing dust emission schemes in four air quality models as described below.

2.4.1 Dust schemes in WRF-Chem v3.9.1

The Weather Research and Forecast community model coupled with a chemistry model (WRF-Chem) is a coupled online community model able to simulate gas and aerosol chemistry simultaneously with the meteorological fields, and is generally used for the prediction and simulation of weather, air quality and regional climate from cloud scales to regional scales (Grell et al., 2005). The WRF-Chem version 3.9.1 was used in this study. Three dust schemes, GOCART, AFWA and UOC, are tested here. The latter scheme was further divided into three dust emission parameterizations with various levels of complexity, namely Shao2001, Shao2004 and Shao2011 (Shao, 2001; Shao, 2004; Shao et al., 2011).

In the GOCART scheme, the dust emission is based on an equivalent empirical formulation by Gillette and Passi (1988) which requires data on the wind speed at 10 m and a threshold velocity to initiate wind erosion, as well as the surface erodibility (Ginoux et al., 2001). Comparing to other dust emission schemes, the dust emission flux in this scheme can be simply and directly calculated via the variables like wind speed, soil moisture and air density (which can be obtained from most numerical weather models) over source emission areas within the dust source map, without the conversion from horizontal to vertical flux (Fig. S1). The AFWA dust scheme is a modified version of Marticorena and Bergametti (1995) dust scheme~~GOCART~~ developed by the Air Force Weather Agency (LeGrand et al., 2019). Unlike GOCART scheme, friction velocity is introduced into this scheme and the d~~D~~ust emission ~~in this scheme~~ is calculated as a saltation flux and vertical uplift dust flux, which is proportional to the horizontal saltation flux, based on the soil clay content (Marticorena and Bergametti, 1995). A soil moisture correction term is also applied to the threshold friction velocity, however, this term in AFWA scheme is calculated according to the method described by Fécan et al. (1999) (Fig. S2), which is different from that used in the GOCART scheme. ~~The UOC (University of Cologne) scheme~~ The UOC (University of Cologne) scheme accounted for the saltation bombardment, aggregate disintegration and volume removal of saltating particles. The vertical dust emission flux is proportional to horizontal saltation flux, but the ratio significantly depends on soil texture and soil plastic pressure (Shao, 2004). The fully disturbed soil particle size distribution was omitted in the simplified scheme of Shao2011 (Shao et al., 2011). The parameterization of ~~Shao2011~~ Shao2004 has been verified by field observations and was therefore adopted in this study. Unlike the GOCART and AFWA dust emission schemes, the threshold friction velocity is obtained via the method from Shao and Lu (2000) rather than Bagnold (1941). Although the equation of moisture correction in UOC scheme is also from Fécan et al. (1999), it is based on the volumetric soil moisture and empirical constants as a function of soil texture (Klose et al., 2014). Furthermore, an additional correction term, roughness correction (or drag partition correction), is also introduced to describe the influence of non-erodible elements (such as vegetation, peddle etc.) on the threshold friction velocity (Raupach, 1992) (Fig. S3). In addition, the UOC scheme only uses the erodible area to constrain the dust source locations instead of scaling dust emissions. Note that the last term in the saltation flux formula in UOC source code is mistakenly expressed as $(1+(\frac{u_{*t}}{u_*})^2)$ in WRF-Chem before the version of 4.0. In this study, the last term in the saltation flux formulait has been debugged corrected from $(1+(\frac{u_{*t}}{u_*})^2)$ into $(1+\frac{u_{*t}}{u_*})^2$ in the UOC source code in WRF-Chem version 3.9.1 according to the description in Shao et al. (2011). This revision would could increase the saltation flux by a factor of 2 or more. More detailed physical descriptions and defects in source codes of above three dust schemes in WRF-Chem model had been explicitly documented, and all schemes had been evaluated over southwest Asia in LeGrand et al. (2019).

2.4.2 Dust schemes in CHIMERE v2017r4

CHIMERE is an Eulerian off-line chemistry-transport model covering local to continental scales (from 1 km to 1 degree resolution). An aerosol module was implemented into CHIMERE in 2004 with further modifications concerning the natural dust emissions and resuspension over the northern Atlantic and Europe (Vautard et al., 2005; Hodzic et al., 2006). Dust emissions have been verified for long-distance transported dust by comparison with long-term and field measurements (Schmechtig et al., 2011; Bessagnet et al., 2017). The CHIMERE version 2017r4 was used in this study.

Three dust emission schemes were employed in the CHIMERE model: the MBW scheme (White, 1986; Marticorena and Bergametti, 1995), AGO scheme (Alfaro and Gomes, 2001; Menut et al., 2005) and KOK scheme (Kok et al., 2014a). Extension of the dust production model to any domain over the globe was available since the model version of chimere2016a, and the KOK scheme was also implemented in this version. In the MBW scheme, the vertically integrated saltation flux was estimated using the equation introduced by White (1986). The vertical dust flux in the second scheme was computed based on the partitioning of the kinetic energy of individual saltating aggregates and the cohesion energy of the populations of dust particles with the assumption that dust emitted by sandblasting is characterized by three modes whose proportion depends on the wind friction velocity. The vertical dust flux in the KOK scheme was estimated directly without converting from horizontal flux to vertical flux but only controlled by dust emission coefficients, namely, bare soil fraction, soil clay fraction, surface friction velocity and threshold friction velocity (Kok et al., 2014a). The dust schemes in CHIMERE follow similar calculating process as the UOC scheme, and the flow chart for these schemes is showed in Fig. S4. Moreover, there are two options for calculation of threshold friction velocity, Iversen and White (1982) and Shao and Lu (2000) in CHIMERE, and it uses the equation from Marticorena et al. (1997) to calculate the roughness correction. Additionally, it needs to note that the friction velocity is calculated independently in this model and equation of Weibull distribution is applied for wind speed adjustment (Cakmur et al., 2004; Pryor et al., 2005). According to the CHIMERE source code, all three schemes needed external land-surface static data (such as land use type, soil type/fraction and vegetation cover) for the erodibility factor calculation.

2.4.3 Dust schemes in CMAQ v5.2.1

The Community Multiscale Air Quality (CMAQ) model is a 3-D Eulerian photochemical dispersion model that allows for an integrated assessment of gaseous and particulate air pollution over many scales ranging from sub-urban to continental (Byun and Schere, 2006). CMAQ version 5.2.1 was used in this study.

The first wind-blown dust emission scheme, named FENGSHA, was implemented into CMAQ version 5.0 in 2012. Four land use types (barren land, shrub-grass land, shrub land and cropland) were treated as potential erodible dust sources instead of dust source maps, and the dust vertical flux was calculated according to a modified Owen's equation (Owen, 1964) when the friction velocity (u_*) exceeded the threshold friction velocity (u_{*t}) (which was set as constant value for each potential erodible land use type and soil texture based on literature and field measurements, such as 0.63 m s^{-1} for clay loam of barren land). The effect of agricultural activities was calculated via a crop calendar, allowing the vegetation fraction to vary with the change of date. In the calculated dust emission from croplands in CMAQ, the vertical flux was improved by adding another two factors: the crusting factor f_{cs} and tillage-ridge factor f_{tr} (Zhang et al., 2015).

Another dust emission scheme was applied in CMAQ version 5.2 in 2017. In this scheme, the vertical dust emission flux was acquired based on the calculated horizontal dust flux (White, 1979) and sandblasting efficiency. The threshold friction velocity was calculated following Shao and Lu (2000), and the friction velocity was calculated based on an updated dynamic relation for the surface roughness length relevant to small-scale dust generation processes (Foroutan et al., 2017) (Fig. S5). Besides the potentially erodible land use types which were the same as those in the original FENGSHA module, the satellite-observed fraction of absorbed photosynthetically active radiation (FPAR) was introduced to act as a surrogate for vegetation cover fraction to constrain dust emission.

2.4.4 Dust scheme in CAMx v6.50

The Comprehensive Air quality Model with extensions (CAMx) is an Eulerian chemistry-transport model that allows for an integrated "one-atmosphere" assessment of gaseous and particulate air pollution ranging from urban to continental scales. The inline dust emission module had not been implemented into the newly-released CAMx version 6.50, but it had been fully developed as a pre-processing program (namely wbdust) which was used to provide binary dust emission files and to merge them with the emissions from other sources into model-ready emission files. The dust emission scheme in CAMx was based on a revised mineral dust emission scheme in the atmospheric chemistry–climate model EMAC (Astitha et al., 2012; Klingmuller et al., 2018). We obtained the source code of this dust scheme through private communication with Dr. Yarwood Greg. The Its vertical dust emission flux was calculated via the saltation flux and sandblasting efficiency when friction velocity exceeds a threshold value (Marticorena and Bergametti, 1995), similar to the MBW scheme in CHIMERE. The major modifications improvements and adjustments were omitting the term supposed to account for the effect of soil moisture omitting the effect of soil moisture on dust emission, adding a topography factor which accounted for enhanced emissions from basins and valleys, filtering the sandblasting efficiency of the soil clay fraction by a Gaussian function with an interquartile range of 5%, and limiting the minimum maximum value of the friction velocity to 0.4 m s^{-1} (Klingmuller et al., 2018). The schematic diagram of dust emission module in CAMx, as well as those in WRF-Chem, CHIMERE and CMAQ, are all provided in Fig. S1~S6 of the supplementary file.

2.5 Model configuration

2.5.1 Physical parameterization

The Weather Research and Forecasting (WRF) model version 3.9.1 was used to conduct the meteorological simulations, then to provide the hourly meteorological output fields to drive the air quality models of CHIMERE, CMAQ and CAMx while the chemistry module of WRF (WRF-Chem) was conducted simultaneously with the meteorological fields. As the surface wind speed was the dominant factor controlling dust blowing and transportation, its accuracy could significantly influence the results of dust modeling. Furthermore, the land surface characteristics played an important role in the WRF surface wind simulation. For the purpose of comparing and selecting the optimal scheme (Table 1) to be used in the following dust emission simulations in the air quality models, two scenarios with different land-surface schemes (Noah-MP scheme and Pleim-Xiu scheme) were chosen for comparison. More detailed comparisons are provided in Section 12 of the Supplementary Information, and Scenario 2 was finally selected for the WRF model.

The FNL and static geographical fields are interpolated to the model domain resolution of 25 km × 25 km by using the WRF preprocessing system (WPS). In addition, the land use and soil category datasets used in this study were obtained from WPS static data file website (http://www2.mmm.ucar.edu/wrf/src/wps_files/), IGBP-Modified MODIS land use data was selected for the simulations of WRF-Chem, CMAQ and CAMx model while USGS dataset was used for CHIMERE as the dust model in CHIMERE could only read USGS data for dust emission calculations.

Table 1. WRF parameterization settings

Physical scheme	Scenario 1	Scenario 2
Microphysics	WRF double moment, 6-class scheme	
Longwave radiation	rrtmg scheme	
Shortwave radiation	rrtmg scheme	
Surface layer	Revised MM5 Monin-Obukhov	Pleim-Xiu scheme
Land-surface	Noah-MP land-surface model	Pleim-Xiu scheme
Number of soil layers	4	2
Boundary layer	YSU scheme	ACM2 (Pleim) scheme
Cumulus parameterization	Grell-Devenyi ensemble scheme	

2.5.2 Chemical parameterization

WRF-Chem v3.9.1 simulations were executed with different source maps (G01, G01_~~1x1~~1.0, K08, G12 and MDB) for each dust scheme, GOCART, AFWA, and UOC_~~Shao2011~~Shao2004. The chemistry scheme with chem_opt of the GOCART simple aerosol scheme was used without anthropogenic emission input. CHIMERE v2017r4 was used with the MELCHIOR chemistry mechanism and MBW, AGO and KOK dust emission schemes along with 3 algorithms of erodible fraction, as well as no surface anthropogenic emissions. The Iversen and White (1982) and Shao and Lu (2000) methods (IW and SL hereafter) were chosen as algorithms for calculation of the threshold friction velocity.

As for CMAQ v5.2.1, the CB6R3 gas-phase mechanism and AE6 aerosol mechanism with sea salt and speciated PM aqueous/cloud chemistry (cb6r3_ae6_aq) were used in this study. The inline dust emission calculation was executed with and without agricultural activity (CTM_ERODE_AGLAND).

The dust emission module used the meteorological output fields to obtain a gridded dust emission flux, and then the emitted dust flux was reformatted and merged with anthropogenic emissions for CAMx. CAMx v6.50 with the CB6r2 gas-phase mechanism and AE6 aerosol mechanism was used in this study.

All simulations covered the 96 hours from 0:00 May 3th to 23:00 May 6th, 2015 and the first 24 hours was regarded as model spin-up. And more detailed configuration information and the namelist files for model simulations were provided on https://doi.org/10.5281/zenodo.3376774~~https://zenodo.org~~.

3. Results and discussion

3.1 Comparison of dust source maps and erodibility fractions

3.1.1 Comparison of dust source maps in WRF-Chem v3.9.1 and CAMx v6.50

The default dust source map (or dust source function) with a horizontal resolution of 0.25°×0.25° used in WRF-Chem v3.9.1, was obtained from static geographical datasets (Ginoux et al., 2001). The map comprises the gridded fraction of alluvium available for wind erosion, calculated from topography and elevation. In this study, we named this source map as G01_0.25 according to first author and published year of relevant literature and its spatial resolution. ~~This source map~~ It ~~(G01_0.25)~~ ~~showed~~ shows only one weakly erodible area in NEC, located in the Horqin sandy land, with erodibility fraction values < 0.2 (Fig. 3a). Because the source map plays a critical role in determining the spatial distribution of dust emission and the calculated magnitude of dust fluxes, here we test five other dust source maps to test with these models (see Table 2 for more details). The source map with resolution of 1°×1° developed by Ginoux et al. (2001) (namely G01_1.0) basing on the same method with G01_0.25 was obtained from the homepage of Dr. Paul Ginoux (https://www.gfdl.noaa.gov/pag-homepage). Except G01_0.25 and G01_1.0, the other source maps were obtained using satellite observations. The map with resolution of 0.25°×0.25° provided by Koven and Fung (2008) (K08_0.25 hereafter) was calculated via the relationships between landscape characteristics, residual landscape roughness and aerosol optical depth ~~(Koven and Fung, 2008)~~, and presented as globally-available erodible fractions. A global-scale high resolution (0.1°) dust source product was derived from a combination of a climatological analysis of MODIS Deep Blue AOD data and land use data (Ginoux et al., 2012). Sources were classified as natural or anthropogenic (primarily agricultural) and their global distributions were described by frequency-of-occurrence (FO). In this study, we established a simple conversion from FO value into erodible fraction: FO(0.05)→0.15, FO(0.1)→0.3, FO(0.2)→0.4, FO(0.25)→0.5, FO(0.4)→0.7, FO(0.5)→0.8, FO(0.6)→0.9. The source map with only natural origins was named G12_0.1_natural while that with both natural and anthropogenic origins was named G12_0.1_ant+nat.

The source map of ~~G01 version with a spatial resolution of 1°×1° (~~G01_1.0~~)~~ evidently had more widespread erodible lands than that of G01_0.25, regardless of region or erodible fraction values (Fig. 3b). Four erodible areas were depicted in the dust source map of K08_0.25: the central plain area of NEC, Hulun Buir sandy land, the northeast corner of NEC and the North China Plain (Fig. 3c). It was obvious that the croplands were identified as dust source areas in K08_0.25, although the erodibility was likely overestimated (e.g., it can not be higher than that of the sandy lands as shown in this map). This is due to the difficulty to distinguish anthropogenic emitted particulates, such as industrial emissions, from the AOD used for retrieving dust sources. Moreover, the erodible fraction values in the northeast corner were significantly overestimated, for example those in Xingkai (Khanka) Lake (45.33 °N, 132.67 °E) and its surroundings did not have any erodible potential according to our ground survey. The spatial distribution of dust sources in G12_0.1_natural in this area seems to miss many dust source areas, such as the Horqin sandy land and Hulun Buir sandy land (Fig. 3d). The dust source map~~ of G12_0.1_seasonal was ~~obtained from the NASA-Unified ~~Weather Research and Forecasting~~WRF (NU-WRF) version 7 ~~model ~~(Kim et al., 2014)~~, and~~ was similar to the former source map but divided into four seasons, therefore it was named as G12_0.1_seasonal in this study (Fig. 3f). The spring spatial distribution of dust sources was used in this study. In comparison, G12_0.1_ant+nat and G12_0.1_seasonal had a similar spatial pattern to that of K08_0.25 but with lower fraction values (Fig. 3c, 3e and 3f), as they were retrieved from the same satellite products. The pattern of the spatial distribution was reasonably similar to the erodible land use type distribution shown in Fig. 1.

Table 2. Information on dust source maps used in WRF-Chem

Name	Method	Region	Resolution	Time	References
G01_0.25	Topographic depression	Global	0.25 °	Constant	Ginoux et al (2001)
G01_1.0	Topographic depression	Global	1 °	Constant	Ginoux et al (2001)
K08_0.25	Satellite AOD, levelness and residual landscape roughness	Global	0.25 °	Constant	Koven and Fung (2008)
G12_0.1_natural	Satellite AOD and frequency of dust occurrence	Global	0.1 °	Constant	Ginoux et al. (2012)
G12_0.1_ant+nat	Satellite AOD	Global	0.1 °	Constant	Ginoux et al. (2012)
G12_0.1_seasonal	Satellite AOD	Global	0.1 °	Seasonal	Ginoux et al. (2012)

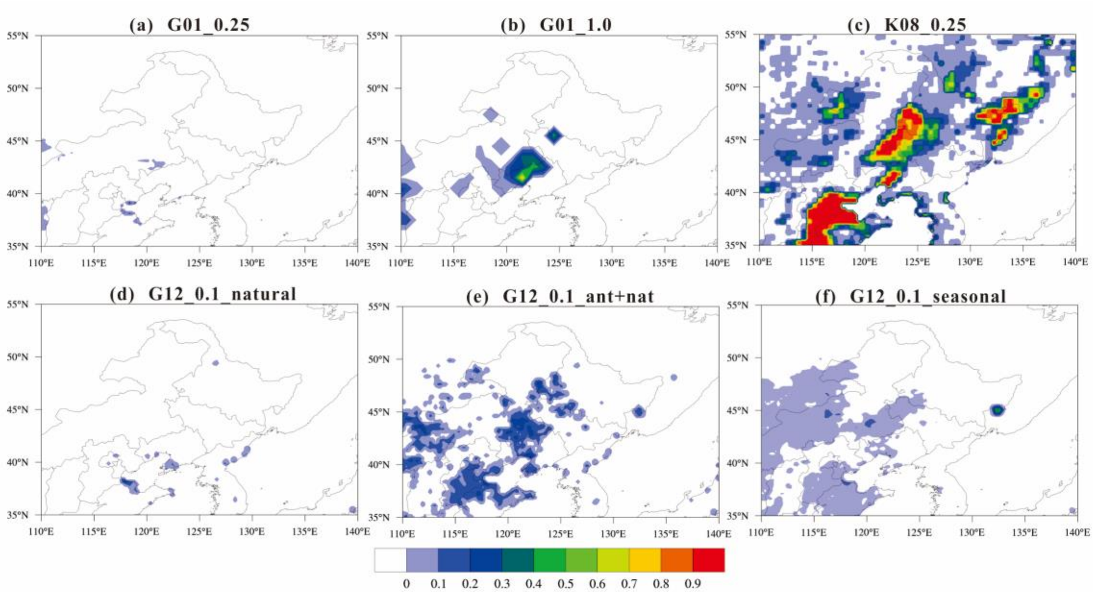


Figure 3. Dust source maps in NEC. (**a**) G01_0.25, (**b**) G01_1.0, (c) K08_0.25, (**d**) G12_0.1_natural, (**e**) G12_0.1_ant_nat, (**f**) G12_0.1_seasonal (spring))

A dust mask file was used in CAMx v6.50, which only had two values: 0 indicating no erodible dust potential while 1 dust emitting capacity in the grid cell. Dust flux was then calculated with the clay fraction-dependent vertical-to-horizontal dust flux ratio (Fig. S~~2~~8a). Unfortunately, no dust erodible area was recorded for the NEC region in the dust mask file (Fig. S~~2~~8b). Therefore, a dust source map will be introduced and used instead of the original dust mask file for ~~no further~~further evaluations ~~was conducted for~~ of the ~~inline~~ dust emission scheme in CAMx.

3.1.2 Comparison of erodible land fractions in CMAQ v5.2.1 and CHIMERE v2017r4

Instead of using a prescribed dust source map, the dust schemes in CMAQ v5.2.1 and CHIMERE v2017r4 estimated erodible land fraction based on land use, crop types, and/or crop calendar. Erodible land fraction in CMAQ was calculated by multiplying the fraction of erodible land use type with an erodibility potential factor assigned to each land use type. In this method, four land use categories~~—~~ (shrub land, shrub grass, cropland and sparse barren land~~—~~) were considered as erodible land types when the Biogenic Emissions Landcover Database version 3 (BELD3) dataset was used during regional simulation in the USA. Otherwise, the global MODIS FPAR data were recommended to represent the vegetation fraction (Flaounas et al., 2017). However, the grassland fraction was not taken into account ~~in either~~when using the USGS or MODIS land use dataset according to the source code of CMAQ, which may lead to an underestimation of dust emission. Thus, this kind of land use type was added into the source code file of LUS_DEFN.F using the same erodible fraction as that of a similar land use type (shrub-grass) in BELD3 (with an erodibility potential value of 0.25), and the final, modified distribution of erodible fraction is depicted in Fig. 4a.

Three methods were used for calculating erodible land fraction in the CHIMERE model. When erodibility option *ierod* was set to 1 in the model, the USGS land use data were used, and the erodible fraction depended on the fractional area without vegetation if the land use type was cropland. The fraction was set to 1 for shrub and barren lands. The second method (*ierod* = 2) simply depended on an erodibility value retrieved from monthly erodibility data derived from MODIS surface reflectance stored in the CHIMERE static dataset. The last method was the mixed usage of USGS and MODIS (*ierod* = 3): when land use type was cropland, the fraction was calculated following method 1, otherwise it was set to the MODIS erodibility for shrub and barren lands. Note that *ierod* = 3 was the default option for dust emission in CHIMERE. Figure. 4(b~d) shows that the erodible fractions acquired from the three methods had similar distributions, although values in method 2 were considerably lower than those of the other two. In addition, the user guide for CHIMERE noted that the USGS land use type must be used to ensure the erodible land fraction is calculated correctly.

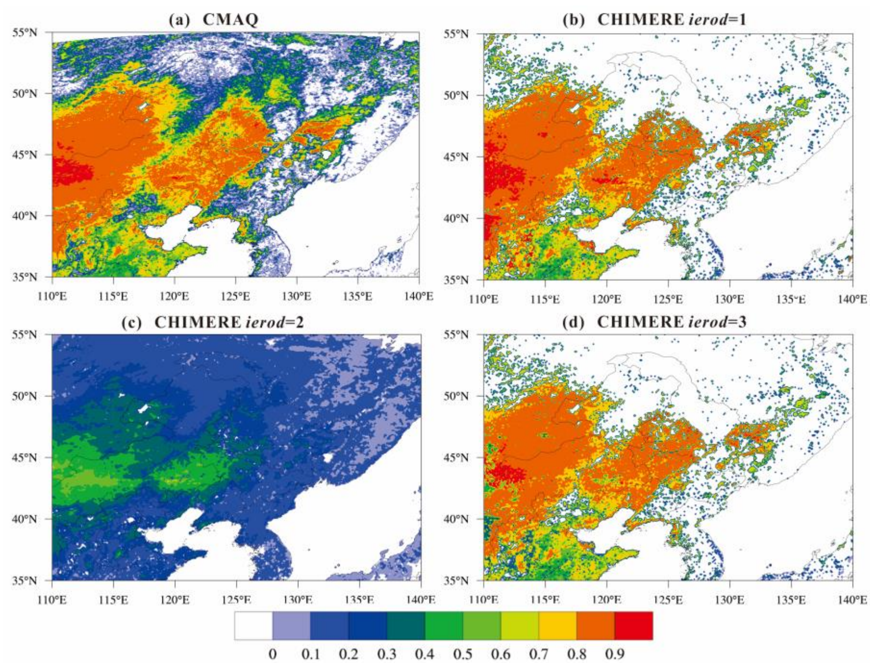


Figure 4. Maps of erodible fractions in NEC in CMAQ (**a**) and CHIMERE model (**b**) uses USGS land use (*ierod*=1), (**c**) uses MODIS surface reflectance (*ierod*=2), (**d**) is a mix of USGS and MODIS (*ierod*=3)

3.2 Performance of WRF-Chem v3.9.1 dust simulation

WRF-Chem v3.9.1 simulations showed large differences among different dust schemes and source maps in terms of both spatial distributions and values of dust emissions. PM_{10} simulated by GOCART scheme with default source map (G01_0.25) presented relatively low concentrations with daily averages less than 40 μg m^{-3}, mainly concentrated in the Horqin sandy land in Tongliao and extending to the east of Liaoning Province with concentrations less than 20 μg m^{-3} (Fig. 5a). In comparison, the spatial distribution of daily PM_{10} concentration simulated by GOCART with the G01_1.0 source showed three dominant dust emission areas: Horqin; the border between western Jilin and Heilongjiang Province; and coastal Liaodong Bay (Fig. 5b). The higher simulated PM_{10} concentration in the latter area was not supported by observations. The high concentration centers around Tongliao and in eastern Liaoning Province reached values of 600 μg m^{-3}, but dust was transported directly to the west without reaching the cities of Changchun and Harbin. The distribution of PM_{10} using the GOCART scheme with the K08_0.25 source map yielded concentrations in Horqin sandy land and Songnen sandy land and their surrounding areas. The simulated concentrations with K08_0.25 were greater than those simulated by the GOCART scheme with other source maps (Fig. 5a~f), but two-fold of the observed values. The result of G12_0.1_natural (Fig. 5d) showed dust emissions in coastal Liaoning Province with quite low dust intensity, showing that this natural source distribution was not applicable in this area. The PM_{10} patterns with G12_0.1_ant+nat and G12_0.1_seasonal indicated that their dust source regions were similar (Fig. 6e and 6f), and the simulated daily concentration of G12_0.1_seasonal was only about 100~200 μg m^{-3}, compared to the observed concentrations of 100~700 μg m^{-3}.

Similar spatial distributions of PM_{10} were obtained using the AFWA and UOC_~~Shao2011~~ Shao2004 schemes with the above 6 NEC dust source maps, but the simulated PM_{10} concentrations using each source map was more than 1 order of magnitude greater than those of GOCART (Fig. 5g~5r). Furthermore, the spatial patterns of dust simulated by UOC_~~Shao2011~~ Shao2004 with the last two source maps (Fig. 5q~5r) extended further northeastwards than those with GOCART and AFWA, and were more consistent with the observations. The overestimation of AFWA scheme might in part be explained by the fact that the AFWA scheme considered vertical dust flux only related to the clay content, unlike the UOC scheme which considered it to be inversely proportional to surface hardness (Kang et al., 2010; Rizza et al., 2016; Rizza et al., 2017). Meanwhile, ~~The overestimation of AFWA scheme might attribute to~~ the misuse in number and distribution of saltation size bins in AFWA might also be part of the reason. The last three bins of the total nine saltation size bins were sand-sized bins and they were also configured to constitutes all of the possible sand mass fractions which indicated that the sand in the soil surface was entirely composed of fine sands, resulting in the increase of the strength of the saltation bin-specific weighting factors and emission of the dust particles (LeGrand et al., 2019). In addition, we found that the simulated dust concentration with these three schemes generally presented over-predictions in this area, this might because of the usage of Pleim-Xiu (PX) land surface scheme. An additional dust simulation with Noah land surface scheme was conducted and the results showed lower PM_{10} concentration (Fig. S9), and it also indicated similar wind speed and higher surface soil moisture (Fig. S10, Table S7) than the simulated values using PX scheme. The higher soil moisture resulted in increasing the value of soil moisture correction, which was used for calculating threshold friction velocity, by about 10%. These discrepancies may result in the differences of estimated dust emissions and it could be the reason of the stronger dust emission when using the PX scheme.

As mentioned above, we found that the distribution and intensity of modeled dust aerosols were sensitive to the dust source maps in use ~~could critically control the distribution and intensity of modeled dust aerosols~~. Further analysis of Figs. 3 and 5 shows that the source maps of G01_0.25 and G12_0.1_natural were not able to reproduce this dust event in NEC. Dust source regions in other source maps were more or less similar, generally located in Horqin sandy land, mid-west Jilin Province and coastal Liaodong Bay. Observations suggest the modeled dust source in Liaodong Bay might be inaccurate, as measured concentrations were relatively low in that area.

Since results obtained by all three dust emission schemes with four source maps (G01_1.0, K08_0.25, G12_0.1_ant+nat, and G12_0.1_seasonal) were in better agreement with observations, these maps were used in subsequent evaluation. Note that the exceedingly high PM_{10} calculated via AFWA and UOC_~~Shao2011~~ Shao2004 or with the K08_0.25 source map indicated that a tuning coefficient was needed to improve the model performance. Considering that the source maps of G12_0.1_ant+nat and G12_0.1_seasonal were obtained via the same methodology but the latter one provides seasonal divisions making it more reasonable and closer to the actual environment, G12_0.1_seasonal~~the latter (with seasonal divisions)~~ was chosen for the next step in the evaluation.

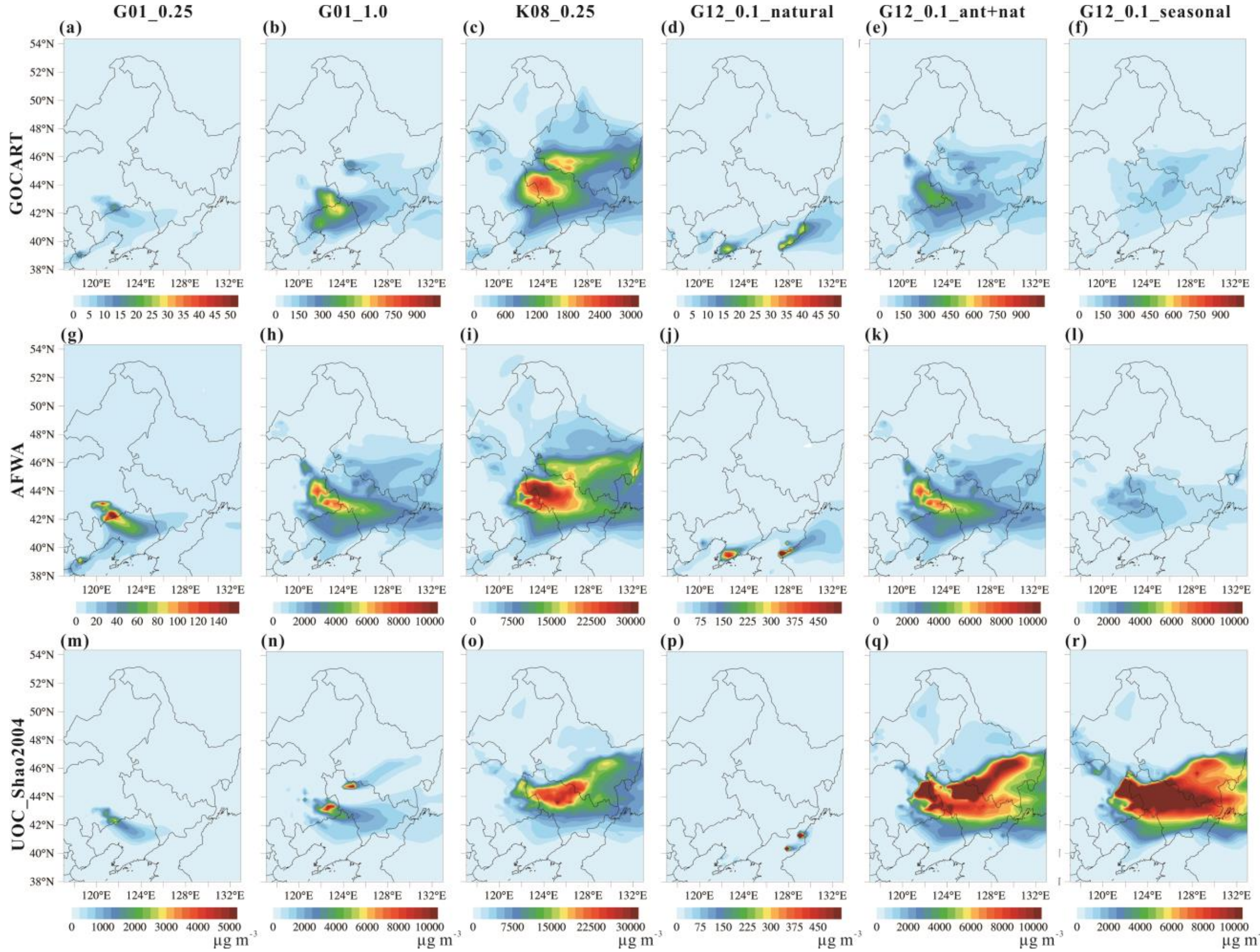


Figure 5. Daily mean PM_{10} distributions in NEC on May 5th, 2015 using GOCART, AFWA and UOC_~~Shao2011~~ Shao2004 with each source map. (**a**)–(**e**): GOCART, (**g**)–(**l**): AFWA, (**m**)–(**r**): UOC_~~Shao2011~~Shao2004.

3.3 Performance of CHIMERE v2017r4 dust simulation

The daily PM_{10} patterns simulated by CHIMERE v2017r4 with different dust schemes (AGO and KOK) and three erodible fraction algorithms are illustrated in Fig. 6. The distributions of AGO PM_{10} with three kinds of erodible fractions present similar patterns: two regions of higher PM_{10} concentration were seen in Horqin sandy land and Wulagai Gobi, respectively. The simulated dust showed its impact on the eastern areas like Jilin and northern Liaoning Province ~~trajectory was eastwards~~ (Fig. 6a~c), while northeastern NEC (such as eastern part of Heilongjiang Province) were also observed to be influenced by this dust episode~~compared to the observed northeastwards trajectory in~~ (Fig. 2). In comparison, there was only one source location, in Wulagai Gobi, with the KOK scheme (Fig. 6d~f), and the same dust source was also presented over

NEC in the global model (Fig. 2c in Kok et al., 2014b). For the simulated results using different threshold friction velocity algorithms, the area of dust source and dust intensity with SL were smaller than those with IW (Fig. 6g~i), indicating that the dust was more difficult to emit using SL. Nevertheless, the dust emissions in Wulagai Gobi were over-predicted. Observations showed relative lower PM_{10} and AOD in that area (Fig. 2).

The most striking ~~feature~~ discrepancy between~~of~~ the model results was in their concentration level. Daily PM_{10} with erodibility derived from USGS and a combination of USGS and MODIS in the source regions exceeded 1200 μg m^{-3}, and ranged from 100 to 800 μg m^{-3} in the transported areas. In comparison, the simulated concentrations with erodibility derived from MODIS were only about half of those values. In the KOK scheme, the simulated PM_{10} concentration was < 50 μg m^{-3} across the whole NEC area, thereby significantly deviating from its actual value. This ~~might~~ difference might~~ay~~ have arisen because the KOK scheme was mainly built on fragmentation theory (dust aggregates are fragmented by saltators into smaller particles and then emitted vertically to the atmosphere), which might be more suitable for desert land and barren land with lower cohesive energy. The strong underestimation of dust emission by the KOK scheme in NEC could be explained by the large areas of cropland with mollisol and grassland, yielding dust aggregates enriched in organic matter (Fan et al., 2010) that resist fragmentation.

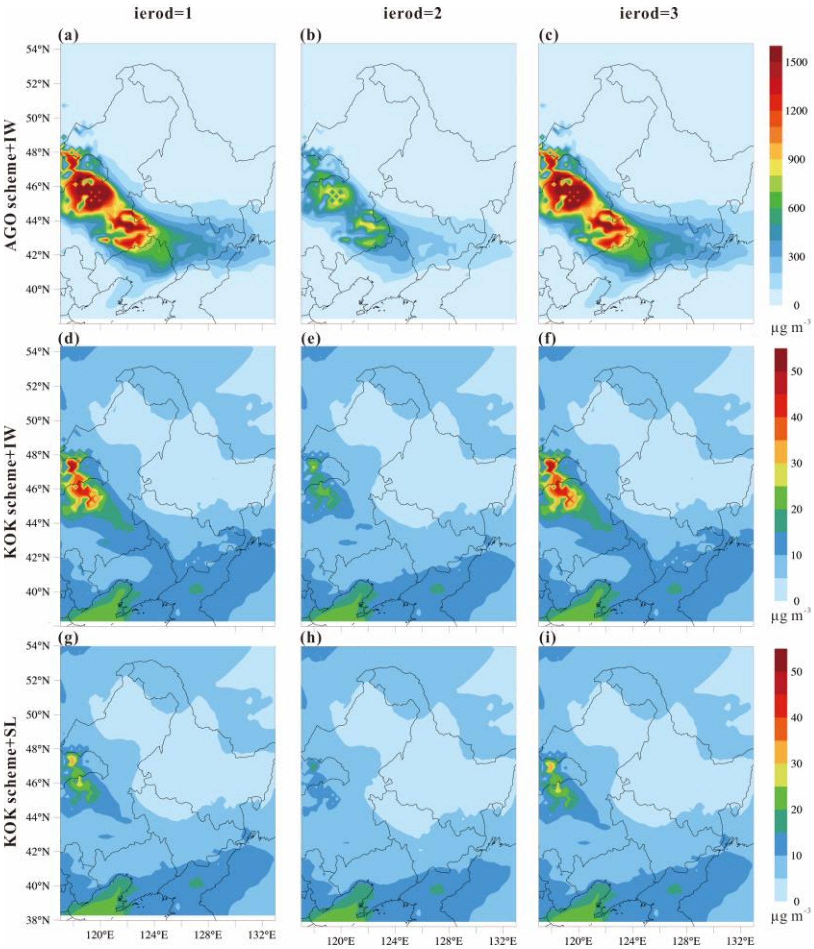


Figure 6. Daily mean PM_{10} distributions in NEC on May 5th, 2015 using AGO (with threshold friction velocity of IW) and KOK (with threshold friction velocity of IW and SL) with erodible fractions *ierod*=1, *ierod* =2 and *ierod* =3.

3.4 Performance of CMAQ v5.2.1 dust simulation

The simulated CMAQ v5.2.1 windblown dust emission is shown in Fig. 7. The distributions with and without dust emissions from cropland both presented the same spatial pattern for daily PM_{10} concentration. Dust was emitted mainly from the Horqin sandy land and a small area in western Jilin Province. ComparingWith further comparison to observations and simulated results of WRF-Chem and CHIMERE, the dust simulated by CMAQ was only short-distance transported southeastwards to parts of Liaoning and Jilin Province, yet had little influence on the areas of Heilongjiang Province, which could be explained as having the lowest simulated dust emissions. The simulated PM_{10} concentrations were about 60 (50) µg m^{-3} with (without) cropland dust emission in source areas, and only ranged from 10 to 20 µg m^{-3} in the transported areas. Comparing these two results, the contribution of anthropogenic wind-blown dust from cropland was only 10 µg m^{-3}, yet there were no further obvious differences.

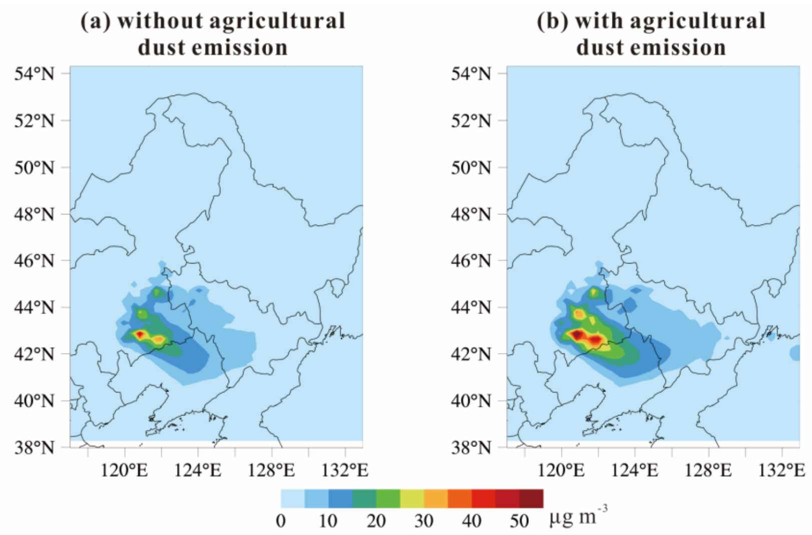


Figure 7. Daily mean PM_{10} distributions in NEC on May 5th, 2015 without (**a**) and with (**b**) agricultural dust emission.

To determine the reason for the underestimate of dust emission flux in CMAQ, the formula and source code of the latest dust emission scheme (LS99-FENGSHA) used in CMAQ version 5.2 were analyzed. According to Equation 13 and its description in Foroutan et al. (2017), the vertical-to-horizontal dust flux ratio (α) which determines the vertically transportable fraction of emitted dust particles is calculated via Equation 24 in Lu and Shao (1999), and parameters in this equation were defined according to Table 2 of Kang et al. (2011). The formula is expressed as follows:

$$\alpha = \frac{F}{Q} = \frac{C_\alpha g f \rho_b}{2p}\left(0.24 + C_\beta u_* \sqrt{\frac{\rho_p}{p}}\right) \tag{1}$$

where f is the fraction of fine particles contained in the soil volume, p is plastic pressure, in the range of 10^3~10^7 N m^{-2} (Gillett, 1977; Callebaut et al., 1985; Rice et al., 1997), ρ_b and ρ_p are the bulk soil and soil particle densities with unit of kg m^{-3}, g is the gravitational constant in m s^{-2}, u_* is friction velocity in m s^{-1}, and C_α and C_β are constants. Here the formula described in Lu and Shao (1999) is named as LS99 and a version of LS99 modified by Kang et al. (2011) and introduced in CMAQ since version 5.2 by Foroutan et al. (2017) is called F17. The formula involving p for calculating α ~~described in~~according to Shao (2004), namely S04, can be described as~~is~~:

$$\alpha = c_y \eta_{f,i}\left[(1-\gamma) + \gamma \frac{p_m(d_i)}{p_f(d_i)}\right] \frac{g}{u_*^2}\left(1 + 12u_*^2 \frac{\rho_b}{p}\left(1 + 14u_* \sqrt{\frac{\rho_b}{p}}\right)\right) \tag{2}$$

w~~W~~here $p_m(d_i)$ and $p_f(d_i)$ are respectively the fully and minimally disturbed dust fraction in bin d_i, and $\eta_{f,i}$ is the fully disturbed dust fraction. ρ_b=1000 kg m^{-3} is bulk soil density. γ is a function specified as $\gamma = \exp[-(u_* - u_{*t})^3]$ where u_{*t} is threshold friction velocity. c_y is a dimensionless coefficient which is set to be 1×10^{-5}, 4×10^{-5}, 5×10^{-5}, 3×10^{-4} for different soil textures and locations in Shao (2004); c_y is a tuning coefficient which is set to be 1×10^{-5}, 4×10^{-5}, 5×10^{-5}, 3×10^{-4} for different soil textures and locations in Shao (2004); then, values of soil plastic pressure p in the range of 10^{2} to 10^{4} N m^{-2} were obtained via matching with observed dust flux and friction velocities. This formula is now used in WRF-Chem v3.9.1. Note that the fitted c_y and p defined above could only be used in formula (2)S04 and not in formula (1)LS99 and F17 with different physical parameters. For example, the fitted value of 5000 for p (silty clay loam) in Table 3 of Shao (2004) was used as p of sand in Kang et al. (2011); moreover, the vertical flux should be calculated with the latter part of formula (2) as $p < 3 \times 10^4$, but the flux was calculated with formula (1) by Kang et al. (2011) and Foroutan et al. (2017). To correct the overestimated p was used in the vertical flux calculation of Lu and Shao (1999)LS99, Kang et al. (2011) reported that a modified C_α was recalculated based upon c_y (which is used in S04). However, to our knowledge, no method based on physical evidence is available to complete this conversion. Moreover, the source code of Shao_2004 in WRF--Chem only considered the former part of formula (2) for calculation of dust vertical flux, using uses prescribed values $p = 3 \times 10^4$ and $c_y = 1 \times 10^{-5}$ without considering the soil textures. As both of their values varied widely over soil types and locations, the mismatch in part of the study domain would lead to difference in magnitude, no matter in CMAQ or WRF-Chem.

In order to further verify the effects of modified C_α and p used by Kang et al. (2011) and Foroutan et al. (2017) on dust vertical flux, the values of α for soil texture of sand, loam and clay were calculated following Lu and Shao (1999) (LS99), Shao (2004) (S04), Foroutan et al. (2017) (F17) and the related formula in Marticorena and Bergametti (1995) (MB95, which was used in the original FENGSHA of CMAQ v5.0), along with measurements from laboratory experiments and field observations, as depicted in Fig. 8. Values of α for sandy soil calculated via four formulae all showed better agreement with observations (Fig. 8a). For loam and sandy clay loam soil, only the result of LS99 was able to match the observed level, while those of S04 and F17 were about two orders of magnitude smaller (Fig. 8b). When considering that the dominant soil textures were loam and clay loam in NEC, this explained the reason for the underestimated dust concentration ion occurred in CMAQ compared with WRF-Chem and CHIMERE. In addition, no α equation (shows in Figure S2, S3 and S5) method could reproduce the observed positive correlation between α and friction velocity (Fig. 8c). Furthermore, comparing to the sandblasting for the clay and clay loam, the dust in originated from aerodynamic entrainment (which was not taken into account by the present dust models) was significantly constituted (by up to 28.3% and 146.4%) relative to sandblasting for the clay and clay loam, respectively (Parajuli et al., 2016). The calculations of vertical dust flux should be further examined in the future to understand its contribution to model bias for key soil types.

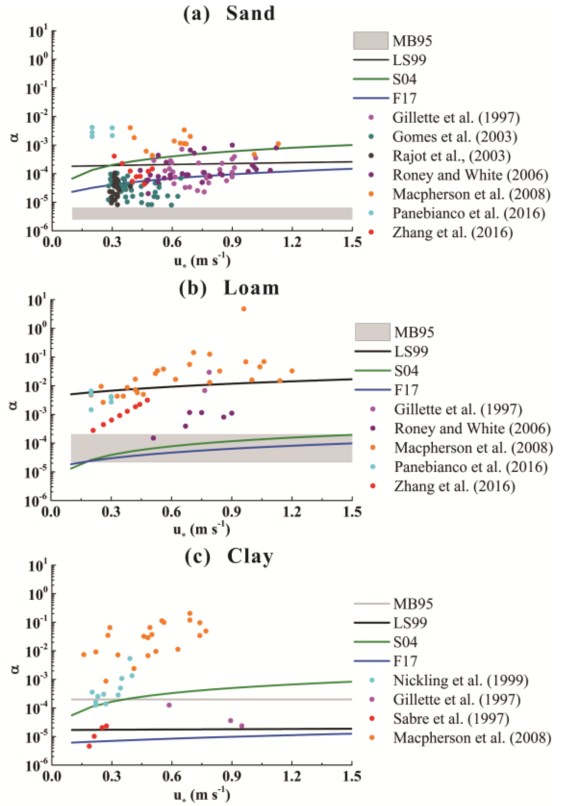
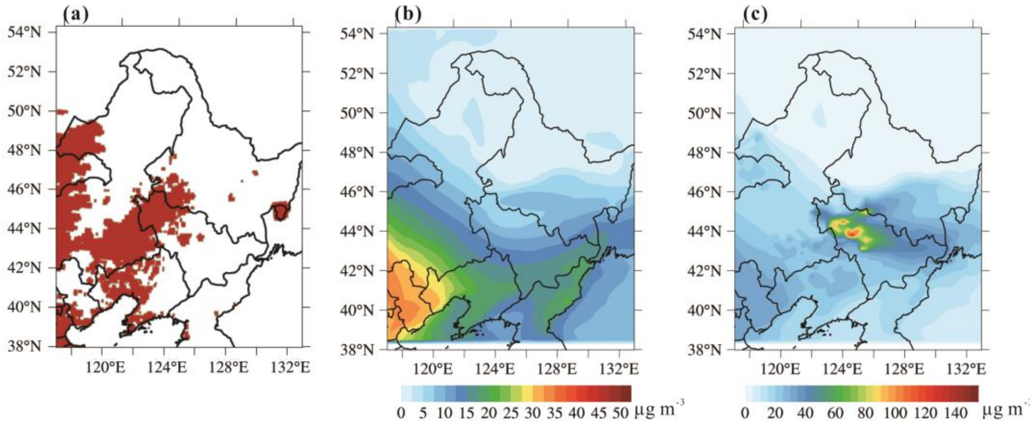


Figure 8. Vertical-to-horizontal dust flux ratio (α) for sand (**a**), loam (**b**) and clay (**c**) as a function of friction velocity (u_*) following Marticorena and Bergametti (1995) (MB95), Lu and Shao (1999) (LS99), Shao (2004) (S04) and Foroutan et al. (2017) (F17), and observations from the literature.

3.5 Performance of CAMx v6.50 dust simulation

As the dust mask used in CAMx showed no coverage in NEC area, the seasonal dust source map (G12_0.1_seasonal) was adapted to replace the original dust mask file as it had the best performance among those source maps in the WRF-Chem model (Fig. 5). The values in source map file were changed to 1 when the erodible fraction > 0 to fit the format of the dust mask file (Fig. 9a). Then the CAMx simulation was implemented and the daily averaged PM_{10} distribution on May 5th, 2015 is presented in Fig. 9b. It shows that the daily PM_{10} concentration simulated by CAMx ranged from 0 to 30 μg m^{-3} with high value in the southwest part of the simulated domain, and there was no dust emitting from any erodible area in NEC. A control simulation without dust emission was also conducted and the PM_{10} pattern was same with Fig. 2b. It means that no dust emission at all and CAMx model failed to reproduce this dust episode occurred in NEC.

Considering the dust mask had been eeupdated and the erodible areas were included in model, the poor performance of CAMx might result from the lower value of friction velocity. In the dust model of CAMx, the friction velocity is limited to a maximum value of 0.4 m s^{-1}, making it keep a low level comparing to the values of other models (Fig. S11). It was difficult to exceed the u_{*t} which was generally larger than 0.4 m s^{-1} (Fig. S12), so no dust emission occurred. Therefore, this limitation value was subsequently removed from the source code (wbdust.f90) and the simulation was conducted again. The distribution of simulated PM_{10} without the u_* limitation was presented in Fig. 9c. It shows that the dust was mainly from western Jilin Province near the Songnen sandy land and transported westward. This pattern could be also observed from ground observations (Fig. 2e). However, there was no simulated dust emitting from Horqin sandy land. Simulated PM_{10} concentrations were generally lower than the observations with about 120 μg m^{-3} in source areas and 10~50 μg m^{-3} in the transported areas. Compared with the simulation with u_* limitation, this result was obviously improved which indicated that the limitation value of u_* in CAMx needs further adjustment on region scale to improve its performance over the areas other than barren and sparsely vegetated area.



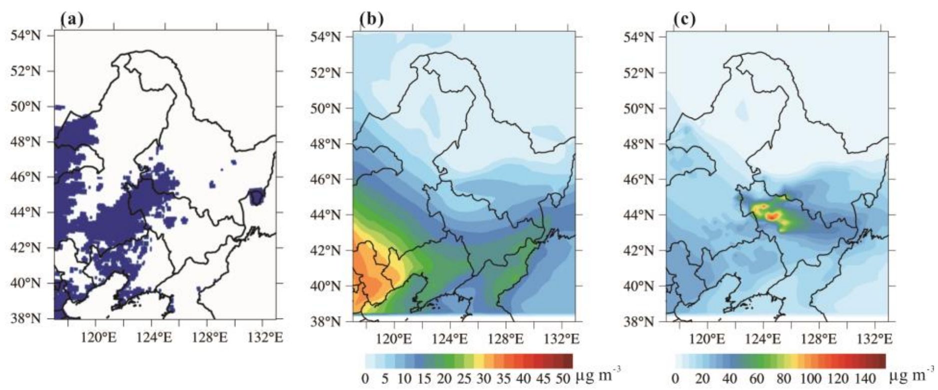


Figure 9. The substituted dust mask (**a**) and daily mean PM_{10} distributions with (**b**) and without (**c**) friction velocity limitation in NEC on May 5th, 2015.

3.~~5~~ 6 Inter-model Comparisons

The distribution and numerical values of PM_{10} in each simulation were described in Sections 3.1~3.~~4~~5, revealing remarkable differences between models. Most models simulated the primary dust source location (Horqin sandy land); however, many could not accurately represent other sources and the dust patterns in other parts of NEC. In this section, quantitative analyses are conducted to validate and evaluate the performances of different air quality models and dust schemes.

Considering the large discrepancies in simulated values (ranging from 10^0 to 10^4 μg m^{-3}), it is difficult to conduct the evaluation on all of the simulations at the same time. Therefore, in this section a scaling factor is applied to the model outputs of WRF-Chem v3.9.1, ~~and~~ CMAQ v5.2.1 and CAMx to allow a meaningful comparison against observed PM_{10} concentrations. The corresponding simulations and scaling factors are summarized in Table 3. Since the patterns modeled with CHIMERE v2017r4 using different erodible fractions were quite similar, here only the outputs simulated by AGO scheme with *ierod*=3 (mixed USGS and MODIS) were chosen for further validation. Subsequently, five statistical parameters (correlation coefficient (CORR), relative mean square error (RMSE), normalized standard deviation (NSD), normalized mean bias (NMB) and normalized mean error (NME)) for the hourly data of ~~12~~ 13 simulations and ~~the~~ observations ~~data~~ at 40 ground-based monitoring sites in NEC were calculated and averaged into four sub-areas of NEC (Fig. 2) for quantitative evaluation.

Table 3. Simulations used for validation and their corresponding tuning coefficients

No.	Simulation	Tuning coefficient
1	chem_gocart_g01	1
2	chem_gocart_k08	0.5
3	chem_gocart_g12	5
4	chem_afwa_g01	0.25
5	chem_afwa_k08	0.04
6	chem_afwa_g12	0.5
7	chem_s04~~11~~_g01	0.5
8	chem_~~s11~~s04_k08	0.05
9	chem_~~s11~~s04_g12	0.07
10	chim_ierod3	1
11	cmaq	70
12	cmaq_agland	70
13	camx	10

Note: ‘chem’ indicates WRF-Chem, ‘chim’ indicates CHIMERE, ~~and~~ ‘cmaq’ indicates CMAQ and ‘camx’ indicates CAMx. ‘gocart’, ‘afwa’ and ‘s11’ indicate the GOCART, AFWA and UOC_Shao2011 Shao2004 schemes, respectively. ‘g01’, ‘k08’ and ‘g12’ indicate the source maps of ‘G01_1.0’, ‘K08_0.25’ and ‘G12_0.1_seasonal’. ‘agland’ indicates that the agricultural dust emission was included in CMAQ.

A Taylor diagram (Taylor, 2001) comparing in-situ observations against simulated concentrations in CTA, SWA and NEA was is shown in Fig. 910. In this diagram, the NSD (ordinate) and CORR indicated each model’s ability to reproduce PM_{10} variability, while the RMSE (distance to point OBSabscissa) measures differences between the modeled and observed PM_{10} within the three sub-areas. WRF-Chem GOCART (labeled 1~3 in Fig. 910) yielded CORR values generally below 0.5 in three sub-areas and differed greatly between different dust source maps, while CORR values with the AFWA scheme (labeled 4~6) indicated stronger correlations in the CTA and NEA areas with values of 0.52~0.79 compared with only 0.17~0.46 in SWA. UOC_Shao2011 Shao2004 yielded the highest CORR values, of up to 0.82, among the four dust schemes in WRF-Chem, and the UOC_Shao2011 Shao2004 simulation with dust source map G12_0.1_seasonal showed the strongest correlation of all. CHIMERE and CMAQ yielded CORR values ranging from 0.43 to 0.76, with good correlations in all three areas. The CORR value of CAMx was similar with those of CHIMERE and CMAQ in NEA, but much lower in SWA (0.18).

The RMSEs and NSDs depended not only on the individual dust source maps, but also, strongly, on the tuning coefficients. Although the CORRs of WRF-Chem with GOCART were the lowest among all schemes, that combination yielded very low NSDs and RMSEs, showing that simulated concentrations were closer to the measurements. AFWA yielded relatively low NSDs and RMSEs in CTA and NEA, but the highest values in sub-area SWA. UOC_Shao2011 Shao2004 in CTA and NEA yielded the highest deviations. The NMBs and NMEs of the WRF-Chem simulations were lower in the CTA and SWA sub-areas than in the other two sub-areas (Fig. 10a11a~b). Finally, CHIMERE yielded the lowest NMB (near zero) and an NME <75%, while the NMB and NME for CMAQ were slightly larger. For CAMx, it had smaller bias and error in CTA and NEA while its NMB and NME were larger than CHIMERE and CMAQ in SWA.

In NWA, which is located in the upwind part of the study area, most simulations considerably overestimated dust: NSDs were 30~90 and CORR values were low (generally < 0.5 and < 0.2) in half the simulations). Thus, NSD NWA was not included in the Taylor diagram. More detailed statistics are provided in Table S1S8.

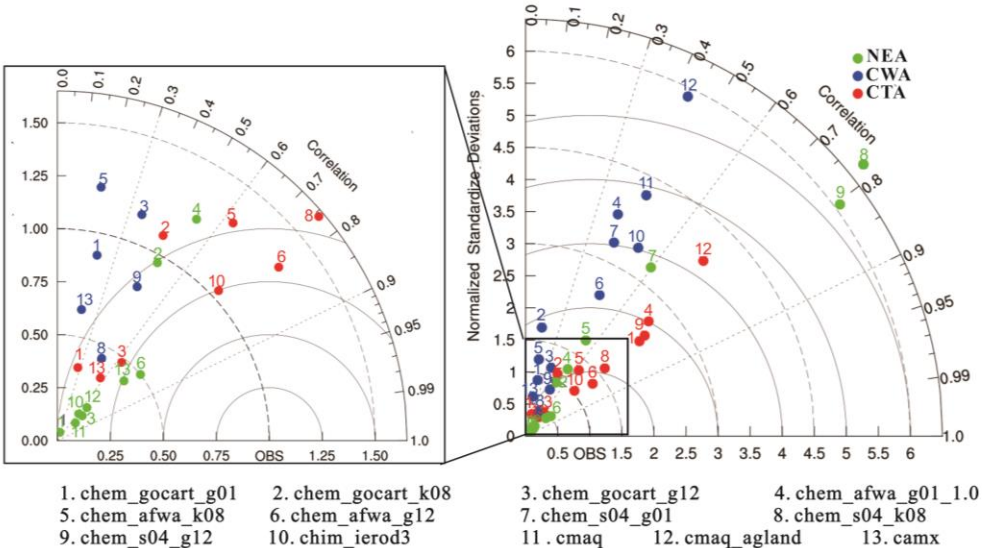


Figure ~~9~~10. Taylor diagram comparing the hourly PM_{10} concentrations of simulations with in-situ measurements for 3 regions of NEC described in Figure 2 (CTA, red circles; CWA, blue; NEA, purple). The numbers correspond to individual simulations summarized in Table 3.

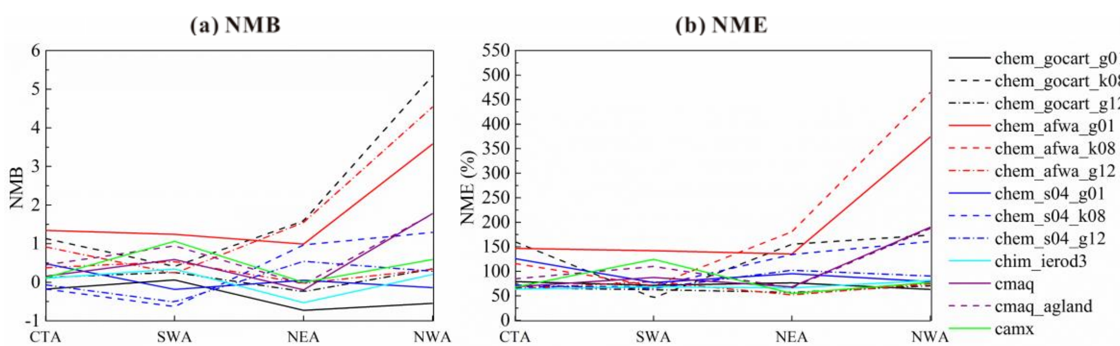


Figure ~~10~~11. Normalized mean simulation bias (**a**) and error (**b**) relative to measurements of PM_{10} during the dust period in the four regions of NEC shown in Fig. 2 (these are CTA, SWA, NEA and NWA). Simulation labels follow Figure 9.

To further illustrate the ability of each model to reproduce the temporal patterns of regional PM_{10}, time series of simulated hourly PM_{10} concentrations and in-situ measurements at four sites in CTA, and at one site in each of NEA, SWA and NWA, are shown in Fig. ~~11~~12. The models were able to reproduce the peak and high value period during May 5th in CTA and NEA (Fig. ~~11a~~12a~e); however, the simulations generally prolonged the period of high PM while often underestimating PM at sites with high dust intensities (e.g. Tongliao) and overestimating PM at lower-intensity sites such as Harbin, Shengyang and Jiamusi. One feature of note is that the PM_{10} concentrations simulated by UOC_~~Shao2011~~ Shao2004 were still overestimated at CTA sites even after parameter tuning (Fig. ~~11a~~12a~~~11d~~12d). CHIMERE underestimated the peaks at most sites and even yielded almost steady levels at Jiamusi, whereas observations showed a moderate dust peak (Fig. ~~11e~~12e). Figure 1~~1~~2f showed that the simulations poorly represented PM_{10} concentrations Jinzhou in SWA: despite showing two PM_{10} peaks on this high-dust day, with strong overestimates and earlier peak timing among the different air quality models. At Hulun Buir in NWA, there was no apparent similarity with measurements among any of the models except CAMx which presented relatively close fluctuating pattern to the observation (Fig. ~~11g~~12g).

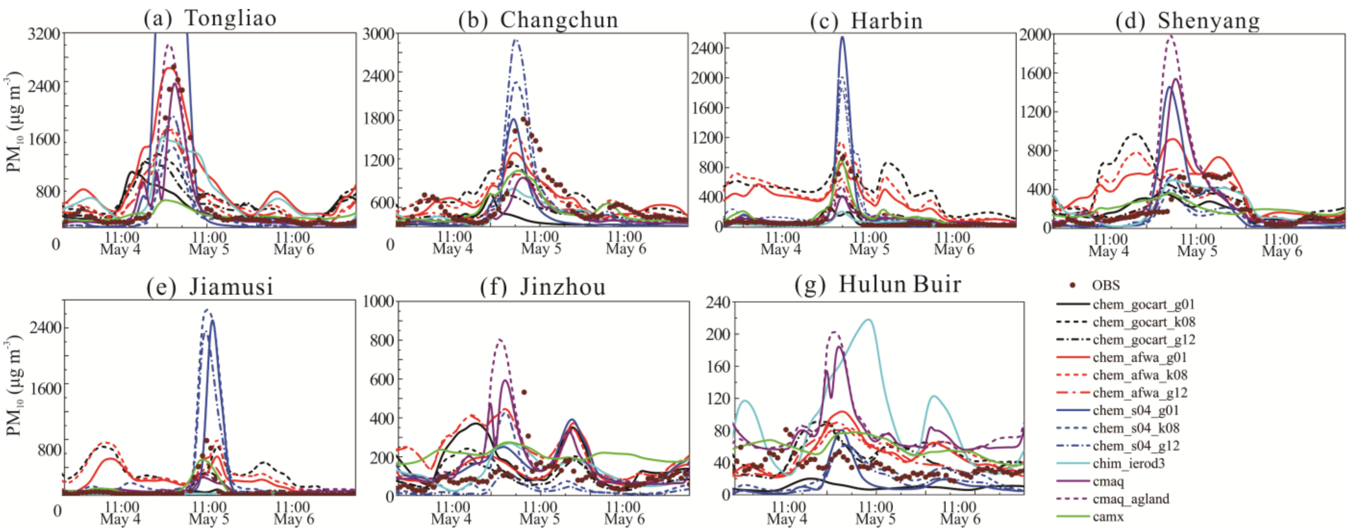


Figure ~~11~~12. Time series (UTC) of hourly PM_{10} for the simulations and measurements at 4 central sites (Tongliao (**a**), Changchun (**b**), Harbin (**c**) and Shenyang (**d**)); 1 NE site, Jiamusi (**e**); 1 SW site, Jinzhou (**f**); and 1 NW site, Hulun Buir (**g**) between May 4th and May 6th, 2015. Simulation labels follow Figure ~~9~~10.

The inter-model comparisons showed that WRF-Chem GOCART underestimated dust levels and yielded the lowest correlation with measurements among all the simulations. AFWA and UOC_~~Shao2011~~ Shao2004 performed well in high dust regions and with different source maps but PM_{10} concentrations were strongly overestimated. Different schemes in WRF-Chem presented correlations in NEC that are comparable with results in other areas, such as East Asia (Su and Fung, 2015), West Asia (Nabavi et al., 2017) and the Mediterranean (Flaounas et al., 2017; Rizza et al., 2017); however, most simulations by WRF-Chem displayed a considerable bias towards higher concentrations. In addition, the newly found source code errors after these simulations, such as bugs in GOCART gravitational settling (module_gocart_settling.F) and optical_prep_gocart routine in WRF-Chem v3.9.1 would also lead to the increase of bias and error.

The statistics for CHIMERE demonstrated its excellent performance in simulating the selected dust episode, with the lowest regional mean bias and error, as well as temporal variations that closely matched the observations (especially in southwestern NEC). Despite the generally good performance, two problems remain. First, the location of dust emissions, as dust emitted from Wulagai Gobi was not observed at the nearby monitoring sites and resulted in the discrepancy with measurements in western NEC. Second, the underestimated PM_{10} in central NEC may have arisen from the omission of sub-regional dust sources or an underestimation of wind speed in central NEC (Fig. S1aS7a) resulting in a lower friction velocity that was less likely to surpass the threshold value. Meanwhile, different algorithms for u_{*t} andMeanwhile, erroneous unit conversion of soil moisture from volumetric to mass percentage (just multiplies a value of 100 from unit of $m^3\ m^{-3}$ to kg kg^{-1} instead of via the equation showed in Figure S2) when calculating u_{*t} in the CHIMERE source code would also have caused a relatively high threshold to influence the dust emission. Moreover, different algorithms for u_{*t} resulted in significant differences in the simulated dust emission. For instance, the variations of u_* in CHIMERE and WRF-Chem are similar (Fig. S11a, b) during the dust episode, however, u_{*t} presented large discrepancies between another (Fig. S12). The value of u_{*t} in CHIMERE is in the upper level of the six kinds of u_{*t} used in the dust emission models while that of AFWA is much lower. This could also be one of the reasons for the overestimation in WRF-Chem AFWA. Besides the relationship between u_{*t} and u_*, the effect of roughness length should also be considered. The high-resolution roughness length data used in this model were from the GARLAP (Global Aeolian Roughness Lengths from ASCAT and PARASOL) dataset and derived from lidar and satellite observations; their relatively low values when compared to those obtained from land use data caused the simulated dust flux to be generally lower than that obtained with land use roughness length (Menut et al., 2013). In addition, different equations were used to calculate the kinetic energy when $u_{*t} < 0.27\ m\ s^{-1}$ and $0.27\ m\ s^{-1} < u_{*t} < 0.55\ m\ s^{-1}$ respectively (Alfaro et al., 1997), however, it was not exactly calculated according to this method in CHIMERE which might lead to deviations of the results when $u_* > 0.27\ m\ s^{-1}$.

Simulations by CMAQ and CHIMERE yielded comparable correlations with observations; however, the largest tuning coefficient (with a value of 70) showed that CMAQ would seriously underestimate dust levels if used without this adjustment. Besides the reasons suggested in Section 3.4, the treatment of u_{*t} as a constant in the original CMAQ FENGSHA and the algorithm of Shao and Lu (2000) in CMAQ LS99-FENGSHA (Table S2S9) would also result in significant differences when calculating horizontal and vertical dust fluxes. In addition, using the dynamic roughness length term when calculating u_* in CMAQ led to lower u_* (with mean value of $0.39\ m\ s^{-1}$, Fig. S11c) comparing to those in WRF-Chem and CHIMERE ($0.58\ m\ s^{-1}$). This would be another reason for its underestimated results.

The dust simulating performance of CAMx with substituted dust mask and no u_* limitation considerably improved comparing to its default configuration. Its correlations and errors to the observations could keep up with the results of CHIMERE and CMAQ, however, failure in reproducing the dust emission in Horqin sandy land accounted for the bad performance over SWA.

Finally, various uncertainties were demonstrated in the physically-based dust emission schemes, and a better understanding of these would aid future model development and enhance the accuracy of dust predictions. First of all, the dust emission research has diverged into several paths and many individual algorithms based on field observations in specific areas have been developed for particular sectors. Combining these algorithms to form a complete dust emission scheme for application in other regions would introduce unavoidable uncertainty. Several inputs are required in the dust schemes; in general, these are wind speed, precipitation, land surface characteristics (e.g. land used type, soil texture, soil moisture, surface roughness, dust source map, vegetation and snow cover), such that the accuracy and quality of external input data and meteorological models cause uncertainties to accumulate in the model prediction. Note that further validation of the meteorological simulating results and the influence of atmospheric conditions on dust emission need to be conducted in the future study. Furthermore, ~~T~~the choices of coefficient/constant values (such as u_{*t}), which are based on the familiarity and experience of the users in the study domain, introduce additional uncertainty. In addition, there is uncertainty related to the spatial resolution of air quality models at regional and global scales (Foroutan et al., 2017). Most importantly, recent studies have suggested that dust emissions by direct aerodynamic entrainment (or convective turbulent dust emission) are about one third of sandblasting dust emissions (Li et al., 2014; Parajuli et al., 2016; Ju et al., 2018), yet the parameterizing of direct aerodynamic entrainment processes has not been implemented in the dust schemes of air quality models. This should be addressed in future work.

4 Summary and conclusion

In this study, we quantitatively evaluated the performance of several physically-based dust emission ~~schemes~~ models in the WRF-Chem v3.9.1, CHIMERE v2017r4, CMAQ v5.2.1 and CAMx v6.50 air quality models to simulate a dust event in Northeastern China. Four dust schemes and four ~~newly~~additionally-introduced dust source maps in WRF-Chem v3.9.1, two schemes in CHIMERE and CMAQ, and one scheme in CAMx were tested for this dust event during May 4th~6th, 2015. For simulations with high overestimates or underestimates, scaling factors were introduced to minimize the bias between the model and observations. Northeastern China was divided into four sub-areas for further quantitative comparisons of the simulations and observed PM_{10} concentrations. These models used either dust source/mask map (WRF-Chem and CAMx), or erodible fraction (CHIMERE and CMAQ) to determine whether dust is emitted from a grid cell. Different algorithms for threshold friction velocity resulted in significant differences in the simulated dust concentration and spatial distribution. Converting the observed wind speed to near-surface friction velocity was one of the most important sources of simulated uncertainties. We demonstrated that a more accurate ratio of horizontal-to-vertical dust flux for each soil texture should be obtained in future field works.

Our evaluation revealed that the PM_{10} simulated by each dust scheme in WRF-Chem yielded similar spatial patterns. AFWA and UOC_~~Shao2011~~ Shao2004 yielded higher correlations than GOCART, but both considerably overestimated surface dust concentrations. Of the dust source maps applied in WRF-Chem, the default G01_0.25 was incapable of reproducing the dust patterns, indicating it was not applicable in Northeastern China. The five newly-introduced dust source maps displayed better performance, and the simulations with G12_0.1_seasonal presented the best relationships with ground-based observations. CHIMERE AGO and CMAQ LS99-FENGSHA were able to reproduce the spatial distribution of the dust plume, but with incorrect dust emission sources (such as Wulagai Gobi in CHIMERE) and strong underestimates in CMAQ. All simulations performed best near ~~to~~the dust source areas and degraded~~but more poorly~~ in accuracy with downstream advection~~dust transport areas~~. Statistical parameters indicated that the strengths of model performances decreased in the order CTA>NEA>SWA>NWA.

In general, if the ~~parameter~~ numerical tuning was included, WRF-Chem AFWA with dust source map G12_0.1_seasonal yielded the best performance among all the simulations, followed by WRF-Chem UOC_~~Shao2011~~Shao2004, CHIMERE AGO, CMAQ and WRF-Chem GOCART. Without tuning the ~~parameter~~concentration, only CHIMERE demonstrated significant correlation with relatively low bias and error, but still presented problems such as the misplaced dust sources and notable underestimates. The tuning coefficient for the outputs of WRF-Chem and CMAQ indicate the dust flux needs to be scaled during the calculation to improve their performance. A dust mask including dust emissions from regions not classified as "barren or sparsely vegetated"~~with higher spatial resolution~~ in CAMx should be developed by refining the land cover mask in future works.~~–~~ and the algorithm of friction velocity needs adjustment and improvement as well. Source code errors in the air quality models need to be further debugged. In addition, a physically-based direct aerodynamic dust entrainment scheme or empirical parameterization should be implemented in the models to enhance the regional air quality forecast ability for particulates.

Code availability. WRF-Chem is an open-source community model. The source code is available at http://www2.mmm.ucar.edu/wrf/users/download/get_source.html. The source code of CHIMERE 2017a along with the corresponding technical documentation can be obtained from the CHIMERE web site at http://www.lmd.polytechnique.fr/chimere/. CMAQ model documentation and released versions of the source code are available on the US EPA modeling site https://www.cmascenter.org/. The source code of CAMx model is available at http://www.camx.com/. All related source codes with modified dust emission files, configuration information and the namelist files for four air quality models, and even the pre- and post-processing scripts used for this study are available online via ZENODO (https://doi.org/10.5281/zenodo.3376774).

Author contributions. MS, XZ and CG performed the majority of the source code reconfiguration of WRF-Chem, CHIMERE CMAQ and CAMx, and initially designed the numerical simulations to carry them out. ~~and CG performed the majority of the source code reconfiguration of WRF-Chem, CHIMERE, CMAQ and CAMx model. MS and XZ initially designed the numerical simulations with help from~~ DQT, AX, WG and CX provided help for the simulation designation. LH provided support for conducting the CAMx model. , HZ and SZ provided advices on the selection and usage of observational data.to carry them out. MS, XZ and DQT led the analysis of the simulations, and GW, SIE, LH, XW, XL and MD provided professional advices. SM and XZ wrote the paper and all authors read, revised, and approved the final manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was financially supported by the National Natural Science Foundation of China (NSFC) (No. 41571063 and 41771071), National key R&D Plan of China (No. 2017YFC0212304) and Hundred-Talent Program (Chinese Academy of Sciences, No. Y8H1021001).The authors are grateful to the website of national air quality history database for collecting and maintaining the ground-based air quality data., and appreciate We also thank NASA for providing the MODIS and CALIPSO datasets, as well as for the datasets maintenance and availability. We also thank to the two anonymous referee reviewers and editors of Samuel Remy and David Ham for their value comments.advice that have helped us to improve our manuscript

References

Alfaro, S. C., Gaudichet, A., Gomes, L. and Maillé, M.: Modeling the size distribution of a soil aerosol produced by sandblasting, J. Geophys. Res. Atmos., 102(D10), 11239–11249, doi:10.1029/97JD00403, 1997.

Astitha, M., Lelieveld, J., Abdel Kader, M., Pozzer, A. and Meij, A. de: Parameterization of dust emissions in the global atmospheric chemistry-climate model EMAC: impact of nudging and soil properties, Atmos. Chem. Phys., 12(22), 11057–11083, doi: 10.5194/acp-12-11057-2012, 2012.

Bagnold, R. A.: The physics of blown sand and desert dunes, Chapmann and Hall, Methuen, London, 265 pp., 1941.

Basart, S., Pérez, C., Nickovic, S., Cuevas, E. and Baldasano, J.: Development and evaluation of the BSC-DREAM8b dust regional model over Northern Africa, the Mediterranean and the Middle East, Tellus B Chem. Phys. Meteorol., 64(1), 18539, doi:10.3402/tellusb.v64i0.18539, 2012.

Bessagnet, B., Menut, L., Colette, A., Couvidat, F., Dan, M., Mailler, S., Létinois, L., Pont, V. and Rou?l, L.: An Evaluation of the CHIMERE Chemistry Transport Model to Simulate Dust Outbreaks across the Northern Hemisphere in March 2014, Atmosphere (Basel)., 8(12), 251, doi:10.3390/atmos8120251, 2017.

Byun, D. and Schere, K. L.: Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system, Appl. Mech. Rev., 59(2), 51–77, doi:10.1115/1.2128636, 2006.

Cahill, T. A.: Generation, characterization and transport of Owens (Dry) Lake dusts, The Division., doi:10.1029/97JD01967, 1994.

Callebaut, F., Gabriels, D., Minjauw, W. and De Boodt, M.: Determination of soil surface strength with a needle-type penetrometer, Soil tillage Res., 5(3), 227–245, doi:10.1016/0167-1987(85)90017-0, 1985.

Cakmur, R. V., R. L. Miller, and O. Torres: Incorporating the effect of small-scale circulations upon dust emission in an atmospheric general circulation model, J. Geophys. Res., 109, D07201, doi:10.1029/2003JD004067, 2004.

Cao, H., Amiraslani, F., Liu, J. and Zhou, N.: Identification of dust storm source areas in West Asia using multiple environmental datasets, Sci. Total Environ., 502, 224–235, doi:10.1016/j.scitotenv.2014.09.025, 2015.

Chappell, A., Webb, N. P., Viscarra Rossel, R. A. and Bui, E.: Australian net (1950s–1990) soil organic carbon erosion: implications for CO_2 emission and land–atmosphere modelling, Biogeosciences, 11(18), 5235–5244, doi:10.5194/bg-11-5235-2014, 2014.

Darmenova, K., Sokolik, I. N., Shao, Y., Marticorena, B. and Bergametti, G.: Development of a physically based dust emission module within the Weather Research and Forecasting (WRF) model: Assessment of dust emission parameterizations and input parameters for source regions in Central and East Asia, J. Geophys. Res. Atmos., 114(D14), doi:10.1029/2008JD011236, 2009.

Dong, X., Fu, J. S., Huang, K., Tong, D. and Zhuang, G.: Model development of dust emission and heterogeneous chemistry within the Community Multiscale Air Quality modeling system and its application over East Asia, Atmos. Chem. Phys., doi:10.5194/acp-16-8157-2016, 2016.

Evan, A. T., Fiedler, S., Zhao, C., Menut, L., Schepanski, K., Flamant, C. and Doherty, O.: Derivation of an observation-based map of North African dust emission, Aeolian Res., 16, 153–162, doi:10.1016/j.aeolia.2015.01.001, 2015.

Fan, R., Liang, A., Yang, X., Zhang, X., Shen, Y. and Shi, X.: Effects of tillage on soil aggregates in black soils in Northeast China., Sci. Agric. Sin., 43(18), 3767–3775, 2010.

Formenti, P., Schütz, L., Balkanski, Y., Desboeufs, K., Ebert, M., Kandler, K., Petzold, A., Scheuvens, D., Weinbruch, S. and Zhang, D.: Recent progress in understanding physical and chemical properties of African and Asian mineral dust, Atmos. Chem. Phys., 11(16), 8231–8256, doi:10.5194/acp-11-8231-2011, 2011.

Flaounas, E., Kotroni, V., Lagouvardos, K., Klose, M., Flamant, C., and Giannaros, T. M.: Sensitivity of the WRF-Chem (V3.6.1) model to different dust emission parametrisation: assessment in the broader Mediterranean region, Geosci. Model Dev., 10, 2925-2945, https://doi.org/10.5194/gmd-10-2925-2017, 2017.

Fécan, F., Marticorena, B., and Bergametti, G.: Parametrization of the increase of the aeolian erosion threshold wind friction velocity due to soil moisture for arid and semi-arid areas, Ann. Geophys., 17, 149–157, https://doi.org/10.1007/s00585-999-0149-7, 1999.

Foroutan, H., Young, J., Napelenok, S., Ran, L., Appel, K. W., Gilliam, R. C. and Pleim, J. E.: Development and evaluation of a physics‐based windblown dust emission scheme implemented in the CMAQ modeling system, J. Adv. Model. Earth Syst., 9(1), 585–608, doi:10.1002/2016MS000823, 2017.

Gillette, D. A. and Passi, R.: Modeling dust emission caused by wind erosion, J. Geophys. Res. Atmos., 93(D11), 14233–14242, doi:10.1029/JD093iD11p14233, 1988.

Gillette, D. A., Fryrear, D. W., Gill, T. E., Ley, T., Cahill, T. A. and Gearhart, E. A.: Relation of vertical flux of particles smaller than 10μm to total aeolian horizontal mass flux at Owens Lake, J. Geophys. Res. Atmos., 102(D22), 26009–26015, doi:10.1029/97JD02252, 1997.

Gillette, D. A., Herbert, G., Stockton, P. H. and Owen, P. R.: Causes of the fetch effect in wind erosion, Earth Surf. Process. Landforms, 21(7), 641–659, doi:10.1002/(SICI)1096-9837(199607), 1996.

Gillette, D. A.: Fine particulate emissions due to wind erosion, Trans. ASAE, 20(5), 890–897, 1977.

Ginoux, P., Chin, M., Tegen, I., Prospero, J. M., Holben, B., Dubovik, O. and Lin, S.: Sources and distributions of dust aerosols simulated with the GOCART model, J. Geophys. Res. Atmos., 106(D17), 20255–20273, doi:10.1029/2000JD000053, 2001.

Ginoux, P., Prospero, J. M., Gill, T. E., Hsu, N. C. and Zhao, M.: Global‐scale attribution of anthropogenic and natural dust sources and their emission rates based on MODIS Deep Blue aerosol products, Rev. Geophys., 50(3), doi:10.1029/2012RG000388, 2012.

Gomes, L., Rajot, J. L., Alfaro, S. C. and Gaudichet, A.: Validation of a dust production model from measurements performed in semi-arid agricultural areas of Spain and Niger, Catena, 52(3–4), 257–271, doi:10.1016/S0341-8162(03)00017-1, 2003.

Goudie, A. S.: Desert dust and human health disorders, Environ. Int., 63, 101–113, doi:10.1016/j.envint.2013.10.011, 2014.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C. and Eder, B.: Fully coupled “online” chemistry within the WRF model, Atmos. Environ., 39(37), 6957–6975, doi:10.1016/j.atmosenv.2005.04.027, 2005.

Guan, X., Huang, J., Zhang, Y., Xie, Y. and Liu, J.: The relationship between anthropogenic dust and population over global semi-arid regions, Atmos. Chem. Phys., 16(8), 5159–5169, doi:10.5194/acp-16-5159-2016, 2016.

Hodzic, A., Bessagnet, B. and Vautard, R.: A model evaluation of coarse-mode nitrate heterogeneous formation on dust particles, Atmos. Environ., 40(22), 4158–4171, doi:10.1016/j.atmosenv.2006.02.015, 2006.

Huneeus, N., Basart, S., Fiedler, S., Morcrette, J.-J., Benedetti, A., Mulcahy, J., Terradellas, E., Garcia-Pando, C. P., Pejanovic, G. and Nickovic, S.: Forecasting the northern African dust outbreak towards Europe in April 2011: a model intercomparison, Atmos. Chem. Phys., 16(8), doi:10.5194/acp-16-4967-2016, 2016.

Huneeus, N., Schulz, M., Balkanski, Y., Griesfeller, J., Prospero, M. J., Kinne, S., Bauer, S., Boucher, O., Chin, M. and Dentener, F.: Global dust model intercomparison in AeroCom phase I, Atmos. Chem. Phys., 11, 7781–7816, doi:10.5194/acp-11-7781-2011, 2011.

Iversen, J. D. and White, B. R.: Saltation threshold on earth, mars and venus, Sedimentology, 29(1), 111–119, doi:10.1111/j.1365-3091.1982.tb01713.x, 1982.

Jish Prakash, P., Stenchikov, G. L., Kalenderski, S., Osipov, S. and Bangalath, H. K.: The impact of dust storms on the Arabian Peninsula and the Red Sea, Atmos. Chem. Phys., doi:10.5194/acp-15-199-2015, 2015.

Ju, T., Li, X., Zhang, H., Cai, X. and Song, Y.: Comparison of two different dust emission mechanisms over the Horqin Sandy Land area: Aerosols contribution and size distributions, Atmos. Environ., 176, 82–90, doi:10.1016/j.atmosenv.2017.12.017, 2018.

Kang, J., Yoon, S., Shao, Y. and Kim, S.: Comparison of vertical dust flux by implementing three dust emission schemes in WRF/Chem, J. Geophys. Res. Atmos., 116(D9), doi:10.1029/2010JD014649, 2011.

Kim, D., Kemp, E. and Chin, M.: A brief description of new dust source functions in NU-WRF version 7, NASA Goddard Sp. Flight Center, 4pp, 2014.

Klingmüller, K., Metzger, S., Abdelkader, M., Karydis, V. A., Stenchikov, G. L., Pozzer, A. and Lelieveld, J.: Revised mineral dust emissions in the atmospheric chemistry–climate model EMAC (MESSy 2.52 DU_Astitha1 KKDU2017 patch), Geosci. Model Dev., 11, 989–1008, doi:10.5194/gmd-11-989-2018, 2018.

Kok, J. F., Albani, S., Mahowald, N. M. and Ward, D. S.: An improved dust emission model–Part 2: Evaluation in the Community Earth System Model, with implications for the use of dust source functions, Atmos. Chem. Phys., 14(23), 13043–13061, doi:10.5194/acp-14-13043-2014, 2014b.

Kok, J. F., Mahowald, N. M., Fratini, G., Gillies, J. A., Ishizuka, M., Leys, J. F., Mikami, M., Park, M.-S., Park, S.-U. and Van Pelt, R. S.: An improved dust emission model–Part 1: Model description and comparison against measurements, Atmos. Chem. Phys., 14(23), 13023–13041, doi:10.5194/acp-14-13023-2014, 2014a.

Koven, C. D. and Fung, I.: Identifying global dust source areas using high‐resolution land surface form, J. Geophys. Res. Atmos., 113(D22), doi:10.1029/2008JD010195, 2008.

Klose, M., Shao, Y., Li, X., Zhang, H., Ishizuka, M., Mikami, M., and Leys, J. F.: Further development of a parameterization for convective turbulent dust emission and evaluation based on field observations, J. Geophys. Res.-Atmos., 119, 10441–10457, https://doi.org/10.1002/2014JD021688, 2014.

Laurent, B., Marticorena, B., Bergametti, G., Chazette, P., Maignan, F. and Schmechtig, C.: Simulation of the mineral dust emission frequencies from desert areas of China and Mongolia using an aerodynamic roughness length map derived from the POLDER/ADEOS 1 surface products, J. Geophys. Res. Atmos., 110(D18), doi:10.1029/2004JD005013, 2005.

LeGrand, S. L., Polashenski, C., Letcher, T. W., Creighton, G. A., Peckham, S. E., and Cetola, J. D.: The AFWA dust emission scheme for the GOCART aerosol model in WRF-Chem v3.8.1, Geosci. Model Dev., 12, 131-166, https://doi.org/10.5194/gmd-12-131-2019, 2019.

Li, J., Wang, Z., Zhuang, G., Luo, G., Sun, Y. and Wang, Q.: Mixing of Asian mineral dust with anthropogenic pollutants over East Asia: a model case study of a super-duststorm in March 2010, Atmos. Chem. Phys., 12(16), 7591–7607, doi:10.5194/acp-12-7591-2012, 2012.

Li, X. L., Klose, M., Shao, Y. and Zhang, H. S.: Convective turbulent dust emission (CTDE) observed over Horqin Sandy Land area and validation of a CTDE scheme, J. Geophys. Res. Atmos., 119(16), 9980–9992, doi:10.1002/2014JD021572, 2014.

Liora, N., Poupkou, A., Giannaros, T. M., Kakosimos, K. E., Stein, O. and Melas, D.: Impacts of natural emission sources on particle pollution levels in Europe, Atmos. Environ., 137, 171–185, doi:10.1016/j.atmosenv.2016.04.040, 2016.

Lu, H. and Shao, Y.: A new model for dust emission by saltation bombardment, J. Geophys. Res. Atmos., 104(D14), 16827–16842, doi:10.1029/1999JD900169, 1999.

Macpherson, T., Nickling, W. G., Gillies, J. A. and Etyemezian, V.: Dust emissions from undisturbed and disturbed supply‐limited desert surfaces, J. Geophys. Res. Earth Surf., 113(F2), doi:10.1029/2007JF000800, 2008.

Mailler, S., Menut, L., Khvorostyanov, D., Valari, M., Couvidat, F., Siour, G., Turquety, S., Briant, R., Tuccella, P. and Bessagnet, B.: CHIMERE-2017: from urban to hemispheric chemistry-transport modeling, Geosci. Model Dev., 10(6), 2397–2423, doi:10.5194/gmd-10-2397-2017, 2017.

Marticorena, B. and Bergametti, G.: Modeling the atmospheric dust cycle: 1. Design of a soil-derived dust emission scheme, J. Geophys. Res.-Atmos., 100, 16415–16430, https://doi.org/10.1029/95JD00690, 1995.

Marticorena, B., Bergametti, G., Aumont, B., Callot, Y., N'doumé, C., and Legrand, M.: Modeling the atmospheric dust cycle: 2. Simulation of Saharan dust sources, J. Geophys. Res.-Atmos., 102, 4387–4404, https://doi.org/10.1029/96JD02964, 1997.

Manders-Groot, A. M. M., Segers, A. J., Jonkers, S., Schaap, M., Timmermans, R., Hendriks, C., Sauter, F., Kruit, R. W., van der Swaluw, E. and Eskes, H.: LOTOS-EUROS v2. 0 reference guide, TNO Rep. TNO2016, 10898, 2016.

Mendez, M. J. and Buschiazzo, D. E.: Wind erosion risk in agricultural soils under different tillage systems in the semiarid Pampas of Argentina, Soil Tillage Res., 106(2), 311–316, doi:10.1016/j.still.2009.10.010, 2010.

Menut, L., Bessagnet, B., Khvorostyanov, D., Beekmann, M., Blond, N., Colette, A., Coll, I., Curci, G., Foret, G. and Hodzic, A.: CHIMERE 2013: a model for regional atmospheric composition modelling, Geosci. Model Dev., 6(4), 981–1028, doi:10.5194/gmd-6-981-2013, 2014.

Menut, L., Pérez, C., Haustein, K., Bessagnet, B., Prigent, C. and Alfaro, S.: Impact of surface roughness and soil texture on mineral dust emission fluxes modeling, J. Geophys. Res. Atmos., 118(12), 6505–6520, doi:10.1002/jgrd.50313, 2013.

Menut, L., Schmechtig, C. and Marticorena, B.: Sensitivity of the sandblasting flux calculations to the soil size distribution accuracy, J. Atmos. Ocean. Technol., 22(12), 1875–1884, doi:10.1175/JTECH1825.1, 2005.

Mokhtari, M., Gomes, L., Tulet, P. and Rezoug, T.: Importance of the surface size distribution of erodible material: an improvement on the Dust Entrainment And Deposition (DEAD) Model, Geosci. Model Dev., 5, 581–598, doi: 10.5194, 2012.

Munkhtsetseg, E., Shinoda, M., Ishizuka, M., Mikami, M., Kimura, R. and Nikolich, G.: Anthropogenic dust emissions due to livestock trampling in a Mongolian temperate grassland, Atmos. Chem. Phys., 17(18), 11389–11401, doi:10.5194/acp-17-11389-2017, 2017.

Nabavi, S. O., Haimberger, L. and Samimi, C.: Sensitivity of WRF-chem predictions to dust source function specification in West Asia, Aeolian Res., 24, 115–131, doi:10.1016/j.aeolia.2016.12.005, 2017.

Nickling, W. G., McTainsh, G. H. and Leys, J. F.: Dust emissions from the Channel Country of western Queensland, Australia, Zeitschrift fur Geomorphol. Suppl., 1–17, 1999.

Owen, P. R.: Saltation of uniform grains in air, J. Fluid Mech., 20(2), 225–242, doi:10.1017/S0022112064001173, 1964.

Panebianco, J. E., Mendez, M. J. and Buschiazzo, D. E.: PM_{10} emission, sandblasting efficiency and vertical entrainment during successive wind-erosion events: a wind-tunnel approach, Boundary-layer Meteorol., 161(2), 335–353, doi:10.1007/s10546-016-0172-7 , 2016.

Parajuli, S. P., Zobeck, T. M., Kocurek, G., Yang, Z. and Stenchikov, G. L.: New insights into the wind‐dust relationship in sandblasting and direct aerodynamic entrainment from wind tunnel experiments, J. Geophys. Res. Atmos., 121(4), 1776–1792, doi:10.1002/2015JD024424, 2016.

Prospero, J. M. and Lamb, P. J.: African droughts and dust transport to the Caribbean: Climate change implications, Science (80-.)., 302(5647), 1024–1027, doi:10.1126/science.1089915, 2003.

Pryor, S., J. Schoof, and R. Barthelmie: Empirical downscaling of wind speed probability distributions, J. Geophys. Res., 110, D19109, doi: 10.1029/2005JD005899, 2005.

Rajot, J. L., Alfaro, S. C., Gomes, L. and Gaudichet, A.: Soil crusting on sandy soils and its influence on wind erosion, Catena, 53(1), doi:10.1016/S0341-8162(02)00201-1, 1–16, 2003.

Raupach, M.: Drag and drag partition on rough surfaces, Bound-Lay. Meteor., 60, 375–395, https://doi.org/10.1007/BF00155203, 1992.

Revel-Rolland, M., De Deckker, P., Delmonte, B., Hesse, P. P., Magee, J. W., Basile-Doelsch, I., Grousset, F. and Bosch, D.: Eastern Australia: a possible source of dust in East Antarctica interglacial ice, Earth Planet. Sci. Lett., 249(1–2), 1–13, doi:10.1016/j.epsl.2006.06.028, 2006.

Ridley, D. A., Heald, C. L., Kok, J. F. and Zhao, C.: An observationally constrained estimate of global dust aerosol optical depth, Atmos. Chem. Phys., 16(23), 15097–15117, doi:10.5194/acp-16-15097-2016, 2016.

Rizza, U., Barnaba, F., Miglietta, M. M., Mangia, C., Di Liberto, L., Dionisi, D., Costabile, F., Grasso, F. and Gobbi, G. P.: WRF-Chem model simulations of a dust outbreak over the central Mediterranean and comparison with multi-sensor desert dust observations, Atmos. Chem. Phys., 17(1), 93, doi:10.5194/acp-17-93-2017, 2017.

Rizza, U., Anabor, V., Mangia, C., Miglietta, M. M., Degrazia, G. A., and Passerini, G.: WRF-Chem Simulation of a saharan dust outbreak over the mediterranean regions, Vol. 38, Special Edition, 330–336, Ciência e Natura, doi:10.5902/2179460X20249, 2016.

Roney, J. A. and White, B. R.: Estimating fugitive dust emission rates using an environmental boundary layer wind tunnel, Atmos. Environ., 40(40), 7668–7685, doi:10.1016/j.atmosenv.2006.08.015, 2006.

Sabre, M., Lopez, M. V, Alfaro, S. C., Rajot, J. L. and Gomes, L.: Characterization of the fine dust particle production process by wind erosion for two types of bare soil surfaces, in Proceedings of Wind Erosion: An International Symposium/Workshop, 3–5, doi: 10.1007/s00254-006-0363-5, 1997.

Schmechtig, C., Marticorena, B., Chatenet, B., Bergametti, G., Rajot, J.-L. and Coman, A.: Simulation of the mineral dust content over Western Africa from the event to the annual scale with the CHIMERE-DUST model, Atmos. Chem. Phys., 11(14), 7185–7207, doi:10.5194/acp-11-7185-2011, 2011.

Schulz, M., Prospero, J. M., Baker, A. R., Dentener, F., Ickes, L., Liss, P. S., Mahowald, N. M., Nickovic, S., Garcia-Pando, C. P. and Rodríguez, S.: Atmospheric transport and deposition of mineral dust to the ocean: implications for research needs, Environ. Sci. Technol., 46(19), 10390–10404, doi:10.1021/es300073u, 2012.

Shao, Y. and Lu, H.: A simple expression for wind erosion threshold friction velocity, J. Geophys. Res. Atmos., 105(D17), 22437–22443, doi:10.1029/2000JD900304, 2000.

Shao, Y., Ishizuka, M., Mikami, M. and Leys, J. F.: Parameterization of size‐resolved dust emission and validation with measurements, J. Geophys. Res. Atmos., 116(D8), doi:10.1029/2010JD014527, 2011.

Shao, Y.: A model for mineral dust emission, J. Geophys. Res. Atmos., 106(D17), 20239–20254, doi:10.1029/2001JD900171, 2001.

Shao, Y.: Simplification of a dust emission scheme and comparison with data, J. Geophys. Res. Atmos., 109(D10), doi:10.1029/2003JD004372, 2004.

Simpson, D., Benedictow, A., Berge, H., Bergstr?m, R., Emberson, L. D., Fagerli, H., Flechard, C. R., Hayman, G. D., Gauss, M. and Jonson, J. E.: The EMEP MSC-W chemical transport model–technical description, Atmos. Chem. Phys., 12(16), 7825–7865, doi:10.5194/acp-12-7825-2012, 2012.

Singh, P., Sharratt, B. and Schillinger, W. F.: Wind erosion and PM10 emission affected by tillage systems in the world's driest rainfed wheat region, Soil Tillage Res., 124, 219–225, doi:10.1016/j.still.2012.06.009, 2012.

Su, L. and Fung, J. C. H.: Sensitivities of WRF‐Chem to dust emission schemes and land surface properties in simulating dust cycles during springtime over East Asia, J. Geophys. Res. Atmos., 120(21), 11–215, doi:10.1002/2015JD023446, 2015.

Todd, M. C., Karam, D. B., Cavazos, C., Bouet, C., Heinold, B., Baldasano, J. M., Cautenet, G., Koren, I., Perez, C. and Solmon, F.: Quantifying uncertainty in estimates of mineral dust flux: An intercomparison of model performance over the Bodélé Depression, northern Chad, J. Geophys. Res. Atmos., 113(D24), doi:10.1029/2008JD010476, 2008.

Tong, D. Q., Dan, M., Wang, T. and Lee, P.: Long-term dust climatology in the western United States reconstructed from routine aerosol ground monitoring, Atmos. Chem. Phys., 12(11), 5189–5205, doi:10.5194/acp-12-5189-2012, 2012.

Tong, D. Q., Wang, J. X. L., Gill, T. E., Lei, H. and Wang, B.: Intensified dust storm activity and Valley fever infection in the southwestern United States, Geophys. Res. Lett., 44(9), 4304–4312, doi:10.1002/2017GL073524, 2017.

Uno, I., Wang, Z., Chiba, M., Chun, Y. S., Gong, S. L., Hara, Y., Jung, E., Lee, S., Liu, M. and Mikami, M.: Dust model intercomparison (DMIP) study over Asia: Overview, J. Geophys. Res. Atmos., 111(D12), doi:10.1029/2005JD006575, 2006.

Varga, G.: Spatio-temporal distribution of dust storms–a global coverage using NASA TOMS aerosol measurements, Hungarian Geogr. Bull., 61(4), 275–298, 2012.

Vautard, R., Bessagnet, B., Chin, M. and Menut, L.: On the contribution of natural Aeolian sources to particulate matter concentrations in Europe: testing hypotheses with a modelling approach, Atmos. Environ., 39(18), 3291–3303, doi:10.1016/j.atmosenv.2005.01.051, 2005.

Wang, H., Xue, M., Zhang, X. Y., Liu, H. L., Zhou, C. H., Tan, S. C., Che, H. Z., Chen, B. and Li, T.: Mesoscale modeling study of the interactions between aerosols and PBL meteorology during a haze episode in Jing–Jin–Ji (China) and its nearby surrounding region–Part 1: Aerosol distributions and meteorological features, Atmos. Chem. Phys., 15(6), 3257–3275, doi:10.5194/acp-15-3257-2015, 2015.

Wang, K., Zhang, Y., Nenes, A. and Fountoukis, C.: Implementation of dust emission and chemistry into the Community

Multiscale Air Quality modeling system and initial application to an Asian dust storm episode, Atmos. Chem. Phys., 12(21), 10209–10237, doi:10.5194/acp-12-10209-2012, 2012.

Wang, R., Chang, C., Peng, S. and Wang, L.: Estimation on farmland wind-erosion and dust emission amount in Bashang of Hebei province by grain composition contrast, Trans. Chinese Soc. Agric. Eng., 29(21), 108–114, 2013.

White, B. R.: Encyclopedia of fluid mechanics, Gulf Publishing, Houston, Tex, USA, 239–282, 1986.

White, B. R.: Soil transport by winds on Mars, J. Geophys. Res. Solid Earth, 84(B9), 4643–4651, doi:10.1029/JB084iB09p04643, 1979.

Xi, X. and Sokolik, I. N.: Quantifying the anthropogenic dust emission from agricultural land use and desiccation of the Aral Sea in Central Asia, J. Geophys. Res. Atmos., 121(20), doi:10.1002/2016JD025556, 2016.

Xi, X. and Sokolik, I. N.: Seasonal dynamics of threshold friction velocity and dust emission in Central Asia, J. Geophys. Res. Atmos., 120(4), 1536–1564, doi:10.1002/2014JD022471, 2015.

Zhang, J., Teng, Z., Huang, N., Guo, L. and Shao, Y.: Surface renewal as a significant mechanism for dust emission, Atmos. Chem. Phys., 16(24), 15517–15528, doi:10.5194/acp-16-15517-2016, 2016.

Zhang, X., Zhao, L., Tong, D. Q., Wu, G., Dan, M. and Teng, B.: A systematic review of global desert dust and associated human health effects, Atmosphere (Basel)., 7(12), 158, doi:10.3390/atmos7120158, 2016.

Zhang, X., Zhou, Q., Chen, W., Wang, Y. and Tong, D. Q.: Observation and modeling of black soil wind-blown erosion from cropland in Northeastern China, Aeolian Res., 19, 153–162, doi:10.1016/j.aeolia.2015.07.009, 2015.

Zhao, C., Liu, X. and Leung, L. R.: Impact of the Desert dust on the summer monsoon system over Southwestern North America, Atmos. Chem. Phys., 12(8), 3717–3731, doi:10.5194/acp-12-3717-2012, 2012.